Early Implementation of the Head Start Designation Renewal System: Volume I

Early Implementation of the Head Start Designation Renewal System: Volume I

Co-principal investigators and senior adviser
Teresa Derrick-Mills, Urban Institute
Margaret R. Burchinal, Frank Porter Graham Child Development Institute
H. Elizabeth Peters, Urban Institute

And in alphabetical order

Allison De Marco and Nina Forestieri, Frank Porter Graham Child Development Institute
Saunji Fyffe, Devlin Hanson, Caroline Heller, Eleanor Pratt, Heather Sandstrom, Tim Triplett, and Tyler Woods, Urban Institute

ABOUT THE URBAN INSTITUTE

The nonprofit Urban Institute is dedicated to elevating the debate on social and economic policy. For nearly five decades, Urban scholars have conducted research and offered evidence-based solutions that improve lives and strengthen communities across a rapidly urbanizing world. Their objective research helps expand opportunities for all, reduce hardship among the most vulnerable, and strengthen the effectiveness of the public sector.

Cover image from the Frank Porter Graham Child Development Institute.

Contents

Tables

Figures

Acknowledgments

Expert Work Group Members

We would like to thank the following members of the *Early Implementation of the Head Start Designation Renewal System* expert work group. The views expressed in this publication do not necessarily reflect the views of these members.

Greg Duncan
Distinguished Professor
School of Education, University of California-Irvine

Stephanie Jones
Marie and Max Kargman Associate Professor in Human Development and Urban Education Advancement
Harvard Graduate School of Education

Christine McWayne
Associate Professor; Director of Early Childhood Education
Eliot-Pearson Department of Child Study and Human Development, Tufts University

Kathryn Newcomer
Professor of Public Policy and Public Administration; Director, Trachtenberg School of Public Policy and Public Administration
The George Washington University

Kathryn Tout
Co-director, Early Childhood Research; Senior Research Scientist
Child Trends

Mildred Warner
Professor, Department of City and Regional Planning
Cornell University

Overview

The Head Start Designation Renewal System (DRS) was created in 2011 as an accountability system to determine whether Head Start grantees are delivering high quality and comprehensive services and to implement a competitive grant making process in communities where they are not. The purpose of this study was to evaluate the early implementation of DRS, describe the mechanisms by which DRS supports program quality improvement, examine how well DRS is able to identify grantees providing lower quality services, and describe the results of competitions. The evaluation uses a mixed-methods design that integrates administrative data and secondary data sources, observational assessments, and interviews to tap multiple dimensions of program quality and experiences with DRS at all phases of the process. Highlights of the study findings include:

Incentivizing Quality Improvement

- In qualitative interviews, selected grantees saw a need for accountability and quality improvement in Head Start and believed most of the DRS conditions are reasonable. However, they expressed concerns about the how well the system, especially the Classroom Assessment Scoring System (CLASS) condition, assesses program quality and with how designation decisions are made, which may undermine DRS incentives for quality improvement.
- Nonetheless, grantees appear to be responding to the incentives in DRS and the CLASS condition in particular. Survey data as well as qualitative interviews suggest grantees are engaging in multiple strategies to improve quality, including coaching for teachers to improve teacher-child interactions, training for managers and supervisors related to CLASS, and creating new staff positions aimed at raising CLASS scores.

Assessing Grantee Quality

- Overall, there were no significant differences between grantees designated for competition and grantees not designated on any of the evaluation's assessments of classroom or center quality.
- However, when looking at the individual DRS conditions, the study found quality of center operations was significantly lower among grantees designated for deficiencies compared to not-designated grantees. No differences were found among grantees designated for CLASS compared to not-designated grantees.
- Although CLASS observers from both the Office of Head Start (OHS) and the evaluation team met the developer's reliability standards, there were substantial differences in CLASS scores collected by the two teams and substantial variability that could be attributed to raters, raising concerns about the precision of CLASS for the purposes of DRS.

Results of and Responses to Competition

- The act of applying for grants opened through designation appears to serve as an incentive for quality improvement even though a majority of competitions had only one applicant. In qualitative interviews, selected grantees saw the competitive application as an opportunity to make changes they would not have otherwise undertaken. Survey and administrative data show both grant applicants and awardees proposed quality improvements such as higher shares of teachers with BAs, lower ratios, and lower enrollment compared to the prior grantee.
- However, qualitative interviews with selected awardees suggest challenges with the negotiation process, which posed barriers to implementing quality improvements and may have led to service disruptions.

Executive Summary

Background

Head Start is a comprehensive early childhood development program serving low-income families and children from birth through age five administered through the U.S. Department of Health and Human Services (HHS), Administration for Children and Families (ACF), Office of Head Start (OHS). Established in 1965 as an anti-poverty initiative, Head Start promotes "the school readiness of low income children by enhancing their cognitive, social, and emotional development" through supportive learning environments and an array of "health, educational, nutritional, social and other services" to children and their families (Sec. 636. [42 U.S.C. 9831]). Today, local services are delivered through about 1,700 public, private nonprofit, and private for profit organizations across the United States and in the US territories and Puerto Rico (OHS 2016a).

In 2011, ACF expanded its accountability provisions for the Head Start program through the implementation of the Head Start Designation Renewal System (DRS). Before 2011, Head Start grants were awarded for indefinite periods. The Improving Head Start for School Readiness Act of 2007 required that Head Start grants be awarded for a five-year period. Furthermore, the Act required ACF to develop a system for identifying which grantees provide high quality comprehensive services and can receive automatic grant renewal at the end of the 5-year period; and which grantees are not and will be subject to open competition to receive renewed funding. The DRS was developed in response to these provisions and identifies seven conditions of designation renewal. If a Head Start grantee does not meet any of the conditions, they receive automatic renewal of their grant.

> **Head Start Designation Renewal System Seven Conditions**
>
> 1. A deficiency (i.e., a systemic or substantial failure) in meeting program performance standards.
> 2. Classroom Assessment Scoring System (CLASS) domain scores below a minimum threshold or in the lowest 10 percent of the distribution.
> 3. Failure to establish, analyze and utilize program-level school readiness goals.
> 4. Revocation of state or local license to operate.
> 5. Suspension by ACF.
> 6. Debarment from receiving funding from another federal or state agency or disqualification from participation in the Child and Adult Care Food Program.
> 7. Determination from an annual audit of being at risk for failing to continue functioning as a "going concern".

If a Head Start grantee meets any one of the seven conditions, the grant is subject to an open competition for future funding. From 2011 to 2015, 453 Head Start grants were designated for competition based on DRS (ACF 2011, 2013, 2014; OHS 2015a).

Because the DRS represents a major change in the accountability structure for Head Start and has significantly changed the way ACF administers and manages the program, at the request of ACF leadership, the ACF Office of Planning, Research, and Evaluation (OPRE) commissioned a study of DRS implementation, including how the system is addressing its goals of transparency, validity, reliability, and stimulating program quality improvement efforts. The study's purpose is to understand the mechanisms by which the DRS might support program quality improvement, ask whether the DRS identifies grantees providing lower-quality services for designation for competition, and examine the results of grants competitions in terms of the types of applicants that received awards and their efforts to improve services in the community. The study is observational and cannot address causal relationships (e.g., does DRS improve quality of services, or do grantees change their policies and practices because of the DRS) and could not document whether quality improved over time within all grantees or even in the sites designated for competition. The goal of the study is to describe early implementation of the DRS and to ask whether the DRS appears to incentivize grantees' efforts to improve quality during monitoring or competition.

Evaluation Approach

In fall 2012, OPRE contracted with the Urban Institute and its subcontractor, the Frank Porter Graham Child Development Institute at the University of North Carolina at Chapel Hill, to design and execute a study of the early implementation of the DRS. The evaluation is based on an assumption that DRS has two primary mechanisms for improving quality in Head Start programs: (1) incentivizing all grantees to improve quality in an effort to avoid designation for competition, and (2) in communities identified as providing lower quality services, introducing grant competitions through which applicants propose quality improvements to be competitive. The evaluation team sought to examine these mechanisms by describing the early implementation of the DRS in terms of how well the conditions of DRS differentiate between grantees providing lower- and higher-quality services, the types of quality improvement activities in which grantees engage before or during the quality assessment process, the extent to which competition exists for Head Start grants, and the quality improvement efforts introduced through the competitive process. Specifically, the evaluation addresses five research questions (RQ):

1. **How do grantees perceive the DRS and prepare for their monitoring and assessment in terms of efforts to improve program quality?**
2. **Does the DRS differentiate higher versus lower quality programs?**
3. **How do grantees perceive and experience the monitoring, assessment, and DRS designation processes in terms of their efforts to improve program quality?**
4. **What does competition look like? How much competition is generated by DRS?**
5. **How do grantees experience the negotiation, award, and transition processes in terms of preparing them to implement quality improvements to existing services?**

To address these questions, the evaluation used a mixed-methods design that integrates quantitative observational assessments, surveys, and administrative data with qualitative interviews. Table ES.1 describes the sample, constructs of interest, data-collection methods and tools for each research question.

To examine how well the DRS differentiates higher- and lower-quality programs (RQ2), we used various quantitative methods to measure the quality of preschool classrooms, teacher-child interactions, health and safety practices, family engagement, program operations and governance, and fiscal management, and compared how the evaluation's assessments of quality align with OHS's assessments of quality in the DRS. A random sample of 71 grantees (35 designated for competition, 36 not designated for competition) participated in the evaluation's on-site quality assessments. Much of this sample was drawn from grantees receiving their Triennial Monitoring Review visit during the time of the evaluation (during the 2013-14 monitoring year; Monitoring Cohort 4). Because of low response rates, the evaluation had to draw some of its sample from the previous monitoring year (2012-13; Monitoring Cohort 3).[1] One of the quality assessments, the financial vulnerability assessment, used a separate sample of 216 nonprofit grantees. See chapter II and technical volume II, appendix B for more information.

To understand how programs perceive and respond to the monitoring and assessment process applying the DRS conditions (RQ1 and RQ3), we conducted a survey of Head Start directors regarding professional development and technical assistance practices, as well as qualitative interviews with staff at multiple levels of the Head Start organization to explore programs' experience with the monitoring and assessment process. Directors from the 71 grantees participating in the quality assessments (RQ2) were invited to participate in the survey; 66 directors completed it. From among them, 35 directors

---

[1] Of the 71 grantees, 61 grantees (29 designated for competition) were from the 2013-14 monitoring year and 10 grantees (6 designated for competition) were from the 2012-13 monitoring year.

were purposively selected for telephone interviews and 15 grantees were purposively selected[2] for on-site qualitative interviews with multiple staff and stakeholders. See chapter II and technical volume II, appendix C for more information.

To examine competition (RQ4 and RQ5), the evaluation conducted a survey of applicants for Head Start grants to understand their proposed plans, analyzed administrative and secondary data to examine applicant and awardee characteristics, and conducted qualitative interviews with key staff at a sample of awardee agencies to understand how the competitive process may relate to quality improvements in Head Start. The sample for this part of the evaluation was drawn from among grantees in DRS Cohort 3 and applicants for the 2014 funding opportunity announcements (FOAs). The evaluation obtained administrative data on the characteristics of all 105 grantees designated for competition in DRS Cohort 3 and 177 of 182 applicants for the 2014 FOAs. Additionally, 120 of the 182 applicants responded to a survey of program features and services described in their grant applications. Finally, nine grantees who received an award resulting from the 2014 FOAs were purposively selected for qualitative interviews with staff at multiple levels of the organization to explore their experience with the competition, award and transition processes. See chapter II and technical volume II, appendix D for more information.

---

[2] Purposive selection focused on ensuring a sample of grantees that varied in terms of their geographic region, designation status, and funded enrollment. See table C.1 in technical volume II, appendix C for more detail on the characteristics of participating grantees.

TABLE ES.1.

Overview of Sample, Constructs, Methods and Tools by Research Question

Research Question	Sample	Construct	Assessment tool
1. How do grantees perceive the DRS and prepare for their monitoring and assessment in terms of efforts to improve program quality?	66 grantees	Professional development and technical assistance activities	[a]Professional Development and Technical Assistance (PDTA) Survey
	35 grantees	Perceptions of and preparations for DRS	[a]Semi-structured telephone interview with Head Start directors
	15 grantees	Perceptions of and preparations for DRS	[a]Semi-structured on-site interviews with program directors and managers, executive directors, governing body members, and Policy Council members
2. Does the DRS differentiate higher versus lower quality programs?	71 grantees	Classroom quality	Classroom Assessment Scoring System (CLASS) Early Childhood Environment Rating Scale—Revised (ECERS-R) Interaction Score Early Childhood Environmental Rating Scale—Extension (ECERS-E) Academic Activities Score Adapted Teacher Styles Rating Scale (TSRS Adapted) Classroom Structure and Management Subscale
		Child health and safety	Health and Safety Checklist (adapted from NAEYC and CA checklists)
		Family and community engagement	Program Administration Scale (PAS) Family Partnership Subscale
		Child development and education	PAS Child Assessment Subscale
		Management, operations, and governance systems	Composite: PAS Governance and Management, Fiscal Management Personnel Cost and Allocation subscales
	216 nonprofit grantees[b]	Fiscal management	Tuckman & Chang Financial Vulnerability Ratios using IRS Form 990 data
3. How do grantees perceive the monitoring, assessment, and DRS designation processes in terms of their efforts to improve program quality?	35 grantees	Perceptions of and preparations for DRS	[a]Semi-structured telephone interview with grantee Head Start directors
	15 grantees	Perceptions of and preparations for DRS	[a]Semi-structured on-site interviews with program directors and managers, executive directors, governing body members, and Policy Council members
4. What does competition look like?	105 grantees designated for competition; 182 applicants for grants	Characteristics of grantees designated for competition and applicants for new awards	Extant data sources including: OHS administrative data on applicants and awardees, Program Information Report (PIR) data, Census data, and other publicly available data

Research Question	Sample	Construct	Assessment tool
	120 applicant respondents	Proposed service delivery	[a]Applicant survey
	9 grantees with new awards	Decisions to compete, proposed services, the competition and award processes	[a]Semi-structured on-site interviews with program directors and managers, executive directors, governing body members, and Policy Council members awardees
5. How do grantees experience the negotiation, award, and transition processes in terms of preparing them to implement quality improvements to existing services?	105 grantees designated for competition; 182 applicants for grants	Characteristics of grantees designated for competition and applicants for new awards	Extant data sources including: OHS administrative data on applicants and awardees, Program Information Report (PIR) data, Census data, and other publicly available data
	120 applicant respondents	Proposed service delivery	[a]Applicant survey
	9 grantees with new awards	Decisions to compete, proposed services, the competition and award processes	[a]Semi-structured on-site interviews with program directors and managers, executive directors, governing body members, and Policy Council members awardees

Note: See chapter II and technical volume II for detailed descriptions of the samples used in these analyses.
[a]Instrument developed for this evaluation.
[b]All nonprofit grantees in dataset were selected for financial vulnerability analysis.

Key Findings

The findings from the evaluation provide a mixed picture of the DRS. There is some qualitative evidence to suggest that the DRS is incentivizing grantees to improve quality, especially the quality of teacher-child interactions. There is also evidence that grant applicants—both new applicants and incumbents—proposed to improve program quality relative to that of the grantees at the time of designation. On the other hand, the evaluation suggests that the two conditions of the DRS studied here (monitoring deficiency and CLASS) are inconsistent in identifying higher- versus lower-quality grantees, that the opportunity for competition does not necessarily generate competition, and the transition to services after new grant awards has challenges. We discuss these findings below relative to the five research questions.

How do grantees perceive the DRS and prepare for their monitoring and assessment in terms of efforts to improve program quality?

Nineteen of 35 Head Start directors participating in telephone interviews agreed that an accountability system such as the DRS was needed to improve program quality in Head Start. Most grantees indicated

that most of the seven conditions were fair, but 74 percent expressed concerns about the CLASS condition, particularly that grantees scoring in the lowest 10 percent of any CLASS domain were designated for competition. All 66 grantees participating in the professional development and technical assistance (PDTA) survey reported engaging in multiple professional development and technical assistance strategies in the 12 months before their monitoring review. Sixty-five percent indicated that these strategies were not any different than what they would have done at any other time. In qualitative interviews with 35 directors, grantees indicated that they are working on improvements all the time so it is difficult to identify which ones are specific to the DRS, but the directors pointed to some changes in program structure and procedures and some added emphasis on CLASS and school readiness goals. In addition, all 15 grantees participating in the on-site interviews indicated that they had been working on improving CLASS scores for the last two to six years, and some indicated this was because they anticipated CLASS would become a requirement based on the passage of the 2007 Head Start Act.

Does the DRS differentiate higher versus lower quality programs?

Comparisons of grantees that were and were not designated for competition, regardless of the condition for which they were designated, did not reveal statistically significant differences in the quality of classroom or center practices as measured by the evaluation. Some evidence of differences emerged, however, when individual DRS designation conditions were examined. Analyses examined 71 randomly selected grantees on measures of classroom quality, health and safety, family engagement, program governance and management, and fiscal operations, and examined the 216 nonprofit grantees on financial vulnerability.[3] First we compared grantees by overall designation status (designated for competition versus not designated). Then we examined grantees designated for competition based on deficiencies in meeting Head Start Performance Standards separately from grantees designated for competition due to low CLASS scores.[4] Grantees designated due to deficiencies had lower quality on three of six measures of center quality compared with grantees that were not designated. In contrast, no analyses indicated that grantees designated due to low CLASS scores differed from grantees that were not designated on any study measure of quality.[5]

---

[3] Evaluation measures were administered at the classroom, center, and grantee levels. For information about which measures applied to which level, see technical volume II, appendix B.
[4] The evaluation focused on these two conditions in DRS because they accounted for 99 percent of designations for competition at the time of the evaluation.
[5] All analyses were conducted with weights to account for low response rates and multiple imputations (a statistical technique used to analyze datasets with missing values) to account for missing data.

To further understand the lack of differences found between grantees designated for competition due to low CLASS scores and not-designated grantees, we undertook additional psychometric analyses to examine how the CLASS performs as a measure of program-level quality, as it is used for the purpose of the DRS. Specifically, we looked at whether program-level CLASS scores were influenced by the number of cycles of CLASS observation or the individual CLASS observer. Using the CLASS data collected by the evaluation team, we found no evidence that scores were significantly different when either the typical four or fewer cycles of observation were compared, suggesting that the DRS practice of collecting two cycles per classroom in the OHS monitoring did not change grantee-level CLASS scores. In contrast, evidence suggested substantial differences among raters in the CLASS scores and in other measures of classroom quality, such as the ECERS-R. Our analyses of CLASS scores collected by the evaluation team suggest that variability among raters may account for up to 45 percent of the variability in the CLASS. Because training of DRS CLASS observers was similar to training of the evaluation team CLASS observers, and both teams of observers were trained to the reliability standards set by the developer, variability among DRS CLASS raters may be similar. Furthermore, there were large differences in the grantee-level mean CLASS scores collected by the OHS CLASS observers and by our data collectors (with OHS observers rating grantees higher than the evaluation's data collectors), and only limited evidence of correlations in scores across the two data-collection teams. These findings were obtained even when controlling for the time elapsed between the two assessments[6] or whether the same or different classrooms[7] were observed by both teams. These findings raise concerns about whether the CLASS is sufficiently precise for the purposes of the DRS.

How do grantees perceive the monitoring, assessment, and DRS designation processes in terms of their efforts to improve program quality?

All 15 Head Start grantees participating in on-site interviews after they knew their designation status (seven designated for competition, eight not designated) indicated they were continuing to work on improving quality. Grantees designated for competition were more likely than not-designated grantees to indicate some of their quality improvement efforts were related to the DRS. Some of their efforts were focused on CLASS, and some were focused on specific areas of need identified by the program, such as training for substitute teachers on positive guidance and establishing new practices for ensuring

---

[6] On average, 1.2 months passed between when OHS CLASS observations and the evaluation team's CLASS observations were conducted, but 10 percent were more than a year apart because of including the earlier monitoring cohort in the sample.
[7] Whether the same or different classrooms were observed was indexed by whether all classrooms within a grantee were observed by both teams.

children are not left alone. Grantees expressed concerns about the early implementation of the monitoring, assessment, and notification processes associated with DRS including concerns about determining quality based on a single visit or incident (20 of 35 grantees), transparency in the designation decision-making process (8 of 15 grantees), and the length of time between monitoring visits and knowing the designation status (6 of 15 grantees).

What does competition look like?

Based on the evaluation team's analysis of OHS data on grant competitions, 42 percent of the 103 DRS Cohort 3 grants opened for competition in 2014 had more than one applicant. Based on administrative data obtained from OHS (PIR and data on grant applicants), 91 percent applied for funding to be able to continue providing Head Start services in their community. Non-incumbent applicants for the grants (labeled new applicants) were either current Head Start grantees or delegates in other service catchment areas (54 percent) or organizations new to Head Start (46 percent).

Factors related to the decision to apply were different for new applicants versus incumbents, as reported during interviews with directors and staff in nine grantees that received a grant award following competition. All four of the incumbents indicated they had already decided to apply and began planning their application before the formal announcement calling for proposals. The interviews with five new awardees indicated that several factors determined whether to apply, including the effort required to adhere to federal reporting standards, the well-being of children in the community, the potential reaction of the community, and the likely competition from other applicants.

Based on surveys sent to all applicants for grants associated with DRS Cohort 3 competitions, most applicants, including incumbent applicants, proposed changes to service delivery including more delegates, higher shares of teachers with bachelor's degrees (BA), and lower Head Start enrollment than the current grantee. Similarly, based on more in-depth qualitative interviews with awardees, eight of nine proposed service changes. All four of the incumbents felt the competitive FOA provided opportunities to propose changes to their programs they could not otherwise have made, including adding or increasing Early Head Start slots and increasing the share of teachers with a BA.

How do grantees perceive the negotiation, award, and transition processes in terms of preparing them to implement quality improvements to existing services?

The DRS competitive process largely resulted in incumbents winning back their grants: 76 percent of awards went to incumbents, 17 percent to other Head Start grantees or delegates,[8] and 6 percent to a new organization. On average we found that awardees proposed higher proportions of teachers with BA's, lower child-teacher ratios, and lower enrollment than the prior grantee. This was true even when we restricted the analysis to FOAs with just one applicant. This suggests that being designated and applying competitively may lead to higher proposed service quality even without the presence of competitors. Finally, in our exploratory analysis of the negotiation, award, and transition processes of the nine awardees, we found that the timing and confidentiality of the negotiation process created challenges in preparing for initial implementation of the proposed changes to service for eight of the nine interviewed awardees.

Study Limitations

Several limitations must be recognized. First, this is a formative evaluation. The qualitative and quantitative components of the study focused on describing the experiences of grantees, and, as such, do not allow for making causal inferences. Second, not all findings from this evaluation may generalize to all Head Start grantees. Grantees that operated only Early Head Start, home-based, American Indian Alaskan Native, Migrant and Seasonal Head Start, or were located outside of continental United States were excluded from the evaluation except for analyses of administrative data on competition and awards and the applicant survey. Third, it was more difficult to recruit grantees into the study than we anticipated. Response rates were lower than expected for the 71 grantees participating in on-site quality assessments (46 percent), the PDTA survey (44 percent),[9] and the survey of the 182 applicants (65 percent), and the other studies drew subsamples of these. This could affect validity of the findings despite the use of sampling weights to account for this. Fourth, the measurement of the center- and grantee-level operations, family engagement, staff qualifications, and child assessments with the Program Administration Scale (PAS) was compromised. Center directors objected to the amount of time required to prepare for these assessments, which impeded classroom and center recruitment, so we selected a single center per grantee for the last half of the data collection. This may have introduced

---

[8] Two of the awardees were Head Start delegate agencies that were not also Head Start grantees in other service areas.
[9] The PDTA survey was collected during the quality assessments data collection.

differences in PAS scores even after correcting for missing data.[10] Therefore, we are less confident about drawing conclusions from this instrument. Finally, grantees were purposively selected for the telephone interviews (35), for on-site interviews about the monitoring and assessment process (15), and for awardee interviews about the competitive process (9). Their interviews are meant to provide various perspectives and experiences across Head Start programs of different sizes, different regions, and operating both birth-to-five and Head Start-only services. The perceptions, experiences, and actions of these grantees may not fully represent the larger population of Head Start grantees, and we overrepresented grantees designated for competition in the telephone interviews and on-site interviews to assure diverse perspectives; similarly, we overrepresented new awardees in the awardee interviews.

---

[10] The evaluation attempted to account for missing data with the PAS using multiple imputations.

Key Terms

Some of the key terms in this document will be familiar to those familiar with Head Start. Other key terms, however, are associated specifically with the designation renewal system as evaluated in this report. Here we provide both kinds of key terms to guide the readers. Terms are provided in alphabetical order.

Agency Director—the person within an organization that is responsible for oversight and leadership of the whole organization, not just the Head Start program. In this evaluation, this refers to someone other than a Head Start Director when the Head Start program is located in an organization that provides services broader than Head Start.

Applicant—an organization that responds to a funding opportunity announcement

Awardee—an organization that is awarded the grant resulting from the funding opportunity announcement

Competition—when OHS releases a funding opportunity announcement (FOA), it is opening up a competition for the grant specified in the FOA. The applicants for the FOA are the competitors

Delegate Agency—"a public, private nonprofit (including a community based organization, as defined in section 9101 of the Elementary and Secondary Education Act of 1965 (20 U.S.C. 7801), or for profit organization or agency to which a grantee has delegated all or part of the responsibility of the grantee for operating a Head Start program" (Sec. 637. [42 W.S.C. 9832] (C)(3))

Designated for Competition—occurs when grantees meet one of the seven conditions of the DRS[11]

Designation Renewal System—the Head Start Designation Renewal System is an accountability and quality improvement mechanism by which grantee quality is monitored and assessed; if grantees meet any one of seven conditions, they are designated for competition

DRS Cohort—the group of grantees designated to compete for continued funding in a given year and the resulting competitions (funding opportunity announcements) associated with the communities they serve. Since 2011, there have been 4 DRS Cohorts: Cohort 1 was announced in December of 2011

---

[11] The 2007 Head Start Act uses the term "designated" to refer to the Secretary's authority to designate an agency as a Head Start program, but "designated" is used colloquially in the Head Start community to refer to grantees designated for competition as a result of DRS. To avoid confusion by readers, the evaluation elected to use "designated for competition" in the manner it is used colloquially rather than in legislation.

(shortly after the final rule published), Cohort 2 was announced in January 2013, Cohort 3 in February 2014, and Cohort 4 in December 2014.

Funding Opportunity Announcement (FOA)—the vehicle through which the Office of Head Start solicits applications for federal Head Start funding

Governing Body—the entity with ultimate fiduciary and oversight responsibilities in Head Start; it may be the same as the governing board or board of directors in organizations where a governing body would exist without Head Start mandate

Grantee—the organization that has the grant with the federal government for the administration of the Head Start and/or Early Head Start grant. This organization may provide services directly or via partnerships/delegate agencies. Designation occurs at the grantee level for failure to meet any one of the seven conditions by any part of the program, at any level.

Head Start Director—the person within a Head Start grantee charged with overseeing its operation

Incumbent—the organization holding the Head Start grant at the time of the release of a funding opportunity announcement

Likely Designated for Competition or Likely Not-Designated—the designation status that the evaluation estimated for grantees based on data available to us before the designation status had been announced; the estimated designation was used for sampling purposes because final designation status had not yet been assessed at the time of sampling; all analyses in this evaluation use the final designation status as determined by the Funding Opportunity Announcements posted for DRS Cohorts 3 and 4

Monitoring Cohort—for the purpose of the evaluation, the Monitoring Cohort refers to all grantees that were monitored and/or received a CLASS review in a given fiscal year (October – September) including all grantees eventually designated for competition (DRS Cohort) in that year and all grantees that received a noncompetitive grant renewal. Since 2011, there have been 4 Monitoring Cohorts: Monitoring Cohort 1 received a monitoring review between June 2009 and November 2011, Monitoring Cohort 2 in FY2012, Monitoring Cohort 3 in FY2013, and Monitoring Cohort 4 in FY2014.

New Applicant—an organization responding to a FOA that does not already hold the grant for that service delivery area

New Awardee—an organization awarded the grant resulting from the FOA that did not previously hold the grant for that service delivery area (i.e., not the incumbent); may be a Head Start grantee from another service delivery area, a previous Head Start delegate, or an organization new to Head Start

Not Designated or Not-Designated—refers to not designated for competition; grantees do not meet any of the DRS conditions and may renew their grants noncompetitively

OHS Monitoring—assesses grantee compliance with the Head Start Program Performance Standards, the Head Start Act, and other regulations. During the time of the evaluation, routine monitoring occurred every three years, but other types of monitoring occurred as needed, and resulted in one of three possible outcomes: compliance, noncompliance, and deficiency which would trigger designation for competition.

Policy Council—a required part of Head Start program governance, the Policy Council is composed of elected parents, constituting the majority, and community members. Per the Head Start Act, it is responsible for the direction of the Head Start program, including program design and operation and long term planning goals and objectives.

Program Managers—individuals in Head Start programs that have some type of managerial or coordination responsibilities below the level of the Head Start director

Service Delivery Area—this is the geographic area to which a Head Start program extends its services to children and families. Typically, children and families must live within a particular geographic area to receive services. The geographic area is defined by the Office of Head Start. It may be as large as a state or as small as a neighborhood.

Chapter I. Introduction

In 2011, the Administration for Children and Families (ACF) expanded its accountability provisions for the Head Start program by implementing the Head Start Designation Renewal System (DRS). In response to requirements established under the reauthorization bill for Head Start, Improving Head Start for School Readiness Act of 2007, the DRS is designed to identify which Head Start grantees are not providing high quality comprehensive services to children and families in their communities based on at least one of seven conditions. Head Start grantees that do not meet any of the conditions receive automatic renewal of their grant. If a Head Start grantee meets any one of the seven conditions, the federal grant is subjected to an open competition for future funding.

Because the DRS represents a major change in the accountability structure for Head Start and has significantly changed how ACF administers and manages the program, the Office of Planning, Research, and Evaluation (OPRE) within the Administration for Children and Families (ACF), U.S. Department of Health and Human Services (HHS), at the request of ACF leadership, commissioned a study of the implementation of the DRS, including how the system is addressing its goals of transparency, validity, reliability, and stimulating program quality improvement efforts. The study purpose is to understand the mechanisms by which the DRS might support program quality improvement, ask whether the DRS appears to be identifying grantees providing lower quality services for designation for competition, and examine the results of grants competitions by the types of applicants that received awards and their efforts to improve services in the community. The study does not, however, look at whether Head Start program quality improved over time, and the study cannot speak to whether DRS causes changes in program policies, actions, or quality. The goal of the evaluation is to describe early implementation of the DRS and to ask whether the DRS appears to incentivize grantees' efforts to improve quality during monitoring or competition.

This introduction provides an overview of Head Start, background on the Head Start Designation Renewal System, the goals of the evaluation, and a road map for the report.

What is Head Start?

Head Start is a comprehensive early childhood development program serving low-income families and children from birth through age five administered through the U.S. Department of Health and Human Services (HHS), Administration for Children and Families (ACF), Office of Head Start (OHS). Established

in 1965 as an antipoverty initiative, Head Start promotes "the school readiness of low income children by enhancing their cognitive, social, and emotional development" through supportive learning environments and the provision of an array of "health, educational, nutritional, social and other services" to children and their families (Sec. 636. [42 U.S.C. 9831]). Local services are delivered through about 1,700 public, private nonprofit, and private for profit organizations across the United States and in the U.S. Territories and Puerto Rico (OHS 2016a). Grantees may offer Head Start only, Early Head Start only, or a combination of Head Start and Early Head Start services. They may operate center-based care, family child care, home-based (home visiting) or a mix of services. Grantees may operate all the services directly, delegate portions or all of their service delivery to delegate agencies, contract for slots in child care and prekindergarten programs, or partner with social service agencies for various support services.

In fiscal year 2014 (FY 2014), the federal government provided nearly $8 billion in funding with about $7.9 billion for local Head Start services and the remainder for support activities. These funds served just over 1 million children and families during the FY 2014 year (OHS 2015b). Head Start is the umbrella term used to refer to the following types of programs:

- Head Start, established in 1965 to serve 3-to-5 year-old children and their families

- Early Head Start, established in 1995 to serve children from birth through age two years, pregnant mothers, and their families

- Migrant and Seasonal Head Start, serving children from birth to age 5 in families migrating because of agricultural work

- American Indian and Alaska Native Head Start administered by Tribal organizations

Qualifying families must earn incomes at or below the federal poverty level, be experiencing homelessness, or the child must be in foster care. Families may earn up to 130 percent of the federal poverty level if their local program can demonstrate they are serving or attempting to serve all families below the federal poverty line. At least 10 percent of Head Start enrollees must be children with disabilities (Services to Children with Disabilities) (OHS 2015c).

What is the Head Start Designation Renewal System?

Since the program's inception in 1965, Head Start grants were awarded with an indefinite project period. This changed with the most recent reauthorization of Head Start, the Improving Head Start for School Readiness Act of 2007 (Pub. L. 10-134) which established a five-year grant period for Head Start programs and further requires that "only grantees delivering high quality services will be given another five-year grant non-competitively" (OHS n.d.(a)). Section 641 of the Act required the Secretary of the U.S. Department of Health and Human Services (HHS) to: develop and implement a system of designation renewal to determine if Head Start grantees are delivering high quality and comprehensive services, and meeting program and financial management requirements and standards; and to convene a panel of experts to "make recommendations to the Secretary on the development of a transparent, reliable, and valid system for designation renewal" (OHS n.d.(a)).

The DRS, established in final rule on November 9, 2011, requires that "for the first time in the history of Head Start, individual grantees whose programs fall short of certain standards will be required to compete with other organizations to continue receiving funding" (OHS 2011). The rule also stipulates that funds shall be awarded "to the organization that demonstrates it can best meet the needs of Head Start children and families" (OHS 2011).

DRS aims to incentivize program planning and quality improvement in the Head Start program overall, and introduce competition in places where grantees are underperforming. The core of the DRS is a set of seven conditions designed to assess whether existing grantees are delivering high-quality, comprehensive services. The specific conditions are as follows:

1. A deficiency (i.e., a systemic or substantial failure) in meeting program performance standards resulting from a triennial, follow-up or unannounced monitoring review
2. Classroom Assessment Scoring System (CLASS) scores below a minimum threshold or in the lowest 10 percent of the cohort in Emotional Support, Classroom Organization, or Instructional Support
3. Failure to establish, analyze, and utilize program-level school readiness goals.
4. Revocation of state or local license to operate
5. Suspension by ACF
6. Debarment from receiving funding from another federal or state agency or disqualification from participation in the Child and Adult Care Food Program (CACFP)
7. Determination from an annual audit of being at risk for failing to continue functioning as a "going concern"

Grantees that meet any one of these conditions are designated for competition. If these grantees want to continue providing Head Start services, they must enter an open competition where other organizations may also apply for the grant. The competitions are announced through Funding Opportunity Announcements (FOAs) posted to the federal Grants.gov website. Each FOA represents a particular service delivery area that may be the same as the service delivery area for the incumbent grantee (the one delivering the grant at the time of designation) or that may be somewhat different, as determined by OHS. Applications are reviewed by independent review panels and the grant is awarded to the organization that best demonstrates it can deliver high quality Head Start services to the children and families it proposes to serve.

DRS Implementation Rollout Period

Initial implementation of the DRS rolled out in four monitoring cohorts (see table I.1). Monitoring Cohort 1 included all grantees with a deficiency from June 2009 to November 2011 to identify the first round of grantees that would have to enter a competitive process. The next cohorts transitioned into DRS, Monitoring Cohort 2 in 2011-2012, Monitoring Cohort 3 in 2012-2013, and Monitoring Cohort 4 in 2013-2014, using their planned triennial monitoring review years.[12] One hundred thirty-two grantees were designated for competition in the first DRS Cohort (ACF 2011). In each subsequent year fewer grantees were designated for competition: 122 grantees in Cohort 2 (ACF 2013), 103 grantees in Cohort 3 (ACF 2014), and 90 grantees in Cohort 4 (OHS 2015a).

---

[12] Before the DRS and the five-year grant award period, grantees were subject to a comprehensive monitoring review once every three years with approximately one-third of grantees reviewed each year. The review examined compliance in meeting all of the Head Start Performance Standards in place at the time of the review.

TABLE I.1.

DRS Implementation Rollout Period

Monitoring Cohort	Monitoring Cohort 1	Monitoring Cohort 2	Monitoring Cohort 3	Monitoring Cohort 4
Tracking Period for DRS Conditions	Deficiencies only from June 2009-Nov. 2011	Dec. 2011-Sept. 2012	Oct. 2012-Sept. 2013	Oct. 2013-Sept. 2014
Notified Must Compete	Dec. 2011	Jan. 2013	Feb. 2014	Dec. 2014
FOA Period	April-July 2012	July-Oct. 2013	Sept.-Nov. 2014	Oct.-Dec. 2015
Grant Winners Notified	April 2013	July 2014	May-July 2015	June 2016
New 5-Year Grants Begin	July 2013	July 2014	July 2015	July 2016

Source: Constructed by evaluation team through various announcements and press releases available on Early Childhood Learning and Knowledge Center (ECLKC).

Notes: Shaded cohorts are those included in the evaluation. Monitoring Cohort is a term used in this evaluation report to mean all of the grantees monitored in the time period indicated, including those designated for competition and those not designated for competition.

Two Key Conditions: Deficiencies and CLASS

Two of the seven conditions—deficiencies and low CLASS scores —have accounted for nearly all the designations for competition during the implementation rollout period: 99 percent in 2012; 98 percent in 2013; and 100 percent in 2014 (Reports to Congress on Head Start Monitoring FY 2014, FY 2013, and FY 2012). Therefore, these two conditions are the focus for the evaluation. Among grantees designated for competition in 2014, 36 percent were designated for deficiencies, 62 percent were designated due to CLASS, and 2 percent were designated for both reasons (Report to Congress on Head Start Monitoring FY 2014). We describe each condition below.

DEFICIENCIES

When grantees have a monitoring review, the review examines the grantee to assure they meet the Head Start Performance Standards in force at the time of the review. Reviews may focus on the whole program or on particular parts of the program based on the reason for the review. Violations may be noted in delegate sites, partner sites, subcontracted slots, etc. The grantee is responsible for all the slots and services they fund; regardless of who might be delivering the services, they are required to meet the same standards. (See the Report to Congress on Head Start Monitoring, Fiscal Year 2014 for more information, https://eclkc.ohs.acf.hhs.gov/hslc/data/rc/fy-2014-head-start-monitoring-report.pdf).

When grantees are assessed in this way, their program receives ratings of compliant, noncompliant, or deficient. Deficiencies result in designation for competition. Noncompliances do not result in competition unless the grantees do not correct the noncompliances within a certain period of time, in which case they become deficient according to the Head Start Act. According to the OHS Report to Congress, in FY 2014 only five percent of noncompliances were raised to a level of deficiency.

In the Head Start Act, a deficiency is defined as follows (Sec. 637. [42 U.S.C. 9832] (2):

> *"(A) A systematic or substantial material failure of an agency in an area of performance that the Secretary determines …. "(including threats to health and safety of children, denial of the roles and responsibilities of parents, failure to comply with early childhood development, health services, family and community partnerships, or program design and management, a misuse of funds, loss of legal status, and debarment from receipt of federal funds) "(B) systemic or material failure of the governing body of an agency to fully exercise its legal and fiduciary responsibilities;"*

or

> *"(C) an unresolved area of noncompliance."*

LOW CLASSROOM ASSESSMENT SCORING SYSTEMS (CLASS) SCORES

The Classroom Assessment Scoring System (CLASS; Pianta, LaParo, and Hamre 2008) measures the quality of teacher-child interactions and is one of the DRS conditions, as the Head Start Reauthorization Act (2007) required that teacher-child interactions be considered in identifying lower and higher-performing programs. The CLASS is one of the most widely used observational measures of the quality of preschool classrooms. Developed originally as a research tool (Pianta et al. 2005), CLASS has become widely used for research, monitoring, and professional development in early childhood education (Teachstone 2016). Different versions of the CLASS assess the quality of infant, toddler, preschool, and elementary school classrooms. Trained observers (i.e., raters certified as reliable by Teachstone) rate the preschool classroom on 11 items (labeled dimensions) in cycles that involve observation followed by rating. Typically, four cycles are collected and are recommended to obtain a reliable estimate of quality of that classroom (Pianta et al. 2008). The 11 items are averaged across cycles and grouped to form three domain scores: Emotional Support, Instructional Support, and Classroom Organization. Scores range from 1 (low quality) to 7 (high quality).

CLASS assessments for OHS monitoring are only performed in center-based 3- and 4-year-old Head Start classrooms. The CLASS assessments for OHS monitoring are scored at the grantee level, averaging by domain across the classrooms selected for assessment. When there are more than 11 eligible classrooms in a program, a randomly selected sample of classrooms are chosen for CLASS assessment; when there are 11 or fewer classrooms in a program, all classrooms are assessed (OHS

2016b). The DRS uses a two-cycle observation which allows more classrooms to be reviewed per grantee and was recommended by the CLASS developer at the University of Virginia to obtain a valid grantee-level score (OHS n.d.(b)). Grantees may be designated due to CLASS either because their scores fall below an absolute threshold for one or more the domains or because their scores fall in the bottom 10 percent of scores for their assessment cohort (see table I.2).[13]

TABLE I.2.

Designation Criteria of the CLASS Condition

CLASS Domain	Absolute Threshold Will be designated for competition if grantee score falls below:	Relative Threshold
Classroom Organization	3	Grantees scoring within the bottom 10 percent of CLASS scores in any of the three domains in their assessment cohort will be designated for competition.
Emotional Support	4	
Instructional Support	2	Except, if domain scores are 6 or 7, the grantee will not be designated for competition based on those scores, even if they are in the bottom 10 percent.

Source: Use of the Classroom Assessment Scoring System (CLASS) in Head Start (retrieved from ECLKC, February 21, 2016: http://eclkc.ohs.acf.hhs.gov/hslc/hs/sr/class).

Evaluation Goals and Purposes

Because the DRS represents a major change in the accountability structure for Head Start and has significantly changed the way the Office of Head Start (OHS) administers and manages the program, the Office of Planning, Research, and Evaluation (OPRE) within the Administration for Children and Families (ACF), U.S. Department of Health and Human Services, at the request of ACF leadership, commissioned an evaluation of the DRS to examine its implementation including how the system is meeting its goals of transparency, validity, reliability, and stimulating program quality improvement efforts. In the fall of 2012, OPRE contracted with the Urban Institute and its subcontractor, the Frank Porter Graham Child Development Institute at the University of North Carolina-Chapel Hill, to design and execute an independent evaluation of the DRS. The study purpose is to understand the mechanisms

---

[13] See the Report to Congress on Head Start Monitoring, Fiscal Year 2014 for more information, https://eclkc.ohs.acf.hhs.gov/hslc/data/rc/fy-2014-head-start-monitoring-report.pdf.

by which the DRS might support program quality improvement, ask whether the DRS appears to be identifying grantees providing lower quality services for designation for competition, and examine the results of grants competitions by the types of applicants that received awards and their efforts to improve services in the community. The study does not, however, look at whether Head Start program quality improved over time, and it cannot speak to whether DRS causes changes in program policies, actions, or quality. The goal of the evaluation is to inform future decision-making for continuous quality improvement. See chapter II for a discussion of the rationale for the questions that this early implementation evaluation did and did not address.

Road Map to the Report

We have written this report with several audiences in mind, including OHS administrators, Head Start practitioners, policy-makers, and other researchers. The chapters that follow include technical information about our methods, measures, and analyses. We also attempt to distill the study findings in a way that will be accessible to multiple audiences. Because we used a mixed methods approach to provide an integrated picture of the DRS, we present findings across multiple data collection efforts and multiple methods in single chapters. This volume primarily contains our rationale for conducting the study, an overview of our methods, and our findings. Additional information about our methods and the measurement instruments may be found in volume II, the technical volume. Volume I proceeds as follows:

- Chapter II describes our research questions, methods, and approach.

- Chapter III explores how grantees perceive the DRS and prepare for their monitoring and assessment in terms of efforts to improve quality using survey data and qualitative interviews.

- Chapter IV examines whether grantees designated for competition provide lower quality services than not-designated grantees based on the evaluation team's quality assessments. It also explores how the CLASS performs as a measure of program-level quality, as it is used for the purpose of DRS, through additional psychometric analyses.

- Chapter V explores how grantees perceive the monitoring, assessment, and DRS designation processes in terms of their efforts to improve program quality on a continuous basis using qualitative interviews.

- Chapter VI describes how much competition there was, what it looked like, and how awardees experienced the competition process using survey and administrative data as well as qualitative interviews.

- Chapter VII explores the negotiation, award, and transition processes in terms of preparing grantees to implement quality improvements to existing services using a survey and administrative data, as well as qualitative interviews.

- Chapter VIII summarizes and discusses findings and notes limitations of the evaluation.

Chapter II. Research Questions, Methods, and Approach

The purpose of this chapter is to describe the study research questions and our approach to addressing them, including the sample, measures, and methods. Technical volume II includes additional technical details on the study design.

Evaluation Approach

The evaluation of the Head Start Designation Renewal System (DRS) was designed based on an assumption that DRS has two primary mechanisms for improving quality: (1) incentivizing all grantees to improve quality in an effort to avoid designation for competition, and (2) in communities in which grantees are identified as providing lower quality services, introducing grant competitions through which applicants propose quality improvements to be competitive. The evaluation team sought to examine these mechanisms by describing the early implementation of the DRS in terms of how well the conditions of DRS differentiate between grantees providing lower and higher quality services, the types of quality improvement activities grantees engage in before or during the quality assessment process, the extent to which competition exists for Head Start grants, and the quality improvement efforts introduced through the competitive process. Specifically, the evaluation was designed to address five research questions:

1. **How do grantees perceive the DRS and prepare for their monitoring and assessment in terms of efforts to improve program quality?** How do grantees respond to the DRS conditions and prepare for monitoring and assessment in terms of their efforts to improve quality?

2. **Does the DRS differentiate higher versus lower quality programs?** Do grantees designated for competition score lower than not-designated grantees on measures of classroom (e.g., teacher-child interactions) and program quality? In addressing this question, we examined the two DRS conditions that led to almost all designations for competition (i.e., deficiencies and CLASS scores). We also explored additional psychometric properties of the CLASS as a measure of program-level quality, as it is used for the purpose of DRS.

3. **How do grantees perceive the monitoring, assessment, and DRS designation processes in terms of their efforts to improve program quality?** What is their perception of how the DRS

conditions are assessed and applied to their program? How do they perceive and respond to their designation status, once learned, and to what extent is quality improvement considered as they decide whether to compete?

4. **What does competition look like? How much competition is generated by DRS?** Who applies for Head Start grants associated with DRS? How do applicants respond in terms of proposing quality improvements?

5. **How do grantees experience the negotiation, award, and transition processes in terms of preparing them to implement quality improvements to existing services?** What are the outcomes of the competitive process? How do both incumbent grantees and new awardees perceive the negotiation and award processes in terms of preparing them to implement quality improvements to existing services?

To address these questions, the evaluation used a mixed-methods design that integrates quantitative observational assessments, surveys, and administrative data with qualitative interviews. In figure II.1 we display our overall approach to the evaluation. Additional details about the sampling frame, sample, measures, and methods by research question are described below and in technical volume II.

Input Into Designing the Study. The evaluation design was developed in collaboration with the project funders in ACF and ASPE, with input from stakeholders and technical advisers. A six-member interdisciplinary expert work group provided input on the study design, methods and measures, and then later on the approach to analyses and interpretation of findings. The evaluation team also conducted various outreach activities with the Head Start community, including webinars and conference presentations soliciting feedback on the study design. As a result of this process, the evaluation team chose to focus its efforts on the research questions we could address most effectively, with the most rigorous methods and measures, and that best meet the needs of ACF and the Head Start community for understanding the implementation of DRS. We include a discussion of the rationale for not answering the question, "Does the DRS improve quality?" and other design decisions at the end of this chapter.

FIGURE II.1

Early Implementation of the Head Start Designation Renewal System Study Approach

The Evaluation Investigated Two Mechanisms for Improving Quality through the Head Start Designation Renewal System

Incentivizing Quality Improvement through Monitoring and Assessment (RQ1, RQ2 and RQ3) See chapters III, IV and V	Introducing Competition to Improve Quality (RQ4 and RQ5) See chapters VI and VII

Sampling Frame: We drew from a sample of 368 Head Start grantees: 339 in Monitoring Cohort 4 (2013-2014) and 29 from a Monitoring Cohort 3 (2012-2013) supplement to assure a sufficient number of grantees designated for deficiency.	Sampling Frame: We drew from the 103 Funding Opportunity Announcements for designated grantees in 2014 (DRS Cohort 3).

RQ2: Differentiating Quality in Grantees Designated for Competition versus Not Designated (Quantitative)

Observational Quality Assessments
In 71 randomly selected grantees with randomly selected classrooms, we performed quality assessments aligned with the deficiency and CLASS conditions of the DRS. Collected February 2014-January 2015.

Extant data
- 2012 tax returns, IRS Form 990 data to assess financial vulnerability for 216 nonprofit grantees in the sampling frame
- Grantee OHS monitoring and CLASS scores were obtained for the 71 grantees participating in the quality assessments and the 216 nonprofit grantees with tax data.

RQ1 and RQ3: Understanding Grantee Quality Improvement Efforts and Perceptions

Professional Development and Technical Assistance Survey (Quantitative, RQ1 only) Survey to grantee directors (*N*=66) in the 71 randomly selected sites for the quality assessments. Collected February 2014-January 2015.

Director Telephone Interviews (Qualitative, RQ1 and RQ3) Semi-structured interviews with 35 of 71 grantee directors participating in the quality assessments. Collected June 2014-March 2015.

On-Site Follow Up Interviews (Qualitative, RQ1 and RQ3) Semi-structured interviews with Policy Council members, governing body members, program and agency directors, and program managers in 15 of the 35 grantees participating in Director Telephone Interviews. Collected April 2015-June 2015.

RQ4 and RQ5: Competition and Award

Applicant Survey (Quantitative, RQ4 and RQ5) Web-based survey of applicants for FOAs and awardees of grants; 120 of 182 applicants and 74 of 105 awardees responded. Collected December 2014-April 2015.

Awardee Interviews (Qualitative, RQ4 and RQ5) Semi-structured interviews with Policy Council members, governing body members, program and agency directors and program managers in 9 sites awarded grants that responded to the applicant survey. Collected October 2015-December 2015.

Extant data (Quantitative, RQ4 and RQ5)
-Census data on communities
-OHS data on competitor and awardee names and service areas
-OHS PIR data on characteristics and services for grantees designated for competition

Research Design

Two Sampling Frames

Two sampling frames were selected to support the study (figure II.1) because of the timing of the DRS processes in relation to the time period of the evaluation. This design allowed the sample for data collection to be drawn from the population of grantees experiencing the relevant phase of DRS (e.g., grantees experiencing monitoring and assessment for RQs1–3; grantees experiencing competition and award for RQs4–5).

The first sampling frame, supporting the examination of questions about the monitoring and assessment process (RQs 1–3), is drawn primarily from Monitoring Cohort 4 where grantees were receiving their Triennial Monitoring Review visit during the evaluation (between October 2013 and September 2014). Due to low numbers of grantees designated for deficiencies and higher than expected rates of refusal to participate in the study, Monitoring Cohort 4 did not include enough grantees to attain the desired sample size target for grantees designated for competition. Thus, the sampling frame was supplemented by grantees in Monitoring Cohort 3 (grantees receiving their Triennial Monitoring Review visit between October 2012 and September 2013). A total of 368 grantees (339 from Monitoring Cohort 4 and 29 from Monitoring Cohort 3) were included in the sampling frame. Grantees providing only Early Head Start services, only home-based services, only Migrant and Seasonal Head Start services, only American Indian Alaskan Native services, and those providing services outside of the continental United States are excluded from the evaluation for reasons explained at the end of the chapter.

The second sampling frame, supporting the examination of questions about the competition and award process associated with DRS (RQs 4–5), was based on DRS Cohort 3 and the 2014 Funding Opportunity Announcements (FOAs) announced in September 2014 and December 2014. It includes the population of the 105 grantees designated for competition in DRS Cohort 3, the 182 applicants for the resulting 103 FOAs, and the 105 awards for the FOAs. This includes designated grantees and applicants for Early Head Start services, Migrant and Seasonal Head Start services, and services provided outside the continental U.S. None of the grantees in DRS Cohort 3 provided American Indian Alaskan Native Head Start services and, thus, none were included in the frame.

RQ2: Examining How the DRS Differentiates Higher versus Lower Quality Grantees

Research question 2 asks: Does the DRS differentiate higher versus lower quality programs? The key design elements for how the evaluation examined this question are discussed here, with additional information available in technical volume II.

As discussed in chapter I, the DRS assesses grantee performance against seven conditions aimed at identifying which grantees are providing high quality comprehensive services and can receive automatic grant renewal; and which grantees are not providing high quality comprehensive services and will be subjected to open competition to receive renewed funding. To understand how well this mechanism of the DRS is differentiating between grantees providing higher versus lower quality services, the evaluation team conducted its own assessment of grantees' quality and assessed tax returns to examine fiscal vulnerability, and tested for differences between grantees designated for competition and those not designated.

Sample. Two samples were used to address this RQ: a random sample of 71 grantees and a review of the 2012 tax returns of 216 nonprofit grantees. Both samples were based on a single sampling frame. The evaluation team employed a sampling strategy to achieve several goals. First, we conducted power analyses to determine the size of the sample and subgroups needed to address the RQ. Those analyses indicated that we needed 70 grantees to participate in the study— half designated for competition and half not designated—to detect moderate to large differences in the quality assessments. Second, to examine differences in quality by reason for designation for competition (i.e., to look at designations for competition based on deficiency separately from those based on CLASS), power analyses indicated we needed 40 percent of our designated-for-competition sample to be grantees designated for deficiency to detect moderately large differences. Third, the evaluation team aimed to conduct data collection in close proximity to the OHS monitoring visit and before grantees knew whether they had been designated for competition or not. To ensure the sample included a sufficient number of grantees designated for competition without knowing the designation status of grantees in the sampling frame, the evaluation team constructed a statistical model to predict grantees' likelihood of being designated. Using data provided by OHS, grantees were classified as "likely to be designated" if their OHS monitoring CLASS scores were in the lowest 10 percent of the distribution[14] or if deficiencies were noted in preliminary reports from monitoring visits[15] (see technical volume II for additional details). The

---

[14] The evaluation team was receiving real-time information from the monitoring assessments about grantee CLASS scores and deficiencies. At the start of the evaluation, we set the benchmark at the previous year's lowest 10 percent of the distribution. Periodically during sampling, we revised the benchmark based on the data coming from the field.

[15] Some grantees changed categories when deficiencies noted during monitoring were dropped when reviewed by OHS or when it was clear that the bottom 10th percentile was higher for Monitoring Cohort 4 than it had been for Monitoring Cohort 3. No

sample was drawn from the first sampling frame, which included 339 grantees in Monitoring Cohort 4 (grantees who had recently experienced a monitoring visit) and supplemented by 29 grantees in Monitoring Cohort 3 (to ensure sufficient representation of grantees designated for competition based on deficiencies).

A random sample of Head Start grantees was selected and stratified based on grantees' likelihood of being designated for competition, as well as by region and grantee size. The result was a stratified random sample of 71 grantees: 35 designated for competition, 36 not designated. The 35 grantees designated for competition included 14 designated for deficiency and 24 designated due to low CLASS scores—with three designated for both deficiencies and low CLASS scores. The evaluation team saw 75 percent of the sampled grantees within 1.5 months of when the OHS monitoring data were collected. For 10 percent of the sample, the time between the OHS assessments and the assessments by this team was a year or more. This was largely due to the inclusion of grantees from Monitoring Cohort 3.

The 46 percent acceptance rate was lower than we anticipated. Comparisons of grantees that did and did not agree to enroll in the study indicated there were significant differences within designation categories by region, child to teacher ratios, proportion of teachers with a BA, program type, and presence of delegates.[16] Sampling weights for each designation category (not designated, designated due to CLASS, designated due to deficiencies) were developed to correct for these differences. The weights were applied in all analyses (see technical volume II, appendix B for details).

Once grantees agreed to participate, we drew a stratified probability sample of classrooms serving 3- to 5-year-olds in centers within the 71 grantees. We used an algorithm similar to the algorithm used by the OHS monitoring team for randomly sampling classrooms based on the number of eligible classrooms within a grantee: selecting all classrooms in grantees with 8 or fewer classrooms, and randomly selecting 8 classrooms for grantees with 9 to 42 classrooms, 9 classrooms for grantees with 43 to 100 classrooms, and 10 classrooms for grantees with more than 100 classrooms. The resulting sample had 549 classrooms in 375 centers.

Recruitment occurred from January 2014 to October 2014. To encourage participation, incentives were provided at $25 for each teacher and $50 for each participating Head Start center within a grantee up to $500 (10 centers). Incentives (gift cards) were given either on-site after data collection or mailed from FPG.

---

differences were detected when quality measures for the grantees that changed from the "designated" to "not-designated" category were compared with grantees in the "not-designated" category.
[16] See table B.7 in appendix B for detailed results of this nonresponse bias analysis.

In addition, the 2012 tax returns of the nonprofit grantees in the sampling frame were examined using Tuckman and Chang (1991) financial vulnerability ratios. Of the 368 grantees in the sampling frame, 248 (67 percent) are nonprofit organizations as indicated by their filing an Internal Revenue Service (IRS) Form 990. Missing data on 32 forms led to a sample of 216 nonprofit grantees. Data from IRS Form 990 was obtained for only the nonprofit grantees because the Tuckman and Chang (1991) method can only be applied to the tax returns for nonprofit organizations and there is no standard method that could be applied to other types of organizations.[17]

Quality Assessments. To assess how well the DRS differentiates between higher versus lower quality Head Start grantees, the evaluation team selected quality assessment tools that closely aligned with the DRS conditions and Head Start Performance Standards used at the time of the evaluation. As noted in chapter I, two of the seven DRS conditions – deficiencies and low CLASS scores – have accounted for nearly all the designations for competition to date and are, thus, the focus of this evaluation. Table II.1 shows how six separate quality constructs tapped by these two DRS conditions have been measured in the evaluation. Data collection included observational measures of classroom quality, as well as structured interviews of center directors regarding the quality of parental involvement, staff qualifications, and governance and management. All of these data (except for the data from IRS Form 990 measuring financial vulnerability) were collected during site visits between February 2014 and January 2015. Site visits lasted up to five days and involved at least two data collectors. The number of data collectors and days on site depended on the number of classrooms selected for observation.

---

[17] The nonprofit grants constituted 67 percent of the entire sampling frame (*N*=368), but they were less likely to be designated for competition (20 percent) than were Head Start grantees of other organization types (33 percent).

TABLE II.1.

Evaluation Measures Used to Assess Program Quality

Quality Construct	DRS Condition	Study Measure
Child Health & Safety	**Monitoring Deficiency**	**Health and Safety Checklist:** The Health and Safety Checklist was adapted from NAEYC's Childcare Health and Safety Checklist (Aronson 2012) and The California Childcare Health Program Health and Safety Checklist.
Family & Community Engagement	**Monitoring Deficiency**	**Program Administration Scale (PAS) Family Partnership Subscale:** The PAS includes a subscale that measures family and community engagement.
Child Development & Education	**Monitoring Deficiency**	**PAS Child Assessment Subscales and Report on Curriculum:** Includes screening and identification of special needs and of assessment in support of learning.
Management, Operations & Governance Systems	**Monitoring Deficiency**	**PAS Governance and Management Subscales:** Includes Program Planning & Evaluation, Center Operations, Human Resource Development, Marketing & Public Relations, and Personnel Cost & Allocation Scales.
Fiscal Integrity & Vulnerability	**Monitoring Deficiency**	**PAS Fiscal Management Subscale and Tuckman & Chang Financial Ratios:** PAS scale includes questions about budget management and accounting practices. Tuckman and Chang (1991) method uses IRS Form 990 data for nonprofit organizations to calculate financial vulnerability ratios.
Classroom Quality	**Monitoring Deficiency**	**PAS Staff Qualification Subscales:** Includes questions about education and training of all staff members.
Classroom Quality	**Low CLASS Emotional Support**	**ECERS-R Interactions Score:** The ECERS-R is a widely used measure of the quality of classroom environment. Included in the Interaction Score are 10 items measuring the quality of the interactions between the teachers and children during the morning of observation.
Classroom Quality	**Low CLASS Instructional Support**	**ECERS-E Academic Activities Score:** The ECERS-E was developed to extend the ECERS-R by describing academic activities in the classroom. The score is based on four items reflecting academic activities during the morning of observation.
Classroom Quality	**Low CLASS Classroom Organization**	**TSRS Adapted Classroom Structure and Management Subscale**: The scale includes three items that are observed in four cycles during the morning of observation.

Note: PAS is the Program Administration Scale. It has multiple subscales that align with various constructs.

The evaluation team selected Health and Safety Checklists, the Program Administration Scale (PAS) (Talan and Bloom 2011), and IRS Form 990 as measures that reflect Head Start Performance Standards that could result in deficiencies. The Health and Safety Checklist and the PAS Subscales were administered to center directors where classrooms were randomly chosen (375 centers) to participate in the study. The PAS is a structured interview consisting of 25 items grouped into 10 subscales that measure leadership, management, and administrative practices of center-based early childhood programs with subscales related to: parent involvement, child assessment, staff qualifications, governance and management, and fiscal operations. PAS was only collected on 135 of the 375 centers because PAS data collection was restricted to one center per grantee during the second half of the data collection after grantees expressed concerns about the time center directors had to spend preparing for the PAS interview. The protocol change was made when centers (and thereby their classrooms) were refusing to participate, further exacerbating concerns about response rates relating to observed classroom quality.

The evaluation team also constructed financial vulnerability ratios based on the Tuckman and Chang (1991) method using data from tax filing of Form 990 for nonprofit programs. The four sets of ratios measure nonprofit organizations' financial integrity based on their assets, diversity of revenue sources, administrative expenses, and operating margins. More information about how the ratios were calculated is provided in technical volume II, appendix B.

The evaluation team selected the ECERS-R, ECERS-E, and TSRS as measures that aligned with the CLASS condition of DRS. We observed 549 classrooms using the Teaching and Interaction score (Pianta et al. 2006) from the Early Environment Rating Scale-Revised (ECERS-R; Harms, Clifford, and Cryer 1990) to reflect teacher sensitivity as measured by the CLASS Emotional Support domain, the Academic Activities Scale from the Early Environment Rating Scale-Extension (ECERS-E; Sylva et al. 2006) to reflect teacher sensitivity as measured by the CLASS Instructional Support domain, and the Teacher Styles Rating Scale (TSRS Adapted; Raver et al. 2012) Classroom Structure and Management Subscale to reflect classroom management as measured by the CLASS Classroom Organization domain. In addition, the evaluation team conducted its own CLASS[18] observations to allow us to compare grantee-level CLASS scores collected by the evaluation team and OHS.

---

[18] The Classroom Assessment Scoring System (CLASS; Pianta, La Paro, and Hamre 2008) provides an assessment of the quality of teacher-child interactions. Its ten dimensions are organized into three domains: Emotional Support, Classroom Organization, and Instructional Support.

RQ1 and RQ3: Understanding Grantee Quality Improvement Efforts and Perceptions

Research question 1 asks: How do grantees perceive the DRS and prepare for their monitoring and assessment in terms of efforts to improve quality? Research Question 3 asks: How do grantees perceive the monitoring, assessment, and DRS designation processes in terms of their efforts to improve program quality? The key design elements of how the evaluation addressed these questions are discussed here, with additional information available in technical volume II, appendix C.

The DRS is designed to encourage all grantees to improve their quality in an effort to avoid designation for competition and receive automatic renewal of grant funding at the end of the 5-year grant cycle. Thus, research questions 1 and 3 investigate: the activities grantees were undertaking to improve quality before and after their designation status was known; how they understood, perceived, and experienced the DRS conditions and processes; and how those experiences may have shaped grantees' efforts to improve quality in response to DRS. It is important to understand both the actions the grantees are taking and the circumstances they are responding to when taking them. Research question 1 was addressed through a mixed methods approach drawing from a quantitative Professional Development and Technical Assistance Survey and two qualitative interview components (the program director telephone interview and on-site follow up interviews) described here. Research question 3 was exploratory in nature and drew from the qualitative interviews with directors and other program leaders.

Professional Development and Technical Assistance (PDTA) Survey. The PDTA Survey asked directors about the types of professional development and technical assistance activities undertaken by grantees in the 12 months before their first monitoring visit applying the DRS conditions. It was completed by 66 directors in the 71 grantees participating in the quality assessments. Directors were asked about professional development or technical assistance activities related to each of the quality constructs listed in table II.1, who participated in the activities, and what type of activities were provided. The PDTA survey was developed by the evaluation team who used the Early Childhood Education Professional Development, Training and Technical Assistance Glossary jointly produced by the National Association for the Education of Young Children (NAEYC) and the National Association of Child Care Resource and Referral Agencies (NAEYC and NACCRA 2011) to frame the PDTA. This survey was administered to the directors as a structured interview during the quality assessment site visits (described under RQ1 above) between February 2014 and January 2015.

The Director Telephone Interviews and On-Site Follow Up Interviews involved a qualitative approach to better understanding how grantees responded to the DRS in terms of specific activities and operational decisions that might support quality improvement, as well as how their perceptions of DRS might explain those activities and decisions (see table II.2). OHS has long encouraged and supported Head Start grantees to strive for ongoing quality improvement through policy and regulation, training and technical assistance resources, and quality improvement funds. This line of inquiry explores whether grantees are motivated to improve quality specifically because of the DRS in an effort, for example, to avoid the reputational effects of designation or the work of participating in competition. Because of the exploratory nature of these questions, the evaluation team determined that a qualitative, in-depth interview approach with a small number of sites was the most appropriate design.

We conducted interviews in two stages to focus first on the responses of the Head Start directors, and then on the responses of other Head Start and organizational leadership staff, Policy Council members, and governing body members. Head Start directors were the focal point for data collection because previous early care and education research has indicated that these individuals may be the "gatekeepers of quality" (Bella and Bloom 2003), and in Head Start the individual with the title of Head Start director typically has the primary responsibility for assuring the quality of the program. In addition, research by Rohacek and colleagues (2010) indicates that the early care and education directors' belief systems, including their beliefs about standards applied to their programs, shapes efforts to improve quality. Because Head Start programs may have many management actors and leaders contributing to decisions about improving program quality, program managers, governing body members, Policy Council members, and other leaders were identified for on-site interviews following the director interviews.

Director Telephone Interviews. The evaluation team conducted semi-structured interviews with 35 Head Start directors in 35 grantees (18 designated for competition, 17 not designated) that had participated in the PDTA survey. The intention was to collect these data prior to grantees knowing their designation status, but just over half may have known their designation status: 8 because they had received deficiencies and 11 because interviews were conducted following the public announcement of designations (6 not designated and 5 designated due to CLASS). On average, Director Telephone Interviews were conducted six months following monitoring. They occurred after the evaluation's quality assessment visits had been conducted. In the interview, we did not ask if they knew their designation status, so we do not know definitively who knew and who did not know their status. The grantees were purposively selected to represent designation status, region, and size of grantee. Grantee directors were queried about their perceptions of the DRS prior to and following monitoring,

their understanding and concerns about DRS conditions and processes, and their efforts to improve program quality before monitoring. Participants were recruited (May 2014—January 2015) and interviewed (June 2014—February 2015) following their site visit for the quality assessment and PDTA Survey. No incentives were offered for this portion of the study. The sample for these interviews was not representative of all Head Start directors and overrepresented grantees designated for competition to ensure a wide variety of perspectives could be explored and to allow for comparisons between grantees designated for competition and those not designated.

On-Site Follow -Up Interviews. The evaluation team conducted semi-structured interviews of key personnel in 15 of the 35 grantees that participated in the Director Telephone Interviews, including 7 grantees designated for competition and 8 not designated. Interviews were conducted during site visits that occurred after the OHS announcement of the Monitoring Cohort 4 grantees designated for competition. The 15 grantees were purposively selected to represent designation status, region, size of grantee, and the diversity of perspectives shared in the 35 telephone interviews. Grantees were recruited (March—May 2015) and interviewed (April—June 2015) following the Director Telephone Interviews. No incentives were offered for participation in this part of the study. The purpose of these interviews was to gain the perspectives of key stakeholders within Head Start grantees to supplement the information provided by the grantee director in the Director Telephone Interview. Interviews were conducted with staff, including nine agency directors, 17 program directors, 41 program managers, and 61 Policy Council or governing body members. As with Director Interviews, the sample for these interviews was not representative of Head Start managers or leadership and overrepresented grantees designated for competition to ensure a wide variety of perspectives were captured and to allow for comparisons between grantees designated for competition and those not designated.

TABLE II.2.

Data Collection Instruments Used to Assess Grantee Perceptions, Experiences, and Quality Improvement Efforts

Construct	Instrument Description
DRS Implementation and Incentives: ■ Grantee perceptions and experiences of the DRS ■ Grantee actions to improve quality (See technical volume II, appendix E for more information about the construction of the instruments.)	**Semi-structured phone interviews with Head Start directors:** Selected Head Start directors were interviewed (n=35, half designated for competition) and were asked about their understanding of the DRS, perceptions and experiences with the DRS, expectations related to possible competition, and DRS-related actions. **Semi-Structured On-site Implementation Interviews:** On-site visits to 15 Head Start grantees (half designated for competition) included semi-structured interviews with program managers, executive directors, governing body members, and Policy Council members that sought their understanding, perceptions, experiences, and expectations regarding the DRS.

RQ4 and RQ5: Examining How the DRS Introduces Competition to Improve Quality

Research question 4 asks: What does competition look like? How much competition is generated by DRS? Research Question 5 asks: How do grantees experience the negotiation, award, and transition processes in terms of preparing them to implement quality improvements to existing services? The key design elements of our approach to addressing these questions are discussed here, with additional information available in technical volume II, appendix D.

The DRS introduces a form of competition that Kincaid (1991) refers to as *mediated competition* — competition initiated and decided through government institutions rather than through the market. Previous research suggests that mediated competitions generate little competition, but facilitate the formation of collaborative community partnerships or increase access to additional resources (Hefetz and Warner 2011; Warner and Hefetz 2003). Thus, in research question 4, we examine the extent to which the DRS generates competition in communities, who competes, the ways that grant applicants propose to improve quality relative to the designated grantee, and through qualitative interviews we explore how competition affects partnerships and collaborations. We also examine predictors of FOAs with more or less competition and explore features of the competition process that may have contributed to competition levels through qualitative interviews. In research question 5, we examine the outcome of the competitions and compare the services provided by the existing grantee (using PIR data) to the services proposed by the awardee in relation to differences in structural elements of quality. Through qualitative interviews, we also explore how the negotiation and award processes may

support or impede implementation of proposed changes to services that resulted from the DRS process. Inquiry into research questions 4 and 5 used a mixed-methods approach involving an applicant survey, qualitative interviews with grant awardees, and extant data described below.

Extant Data on the Grantees Designated for Competition in 2014 and Applicants for Competed Grants. The evaluation team obtained OHS administrative data, PIR data, Census data and monitoring data for 105 grantees designated in 2014 (DRS Cohort 3), the 182 applicants (including 54 percent who were incumbents)[19] for the competition, and 105 awardees for the competition.[20] The OHS administrative data provided a list of applicants and awardees for each FOA. We examined 2013 PIR data to determine whether applicants had recent experience with Head Start either as a Head Start grantee or delegate in the same service area or a different service area. The PIR describes the characteristics and services of all funded grantees and was accessed to describe the grantees designated for competition and any applicant that was already a Head Start grantee. We focus on PIR data to describe the current enrollment, number of teachers, and qualifications of teachers. The census data provides community level characteristics on the service areas of Head Start, including population, region, and density. The monitoring data provides information on why grantees were designated for competition. Websites were accessed to supplement these data sources when information was missing.

Applicant Survey. The Applicant Survey was developed by the evaluation team to assess the proposed services of applicants to the competitive FOAs studied in this evaluation. The survey was developed to collect parallel information as is available in the PIR on Head Start grantees. We recruited respondents between December 15, 2014 and April 30, 2015. No incentives were offered for this portion of the study. The survey had 16 closed-ended questions about the applicant's proposed services, including program options, enrollment, teacher qualifications, program schedule, and partnerships. All applicants (*N*=182) responding to the 2014 competitive FOAs were invited to participate.[21] A total of 120 applicants completed the survey (66 percent response rate).[22] In some analyses, only 118 surveys are used due to missing data. Data from the Applicant Survey and OHS PIR were used together to compare the characteristics and services offered by the DRS Cohort 3 grantees

---

[19] We only have data on 177 of the 182 applicants.
[20] These included data on one Early Head Start grantee designated for competition and three applicants for that FOA; two Migrant and Seasonal Head Start grantees designated for competition and four applicants for those FOAs; and two FOAs issued for grantees providing services outside the continental United States and the three applicants for those FOAs.
[21] This could include applicants for Early Head Start only, Migrant and Seasonal Head Start only, home-based only services and applicants for FOAs issued for service areas outside the continental U.S. There were no FOAs for American Indian/Alaskan Native services because these services are not subject to the DRS.
[22] This included two applicants for Early Head Start services, four applicants for Migrant and Seasonal Head Start services, and one applicant for services to be provided outside the continental U.S.

designated for competition to those proposed by applicants (for RQ4) and to compare the characteristics and services proposed by incumbents and awardees (RQ5).

Awardee Interviews. The Awardee Interview protocols were designed by the evaluation team based on competition research (Hefetz and Warner, 2011; Warner and Hefetz, 2003) and in consultation with expert work group members. The interview guides were tailored for the type of person interviewed and for whether the organization was the incumbent organization who won back their grant or a new organization obtaining the grant. This was a qualitative component of the study and, thus, we purposively sampled using observable characteristics that may create different organizational experiences and different responses as the basis for the purposive site selection. These characteristics included incumbent or new awardee status, whether the new awardee was new to Head Start or just new to the grant, region of the country, ACF enrollment of the incumbent organization, type of agency, urbanicity, number of applicants for the FOA, FOA type (e.g., Head Start only, Early Head Start only or Birth to Five), and proposed service mix as indicated on their Applicant Survey response. We recruited participants from August to October 2015 and collected data from October to December 2015. No incentives were offered in this portion of the study. Interviews were conducted in nine sites (four incumbents; five new awardees), selected from the 120 grantees that responded to the Applicant Survey who were awarded grants as a result of the competitive FOAs for DRS Cohort 3.[23] Interviews were conducted with grantee staff and stakeholders including nine agency directors, 17 program directors, 16 program managers, and 17 Policy Council or governing body members. New awardees are overrepresented in the sample to assure that diverse perspectives across multiple new awardees and incumbents may be captured.

Table II.3 summarizes the data sources and sample by research question. All research questions are addressed using data from more than one source, and most questions were addressed using both qualitative and quantitative methods and supplemented with extant data.

---

[23] Early Head Start only, Migrant and Seasonal Head Start only, home-based only services and grantees outside the continental United States were excluded.

TABLE II.3.

Summary of Research Design: Research Questions, Sampling and Recruitment, and Timing of Data Collection, by Data Type and Source

	Quantitative Primary Data Collection			Qualitative Primary Data Collection			Extant Data	
	Quality Assessments [a]	Professional Development/Technical Assistance Survey	Applicant Survey	Director Telephone Interviews	On-site Follow Up Interviews	Awardee Interviews	Financial Vulnerability – IRS Form 990	2014 Applicants -OHS Administrative Data, Census, and Grantee Website Data
Research Question								
1. How do grantees perceive the DRS and prepare for monitoring and assessment in terms of efforts to improve quality?		X		X	X			
2. Does DRS differentiate higher versus lower quality programs?	X						X	
3. How do grantees perceive the monitoring, assessment, and DRS designation processes in terms of efforts to improve program quality?				X	X			
4. What does competition look like? How much competition is generated by DRS?			X			X		X
5. How do grantees experience the negotiation, award, and transition processes in terms of preparing to implement quality improvements?			X			X		X
Sampling and Recruitment								
Sample	71 grantees from 2014 Monitoring Cohort[a]	66 of the 71 grantees from Quality Assessment	120 applicants to 2014 FOAs	35 of the 71 grantees from Quality Assessment	15 of the 35 grantees from Director Telephone Interviews	9 of the 120 grantees from Applicant Surveys	216 nonprofit grantees from 2014 Monitoring Cohort*	105 grantees designated for competition in DRS Cohort 3; 182 Applicants to 2014 FOAs
Response Rate [b]	46%	44%	66%	na	na	na	87%	na
Data Collection Period	2/2014 – 1/2015	2/2014 – 1/2015	12/2014-4/2015	5/2014 – 2/2015	4/2015 - 62015	10/2015-12/2015		

[a] Monitoring Year 2014 is the focal cohort, but sample is supplemented with 29 grantees from Monitoring Year 2013.

[b] Response rate is presented at the grantee level. See technical volume II, appendix B for information about response bias analysis, response rates for individual instruments, and weighting to account for nonresponse.

Data Collector Hiring and Training

This evaluation used two different types of data collectors to assure that data collectors were well-trained and suited for the types of data collection they were engaged in. The Frank Porter Graham Child Development Institute at the University of North Carolina at Chapel Hill (FPG), the subcontracted partner in this evaluation, hired, trained, and supervised the data collectors for the quality assessments and PDTA survey as described below. The remaining data collectors were employees of the Urban Institute and were trained and supervised by the project principal investigator at the Urban Institute as described below.

Quality Assessments

The Frank Porter Graham Child Development Institute at the University of North Carolina at Chapel Hill, the subcontracted partner in this evaluation, hired, trained, and supervised the data collectors for the observations of program quality that occurred during the same site visits to programs. FPG recruited staff from around the country to minimize travel time and cost during data collection. To avoid potential conflicts of interest, the evaluation's data collection staff were not recruited from the pool of OHS monitoring reviewers or OHS CLASS observers. Half the data collectors were trained on the Environment Rating Scales (ECERS-R, ECERS-E) and the Child Care Health and Safety Checklist and half were trained on the Teacher Style Rating Scale (Adapted TSRS), and the CLASS. A third of data collectors were already certified Program Administration Scale (PAS) assessors and were identified through the McCormick Center for Early Childhood Leadership.

Data collector training started with a group meeting before data collection commenced to orient staff to the purposes of the study, the procedures required of them, and ethical principles of assessment and data handling, followed by specific training on the measures. The principal investigators from both the prime contractor (Urban Institute) and the subcontractor (FPG) kicked off the meeting to explain the study, its importance, and to emphasize the importance of their roles. Data collectors were briefed on the provisions for the protection of human subjects approved by the IRB, including procedures for informed consent, confidentiality and privacy (including the legal requirement to report abuse or neglect and applicable procedures), and data security. Data collectors were trained to omit identifying information from any notes, even if respondents used identifying information in response to questions. Senior subcontractor staff experienced in both primary data collection and supervision of data collectors provided this initial training.

Each of the observation assessment instruments selected for the study has a standardized training process designed and typically facilitated by the developers of the scales. For the ECERS-R and ECERS-E, as has been done in previous studies, an experienced member of the evaluation team was trained to reliability by the authors of the measure. This staff member then provided the training for the data collectors and was the anchor for reliability. Each data collector met a reliability criterion of 80 percent agreement, calculated by dividing the number of items that were within one scale point of the trainer (i.e., an experienced data collector who had been trained by the developer or one of their official representatives) by the total number of items. For example, if the data collector is within one scale point of the trainer's score on 20 out of the 22 items, the reliability score would be 90 percent. Certification of data collectors was based on observer's reliability scores, as well as the measure trainer's qualitative evaluations of each observer. Additional days for reliability were scheduled for individuals who did not meet the reliability standard after the initial scheduled days. These procedures are similar to procedures that the study subcontractor used in previous studies, such as the Quality Initiatives for Early Care and Education and the National Center for Early Development and Learning. In conjunction with this training, the consultants were also trained on the Child Care Health and Safety Checklist. Data collectors practiced using the checklist and were certified with the master coder.

For CLASS, a trainer from Teachstone provided a two-day CLASS observation training to prepare data collectors to take the test for the CLASS Observer Certification. Data collectors reviewed materials before reliability testing to increase proficiency and accuracy. Data collectors had 30 days after training to take the test. Once data collectors pass the CLASS reliability test, they are certified for one year. For data collectors who were already certified in CLASS, the project coordinator confirmed their certification. Training for the Adapted TSRS followed standard procedures developed for previous studies with similar reliability procedures as the CLASS observation.

Data collectors hired to conduct the PAS had all previously attended a four-day reliability training provided by the McCormick Center for Early Childhood Leadership. This intensive training provides an overview of the reliability and validity of the PAS, rating indicators and scoring items, interview protocol for collecting data, verifying documentation, and establishing and maintaining reliability. Certification is valid for two years.

The data collectors who administered the PAS to the grantee directors were also trained to administer the PDTA survey. In the training, the general content of the survey was reviewed, and each data collector practiced administering the survey using the computer-assisted interviews on their laptops.

Understanding Grantee Quality Improvement Efforts and Competition and Award

Staff at the Urban Institute carried out the remaining data collection strategies. Before participant recruitment for each data collection strategy, the principal investigator at the Urban Institute convened a training of all staff who would be involved in the recruitment and data collection for that component of the study. All staff were required to attend each kick-off training even if they had participated in a previous study component. In each training, the principal investigator from Urban explained the evaluation, its importance, and emphasized the importance of the staff roles. The training included recruitment procedures and principles, provisions for the protection of human subjects approved by the IRB, including procedures for informed consent, confidentiality and privacy (including the legal requirement to report abuse or neglect and applicable procedures), and data security. The training also included a walk-through of the instruments that would be used for that portion of the data collection. For the semi-structured interviews, we talked about the purpose and meaning of each question, and when and what to probe on. We discussed the types of people we would be interviewing and their expected roles within Head Start. We also discussed the stage of the DRS process they would be experiencing at the time of the interviews, the magnitude of the change that the DRS represented for programs, and sensitivities related to DRS.

Considering Design Options

The evaluation team in collaboration with federal stakeholders and expert workgroup members considered several design options for addressing the questions of interest. We provide a brief overview of those considerations here including the decision not to evaluate whether quality improved because of the DRS, the types of Head Start programs included in the evaluation, and the cohorts on which to focus the evaluation.

Examining Changes in Program Quality Over Time. We discussed the possibility of examining program quality over time to determine if the DRS is improving program quality. We did *not* choose that evaluation design because (1) it would have been premature to examine if the DRS were improving quality over time without first assessing how the mechanisms for improving quality in the DRS were implemented and if grantees were making efforts to improve quality in response to the DRS; (2) research designs that would permit us to draw causal conclusions about the effect of DRS on program quality were not feasible because all Head Start grantees are subject to DRS (i.e., no available control or comparison group); and (3) results from nonexperimental, descriptive analyses of changes in Head Start

program quality over time would be available through other research projects already underway and, thus, available on a shorter timeframe than could be accomplished by the evaluation team.

Types of Head Start Programs to Include. In the design exploration, we also discussed the focus of the evaluation in terms of the types of Head Start programs or delivery designs that should or should not be included. We decided not to include Early Head Start-only programs, home-based-only programs, or Migrant and Seasonal Head Start (MSHS) programs in examining how DRS incentivizes quality improvement through monitoring and assessment (research questions 1—3). Early Head Start-only programs and home-based-only programs were excluded because they are not assessed using the CLASS and there are different monitoring protocols for Head Start and Early Head Start grantees. Classrooms serving children younger than age three and home-based programs would also require use of different assessment instruments by the evaluation team and a substantial increase in the sample size. Similarly, we did not include MSHS programs in this part of the study because of measurement limitations (i.e., the validity of quality measures for use with this population is unknown). We included Early Head Start, home-based-only and MSHS programs in the examination of the competition and award processes (research questions 4—5) because the mechanisms for competition and award are the same and measurement issues were not a barrier to inclusion in this part of the study. American Indian/Alaskan Native Head Start grantees are not included in any portion of the evaluation because they are subject to a different process in DRS (i.e., tribal programs participate in government-to-government consultation to establish a quality improvement plan before being designated for competition). Grantees located outside the continental United States were excluded because of the costs of traveling to those locations.

Focal Cohorts for Sampling. In the study design phase, we discussed the sampling frame in terms of the pros and cons of studying grantees in one cohort or across multiple cohorts. Our primary sampling goal was to ensure data for the evaluation were collected in close proximity to when participating grantees were experiencing the stage of DRS under study (i.e., monitoring and assessment, competition, award). For example, we sought to gather our quality assessments and grantee perceptions of the process close to when OHS monitoring occurred and to assess the competitive process soon after award of new grants. At the time we were designing the evaluation, Monitoring Cohort 3 (i.e., grantees monitored October 2012—September 2013) was already in their monitoring and assessment process, but if we focused solely on Monitoring Cohort 4 (i.e., grantees monitored October 2013—September 2014), the data collection period would have to span at least three years (from the start of a monitoring year for a given cohort through award and implementation of new grants). A focus on DRS Cohort 3 (i.e., grantees designated for competition in the FOAs released in September and November 2014) to study

the competition and award process and Monitoring Cohort 4 to study monitoring and assessment allowed us to limit the bias that occurs in time lags for both quality assessments and respondent recall. It should be noted, however, that for a portion of the sample, there was a substantial time lag between grantees' monitoring visits and the evaluation's quality assessments. Due to a lower than anticipated rate of response among grantees in Monitoring Cohort 4, the sampling frame for the quality assessments was supplemented with 29 grantees from Monitoring Cohort 3 to ensure sufficient representation of grantees designated due to deficiencies. For these 29 grantees, the time lag between OHS's monitoring and the evaluation team's quality assessment averaged 1 year, compared to an average of 1.2 months for grantees in the rest of the sample. However, follow up analyses suggested the time-lag between assessments was not related to differences in scores and bias associated with response rate was addressed using weights (see technical volume II, appendix B).

Using Predictive Sampling. Because the evaluation team sought to conduct its assessments of program quality before designation status was known, while ensuring our sample included an equal number of grantees designated and not designated for competition, we used preliminary monitoring findings to statistically predict which grantees were likely to be designated for competition and which were not in our sampling procedure. For example, grantees in monitoring cohort 4 received their triennial monitoring reviews between October 1, 2013 and September 30, 2014. Designation status for grantees in that cohort was announced in December 2014. If we had waited for designation status to be announced, the evaluation data collection would have occurred more than a year after OHS assessments for most grantees. OHS agreed to give the evaluation team access to the monitoring and assessment data as it was available from the field and before it underwent OHS's review process to expedite the sampling and recruitment processes. This allowed us to conduct our quality assessments closely behind the OHS monitoring visits (see technical volume II for more information about this predictive sampling process).

Analytic Approaches

We use several analytic approaches in this evaluation, varying by the type of data collected and the purpose for which it was collected. We provide an overview of those analytic approaches here with additional detail about the approach for each research question provided in chapters III—VII and in technical volume II.

Observational Quality Assessments

Before conducting analyses to answer research question 2 (i.e., whether there are differences in the evaluation team's quality rating between grantees designated for competition and grantees not designated, as well as additional psychometric analyses of CLASS scores), we first conducted response bias analyses to examine the effect of low response rates on our sample. We then constructed sampling weights to account for differential response rates which were used in all analyses using the quality assessment data. Comparisons of grantees that did and did not agree to participate were conducted on characteristics of the grantees and their monitoring CLASS scores within recruitment groups (not-designated, designated for deficiencies, and designated for low CLASS), and sampling weights were created to adjust for significant difference that emerged. Details on the response bias analyses and sample weight construction are available in technical volume II, appendix B. We also used multiple imputations[24] to account for missing data at the center level on the Program Administration Scale (PAS).

We then computed grantee-level means for all quality measures and compared the grantee-level quality scores of grantees designated and not designated for competition using *t*-tests. We also used *t*-tests to examine differences by the reason for designation (i.e., whether there were differences between grantees designated for competition based on deficiencies versus not designated, and then grantees designated for competition based on CLASS versus not designated). See chapter IV for results from these analyses. Finally, we conducted sensitivity analyses to examine power to detect differences when data were analyzed at the classroom or center level using multi-level models. See technical volume II, appendix B for results from these analyses.

Finally, we conducted additional psychometric analyses to better understand how the CLASS performs as a measure of grantee-level quality and whether features of the way CLASS is implemented in DRS may influence scores. Specifically, we examined four issues: (1) using *t*-tests and correlations, compared grantee-level CLASS scores from OHS observers and from the evaluation team to examine stability across time and rating teams in grantee-level CLASS scores; (2) estimated the proportion of variability in CLASS scores attributable to differences among observers and among grantees to examine rater bias,[25] (3) used paired *t*-tests to determine whether differences in the number of observation

---

[24] *N*=40 imputed datasets.
[25] Ideally, we would have also looked at variability within center, but there were too many centers with a single classroom because classrooms, not centers, were randomly selected. It was not possible to estimate center-level variances.

cycles (i.e., two cycles versus four) was associated with differences in CLASS scores; and (4) used *t*-tests to examine whether sampling classrooms within grantees, rather than conducting CLASS assessments with all classrooms, was related to differences in CLASS scores. Results from these analyses are reported in chapter IV.

Financial Vulnerability Data

The evaluation team obtained tax data from IRS Form 990 EZ or Form 990 and constructed financial vulnerability ratios developed by Tuckman and Chang (1991) for 216 nonprofit grantees in the sampling frame. We used *t*-tests and chi-square tests to look for differences between grantees designated for competition and those not designated across four ratios: equity, revenue concentration, administrative cost, and operating margin. More information about how the ratios are calculated is provided in technical volume II, appendix B. Results from analyses are reported in chapter IV.

Professional Development and Technical Assistance (PDTA) Survey

For each type of professional development or technical assistance activity identified in the PDTA survey, we generated summary statistics of the frequency of respondents indicating that "yes" they engaged in that activity during the prior 12 months. After determining there were no significant differences between grantees that were and were not designated for competition, we computed descriptive statistics for the PDTA survey for the sample as a whole. These analyses included the weights that account for potential bias related to the low response rates. Results of analyses are reported in chapter III. See technical volume II, appendix E for the survey instrument.

Applicant Survey and Extant Data Sources On Competition

The evaluation team used data from the applicant survey, PIR, and other extant data sources to assess levels of competition, characteristics of applicants and awardees, predictors of competition levels, probabilities of winning, and the ways that competition and awards may improve quality. We used several quantitative methods including summary statistics on characteristics and levels of competition and award, *t*-tests comparing incumbent applicants and new applicants characteristics and proposed quality improvement activities, paired *t*-tests comparing the prior grantees and awardees characteristics and structural quality features, a probit regression model to predict the likelihood of competition based on grantee and service area characteristics, and a probit regression model to determine the probability of winning a grant competition based on applicant characteristics. Some analyses were performed at the FOA level, while other analyses were performed at the applicant level. Results of analyses are reported in chapters VI and VII.

ANALYTIC APPROACHES TO THE QUALITATIVE DATA

Qualitative, semi-structured interviews gathered information about grantee perceptions and actions related to the DRS. These include telephone interviews with 35 program directors, on-site follow up interviews with leaders in 15 grantees, and on-site awardee interviews with nine grantees post-competition. Here, we provide an overview of the philosophy of the analytic approach that begins with the way the interview questions were asked, continues through coding, and results in analytic theme building before summarizing the information for the report. We provide an example based on one interview question to demonstrate the iterative process we follow. Results from the analyses described below are reported throughout chapters III, V, VI, and VII and additional information about the methods and analyses are available in technical volume II, appendices C and D.

Asking the Interview Questions

We designed the interviews as semi-structured instruments. Interviewers were trained to elicit stories from the interviewees using the questions as prompts if the interviewee did not address all the points that we were trying to capture. This approach to interviewing has several advantages: interviewees become more comfortable because the interview takes on a conversation-like quality, interviewees can link concepts and ideas together in ways that make sense to them, and the information from the interviews is more contextual providing the why and how, not just the what. The primary disadvantage is that data coding and analyzing become more complex because the interview information is not neatly grouped by question. This form of interviewing also requires a more skilled interviewer because the interviewer must be well versed in the information the interview is seeking, able to listen to the interviewee explain his or her story, and discern where and when to ask the interviewee for additional information. Because we used this approach, we used teams of two interviewers—a senior interviewer who asks the questions and a junior interviewer who takes the notes.

Figure II.2 displays an example question from the Director Telephone Interview protocol and illustrates how the instructions to the interviewer and probes are used to draw out the viewpoint of the interviewee (see technical volume II, appendix E for all the interview protocols). This process is designed to provide a balanced representation of the interviewees' perspective. We learn what conditions of the DRS they are concerned about, which ones they are not concerned about, and if they have a concern, why they have that concern. They define the concerns in their own terms, not ours.

FIGURE II.2.

Example Question from the Director Telephone Interview Protocol
In the interview protocol, this question follows one where the interviewer asks the interviewee to describe his or her understanding of what the DRS is. The interviewer asks the interviewee the questions in italics below; nonitalicized text represents notes to the interviewer about how to prompt the interviewee if he or she is having difficulty answering the question (notes 1 and 2) or if the interviewee asks for clarification about the purpose of the question (note 3). The ordering of the notes is designed to solicit responses with the least amount of framing from the interviewer as possible. We attempt to minimize the extent to which the interviewee's response is shaped by the way we ask the questions.

Now thinking (more) specifically about the DRS conditions, what do you think of those seven conditions?
(Note 1 to interviewer, clarify if necessary: The seven conditions that OHS uses to identify grantees eligible for noncompetitive five-year grants and to identify grantees that will have to compete for ongoing funding).

(Note 2 to interviewer, clarify if needed: The conditions include deficiencies on monitoring reviews, CLASS scores, requirements around school readiness goals, certain audit findings, licensing status, and others.)

a) *Are there certain conditions you see as more problematic than others?*
b) If yes, *which conditions and why?*
c) If no, *is that because you view all the conditions as problematic (explain how) or because you don't view any of them as problematic?*

(Note 3 to interviewer: We are interested in "problematic" in whatever way the respondent defines it, whether in terms of areas their program may have difficulty meeting or in terms of being invalid or unreliable measures of quality in general, or some other definition.)

Coding and Analyzing the Interview Data

We used the same techniques for coding and analyzing all the qualitative interview data collected in this study. We followed five basic principles in the coding, analysis, and presentation of the qualitative data: (1) we sought to be as literal as possible using the interviewees' own words to identify themes when possible; (2) we provided limited interpretation of the data and used interviewee words when possible; (3) we provided some interpretation when needed (e.g. when grantees said there is nothing they are confused about but their answers suggest they may be confused based on an actual rule); (4) we provided the breadth of perspectives even when this meant reporting themes with a small number of respondents (our purposive sampling was designed to create a diversity of perspectives); and (5) we protected the privacy of the interviewees by providing just enough context to make sentiments understandable but took out details that may reveal who they are.

We used a system of "interrogating the data" (Coffey and Atkinson 1996; Mason 2006) to generate codes and themes. We used both *a priori* codes that are directly linked to the interview protocol and emergent codes that were generated from the responses that interviewees provide. Emergent codes were generated in two ways. First, following each interview or site visit, the interview team reflected on the interview or set of interviews regarding the kinds of information the interviewee(s) expressed. We compared the features of the interviewees' responses with the *a priori* codes and we began a list of other possible emergent codes. We used those possible emergent codes as we reflected on future interviews. If future interviews indicated similar features across interviews, we added the emergent to the coding process. After each interview component was completed, all the senior and junior researchers who participated in that round of interviews met to discuss features to search for across all interviews. These became additional emergent codes added to the coding structure before coding began. Emergent codes were also identified during the coding process in two additional ways: (1) the coders indicated that there were pieces of text that appeared important but were not coded or (2) as subcodes to further explore a particular issue.

The purpose of theme building in qualitative data is to create organized ways of analyzing and presenting the textual data collected. It represents a way of relating the codes to each other and the research questions. In theme building, the data is broken apart and recombined iteratively (Coffey and Atkinson 1996) to make sense of the information. Returning to the data generated about what grantees think about the seven conditions of the DRS, we provide an example of this process of "interrogating the data" in table II.4.

TABLE II.4.

Interrogating the Qualitative Data to Generate Codes and Themes

A Priori code: Concerns about the DRS		
Interrogating the Code (breaking apart the data)	**Results of Interrogation—Emergent Codes**	
How do the concerns relate to the DRS?	Some are concerns about the conditions	Some are concerns about implementation of the conditions
What conditions are they concerned about or not concerned about?	Code based on specific conditions	
In what ways are they concerned about implementation?		Code based on specific implementation concerns
What patterns do we observe in their responses to generate themes? (recombining the data)		
Conditions	The responses around the conditions of the DRS are straightforward. We group by condition, clarifying the reason for the concern. These are discussed in chapter III.	
Implementation of the Conditions	The previous step generated 17 different codes about concerns. In this step, we examine those 17 codes to see which ones could be related to each other based on the words that grantees use. This generates the three themes indicated below (presented in chapter V).	
	Concerns about the Assessment Process	*Concerns about a Single Visit or Incident*	*Transparency in the Decision-making Process*

We used qualitative coding software to help us manage this process. This software allows us to assign segments of text to codes, to create subcodes, and to group codes into themes. Senior staff generated the coding structure and trained junior staff to perform the initial round of coding under their supervision. All coders were selected from among trained interviewers, but the coders are kept to two to three per study component. Senior and junior researchers worked together to develop the themes. The themes were shared and discussed with the full team of interviewers to determine if those themes resonated with them based on the interviews they conducted and if they felt like any themes were missing. The interview team generated a shared understanding of the key themes and agreed on how to represent them. Senior researchers and junior researchers worked together to select text representative of the themes, and senior researchers wrote the report.

We also coded the data based on the designation status of the grantee to allow us to explore the extent to which grantees designated and not designated for competition vary in their perspectives. We further coded by the type of respondent for the site visit interviews where we have multiple types of respondents. For the site visit interviews, we considered the perspectives of the Head Start director as the anchor perspective and then compared the perspectives of the other interviewees with that perspective. If we found differences in perspectives across types of interviewees, we reported those differences. Codes and themes are presented in technical volume II, appendices B and C.

Chapter III. How do grantees perceive the DRS and prepare for their monitoring and assessment in terms of efforts to improve program quality?

A key goal in this study was to better understand how grantees responded to the DRS in terms of specific activities and operational decisions that might affect quality. In this chapter, we examine the ways Head Start grantees sought to improve quality through professional development and technical assistance activities in the year before their first experience with DRS. Additionally, we explore whether or how grantees' decision to engage in quality improvement may have been shaped by their perceptions and understanding of DRS. This chapter explores how DRS might stimulate grantee quality improvement efforts before their scheduled monitoring review by looking at the types of quality improvement activities grantees engage in and what motivated them.

Overview of Methods

This portion of the study was supported by a mixed methods design. Seventy-one grantees (half designated for competition, half not) were sampled from among grantees receiving a monitoring visit in the 2013—14 year (Monitoring Cohort 4), supplemented with some grantees monitored in 2012—13.[26] Of those, program directors from 66 grantees completed the Professional Development and Technical Analysis (PDTA) Survey to measure the quality improvement activities grantees engaged in during the previous 12 months. The survey asked directors to report on the topics addressed by the quality improvement activities (e.g., child health and safety, teacher-child interactions, management, fiscal operations), the mechanisms through which they were delivered (e.g., coaching, in-person TA, distance

---

[26] Sixty-one grantees (29 designated for competition) were sampled from among grantees receiving monitoring visits during the 2013-2014 monitoring year. Due to low response rates and to ensure the sample had adequate power to detect differences between grantees designated and not designated for competition, 10 grantees (6 designated) were sampled from among grantees that received a monitoring visit during the 2012-2013 monitoring year. See technical volume II, appendix B for additional information.

TA), and the recipients (e.g., teachers, managers, governing board). A subset of those grantees also participated in qualitative interviews exploring how DRS may be incentivizing grantees to improve quality. First, 35 program directors participated in semi-structured telephone interviews shortly after their monitoring visit in 2013—2014.[27] Then, after all grantees learned whether they were designated for competition but before funding opportunity announcements were issued, 15 grantees participated in on-site follow up interviews with key personnel, including directors, managers, governing body members, and Policy Council members. Interviewers asked about grantees' understanding of the purpose of the DRS, its conditions, and the monitoring and assessment process, and in what ways they had responded to the DRS through quality improvement efforts.

Descriptive statistics and frequencies were calculated for the responses to the PDTA Survey. We used *t*-tests to assess differences in the quality improvement activities reported by grantees designated for competition and grantees not designated. We found no statistically significant differences and, therefore report results across both groups using a sampling weight to adjust for the oversampling of grantees designated for competition. Thus, the quality improvement activities reported in this chapter are representative of the sampling frame for this portion of the study,[28] including both grantees designated for competition and those not designated.

Qualitative interviews were coded to identify themes about respondents' understanding of the DRS, actions taken to improve quality before their monitoring review, and the ways those actions were motivated by the DRS. We coded by designation status and by type of respondent and looked for variations in themes across groups, as well as within and across sites. We report themes at the grantee level, but point out where within-grantee differences or other variations of note occurred. Our goal was to try to understand a range of perceptions and experiences of DRS and not necessarily to represent the population of all grantees. Throughout this chapter, we provide information about the proportion of grantees we spoke to that held similar perceptions of their experience with DRS, but this does not suggest the same would be found within the general population of Head Start grantees. In other words, if half of grantees responded similarly in this study, it may be that less than half or more than half of all

---

[27] The goal of the evaluation was to conduct the Director Telephone Interviews before grantees knew their designation status. On average, interviews were conducted six months following monitoring, and they occurred after the quality assessment visits had been conducted. A total of 19 interviewees may have known their designation status at the time of the interview: eight because they had received deficiencies and 11 because interviews were conducted following the public announcement of designations (six not designated and five designated due to CLASS). In the interview, we did not ask if they knew their designation status as we did not want interviewees focused on that, so we do not know definitively who knew and who did not know their status.

[28] This weighting adjustment adjusts the value of the analytical weight (the one applied to adjust for nonresponse bias) so that each sample domain ends up with the same proportion of grantees that exists in the sampling frame (i.e. 75.2 percent not designated for competition; 12.8 percent designated due to CLASS; 8.9 percent designated due to deficiency; and 3.0 percent designated due to CLASS and deficiency.

Head Start grantees that experienced the DRS would have responded similarly. We provide exemplar quotes to illustrate the themes. See chapter II and technical volume II for additional details on the study design, sample, measures, and analyses.

Findings

These analyses of the PDTA Survey and qualitative interviews revealed three key findings:

1. *Grantee understanding and perceptions of the DRS are mixed.*
2. *Grantees report engaging in multiple strategies to improve program quality before their monitoring visit and many of those strategies are related to the CLASS condition of the DRS.*
3. *Some grantees report that they did make changes in their quality improvement strategies in response to the DRS while others indicate that their efforts to improve quality are not related to the DRS.*

Each of these is discussed in turn with some exemplar quotes provided. The designation status of the grantee providing the quote is indicated to provide some sense of diversity of respondents, but other characteristics are not indicated to ensure confidentiality.

Finding 1: Grantee understanding and perceptions of the DRS are mixed.

In the telephone interviews, we asked program directors about their understanding and perceptions of the DRS as a whole and the DRS conditions individually. We also asked about what they saw as positive and negative aspects of the DRS. One reason for asking these questions was to determine if their understanding and perceptions might influences the types of actions they took in response to the DRS.

Many responding grantees acknowledged the need for accountability and quality improvement in Head Start.

All 35 program directors who participated in the telephone interviews conveyed an understanding of the intended purpose of DRS as a mechanism for accountability. About half (16 of 35) had positive initial impressions of the DRS and viewed it as a system designed to raise standards of quality throughout

Head Start by identifying low-quality programs and requiring them to compete for their grant. Others (17 of 35) also viewed the DRS as a monitoring system, but saw it as another strategy to ensure compliance across grantees and initially had many concerns about the new system. A few Head Start directors (5 of 35) stated that they were unclear about the DRS when they first heard about it, but as they received more information, they started to view it in a more positive light, as a tool to improve accountability and identify low-performing programs. One Head Start director shared that when she initially heard about the DRS, she thought it was "overkill," but over time she developed a more positive outlook because

> *We have to become better at showing outcomes and results that we're getting. This is forcing us to learn how to do that better...I don't want to use the word favorable, but I do feel like it is bringing everybody to the point of having to improve and prove accountability. —Not Designated*

Many responding Head Start directors (18 of 35) described seeing the value of the DRS in making grantees receiving federal funds accountable for their performance and quality of services. They recognized that not all grantees were performing to standard and the DRS could be a way to rectify that.

> *I said, 'About time.' That's it, I thought it was very necessary...when they decided on CLASS, I said, 'Great!'...if I was the person giving this money, there would need to be a whole lot of accountability. I felt, this is my taxpaying money. You need to get this right. —Designated for Competition*

Some grantees (11 of 35) also felt the DRS would be beneficial because it emphasizes quality improvement and "*gives everyone a purpose to continue working and have a successful program*." "*Sometimes if you're in the same spot, the same position, you get complacent. This way it shakes you up...If there's an area of noncompliance, they know sooner and can make that correction*" (Not Designated). Some grantees (13 of 35) were pleased that the Office of Head Start (OHS) created a system to identify and defund low-quality programs, to "*ensure that there are more quality programs and less of those programs that are questionable in their services.*"

A minority of responding Head Start directors (3 of 35) thought that the triennial monitoring should have been sufficient in identifying poor quality, underperforming programs and that DRS was reinventing the wheel.

> *My first thoughts were, 'Why are we doing this? There is already a system in place that could be utilized to defund, or whatever word you want to use, to take the Head Start grant away from programs that have proven that they cannot operate it efficiently or effectively. Why are we recreating the wheel?" —Designated for Competition*

Many grantees (19 of 35) agreed that the overarching goal of the DRS to improve accountability was a good idea and *"in principle, on paper, it's a good step in the right direction."* At the same time, many had concerns.

Responding grantees tended to think most of the DRS conditions are reasonable and comprehensive, but many grantees are concerned about the lowest 10 percent requirement of the CLASS condition.

We asked the 35 Head Start directors about their understanding and perceptions of the conditions of the DRS. If respondents could not remember the conditions, the interviewer read the conditions over the phone. Directors tended to respond in similar ways. None of the responding directors were concerned about what they referred to as the "financial conditions" but many directors (74 percent) were concerned about the CLASS condition and the lowest 10 percent stipulation in particular. Three responding directors were not concerned about any of the conditions.

TABLE III.1.

Program Directors' Concerns about DRS Conditions

DRS Condition	Percentage of Telephone Respondents Concerned About Condition (N=35)
A deficiency (i.e., a systemic or substantial failure) in meeting program performance standards resulting from a triennial, follow-up or unannounced monitoring review.	23% These directors were concerned that one deficiency would lead to designation, but not about the category of designations as a whole.
Classroom Assessment Scoring System (CLASS) scores below a minimum threshold or in the lowest 10% in Emotional Support, Classroom Organization, or Instructional Support.	74% These directors were concerned primarily with the lowest 10% threshold
Failure to establish, analyze and utilize program-level school readiness goals.	6% Two directors were concerned that the standard was not clear enough or that it should be based on improvement.
Revocation of state or local license to operate.	14% These directors expressed concern that not all Head Start grantees were subject to state licensing and that sometimes the term "revocation" was misinterpreted
Suspension by ACF; Debarment from receiving funding from another federal or state agency or disqualification from participation in the Child and Adult Care Food Program (CACFP); or Determination from an annual audit of being at risk for failing to continue functioning as a "going concern."	0% Responding directors referred to these three conditions as the "financial conditions" but none expressed concerns about them

Responding directors tended to indicate that the fiscal conditions were a "*no brainer*" because everyone should be accountable for taking care of the government funds they are given. No grantees objected to the principle behind any condition. Rather, their concerns focused on particular elements of conditions as indicated in table III.1. Those that were concerned about the lowest 10 percent portion of the CLASS condition focused primarily on it being a moving target (i.e., CLASS scores falling in the lowest 10 percent vary from one cohort to the next). However, some of the comments regarding concerns about the lowest 10 percent portion of the CLASS condition suggest that some directors may not fully understand it. One director remarked, *"Eventually, all the grantees will be in the upper level of the CLASS scores, and to have them recompete is a horrible waste of government money and people's*

time and effort." The CLASS condition does specify that once grantees reach a score of 6 on the 7-point scale, they will not be designated for competition, even if their score is in the bottom 10 percent.

We also asked the 35 program directors participating in the telephone interviews about whether they thought the DRS captured their program quality well, and if not, what aspects of program quality did it not capture well. The majority of directors (86 percent) indicated that they thought the DRS was very comprehensive and covered all aspects of program quality. Only five of the 35 respondents thought that the DRS did not capture all aspects of program quality. These directors were concerned that it did not capture the following areas of quality well: family support/parent engagement (3), partnerships (2), and degreed teachers (2). They indicated their response was based on their recent monitoring visit which they perceived did not examine these elements of their program.

When we conducted the on-site follow-up interviews in 15 grantees with program managers, agency directors, governing body members and Policy Council members, we asked them the same types of questions to see how their perspectives might be similar to or different from the Head Start director's perspective. Perspectives of other leadership staff tended to be similar. It was not clear how much Policy Council members understood the DRS, but they agreed with the general concept of accountability. Governing body members generally thought an accountability system like the DRS was important, but they were mixed in the level of their understanding. Some understood that only some grantees would have to compete to maintain grants, but others thought every grantee would have to compete (and, thus, were not upset about their own organization needing to compete). Governing body members at 2 of the 15 sites mentioned concerns about the CLASS condition. In one case, they were concerned that attaining high CLASS scores would always be hard for them because they struggled to attract the highest quality teachers because of the pay differential between Head Start and prekindergarten programs in their community. In the other case, they felt that a national standard of the bottom 10 percent did not take into account their local situation and preferred to focus efforts on locally determined priorities not measured by the CLASS, such as whether the child and family are ready for kindergarten, and the extent to which families are accessing other services they need. Although the governing body member understood that the CLASS condition did not preclude them from working on these other things, the comments indicated concern that the CLASS condition was causing the program to focus more on national priorities than local ones.

Some grantees were concerned about what they perceived as a changing relationship with their regional OHS representatives.

We did not ask interviewees in the telephone or on-site interviews about their relationships with their regional OHS representatives. We asked about grantees' views of the positive and negative results of the DRS and 29 percent of the Head Start directors responding to the telephone interviews expressed concerns with what they perceived to be a change in relationship with the regional OHS representatives. These directors indicated that historically, they had positive relationships with the regional offices and perceived them to be collaborative and supportive. But with the introduction of the DRS, the directors felt those relationships change in various ways, ranging from an inability to provide them with help related to the DRS (perceived by the grantees as the regional offices not having enough information to give them) to an air of distrust where grantees were afraid to seek help from the regional office because they might say something that would lead to a designation for competition.

> *I just think how everything was rolled out could have been better. [There could have been] information provided to regional office so they were more equipped. Our regional office is top notch…but their hands are tied by the information they have to give out…They have to get better at that. Just ramming things down people's throats on the fly is not a good practice.* —Designated for Competition

> *There's a culture of fear that didn't used to be there; that's been detrimental. There are all these stories of a program asking for help and then getting recompeted, so now you don't ask the [Head Start] regional office for help anymore.* —Not Designated

Finding 2: Grantees report engaging in multiple strategies to improve program quality.

In the PDTA Survey, grantee directors reported on their program's professional development and technical assistance activities during the 12 months before their first monitoring visit after implementation of DRS. We focused on the types of activities designed to improve quality and asked about the topics and staff targeted in training, coaching, technical assistance, and other activities. We also conducted qualitative interviews with a small subset of these grantees (15) to explore ways they may have worked to improve program- or systems-level quality, such as changes to their administrative structure, fiscal controls or processes, and procedures for meeting various standards. Below we describe a range of quality improvement activities that programs engaged in before their first monitoring visit after implementation of DRS. In the next section (finding 3), we discuss exploratory findings about grantees' motivations for seeking quality improvement.

All grantees reported engaging in multiple professional development and technical assistance activities in the 12 months before their first experience with DRS.

All the program directors reported that in the 12 months before the first monitoring visit applying the DRS conditions, teaching staff participated in various professional development and technical assistance activities, and nearly all reported that directors and coordinators (99 percent) and family support staff (95 percent) also participated in those activities. Similarly, 67 percent indicated that other program staff (e.g., health, transportation, and office staff) also participated in training. Nearly two-thirds of surveyed grantees (65 percent) indicated that they did not focus their professional development and technical assistance efforts any differently in the 12 months leading up to their monitoring visit compared with earlier periods of time.

We asked the Head Start directors how professional development and technical assistance decisions are made in their programs, giving them the choices of "decisions are made centrally by the Head Start program director," "each delegate agency decides on what is needed for the sites they operate," "each center makes decisions about what they need," or "another approach, please explain." Most of the respondents (78 percent) indicated they use another approach; nearly all of these (96 percent) said they use an individual teacher education plan to determine what each teacher needs.

Directors reported that quality improvement activities cover a variety of topics. Topic areas covered through professional development and technical assistance during the previous 12 months included: child health and safety (97 percent), family engagement (87 percent), enrollment practices (81 percent), and family referral (75 percent). Nearly all grantees also reported engaging in professional development activities related to the curriculum they were implementing (96 percent) and improving teacher-child interactions (95 percent).

We also asked grantees to report on types of technical assistance strategies used in the previous 12 months including phone technical assistance, on-site technical assistance, mentoring, and coaching. Figure III.1 shows that coaching was the approach most used for technical assistance activities related to CLASS (74 percent), health and safety (55 percent), and enrollment practices (46 percent), but on-site technical assistance was most used for school readiness goals (73 percent) and family engagement practices (68 percent). Technical assistance provided by phone was used by fewer than 50 percent of the grantees in all topic areas.

FIGURE III.1.

Percent of Grantees That Report Receiving Pre-Monitoring Technical Assistance, by Topic and Mode
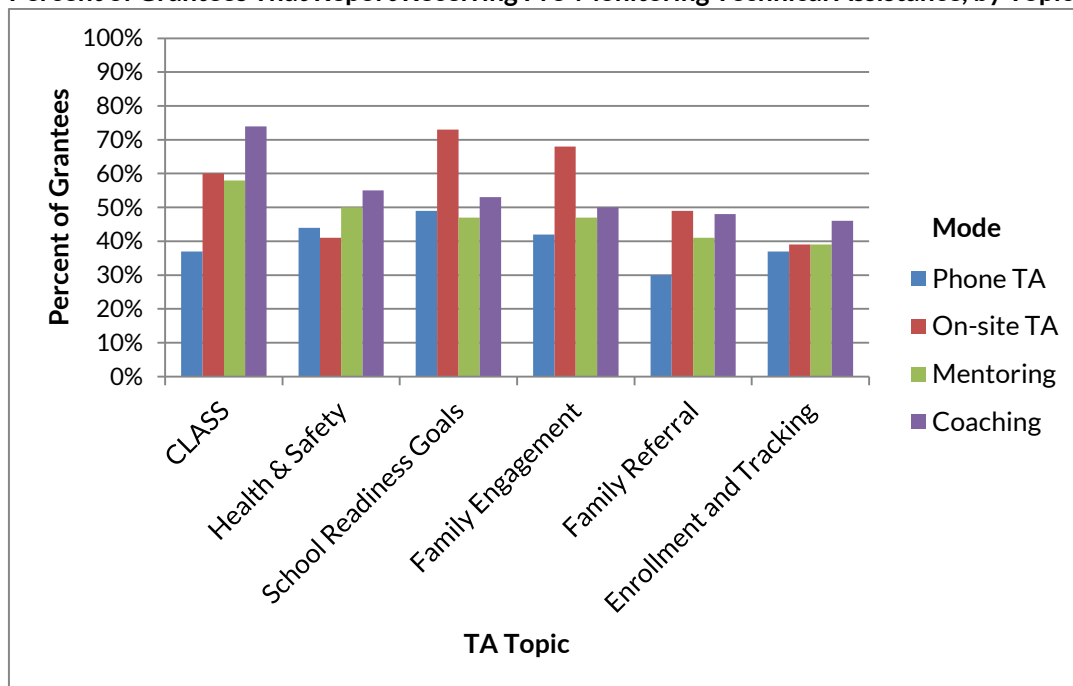
Source: Professional Development and Technical Assistance (PDTA) survey of 66 grantee directors.

We asked program directors some specific questions about their efforts to improve CLASS scores or teacher-child interactions because the CLASS condition was new to Head Start monitoring when it was introduced by the DRS. One hundred percent of the grantees indicated that they had obtained some type of professional development or technical assistance for their staff either specifically on the CLASS or on child-teacher interactions in general in the 12 months before their monitoring review. The methods they used to provide this support to their staff varied:

- 91 percent reported teacher-child interactions were the focus of coaching for the purpose of professional development

- 74 percent reported teacher-child interactions were the focus of coaching for the purpose of technical assistance (as noted above in figure III.1)

- 86 percent conducted internal group trainings on one or more topics to help teaching staff improve CLASS scores and 67 percent conducted similar internal trainings for program supervisors, managers, and coordinators

- 60 percent reported bringing in someone to perform mock-CLASS assessments

- 49 percent of grantees sent supervisory staff, program coordinators, and managers to get official CLASS certification at Teachstone[29] so they could conduct internal CLASS assessments and train other staff; 38 percent of grantees also sent teachers to these types of CLASS certification trainings

We also asked program directors about staff engagement in college coursework toward obtaining a degree (61 percent of grantees had teaching staff, and 41 percent of grantees had supervisory staff engaged in coursework), training on topics related to previous noncompliances (29 percent of grantees had teaching staff, and 33 percent of grantees had supervisory staff that engage in this type of training), and training based on particular staff weaknesses but not closely tied to monitoring criteria (82 percent of grantees had teaching staff and 88 percent of grantees had supervisory staff that engage in this type of training).

In addition to the professional development and technical assistance activities already discussed in the survey results, grantees were most likely to describe two program- or systems-level activities to improve quality that emerged from qualitative interviews with staff and leaders in 15 grantees: (1) creating new staff positions and responsibilities and (2) refining operational systems and procedures to meet Head Start Program Performance Standards.

Grantees created new staff positions and responsibilities to improve quality.

Nine of the 15 grantees participating in qualitative interviews mentioned that they reorganized their management structure by adding or removing staff functions. While the titles used by grantees varied, many grantees (9 of 15) reported creating management positions comparable to an education manager or center manager[30] with the explicit goal of raising CLASS scores. Three grantees (all designated for competition) spoke of adding a compliance-related position and one grantee added health and nutrition coordinators. One program restructured so one staff member would be in charge of transportation and another would be in charge of facilities. Restructuring was always related to the need to have staff with targeted expertise in the areas they were overseeing.

---

[29] Teachstone is the organization that provides training for and certifies individuals in Classroom Assessment Scoring System (CLASS) assessments.

[30] Grantees used various terminologies to describe their staffing positions. All of the grantees participating in site visits had someone focused on guiding and overseeing the education portion of their Head Start programs. These people were frequently called education directors, managers or coordinators. Sometimes, they situated these activities within their center director or center manager functions. Either way, these staff had a role in guiding or supervising classroom teachers.

A Head Start director in a grantee designated for competition reported how *"…the compliance officer is the one position risen out of DRS."* Another director (also in a grantee designated for competition) described the responsibilities of her compliance monitoring director, a position created two years before the grantee's review.

> *She's also the trainer for CLASS and CDA.[31] She's our expert on meeting the requirements for child care licensing. She makes sure everyone uses and follows the right monitoring tools. We made sure our systems were in place—governance, training, ERSEA.[32] —Designated for Competition*

Programs that changed job descriptions generally added coaching, mentoring, and training responsibilities to the role of their education managers, center managers, or other senior staff (9 of 15). Program leaders in these sites described how they needed to reshape their staffing approach so staff in these management-level positions would know the educational requirements and could offer teachers supervision and training. As one grantee said,

> *We have reorganized our education unit as a result of DRS, specifically CLASS. We have put in a new position of education manager to make sure supervisors and centers have what they need to have higher CLASS scores. —Not Designated*

In terms of changing staff functions, 3 of the 15 grantees (all designated for competition) said they dismissed staff members, including teachers, administrators, and transportation staff. Reasons for dismissal varied, and interviewers did not probe on this topic. In these cases, programs hired degreed teachers or other staff who had the skills necessary to effectively handle tasks and comply with Head Start requirements. In one case, the program eliminated the transportation services and staff to reduce the risk of deficiency, because they had heard other grantees had been designated for competition for issues related to transportation; they considered it a proactive risk-management strategy.

Grantees refined operational systems and processes to better meet standards.

Some grantees (5 designated for competition and 3 not designated of the 15 grantees interviewed) described reviewing their policies, procedures, record-keeping, and data to make sure "our systems were in place—governance, training, ERSEA." These efforts attempted to meet standards and avoid any violations that would put them in jeopardy of being designated for competition. One Head Start director discussed using data more frequently and intentionally now.

---

[31] CDA is the Child Development Associate credential. It is a nationally available teaching credential.
[32] ERSEA is Head Start's set of requirements regarding enrollment. It stands for Eligibility, Recruitment, Selection, Enrollment and Attendance.

We've certainly done more with data collection and analysis on CLASS, school readiness goals. We're doing way more than we did before. Before it was done, but almost done intuitively. We had the data, we used the data… I'm not sure all the factors have made us go, 'Oh my, there's something we didn't know.' It's more the process of having this stuff available to someone else should they come in and ask for it. Making sure that we are meeting those standards. They want you to be able to concretely show that you are doing data analysis. —Not Designated

Finding 3: Some grantees report that they made changes in their quality improvement strategies in response to the DRS, while others indicate that their efforts to improve quality were not related to the DRS.

It is difficult to untangle which quality improvement activities grantees engaged in because of the DRS and which ones they would have undertaken any way. Telephone interviews with directors of 35 Head Start grantees and on-site follow-up interviews with staff and leaders in 15 of those grantees explored grantees' motivations for engaging in quality improvement activities and their perceptions of how those activities related to the provisions of DRS. Below we describe lessons from these interviews about whether grantees attributed their quality improvement efforts to the DRS.

In the telephone interviews, program directors typically reported that the changes they made to improve their program's quality were not because of the DRS. They indicated that they were always making changes in efforts toward continuous quality improvement and had a fairly rigorous routine preparing for past monitoring visits that they continued to use in preparation for their most recent monitoring visit. As one grantee expressed:

We prepared for our review like we always do. We were mindful of those seven items [the DRS conditions]. We made sure to hit up on those so we didn't go into recompetition. I think we prepared for it. Before DRS we had always prepared for the review using the guidelines and the Head Start Act. We prepare the same way; we just had to make sure the paper trail proved it. —Not Designated

In the follow-up site visits, we asked grantee staff and leaders about the quality improvement strategies directors either mentioned in the telephone interviews or reported in the PDTA surveys. We asked whether activities to improve quality before their monitoring visits were related to the DRS. Their responses suggested that grantees are making some changes to improve quality because of the DRS, but they are reluctant to attribute the changes to the DRS. When we asked whether quality improvement activities aimed at improving CLASS scores or school readiness goals were related to the DRS, 11 of 15 grantees indicated that they engaged in some of these activities in response to the DRS.

Some of them indicated they were more intentional about linking together their various strategies, as seen in the quotes below from three grantees about using the school readiness goals and CLASS, Creative Curriculum and CLASS, or Conscious Discipline and CLASS as paired strategies for improvement:

> *Training is definitely more of a focus. School readiness goals are definitely at the forefront. Our classrooms have improved as far as teacher implementation and the structure, routines, behavior management, since the DRS and especially with CLASS being implemented. —Not Designated*

> *We did purchase the Creative Curriculum kits because we heard that other HS programs that had those had their CLASS scores go up. We have also had some CLASS score trainings last year and will again this year, and we have seen our scores go up. I think those kits helped not only our CLASS scores, but we've seen the 4 year olds progress more than they used to, and we think it's because of the training and the kits. —Not Designated*

> *Well, for the Conscious Discipline Learning Guidance connection with CLASS, we had our trainer who comes for pre-service and in-service, we had her review the CLASS framework, and she saw some of the things she'd already been training them on and how it related, so she helped them connect the dots on the social-emotional, instructional support, and whatever the other one was. —Not Designated*

Grantee responses to the PDTA survey and follow-up interviews suggest changes in how grantees prepare for their monitoring visit occurred more than 12 months before their first monitoring visit applying the DRS conditions. That is, the PDTA survey asked about quality improvement strategies and changes implemented in the 12 months before grantees' most recent monitoring visit; follow up interviews suggest changes in grantees' quality improvement strategies may have been implemented much earlier. For example, all 15 grantees participating in the site visit interviews said they had been using the CLASS for two to six years, with at least one respondent reporting they were trained in CLASS as soon as the revised CLASS instrument was available. Three of 15 grantees reported anticipating that CLASS was coming as a Head Start standard with the passage of the Head Start Act in 2007, which required that Head Start use an assessment of teacher-child interactions for monitoring grantee performance at the federal level and to inform professional development of teachers at the local level. Those three grantees also indicated that they sent one or more staff to the initial CLASS trainings offered by OHS in preparation for DRS implementation.[33] Other grantees did not anticipate that CLASS was on its way to becoming part of the DRS, but began to focus on the CLASS nonetheless because they perceived CLASS to be a useful professional development and coaching tool. This led grantees to certify staff on the instrument (10 of 15) and to use a coaching model to work with teachers (9 of 15).

---

[33] OHS offered CLASS trainings in 2009 prior to the DRS regulations going into effect in 2011 (Classroom Assessment Scoring System ACF-IM-HS-08-11). The interview did not specifically ask whether sites participated in these trainings, which means that more interviewed sites may have participated but did not mention it in their responses.

Summary

In sum, Head Start grantees tended to agree that a system of accountability like the DRS was needed to improve program quality in Head Start. Most grantees indicated that most of the seven conditions were fair, but many grantees expressed concerns about the CLASS condition, particularly the lowest 10 percent threshold. All grantees engaged in multiple professional development and technical assistance strategies in the 12 months before their monitoring review. In the PDTA survey, however, many grantees indicated that these strategies were not any different from what they would have done any other time. In interviews, grantees indicated that they are working on improvements all the time, so it is difficult to identify which improvements are specific to the DRS, but the grantees pointed to some changes in program structure and procedures and some added emphasis on CLASS and school readiness goals.

Next, we examine how well the DRS identified higher versus lower quality grantees.

Chapter IV: Does the DRS differentiate higher versus lower quality programs?

A primary goal of this study was to determine the extent to which the DRS is able to identify which Head Start grantees are providing high quality comprehensive services and are, thus, eligible to receive automatic renewal of their Head Start grant; and which grantees are not providing high quality comprehensive services and will be subject to open competition for renewed funding. Specifically, in this chapter, we examine whether there are differences between grantees designated for competition and grantees not designated across a range of quality measures. We further examined quality differences for grantees designated for competition based on deficiencies in meeting Head Start Performance Standards separately from grantees designated for competition due to low CLASS scores.[34] Finally, we examined how well the CLASS performs as a measure of grantee-level quality, as it is used for the purpose of DRS, and whether CLASS scores were influenced by such factors as number of observation cycles or differences among individual observers. This additional focus on the CLASS condition was stimulated by an interest in understanding this new feature of Head Start monitoring and used additional analyses to explore possible reasons for the results. In sum, this chapter focuses first on the central tenet of the DRS and asks whether grantees designated for competition provided lower quality services compared with not-designated grantees. Second, we look at the extent to which the CLASS provides reliable assessments of teacher-child interactions at the grantee level, as it is used for the purpose of the DRS.

Overview of Methods

We addressed this research question using several quantitative methods to measure multiple features of program and classroom quality. We recruited a stratified random sample of 71 grantees (35 designated for competition, 36 not designated) from Monitoring Cohort 4 with supplementation from

---

[34] The evaluation focused on these two conditions in DRS only because they accounted for 99 percent of designations for competition at the time of the evaluation.

Monitoring Cohort 3 to participate in the quality assessments.[35] The reasons for designation included deficiencies (40 percent) and low CLASS scores (69 percent).[36] As expected, the designated grantees (regardless of reason for designation) were rated by the OHS monitors as scoring significantly lower on the CLASS (see appendix B in technical volume II).

Within grantees, we recruited 554 classrooms serving 3- and 4-years-olds.[37] Table IV.1 describes the measures used to assess the quality of classroom, center, and grantee level services through observations of classrooms and interviews with center directors. We selected quality measures to align with the two DRS conditions that account for over 99 percent of designations—deficiencies and low CLASS scores. These include measures of the quality of health and safety practices, parent involvement, staff qualifications, management, governance, and fiscal operations[38] to measure practices within grantees that could result in designations for competition due to deficiencies. We selected measures of the quality of teacher-child interactions involving emotional support, instructional support, and classroom management to measure practices within grantees that could result in designation for competition due to low CLASS scores. Finally, we also administered the CLASS to allow us to compare CLASS scores of the OHS monitors with those of the evaluation team.

Analyses tested for mean differences in quality ratings at the grantee-level between grantees designated for competition and not designated, overall and separately by the CLASS condition and deficiency condition.[39] Finally, we conducted additional psychometric analyses to better understand the reliability of CLASS, as used for the purpose of DRS. Our analytic approach is described in more detail below. Also see chapter II and appendix B in technical volume II for additional details on the study design, sample, measures, and analyses.

---

[35] The sample drawn from Monitoring Cohort 4 was supplemented with grantees from Monitoring Cohort 3 due to low response rates and a lower than expected number of grantees being designated for competition based on deficiencies in Monitoring Cohort 4. For a complete discussion of the sample, see technical volume II, appendix B.
[36] Percentages do not add up to 100 because some grantees were designated for competition for both reasons. The evaluation team calculated this information from OHS administrative data on reasons for deficiencies.
[37] In grantees with eight or fewer classrooms, all eight classrooms were sampled.
[38] Quality of financial operations was assessed using data from IRS Form 990 for all nonprofit grantees in the sampling frame and not only those 71 recruited to participate in the classroom observations and interviews. For additional details, see chapter II and technical volume II, appendix B.
[39] The evaluation did not test for differences in quality ratings by type of deficiency because the sample of grantees with deficiencies was too small for further subgroup analysis.

TABLE IV.1.

Evaluation Measures Used to Assess Program Quality

Quality Construct	DRS Condition	Study Measure
Child Health & Safety	Monitoring Deficiency	**Health and Safety Checklist:** The Health and Safety Checklist was adapted from NAEYC's Childcare Health and Safety Checklist (Aronson 2012) and The California Childcare Health Program Health and Safety Checklist.
Family & Community Engagement	Monitoring Deficiency	**Program Administration Scale (PAS) Family Partnership Subscale:** The PAS includes a subscale that measures family and community engagement.
Child Development & Education	Monitoring Deficiency	**PAS Child Assessment Subscales and Report on Curriculum:** Includes screening and identification of special needs and of assessment in support of learning.
Management, Operations & Governance Systems	Monitoring Deficiency	**PAS Governance and Management Subscales:** Includes Program Planning & Evaluation, Center Operations, Human Resource Development, Marketing & Public Relations, and Personnel Cost & Allocation Scales.
Fiscal Integrity & Vulnerability	Monitoring Deficiency	**PAS Fiscal Management Subscale and Tuckman & Chang Financial Ratios:** PAS scale includes questions about budget management and accounting practices. Tuckman and Chang (1991) method uses IRS Form 990 data for nonprofit organizations to calculate financial vulnerability ratios.
Classroom Quality	Monitoring Deficiency	**PAS Staff Qualification Subscales:** Includes questions about education and training of all staff members.
Classroom Quality	Low CLASS Emotional Support	**ECERS-R Interactions Score:** The ECERS-R is a widely used measure of the quality of classroom environment. Included in the Interaction Score are 10 items measuring the quality of the interactions between the teachers and children during the morning of observation.
Classroom Quality	Low CLASS Instructional Support	**ECERS-E Academic Activities Score:** The ECERS-E was developed to extend the ECERS-R by describing academic activities in the classroom. The score is based on four items reflecting academic activities during the morning of observation.
Classroom Quality	Low CLASS Classroom Organization	**TSRS Adapted Classroom Structure and Management Subscale**: The scale includes three items that are observed in four cycles during the morning of observation.

Notes: PAS is the Program Administration Scale. It has multiple sub-scales that align with various constructs.

Findings

The major findings include:

1. *The evaluation's classroom and center quality scores were not significantly different for grantees that were and were not designated. When we looked at reasons for designation, however, we found quality of center operations was significantly lower among grantees designated for deficiencies compared with not-designated grantees.*
2. *The CLASS scores were significantly higher when collected by OHS monitors than when collected by the evaluation team. In the data collected by the evaluation team, there were no significant differences related to the number of cycles collected per classroom or whether classrooms were sampled, but there were substantial differences among raters.*

Finding 1: The evaluation's classroom and center quality scores were not significantly different for grantees that were and were not designated overall. When we looked at reasons for designation, however, we found quality of center operations was significantly lower among grantees designated for deficiencies compared with not-designated grantees.

Comparison of Quality Measures by Designation Status

Analyses first compared the quality of services, as measured by the evaluation, of grantees designated for competition with those not designated. We computed the mean of classroom- and center-level quality measures for each grantee and compared these grantee-level scores using *t*-tests. We weighted analyses to account for differential recruitment rates (see appendix B in technical volume II for details on the weight construction). Our analyses involved multiple imputations[40] to account for differential participation rates on the measure of center quality, the Program Administration Scale (PAS). Table IV.2 shows the results of these analyses.

---

[40] *N*=40 imputed datasets.

Overall, grantees designated for competition and not designated did not reliably differ on any of the evaluation's measures of quality.[41] There was one "moderate" effect size with a p-value between .1 and .05 that suggested the study might be underpowered to detect lower levels of staff qualifications among designated grantees.

TABLE IV.2.

Comparing Grantee Quality Measures by Designation Status

Quality Measure (collected by evaluation team)	Not Designated			Designated for Competition			t-test	
	N	Mean	SD	N	Mean	SD	P	Effect Size[a]
Classroom Quality, Aggregated at Grantee Level								
CLASS Emotional Support	36	5.33	0.42	35	5.33	0.46	ns	.00
CLASS Instructional Support	36	2.48	0.73	35	2.68	0.84	ns	-.25
CLASS Classroom Organization	36	4.63	0.61	35	4.61	0.59	ns	.03
ECERS-R Interaction	36	5.19	0.65	35	5.25	0.61	ns	-.10
ECERS-E Quality of Academic Activities	36	3.03	0.57	35	3.22	0.59	ns	-.33
TSRS Class Structure & Management	36	3.73	0.52	35	3.79	0.43	ns	-.13
Center Quality[b], Aggregated at Grantee Level								
PAS Staff Qualifications	36	2.80	0.80	35	2.48	0.75	+	.42[+]
PAS Personnel Cost/Allocation	36	3.17	0.74	35	3.08	0.89	ns	.11
PAS Child Assessment	36	5.87	0.98	35	6.12	1.15	ns	-.23
PAS Fiscal Management	36	3.57	1.39	35	3.48	1.54	ns	.06
PAS Family Partnerships	36	4.31	1.17	35	4.49	1.18	ns	-.15
PAS Governance and Management	36	3.06	0.91	35	3.27	0.80	ns	-.24
Health and Safety Checklist: Indoor	36	0.86	0.07	35	0.85	0.07	ns	.12
Health and Safety Checklist: Outdoor	36	0.80	0.10	35	0.79	0.10	ns	.07

Source: Data collected by the evaluation team during week long site visits involving classroom observations and center director interviews.

Notes: Analyses tested for differences between the designated for competition and not-designated groups on quality scores aggregated to the grantee level. Analyses were weighted to account for nonresponse and included multiple imputations for missing data.

Significance noted as: + .1>p>.05; * p<.05; ** p<.01; *** p<.001; ns=not significant.

[a] Effect sizes computed as the difference between the means divided by the sample standard deviation.

[b] Analyses of the PAS data involved multiple imputations to account for missing data.

---

[41] Aggregating the data to the grantee level may have reduced power or obscured differences, so analyses were also conducted using the scores as collected at the classroom or center level as a sensitivity analysis; no reliable differences between designated and not-designated grantees emerged in those analyses either (see appendix B in technical volume II).

Table IV.3 shows the results of our analysis of the four measures of financial vulnerability based on the tax returns for 216 nonprofit grantees in the sampling frame.[42] The means and standard deviations by designation group, and results of the *t*-tests are shown, listing the *p*-value from the comparison of the means and effect size for those comparisons. The grantees designated and not designated for competition did not reliably differ on any of the measures of financial vulnerability. There is a suggestive finding that designated grantees had modestly higher revenue concentration ratios when compared with not-designated grantees.

TABLE IV.3.

Comparing Financial Vulnerability Ratios of Nonprofit Grantees by Designation Status

| Financial Vulnerability Ratio | Not Designated | | | Designated for Competition | | | *t*-test | |
	N	Mean	SD	N	Mean	SD	P	Effect Size[a]
Equity Ratio[b]	174	0.35	0.67	42	0.50	1.77	ns	-.15
Revenue Concentration[c]	174	0.81	0.19	42	0.85	0.17	+	-.21
Administrative Cost Ratio[b]	174	0.08	0.07	42	0.07	0.05	ns	.14
Operating Margin[b]	174	-0.01	0.18	42	0.00	0.03	ns	-.06

Sources: Data provided by the Office of Head Start and obtained from Internal Revenue Service Form 990 filings for 2012 available through the National Center for Charitable Statistics.
Notes: Significance noted as: + *p*<.10; ns=not significant.
[a] Effect sizes for comparison of means computed as the difference between the means divided by the pooled sample standard deviation.
[b] A higher number on the equity ratio, administrative cost, and operating margin denotes a lower level of financial vulnerability.
[c] A lower number on the revenue concentration denotes a lower level of financial vulnerability.

Comparisons by Specific Designation Conditions

The next set of analyses examined the two conditions that most commonly triggered designation for competition. Because about 99 percent of grantees are designated for competition based on either a deficiency or a low CLASS score, we conducted two analyses: (1) we compared 36 grantees not designated with 24 grantees designated for competition due to low CLASS scores and (2) we compared 36 grantees not designated with 14 grantees designated for competition based on deficiencies.[43] As

---

[42] Of the 368 grantees in the sampling frame, 248 (67 percent) are nonprofit organizations, indicated by their filing of an IRS Form 990 for 2012. Of those 248, 32 had missing information on their IRS Form 990 that prevented calculating the financial vulnerability analysis; thus, 216 were included in our analysis.

[43] We examine the individual conditions for designation because there was some, but not a lot, of overlap between grantees designated for CLASS and for deficiencies, (X^2(1, *N*=368) = 5.43, p=.02). In the entire monitoring cohort, 30 percent of grantees designated for deficiency were also designated for a low CLASS score, and 15 percent of grantees designated for low CLASS

before, we conducted these analyses using data aggregated to the grantee level, weighted the data to account for differential enrollment rates, and used multiple imputations to account for missing data.

Designation Due to Low CLASS Scores. First, we compared the quality of grantees that were designated because of low CLASS score to grantees that were not designated. Compared with the not-designated grantees, the grantees designated due to low CLASS scores did not have reliably lower scores on any of the evaluation's quality measures and, unexpectedly, were rated as having higher quality practices on the PAS Child Assessment Scale (table IV.4).[44]

TABLE IV.4.

Comparing Quality Scores for Grantees Not Designated and Grantees Designated for Low CLASS Scores

Quality Measure (collected by evaluation team)	Not Designated			Designated for Low CLASS Score			t-test	
	N	Mean	SD	*N*	Mean	SD	*P*	Effect Size[a]
Classroom Quality, Aggregated at Grantee Level								
CLASS Emotional Support	36	5.33	0.42	24	5.34	0.48	ns	-.02
CLASS Instructional Support	36	2.48	0.73	24	2.63	0.91	ns	-.18
CLASS Classroom Organization	36	4.63	0.61	24	4.59	0.64	ns	.06
ECERS-R Interaction	36	5.19	0.65	24	5.23	0.67	ns	.-06
ECERS-E Academic Activity	36	3.03	0.57	24	3.16	0.59	ns	-.22
TSRS Class Structure	36	3.73	0.52	24	3.80	0.46	ns	-.14
Center Quality, Aggregated at Grantee Level								
PAS Staff Qualifications	36	2.80	0.80	24	2.56	0.68	ns	.32
PAS Personnel Cost	36	3.17	0.74	24	3.25	0.93	ns	-.10
PAS Child Assessment	36	5.87	.98	24	6.37	0.87	*	-.54
PAS Fiscal Management	36	3.57	1.39	24	3.71	1.54	ns	-.10
PAS Family Partnerships	36	4.31	1.17	24	4.48	1.11	ns	-.15
PAS Governance and Management	36	3.06	0.89	24	3.39	0.76	ns	-.40
Health and Safety: Indoor	36	0.86	0.07	24	0.85	0.08	ns	.13
Health and Safety: Outdoor	36	0.80	0.10	24	0.79	0.11	ns	.10

Source: Data collected by evaluation team.
Notes: Analyses compared grantee-level mean scores, involved imputation for missing data, and were weighted for nonresponse. Significance noted as: + .1>*p*>.05; * *p*<.05; ns=not significant.
[a] Effect sizes computed as the difference between the means divided by the pooled sample standard deviation.

---

scores were also designated for deficiencies. Although the overlap is statistically significant, only three grantees were designated for both low CLASS scores and deficiencies. For this reason, we could not examine them as a separate group. They were included in both the comparisons examining designation for deficiencies and designation for low CLASS scores.
[44] Sensitivity analyses involving multi-level analyses also provided no evidence that grantees designated for low CLASS scores differed from not-designated grantees on classroom or center quality scores (see appendix B in technical volume II).

Similarly, on our measures of financial vulnerability, we found no reliable differences between not-designated nonprofit grantees and nonprofit grantees designated due to low CLASS scores (see table IV.5).

TABLE IV.5.

Comparing Financial Vulnerability Ratios for Nonprofit Grantees Not Designated and Nonprofit Grantees Designated Due to CLASS

Financial Vulnerability Ratios, All Nonprofit Grantees	Not Designated			Designated Due to CLASS			*t*-test	
	N	Mean	SD	*N*	Mean	SD	*P*	Effect Size[a]
Equity Ratio[b]	174	0.35	0.67	33	0.23	0.21	ns	-0.19
Revenue Concentration[c]	174	0.81	0.19	33	0.86	0.17	ns	0.26
Administrative Cost Ratio[b]	174	0.08	0.07	33	0.07	0.05	ns	-0.14
Operating Margin[c]	174	-0.01	0.18	33	-0.00	0.02	ns	0.04

Sources: Data provided by the Office of Head Start and obtained from Internal Revenue Service Form 990 filings for 2012 available through the National Center for Charitable Statistics.

Notes: Significance noted as: ns=not significant at the .10 level.

[a] Effect sizes for comparison of means computed as the difference between the means divided by the pooled sample standard deviation.

[b] A higher number on the equity ratio, administrative cost, and operating margin denotes a lower level of financial vulnerability.

[c] A lower number on the revenue concentration denotes a lower level of financial vulnerability.

Designation Due To Deficiencies. Next, we compared the quality of grantees that were designated because of deficiencies to grantees that were not designated. Analyses revealed grantees designated due to deficiencies had lower levels of quality than not-designated grantees on three out of six measures of center quality: PAS subscale of Staff Qualifications, Personnel Costs, and Fiscal Management (figure IV.1 and table IV.6). They did not differ from the not-designated grantees on measures of classroom quality.[45] These findings—on areas in which grantees designated due to deficiencies had lower levels of quality according to the PAS—are consistent with the specific types of deficiencies found among grantees in the sample. That is, OHS monitoring data[46] indicate that many grantees designated for deficiencies in our sample had issues with staff qualifications or fiscal management.

---

[45] Multi-level analyses of the analyses yielded the same findings when analyses involved imputation for missing PAS scores (see appendix B of technical volume II).

[46] The evaluation team examined OHS Administrative data tracking deficiency citations.

FIGURE IV.1.

Evaluation Team Measures of Center Quality Comparing Grantees Not Designated with Grantees Designated for Deficiency
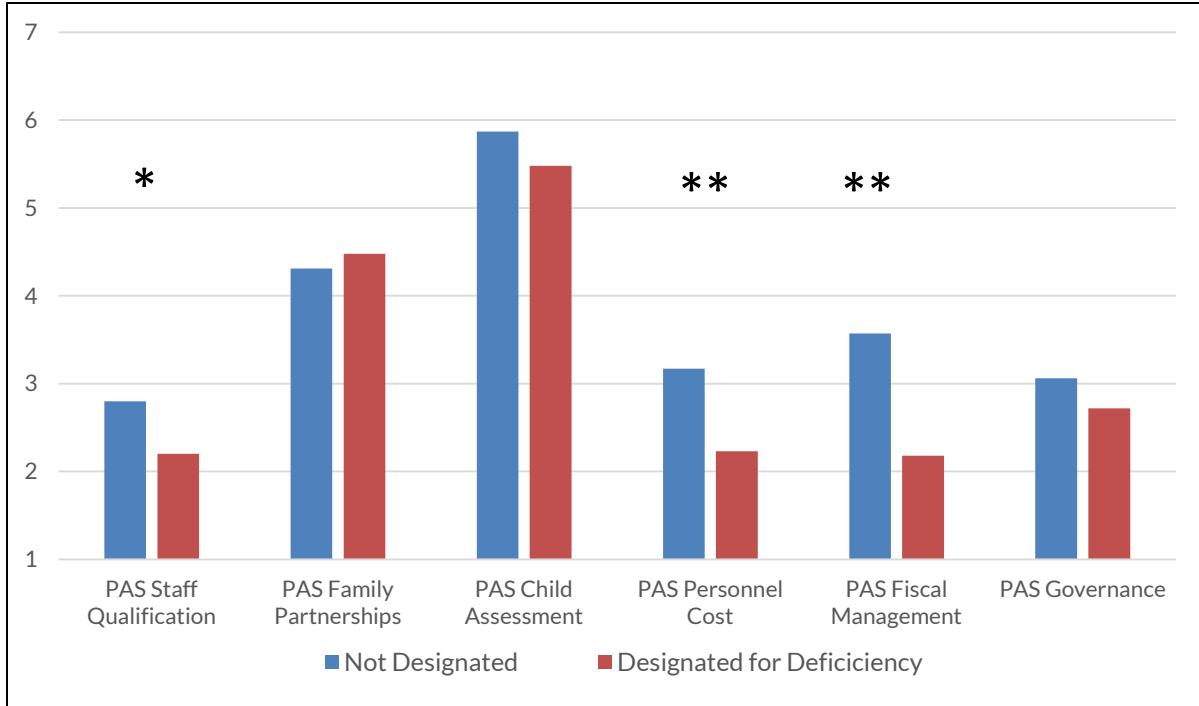
Sources: Data collected by evaluation team.

Notes: Analyses compared grantee-level mean scores, involved imputation for missing data, and were weighted for nonresponse.

Significance noted as: +.1>p>.05; * p<.05; ** p<.01; *** p<.001.

TABLE IV.6.

Comparing Quality Scores for Grantees Not Designated and Grantees Designated for Deficiencies

Quality Measure (collected by evaluation team)	Not Designated			Designated for Deficiencies			t-tests	
	N	Mean	SD	N	Mean	SD	P	Effect Size[a]
Classroom Quality, Aggregated at Grantee Level								
CLASS Emotional Support	36	5.33	0.42	14	5.19	0.43	ns	.32
CLASS Instructional Support	36	2.48	0.73	14	2.73	0.64	ns	-.35
CLASS Classroom Organization	36	4.63	0.61	14	4.55	0.49	ns	.14
ECERS-R Interaction	36	5.19	0.65	14	5.31	0.44	ns	-.20
ECERS-E Academic Activity	36	3.03	0.57	14	3.25	0.59	ns	-.37
TSRS Class Structure	36	3.73	0.52	14	3.77	0.34	ns	-.06
Center Quality, Aggregated at Grantee Level								
PAS Staff Qualifications	36	2.80	0.80	14	2.20	0.85	*	.72
PAS Personnel Cost	36	3.17	0.74	14	2.23	0.72	**	1.25
PAS Child Assessment	36	5.87	0.98	14	5.48	1.34	ns	.35
PAS Fiscal Management	36	3.57	1.39	14	2.18	1.48	**	.96
PAS Family Partnerships	36	4.31	1.17	14	4.48	1.33	ns	-.14
PAS Governance	36	3.06	0.89	14	2.72	0.80	ns	.38
Health and Safety: Indoor	36	0.86	0.07	14	0.87	0.05	ns	-.15
Health and Safety: Outdoor	36	0.80	0.10	14	0.80	0.09	ns	.00

Sources: Data collected by evaluation team.

Notes: Analyses compared grantee-level mean scores, involved imputation for missing data, and were weighted for nonresponse. Significance noted as: $+.1 > p > .05$; $* \ p < .05$; $** \ p < .01$; $*** \ p < .001$; ns=not significant.

[a] Effect sizes computed as the difference between the means divided by the pooled sample standard deviation.

We also conducted analyses of the financial vulnerability ratios of the nonprofit grantees to test for differences related to designation due to deficiencies. Comparisons indicated differences on one of four measures —the equity ratio that captures the grantee's level of assets available as a cushion. Nonprofit grantees designated due to deficiencies had higher equity ratios (and thus lower levels of financial vulnerability) than nonprofit grantees that were not designated, but an outlier in the distribution raised questions about this finding. There were no significant differences in equity ratios when the outlying score was trimmed. The small sample size for the designated for deficiency group makes their mean vulnerable to extreme scores.[47]

---

[47] The outlier equity ratio is for a grantee designated for competition. The ratio was 60 times larger than the median value.

TABLE IV.7.

Comparing Financial Vulnerability Ratios for Nonprofit Grantees Not Designated and Nonprofit Grantees Designated Due to Deficiencies

Financial Vulnerability Ratios, All Nonprofit Grantees	Not Designated			Designated for Deficiency			*t*-test	
	N	Mean	SD	*N*	Mean	SD	*P*	Effect Size[a]
Equity Ratio[b]	174	0.35	0.67	13	1.07	3.18	**[d]	0.68
Revenue Concentration[c]	174	0.81	0.19	13	0.84	0.20	ns	0.17
Administrative Cost Ratio[b]	174	0.08	0.07	13	0.07	0.06	ns	-0.13
Operating Margin[b]	174	-0.01	0.18	13	0.00	0.03	ns	0.08

Sources: Data provided by the Office of Head Start and obtained from Internal Revenue Service Form 990 filings for 2012 available through the National Center for Charitable Statistics.

Note: Significance noted as: + $p<.10$; * $p<.05$; ** $p<.01$; ns=not significant at the.10 level.

[a] Effect sizes for comparison of means computed as the difference between the means divided by the pooled sample standard deviation.

[b] A higher number on the equity ratio, administrative cost, and operating margin denotes a lower level of financial vulnerability.

[c] A lower number on the revenue concentration denotes a lower level of financial vulnerability.

[d] The equity ratio for one grantee was extremely different from those of all other grantees. The differences between the grantees designated and not designated for competition was smaller and not statistically significant when the data for that grantee was trimmed.

Next, we examine the CLASS condition more closely in terms of its performance as a measure of grantee-level quality.

Finding 2: The CLASS scores were significantly higher when collected by OHS monitors than by the evaluation team. There were no significant differences in CLASS scores related to the number of cycles collected per classroom or whether classrooms were sampled, but there were substantial differences among raters.

In addition to examining whether grantees that were designated for competition differed from grantees that were not designated on independent measures of quality, we undertook additional psychometric analyses of the CLASS. The addition of the CLASS to Head Start monitoring represents a substantial change in the way ACF monitors Head Start grantees and previous research has not examined the

psychometric performance of CLASS as a measure of quality at the grantee level. For this reason, we examined the reliability of the grantee-level CLASS scores.[48] Reliability reflects the extent to which the observed scores provide stable estimates across raters and time, examining whether comparable information would be obtained regardless of the individual observer and when the observation is conducted. Furthermore, the results from the analyses reported above suggested that grantees designated due to low CLASS scores did not significantly differ from not-designated grantees on any independently collected quality measure, including the CLASS and other classroom quality measures as measured by the evaluation team. Thus, a deeper examination of the CLASS scores is warranted. We examined four issues: (1) the stability of grantee-level CLASS scores as evidenced by similarities between OHS CLASS observer scores and the evaluation team's observer scores; (2) the proportion of variability in CLASS scores attributable to differences among observers, relative to differences among grantees; (3) the use of two rather than four cycles for classroom observations to determine whether changes in the standard protocol for the CLASS in measuring quality at the classroom level affected scores; and (4) the use of sampling classrooms to determine whether sampling, rather than observing all classrooms within a grantee, resulted in bias in measuring quality at the grantee level.

The evaluation team used similar training and certification procedures for the CLASS observers and similar sampling strategies for determining which classrooms to assess as did the OHS monitoring team. These similarities were designed to limit the possibility that comparisons might result in differences between observers due to differences in methods and procedure.

Similarities between OHS and Evaluation Study Team CLASS Scores

We examined the similarities and differences between the CLASS scores as collected at the grantee level during OHS monitoring visits and during the evaluation study. We compared the grantee-level scores as reported for the 70 grantees[49] with CLASS data from both sources to ask whether the scores from the two sources agreed in terms of the level of quality of teacher-child interactions (i.e., mean differences) or in terms of their variability (i.e., correlations). The correlations describe the extent to which the relative ranking of the grantee is maintained across the two sources of CLASS data, whereas

---

[48] Reliability estimates should include all possible sources of variability. We typically measure variability among raters and day-to-day variability in classroom activities or teacher behaviors that need to be observed for accurate assessments separately, but each is just one component of the overall reliability of a measure and neither should be regarded as representing the overall reliability of measurement of that instrument.
[49] One grantee that was designated due to deficiencies did not have a CLASS score, resulting in only 70 of the 71 grantees with a CLASS score.

the mean comparisons of the grantee's scores addresses questions regarding whether the two teams of observers provided similar CLASS scores overall.

Results suggest there is considerable lack of agreement in CLASS scores across the two rating teams. Results are presented in table IV.8 separately by designation status.[50] Correlations between the grantee-level ratings from the two sources tended to be modest and nonsignificant, except for associations for the CLASS Instructional Support and Classroom Organization in the not-designated for competition group. These significant correlations indicate moderate agreement in the variability of these two CLASS domain scores among grantees not designated for competition. That is, grantees that OHS raters rated higher were also rated higher by the evaluation team. With the *t*-test (matched-pairs *t*-test), we asked whether the average CLASS domain scores are reliably different across teams of CLASS observers. The results show that among grantees designated for competition and grantees not designated, the OHS CLASS observers rated grantees significantly higher than did the evaluation team. Specifically, OHS CLASS observers rated grantees not designated for competition significantly higher on all three CLASS domains than the evaluation team. For grantees designated for competition, OHS CLASS observations were significantly higher on two of the three CLASS domains (see figure IV.2). The differences were quite large—ranging from .5 to 2.0 standard deviations. It is not possible to say whether the ratings of one team are more accurate, only that there is a notable difference.

---

[50] Analyses were conducted separately by designation status because the sample was stratified by designation status and, thus, only provides generalization from the sample to the population within the stratification groups. Weighting the grantees designated for competition to represent all designated grantees may not be possible given the relatively small sample of designated grantees, and combining the groups without weighting would misrepresent the data from the full sample of grantees due to overrepresentation of the designated group (i.e., 50 percent in evaluation sample instead of 30 percent in sampling frame).

TABLE IV.8.

Examining Similarities by Designation Status between CLASS Scores Collected by OHS Observers and the Evaluation Study

	OHS CLASS Observers Mean (SD)	Evaluation Study Mean (SD)	Correlation	*t*-test	Effect size
Not Designated for Competition (*N*=35)					
CLASS Emotional Support	6.13 (0.23)	5.33 (0.42)	.20	10.7***	1.90
CLASS Instructional Support	2.90 (0.49)	2.50 (0.73)	.53**	3.67***	0.54
CLASS Classroom Organization	5.88 (0.32)	4.64 (0.61)	.46**	13.4***	2.03
Designated for Competition (*N*=35)					
CLASS Emotional Support	5.77 (0.31)	5.34 (0.49)	.21	4.78***	0.88
CLASS Instructional Support	2.46 (0.52)	2.71 (0.93)	.04	1.39	-0.26
CLASS Classroom Organization	5.43 (0.40)	4.65 (0.60)	.16	6.92***	1.30

Source: The table includes data from the grantees included in the evaluation study collected by the OHS CLASS observers and by the evaluation team.

Notes: All data were summarized at the grantee level. Significance noted as: +.1>*p*>.05; * *p*<.05; ** *p*<.01; *** *p*<.001.

FIGURE IV.2.

CLASS Scores: Comparing OHS Monitoring Scores and Evaluation Study Scores
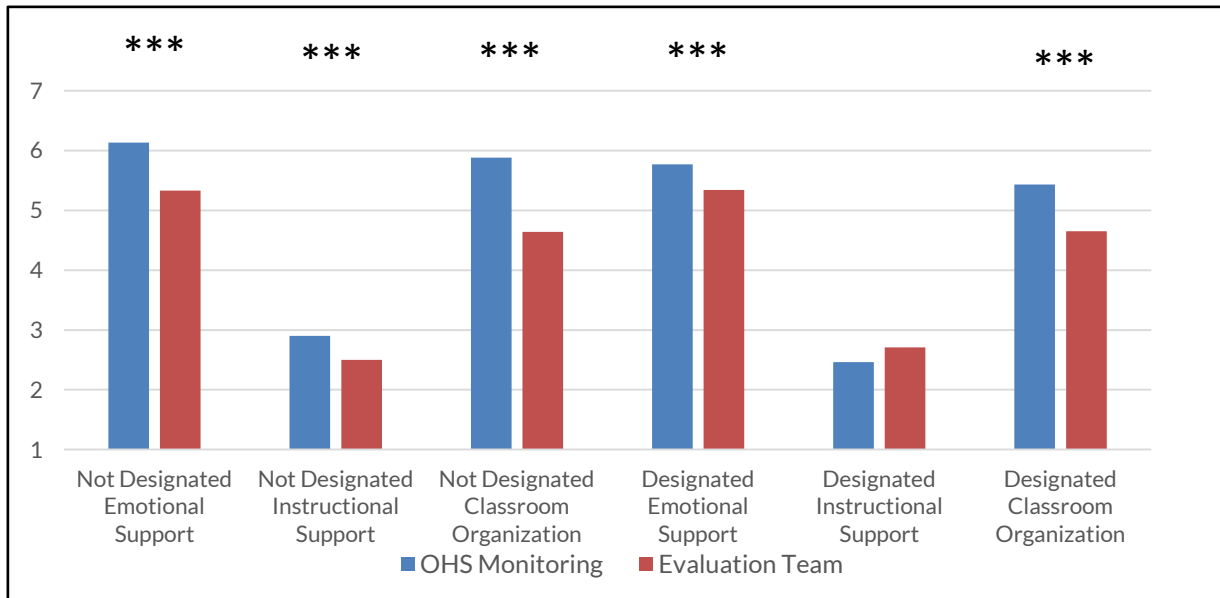
Source: The figure includes data from the grantees included in the evaluation study collected by the OHS CLASS observers and by the evaluation team.

Notes: All data were summarized at the grantee level. Significance noted as: +.1>*p*>.05; * *p*<.05; ** *p*<.01; *** *p*<.001.

Next, we explored whether these discrepancies in CLASS scores between the two teams of observers could reflect changes in CLASS scores that occurred between when the OHS monitors and the evaluation team collected their data. We sought to minimize the time between the OHS and evaluation teams' data collection, but there was considerable variability across grantees. Sensitivity analyses indicated that the difference in when the data were collected by the two teams did not account for the observed differences between the OHS monitors and the evaluation team (see technical volume II, appendix B for details).

Variability among Observers and Grantees in Evaluation Team CLASS Scores

The evaluation team also looked at the reliability of CLASS scores as a measure of the quality of teacher-child interactions at the grantee level by the proportion of variability among grantee scores that is due to differences among CLASS observers. Using the evaluation study data, we computed intra-class correlations to describe the degree of variability among grantees and among observers. Two observers collected CLASS data across the selected classrooms within most grantees in our sample (ranging from one for small grantees to two to three observers for large grantees). We conducted a general linear mixed-model analysis to estimate variance components related to observer and grantee. We see considerable variability among observers within grantees (table IV.9). If we assume that observers were assigned to grantees randomly (or in a manner that is roughly equivalent to randomly), then findings suggest that variability among observers accounts for up to 45 percent of the variability in the CLASS scores and other classroom and center measures of variability among grantees explained little of the variation on all of these observational quality measures. Follow-up analyses that demonstrated much higher correlations among measures collected by the same raters than among measures assessing comparable constructs also support concern about substantial rater effects (see appendix B of technical volume II for details).

TABLE IV.9.

Proportion of Variation in Evaluation Study Scores Attributable to Differences among Observers and Grantees

Quality Measure	Variability Among Observers	Variability Among Grantees	Residual Variability
CLASS Emotional Support	41%	10%	49%
CLASS Instructional Support	45%	9%	46%
CLASS Classroom Organization	43%	10%	47%
ECERS-R Interactions	46%	9%	44%
ECERS-E Academic Activities	44%	25%	30%
TSRS Management	40%	48%	13%
PAS Total [a]	40%	10%	49%

Source: Data were collected by the evaluation team.
Note: Analyses included CLASS, ECERS-R, ECERS-E, and TSRS scores from 549 centers in 71 grantees.
[a] These analyses included data from 40 grantees with PAS data collected for multiple centers within that grantee.

Number of CLASS Cycles

Another possible threat to reliability involved a change in the standard CLASS protocol in the way OHS monitoring conducts CLASS data collection. After consultation with the CLASS developers at the University of Virginia, OHS decided that OHS CLASS observers would collect two cycles of data in preschool classrooms, rather than the four cycles typically recommended by the CLASS developer, to allow for observations of more classrooms per grantee (thereby capturing more variability within grantees, because CLASS is used at the grantee level in DRS). Since the evaluation study collected four cycles per classroom, we used our data to compare ratings for each classroom based on two versus four cycles. We randomly identified the first, second, or third cycle in the evaluation classroom data, and then computed CLASS scores based on two consecutive cycles beginning with the randomly selected cycle. We computed CLASS scores for each classroom in the evaluation study based on either two or four cycles. We then compared the scores in two ways to see if it appears that using two cycles biased findings or reduced precision.

First, we asked whether the scores for the classrooms in the evaluation study differed when computed based on two or four cycles. Repeated measures analyses tested whether the scores differed, using methods that took into account the nesting of classrooms in grantees. No statistically significant differences were detected in these analyses (table IV.10). No evidence emerged suggesting that using two cycles resulted in bias in the CLASS data. Findings suggest evaluation ratings are the same when

classrooms are observed for two cycles or four cycles. However, these findings suggest that use of two cycles may have slightly increased variability at the grantee level (and thus reduced overall reliability) when the same number of classrooms are included. On the other hand, the use of two cycles may not have increased the variability in the OHS ratings in grantees with more than eight classrooms, because the use of two cycles allows OHS to include more classrooms when estimating the grantee-level score, and increasing the number of classrooms will reduce variability even if each classroom score is slightly less reliable (Lambert et al. 2006).

TABLE IV.10.

Comparing CLASS Scores Based on Number of Classroom Observation Cycles

	2 Cycles		**4 Cycles**		**Paired *t*-test**
CLASS Domain **(collected by evaluation team)**	**Mean**	**SD**	**Mean**	**SD**	*P*
Emotional Support	5.34	0.82	5.33	0.76	ns
Instructional Support	2.58	1.26	2.56	1.18	ns
Classroom Organization	4.63	1.14	4.63	1.05	ns

Source: All data were collected by the evaluation team in 549 classrooms.
Note: Significance noted as: +.1>*p*>.05; * *p*<.05; ** *p*<.01; *** *p*<.001; ns=not significant.

Second, we asked whether the CLASS scores from the evaluation study correlated differently with the CLASS scores from OHS CLASS observers when the evaluation study CLASS scores were computed based on two or four cycles of observations. These correlations were conducted separately by designation status.[51] Again, no statistically significant differences were detected in these analyses (table IV.11) suggesting that using two cycles rather than four cycles does not affect the CLASS scores and does not seem to account for the low levels of agreement between the two data collection teams.

---

[51] The variability of OHS CLASS scores was restricted among grantees designated for competition based on the CLASS because they all had scores that fell below a specified cut point.

TABLE IV.11.

Comparing Correlations between OHS Monitoring Scores and Evaluation Study Scores on the CLASS with Evaluation Study Scores Computed Based on Either Two or Four Cycles, by Designation Status

	Not Designated			Designated for Competition		
CLASS Domain	2 Cycles	4 Cycles	Comparisons	2 Cycles	4 Cycles	Comparisons
Emotional Support	.21	.20	ns	.28	.23	ns
Instructional Support	.54***	.53**	ns	.11	.05	ns
Classroom Organization	.50**	.46**	ns	.23	.18	ns

Source: The table includes data from the grantees included in the evaluation study collected by the OHS CLASS observers and by the evaluation team.

Notes: Table presents the correlations between CLASS scores from OHS monitoring and CLASS scores from the evaluation team. Separate correlations are presented for evaluation team CLASS scores computed using data from all four cycles, and from only two of the four cycles of evaluation team classroom observation. Results are presented separately by designation status because of the potential that the range of OHS CLASS scores is restricted among grantees designated based on the CLASS. Significance noted as: $+ .1 > p > .05$; $* p < .05$; $** p < .01$; $*** p < .001$; ns=not significant.

Sampling Classrooms for CLASS Assessments

Another possible threat to the reliability of the grantee-level CLASS scores involved the use of sampling to determine which classrooms were observed during OHS monitoring visits. To address this concern, we estimated the correlations between the OHS CLASS scores and the evaluation CLASS scores for grantees in which we did and did not observe all classrooms. If sampling led to differences in mean scores or substantial reduction in precision, the correlations for the grantees in which all classrooms were observed should be substantially larger. We see no evidence that sampling classrooms appeared to systematically dilute these correlations (table IV.12). Furthermore, it does not appear that whether the two teams observed the same classrooms accounts for the low levels of agreement between the OHS and this study's data collection teams.

TABLE IV.12.

Comparing Correlations between OHS Monitoring Scores and Evaluation Study Scores on the CLASS, by Whether Classrooms Were Sampled or Not

CLASS Measures	No Sampling of Classrooms (*N*=14)	Classrooms were Sampled (*N*=56)	Comparisons
CLASS Emotional Support	-.02	.25	ns
CLASS Instructional Support	.21	.33	ns
CLASS Classroom Organization	.49	.25	ns

Source: The table includes data from the grantees included in the evaluation study collected by the OHS CLASS observers and by the evaluation team.

Notes: Table presents the correlations between CLASS scores from OHS monitoring and CLASS scores from the evaluation team and compares those correlations between grantees in which all classrooms were observed and grantees in which a subset of classrooms was randomly selected for observation.

Significance noted as: +.1>*p*>.05; * *p*<.05; ** *p*<.01; *** *p*<.001; ns=not significant.

Summary

The classroom and center quality of grantees designated for competition for any reason was not significantly different from the quality of not-designated grantees, but quality of overall center operations was significantly lower in grantees designated for deficiencies than in not-designated grantees. Among grantees designated for deficiencies, the PAS subscales measuring the quality of staff qualifications, personnel costs and allocation, and fiscal management were rated lower than not-designated grantees. Grantees designated for low CLASS scores did not have significantly lower scores than not-designated grantees on any quality measures assessed in this study. Further examination of CLASS scores indicated that the OHS monitors rated grantees significantly higher than did the evaluation team, and scores across the two teams were only modestly associated. Within the evaluation team data, analyses indicated that use of two instead of four cycles and the use of sampling of classrooms did not seem to change grantee's CLASS scores, but that up to 45 percent of the variance in quality scores was related to differences among raters.

Next, we explore how grantees perceive the monitoring, assessment, and DRS designation processes in terms of their efforts to improve program quality.

Chapter V. How do grantees perceive the monitoring, assessment, and DRS designation processes in terms of their efforts to improve program quality?

In this chapter, we explore how grantees experienced the early implementation of the monitoring and DRS designation process, and the types of quality improvement activities they undertook in response to participating in that process. Whereas the findings in chapter III were focused on the perceptions and actions before monitoring (how did grantees prepare for DRS monitoring?), the findings in this chapter are focused on the perceptions and actions after monitoring and when designation status was known. This chapter focuses on understanding how the DRS may be incentivizing ongoing and continuous quality improvement activities and not just those in preparation for monitoring. That is, grantees' understanding of the monitoring and assessment process and their perceptions of how well the DRS conditions reflect the quality of their program may support or impede incentives to continuously improve quality.

Overview of Methods

Our approach to addressing this research question was qualitative and exploratory. Our goal was to try to understand a range of perceptions and experiences of the early implementation of the DRS and not necessarily to represent the population of grantees that participated in the DRS assessments. Throughout this chapter, we provide information about the proportion of grantees we spoke to that held similar perceptions of their experience with DRS, but this does not suggest the same would be found within the general population of Head Start grantees. In other words, if half of grantees responded similarly in this study, it may be the case that less than half or more than half of all Head Start grantees that experienced the DRS would have responded similarly.

We examined the perspectives of grantees within this early implementation of the DRS through two sets of qualitative interviews. First, 35 program directors (half from grantees designated for competition; half from grantees not designated) participated in semi-structured qualitative telephone interviews shortly after their monitoring visit in 2013—14.[52] Then, shortly after all grantees learned their designation status but before funding opportunity announcements were issued, a subset of those grantees (*n*=15; half designated for competition, half not) participated in on-site follow-up interviews with key personnel, including directors, managers, governing body members, and Policy Council members. We asked grantees about their understanding of the purpose of the DRS, its conditions, and the monitoring and assessment process, as well as what ways they had responded to the DRS through quality improvement efforts.

Qualitative interviews were coded to identify themes about respondents' understanding of the DRS, actions taken to improve quality after their monitoring review, and the ways that those actions were motivated by the DRS. In this chapter, we focus on themes that emerged related to grantees' experiences with the monitoring and assessment process up through learning their designation status. See chapter III for a focus on the time period before their first monitoring review for the DRS, chapter VI for a discussion of grantees' experience with the competitive process, and chapter VII for a discussion of grantees' experiences with the negotiation, award, and transition processes.

We coded by designation status and by type of respondent and looked for variations in themes across groups, as well as within and across sites. We report themes at the grantee level, but point out where within-grantee differences or other variations of note occurred. See chapter II and technical volume II, appendix C for additional details on the study design, sample, measures, and analyses.

Findings

Two key findings emerged:

1. ***All 15 of the Head Start grantees interviewed after they knew their designation status indicated they were continuing to work on improving quality. Grantees designated for***

---

[52] The goal of the evaluation was to conduct the Director Telephone Interviews before grantees knew their designation status. On average, Director Telephone Interviews were conducted six months following monitoring and after the quality assessment visits had been conducted. A total of 19 interviewees may have known their designation status at the time of the interview: eight because they had received deficiencies and 11 because interviews were conducted following the public announcement of designations (six not designated and five designated due to CLASS). In the interview, we did not ask if they knew their designation status, as we did not want interviewees focused on that, so we do not know definitively who knew and who did not know their status.

competition were more likely to indicate some of their quality improvement efforts were related to the DRS.

2. *Grantees expressed a number of concerns about the early implementation of the monitoring, assessment, and notification processes.*

Each of these findings is discussed in more detail below with some exemplar quotes. The designation status of the grantee providing the quote is indicated to provide some sense of diversity of respondents, but other characteristics are not indicated to ensure confidentiality.

Finding 1: All 15 of the Head Start grantees interviewed after they knew their designation status indicated they were continuing to work on improving quality. Grantees designated for competition were more likely to indicate some of their quality improvement efforts were related to the DRS.

In the on-site interviews, we asked questions about the types of changes grantees may have made in response to the DRS *after* their program's designation status was known (either designated for competition or not). We asked grantees to reflect on programmatic or administrative changes, including changes to their administrative structure, fiscal controls or processes, provisions and procedures for meeting health and safety standards, record-keeping practices, types of technical assistance, and any other changes. All the grantees were uniform in stating that they did not make any changes to fiscal controls or processes, or record-keeping practices. Grantees designated for competition were more likely to indicate changes post-designation status that related to the DRS than grantees not designated (table V.1). In the interviews, the grantees designated for competition typically related the changes directly to their reason for designation.

Among Head Start directors in grantees designated for competition, six of seven indicated that since designation for competition, they had been working to make some quality improvements in their program motivated by the DRS. The other Head Start director in a grantee designated for competition indicated they were continuing quality improvement efforts they had begun almost a decade earlier, but they were not motivated by the DRS. Among Head Start directors in grantees not designated for competition, only two of eight indicated they were focusing on some quality improvement activities

specifically because of the DRS. One other Head Start director indicated her site was making some changes to improve quality in relation to areas that the Office of Head Start had recently been emphasizing (e.g. improving CLASS scores, improving school readiness scores, and improving parent engagement), but she did not view these changes as a direct response to the DRS. The remaining five sites indicated that they are always working to improve quality but did not provide specific examples.

Two of the seven grantees designated for competition reflected that the DRS brought a change in the way they think about planning and delivering services. One indicated that her mindset had changed from quantity (serving as many children as possible) to quality. Another said,

> *It makes you look at things in a different way. It changes your scope. We have to consciously think about how we will plan to do things in a different way. …We've learned what to stop doing. That's a critical piece you sometimes forget. Sometimes you're doing things just to do them because you've always done them. We have to sit back and say we're going to stop doing it because we don't need to do that anymore. —Designated for Competition*

The Head Start director who indicated they were creating a new position to focus on compliance echoed a similar sentiment,

> *What I found through entire process was that no matter how good your systems are, and how much oversight you have, the people closest to it don't always see what —aren't always best to evaluate little things. Center directors love the teachers who work for them, and they miss some of the subtleties. We needed some mechanism to have someone not involved in day to day being able to do spot checks. —Designated for Competition*

TABLE V.1.

Grantees' Post-Designation-Status Quality Improvement Plans, by Designation Status

Planned Quality Improvement Effort	Grantees Designated for Competition (*N*=7)[a]	Grantees Not Designated for Competition (*N*=8)[b]
Efforts Specific to CLASS (*n*=6)		
Provide coaching or improve coaching related to CLASS	3	1
Start using the Making the Most of Classroom Interactions (MMCI)	1	0
Establish CLASS domain score thresholds in job descriptions	1	0
Train support staff on CLASS	0	1
Efforts Not Indicated as Related to CLASS (*n*=5)		
Increase individualization of staff training	2	0
Training for substitutes on positive guidance	1	0
Rethinking staff orientation plan	1	0
Establishing an internal position focused on ensuring compliance	1	0
Establishing new practices for ensuring children are not left alone	1	0
Providing a 50 cent per hour raise for direct service staff	1	0

Source: Head Start director responses from the On-Site Follow-Up Interviews when asked if their program had made or had plans to make quality improvements related to the DRS.

Notes: Interviews were conducted after grantees knew their designation status.

[a] Six of seven grantees indicated changes. Some indicated more than one change.

[b] Only two of eight grantees indicated changes related to the DRS.

Four of the seven designated grantees also viewed the competitive application as an opportunity to redesign their service delivery to better serve the community. The grantees perceived it as easier to propose changes to the ages of children served and to the types of services. Three of these four grantees indicated that they planned to add Early Head Start services for the first time or add slots to existing Early Head Start services. One grantee planned to increase 3-year olds served and decrease 4-year olds served due to expansion of state-funded prekindergarten in the community. One of the four planned to reduce the level of home-based services because the needs in the community had changed. Interviews with new awardees reported in chapter VI yielded similar findings.

Finding 2: Both grantees designated and not designated for competition expressed a number of concerns about the early implementation of the monitoring, assessment, and notification processes.

In the 35 telephone interviews and 15 site visits, interviewees expressed a number of concerns about the monitoring, assessment, and notification processes of DRS early implementation. We grouped the concerns that grantees expressed into four categories: (1) concerns about the assessment process, (2) concerns about determining quality based on a single site visit, (3) concerns about transparency, and (4) concerns about not knowing their designation status. Typically, the Head Start director and other people we spoke with echoed the same types of concerns, which is not surprising given that interviewees indicated the Head Start director is typically the primary person in the organization who explains and shares information about the DRS to other staff, Policy Council, and governing body members. Governing body members expressed some distinct concerns noted at the end of the chapter.

Some grantees expressed concern about the OHS assessment process

Twenty-two directors expressed concerns in either the 35 telephone interviews or the 15 on-site visits about the assessment process, including directors from grantees that were and were not designated for competition. Most of those concerns focused on the CLASS assessment process, but some directors also expressed concerns about the deficiency assessment process (table V.2). We report here the concerns expressed by directors and note that some of these situations, as described, would be violations of assessment review protocols.

TABLE V.2.

Grantee Concerns about the OHS Assessment Process

Assessment Concern	Number Expressing Concern
Concerns Related to CLASS Assessments	
Timing of the Visit[a]	2 of 35 grantees in telephone interviews
Regrouping of Children[a]	3 of 15 grantees in on-site visits
Concerns about Culture and Language[a]	2 of 35 grantees in telephone interviews
Observer Actions and Behaviors	4 of 15 grantees in on-site visits
Inadequate Training or Knowledge of Monitoring and Assessment Processes	6 of 15 grantees in on-site visits
Variations in Deficiency Classification	5 of 35 grantees in telephone interviews

Source: Telephone and On-Site Follow-Up Interviews with grantees.

[a] The CLASS protocol provides specific guidance around the allowable timing of assessments in regard to weeks after the start of the program year, weeks before the end of the program year, and weeks after the start of a new teacher (see the CLASS Field Guide at https://eclkc.ohs.acf.hhs.gov/hslc/grants/monitoring/docs/fy2016-CLASS.pdf and https://eclkc.ohs.acf.hhs.gov/hslc/hs/sr/class (see how CLASS is used in Head Start reviews)).

When grantees expressed concerns about the CLASS assessments, they were expressing those concerns based on what they observed in their own program sites. These grantees expressed to the evaluation team that they had visits where they questioned the timing (one based on a date too close to the start of the program year and one for a new teacher), saw regrouping of children (e.g., bringing children from multiple classrooms into a single room), observed that a reviewer could not speak the language used in the classroom, felt an observer did not understand the cultural dynamics of the Native American children in the classroom, and observed other behaviors and actions that caused them to question how well trained and prepared assessors were for their visits. OHS's CLASS protocol provides specific guidance around many of these issues, including the allowable timing of assessments with regard to weeks after the start and before the end of the program year, as well as weeks after the start of a new teacher.[53] In some cases, grantees raised concerns about behaviors consistent with the OHS CLASS protocol, while other concerns suggest there were violations of the protocol. The concerns expressed related to deficiencies, however, are based more on what grantees hear from their peers about what could count as a deficiency than based on their own experiences; we do not report those here.

---

[53] See https://eclkc.ohs.acf.hhs.gov/hslc/grants/monitoring/docs/fy2016-CLASS.pdf and https://eclkc.ohs.acf.hhs.gov/hslc/hs/sr/class (see how CLASS is used in Head Start reviews).

Some grantees expressed concerns about determining quality based on a single visit or incident.

Twenty grantees, both grantees designated and not designated for competition, expressed concerns about the importance of a single visit to their program or single incident reported in assessing their program quality and determining whether the grant would be designated for competition (table V.3). Some grantees thought a corrective action approach would prove more helpful in both assessing and improving quality. Some expressed that the visits did not provide enough time for them to demonstrate all their quality efforts or that reviewers were not interested in looking at program elements that really showed their quality efforts. Some expressed concern about the influence a single CLASS reviewer could have in determining the designation status of their programs. Grantees also expressed concern that single incidents could cause designation even if programs had a history of compliance. Some grantees were concerned about the use of self-reported incidents for deficiency findings. Some grantees suggested that teachers and kids have bad days and therefore a review on a single day may not be representative of every day quality. Grantees also suggested that DRS conditions should be differentially weighted and a few suggested that everyone should have to compete for funding.

TABLE V.3.

Grantee Concerns about a Single Visit or Single Incident Determining Designation Status

Type of Concern	Number Expressing Concern
Corrective Action Plans Would Be Better	19 of 35 grantees in telephone interviews
Focused on Weaknesses Rather Than Strengths	18 of 35 grantees in telephone interviews
Single CLASS Review	11 of 35 grantees in telephone interviews
Self-Reports Cause Designation	5 of 35 grantees in telephone interviews
History of Program Compliance Not Considered	7 of 15 grantees in on-site visits
Teachers and Kids Have Bad Days	5 of 15 grantees in on-site visits
DRS Conditions Should Not All Have Same Weight	5 of 15 grantees in on-site visits
Everyone Should Compete	6 of 35 grantees in telephone interviews

Source: Telephone and On-Site Follow-Up Interviews with grantees.

TYPES OF CONCERNS

Corrective Action Plans Would Be More Effective. Some Head Start directors (19 of 35 telephone interviews) and program managers (4 of 15 sites visited) felt that the DRS should offer a corrective action process and probationary period to help grantees improve quality or rectify infractions. An improvement plan would be better than, *"throwing them into recompeting, spending time and energy and money writing grants... It's a lot of work on the staff, a lot of stress, and that affects teachers and all the way down to children."* It should be noted that, although some findings reach deficiency status immediately, corrective action periods already exist for findings of noncompliance in monitoring. That

is, grantees are given a period of time to take corrective action to address noncompliances and only unresolved areas of noncompliance become deficiencies which would lead to designation for competition in DRS (Head Start Act, Sec. 637 [42 U.S.C. 9832]). In fact, program managers at a site not designated for competition shared that they developed a correction plan because of a previous noncompliance issue. They believe developing this plan and the noncompliance experience made them a stronger program and better prepared for their monitoring review.

Focused on Weaknesses Rather Than Strengths. About half the Head Start directors that participated in the 35 telephone interviews (18 of 35) stated that they did not feel the DRS assesses program quality well. Six Head Start directors thought the DRS focuses more on program weaknesses and does not allow a program to highlight its strengths. One Head Start director stated, *"I think we need to look more strengths-based instead of looking at what's bad in a program."* These Head Start directors felt the DRS failed to measure their programs' unique strengths, such as their ability to analyze child data, engage parents, and foster partnerships within their school districts or communities. When asked what strengths the DRS had not measured, managers at one site responded:

> *The fact that you have these partners where you can multiply the money and impact. You can maximize the dollars…that's not mentioned…I don't think there's anything about family engagement. We hire our parents. We're at 37 percent of staff as current or former parents. We send them to school. We want to break the cycle of poverty. Everything we do is not just hiring professional people but also families. That's nowhere in this system. They're not looking at what you're doing for families and making them self-sufficient…The more you work you can see the impact. When we get the parents off welfare that impacts the community we serve. They (the DRS) only focuses on one thing. —Designated for Competition*

This sentiment also was raised during the site visits as staff at 4 of 15 grantee sites visited (both designated and not designated for competition) stated that they felt DRS did not adequately capture aspects of their program's quality. In describing their experience, two of these grantees at sites not designated for competition shared that they went above and beyond to prepare for the monitoring review, but they felt that the reviewers did not thoroughly review their files or records or offer their programs an opportunity to demonstrate quality. Program managers at one site (not designated for competition) described feeling rushed in that the reviewers did not have enough time to thoroughly look at their files and did not capture their program's quality well:

> *I think there are a lot of things we do that go above and beyond that aren't looked at. In the education section, there were questions I was waiting to be asked and nobody really cared. The way I tracked and monitored lesson plans, education activities that I sent home to families, that's fine but… We did a lot of work with the parents and staff—it sunk in this past year exactly what we were supposed to do. We did a survey of people in the health community involved. Nobody said anything about it. It was good it helped our children and families, but nobody cared… They were just interested in the files and the dates and the screenings. That and doing their safety checks. —Not Designated*

Eight of 15 grantees that participated in the on-site visits expressed concern that the DRS does not measure program quality well because the monitoring reviews capture only a *"small snapshot"* rather than looking *"at the big picture"* of program quality. These Head Start directors shared the opinion that time spent on-site conducting the DRS assessments did not allow them to show the reviewers all the things they wanted to share, exhibit best practices, or encourage programs to do innovative things.

> *Sometimes I think they need to look at the bigger picture instead of tiny things. You may have one little thing missing—a piece of paper or sign—you could have a sign-in sheet paper missing, but they're not looking at the big picture and all the work that was put in…it's a shame to be penalized for one tiny thing when you've done a million wonderful things. —Not Designated*

The experience of Head Start directors and program managers at 6 of 15 grantee sites left them feeling the monitoring reviewers were not looking for the right things related to program quality; in many instances, the reviewers were only looking for compliance. Head Start directors mentioned that the DRS does not give them credit for their efforts to hire well-educated staff, leverage additional dollars, and engage parents and the community. Per the telephone interview data, one Head Start director indicated her program was cited for not mentioning transportation in one brochure, yet they transport about half of their children. As expressed by that Head Start director, "The concern did not reflect the reality of the program." Another Head Start director expressed concern about how the school readiness goals condition was assessed in her program:

> *We're all required to have school readiness goals. But on the review, all they asked was do you have goals and were parents involved in writing them and show me the documentation. They didn't spend any time looking at the actual goals. Is that a measure of quality? I could have really crappy goals and be deemed a quality program. —Designated for Competition*

Single CLASS Review. During the telephone interviews about one-third of Head Start directors interviewed (11 of 35) expressed concerns that just one person—the OHS CLASS observer—could determine the designation status of their program. One program manager at a site not designated for competition suggested that observers need to collect a large amount of information and at times may make mistakes.

> *The observer that comes in comes from another walk of life and another state… I'm not sure whether the people who come from CLASS get their information right. I wonder whether they might make a mistake —human error—and then put a program up that isn't actually supposed to be on there. —Not Designated*

History of Program Compliance. Of particular concern among 7 of 15 grantees was the perception that one infraction on one day can cause a program that has had no history of problems (or has a history of positive reviews) to be designated for competition. The Head Start director at a not-designated grantee shared these comments:

I also think there should be some recognition that humans run these programs and make mistakes. Where does this rise on the continuum of competition? Is it serious enough to recompete given a long history of quality? Things can happen and I don't think there is enough recognition looking at the history of the quality of the program. I think that worries people. Someone could walk in and see one thing that either someone didn't notice or doesn't typically happen and you could get recompeted. That's what I think our staff worry about and puts a lot of pressure on people. —Not Designated

Self-Reports Cause Designation. Another area of concern is the treatment of self-reports for health and safety incidents. Grantees perceive that self-reports are treated the same as external reports or monitoring discoveries. Head Start directors in 5 of 35 grantees expressed concerns that self-reporting guarantees that a program will be designated for a deficiency. Hence, the general sentiment seemed to be that following the self-reporting rules can unfairly get a program designated. Those who were designated because of this felt punished for following the rules. A few respondents (6 of 35 grantees) even suggested that many programs were beginning to (or would in the future) fail to report these incidents to avoid designation. As one Head Start director from a designated grantee shared,

The whole idea of self-reporting for health and safety, I think that gets a little dicey. I don't know the answer. Those who follow it get up for recompetition and those who don't get by it... I'm saying that you can't count on everyone. It's not applied equally because not everyone jumps to self-report. And those who do certainly pay a greater price than those who don't. —Designated for Competition

One of the things you find going through the DRS process is that we always had a strong partnership with the regional office and T&TA folks. We helped train other programs. We have provided training at the regional level. When you go through recompetition it drives a wedge between you and Early Head Start/Head Start funders because now you don't trust them. So we're rebuilding that. It doesn't mean you're disrespectful but you are extremely much more cautious. Last year at the regional office [they] said during the regional conference, 'You need to be telling us about any sort of incident with a child.' I was sitting at back of the room to watch body language. I saw some people from our state looking at each other. Three of the program directors were waiting for me outside the room. They said 'We're not reporting anything. Are you kidding me?' It really creates a wedge between people. —Designated for Competition

Teachers and Kids Occasionally Have Bad Days. In addition to concerns about program compliance, 5 of 15 respondents also remarked that DRS does not take into consideration that "kids have bad days" that could adversely affect the outcome of the monitoring review. A Head Start director at a grantee site not designated for competition shared that during their monitoring review, there was an incident with a child misbehaving. The child was experiencing homelessness, had been abused, and had taken the bus to school for the first time that day. The Head Start director explained that they cannot exclude troubled or difficult children for the benefit of their CLASS scores—that goes against their mission and Head Start's mission of supporting all children, including those with a history of trauma. Teachers keep these children in their classrooms, putting their CLASS scores at risk.

It's not always about having 6's and 7's on the CLASS scores, but about what the teachers do. These assessments often times lose track of the true need of the children. Not every child comes emotionally stable. —Not Designated

DRS Conditions Should Not All Count the Same. Five of 15 Head Start directors suggested that it would be fairer if not all seven conditions carried the same weight in designating grants. They suggested that perhaps some type of weighting might be used.

I think DRS should be weighted. Many of the conditions only apply to certain states – the licensing conditions apply only where licensing exists. It's like those states have a higher bar. The states with limited licensing laws, those programs aren't being checked in those areas. I think there are things more serious than others. —Not Designated

Everyone Should Compete. A few Head Start directors (6 of 35 telephone interviews, 4 designated and 2 not designated for competition) suggested that it would be a more equitable system if all Head Start programs had to compete for their grant every five years.

I don't think any program has a problem being on the DRS, but it's just the stupidity. I think any director will tell you—just tell me when I'll be on it, I'll apply for it. But let's not make us look bad in the process. We've said as directors, we have a grant number, just use that, odds and evens, to decide who goes on. Let's still have our audits. But let's not use it to find some issue to get us on the list. If we were all passing these reviews, there would still be 10 percent on the list anyway. So why not just make it random, every year 10 percent [recompete]. —Designated for Competition

Some grantees expressed concerns about transparency in the designation decision-making process.

Eight of 15 grantees perceive a lack of transparency at how decisions regarding designations are made within OHS. They stated that they felt a lack of clarity and have unanswered questions surrounding a process to appeal designation status or provide additional information on program context or to explain certain findings or circumstances. As one Head Start director expressed during their telephone interview, *"Information gets sent to someone and you don't know how it was given to the decision makers and you can't appeal the decision."* Respondents at five designated grantee sites (five of seven designated grantee site visits) were angry that their program had been designated and felt it was unfair that no appeal process was made available to them. As one Head Start director at a grantee designated for competition expressed,

One incident that can throw you into recompetition. I reported our incident to OHS. They didn't even investigate. The state even said, 'We're out of this.' So how did that end up then putting us into recompetition and there was no appeal process? When Washington asked the Regional Office about our policies and procedures, the Regional Office said they have all that and sent it to Washington. It was just really unjust. —Designated for Competition

Four Head Start directors indicated that they had attempted to appeal findings from reviewers, but had made no headway and were ultimately designated for competition. One Head Start director talked about their OHS regional office assuring them that a particular incident was not a deficiency, only for them to find out several months later that they had in fact been designated.

Some grantees expressed concerns that uncertainties in the DRS process damaged staff morale, and they worried about staff leaving.

Grantees we spoke to understood that receiving a deficiency meant they would need to compete for their grant, even if they had not received a notification letter from OHS immediately. Grantees not designated for competition and grantees designated for competition based on CLASS must sometimes wait more than a year to be notified of their designation status (see table I.1 on the DRS implementation rollout period). That is, in order to determine which grantees in a given cohort scored in the bottom 10 percent on CLASS, CLASS assessments need to have been conducted for all grantees in that cohort. Some grantees could be monitored in October or November of one year, for example, and have to wait until December of the following year to learn their status.

Head Start directors and program managers at 6 of the 15 grantees participating in site visits expressed concerns with how long they had to wait to receive notification of their designation status after their monitoring visits ended (three were designated for competition and three were not designated, and four of six had reviews in the first half of their monitoring years). Grantees that passed the review with no deficiencies still had to wait for CLASS scores to be calculated before knowing their designation status. One Head Start director said,

> We came through our review okay, but we're awaiting our final CLASS scores. We feel we'll be okay because we're above the absolute standard, but with bottom 10 percent you can't feel safe until you see the national averages come out. —Designated for Competition

Two Head Start directors (one at a site designated for competition and the other at a site not designated) discussed the difficulty of continuing program operations at the same level of quality during the waiting period. Many directors spoke of the emotional costs of the extended waiting period, highlighting their own stress levels, and their staff's concerns and fears.

> And the fact that you don't know until you get that letter, are my staff, are they going to have jobs? Are we going to survive this? It's so high-stakes. And I think I felt it the most. I tried really hard to be honest and transparent about the stakes, but I tried to carry the burden of that myself. So, at the same time as I'm saying stakes are high, I'm also saying to my staff, 'Look at your CLASS scores, you guys are doing really, really well. We are a great program, we should fly through this.' But, in the back of my head I'm thinking, 'Well, I'm in directors' meetings with people who have lost their grant.' —Not Designated

I think what was most damaging is that there is no indication of how long you have to wait for anything. The agency is put into stasis. You're trying to help your staff, who are underpaid and nervous, to feel confident, but you can't give them any information. You can't say, "We're doing this. On such and such a date, the RFA will come out and we can respond." No, you just wait and wait. That kind of lack of communication is quite dysfunctional, and it creates a sense of suspicion of the national organization. – Designated for Competition

Because of the potential loss of the Head Start grant, all 15 Head Start directors indicated that they exercised caution and sensitivity when informing their staff about the DRS process and possible consequences of competition.[54] They also discussed encouraging staff to continue performing their jobs at the same level of quality or higher. Program managers (five of seven designated grantees) discussed reinforcing what their Head Start director stated in the staff meeting and delivering a consistent message to their direct reports. As one program manager explained, *"Our job, me and the managers, was to tell staff 'we are quality, we will continue to be quality, we will be okay.' We may have lost one staff member who was threatened by the process."*

Summary

In this chapter, we conducted qualitative explorations into how grantees' experiences with the monitoring, assessment, and designation notification process may shape their efforts to improve program quality. In summary, all 15 of the Head Start grantees interviewed after they knew their designation status indicated they were continuing to work on improving quality. Grantees designated for competition were more likely to indicate some of their quality improvement efforts were related to the DRS. Some efforts were focused on CLASS, but some were focused more broadly on areas of need identified by the program. Grantees expressed a number of concerns about the early implementation of the monitoring, assessment, and notification processes, including concerns about determining quality based on a single visit or incident (20 of 35 grantees), transparency in the designation decision making process (8 of 15 grantees), and the length of time for not knowing the designation status (6 of 15 grantees).

---

[54] Although grantee interviews suggest that the Head Start director is typically tasked with sharing information about the DRS within the Head Start program, OHS indicates that all DRS notifications (and all official actions) are sent from OHS to the governing body chairperson, as the governing body has the legal and fiduciary responsibility for the Head Start funds as specified in the Head Start Act.

Chapter VI. What does competition look like?

The goal of the competitive process of the DRS is to improve the quality of Head Start services. The design of this evaluation is based on the assumption that, once a grantee has been designated for competition, the competitive process of the DRS can improve quality if the incumbent is awarded the grant after proposing to increase the quality of their services or if the incumbent is replaced by a new awardee that proposes higher quality services. This framework suggests that the extent to which competition is an effective mechanism for quality improvement depends on three major factors: (1) the extent to which incumbents perceive competition as a reason or means to propose higher quality services, (2) the extent to which other organizations enter the competition and propose higher quality services, and (3) the degree to which proposed quality improvements in the awarded grant can be implemented.

In this chapter, we examine these issues by discussing the level of competition and characteristics of the applicants for the Funding Opportunity Announcements (FOAs) associated with DRS Cohort 3 and released in 2014. We also describe factors that may influence the level of competition, including the characteristics of the incumbent or of the service area, and barriers and incentives perceived by applicants. We describe both the incumbent and new awardees' perceptions of whether there was competition. Finally, we document and discuss the quality improvements that both incumbents and new awardees proposed as a result of this process. Thus, we explore the research question, "What does competition look like?"

Overview of Methods

The evaluation team examined competition using various sources of information. Our quantitative analyses of the extent of competition and characteristics associated with competition began with administrative data provided by the Office of Head Start identifying the organizations that applied for the 2014 FOAs. These data were then matched with other sources of data including (1) PIR data to obtain information about the characteristics of existing grantees (including some new applicants who

were already providing Head Start services in another area),[55] (2) data from the web-based Applicant Survey we administered to obtain information about the characteristics of new applicants (see below for more information about the Applicant Survey), (3) supplemental information collected from publically available sources about the characteristics of other new applicants who did not respond to the survey or were not already in the PIR, and (4) Census data to describe the demographic characteristics of the regions for each FOA. These combined sources of data provided information about the characteristics of 177 of the 182 applicants (96 incumbents and 81 new applicants).

To obtain information about the services and program structure proposed by applicants, we asked incumbents and new applicants to respond to a web-based survey (Applicant Survey), and we used the PIR to obtain information about the prior services and program structure of prior grantees that were designated for competition. Overall, 67 percent of the 182 total applicants responded to the survey. Note that some of our quantitative analyses are performed at the FOA level (e.g., both descriptive information about the degree of competition across FOAs and probit regressions predicting the likelihood of competition), while other analyses are performed at the applicant level (e.g., paired t-tests comparing the characteristics of incumbents with new applicants). Our third analysis used qualitative methods to analyze information from semi-structured interviews of nine grantees after they were awarded grants through this competition. These interviews explored the reasons they chose to compete, the challenges and opportunities of the competitive process, and their experiences with the award and negotiation processes; in this chapter we report on their perceptions and experiences related to the competitive process. See chapter II and technical volume II for more information regarding the sample, methods, measures, and analytic process for the qualitative interviews discussed in this chapter.

Findings

The findings include the following:

1. ***Just over half of the FOAs (58 percent) had only one applicant, and in most cases, that applicant was the incumbent, as almost all incumbents (91 percent) re-applied for their grant.***

---

[55] The PIR data contain information about characteristics of all organizations providing Head Start services. Those organizations can be incumbent applicants for a given FOA who were designated for competition, but they can also be new applicants who are already providing Head Start services in another service area.

2. *We found that applicants, including incumbent applicants, proposed changes to service delivery, including more delegates, higher shares of teachers with BAs, lower child-teacher ratios, and lower enrollment than the current grantee.*

3. *In open-ended interviews with awardees, incumbent awardees identified the threat of competition and the fear of being designated for competition again as factors that spurred them to make changes, but the new awardees were generally motivated by the idea that their proposed changes would improve upon the current service provision.*

Finding 1a: Just over half of the FOAs (58 percent) had only one applicant, and in most cases, that applicant was the incumbent as almost all incumbents (91 percent) re-applied for their grant.

What does competition look like, and how do applicants experience it?

The administrative data from the Office of Head Start reported there were 103 FOAs that posted requests for applications for the Head Start grants opened for competition in 2014.[56] The FOAs covered both rural and urban areas in 39 states and Puerto Rico. Most of the FOAs (95 of the 103) allowed for flexibility in proposing Head Start and/or Early Head Start Services (labeled Birth to Five). Only seven FOAs called for Head Start services only, and only one called for Early Head Start services only. Findings are presented below.

---

[56] Although 105 grantees had been designated for competition, the service areas of four designated grantees were combined into two FOAs, resulting in 103 FOAs.

FIGURE VI.1.

Distribution of FOAs by Number of Applicants
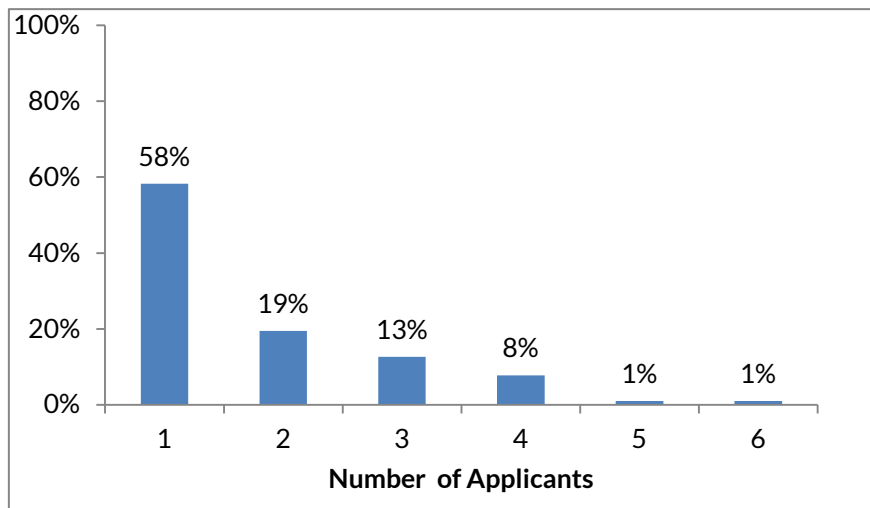
Source: Office of Head Start administrative data on all applicants (*N*=103 FOAs).

There were 182 applications for the 103 FOAs. All FOA competitions had at least one applicant organization that responded to the FOA, but less than half (42 percent) had two or more applicants (figure VI.1). In all but nine FOAs, the incumbent grantee (the organization that held the grant at the time it was designated for competition) applied to win back the grant.[57]

Who were the new applicants, and how did they compare with the incumbents?

We compared the characteristics of the incumbents and new applicants (organizations that applied to the FOA but were not the incumbent) using a *t*-test for difference in means (table VI.1).[58] The new applicants differed from the existing incumbents in some ways. Almost half (46 percent) of all new applicants were new to Head Start, whereas the other 54 percent were currently involved in Head Start. Of the 54 percent who were currently involved in Head Start, 68 percent were grantees in other areas, 27 percent were both grantees and delegates, and 5 percent were only delegates. New applicants were less likely to be Community Action Agencies and more likely to be other kinds of nonprofit or for profit organizations. Although the majority of new applicants were currently serving children, 10 percent were not.

---

[57] Both incumbents applied in the two cases in which the service areas were combined.
[58] Information about characteristics of incumbent applicants came from the PIR, and information about new applicants came from our Applicant Survey, from the PIR (for new applicants who were in the PIR, because they were already providing Head Start in another service area), and from publically available data (for new applicants who either did not respond to our Applicant Survey or were not already in the PIR).

TABLE VI.1.

2014 FOA Applicant Characteristics, by Incumbent Status

	Total (*N*=177)	*t*-test	Incumbent Applicants (*N*=96)	New Applicants (*N*=81)
Organization Auspice				
Nonprofit	79%		81%	75%
Public	17%		19%	15%
For Profit Organization	5%	**	0%	10%
Organization Type				
Community Action Agency	33%	**	48%	15%
Private/Public Nonprofit (Non-CAA)	45%	**	30%	62%
Government Agency (Non-CAA)	7%		8%	6%
School System	10%	*	14%	6%
For Profit Organization	5%	**	0%	10%
Ages Currently Serving				
Not Currently Serving Children	5%	**	0%	10%
Currently Serving Children	95%	**	100%	90%
Serve Children Ages 2 and Under[a]	81%	*	74%	88%
Serve Children Ages 3 to 5 Years[a]	97%	**	100%	93%
Head Start History				
Not In Current Head Start Grantee	21%	**	0%	46%
Current Grantee or Delegate[a]	79%	**	100%	54%
Current Head Start Delegate[a]	10%	**	3%	17%
Current Head Start Grantee[a]	78%	**	100%	52%
Head Start Type				
Current Head Start Grantee	78%	**	100%	52%
Head Start[a]	72%	**	98%	41%
Early Head Start[a]	44%	+	50%	37%
Migrant/Seasonal Head Start[a]	1%		1%	1%

Sources: Office of Head Start administrative data on applicants, PIR data on applicants (for incumbent applicants and for new applicants who were already providing Head Start services in another area), information collected from the Applicant Survey and publically available information on applicant organizations (for new applicants not in the PIR).

Notes: Results of significance tests indicate whether the characteristic is significantly different between the incumbent and the new applicants. Five applicants for whom we could not locate any information are excluded from the analysis. Significance noted as: + $p<.10$; * $p<.05$; ** $p<.01$.

[a] Subcategories not mutually exclusive.

Finding 1b: Areas with higher ACF-funded enrollment and areas located in the South were more likely to have competition. Competition was less likely if the incumbent was a school system.

What characteristics of the incumbent grantee and of the service area were associated with competition?

Next, we compare the communities where competition did and did not occur when FOAs for designated grants were released. Using PIR and census data related to all FOAs, we estimated a probit regression to identify the characteristics of the prior grantee that was designated for competition and characteristics of the location of the FOA that were associated with competition. Table VI.2 shows the results of those analyses. The characteristics included in the probit regression accounted for about 15 percent of the variability in competition and suggested that competition was more likely when the prior grantee enrolled more students and was not a school system and when the location was in the South.

TABLE VI.2.

Probit Regression Model Estimating the Likelihood of Competition (DRS Cohort 3, _N_=103)

	Mean		Prediction Model
	FOAs without Competition	**FOAs with Competition**	**Marginal Effect[a]**
Characteristic of the Incumbent Grantee[b]			
ACF Enrollment	421 (386)	949 (1,220)	0.015+[c] (0.010)
Incumbent is School System	20%	7%	-0.253+ (0.136)
Incumbent Designated for Competition Due to CLASS	56%	44%	-0.122 (0.091)
Characteristic of the Location			
Population	245,895 (460,895)	528,639 (599,697)	0.013[d] (0.010)
Northeast	22%	14%	Reference Category
Midwest	22%	16%	0.116 (0.134)
South	40%	56%	0.242* (0.115)
West	15%	12%	0.149 (0.156)
Puerto Rico	2%	2%	0.311 (0.311)

Sources: Office of Head Start data on applicants, PIR data on incumbents and Census Bureau data on regions. These data are available for all 103 FOAs.

Notes: Standard deviations in parentheses for means and standard errors in parentheses for marginal effects. Significance noted as: + p<.10; * p<.05; ** p<.01.

[a] Marginal effect is the change in the likelihood of competition associated with a one unit change in the independent variable.

[b] Head Start combined the service areas of four grantees designated for competition, resulting in two FOAs, each of which included two prior grantees. Because the regression estimates the association between the characteristics of the prior grantee with the likelihood of competition, we need to combine the characteristics of the two prior grantees to reflect the characteristics of a single representative prior grantee in the two FOAs with two prior grantees. We defined total enrollment as the sum of the enrollment for the two prior grantees. When the two prior grantees within a FOA had the same characteristics, it is clear how to combine those characteristics. This was true for characteristics of the location (by definition), and within each of the two FOAs the two prior grantees also were designated for competition for the same reason. In one of the FOAs, both prior grantees were the same auspice, but in the other FOA, one of the prior grantees was a school system, but the other was not; for this FOA we defined the prior grantee as being a school system. The results are not sensitive to this classification decision.

[c] Marginal effect coefficient is interpreted as follows: for every increase in enrollment of 100 children, the likelihood of competition will increase by 1.5 percentage points.

[d] Marginal effect coefficient is interpreted as follows: for every increase in population of 100,000 people, the likelihood of competition will increase by 1.3 percentage points.

How did incumbents and new applicants decide to apply? What are the potential barriers to application?

In this section, we present the results of interviews conducted on site visits with nine awardees including four incumbent awardees and five new awardees. Two of the new awardees were current Head Start grantees in another location and two of the new awardees were not competing against an incumbent. During these interviews, awardees were asked about their decision to apply and the application process. The four incumbent awardees we interviewed all said their decision to apply was automatic and that they started planning for the application well in advance of the FOA posting.

> [T]here was no doubt in our minds that we would reapply. That was an unfortunate situation, but we're a good program and know it. There wasn't even a second we weren't sure. –Incumbent Awardee

In contrast, our interviews with new awardees revealed many factors that they considered when applying and all five spent time deciding whether they would apply. Factors mentioned in the decision to apply were adhering to federal reporting standards, the well-being of the children in the community, potential reaction of the community, and outside competitors applying.

All of the five new awardees had a more formal decision making process for the application than did the incumbents. None of them said the decision to apply was automatic. The new awardees had internal and external discussions on the feasibility of applying and providing high quality services. The four new awardees, who anticipated that an FOA would be released, spent time before the FOA release internally discussing what the application should look like. Two of the three new awardees that were not current Head Start grantees were concerned about the additional burden of federal reporting. Both of the new awardees where the incumbent did not apply said that concern about competitors outside the community taking over services and concern for the well-being of the children motivated their application. One of the new awardees worried about the reaction of the community because they were not well-known in the community. Four of the five new awardees saw it as an opportunity to expand into a new area.

> [The incumbent organization] actually called me in for a visit one day, and … told me that they weren't interested in continuing as the grantee for Head Start, and would we be interested in applying for the grant… I immediately said no, because we have a lot of programs here, and federal reporting was not something we were doing …I was not really interested, to be perfectly honest…and I left the meeting saying, 'No,' and then thought for a long time and did some research on what would happen if we didn't apply. I learned at that point that there were folks in the US that were applying for Head Start grants not even affiliated with the county… and that scared me. I think Head Start should be deeply imbedded in the community…. After about a week, I went to the board and told them that [applying] was the right thing to do.—New Awardee

We analyzed the data from our interviews to understand the challenges and opportunities posed by the application process for the grants opened through competition. First, we asked incumbent awardees to compare preparing this proposal with preparing for previous noncompeted continuation proposals. Second, we asked about the application process to see if incumbent applicants used this proposal to implement quality improvement, and if so, why.[59] All incumbent awardees stated that releasing the FOAs about a month and a half after they were forecasted to have been released created challenges. They felt they could not plan their time well and were distracted by not knowing when it would be released.

> *It was a long time until the RFP came out it was a lot of just waiting…Our timeframe was small to write a whole grant…[it was] a lot of research, which we'd done some ahead of time, but it's hard until you actually see the RFP and are putting things on paper.— Incumbent Awardee*

> *I think the biggest challenge was not knowing when the grant would be released. We had heard initially that there was word it would be released in March 2014; nothing came out. We waited, waited, waited. Then we thought it would be about July or August, and it didn't come out. … It wasn't until the end of September that the grant was actually released. —New Awardee*

Eight of the nine awardees we interviewed felt that the application turnaround time was shorter than expected and they were not given enough time to prepare their proposals.

> *One of the things that could help foster competition and participation among non-Head Start existing providers is giving them more guidance and time around that. It's a pretty accelerated timeline to build some pretty strong relationships with [community partners] who may not have heard from anybody outside of whoever's providing the services in 25 years. It's very much a set of conversations that are deep and robust and that need to happen in artificially constrained timelines in a lot of these situations. —New Awardee*

Most awardees felt that this proposal was more complicated than the typical noncompeted renewal proposal (four of six who had experience with the typical noncompeted renewal process).

> *… it was a more complex grant. We had prepared for that. We'd been looking at previous years' applications to prepare. We knew what the content would be. —Incumbent Awardee*

All the awardees indicated that applying for the grants opened through competition due to designation required a substantial investment of staff time and resources. They reported that the grant proposal required developing plans and collecting data about service deliveries and developing complex budgets. This process required considerable staff time, even in the seven of the nine grantees that hired consultants and/or professional grant writers to assist them. Estimates of the required resources ranged from $17,000 to $50,000 for outside consulting and grant writing services. These additional

---

[59] The data relating to this second question are discussed at the end of this chapter.

demands on staff time reduced the staff's ability to focus on the typical tasks, according to three of the incumbents and two of the new awardees.

> *Everything else was put on hold. All your energy was into this process, and we were supposed to be running a program where every child's experience mattered and you couldn't focus on that, because your energies were focused on the grant. —Incumbent Awardee*

> *We worked seven days a week… long hours, and were here sometimes until 10 or 10:30 at night. My boss said, 'This is your life for next two and a half months.' He said, 'Whatever you need to do, clear your plate.' It was hard for me and my department … It was manpower, and beyond that it was a fiscal commitment … —New Awardee*

Overall, the grant application process required both incumbents and new applicants to invest substantial time and resources, and they felt this was exacerbated by what they perceived as too short an application period, following uncertainty about when the FOAs were to be released.

How aware were awardees about who their competitors were?

Regardless of the actual level of competition, if incumbents and new applicants thought there would be competition, it could provide an incentive to propose changes to existing services. We asked the nine grantees that had recently been awarded a grant (four incumbent awardees and five new awardees) about what their expectations had been before participating in the competition. All these awardees stated that at the time they decided to apply for the grant, they believed that there would be some competition. Six of the nine awardees we interviewed had at least one competitor. But even those three awardees that never faced any competition said that they believed there would be competition.

While at least one person at each site believed there would be competition, there was sometimes disagreement within organizations about whether they would actually have competitors. Staff at three of the four incumbent sites said they believed that they would have competition because the Head Start grant was a source of funding in a funding landscape that they viewed as increasingly sparse.

> *You have 10 million dollars out there when no one has money. You're crazy if you think no one will apply. —Incumbent Awardee*

In contrast, staff at two of the same incumbent sites and one additional incumbent site said they had doubts about competitors' ability and desire to start a Head Start program in a new service area given that the grant would not pay for start-up costs.

> *We were surprised that anyone else would apply, given the criteria that they weren't going to pay for startup. Just to have early childhood spaces was a huge start-up, and that would be a huge deterrent.*

How would an applicant come up with the places and spaces to run this? In the end, I realized no one had it when they were awarded. —Incumbent Awardee

They also speculated that competitors outside of Head Start would not be willing to adhere to all of the Head Start guidelines and rules.

You don't just run Head Start. Part of my reason for thinking there were few [competitors] is that you just don't run a government contract of this magnitude without experience and capacity. —Incumbent Awardee

All six of the awardees in our sample with competitors eventually found out that they had competition, and in many cases, they learned exactly who their competitors were during the application process. Grantees had diverse experiences finding out about their competition. In one site, an incumbent awardee discovered one of their partners was applying during the application process because the partner called to ask for information from the incumbent for their application. In another site, the competitors not only knew about each other but wrote each other letters of support, saying they would support whoever won.

Finding 2: Both incumbent and new applicants proposed changes to service delivery, including more delegates, higher shares of teachers with BAs, lower child-teacher ratios, and lower enrollment than the current grantee.

We examined the extent to which quality improvement was likely to result from the competition for Head Start grants by comparing the services proposed by the applicants for a given FOA with the prior services provided by the grantee that was designated for competition (referred to as the prior grantee). This analysis used two sources of data: PIR data describing the prior services offered by current grantees designated for competition, and the Applicant Survey of the incumbent and new applicants describing the proposed services for the grants opened through that competition. We conducted analyses using paired t-tests and clustered standard errors to account for lack of independence.[60,61] We

---

[60] When there was more than one applicant, we compared the proposed services of each applicant with the prior services provided by the grantee that was designated for competition, and we clustered by FOA.

[61] Head Start decided to combine the service areas of four grantees that were designated for competition, resulting in two FOAs, each of which had two prior grantees. For the paired t-test of applicants with the prior grantee, we created a hybrid prior grantee where the total enrollment is the sum of the enrollment for the two prior grantees, and values for the other variables in this table—share of teachers with a BA and average child-teacher ratio—are averaged across the two prior grantees. For defining the use of delegates, in one FOA both prior grantees did not use delegates, but in the other FOA one prior grantee had delegates and the other did not, and we defined the delegate status for that FOA as having delegates.

first compared the proposed services of all applicants that responded to the Applicant Survey to the service provision in 2013 of the corresponding prior grantee designated for competition. Then, we assessed how these differences vary separately for areas with and without competition. We examined the use of delegates and enrollment as measures of program structure and operation, and teacher qualifications and child-teacher ratios as measures of the quality of the classroom-level services.

Comparisons of the 2013 services to services proposed by all applicants indicated that applicants proposed changes in program structure and classroom services. Applicants were more likely to propose having delegates than the prior grantee, with 35 percent of applicants proposing to use delegates for all of their service provision compared with only 2 percent of prior grantees (table VI.3). Applicants also proposed increased proportions of teachers with BAs. Applicants, on average, proposed a teaching staff where 77 percent of their teachers had BAs compared with the prior grantee, where only 55 percent of their teachers have BAs. Applicants also proposed lower child-teacher ratios and lower enrollment, but those differences were observed in the Head Start programs and not the Early Head Start programs. These results were generally similar for both incumbents and new applicants.

TABLE VI.3.

Comparison of Program and Classroom Characteristics Proposed by Incumbent and New Applicants with the Characteristics of the Existing Grantee Before DRS Competition

Characteristic	All				Incumbents				New Applicants			
	Existing Grantee[b]	All Applicants	N	t-test	Existing Grantee[b]	Incumbent Applicants	N	t-test	Existing Grantee[b]	New Applicants	t-test	N
Delegates												
No Delegates	83%	54%	106	**	88%	67%	57	**	78%	39%	**	49
Some Delegates	15%	11%	106		11%	9%	57		20%	14%		49
All Delegates	2%	35%	106	**	2%	25%	57	**	2%	47%	**	49
Share of Teachers with a BA	0.55 (0.29)	0.77 (0.30)	105	**	0.57 (0.30)	0.79 (0.27)	58	**	0.53 (0.27)	0.75 (0.33)	**	47
Average Child-Teacher Ratio[a]												
Head Start	17.9 (4.4)	15.3 (6.6)	94	**	18.3 (4.1)	16.8 (6.4)	56		17.3 (4.7)	13.1 (6.5)	**	38
Early Head Start	7.4 (6.4)	7.1 (6.4)	43		8.8 (8.0)	7.6 (8.0)	24		5.7 (2.8)	6.6 (3.6)		19
Total Enrollment[a]												
Head Start	765.1 (1057.6)	518.6 (850.5)	103	*	635.6 (891.2)	533.4 (845.9)	60	*	945.8 (1242.3)	498.0 (866.6)		43
Early Head Start	152.4 (119.4)	153.2 (157.4)	55		140.4 (107.6)	145.7 (129.2)	30		166.8 (133.1)	162.3 (188.2)		25

Sources: Data on the applicants come from the evaluation's Applicant Survey; data on the current grantee comes from 2013 OHS PIR data.

Notes: Significance noted as: + $p<.10$; * $p<.05$; ** $p<.01$. Results of significance tests indicate whether there is a difference between the applicant and the existing grantee based on a paired t-test. Existing grantees with no corresponding information from the survey on applicants who applied to the FOA were excluded from the analysis. Sample sizes differ because of nonresponse to the survey question and different shares providing service for Head Start and Early Head Start.

[a] We report average child-teacher ratio and total enrollment separately for Head Start and Early Head Start programs because these program characteristics are generally very different for Head Start compared with Early Head Start programs.

[b] Office of Head Start combined the service areas of four grantees that were designated for competition, resulting in two FOAs, each of which had two prior grantees. For the paired t-test of applicants with the prior grantee, we created a hybrid prior grantee, where the total enrollment is the sum of the enrollment for the two prior grantees, and values for the other variables in this table—share of teachers with a BA and average child-teacher ratio—are averaged across the two prior grantees. For defining the use of delegates, in one FOA both prior grantees did not use delegates, but in the other FOA, one prior grantee had delegates and the other did not, and we defined the delegate status for that FOA as having delegates.

We also tested whether applicants proposed larger changes in services when there was competition compared with applicants facing no competition. We found no significant differences in the extent of changes proposed between areas with no competition and those with some competition. However, the number of cases examined here is small, which confounds interpretation, and thus we do not present the analysis in the report.

Finding 3: All four incumbent awardees stated that the FOA, the threat of competition, and the fear of being designated again spurred them to make changes. Four of the five new awardees believed their proposed changes would improve upon the current service provision.

Next, we explored the types of changes proposed by grantees during competition and their motivations for those changes. The findings discussed below are based on the qualitative interviews conducted with nine recent awardees of Head Start grants, including four incumbent awardees and five new awardees. During these interviews, awardees were asked about any changes they proposed in their application to the existing service delivery and what motivated those changes.

All but one of the grantees we spoke with proposed making changes to existing services in the community. The four incumbent awardees all stated that the competitive FOA allowed more flexibility in changing their program and that the threat of competition spurred them to reflect on how to improve their program.

> It was a stressful process, but one of the benefits is that you are forced to look at your agency and your program. Prior to this process, you know what you need to do but in this process you're forced to take a hard look in the mirror and fix it and address it. It allowed us to reset and put systems in place, across the whole program.—Incumbent Awardee

> [The most beneficial aspect of DRS is] making us look at our program, really examine what we want to do, how we can do it, what are our flaws, we have to really look at it. We can't just put the grant in exactly the same way of doing business. That's good.—Incumbent Awardee

> I'm glad we rewrote [to add] Early Head Start. That was a blessing, a good thing to get out of it. I'm very much a glass-half-full kind of person, so I do look at those things… the opportunity to get Early Head Start and redesign our services was a blessing ….—Incumbent Awardee

All four incumbent awardees looked to needs assessments of the community and the reasons they were designated for competition to determine what changes to make. Three of the four incumbent

awardees talked about moving into or expanding their Early Head Start services, both because they saw a need in the community and they believed that 4-year olds were being served already by the expanded state or local prekindergarten efforts.

> *We converted some slots to full-day, converted some to Early Head Start, just looked at locations and one area we were having more problems with enrollment, so we elected to close some and moved those slots somewhere else. —Incumbent Awardee*

> *I think we'd heard throughout the network that Early Head Start was the place to go, and that's where we had to dip our toes and explore. That was the biggest focus; how can we do this successfully? —Incumbent Awardee*

Others spoke of increasing the share of teachers with a BA to improve teacher quality.

> *We increased the caliber of internal programs and credentials. We no longer hire teachers who don't have bachelor's degrees. We raised salary level…[to be] competitive with the district. —Incumbent Awardee*

All four incumbent awardees proposed changes related to their reason for designation. The two sites designated for competition for CLASS scores proposed trainings related to CLASS score improvement. One of the sites that was designated for competition based on the actions of their delegates spoke of using the DRS as an opportunity to change the delegate structure to a partnership structure where they would have greater control over compliance and fiscal issues. Another site, seeing a greater need for oversight and family advocacy, made substantial organizational changes to the structure of the classroom.

Four of the five new awardees proposed changes to the current grantees' service provision. Many of the changes they proposed were similar to those that the incumbents proposed. One site proposed expanding Early Head Start. Another new awardee, in a location where the original grantee was designated for competition for the actions of their delegate, changed the delegate structure to a partnership structure to ensure more control over the delegate agencies. One of the new awardees, however, proposed no changes to the service delivery that the former grantee had offered.

> *We proposed a specialized classroom for kids with challenging behaviors. We knew that from being a consultant that there are lots of kids with Head Start. We wanted to design and implement a unique component. We wanted to add a home-based Early Head Start component that didn't exist locally. We plotted on a county map where quality preschool services were available through Head Start and other providers and saw lots of holes so we added the home-based Early Head Start. We knew the family support side needed significant works. So we knew we needed to strengthen leadership in that department. And we knew we wanted to do more full-day. We proposed an increase in full-day classrooms because we knew that's what parents want. Half-day is not convenient.—New Awardee*

> *I got rid of my part-days. They only stayed opened 'til 1:30, and that doesn't work for parents. We do 8:30 a.m. to 4:00 p.m. And we do wraparound so that if a parent needs to drop them off early or if they*

need to get back late. We make sure to provide the service on both ends. That's something we'd like to expand —our wraparound services.—New Awardee

Both new and incumbent awardees established new partnering relationships, and new and incumbent awardees also formed partnerships with organizations that had existing relationships with the current grantee. Other than the time period for forming relationships, most interviewees reported no problems with forming the desired partnerships.

Two of the four incumbent awardees reported forming new partnerships as a result of the competition process, with one using competition as a chance to develop partnerships they had previously considered, and the other reaching out to new partners to help increase their enrollment. One incumbent awardee explored new potential partnerships, but ultimately decided not to partner with any new organizations.

We had months of conversations with [two potential partners] before it came to be that it wasn't going to work in either. Head Start requirements and those programs were too disparate. We talked for months to see if we could get around, see if we can do a locally designed option for Head Start, it just, neither partner felt like it would work. —Incumbent Awardee

Four of five new awardees partnered with organizations that had worked with the previous grantee. However, four out of five also explored new partnerships, reaching out during grant writing and after the grant was awarded. One new awardee reported no change in partnerships from the former grantee.

We met with the mayor, common council members, clergy, child care providers who'd worked with the previous grantee to provide services for children with disabilities, mental health, people who provided resources to the former grantee. We immediately reached out to them. We met no resistance.—New Awardee

While three new awardees sought out partnerships by asking for letters of support, one new awardee had a formalized application process for potential partners, who were asked to submit information on proposed services and finances.

We put information out to anyone who wanted to play. It has to be their decision too; it meant their taking on staff and responsibilities. From ones that were willing to play, they had to submit their application, financials, budgets, audits, and we put them through an interview if they had any audit findings or litigation. We had a pretty extensive application.—New Awardee

Two new awardees took advantage of their existing partnerships and infrastructure in recruiting new partners. Additionally, one of these awardees utilized staff with connections in the community to initiate relationships.

Interviewer: Did you know who had been working with the former grantee?

Respondent: They came forward. Vendors came forward. Security contracts, telephones, food, contracts for computer, technology. Anyone who did any work came forward quickly to engage with us. Some we retained, and some we were already working with and we folded them into the existing program.—New Awardee

We had no relationships with organizations in [service area], other than my employees that live over there. They helped us get in the door because they knew folks and lived in the community.—New Awardee

Finally, one new awardee believed that new awardees and incumbent organizations face different concerns and considerations in seeking out new partnerships.

When you're coming in as a challenger, especially as a group that hasn't been a Head Start grantee before, there's a lot of awareness-building that needs to happen both within their organization and with the community about their organization. Even though they have facilities on the ground, they don't necessarily tie in with the same comprehensive service providers. They don't necessarily have the same broad base of relationships that a well-oiled Head Start machine would have. Building those relationships and bridges, helping the team decide who to target, getting to that network of local support in advance of the grant submission and also throughout the process of implementing is very important for new providers in a way that is not—if you're an existing grantee, even if you're in a new geographical area you know all that stuff already.—New Awardee

Summary

The goal of the competitive process of the DRS is to improve the quality of Head Start services. Our analyses show that there was no competition in 58 percent of the sites. However, when we explored through qualitative interviews whether awardees of grants were proposing quality improvements to services, both incumbent and new awardees proposed changes to service provision. All nine awardees we spoke with believed there would be competition, although only six actually faced competition. In addition, awardees viewed the DRS as providing the opportunity to propose changes to services that would not have been possible outside this process. In the next chapter, we explore how the negotiation, award, and transition process might facilitate or pose challenges to implementing the changes proposed by awardees.

Chapter VII. How do grantees experience the negotiation, award, and transition processes in terms of preparing them to implement quality improvements to existing services?

Following submission of the proposals to OHS in response to the DRS FOAs, the applicants wait to hear which organizations won the competitions to become the Head Start grantees, how much their awards are, when they can begin delivering services, and how to make the transition to service delivery. Upon receiving an award, the applicant must negotiate and transition service delivery. New applicants must establish all operations and services, and incumbent applicants almost always have to negotiate changes to at least some services as proposed in their applications. Little is known about how this process—the negotiation, award, and transition of services processes—may support implementation of the quality improvements proposed in Head Start grantee applications, or present implementation challenges. This chapter describes the outcomes of the 2014 DRS competitions and explores these issues through the research question, "*How do grantees experience the negotiation, award, and transition processes in terms of preparing them to implement quality improvements to existing services?*"

Overview of Methods

This portion of the study was supported by a mixed-methods approach. This chapter begins with a quantitative analysis of the award outcomes from the 2014 competitions. The Office of Head Start provided administrative data identifying the organizations that applied for the FOAs and those that won the award. To obtain information about the characteristics of the prior grantee, applicants and awardees, these data were then matched with other sources of data including (1) PIR data to obtain information about the characteristics of existing Head Start grantees (including some new applicants

who were already providing Head Start services in another area),[62] and (2) data from a web-based survey we administered to obtain information about the characteristics of new applicants (Applicant Survey). These combined sources of data provided information about the 105 awardees (80 incumbents and 25 new awardees). We provide descriptive statistics about the characteristics of the grantees who received awards, compare the services proposed by awardees and the prior grantee using paired *t*-tests and examine associations between proposed operations and services and the probability of obtaining an award using a probit regression.

In addition, we used qualitative methods to explore the negotiation and award process and how that affected the implementation of proposed changes to services that resulted from the DRS process. Nine awardees participated in semi-structured on-site interviews (four incumbent awardees and five new awardees, including agency and program directors, program managers, and members of the governing body or Policy Council) about two to three months after they had received their final award notification; all were in the process of starting up their newly funded services. Qualitative interviews were coded to explore the reasons they chose to compete, the challenges and opportunities of the competitive process, and their experiences with the negotiation, award, and transition processes. We coded by status of awardee—whether they were the incumbent re-awarded the grant or a new awardee. We report themes at the awardee level, but point out where within-awardee differences or other variations of note occurred. Throughout this chapter, we provide information about the proportion of grantees we spoke to that held similar perceptions of their experience with DRS negotiation, award, and transition processes, but this does not suggest the same would be found within the general population of Head Start grantees. In other words, if half of grantees responded similarly in this study, it may be that less than half or more than half of all Head Start grantees that experienced the DRS would have responded similarly. We provide exemplar quotes to illustrate the themes.

See more information in chapter II and technical volume II, appendix D regarding the sample, methods, measures, and analytic process for the qualitative interviews discussed in this chapter.

Findings

The three major findings discussed in this chapter include the following:

---

[62] The PIR data contain information about characteristics of all organizations that are providing Head Start services. Those organizations can be incumbent applicants for a given FOA who were designated for competition, but they can also be new applicants who are already providing Head Start Services in another service area.

1. *The majority of awards were made to incumbents. Most awards to new applicants went to Head Start grantees from other areas.*

2. *The Head Start grant awardees proposed more delegates, increased proportions of teachers with BAs, lower child-teacher ratios and lower enrollment than the existing grantees.*

3. *The timing and confidentiality of the negotiation process created challenges in preparing for implementation of the proposed changes to services for most of the interviewed awardees.*

Finding 1: The majority of awards were made to incumbents; most awards to new applicants went to Head Start grantees from other areas.

The 2014 grant competition resulted in 105 awards from 103 FOAs. Some FOAs resulted in no award when OHS found that none of the applicants proposed services of sufficient quality for funding,[63] while other FOAs split the awards, resulting in multiple awards in a given service area. Of the eight FOAs that were not awarded to an applicant, six were reposted with on-going services provided by the incumbent or Community Development Institute Head Start (CDI),[64] one had not been decided at the time of analysis, and one was awarded to an organization that did not apply to that FOA but applied for and was awarded a grant for a nearby service area. Of the nine FOAs that resulted in multiple awards, eight FOAs ended up with two awards per FOA, and one FOA competition was split into three awards, producing 19 awards in these FOAs.

The majority of awards were made to incumbent grantees and most awards to new applicants went to Head Start grantees or delegate agencies from other service areas. According to administrative data from the Office of Head Start matched with PIR data, 76 percent of awards went to incumbents, 17 percent to other Head Start grantees or delegates,[65] and only 6 percent to a new organization. These evaluation findings are similar to those noted by a National Head Start Association analysis of the first three DRS cohorts (Dropkin 2015). When there was no competition, 91 percent of the awards went to the incumbent grantee, 9 percent were awarded to a Head Start grantee from another service area, and none were awarded to organizations not currently providing Head Start. In contrast, for FOAs that involved competition, 59 percent were awarded to the incumbent grantee, 27 percent were awarded to

---

[63] All the FOAs had one or more applicants.

[64] CDI is the agency OHS uses to temporarily operate Head Start programs when the grantee has been deemed unfit to run the program or has relinquished the grant and another permanent grantee has not yet been found.

[65] Two of the awardees were Head Start delegate agencies that were not also Head Start grantees in other service areas.

other Head Start grantees or delegates, and 14 percent were awarded to organizations not currently providing Head Start.

Finding 2: The Head Start grant awardees proposed more delegates, higher proportions of teachers with BAs, lower child-teacher ratios and lower enrollment than the existing grantees.

We compared the program features and services proposed by grantees that were awarded the new Head Start grant with the program features and services provided by the prior grantees (i.e., the grantee in their service area that had the grant in 2013.) For this comparison, we used a paired *t*-test and clustered standard errors to account for lack of independence.[66,67] We found awardees proposed to use more delegates, higher shares of teachers with BAs, lower Head Start child-teacher ratios, and lower Head Start enrollment than the prior grantees. No differences were found for child-teacher ratios in Early Head Start classrooms or for total enrollment in Early Head Start services. Results are similar for FOAs with only one applicant. For FOAs with two or more applicants, however, there were no differences in Head Start teacher-child ratios or Head Start enrollment between the awardees and prior grantees.[68] Thus in both areas with and without competition, there were changes in proposed service provision. Table VII.1 shows these comparisons by the number of applicants who competed.

[66] When there was more than one awardee from an FOA, we compared the proposed services of each awardee with the prior services provided by the grantee that was designated for competition, and we clustered by FOA.
[67] The Office of Head Start combined the service areas of four grantees that were designated for competition, resulting in two FOAs, each of which had two prior grantees. For the paired *t*-test of awardees with the prior grantee, we created a hybrid prior grantee where the total enrollment is the sum of the enrollment for the two prior grantees, and values for the other variables in table VII.1—share of teachers with a BA and average child-teacher ratio—are averaged across the two prior grantees.
[68] The sample involving comparisons of new and prior grantees in communities with two or more applicants is small and may lack power to detect differences between groups.

TABLE VII.1.

Service Provision of Awardees versus Prior Grantees of the 2014 DRS Competition Awardees

	All				1 Applicant				2+ Applicants			
	Prior Grantee[a]	Awardee	N	t-test	Prior Grantee[a]	Awardee	N	t-test	Prior Grantee[a]	Awardee	N	t- test
Delegates												
No Delegates	85%	63%	67	**	94%	63%	35	**	75%	63%	32	**
Some Delegates	12%	10%	67		3%	9%	35		22%	13%	32	
All Delegates	03%	27%	67	**	3%	29%	35	**	3%	25%	32	**
Proportion Teachers with BA	0.57 (0.29)	0.78 (0.28)	66	**	0.60 (0.31)	0.80 (0.30)	37	**	0.53 (0.28)	0.75 (0.27)	29	*
Average Child-Teacher Ratio[b]												
Head Start	18.6 (4.5)	16.4 (6.4)	62	*	18.0 (3.6)	16.5 (4.8)	36	+	19.4 (5.5)	16.2 (8.2)	26	
Early Head Start	8.3 (7.5)	7.2 (7.4)	25		9.8 (9.8)	8.0 (9.1)	15		6.7 (2.7)	6.2 (5.1)	13	
Total Enrollment[b]												
Head Start	678.7 (875.2)	512.5 (804.5)	67	**	424.1 (401.4)	364.6 (368.4)	37	**	992.6 (1165.7)	694.9 (1114.6)	30	
Early Head Start	152.7 (119.8)	158.1 (150.1)	35		106.9 (72.0)	108.9 (74.4)	18		201.3 (141.9)	210.1 (190.9)	17	

Sources: Data on the awardee come from our Applicant Survey; data on the prior grantee come from 2013 PIR data.

Notes: Results of significance tests indicate whether there is a difference between the awardee and the prior grantee based on a paired t-test. The table only includes observations where we have information about service provision for both the prior grantee and the awardee. Sample sizes differ due to nonresponse to the survey question and different shares providing service for Head Start and Early Head Start.

Significance noted as: + $p<.10$; * $p<.05$; ** $p<.01$.

[a] Office of Head Start combined the service areas of four grantees that were designated for competition, resulting in two FOAs, each of which had two prior grantees. For the purposes of the paired t-test of applicants with the prior grantee, we created a hybrid prior grantee where the total enrollment is the sum of the enrollment for the two prior grantees, and values for the other variables in table VII.1—share of teachers with a BA and average child-teacher ratio—are averaged across the two prior grantees.

[b] We report average child-teacher ratio and total enrollment separately for Head Start and Early Head Start programs, because these program characteristics are generally very different for Head Start compared with Early Head Start programs.

Our analyses also examined whether characteristics of the applicant (i.e., being an incumbent, and whether the agency was a public, for profit, or nonprofit organization) and the proposed services (i.e., use of delegates, proportion of teachers with BAs, average child-teacher ratio, total enrollment) predicted whether an applicant won the award. Table VII.2 reports the results of a probit regression identifying factors associated with winning. In the "All" column, we examine whether agency characteristics predicted winning the award. This analysis includes the 103 applicants to FOAs with competition (i.e., more than one applicant) for whom we have data on applicant characteristics (from PIR, survey data, or publicly available information). The analysis shown in the "Survey" column further restricts the sample to 50 of those who responded to our Applicant Survey to allow us also to examine whether the proposed use of delegates, proportion of teachers with BA, and ratios and enrollment predicted whether an applicant won the award. The results indicate that relatively few factors predicted winning the award (table VII.2). In both specifications, incumbent status and being a public entity (compared with the reference category of being a nonprofit) are associated with the probability of winning, and the magnitudes of the coefficients on these variables are similar for the two specifications. In the "Survey" column, none of the measures of proposed service provision were significant predictors of whether an applicant won the award when controlling for applicant characteristics.[69]

---

[69] Because of survey nonresponse, the sample size in the second analysis is small, which may hinder our ability to detect significant impacts.

TABLE VII.2.

Probability of Winning When There is More than One Applicant

Variable	Mean/Percent		Marginal Effects[a]	
	Awardee	**Non-Awardee**	**All[b]**	**Survey[c]**
Number of Applicants	2.82 (0.88)	3.22 (0.91)	-0.05 (0.04)	-0.02 (0.09)
Agency Characteristics				
Incumbent	59%	15%	0.46** (0.13)	0.40* (0.19)
Auspice				
Nonprofit	78%	75%	Reference Category	Reference Category
Public	18%	7%	0.30* (0.13)	0.35+ (0.21)
For Profit	4%	10%	0.02 (0.16)	-0.07 (0.28)
Proposed Services				
Delegates	39%	36%		-0.00 (0.17)
Prop. of Teachers w/ BA	0.76 (0.27)	0.75 (0.31)		-0.04 (0.24)
Head Start Child-Teacher Ratio	20.37 (20.93)	12.79 (6.87)		0.00 (0.00)
Early Head Start Child-Teacher Ratio	6.93 (6.67)	7.36 (3.85)		-0.02 (0.02)
Head Start Total Enrollment[d] (Compared with Existing Grantee)	-0.21 (0.42)	-0.22 (1.22)		0.00 (0.01)
Early Head Start Total Enrollment[d] (Compared with Existing Grantee)	0.47 (2.65)	0.22 (1.16)		0.03 (0.07)
Observations			103	50
Pseudo R^2			0.25	0.31

Sources: Data on agency characteristics come from the OHS administrative data supplemented with PIR data, survey data and publicly available data. Data on proposed services comes from the evaluation's Applicant Survey.

Notes: Standard errors in parentheses are clustered at the FOA level. Restricted to competitions with more than one applicant.

Significance noted as: + $p<.10$; * p<.05; ** $p<.01$.

[a] A marginal effect is the change in the probability of winning associated with a one-unit change in the applicant characteristic or proposed services.

[b] Includes applicants who applied to FOAs with competition for whom we have data from either the PIR, survey or publicly available data.

[c] Restricted to applicants who applied to FOAs with competition who responded to questions about service provision in the Applicant Survey.

[d] Normalized as the difference between the proposed enrollment of the applicant and the enrollment of the prior grantee divided by 100.

Next, we discuss qualitative findings related to how awardees experienced the negotiation and award processes and explore how these experiences might support or hinder the implementation of changes to services proposed in the applications.

Finding 3: The timing and confidentiality of the negotiation process created challenges in preparing for implementation of the proposed changes to services for most of the interviewed awardees.

Staff from all nine of the grantees participating in the awardee interviews reported they were first unofficially notified of their awards through a phone call from their Head Start regional office in April, May, or June. The awardees reported that they received their official award letter around July. Awardees reported that during the initial phone call, they were told they were not allowed to tell anyone outside of a select group of leadership staff that they had won the award until the official letter arrived. The awardees also reported that they were all required to sign a confidentiality agreement agreeing to these terms, although at least one grantee was unsure that the confidentiality agreement was binding. After they were notified of the award by phone, they reported that there were approximately two months of negotiation before the official award notification.

The negotiation processes sometimes involved making changes to their proposed service provision, structure, and budgets. Five of the nine sites said they had to change their proposed service provision in response to these negotiations either due to requests for reduced budgets or requests for specific changes.

> We had wanted to do more center-based Early Head Start, but our cost per child was too high, and we couldn't make it work with that many Early Head Start slots, so we converted them back to home visiting.—Incumbent Awardee

Awardees reported that during this negotiation period, they were not allowed to tell their staff or partners that they had any news about the award status. All nine awardees expressed frustration at the confidentiality required during the negotiation process. Grantees reported that secrecy created three primary issues: it strained relationships with staff; it strained relationships with partners; and it delayed preparation for implementation.

Incumbents, in particular, faced challenges related to staff relationships and morale. Three of the four incumbent awardees and their staff experienced uncertainty about the security of their employment for more than two years. The delay in the ability to tell staff, especially staff that were out of the office over the summer, strained their relationship with staff, and two incumbent awardees reported losing staff as a result.

> *We couldn't communicate that to staff. The majority of staff are off for the summer, and we could not communicate when they left whether they would have jobs when they came back. It was very difficult. [We] had that on our shoulders for two months. Staff would ask daily, and it felt so wrong to basically not be forthright with people when you're talking about their livelihood. —Incumbent Awardee*

> *It's definitely been challenging from a senior leadership perspective, from an overall morale perspective, at staffing levels, the folks that run these programs.—Incumbent Awardee*

While these negotiations affected their partners or delegates, grantees reported they were not allowed to consult with them. Five of the nine awardees reported this as an issue with the negotiation process. One new awardee was told in the negotiation process that they could not delegate to a particular organization that they had written into their proposal. That delegate was supposed to serve hundreds of children. At the same time, the new awardee needed to find new slots for the children, but found the confidentiality clause made that difficult: "How do I find homes for the [hundreds] of slots if I can't tell people we need to?"

Four of the five new awardees described similar issues related to confidentiality, as it prevented them from managing reactions to their award, building relationships, and preparing for implementation. Some problems occurred because not all the parties that were supposed to keep the information confidential did so.

> *We knew they'd [the prior grantee] closed their doors and laid people off even before it was formally announced—we knew and [the prior grantee] knew that this was going to happen and we had [the prior grantee] staff calling our office. So they'd say, 'I've been laid off and I heard y'all will handle operations.' And at that point we couldn't tell them; we had to say, 'You're getting word ahead of what's been formally announced by Head Start, but if we win, we'd love to hire you,' or something. They let people go before even the formal announcement was made. —New Awardee*

> *The problematic aspect is that you can't control what gets out in the community, so you have some people who know for certain you will be the grantee, but you can't tell them. That's a problem. It creates bad will with people before you have a chance to create good will. It messes up the relationships you want to create with people if you can't talk to them about being the provider who will get things done.— New Awardee*

> *There were lots of things, lots of moving pieces…and we didn't hire a [Head Start] director until the end of July, and she started in August. For a long time it was just [an interim Head Start director], which was hard. Same thing for educational partners. They can't hire until they have money. And we didn't get*

notice of award until July, and then went through multiple versions of a contract before finalizing. — New Awardee

Once they received the official award, the awardees we interviewed reported that the timing of the award process presented many challenges. The awardees indicated receiving their official award notice in July, but two of them said they received the award letter after they were supposed to start providing services.

Interviewer: When did you finally found out given award?
Respondent: July 7 but that grant year was July 1. —Incumbent Awardee

Even just getting the actual notice of award was challenging. For whatever reason, the numbers didn't match, so our grant was effective July 1, but we didn't receive award notice until the end of July. Prior to that, we had been told in April or May that our fiscal year would be one thing, but it changed multiple times before arriving at the current one, which is February to January. —New Awardee

All five of the new awardees we interviewed reported multiple challenges in transitioning. The primary challenges they cited were trouble taking over staff, buildings, and equipment from the old grantee, and state child care licensing. When applicants wrote their proposals, they had access to information provided by the prior grantee about assets purchased with Head Start funds that would accrue to a new awardee and information about the spaces occupied by the prior grantee that the new grantee might take over. The applicants also indicated how much time they thought they needed to make the transition of services from the prior grantee to the new awardee; new awardees we interviewed reported those estimates were based on easy acquisition of the federally funded equipment and space. They also anticipated initially staffing their programs with the staff from the old grantee given limited labor pools for qualified individuals and the short start-up time expected.

Thus, challenges acquiring the former grantees' equipment, space, or staff interfered with implementing new services. The extent of these challenges seemed to depend on their relationship with the former grantee and the extent of the change. Three of the five new awardees had good relationships with the former Head Start grantee and therefore were more informed about the resources available to them and faced less resistance in the transition. While those with a better relationship with the former grantee faced fewer challenges, coordination issues still arose.

We had all these months trying to plan in negotiations but we couldn't do things because they hadn't notified [the delegate of the prior grantee] yet that they weren't going to be able to play. So we couldn't see facilities or warehouse, and our hands were tied, yet we were having these meetings that you had to get this done, but we can't because there hadn't been this notification process. They needed a very effective time schedule to make this happen, and since that was all delayed until the end of the year, we were already into July and we don't have an official letter yet, but were unofficially told.—New Awardee

The two new awardees without good relationships with the prior grantee both faced resistance from the prior grantee, which made it difficult to take over buildings, hire staff, and transition children.

> *We wrote a grant [the FOA proposal] saying we'd step in their shoes and operate their facilities… Anyway, [the former grantee's refusal to give up the space has] been really, really, really problematic. If we don't meet our funded enrollment goals this year, it's going to be because we don't have the classroom capacity.—New Awardee*

> *For confidentiality reasons, they wouldn't give up the children's records. We had to conduct a mass mailing to get authorization complete for them to turn over those records, and then when they turned them over they were incomplete. I don't know if they were incomplete because they were incomplete or if they're not sharing. We can't, as the new grantee, walk into their office and say, 'You need to do xyz today or tomorrow.' It seems like OHS needs to be the heavy and get the former grantee in line and be very directive of it.—New Awardee*

> *They wiped everything clean, destroyed the paperwork so there was nothing for us. One of the challenges to be addressed going forward, if Office of Head Start realizes the existing grantee won't get grant back, they need to ensure that the paperwork will be there. The documentation was gone. They wiped everything clean. They moved out all the records, cleaned out all the computers. That's huge for a grantee coming in. To protect [the records]. I think one of the reasons they didn't is because they gave the staff records to staff. Anything related to children, we only had a hard copy document so we had to put that data back in. We kept plugging away until we got it up and running.—New Awardee*

All five of the new awardees cited issues getting their state licensing; they said the timeline from award to the expected start of service delivery did not account for the time it takes to get licensed. Even the new awardees already providing Head Start in other areas had issues because they were either opening new sites or required licensing in the new region.

> *We couldn't jump in and take over, we had to follow processes. We had to get all of our centers licensed. We weren't grandfathered in. We had to start that process over. —New Awardee*

> *To get this [in early July] and to be expected to be up and running [in early September] and at full enrollment, and you don't have a lease signed and you weren't allowed to speak to anyone before 4:00 p.m. that day and that was 4th of July weekend and so you already lost three days…that sounds absolutely ridiculous. But that was putting bids out for food, licensing, inventory, I mean just everything. It's so much of a blur, I mean those weeks because I worked 65to70 hours for most of July, there just weren't enough weeks to do everything. We worked weekends and such. We had to have everything re-licensed, every single site. Who would start a $7 million corporation and expect it to be up and running in two months? I was amazed we did it.—New Awardee*

Three of five new awardees felt that the regional office did not provide enough support. They faced particular challenges in preparing to take over for the former grantee. The lack of communication meant that getting contracts into place with partners, being able to assess the quality of the prior grantees' buildings and equipment, hiring staff, etc. all had to be delayed until the official award was announced.

I can't tell you how much I wish that OHS or the regional office would have sent in a team to come help us through transition. Not just for us, but the former grantee and delegate, closing out things and transitioning them to us. It would have made a world of difference to have a couple of people from the regional office here for a couple of weeks to help us through some of the things we went through. —New Awardee

I don't always get straight answers when I need help. If I hear one more time that it needs to be developed for your county, there has to be a guideline. What's reasonable and what's not? To a new grantee, you might think it's reasonable, but if you haven't been in the system before, you don't know what's reasonable. If we had more guidance from the regional folks, it would have been so much easier. Just to give us the DRS transition notebook was not helpful in prioritizing. —New Awardee

On the other hand, two felt they were adequately supported.

As for the regional office, I cannot say enough positive things about them and the support we got and continue to get from them. —New Awardee

The four incumbent awardees said that their greatest challenge was implementing their proposed changes in the short time frame. All the incumbent awardees had proposed major changes to their service provision to improve quality, including restructuring and providing new services (Early Head Start). The delay in being able to announce the official award meant that these proposed changes had to be made in a compressed time frame leading to concerns about the quality of implementation.

It forced us to hire in positions they didn't have experience for. There weren't specific qualifications, but they really weren't suited for this job. It affected this year. Now you have family advocates that are out there that don't know what they're doing. They have a huge learning curve, and it builds on itself. That's one judgment that would have changed so much had we been given five or six more months.—Incumbent Awardee

My site isn't what it was. Kids aren't getting food on time, teachers don't have time for circle time. With that now, we have teachers doing food services, and it's taking away from time and quality for children. —Incumbent Awardee

Four of the nine awardees did not receive the amount of funding they had expected. One incumbent awardee had its fiscal year changed from April to March to July to June. They received some funding to continue to provide services from April to July, when the award was made. This amount was subtracted from the total award for the current fiscal year, meaning that for this year, they have three months' less funding. Another grantee said they received less for this year because the former grantee spent more than they were supposed to. This means that for this year, they have to provide the same level of services but with a reduced budget.

We used to go from April 1 to March 31, now the funding cycle is from July 1 to June 30. So, we were under the last year of this three-year plan, and my funding is supposed to run out on March 31, but my school year doesn't end until the end of May. We get this letter while we're still waiting on whether we won, saying that we can apply for some continuation money. This is troubling for me. You said apply for

continuation money, and it took us maybe a three-month period, right at $500,000. You told me I got this money for continuation, but then turn around and tell me that this money is going to be deducted from the grant. So March 31, I'm out of money again. So I actually didn't have $2.2 million to operate the program from July 1 to June 30. —Incumbent Awardee

It wasn't what we were told to expect. When we got the grant award, it was a different number, because it was the amount leftover from [the former delegate], and the fact that they went over took away from us. The numbers kept changing and it made it very difficult. —New Awardee

These challenges led to service interruptions in four of the nine sites, three of the new awardees, and one of the incumbent awardees. However, for one new grantee the transition was so smooth that they believe that most of the community is still unaware that there was a change in provider.

We did not offer a summer program. Our plan at the beginning of the year was to offer a summer program, and we had communicated that to staff and family. So when we were re-awarded in May 2015 and it was reduced, we knew we couldn't offer a summer program. Any disruption happened over the summer.—Incumbent Awardee

For some we weren't able to open it. Our president said, 'We will open September 1,' and I said, 'No, no, no, don't commit to a date!' And we weren't able to get all facilities open, and the media was on us. We got some open by working day and night, calling politicians [to speed up the licensing process]. Our infant and toddlers facility, we're still working to get it up. We don't have the staff yet. You can't look for staff before July 1. We couldn't do anything prior to that.—New Awardee

The regional office expressed that top priority was no disruption to kids in full-day, full-year option. Those kids went uninterrupted; it was seamless for those kids. Those families didn't see any interruptions in services. But the part-day programs—normally services would have begun in August or September, but we've had licensing challenges and facilities challenges. Also, just the delay in getting grant award was a challenge, because we couldn't issue contracts.—New Awardee

I don't think that people outside of the table managing, the internal stakeholders, I don't think the end users were aware of that, as long as services continued. To the people around the table it was a lot to deal with it, but for folks in the community I don't even think they're aware, and if you ask them right now I don't think they'd realize there has been a change.—New Awardee

Summary

The DRS competitive process largely resulted in the incumbents winning back their grants: 76 percent of awards went to incumbents, 17 percent to other Head Start grantees or delegates,[70] and 6 percent to a new organization. On average, we found that awardees proposed higher proportions of teachers with BAs, lower child-teacher ratios, and lower enrollment than the prior grantee. This was true even when

---

[70] Two of the awardees were Head Start delegate agencies that were not also Head Start grantees in other service areas.

we restricted the analysis to FOAs with just one applicant. This suggests that being designated for competition and applying competitively may lead to higher proposed service quality, even without the presence of competitors. Finally, in our exploratory analysis of the negotiation, award, and transition processes of nine awardees, all but one of the nine reported the timing and confidentiality of the negotiation process created challenges in preparing for, and initial implementation of, the proposed changes to services.

Chapter VIII: Summary, Discussion, and Next Steps

The goal of this study is to identify how well the elements of the Head Start Designation Renewal System (DRS) are operating as intended in their early implementation to improve accountability and quality, and to describe how grantees and new applicants experience the system. The following sections summarize the findings and then discuss how they may be helpful in refining the monitoring system, using the DRS to improve Head Start quality, and improving the grantee experiences during the monitoring and competition phases of the system.

Summary of Findings

How do grantees perceive the DRS and prepare for their monitoring and assessment in terms of efforts to improve program quality?

Nineteen of 35 Head Start directors participating in telephone interviews agreed that a system of accountability such as the DRS was needed to improve program quality in Head Start. Most grantees indicated that most of the seven conditions were fair, but 74 percent expressed concerns about the CLASS condition, particularly the lowest 10 percent specification of the condition. All 66 grantees participating in the PDTA survey reported engaging in multiple professional development and technical assistance strategies in the 12 months before their monitoring review, but in the PDTA survey, 65 percent indicated that these strategies were not any different than what they would have done at any other time. In qualitative interviews with 35 directors, grantees indicated that they are working on improvements all the time so it is difficult to identify which improvement efforts are specific to the DRS, but the directors pointed to some changes in program structure and procedures and some added emphasis on CLASS and School Readiness Goals. In addition, in the on-site interviews, all 15 grantees indicated that they had been working on improving CLASS scores for the last two to six years, with some of them anticipating that CLASS would become a requirement based on the passage of the 2007 Head Start Act. Given the long time horizon that grantees express in preparing for the DRS, it is difficult

to pinpoint changes made specifically for the DRS and changes made to improve quality that are not related to the DRS.

Does the DRS differentiate higher versus lower quality programs?

Comparisons of grantees that were and were not designated for competition did not reveal reliable differences in the quality of classroom or center practices, but some evidence of differences did emerge when individual DRS designation conditions were examined. Analyses compared the 71 randomly selected grantees on measures of classroom quality, health and safety, family engagement, program governance and management, and fiscal operations and the 216 nonprofit grantees on financial vulnerability. Analyses first compared grantees by overall designation status. Then, to evaluate whether specific DRS conditions are differentiating higher versus lower quality programs, analyses contrasted grantees designated for competition due to deficiencies in meeting Head Start Performance Standards with grantees designated due to low CLASS scores. These two designation criteria account for 99 percent of designations.

The first set of analyses, comparing grantees by overall designation status, did not show significant differences. However, grantees designated due to deficiencies had lower quality on three out of six measures of center quality compared to grantees that were not designated in analyses that accounted for missing data. In contrast, no analyses indicated that grantees designed due to low CLASS scores differed from grantees that were not designated on any study measure of quality.

In addition to asking whether grantees that were or were not designated for competition differed on observed quality measures, we also examined the use of CLASS more closely because it was the newest, and perhaps most controversial, condition for monitoring Head Start programs. This may be the first time that a monitoring system combined classroom observational quality measures within a grantee to represent the whole grantee. Examining the psychometric properties of the CLASS as used within the DRS provided a mixed pattern of results. There was no evidence that scores were reliably different when either the typical four or fewer cycles of observation were compared, suggesting that the practice of collecting two cycles per classroom in the OHS monitoring did not change grantee-level CLASS scores. In contrast, evidence suggested substantial differences among raters in the CLASS scores and in other measures of classroom quality such as the ECERS-R. For example, our analyses of CLASS scores collected by the evaluation team suggest that variability among raters may account for up to 45 percent of the variability in the CLASS. Because training of DRS CLASS observers was similar to training

of the evaluation team CLASS observers, variability among DRS CLASS raters may be similar. Furthermore, there were large differences in the CLASS scores collected by the OHS Monitors and by our data collectors, and only limited evidence of correlations in scores across the two data collection teams. These findings were obtained even when the time elapsed between the two assessments (mean = 1.2 months) was considered or whether the same or different classrooms[71] were observed by both teams.

How do grantees perceive the monitoring, assessment, and DRS designation processes in terms of preparing them to implement quality improvements to existing services?

All 15 Head Start grantees interviewed in on-site follow-up visits after they knew their designation status indicated they were continuing to work on improving quality. Grantees designated for competition were more likely than not-designated grantees to indicate some of their quality improvement efforts were related to the DRS. Some efforts were focused on CLASS, and some were focused on specific areas of need identified by the program, such as training for substitute teachers on positive guidance and establishing new practices for ensuring children are not left alone. Grantees expressed concerns about the early implementation of the monitoring, assessment, and notification processes including concerns about determining quality based on a single visit or incident (20 of 35 grantees), transparency in the designation decision-making process (8 of 15), and the length of time for not knowing the designation status (6 of 15).

What does competition look like?

Based on the evaluation team's analysis of OHS data on grant competitions, 42 percent of the 103 Head Start grants opened for competition in 2014 had more than one applicant. Merging the OHS data about applicants with PIR data on existing Head Start grantees shows that 91 percent of incumbents applied. The applicants for the grants that were not incumbents (labeled new applicants) were almost equally

---

[71] Whether the same or different classrooms were observed was indexed by whether all classrooms within a grantee were observed by both teams.

divided between current Head Start grantees or delegates in other service catchment areas (54 percent) and applicants that were new to Head Start[72] (46 percent).

Factors related to the decision to apply were different for new applicants and incumbents, as reported during interviews with directors and staff. The incumbents indicated they had not expected competition prior to monitoring, but many were concerned there might be competition after being designated. All incumbents designated for competition indicated they had already decided to apply before the formal announcement calling for proposals. The interviews with five new awardees indicated that several factors determined whether they applied, including the effort required to adhere to federal reporting standards, the well-being of children in the community, the potential reaction of the community, and likely competition.

Based on surveys sent to all applicants for the grants opened through competition in 2014, most applicants—including incumbent applicants—proposed changes to service delivery, including more delegates, higher shares of teachers with BAs, and lower Head Start enrollment than the current grantee. Similarly, based on more in-depth qualitative interviews with 9 awardees, all but one proposed service changes. All four of the incumbents felt the competitive FOA provided opportunities to propose changes to their programs they could not otherwise have made, including adding or increasing Early Head Start slots and increasing the share of teachers with a BA.

How do grantees perceive the negotiation, award, and transition processes in terms of preparing them to implement quality improvements to existing services?

The DRS competitive process largely resulted in the incumbents winning back their grants: 76 percent of awards went to incumbents, 17 percent to other Head Start grantees or delegates,[73] and 6 percent to a new organization. On average we found that awardees proposed higher proportions of teachers with BAs, lower child-teacher ratios, and lower enrollment than the prior grantee. This was true even when we restricted the analysis to FOAs with just one applicant. This suggests that being designated and applying competitively may lead to higher proposed service quality even without the presence of competitors. Finally, in our exploratory analysis of the negotiation, award, and transition processes of the nine awardees, we found that the timing and confidentiality of the negotiation process created

---

[72] We determined if an applicant was new to Head Start in two ways. If they answered the survey, they were asked if they had ever had a Head Start or Early Head Start grant or if they had ever been a delegate agency. For organizations that did not participate in the survey, we examined the 2013 PIR data; if they were not found in the PIR data, then they were labeled as new.
[73] Two of the awardees were Head Start delegate agencies that were not also Head Start grantees in other service areas.

challenges in preparing for, and initial implementation of the proposed changes to service for eight of the nine interviewed awardees.

Study Limitations

Before discussing the findings, we must acknowledge the study limitations.

First, this is a formative evaluation. The qualitative and quantitative components of the study focus on describing the experiences of grantees, and, as such, do not allow for making causal inferences. The goal is to identify whether the Head Start Designation Renewal System is able to differentiate grantees with lower quality services from other grantees, and to explore the experiences of a small number of grantees during monitoring, and experiences of designated grantees and new applicants during competition. The findings from this study help identify DRS components or practices that may (or may not) serve as mechanisms for improving quality in Head Start. The study, however, cannot answer questions linking DRS to program quality improvement or child and family outcomes.

Second, parts of this study may not generalize to all Head Start grantees monitored using the DRS. We decided not to include Early Head Start-only programs, home-based-only programs, or Migrant and Seasonal Head Start (MSHS) programs in the examination of how DRS incentivizes quality improvement through monitoring and assessment (research questions 1—3). Early Head Start-only programs and home-based-only programs were excluded because they are not assessed using the CLASS and there are different monitoring protocols for Head Start and Early Head Start grantees. Classrooms serving children younger than three years and home-based programs also would require different assessment instruments by the evaluation team and a substantial increase in the sample size. Similarly, we did not include MSHS programs in this part of the study because of measurement limitations (i.e., the validity of quality measures for use with this population is unknown). We included Early Head Start-only, home-based-only and MSHS programs in the examination of the competition and award process (research questions 4—5) because the mechanisms for competition and award are the same and measurement issues were not a barrier to inclusion in this part of the study. American Indian Alaskan Native Head Start grantees are not included in any portion of the evaluation because they are subject to a different process in DRS (i.e., tribal programs participate in government-to-government consultation to establish a quality improvement plan before being designated for competition). Grantees located outside the continental United States were excluded because of the costs of traveling to those locations.

Third, it was more difficult to recruit grantees into the study than we anticipated. Response rates were lower than expected for the quality assessments conducted with 71 grantees (46 percent), as well as the applicant survey sample of the 182 applicants (65 percent), and we drew subsamples of these for the other measures in the study. The recruitment rates were lower than prior Head Start evaluations, such as the Head Start Impact Study and FACES, but were similar to or better than those reported in recent state-level evaluations of Quality Rating and Improvement Systems (Yazejian 2016 private communication regarding the NC and IL QRIS evaluations). A comprehensive response bias analysis indicated some differences among both the designated and the not-designated grantees that did and did not agree to participate, and weights were applied to account for those differences. Nevertheless, it is not possible to eliminate the threat of bias due to issues in recruitment. See technical volume II for more information.

Fourth, the measurement of the center- and grantee-level operations, family engagement, staff qualifications, and child assessments with the PAS was problematic. We changed the data collection plan for the PAS during the middle of data collection. Originally, the PAS was collected in all centers where classrooms had been sampled for evaluation observations, but Head Start directors objected to this measure due to the time required to prepare for these visits in multiple centers. To secure cooperation from the Head Start directors, we selected a single center per grantee for the last half of the data collection. Multiple imputations were used to impute missing PAS scores. Analyses testing for differences in data-collection protocol indicated that this change in protocol may have introduced bias in our analyses of the PAS data scores that may or may not have been adequately accounted for by using multiple imputations.[74] Accordingly, we are less confident about drawing conclusions from this instrument than from our other quality measures.

Fifth, grantees were purposively selected for the telephone interviews, on-site follow up interviews, and awardee interviews. The goal of these interviews was to explore the experiences of a small number of grantees to better understand potential mechanisms through which DRS may support (or impede) quality improvement in Head Start. The perceptions, experiences, and actions of these grantees may not be representative of the larger population of Head Start grantees. The grantees were selected to provide a variety of perspectives and experiences, rather than represent the overall or average experience. The samples included Head Start programs of different sizes, from different regions, and operating both birth-to-five and Head Start-only services. We overrepresented grantees designated for competition among the telephone and on-site interviews to provide more variations on the experiences

---

[74] See technical volume II, appendix B for these analyses and discussion of the issues.

of designated grantees. Similarly, we overrepresented new awardees in the awardee interviews to capture a variety of perspectives of new awardees because transitioning grants from one organization to another is a less frequent activity in Head Start and therefore may have more challenges associated with it and more variations of challenges based on organizational and community characteristics.

Discussion and Suggestions for Future Research

Very few programs, services, or systems are totally successful in their initial implementation, and most are improved over time using the evidence regarding what appears to work and what may need further refinement. We hope this study provides some of that evidence. Below, we put our findings in the context of other research to help readers consider the meaning of the findings. We also suggest future research that could lead to additional insights and improvements.

Identifying lower quality grantees using the DRS

This study found limited evidence that the DRS was successfully identifying lower quality grantees. Null findings do not imply that differences do not exist, and cannot be interpreted as indicating that grantees designated and not designated for competition provide the same level of quality. However, the results do raise questions about the current system, especially about the manner in which the CLASS is being used within the DRS.

Some evidence emerged suggesting that monitoring using the Head Start Performance Standards was able to identify programs that may be experiencing issues with program and center-level operations. This provides some, albeit limited, evidence that designations based on the monitoring of Head Start Performance Standards may be identifying grantees with issues at least in program and center-level operations. In contrast, classroom quality measured by the evaluation team was not systematically or statistically significantly different when not-designated grantees were compared to all designated grantees, grantees designated for low CLASS scores, or grantees designated for deficiencies.

OHS and the evaluation team used identical standards for certifying data collectors, and similar strategies for training data collectors and for sampling classrooms within grantees, so CLASS differences between the monitoring and evaluation teams were not likely due to those factors. Furthermore, analyses suggest these differences were not due to differences in the format of the

CLASS—that is, using of two instead of the usual four cycles per classroom, the amount of time between the two evaluations, or in using different sampling ratios for different sized grantees. Instead, there was some evidence that all of the classroom observation measures may suffer from substantial rater effects. These rater effects were observed not just for the CLASS, but also for the other classroom quality measures—the ECERS-E, ECERS-R, and TSRS. All data collectors were certified on the measures they collected—typically meeting inter-rater reliability criteria of agreement within one point on at least 80 to 85 percent of the items with a trainer or the anchor score on the certification tapes. Therefore, a likely explanation for the discrepancies between OHS CLASS scores and those of the evaluation team observers is that this criterion for reliability (i.e., agreement within one point on 80 percent of items) is too liberal to ensure adequate agreement and, possibly, that observation on a single day is not sufficient to provide a generalizable description of quality.

Using the proportion of items with agreement within one point to determine inter-rater reliability is the standard used for almost all preschool classroom quality measures (Halle, Whittaker, and Anderson 2010). Training data collectors is complex and it is almost always difficult for at least some people to meet this criterion. This is likely problematic, however, because the standard deviation (SD) on the summary scores for preschool quality measures is also often about 1. In the 549 classrooms observed in the evaluation study, the SDs ranged from 0.73 for CLASS Emotional Support to 1.14 for CLASS Instructional Support, with SDs of 0.82 for ECERS-E Academic Activities, 0.84 for TSRS Class Structure/Management Scale, 1.04 for CLASS Classroom Organization, and 1.09 for ECERS-R Interaction score. Thus, data collectors in this study were certified as reliable when they scored within about 1 SD of the gold standard.

The Measuring Effective Teachers study reported substantial variation in K-12 classroom observational measures linked both to rater and time (Kane and Straiger 2012). They found the reliability of the CLASS (i.e., proportion of variance due to systematic differences among teachers) and other classroom observation measures to be about 35 percent when a classroom was observed once by one rater, but reliability could be increased to 53 percent if the classroom was rated by two observers on different days and 65 percent if rated by four observers on four different days. They suggested that reliable classroom observations should involve having the classroom observed by multiple raters on multiple days. Furthermore, they suggest that rater effects can be reduced by using video-tapes of classrooms that are coded by a small number of raters who were highly trained and who had to pass a certification test before rating classroom videos each day.

The DRS might be the first monitoring system to include classroom observation scores computed to represent programs and large grantees, and this study is almost certainly the first to examine across-time reliability of CLASS scores computed at a level beyond the individual classroom. This study provided an opportunity to examine the use of the CLASS as a grantee-level measure of quality in several ways, and to address some of the criticisms of its use with the DRS by: comparing scores when two versus four cycles were observed per classroom, estimating the degree to which there are systematic differences in scores collected by different raters, and comparing scores as collected by the OHS monitors and by our data collectors. We found no systematic evidence that number of cycles changed scores,[75] but found substantial differences among raters and between the two teams of data collectors. The OHS monitors rated grantees significantly higher on almost all CLASS domain scores than did our data collectors. The CLASS scores across the two data collection teams were not consistently correlated with each other. The lack of correspondence across the two data collection teams was not accounted for by the time between data collection (mean = 1.2 months) or by the fact that different classrooms might be included in the grantee-level summary score as computed by the two data collection teams.[76]

These findings raise questions about how classroom quality measures are being used within monitoring systems. To date, the focus of various monitoring, accountability, and quality rating systems using CLASS and other observational measures has been on certifying data collectors with the developers or the designated trainers. It is likely the criteria to certify data collectors are too liberal, and that might account for the large discrepancies observed across data collection teams in this study. Soon the evaluations of the Quality Rating and Improvement Systems in the Race-to-the-Top Early-Learning First states will also be able to compare ratings by the monitors and by the evaluators to examine the extent to which the systematic differences are observed in those evaluations as well.

Many facets of this evaluation require further study to draw more definitive conclusions. First, the issue of the reliability of classroom observations for the purpose of monitoring needs further examination. Perhaps future studies could examine whether the stability and consistency of grantee-level averages of CLASS or other classroom-level measures of quality might be improved if data collectors were required to meet higher standards for reliability or if a large number of raters collected data for each grantee. Second, future research is needed to identify valid and reliable measures of program- and center-level operations and quality. The PAS was the only measure we could identify to

---

[75] These analyses were based on CLASS data collected by the evaluation team.
[76] See sensitivity analyses in technical volume II, appendix B.

examine center-level operations and governance, but we found it difficult to administer and it has not been widely used in other research. Perhaps future studies could identify or develop measures to closely align with the specific types of Head Start Performance Standard violations that are most likely to result in designations for competition and relate them directly to grantee and center operations and governance.

Competition as a Mechanism to Improve Accountability and Quality

Forty-two percent of the FOAs garnered more than one competitor, and the competition ranged from two to six applicants. Previous research suggests that this level of competition may be what we should expect.

The DRS introduces a form of competition that Kincaid (1991) refers to as mediated competition – competition that is initiated and decided through the institutions of government, rather than through the market. Previous research suggests that mediated competitions generate little competition, but facilitate the formation of collaborative community partnerships and/or increase access to additional resources (Hefetz and Warner 2011; Warner and Hefetz 2003). According to a survey of local governments by Hefetz and Warner (2011), competition varies by type of service and the urbanicity of the municipality, with metro and suburban cases experiencing the highest competition. In their study, child care facilities' operation would be the service most closely comparable to Head Start. Competition for that service tends to be among the highest in local government procurement, with a typical competition yielding 3.7 competitors in metro areas, 3.3 competitors in suburban areas, and 3.5 competitors in rural areas. Running a Head Start program, however, is more complicated than running the typical child care program because most grantees operate multiple centers, because of the need to meet state and federal requirements for the child care portion of the program, and because of the additional comprehensive and parent-centered services that are mandatory elements of Head Start programs. Thus, actual competition for Head Start programs is likely to be less than competition for child care administration. In addition, it is important to note that many Head Start programs operate in rural or isolated jurisdictions where there may not be a lot of potential applicants operating. In fact, 46 of the 103 FOAs from the 2014 Head Start competition were located in counties with a density of 100 or fewer people per square mile, and in those low-density counties, only 30 percent had any

competition. In contrast, in the 25 FOAs in high-density counties of more than 500 people per square mile, 48 percent had some competition.[77]

Our study assessed whether FOAs with more competitors garnered proposals with more structural quality changes than FOAs with fewer competitors, but the results showed no significant differences. Interpreting those results is difficult because of the small number of FOAs with more than one applicant. Interview results indicate that perceptions of competition rather than actual competition are more likely driving the incentive structure (six of nine grantees interviewed faced competition in their FOAs, but all nine thought they would). Future research across DRS cohorts may better ascertain whether more competition leads to stronger incentives to improve quality.

DRS and Quality Improvement in Head Start

Structural Quality. Both incumbent and new applicants that competed for the FOAs indicated that they proposed changes linked to quality improvement. Compared with the grantee designated for competition, the proposals for the new grant tended to include higher proportions of teachers with bachelor degrees and lower Head Start (not Early Head Start) child-teacher ratios. These structural quality indicators are widely viewed as promoting higher classroom quality and teacher-child interactions (Burchinal et al., 2015). This study could not document whether these proposed changes were implemented when the new grant was awarded and whether the quality of teacher-child interactions was improved, but these proposed changes are encouraging and consistent with the quality improvement efforts within state and local quality rating and improvement systems (Build Initiative and Child Trends 2014). Future research may be able to document these changes and determine whether the new awardee was providing higher quality services than did the grantee designated for competition.

Quality Improvement and Child Outcomes. The focus on quality improvement within the DRS is consistent with that focus within the early childhood education literature, yet there are growing concerns that current concepts of quality are so modestly related to child outcomes that quality improvement efforts may not improve child development and learning (see Burchinal et al. 2015 for review). Recent analyses across multiple studies suggest significant but modest associations between the quality of teacher-child interactions or global classroom environmental quality and child outcomes (e.g., typically partial correlations of 0.07 or less). Perhaps the largest impacts on child outcomes have

---

[77] These results are based on OHS data about applicants for the 2014 FOAs, combined with census data about the counties in which the FOAs are located.

been observed when professional development and technical assistance focused on the quality and content of instruction (Duncan and Murnane 2011; Powell and Diamond 2011; Wasik 2010; Wasik and Hindman 2011). The research designed to identify the aspects of program quality that relate to academic and social skills among children and family engagement and involvement is ongoing (e.g., the Institute of Education Science's Early Learning Network and other professional development studies in Head Start).

Possible Adverse Effects of DRS. Grantees expressed concerns that the DRS might adversely be affecting quality in their programs because of disruptions of service during the transition from one grantee to another and the compressed timeline for start-up of new services when grant awards are made—both for incumbents proposing changes and new grantees starting up services. Future research might determine the net effect of these concerns by exploring the following questions: To what extent are the positive efforts to improve quality diminished by these issues? Are some programs more susceptible to these issues than others? Are there particular strategies that could be used to ameliorate these concerns or are these simply issues that travel hand in hand with accountability systems and transitions? We discuss these questions in the context of the literature briefly here.

Lost Days of Service. According to recent early education research, full-day and extended-year preschool programs offer greater benefits to children and their families (particularly those with working parents) than programs offering shorter program lengths (Ackerman, Barnett, and Robin 2005; Robin, Frede, and Barnett 2006; Tout et al. 2013). Research on preschool absenteeism also suggests a strong positive association between the amount of seat-time and children's school readiness (Ehrlich et al., 2014), and at least one large study in which teachers reported attendance weekly found an association between attendance and gains in child outcomes in analyses that included entry-skill levels as covariates (Burchinal, Zaslow, and Tarullo 2016). Delays in program start dates as a result of delays in grantee notification and other grant transition challenges are not only problematic for program administrators and staff who are faced with handling a big transition with little time, but such delays may also weaken the potential benefits of the program.

Science of Implementation and Transitioning to a New Grant. Implementation science indicates that successful implementation of new strategies requires at least four stages (Metz et al. 2015): exploration, installation, initial implementation, and full implementation. Within each of those stages, successful organizations utilize implementation teams, construct data and feedback loops, and build implementation infrastructure (Metz et al. 2015). If we apply this process to the efforts of incumbents and new applicants to improve quality of Head Start programs during the competition and transition

stages, we see that the first stage—exploration—occurs when the applicants are contemplating and responding to the FOA; they assess needs and resources, consider options, and propose implementation in their FOA responses. The next stages are put on hold until notice of FOA award arrives and they are cleared to take action. Awardees need to take the steps that are part of the installation and initial implementation stages, such as securing and updating spaces, securing and training staff, installing communications systems, etc. But the timeline between notice of award and expected start date compresses the installation and initial implementation stages into very short time frames. This could lead to failure to implement all proposed quality improvements because of inadequate time and resources in installation and initial implementation stages, and it is consistent with comments awardees made during interviews.

In Conclusion

This study of the early implementation of the Head Start Designation Renewal System provides a preliminary understanding of how the DRS may be affecting quality improvement efforts within Head Start, grantee experiences and perceptions of the system, and the way that the DRS is identifying lower quality grantees. Grantee and applicant responses to surveys and interviews indicate that the process of the DRS is incentivizing quality improvements. Analyses of the DRS assessment results, however, indicate that the identification of lower quality programs is not consistent. Future research is needed to understand the net positive benefits of the DRS, and the extent to which the quality improvement efforts it engenders make a difference in actual quality and outcomes for children.

References

Ackerman, D. J., W. S. Barnett, and K.B. Robin. 2005. "Making the most of kindergarten: Present trends and future issues in the provision of full-day programs." New Brunswick, NJ: NIEER.

Administration for Children and Families. 2008. Classroom Assessment Scoring System (CLASS) ACF-IM-HS-08-11. https://eclkc.ohs.acf.hhs.gov/policy/im/acf-im-hs-08-11.

Administration for Children and Families. 2011. Press Release on 132 grants designated for competition. https://eclkc.ohs.acf.hhs.gov/hslc/grants/dr/pdf/Grantees-notified-of-requirement-to-compete-for-continued-funding.pdf.

Administration for Children and Families. 2013. "Head Start Continues Quality Reforms with Second Group of Competitions." http://www.acf.hhs.gov/media/press/2013/head-start-continues-quality-reforms-with-second-group-of-competitions.

Administration for Children and Families. 2014. "Head Start Continues Drive to Raise Standard of its Providers." http://www.acf.hhs.gov/media/press/2014/head-start-third-cohort.

Aronson, S.S., ed. 2012. *Healthy young children: A manual for programs (5th ed).* Washington, DC: National Association for the Education of Young Children.

Bella, J., and P.J. Bloom. 2003. *ZOOM: The Impact of Early Childhood Leadership Training on Role Perceptions, Job Performance, and Career Decisions*. Wheeling, IL: The Center for Early Childhood Leadership, National-Louis University.

Burchinal, M., K. Magnuson, D. Powell, and S.S. Hong. 2015. "Early Child Care and Education and Child Development." In *Handbook of Child Psychology and Developmental Science*, edited by M. Bornstein, R. Lerner, and T. Leventhal. Hoboken, NJ: John Wiley & Sons, Inc.

Burchinal, M., M. Zaslow, and L. Tarullo. 2016. "Quality Thresholds, Features, and Dosage in Early Care and Education: Secondary Data Analyses of Child Outcomes." *Monographs of SRCD* 81 (2), DOI: 10.1111/mono.12236.

Coffey, A., and P. Atkinson. 1996. *Making Sense of Qualitative Data: Complementary Research Strategies.* Thousand Oaks, CA: Sage Publications.

Dropkin, E. 2015. "Analysis of the Designation Renewal System: Cohorts One to Three." Washington, DC: National Head Start Association. https://www.nhsa.org/files/resources/nhsa_drs_analysis_report_-_september_2015.pdf.

Duncan, G. J., and R.J. Murnane, eds. 2011. *Whither Opportunity? Rising Inequality, Schools and Children's Life Choices.* New York, NY: Russell Sage Foundation.

Ehrlich, S.B., J.A. Gwynne, A.S. Pareja, and E.M. Allensworth. 2014. *Preschool Attendance in Chicago Public Schools: Relationships with Learning Outcomes and Reasons for Absences.* The University of Chicago, Consortium on Chicago School Research.

Halle, T., J. E. Vick Whittaker, and R. Anderson. 2010. *Quality in Early Childhood Care and Education Settings: A Compendium of Measures, Second Edition*. Washington, DC: Child Trends. Prepared by Child Trends for the Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.

Hefetz, A., and M. Warner. 2011. "Contracting or Public Delivery? The Importance of Service, Market and Management Characteristics." *Journal of Public Administration Research and Theory* 22 (2): 289–317.

Kane, T. J., and D.O. Staiger. 2012. Gathering Feedback for Teachers: Combining High-Quality Observations with Student Surveys and Achievement Gains. Seattle, WA: The Bill and Melinda Gates Foundation.

Kincaid, J. 1991. "The Competitive Challenge to Cooperative Federalism: A Theory of Federal Democracy." Pp. 87–114 in *Competition Among State and Local Governments,* edited by D. Kenyon and J. Kincaid. Washington, D.C.: Urban Institute Press.

Lambert, R. G., L. Nelson, D. Brewer, and M. Burchinal. 2006. "Measurement issues and psychometric methods in developmental psychology." *Monograph of the Society for Research in Child Development* 71 (3): 24–41.

Mason, J. 2006. *Qualitative Researching, 2nd Edition*. Thousand Oaks, CA: Sage Publications.

Metz, A., S.F. Naoom, T. Halle, and L. Bartley. 2015. *An Integrated Stage-Based Framework for Implementation of Early Childhood Programs and Systems*. (OPRE Research Brief OPRE 201548). Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.

NAEYC and NACCRRA. 2011. *Early Childhood Education Professional Development: Training and Technical Assistance Glossary*.

Office of Head Start. 2011. "Report to Congress on the Final Head Start Program Designation Renewal System." Washington, DC: Administration for Children and Families, U.S. Department of Health and Human Services. Retrieved from http://eclkc.ohs.acf.hhs.gov/hslc/mr/rc/Head_Start_Designation_Renewal_System_Final_Rule.pdf.

Office of Head Start. "Report to Congress on Head Start Monitoring. Fiscal Year 2012." Washington, DC: Administration for Children and Families, U.S. Department of Health and Human Services. http://eclkc.ohs.acf.hhs.gov/hslc/data/rc/report-congress-ohs-monitoring-fy2012.pdf.

Office of Head Start. "Report to Congress on Head Start Monitoring. Fiscal Year 2013." Washington, DC: Administration for Children and Families, U.S. Department of Health and Human Services. http://eclkc.ohs.acf.hhs.gov/hslc/data/rc/fy-2013-head-start-monitoring-report.pdf.

Office of Head Start. "Report to Congress on Head Start Monitoring. Fiscal Year 2014." Washington, DC: Administration for Children and Families, U.S. Department of Health and Human Services. http://eclkc.ohs.acf.hhs.gov/hslc/data/rc/fy-2014-head-start-monitoring-report.pdf.

Office of Head Start. "Services to Children with Disabilities." http://eclkc.ohs.acf.hhs.gov/hslc/tta-system/teaching/Disabilities/Services%20to%20Children%20with%20Disabilities.

Office of Head Start. (n.d.(a)). "Report to Congress on the Proposed Head Start Program Designation Renewal System." Washington, DC: Administration for Children and Families, U.S. Department of Health and Human Services. http://eclkc.ohs.acf.hhs.gov/hslc/data/rc/Head_Start_Proposed_Designation_Renewal_System.pdf.

Office of Head Start. (n.d.(b)). "Use of Classroom Assessment Scoring System (CLASS) in Head Start." https://eclkc.ohs.acf.hhs.gov/hslc/hs/sr/class/use-of-class.pdf.

Office of Head Start. 2015a. "Grantees Required to Compete for Continued Funding." http://eclkc.ohs.acf.hhs.gov/hslc/grants/dr/cohort-4-listsby-state.html.

Office of Head Start. 2015b. "Head Start Program Facts Fiscal Year 2014." http://eclkc.ohs.acf.hhs.gov/hslc/data/factsheets/2014-hs-program-factsheet.html.

Office of Head Start. 2015c. "45 CFR 1301-1311. Head Start Program Performance Standards and Other Regulations.

Office of Head Start. 2016a. "About Us." http://eclkc.ohs.acf.hhs.gov/hslc/hs/about#about.

Office of Head Start. 2016b. "FY 2016 CLASS Field Guide." http://eclkc.ohs.acf.hhs.gov/hslc/grants/monitoring/docs/fy2016-CLASS.pdf.

Personal communication with Noreen Yazejian in January 2016 regarding the NC and IL QRIS Evaluations.

Pianta, R., C. Howes, M. Burchinal, D. Bryant, R. Clifford, D. Early, and O. Barbarin. 2005. "Features of pre-kindergarten programs, classrooms, and teachers: Do they predict observed classroom quality and child-teacher interactions?" *Applied Developmental Science* 9 (3): 144–159.

Pianta, R. C., K.M. La Paro, and B.K. Hamre. 2008. *Classroom Assessment Scoring System (CLASS) Manual, Pre-K*. Baltimore, MD: Paul H. Brookes Publishing Company.

Pianta, R. C., A.J. Mashburn, J.T. Downer, B.K. Hamre and L. Justice. 2008. "Effects of Web-Mediated Professional Development Resources on Teacher–Child Interactions in Pre-Kindergarten Classrooms." *Early Childhood Research Quarterly* 23 (4): 431–451.

Powell, D. R., and K. E. Diamond. 2011. "Improving the outcomes of coaching-based professional development interventions." Pp. 295-307 in *Handbook of Literacy Research: Vol. 3*, edited by S. B. Neuman and D. K. Dickinson. NY: Guilford.

Raver, C.C., C. Domitrovich, M. Greenberg, P.A. Morris and S. Mattera. 2012. *Adapted Teaching Style Rating Scale*. New York, NY: MDRC.

Robin, K. B., E.C. Frede and W.S. Barnett. 2006. *Is More Better? The Effects of Full-Day vs. Half-Day Preschool on Early School Achievement*. New Brunswick, NJ: NIEER. http://nieer.org/resources/research/IsMoreBetter.pdf.

Rohacek, M., G. Adams, E. Kisker, A. Danziger, T. Derrick-Mills, and H. Johnson. 2010. *Understanding Quality in Context: Child Care Centers, Communities, Markets, and Public Policy*. Washington, DC: The Urban Institute. Retrieved from: http://www.urban.org/research/publication/understanding-quality-context-child-care-centers-communities-markets-and-public-policy.

Sylva, K., I. Siraj-Blatchford, and B. Taggart. 2003. *Assessing Quality in the Early Years: Early Childhood Environment Rating Scale-Extension (ECERS-E): Four Curricular Subscales*. Stoke-on Trent, UK: Trentham Books.

Talan, T. N. and P.J. Bloom. 2011. *Program Administration Scale, Second Edition*. New York, NY: Teachers College Press.

Teachstone. 2016. *Research is in our DNA*. Downloaded April 27, 2016 from http://teachstone.com/resources/research/.

Tout, K., T. Halle, S. Daily, L. Albertson-Junkans, and S. Moodie. 2013. *The Research Base for a Birth Through Age Eight State Policy Framework*. Washington, DC: Alliance for Early Success and Child Trends.

Tuckman, H., and C. Chang. 1991. "Methodology for Measuring the Financial Vulnerability of Charitable Nonprofit Organization." *Nonprofit and Voluntary Sector Quarterly* 20 (4): 445–460.

Warner, M., and A. Hefetz. 2003. "Rural-urban differences in privatization: limits to the competitive state." *Environment and Planning* 21 (5): 703–718.

Wasik, B. A. 2010. "What Teachers Can do to Promote Preschoolers' Vocabulary Development: Strategies From an Effective Language and Literacy Professional Development Coaching Model." *Reading Teacher* 63 (8): 621–633.

Wasik, B. A., and A.H. Hindman. 2011. "Improving Vocabulary and Pre-Literacy Skills of At-Risk Preschoolers Through Teacher Professional Development." *Journal of Educational Psychology* 103 (2): 455–469.

About the Authors

Teresa Derrick-Mills is a senior research associate in the Center on Labor, Human Services, Population and the Center on Nonprofits and Philanthropy at the Urban Institute. Her research focuses on early care and education policy and practice, workforce development policy and practice, performance measurement and management, and the relationships between nonprofits and government. She has a Ph.D. in public policy and public administration with a focus in program evaluation from The George Washington University where she is an adjunct professor in the Trachtenberg School of Public Policy and Public Administration.

Margaret R. Burchinal is a senior scientist at the Frank Porter Graham Child Development Institute at the University of North Carolina at Chapel Hill and an adjunct professor in the Department of Education at the University of California, Irvine. She has been a leading researcher and statistician in child care research and a widely recognized applied statistician. She has authored or co-authored over 150 papers published in peer-reviewed journals, including *Child Development*, *Developmental Psychology*, and *Science*. She currently serves on the editorial boards for *Child Development* and *Early Childhood Research Quarterly*. She has a Ph.D. in quantitative psychology from the University of North Carolina at Chapel Hill.

H. Elizabeth Peters is the director of the Center on Labor, Human Services, and Population at the Urban Institute. Her research focuses on family economics and family policy, specifically addressing topics related to child and family well-being, father involvement, family formation and dissolution, child support policy, child care choice and child care policy, work-family balance, and family decision-making. Her work has appeared in peer-reviewed journals in economics, demography, and public policy, and she is the co-editor of *From Welfare to Childcare: What happens to Infants and Toddlers when Single Mothers Exchange Welfare for Work* (2006). She has a Ph.D. in economics from the University of Chicago.