Developing Measures of the Implementation and Cost of High-Quality Early Care and Education

iCHQ

Implementation and Costs of High Quality

EARLY CARE & EDUCATION

Developing Measures of the Implementation and Cost of High-Quality Early Care and Education

Submitted to:
Ivelisse Martinez-Beck, Meryl Barofsky, and Tracy Carter Clopet, *Project Officers*
Office of Planning, Research, and Evaluation
Administration for Children and Families
U.S. Department of Health and Human Services

Submitted by:
Gretchen Kirby
Pia Caronongan
Andrew Burwick
Shannon Monahan
Dmitriy Poznyak
Theresa Schulte
Julia Lyskawa
Annalee Kelly

Project Director:
Gretchen Kirby, Mathematica® Inc.
1100 1st Street, NE, 12th Floor
Washington, DC 20002-4221

Disclaimer:

Suggested citation:

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

Sign-up for the OPRE Newsletter

Follow OPRE on Twitter @OPRE_ACF

Like OPRE on Facebook facebook.com/OPRE.ACF

Follow OPRE on Instagram @opre_acf

Follow OPRE on LinkedIn linkedin.com/company/opreacf

OPRE

Mathematica®
Progress Together

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

Contents

Exhibits

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

Executive Summary

Evidence suggests that high quality early care and education (ECE) can benefit young children, particularly children in families with low incomes. This evidence has prompted increased state and federal investment in quality improvement initiatives and placed emphasis on helping families with low incomes gain access to high quality ECE. Policymakers, administrators, and program and center directors must weigh competing demands and limited resources for program improvement, but they lack information about the cost of high quality care and the best ways to use resources at the center level to meet expectations for quality.

A. Introduction

The Office of Planning, Research, and Evaluation (OPRE) in the Administration for Children and Families contracted with Mathematica to develop measures that can inform pathways to quality, through the Assessing the Implementation and Cost of High Quality Early Care and Education (ECE-ICHQ) project. The ICHQ (pronounced I-check) measures capture (1) ***implementation*** of activities that can support quality in center-based early care and education settings serving children from birth to age 5 (not yet in kindergarten) and (2) the ***cost*** of providing care in these settings. The ICHQ implementation and cost measures are framed around five key functions of center operations: (1) Structural Supports for Instruction and Caregiving; (2) Instructional Planning, Coordination, and Child Assessment; (3) Center Administration and Planning; (4) Workforce Development; and (5) Child and Family Support.

The measures can help us look at different pathways to quality and inform decisions about how centers can use resources effectively to deliver high quality ECE. The project has three components: (1) foundational work of a literature review and conceptual framework; (2) a phased

> The final **instrument** will include all the data collection tools, guidance for producing the measures from the collected data, and instructions for interpreting and using the measures.

multi-case study of centers to collect data for use in creating draft measures; and (3) a field test to validate the measures and finalize the instrument. This report focuses on the results from the multi-case study, describing the development of the draft measures of implementation and cost and how they perform.

B. Purpose and primary research questions

The goals of ICHQ are to (1) produce technically sound, systematic measures of the implementation and costs of education and care in center-based settings that serve children from birth to age 5; (2) produce implementation and cost measures to examine the variation in ECE center capacities and resources that can make a difference in quality and the experiences of children; and (3) develop a feasible and useful instrument to guide the collection, development, and reporting of the measures. The measures are intended for use in research to examine and explain variations in quality, in practice to inform quality improvement, and in policy to inform funding and technical assistance.

Exhibit ES.1. ECE quality and ECE-ICHQ measures
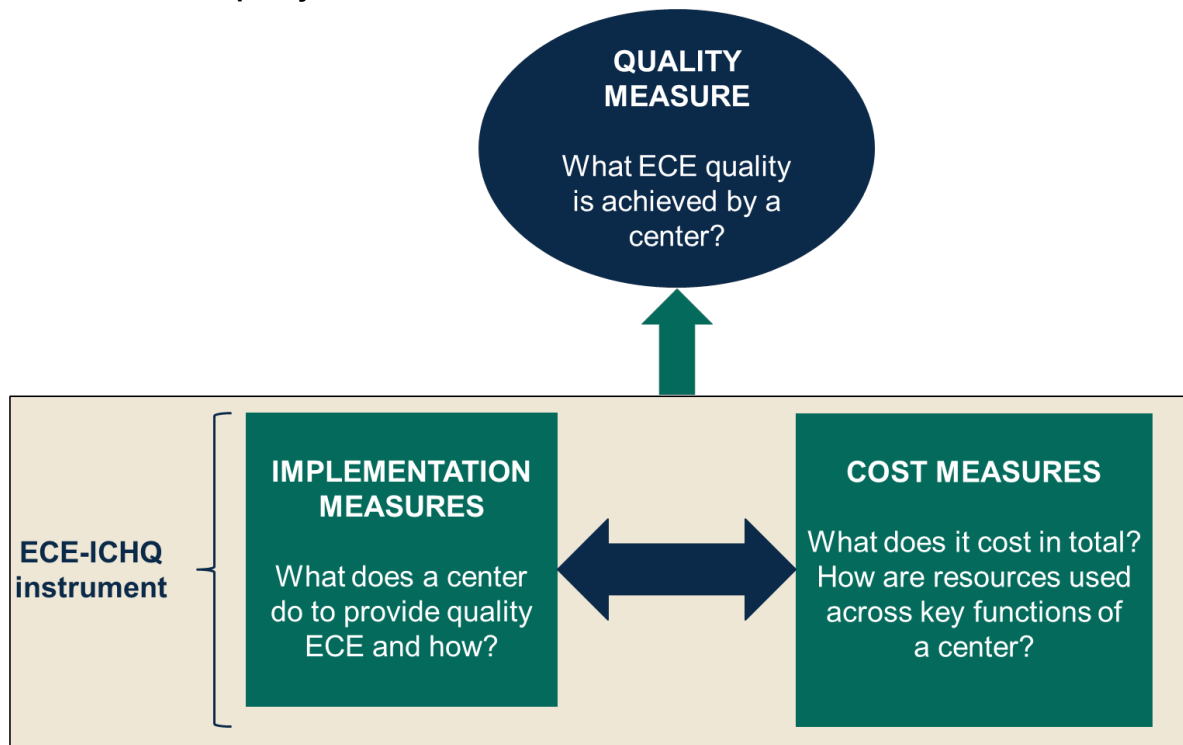
This report describes the approach to the multi-case study, details the development of the implementation and cost measures, and presents next steps for testing and using the measures. Five research questions guided the work of the multi-case study:

1. What implementation activities support each key function and can we reliably measure implementation of each function?

2. How do staff members use their time in support of key functions within the center?

3. What are the costs associated with the implementation of key functions?

4. Are differences in center characteristics related to implementation and costs?

5. Can implementation and cost measures be aligned to inform decisions to guide quality improvement?

C. Key findings and highlights

Early testing with a purposive sample of 30 centers suggests that the ICHQ measures are capturing important information about implementation and costs and the variation that exists across centers. The measures are doing what they are intended to, and they hold promise for working together to inform pathways to quality.

Useful and relevant measures for interpreting implementation—what a center does to support quality and how —need to comprise a small set of items that can be used consistently across a range of center-based ECE settings. Our aim was to create measures with good psychometric properties that would be ready for more formal testing as part of a field test.

- **The implementation scores for each of the five key center functions produced reliable scales.** The items for each function measure the same thing in different ways and together they capture the essence of the function (they hold together). The measures appear to be dependable for assessing implementation of each key function.

- **Implementation scores for each key function have distributions that can detect differences across centers.** The scores for each key function have substantial variation; we are able to tell centers apart in the middle of the distribution as well as along the tails.

- **Implementation scores for the key functions are related to each other, but each also provides distinct information.** We found moderate, significant associations among the implementation scores of the five key functions. The scores for Center Administration and Planning are positively and significantly correlated with the scores of the other four functions, suggesting the central nature of this function in driving implementation.

> **Psychometric properties** show how reliable and valid a measure is based on the purposes for which it is designed and used.
>
> **Reliability.** Indicators of reliability show how dependable a measure is for the purpose for which it is used. Reliable measures are stable over time and include items that measure the same thing in different ways.
>
> **Validity.** The degree to which an assessment accurately measures what it is designed to measure. In the early stages of measures development, validity can be examined based on patterns with other observable variables or characteristics.

The measures demonstrate validity based on patterns with observable characteristics and what we know from previous cost studies.

- **On average, center staff reported spending their time in expected ways.** Directors reported spending more of their time on Center Administration and Planning than other staff. Lead teachers, head teachers, or co-teachers reported spending more of their time on Instructional Planning, Coordination, and Child Assessment relative to assistant teachers, who reported spending more of their time on Instruction and Caregiving.

- **The cost measures show that the costs of staff compensation and facilities made up nearly 80 percent of total costs, on average.** Centers invest in staff compensation (salary and fringe benefits) as the largest proportion of costs (62 percent, on average). Facilities, and supplies and materials account for the next largest proportions of costs (16 percent and 9 percent, respectively, on average). This allocation is in line with past cost studies in ECE.

- **Centers spent the largest portion of costs on the Instruction and Caregiving function.** Centers in the sample directed 33 percent of costs to the Instruction and Caregiving function, on average, and 22 percent of costs to the Instructional Planning, Coordination, and Child Assessment function. The centers spent just over one-quarter of their total costs (26 percent, on average) on the Center Administration and Planning function. Spending on the Workforce Development function accounted for 8 percent of total costs, on average, and spending on the Child and Family Support function accounted for 10 percent of total costs.

- **Implementation and cost measures vary with center characteristics in ways we would expect.** In our sample, centers with high QRIS ratings have higher implementation scores and costs per child care hour, on average, on each key function than centers with low QRIS ratings. We also saw in the sample, for example, that the average cost per child care hour was lower for centers that served school-age children in addition to children ages 0 to 5 than for centers that only served children ages 0 to 5. We could not conduct statistical tests for significance between the mean scores of centers with different characteristics, given the small number of centers within each characteristic subgroup.

The measures appear to work together in ways that can identify and inform relationships with quality.

- **The implementation and cost measures are related.** We found positive associations between four of the five implementation scores and total cost per child care hour. The implementation scores and cost per child care hour for each key function are also positively related, meaning the higher the implementation score, the higher the cost of the function. These relationships are expected and important in validating the measures.

- **There is a good deal of variation in both implementation and costs at the center level.** There is more to the story beyond observing that implementation and higher costs are generally higher in centers with high QRIS levels. Scatterplots of the measures across the functions show that QRIS levels do not completely predict center implementation scores and costs. For example, we observe some low quality level centers with above median implementation function scores. It was rare, however, to observe above median costs in centers with low quality levels.

D. Methods

At the start of the project, the study team conducted a literature review to inform a conceptual framework, guide data collection in the multi-case study, and support the development of measures (see Caronongan et al. 2016).

A phased approach to data collection in the multi-case study provided opportunities to refine the measurement constructs and data collection tools and processes, and to develop draft measures. The multi-case study included (1) conducting semistructured interviews about the intentionality, structure, and consistency of *implementation* of key functions of a center that can support quality; (2) collecting *cost* data through Excel workbooks to assess center-level costs by key function and how resources are used across functions; and (3) administering surveys about *staff time use*—essential information for allocating costs to key functions, since labor is a large driver of costs in ECE centers (Exhibit ES.2).

Exhibit ES.2. Overview of the ECE-ICHQ multi-case study

- Convenience sample of 3 centers
- On-site data collection

- Purposive sample of 15 centers in 3 states
- On-site data collection

- Purposive sample of 30 centers in 3 states
- Remote data collection

Data collection methods across phases

| Short staff survey of time use | Semistructured interview about implementation | Excel workbook for cost reporting |
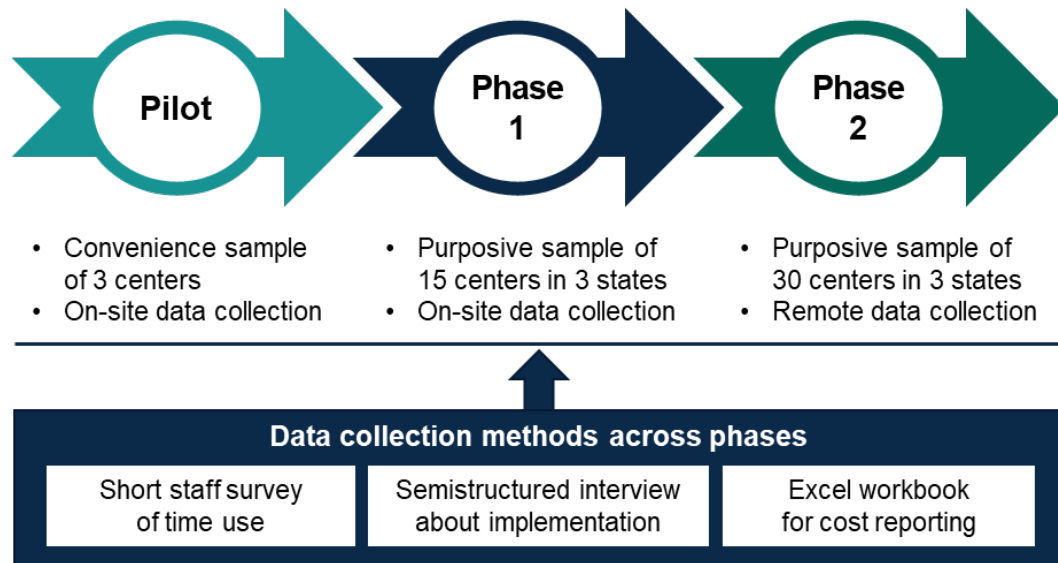
The length of the data collection tools, the methods of collection, and the number of participating centers differed across the phases. The multi-case study began with an ***exploratory*** pilot study with three centers in the fall of 2015 to test the data collection approach and make an initial assessment about the relevance, appropriateness, and connection of the topics covered in each data collection tool. We then recruited 15 centers to participate in a comprehensive ***formative*** phase of data collection in the fall of 2016 to test the feasibility of data collection tools and methods, conduct cognitive interviewing to obtain feedback from the respondents about the topics and methods of collecting the information (such as relevance, clarity, and ease), and refine the tools. Based on what we found in Phase 1, we developed a set of measures to test in Phase 2. Phase 2 consisted of ***early testing*** of the new measures to ensure that they are technically sound (meaning they are reliable and the measures of implementation and cost are related) and can be used consistently across a range of ECE center-based settings.

We structured data collection in Phase 2 around the functions in the conceptual framework and the functions connected to the implementation and cost data collection (Exhibit ES.3).

Exhibit ES.3. Implementation and cost data collection to understand key functions of a center

E.   Implications for next steps

The findings from the phased multi-case study provide preliminary evidence that the measures are capturing variations in implementation and costs among centers and are working together in ways that have the potential to inform pathways to quality. The measures are not yet fully validated, meaning they have not been tested in a large, representative sample of centers to look at associations with center quality or children's outcomes.

We are pursuing a step in this direction through a field test with a purposive sample of 80 centers in 4 states. Using data collected in the field test, we will create the implementation scores and cost measures, assess their functionality and validity, and examine associations among implementation, cost, and quality. Specifically, we will: (1) examine whether the cost and implementation measures explain variation in quality beyond what is explained by center characteristics; (2) examine the degree to which each set of measures (cost and implementation) is related to quality, independent of the other; and (3) test whether the relationship between cost and quality varies by levels of implementation (that is, whether implementation moderates the relationship between cost and quality).

The ICHQ measures have the potential to support the efforts of researchers, practitioners, and policymakers to better understand pathways to achieving high quality. These insights can inform decisions about the level of resources needed and how to best use resources across functions at the center level to deliver high quality early care and education. Examples of potential uses of the measures include:

- Practitioners could use the measures to examine implementation and costs of quality within a single center to understand and guide quality improvement or needs for technical assistance.

- Federal and state administrators could use the measures to examine a group of centers within a state or across states to inform decisions about quality improvement initiatives or subsidy rates.

- Federal policymakers could use data collected with the ECE-ICHQ instrument to inform cost modeling for efforts to expand access to high quality early care and education.

- State leaders might use measures of the cost of high quality care to inform subsidy payments and policymaking that promotes cost-efficient paths to enhance program quality. If observations vary by geographic location or center characteristics, policymakers and administrators might target funding and quality improvement efforts appropriately.

- Researchers could use the measures to compare cost and implementation patterns among different ECE program models, explore pathways to quality, or make them part of a cost-benefit analysis.

- More generally, the ICHQ measures could contribute to equity considerations by measuring the services centers provide and the cost of services by the characteristics of the children and families served.

Glossary/key terms

This report and the ECE-ICHQ project focus on implementation, cost, and quality measures within center-based ECE settings. A **center** refers to a specific physical location where ECE classroom-based services are provided to children ages 0 to 5.

Implementation measures assess what a center does and how it does it. These elements include the package of services a center provides, including structural features and adopted practices.

Cost measures assess the resources that provide the package of services and resource use across functions.

High quality center-based ECE refers to centers that achieve positive child outcomes through positive teacher-child interactions and a positive learning environment.

Key functions are the tasks in which all center-based ECE providers engage. The key functions drive how a center operates to achieve quality. Each of the five key functions has a set of activities and practices that allow us to measure implementation and costs for each function distinctly.

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

I. Introduction

Evidence suggests that high quality early care and education (ECE) can benefit young children, particularly children whose families have low incomes (Yoshikawa et al. 2013). This relationship has led to more investment in quality improvement initiatives and greater emphasis on getting low-income children into high quality ECE, through such federal or state partnerships as the Race to the Top Early Learning Challenge and Preschool Development Grants. Policymakers, administrators, and program and center directors must weigh competing demands and limited resources for program improvement, but they lack information about the cost of high quality care and the best ways to allocate resources at the center level to align with expectations for quality.

The Office of Planning, Research, and Evaluation (OPRE) in the Administration for Children and Families (ACF) at the U.S. Department of Health and Human Services (DHHS) is supporting the Assessing the Implementation and Cost of High Quality ECE (ECE-ICHQ) project to develop measures to inform research, policy, and practice. The goal of the project is to improve understanding of variations in what centers do to support quality, the associated costs, and how resources for ECE may be better aligned with expectations for quality. OPRE contracted with Mathematica and its consultant Elizabeth Davis, of the University of Minnesota, to conduct the project.

The goal of ECE-ICHQ (pronounced "I-check") is to produce center-level measures of cost and implementation that can be used to examine and explain variations in quality (Exhibit I.1). Implementation measures capture *what* a center does to support quality, including structural features (for example, teacher–child ratios, group size, and staff qualifications) and adopted practices (for example, curriculum use and child assessment).

Exhibit I.1. ECE quality and ECE-ICHQ measures
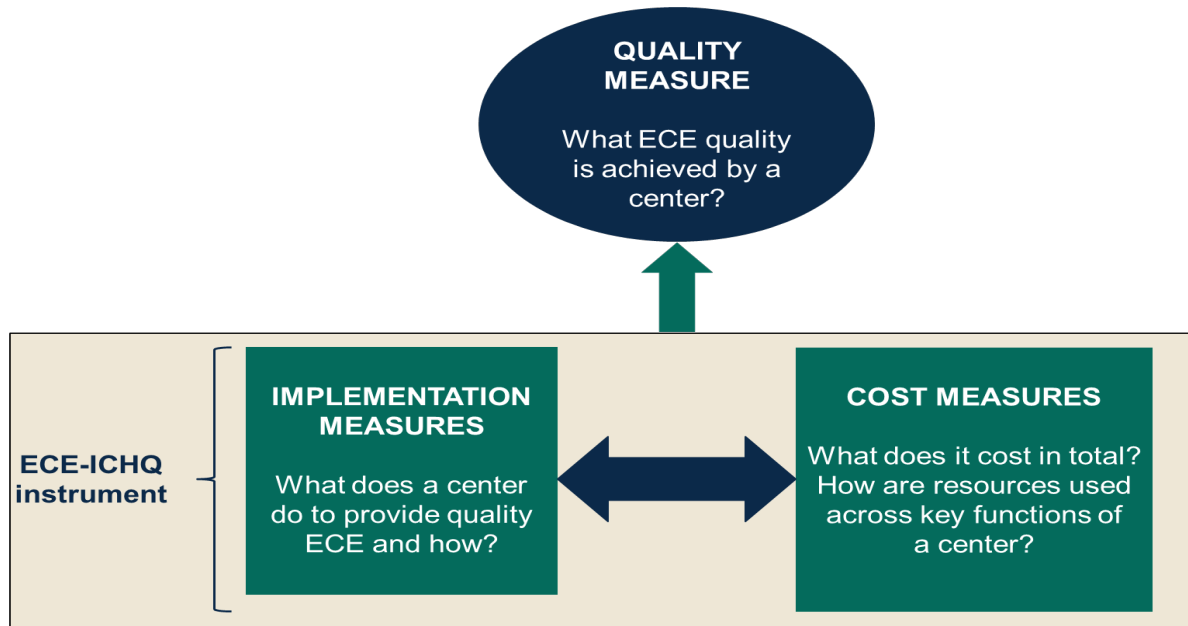
Implementation measures also reflect *how* structural features and adopted practices are supported and implemented within a center. Cost measures capture the total resources needed to support the ECE program (or package of services) a center provides. These resources include costs the center needs to support all operations and programs, as well as the value of in-kind contributions such as space or equipment. The measures also capture how resources are used across different functions within a center.

Specifically, the work of ECE-ICHQ will (1) produce technically sound, systematic measures of the implementation and costs of education and care in center-based settings that serve children from birth to age 5; (2) produce implementation and cost measures to examine the variation in ECE center capacities and resources that can make a difference in the experiences of children; and (3) develop a feasible and useful instrument to guide the collection, development, and reporting of the measures for a broad range of users.

> **Terms associated with providing ECE services**
>
> **ECE center:** A specific physical location where early care and education classroom-based services are provided to children ages 0 to 5
>
> **ECE provider:** An organization or agency that delivers ECE services
>
> **ECE program:** The package of services provided at a particular ECE center
>
> **Center function:** Activities in support of a specific purpose (detailed in Chapter II)

> ECE-ICHQ will develop a set of **measures** for key functions of a center that include the following:
>
> - **Implementation measures** of what a center does (the package of services a center provides, including structural features and adopted practices) and how it does it
>
> - **Cost measures** of resources used to provide the package of services and resource use across functions
>
> **Data collection tools** collect the information from center staff that is needed to build the measures.
>
> The final **instrument** will include all the data collection tools, guidance for producing the measures from the collected data, and instructions for interpreting and using the measures.

The measures are intended to be broadly useful for a range of stakeholders, such as the following:

- *Practitioners* could use the measures to examine implementation and costs of quality within a single center to understand and guide quality improvement or needs for technical assistance.

- *Federal and state administrators* could use the measures to examine a group of centers within a state or across states to inform decisions about quality improvement initiatives or subsidy rates.

- *Researchers* could use the measures to study a large sample of programs to describe implementation and costs systematically or explore pathways to quality, or as part of a cost-benefit analysis.

A.   What ECE-ICHQ measures

Research on the association between quality and costs in ECE center-based programs is limited, but there is some evidence of a relationship. Studies of the association between quality and costs in ECE center-based programs have consistently found that higher quality centers tend to have higher costs (Helburn 1995; Blau and Mocan 2002; Marshall et al. 2001; Marshall et al. 2004a; Marshall et al. 2004b; Belfield and Schwartz 2007; Levin and Schwartz 2007). However, the magnitude of the relationship between cost

and quality ranges considerably across studies. In addition, center costs can also vary based on enrollment level, hours of operation, staffing structure, and the ages of children served.

Because most studies have focused on measures of the total program cost, they do not capture important information about how centers use resources in ways that could affect quality. Finer-grained measures of costs that map to what a center does to support quality would help us better understand the relationship between costs and quality. Centers may be able to achieve quality through different pathways that relate to differences in how center features and practices are combined, implemented, and supported.

Current measures of ECE quality largely focus on the classroom, but support for quality at the center level can also contribute to the level of quality ECE centers can achieve. The ECE-ICHQ measures use an implementation science lens to assess differences in context, capacities, and implementation of quality features at the classroom and center level. These differences may be reflected in total costs and how resources are used within a center and, ultimately, in the level of quality a center achieves.

> "**Implementation science studies programs and practices that achieve their intended outcomes to identify programmatic and practitioner factors that are important for successful implementation.**"
> *(Berkel et al. 2011; Fixsen et al. 2005)*

The implementation measures developed in ECE-ICHQ provide a way to more closely examine supports for quality and test the assumption that higher quality always costs more. For example, if Center 1 and Center 2 have the same levels of quality but Center 2 has lower costs, the ECE-ICHQ measures will help identify what Center 2 is doing differently to achieve quality. The difference between the two centers may lie in implementation. That is, what one center does to support quality, and how, may result in a different allocation of resources but not necessarily at a higher cost (Exhibit I.2).

Exhibit I.2. ECE-ICHQ measurement scenario: Quality is the same, total costs differ

Higher levels of implementation may also help a center achieve higher quality without facing higher costs, as in Exhibit I.3. Higher levels of implementation might mean doing more things that can support quality such as supporting coaching for teachers beyond stand alone professional development or having a strategic plan that involves leadership, teaching staff, and parents compared to no strategic plan for the center. The ECE-ICHQ measures may again show that the difference stems from what Center 2 does

differently to support quality—through the features and practices it has in place and how they are supported and implemented, as well as how it allocates its resources.

Exhibit I.3. ECE-ICHQ measurement scenario: Total costs are the same, quality differs

B. Project overview

The ECE-ICHQ project has three main components: (1) foundational work of a literature review and conceptual framework; (2) a phased multi-case study of centers to collect data for use in developing draft measures and creating the building blocks for the ECE-ICHQ instrument; and (3) a field test to validate the measures and finalize the instrument and user's manual. The project is organized around seven key tasks, as shown in Exhibit I.4. Throughout the project, the team consulted with a range of experts. We had an initial 10-member technical expert panel of research and practice experts in early childhood who advised on the conceptualization of the work and developing the initial conceptual framework. We sought review and input on specific data collection and analysis topics from individuals with certain expertise within and beyond our initial expert panel.

At the start of the project, the study team conducted a literature review to inform a conceptual framework, guide data collection in the multi-case study, and support the development of measures (see Caronongan et al. 2016). A phased approach to data collection in the multi-case study provided opportunities to refine the measurement constructs and data collection tools and processes, and to develop draft measures. The multi-case study began with an exploratory pilot study with three centers in the fall of 2015. We then recruited 15 centers to participate in the first phase of data collection for the multi-case study and conducted data collection with these centers in fall of 2016. We refined the data collection tools and plans for building the measures based on experiences and findings from Phase 1. In the fall of 2017 and spring of 2018, we collected data from 30 centers that participated in Phase 2 of the multi-case study and analyzed the data to create draft measures of implementation and cost. In this report, we focus on the development of the measures through the multi-case study.

Exhibit I.4. Key tasks of the ECE-ICHQ project
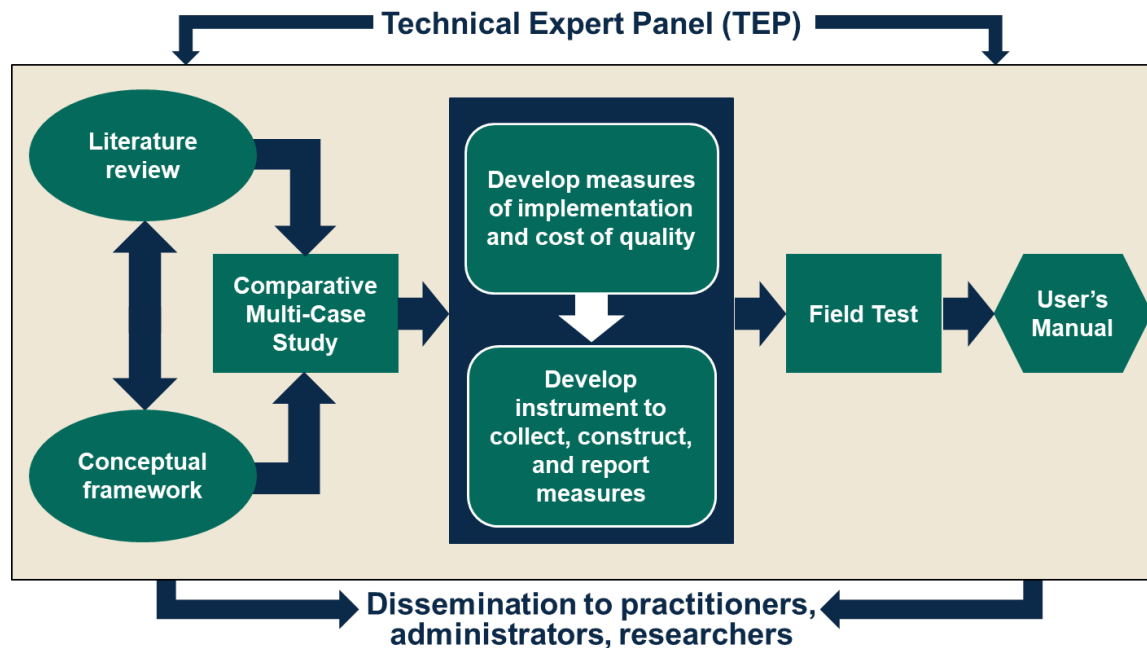
In the last component of the work, we will conduct a field test to assess consistency and reliability in measurement across constructs and users. We will validate the measures with measures of quality or other existing, validated measures that assess constructs similar to the ECE-ICHQ implementation measures. The end product of the ECE-ICHQ project is a set of final measures and a complete instrument with a user's manual for collection, analysis, and application of the full set of measures.

C.   Road map of the report

In this report, we describe the approach to the multi-case study, the development of the implementation and cost measures, and the next steps for testing and use of the measures. In Chapter II, we introduce and describe the ECE-ICHQ conceptual framework that has guided data collection and measurement. In Chapter III, we provide an overview of the methodology of the multi-case study, including the purpose of each phase, and the criteria for the selection of states and centers. In Chapters IV and V, we detail the development of the implementation and cost measures, respectively. In Chapter VI, we present findings about the associations between the implementation and cost measures. We conclude, in Chapter VII, with implications for next steps in development, testing, and use of the ECE-ICHQ measures. We include more detailed technical information in a series of appendices, referenced in the related chapters.

II. ECE-ICHQ Conceptual Framework

A conceptual framework is essential to guide the measures development work of ECE-ICHQ, by identifying what to measure based on hypothesized pathways to quality and improved outcomes for children. We developed a conceptual framework collaboratively with experts and OPRE using the research base and revised it iteratively in response to findings from the multi-case study. At the start of the project, the project team, together with OPRE and a 10-member technical expert panel, developed a draft conceptual framework that was informed by the ECE-ICHQ literature review (Caronongan et al. 2016). We used the framework to guide data collection and measures development in the multi-case study. With OPRE and select experts, we revised the draft conceptual framework based on findings from the multi-case study. In this chapter, we present the current conceptual framework, describe its elements, and highlight key revisions we made over time.

A. Overview of the conceptual framework

The ECE-ICHQ conceptual framework shows that centers carry out a set of key functions in pursuit of high quality care and better outcomes for children (Exhibit II.1). The key functions in the framework occur in the variety of center-based ECE settings (including community-based centers and Head Start programs) that provide services to young children and their families. The key functions are mutually exclusive but complementary and are all necessary in delivering high quality ECE. The core idea behind the ECE-ICHQ project is that centers vary in their investment in and capacity to implement the key functions in ways that promote quality. We seek to develop measures that capture this variance and deepen our understanding of how it relates to quality.

Exhibit II.1. Conceptual framework guiding measures development for the ECE-ICHQ project
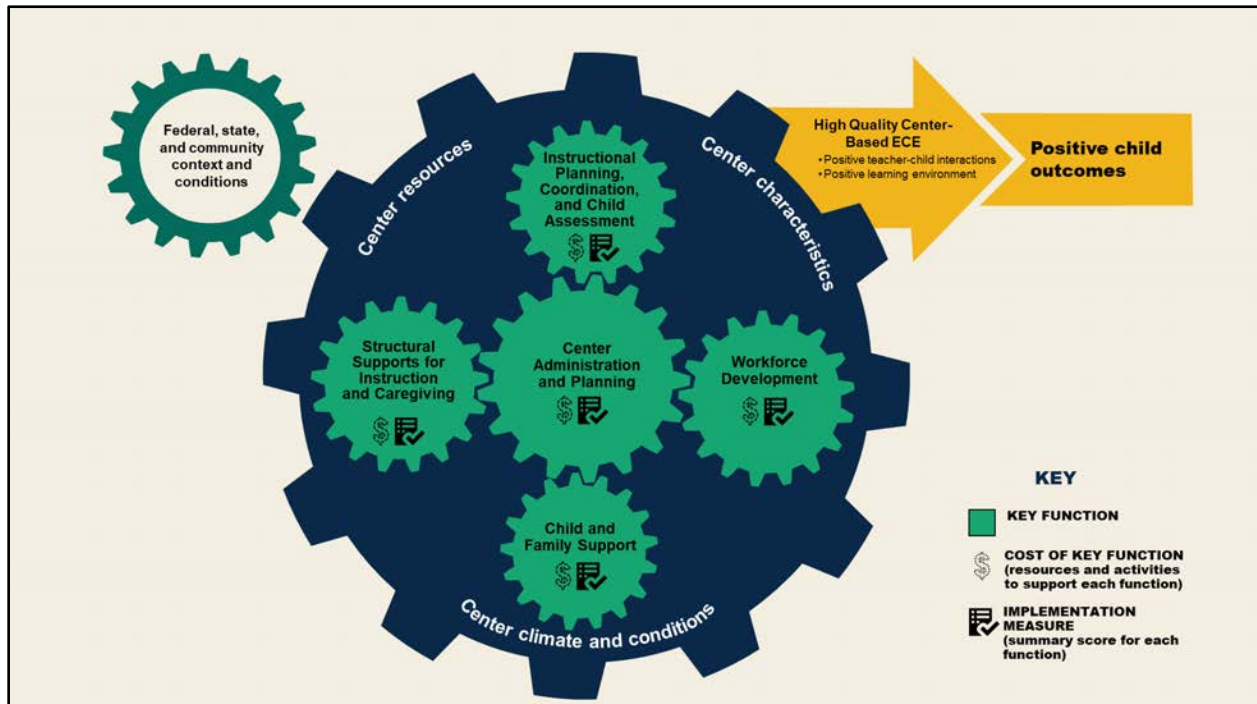
Findings from the literature review and discussions with the technical expert panel suggested that a center with particular features and practices in place, and ways to support them, has greater potential to provide high quality ECE. In this framework, we define high quality center-based ECE in line with common observational measures of quality that emphasize positive teacher–child interactions and a positive learning environment (such as the Classroom Assessment Scoring System [CLASS; Pianta et al. 2006] and the Early Childhood Environment Rating Scale [ECERS-R; Harms et al. 1998]). The expected result of children's high quality ECE experiences is improved child outcomes. That is, high quality early care and education should prepare all children with the social, emotional, and cognitive skills to enter kindergarten ready to learn, helping to close the achievement gap for disadvantaged children in low-income households.

We present the main elements of the framework as gears because they drive how a center operates to achieve quality. They also interact with and influence each other. The small, green gears within the main gear represent the key functions that we expect to find across all center-based ECE providers, although the relative role and emphasis of each function will vary across centers. Centers operate within the federal, state, and community contexts and conditions that set requirements, standards, and policies (such as through licensing, Quality Rating and Improvement Systems [QRIS], and Head Start, state pre-K, and child care subsidy programs); professional development and quality improvement supports; and wage rates, real estate costs, and consumer prices.

B.   The conceptual framework as a guide for data collection and measures development

We set out to develop measures of implementation and cost for each of the key functions. Implementation refers to the features and practices that are in place to support the function, and how they are carried out. Cost refers to the value of the resources that are devoted to supporting each function. To achieve this, we needed to develop items to measure each of the functions. We also needed to understand other elements— including center characteristics and resources, and the context in which a center operates—that can affect what centers do and how they use their resources.

Key functions. Key functions reflect tasks that all center-based ECE providers engage in and for which there are clear means of measuring implementation and costs. The five key functions include (1) Structural Supports for Instruction and Caregiving (referred to throughout the report as Instruction and Caregiving); (2) Instructional Planning, Coordination, and Child Assessment; (3) Center Administration and Planning; (4) Workforce Development; and (5) Child and Family Support (described in more detail in Exhibit II.2). To identify key functions, we compared cost and quality categories defined within several sources reviewed in the literature review, including the QRIS Compendium (Tout et al. 2010) and the Program Administration Scale (Talan and Bloom 2004).

Exhibit II.2. Descriptions of key functions

Key function	Types of activities
Structural Supports for Instruction and Caregiving	• Classroom ratios and group sizes • Materials and equipment • Staff qualifications
Instructional Planning, Coordination, and Child Assessment	• Planning time for curriculum, activities, and lessons • Conducting screening and child assessments • Communicating with families about children's learning
Center Administration and Planning	• Planning and goal setting • Regulatory compliance and reporting • Managing center finances • Managing and maintaining center operations
Workforce Development	• Recruiting and hiring teaching staff • Supervising staff • Conducting staff meetings • Evaluating staff performance • Planning and providing training and professional development activities • Using classroom observations to support teaching staff and improve practice • Providing (and receiving) coaching, mentoring, or technical assistance
Child and Family Support	• Providing services or access to services that support children and families beyond instruction and caregiving (such as social services, therapeutic services, health screenings, and counseling, among others)

Center characteristics and resources. Center characteristics and resources are important elements for measurement because they contribute to decisions that affect the implementation and cost of the key functions. Center characteristics include the ages of children served, profit status, whether a center is independent or embedded in a larger organization or chain, funding mix (the level and combination of public and private funding a center receives), and size (number of children served), among others. Center resources mean funding sources (the actual sources of funding), donated space, labor, and materials that can support what the center is able to do (and how).

Context and conditions. Centers operate within federal, state, and community contexts and conditions. We need to capture information about context and conditions for two purposes. First, we need to account for some conditions (such as differences in the costs of doing business as reflected in wages and facilities costs) explicitly in the cost measures. Second, some contextual information is important for interpreting patterns in the measures and ensuring applicability in a variety of contexts. For example, licensing requirements that set a floor for quality differ across states. In addition, the extent of professional development or quality improvement supports that exist within a state or locality can influence what a center is able to do to support quality. Also, the standards to which a center is held accountable (based on funding sources or program participation)—such as Head Start Program Performance Standards, QRIS standards, or standards for public pre-K—can also influence the ways it pursues quality.

C.   Revisions to the conceptual framework to reflect the ECE-ICHQ measures

We started with six key functions. Center administration and planning were originally broken out into (1) center administration and (2) center leadership, planning, and evaluation activities. Based on findings from the multi-case study, we combined these into one function.[1] We also altered the placement of the functions from what we had in the original version of the framework. In the original version, we treated the functions as equal, all shown in the same circle. Using data from the multi-case study, we found correlations among the implementation measures and between the implementation and cost measures that suggest the central role of the center administration and planning function. This function is placed in the center of the other functions to represent the associations we found. (We detail these findings in Chapters IV and V.)

The draft framework presented center resources and characteristics as outer gears along with context and conditions. We moved these elements into the main gear as they are explicit elements for measurement that influence the implementation and costs of center functions and how they are comprised in each center. In doing so, the central blue gear presents the information that is essential for developing and using the ECE-ICHQ measures to explore pathways to quality in center-based ECE settings.

---

[1] We tested multiple iterations of implementation items within and across these functions and found that the correlations among the items and the reliability of the scale were optimized when the items were combined. In addition, the costs for center leadership, planning, and evaluation activities were one percent or less across centers and could not differentiate centers. We hypothesize that costs associated with center leadership, planning, and evaluation activities are embedded within those that support the overall administration of the center.

III. Methodology of the Multi-Case Study

To inform measures development, we carried out a multi-case study that included (1) a qualitative study of the implementation of key functions of center-based ECE providers and (2) an analysis of center-level costs informed by how staff use their time. In this chapter, we describe the approach to the multi-case study and the goals of its different phases. We then discuss, in turn, the development, refinement, and testing of the data collection tools and methods. In discussing how we tested the tools and methods, we provide information about the selection and characteristics of states and centers. State context and center characteristics provide information useful in interpreting the draft measures that are the focus of the later chapters of this report.

A. Approach to the multi-case study

The multi-case study was designed to answer empirical questions about implementation and costs in ECE centers, as well as questions about data collection and measurement approaches that are feasible and relevant for a range of center-based ECE settings (Exhibit III.1).

Exhibit III.1. Research questions for the ECE-ICHQ multi-case study

Questions focused on ECE centers:
Are differences in center characteristics, contexts, and conditions related to implementation and costs?
What key functions do center-based ECE providers pursue, and what implementation activities support each function?
What are the costs associated with the implementation of key functions?
How do staff members use their time in support of key functions within the center?
Questions focused on measures development:
How can time-use data be collected efficiently from selected staff and analyzed to allocate labor costs into distinct cost categories by key function?
What data collection approaches can support an efficient and feasible instrument for broad use?
Questions focused on the purpose and relevance of the measures for policy and practice:
How can implementation and cost data be aligned to produce relevant and useful evidence to inform decisions about implementation activities and key functions likely to lead to quality improvement?

The multi-case study was structured in phases (Exhibit III.2) to provide the study team with the opportunity to refine, over time, the data collection processes and tools needed to build the measures. In each phase of the multi-case study, we relied on three data collection tools: (1) a staff survey of time use, (2) a semistructured interview about implementation, and (3) an Excel workbook for cost reporting. The length of the tools and the methods of collection (for example, on-site or remote) evolved over the phases. The number of centers increased with each phase as we refined the data collection tools and processes and decreased the burden of data collection on center staff.

Exhibit III.2. Overview of the ECE-ICHQ multi-case study
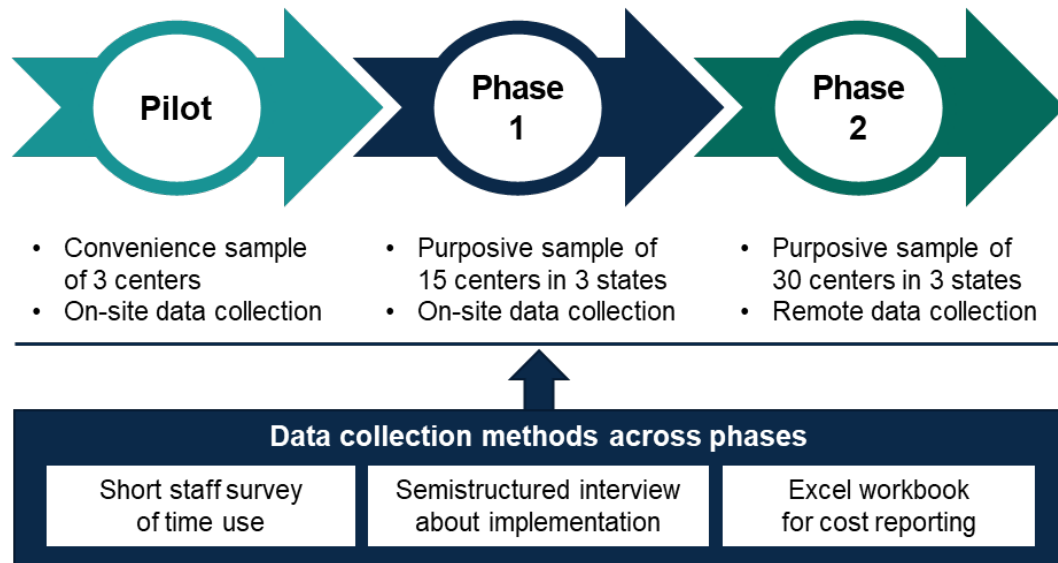
The goals for the phases of the multi-case study are aligned with the steps in the measures development process. Data collection began in an ***exploratory*** stage with site visits to three pilot centers to test the data collection approach and make an initial assessment about the relevance, appropriateness, and connection of the topics covered in each data collection tool (Exhibit III.3). Phase 1 was a comprehensive ***formative*** data collection effort to test the feasibility of data collection tools and methods, conduct cognitive interviewing to obtain feedback from the respondents about the topics and methods of collecting the information (such as relevance, clarity, and ease), and refine the tools for later phases. From the formative work in Phase 1, we (1) identified a range of approaches and intensities in what centers do and how, (2) developed a preliminary understanding of costs and allocations across functions, and (3) determined that we could measure costs and implementation around the functions of a center.

> **Formative** approaches are used for learning in an early part of development of a process, program, or measure. *Formative*, as an adjective, means giving shape to or molding something.

Based on what we found in Phase 1, we developed a set of measures to test in Phase 2. Phase 2 consisted of ***early testing*** of the new measures to ensure that they are technically sound (meaning they are reliable and the measures of implementation and cost are related) and can be used consistently across a range of ECE center-based settings. Phase 2 created a bridge between the formative work in Phase 1 and a planned field test, which will allow for more formal vetting and testing of draft measures and development of a final instrument.

Exhibit III.3. Goals for the phases of the ECE-ICHQ multi-case study to guide measures development

Pilot Exploratory (3 centers)	Phase 1 Formative (15 centers)	Phase 2 Early testing (30 centers)
• Explore the feasibility of the data collection approach through on-site visits • Assess the relevance and appropriateness of the topics covered in the data collection tools	• Test the feasibility of data collection tools and processes • Identify the range of implementation activities, costs, and key functions that occur in an ECE center to support quality • Explore options for constructing measures of implementation and costs that can work together around key functions	• Test the usability and efficiency of refined data collection tools across a range of early care and education centers • Explore whether items are working similarly (or varying in ways we would expect) across different settings • Assess whether implementation, cost, and time use are related in ways that we expect (or in ways that can be explained by the data) in preparation for further testing in a field test

B.   Developing data collection tools and methods

We developed the data collection tools through an iterative process with each phase of the multi-case study. Our goal was to select, adapt, or create items that could be measured in an objective way to achieve the following:

1. Provide information about the intentionality, structure, and consistency of ***implementation*** of key functions of a center that can support quality

2. Assess the ***costs*** by key function and how resources are allocated across functions

3. Gather information about staff ***time use***—an essential step in allocating costs to key functions, since labor is a large driver of costs in ECE centers

We drafted a preliminary set of data collection tools for the pilot and Phase 1 based on the existing knowledge base about implementation and costs in ECE using the literature review, discussions with experts, and a review of select measures and protocols. We refined data collection tools and methods over the course of the multi-case study to reduce burden, increase efficiency, improve data quality, and increase accuracy in measurement. From the outset, the plan was to start with broad, on-site data collection and transition to remote, more focused methods of data collection over the course of the multi-case study (Exhibit III.4). The type of respondents for different types of data collection remained the same throughout the multi-case study:

• The center director or person most knowledgeable about the early childhood program was the respondent for the ***implementation interview***.

• The person most knowledgeable about center finances was the respondent for the ***cost workbook***.

• All paid teaching staff of children ages 0 to 5 (including lead teachers, head teachers, or co-teachers, as well as assistants) and select administrators (center director, education program lead, and other administrators or managers involved in center oversight) were the respondents for the ***time-use survey***.

Exhibit III.4. Methods and respondents by data collection phase

Implementation Respondents: Center directors, person most knowledgeable about the early childhood program, or both				Cost Respondents: Person most knowledgeable about center finances			Time use Respondents: All teaching staff and select administrators	
Semi-structured interviews	Electronic or paper self-administered questionnaires	On-site data collection	Off-site data collection	Paper cost workbooks	Electronic cost workbooks	Phone discussions and email follow-ups	Paper survey[a]	Web-based survey[a]
Pilot phase: 3 centers								
X	X	X		X	X	X	$\bar{x} = 3$	
Phase 1: 15 centers								
X		X		X	X	X	$\bar{x} = 7$	
Phase 2: 30 centers								
X			X		X	X		$\bar{x} = 14$

[a] Average number of completed surveys per center.

1.   Implementation data collection tool

We developed a ***semistructured implementation interview*** guide to collect information about activities identified in the implementation science literature as important for effective implementation (summarized in Caronongan et al. 2016). We examined eight frameworks that convey the core principles and factors common to implementation of evidence-based programs and practices, based on literature reviews of empirical studies and theory development across a variety of service areas. Each framework presented an approach to measuring and evaluating the implementation of evidence-based programs. Through reviewing these implementation science frameworks, we identified five core activities that are common among successfully implemented programs:

1.   Recruitment, hiring, and selection of staff with the required skills and competencies

2.   Selection and use of tools that clearly convey key concepts, principles, procedures, and practices

3.   Training that delivers content knowledge to staff

4.   Technical assistance (TA) or coaching that includes observation and feedback

5.   A quality assurance (QA) and quality improvement (QI) process

The initial implementation interview guide was structured around the five core activities for each function, drawing on items from the frameworks and existing protocols from other OPRE and Mathematica projects. After Phase 1 data collection, we organized the collected data into rubrics on select topics by key function, transforming the organization and analysis of the information from the implementation core activities to the key center functions.

<div style="border: 1px solid teal; padding: 10px;">

Implementation science frameworks used to inform implementation data collection

Conceptual Model for Considering the Determinants of Diffusion, Dissemination, and Implementation of Innovation in Health Service Delivery and Organizations (Greenhalgh et al. 2004)

Framework for Implementation; Core Implementation Components; Multilevel Influences on Successful Implementation (Fixsen et al. 2005)

Interactive Systems Framework for Dissemination and Implementation (Wandersman et al. 2008)

Consolidated Framework for Implementation Research (Damschroder et al. 2009)

Framework for Effective Implementation (Durlak and DuPre 2008)

Quality Implementation Framework (Meyers et al. 2012)

Evidence-Based System for Innovation Support (Wandersman et al. 2012)

Active Implementation Frameworks for Program Success (Metz et al. 2013)

</div>

2.   Cost data collection tool

We developed an *Excel-based electronic cost workbook* to collect information on ECE program expenses. The workbook asks respondents to report costs for resource categories that are common in cost analyses—such as staff salaries and benefits, facilities, and supplies and equipment—as shown in Exhibit III.5.

Exhibit III.5. Resource categories collected in the ECE-ICHQ cost workbook

Resource category	Description/definition
Salaries and fringe	Salaries, taxes, and fringe benefits for regular, paid staff
Staff training and education	Costs for training and education provided to center staff members
Contracted services	Services purchased from organizations and/or people who operate independently from the center
Volunteers	Volunteers working at the center during the reporting year
Facilities	Facilities-related costs incurred during report period, including rent and utilities
Supplies and materials	Resources that cost under $1,000 that are used/replenished regularly
Equipment	Resources with a cost of more than $100 and an expected useful life of more than one year
Other/miscellaneous	Items and services that are not captured under any other categories
Payments/overhead	Payments (or amount calculated or allocated) for all or any of the resources the center received from a larger organization or entity, or for operating as part of a larger organization or entity

We built from a strong foundation of prior cost studies and surveys to develop the cost workbook. Sections of the workbook are informed by data collection tools used in past cost studies, including the Cost Study of the Saint Paul Early Childhood Scholarship Program (Schwartz and Karoly 2011) and the Evidence-Based Home Visiting study (Burwick et al. 2014). We also reviewed the director survey from the study of Cost, Quality and Outcomes (Helburn 1995) and the Provider Cost of Quality Calculator (Mitchell et al. 2014).

<div style="border:1px solid;padding:10px;">

Cost studies and surveys used to develop the cost workbook

Cost Study of the Saint Paul Early Childhood Scholarship Program (Schwartz and Karoly 2011)

Evidence-Based Home Visiting study (Burwick et al. 2014)

2012 National Survey of Early Care and Education Center-Based Provider Questionnaire (NORC 2011)

Program Administration Scale (Talan and Bloom 2011)

Cost, Quality and Outcomes (CQO) study (Helburn 1995)

Provider Cost of Quality Calculator (PCQC) (Mitchell et al. 2014)

</div>

3.   Staff time-use survey

We developed a ***time-use survey*** to collect information on how center staff use their time to allocate labor costs to key functions. The time-use survey asks about how much time respondents spent on a range of activities—some that happen frequently (such as daily or weekly) and others that happen only periodically (such as monthly, quarterly, or yearly) (Exhibit III.6). Teaching staff and select administrators are asked to report their total number of paid and unpaid working hours in a typical week, and to estimate the number of hours they spend on each activity.

Existing surveys or tools provided a foundation from which we could tailor questions or items that capture a distinct concept, to ensure they were relevant for ECE administrators and classroom staff (including lead teachers, head teachers, or co-teachers, as well as assistant teachers). The items we included were informed by prior time-use surveys used for cost analysis purposes, but to our knowledge, this approach had not been used in center-based ECE settings. For example, we looked to items from studies such as the Evidence-Based Home Visiting study (Burwick et al. 2014) and the Distributed Leadership Study (Spillane et al. 2004). We also consulted the 2012 National Survey of Early Care and Education Center-Based Provider Questionnaire (NORC 2011) for professional development topic categories, and the Program Administration Scale (Talan and Bloom 2011) to inform topics around program planning.

We considered other approaches for collecting information on time use, such as a daily log or time diary. However, using a survey format allowed us to (1) collect information about specific types of activities, including those that occur sporadically over the course of a year that were important to building the cost estimates; and (2) minimize burden on center staff.

Exhibit III.6. Time-use survey activities, by key function

Key function	Activity in time-use survey
Instruction and Caregiving	Providing instruction or care
Instructional Planning, Coordination, and Child Assessment	Planning curriculum, activities, and lessons
	Conducting child assessments during or outside of classroom time
	Parent communication
	Periodic child assessments and parent conferences
	Curriculum planning and development

Key function	Activity in time-use survey
Center Administration and Planning	Planning and goal setting
	Ongoing regulatory compliance and reporting
	Managing center finances
	Managing and maintaining center operations
	Periodic regulatory compliance and reporting
Workforce Development	Staff supervision, coaching, or mentoring
	Staff meetings
	Recruiting and hiring teaching staff
	Evaluating staff performance
	Training and professional development activities
	Periodic coaching, mentoring, or technical assistance
Child and Family Support	Planning and conducting family engagement or family support activities

C. Refining data collection tools and methods

We learned from the experiences in the pilot and Phase 1 about ways to improve and refine the data collection tools and processes in preparation for the early testing of the measures in Phase 2. These changes included the following:

Shorten and simplify data collection tools and processes for respondents. Data collection was time intensive. During Phase 1, it was difficult to schedule a long enough period of time to complete all implementation interviews while on-site. It also took most respondents more than an hour to complete the time-use survey. In Phase 2, the implementation interviews occurred over the phone and took just under four hours, on average, and spread across two or three interview periods; the time-use survey was substantially reduced to take 10 minutes to complete.

Limit (or remove) preparation and gathering of supporting documentation for the implementation interview. Center staff have limited time to participate in—or prepare for—data collection activities. To make the best use of center staff's limited time and reduce burden, we refined our approach so that respondents for the implementation interview could respond to most interview questions without having to prepare in advance, reference other documents, or consult with other staff.

Refine tools to improve clarity and data quality for analysis. The study team drew on lessons learned in Phase 1 to maximize opportunities to improve each data collection tool. For the implementation interview, we created implementation rubrics that would elicit more structured responses and ease analysis in Phase 2. We grouped implementation activities by key function to create rubrics focused on distinct subtopics. We then organized the information collected from each question in the Phase 1 implementation interview into the rubrics. The rubrics shed light on patterns in responses within a center and variations across centers. These patterns showed promise in making distinctions about what a center does and how it implements each key function. We used the rubrics to reorganize the implementation interview around the key functions to reduce redundancies in data collection, improve flow, and support analysis of data from a larger number of centers in Phase 2.

We streamlined the time-use survey by focusing on certain activities that are core to each key function. For example, we reduced the level of detail collected about training. We also dropped items asking about

activities for which we expect few differences between staff within a center, such as center wide staff trainings.

To improve cost data collection and reduce burden, we revised the cost workbook in several ways. First, we simplified the questions on some types of resources, including salaries and fringe, facilities, volunteer labor, equipment, and overhead. For example, the revised workbook does not request detailed information on the type of work and number of hours contributed by individual volunteers. Second, we added drop-down lists to some elements of the workbook to facilitate the allocation of line items to key functions. Third, to support more precise estimates of the cost per child care hour, we added items related to the number of hours of child care the center provided. Finally, we removed question items related to donated supplies and equipment because the estimated value of these donations tended to be very small relative to a center's total costs.

Although the tools became more streamlined and structured over the course of the multi-case study, the implementation interview and cost workbook still require administration and support by trained data collectors. Based on the experience in Phase 2 (discussed in more detail in a later section of this chapter), we determined that remote data collection is feasible for the implementation interview and cost data collection. We found, however, that on-site distribution and collection of the time-use surveys was helpful in achieving a high response rate from busy teaching staff who often wanted to complete the survey in sections, during breaks throughout the day.

D.   Testing tools and measures in Phase 2

We tested the revised data collection tools in Phase 2 and used the data to further revise and create a set of draft measures for validation in a field test. In this section, we describe the data collection tools and process we used in Phase 2 as background for understanding the measures we ultimately developed. We first discuss how we selected states and centers to achieve the variation in context and conditions necessary to support measures development.

1.   Using a purposive sample to achieve variation in context, conditions, and center characteristics

States and centers were purposively sampled to ensure variation on certain characteristics to inform measures development and test performance of the data collection tools across center-based settings. Throughout the multi-case study, it was important to conduct data collection with centers that varied in context and characteristics. We sought to develop measures that appropriately account for differences in state and community ECE policies and contexts, center characteristics, and regional and urban variations in the costs of doing business (such as salaries and rent).

We selected *states* for Phase 2 based on the following:

- Sizeable number of ECE centers participating in the state's Quality Rating and Improvement System (QRIS) (at least 300 centers participating) and centers reaching the highest rating level
- Variation in the stringency of licensing requirements
- Variation in state and regional contexts and conditions
- Offer state pre-K in community-based ECE centers
- A high proportion of all licensed centers participating in the QRIS

Based on these criteria, we selected Arizona, Arkansas, and Pennsylvania.[2] Exhibit III.7 summarizes characteristics of the states in which we recruited centers for Phase 2.

Exhibit III.7. Select characteristics of states included in Phase 2

	Arizona	Arkansas	Pennsylvania
Percentage (and number) of licensed center-based programs participating in QRIS	84% (701)	67% (1,240)	71% (3,007)
Number of center-based programs at highest QRIS level	34	643	629
Participation requirements in QRIS for center-based programs	Voluntary	Mandatory only for centers that receive CCDF subsidies	Voluntary
Number of rating levels in the QRIS	5	3	4
Tiered reimbursement for child care subsidies tied to QRIS participation[a]	No	Yes	Yes
Stringency of licensing requirements based on Child Care Aware ranking[b]	22	41	22
Region	West	South	Northeast

Source: Participation data and other QRIS characteristics were obtained from the QRIS Online Compendium and were current as of October 31, 2016. Child Care Aware rankings based on 2013 data (the most recent available) and were accessed at http://www.naccrra.org/sites/default/files/default_site_pages/2013/wcdb20rankings2020041013.pdf

[a] Tiered reimbursement rates increase the amount of payment a center receives per hour to subsidize the care of young children from low-income families as its QRIS level increases.

[b] Child Care Aware of America ranking is based on 15 benchmarks that reflect state regulations and policies, including requirements on staff credentialing and training, teacher-child ratios, group size, health and safety practices, and more. States are ranked based on their performance against the benchmarks, with a higher number indicating less stringent licensing requirements (for example a state with a ranking of 20 has more stringent requirements than a state ranked 40).

CCDF = Child Care and Development Fund.

We selected *centers* based on QRIS rating level, whether a center was community based or Head Start, and funding sources. We categorized centers into high and low QRIS ratings based on the requirements for the different levels in each state in the multi-case study. We included the top one or two rating levels in the high category, depending on the total number of rating levels and the definitions of high quality set by each state. We excluded the first rating level from the low category in two of the three states because there was no assessment or gauge of quality made at entry. We also excluded middle rating levels in two of the states to get a better distinction between high and low quality based on the QRIS requirements.

During recruitment, we gathered additional information to maximize variation in the ages of children served, center size (based on licensed capacity), profit status, and whether they were embedded in a larger organization. We show the characteristics of the 30 centers included in Phase 2 in Exhibit III.8.

---

[2] In Phase 1, we collected data from centers in Arkansas, Pennsylvania, and Wisconsin. To increase our ability to include centers that did not rely on public funding in Phase 2, we replaced Wisconsin with Arizona because Arizona had a higher percentage of participation in the QRIS among all licensed centers.

Exhibit III.8. Characteristics of centers in Phase 2 of the ECE-ICHQ multi-case study

Characteristic	Number of centers
QRIS rating[a]	
High	20
Low	10
Age groups served	
Ages 0 to 5 and school age	11
Ages 0 to 5 only	12
Ages 0 to 3 or ages 3 to 5 only	7
Program size	
Large center (serving 75 children or more)	18
Small center	12
Profit status	
For profit	11
Not for profit	19
Embedded in a larger organization or part of a chain[b]	
Yes	14
No	16
Funding	
High subsidy[c]	8
Majority Head Start or state pre-K[d]	10
Majority private tuition[e]	8
Mixed public or mixed public/private	4
Total sample size	**30**

Source: Data collected for the Assessing Implementation and Cost of High Quality Early Care and Education (ECE-ICHQ) Project, Phase 2.

[a] The high QRIS category includes Arkansas centers with a rating of 3, Arizona centers with a rating of 4 or 5, Pennsylvania centers with a rating of 4, and all Head Start centers in each state. The low QRIS category includes Arkansas centers with a rating of 1, Arizona centers with a rating of 2 or 3, and Pennsylvania centers with a rating of 2.

[b] Part of a larger organization means the center is part of, and shares resources/staff with, a larger organization, such as a YMCA, university, or other social services agency.

[c] Includes centers in which 50 percent or more of the children are supported by funding from CCDF and less than 30 percent of children are supported by funding from other federal, state, or local government sources.

[d] Includes centers in which less than 30 percent of the children are supported with funding from CCDF.

[e] Includes centers in which 90 percent or more of the children are supported through private tuition paid by their parents or guardians without any public funding.

2.   Collecting data around the key functions

We structured data collection in Phase 2 around the functions in the conceptual framework and the functions connected to the implementation and cost data collection (Exhibit III.9). The data collection process for Phase 2 began with the first implementation interview. During this interview, we collected information on the center's staffing structure that informed selection of the staff time-use survey

respondents. The staffing structure information, along with other items collected during the implementation interview (such as use of a curriculum, types of training), also helped the cost data collection team check alignment and coverage in the cost workbook.

Exhibit III.9. Implementation and cost data collection to understand key functions of a center

3. Testing implementation data collection

We used a semistructured interview to collect qualitative information about the implementation of each key function, including the features or practices in place (the "what"), and the intentionality, accountability, and consistency with which each is implemented (the "how"). We revised the implementation interview tool for use in Phase 2, using the rubrics we created from the Phase 1 data. We created an editable PDF that provided a structured format for trained interviewers to code information into closed-ended categories during the interview (rather than taking extensive notes). The PDF still allowed for embedding notes, as needed, to clarify or expand on responses. We created this type of tool to make data collection and entry of a large quantity of information efficient and to ensure the data are ready for cleaning and analysis. Interviews were recorded so data collectors could check completeness and accuracy of the data and for use in reliability checks by the implementation data collection lead.

The implementation data collection tool was not meant to function as a traditional survey (meaning interviewers did not read the answer categories for each item to respondents). We trained the interviewers to understand the intent of the questions. They adjusted the question phrasing and order as needed to respond to the semistructured nature of the interview and the flow of the conversation, and they probed to obtain, interpret, and confirm the responses. We allowed for open-ended responses to questions to ensure we gathered information capturing a range of activities across centers. In addition, interviewers made notes if responses were not captured in the existing categories.

We did not ask center staff to prepare for the interviews or provide materials in advance. We did suggest a few materials to have readily available for the interviews (such as job descriptions). The implementation data collection lead conducted initial reliability checks by listening in on the interviews each interviewer

conducted with their first center and coding alongside the interviewer. The implementation data collection lead conducted another reliability check with each interviewer halfway through data collection. These checks helped to ensure that the interviewers collected high quality data and coded the data in consistent ways. We provided additional ongoing support to resolve questions or discrepancies in data collection through weekly interviewer meetings.

Gathering implementation information from centers was time-consuming and required multiple interviews. For most centers, the interviews were conducted with one respondent. In seven centers, there were multiple respondents (Exhibit III.10). Total interview time per center was just under four hours and took between two and six interviewing sessions to complete; the average center needed at least three sessions to complete the interview. Because of center staff availability and scheduling constraints, there were often several days or weeks between the first and final interview sessions for a center. The time span between the first and final sessions ranged from 2 to 67 days (with an average of 20 days). With 13 centers, we experienced cancellations or no-shows for originally scheduled interviews, though we were ultimately able to conduct the interviews. Many respondents had numerous demands on their attention during each interview.

4. Testing cost data collection

We collected cost data through the cost workbook, an Excel-based electronic spreadsheet that was refined after Phase 1. Respondents entered information on categories of resources used to deliver services during the most recently completed fiscal year. Each type of resource (Exhibit III.5) appeared on a separate tab in the workbook. Respondents were asked to enter actual expenditures (or, in some cases, estimated values) related to each category of resource. In addition, the cost workbook requested information on the number of hours of care the center provides for children in different enrollment options (for example, full day, partial day) to support the estimation of cost per child care hour. The electronic spreadsheet format promoted accurate data entry through validation of values entered into cells (a feature that limits the type or range of values entered) and facilitated data extraction for analysis. The data collection process started with an orientation to the workbook for respondents, allowed time for respondents to complete the workbook independently, and then was supported by follow-up contacts to clarify responses or work with respondents to clarify or complete data, as needed.

Collecting cost data remotely from ECE centers required support and follow-up from trained data collectors. With follow-up, we collected complete cost data from more than 80 percent of centers in our sample. The completeness of cost data we initially received varied among centers and by resource type. More than half of the 30 centers in our sample provided incomplete or no information, when initially submitting the cost workbook, about the following types of resources: staff compensation, facilities, supplies and materials, overhead, and child care hours. After follow-up, 26 centers provided complete responses to the workbook.[3]

---

[3] The study team determined that data received from four centers was not complete enough to be included in the analysis, specifically because there was insufficient information on substantial resources like salaries or facilities.

The cost data collection required an intensive follow-up effort. On average, we made 15 contacts with each center over a period of 69 days (Exhibit III.10). In general, smaller centers, for-profit centers, and centers in which private tuition made up most of the revenue required fewer days to complete data collection than centers that did not have these characteristics. These types of centers may have had less complex financial structures and required less coordination among different individuals to provide cost data. Generally, it was possible to allocate individual expenditures to key functions. However, we needed more information about the purpose of expenditures reported in some sections of the workbook—such as facilities, equipment, and supplies and materials—to facilitate allocating these costs to specific functions. We collected the additional information through follow-up phone calls and emails.

Exhibit III.10. Data collection metrics for ECE-ICHQ Phase 2, by center characteristics

	Implementation data collection				Cost data collection			Staff time-use surveys	
	Number of centers	Number of centers with more than one respondent	Total interview length in minutes	Time from first interview to completion in days	Number of centers	Number of times contacted	Days to complete (orientation call to final workbook)	Number of centers	Mean percentage response
QRIS rating[a]									
High	20	7	236.6	20.1	19	16.1	71.6	20	91.2
Low	10	0	219.0	19.1	7	11.3	73.9	10	89.5
Age groups served									
Ages 0 to 5 and school age	11	0	228.6	17.4	9	14.0	81.9	11	87.6
Ages 0 to 5 only	12	6	239.2	22.8	10	14.2	64.3	12	89.0
Ages 0 to 3 or ages 3 to 5 only	7	1	219.4	18.3	7	16.7	71.1	7	98.3
Program size									
Large center (serving 75 children or more)	18	5	241.7	21.8	15	16.7	75.1	18	89.8
Small center	12	2	214.3	16.8	11	12.2	68.3	12	92.0
Profit status									
For profit	11	1	195.5	15.5	10	12.9	59.7	11	88.6
Not for profit	19	6	251.1	22.2	16	16.0	80.1	19	91.8
Embedded in a larger organization or part of a chain[b]									
Yes	14	5	247.1	18.8	13	16.5	77.2	14	92.2
No	16	2	216.3	20.6	13	13.2	67.3	16	89.3
Funding									
High subsidy[c]	8	1	209.4	21.9	6	11.8	86.0	8	94.2
Majority Head Start or state pre-K[d]	10	6	262.5	25.1	9	19.7	74.7	10	95.9
Majority private tuition[e]	8	0	210.8	15.0	7	9.7	48.6	8	82.8
Mixed public or mixed public/private	4	0	233.8	11.8	4	17.3	87.5	4	85.9
All centers	**30**	**7**	**230.7**	**19.8**	**26**	**14.8**	**72.2**	**30**	**90**

Source: Data collected for the ECE-ICHQ project, Phase 2.

[a] The high QRIS category includes Arkansas centers with a rating of 3, Arizona centers with a rating of 4 or 5, Pennsylvania centers with a rating of 4, and all Head Start centers. The low QRIS category includes Arkansas centers with a rating of 1, Arizona centers with a rating of 2 or 3, and Pennsylvania centers with a rating of 2.

[b] Part of a larger organization means the center is part of, and shares resources/staff with, a larger organization, such as a YMCA, university, or other social services agency.

[c] Includes centers in which 50 percent or more of the children are supported by funding from CCDF and less than 30 percent of children are supported by funding from other federal, state, or local government sources.

[d] Includes centers in which less than 30 percent of the children are supported with funding from CCDF.

[e] Includes centers in which 90 percent or more of the children are supported through private tuition paid by their parents or guardians without any public funding.

CCDF = Child Care and Development Fund.

5. Testing time-use data collection

To estimate costs by key function, we fielded a staff time-use survey to support allocating labor costs (the biggest resource category) to different key functions. We identified potential respondents during the first implementation interview, including (1) all teaching staff of children ages 0 to 5 and (2) key administrators who oversaw the center's operations (such as the center director and assistant director) and the educational program lead. Field staff confirmed all respondents on-site and distributed information about the surveys; they stayed on-site through the day or returned the next day to answer questions or collect completed surveys. We identified 483 staff in roles relevant for completing the time-use survey across the 30 centers; 430 responded to the survey, for an overall response rate of 89 percent. The average center response rate was 90 percent, ranging from 63 to 100 percent across centers; response rates did not vary systematically by center characteristics (Exhibit III.10). Respondents had the option to complete a web-based or paper survey; 71 percent opted to complete paper surveys (nearly all teachers preferred this method). The average number of respondents per center was 14 but ranged from 3 to 45 based on the size of the center. We used 421 of 430 surveys in the analysis based on completeness and quality of the data; 359 of the 421 surveys were completed by teaching staff.

In the remainder of the report, we focus on the development of the draft measures produced using the data collected from the 30 centers in three states included in Phase 2.

IV. Developing Implementation Measures

Useful and relevant measures for interpreting implementation—what a center does to support quality and how —need to comprise a small set of items that can be used consistently across a range of center-based ECE settings. Our task following the multi-case study was to substantially reduce the burden of data collection on center staff and build practice-relevant summary measures that are acceptable, by technical standards, in preliminary measures development. To accomplish this goal, we used the comprehensive Phase 2 data to identify items that would provide the greatest efficiency and contribute the highest relevance in creating the measures (implementation scores). In this chapter, we describe the methods we followed to develop the draft implementation measures for each of the key functions in our conceptual framework. Our aim was to create strong measures that would be ready for more formal testing as part of a field test.

Implementation measures reflect …

***What* a center does to support quality, including:**

- Structural features that are in place (teacher–child ratios, group size, staff qualifications)
- Adopted practices (curriculum use, child assessment)

***How* quality features and practices are supported and implemented within a center through implementation drivers, including (Caronongan et al. 2016):**

- Staff selection and hiring
- Use of evidence-based tools
- Training to support practice and use of tools
- Coaching to support practice and quality improvement
- Quality assurance and quality improvement processes

A. Defining implementation measures

Following the multi-case study, we developed an implementation measure for each of the five key functions in the conceptual framework, to represent what a center does within the function and how the work is carried out and supported. The measures are scored from 0 to 1. A higher score (closer to 1) indicates stronger structural features of care, intentional practices that are associated with quality, and attention to implementation drivers that are typically present in programs that produce positive outcomes (described in Chapter I; see Caronongan et al. 2016 for further details).

Implementation measures result in scores for:

- Structural Supports for Instruction and Caregiving
- Instructional Planning, Coordination, and Child Assessment
- Center Administration and Planning
- Workforce Development
- Child and Family Support

We focused implementation data collection on center-level information that could be reported, for the most part, by one respondent. We conducted semistructured interviews with the center director and/or the

person they identified as being most knowledgeable about the educational program in the center. We wanted to focus implementation data collection with center directors or educational program managers and not place an undue burden on teaching staff. (We needed to prioritize administering the time-use survey to teaching staff to produce the cost estimates, as discussed in Chapter V.)

We focused data collection on descriptive, objective information, because we could not triangulate information across respondents. We gathered information, by function, that was comprehensive enough to build a picture of what a center does and how; we wanted this information to correspond with costs and inform pathways to quality. For example, we wanted to learn what curriculum is primarily used in preschool classrooms in a center (the adopted practice to capture what a center does), as well as what informed the selection of the curriculum and how teaching staff are trained on the curriculum (the implementation drivers to capture how the work is carried out). As this was an effort to measure implementation rather than quality, we gathered information on how the center implemented specific structures, practices, or activities, but not how well they did it.

A limitation of our approach is that by relying on only one perspective, the interview may have captured information on intended structures or practices (for example, frequency or consistency) rather than what was actually happening (adherence to intent). There might be differences between what a center director reports and what staff report is happening in the center. However, we have no reason to believe that this approach affects centers differentially. We found a considerable amount of variation on the implementation scores, suggesting that the scores are capturing useful information that can distinguish implementation between centers (as discussed later in this chapter).

B. Developing the implementation measures

A sample of 30 centers allowed us to conduct preliminary analyses to develop and assess the psychometric properties of the implementation measures. We developed the measures (or scores) for each of the five key functions in the conceptual framework through an iterative process.

> **Construct:** An idea, underlying theme, or subject matter measured by related items
>
> **Theory-driven:** Grouping items based on their content
>
> **Data-driven:** Grouping items based on empirical evidence through statistical analysis

We attempted to create the implementation measures of each of the five key functions in a few different ways, ending with an approach that balanced theory-driven and data-driven methods. We started with approaches that relied on theoretical groupings of items within each of the key functions (for example, grouped by our constructs of "what" and "how," and by implementation principles such as intentionality, accountability, and consistency). We also explored purely data-driven approaches to determine solutions for the item sets for each key function through exploratory and confirmatory factor analysis. Both the purely theory-driven and purely data-driven approaches fell short of our goals because the resulting item sets were small, did not group items together well conceptually, or did not achieve the levels of reliability we were seeking. The data-driven approach did result in function scores that logically related to center characteristics, so we hoped to retain similar item sets in a revised approach that balanced theory and data-driven methods.

> **Psychometric properties** show how reliable and valid a measure is based on the purposes for which it is designed and used.
>
> **Reliability.** Indicators of reliability show how dependable a measure is for the purpose for which it is used. Reliable measures are stable over time and include items that measure the same thing in different ways. Statistical measures of reliability are typically reported as coefficients, which range from -1.0 to 1.0, with a greater value reflecting greater reliability. Many researchers and assessment developers require that measures have reliability values of 0.7 or higher.
>
> **Validity.** The degree to which an assessment accurately measures what it is designed to measure. In the early stages of measures development, validity can be examined based on patterns with other observable variables or characteristics. Validity is sometimes measured in comparison to other instruments established to measure the same or similar constructs.

To balance theory-driven and data-driven approaches, we used a method to produce a single item set for each of the five key functions. We built up an item set for each key function by examining item-total correlations (ITC) together with our theory, based on the literature review and consultations with experts, about items that comprise the "what" and "how" of each key function. The advantage of this approach is that the resulting scale is not as dependent on the idiosyncrasies of the data from a small sample of 30 centers as a purely data-driven solution (developed through factor analysis) may be. In addition, this approach allowed us to organize items into theoretical groupings within the one scale, which can support interpretation and use.

We followed a series of steps to identify item sets for each of the key function implementation measures (or scores). (Appendix A provides additional technical information about the approach.) We used data from implementation interviews with 29 to 31 centers to develop the measures for the five key functions.[4]

> **Factor analysis.** Statistical analysis that examines the pattern of relationships among items to form related groups of items (or factors). Factor analysis may be exploratory (looking at how items group together in the data) or confirmatory (examining whether the relationships among items are consistent with a predetermined structure).
>
> **Factor solution.** Findings from a factor analysis to define (in exploratory factor analysis) or confirm (in confirmatory factor analysis) the number of item groupings—the factors. One factor or multifactor solutions are possible.
>
> **Correlation.** The degree to which two items vary together, ranging from -1.0 (a perfect negative relationship) to 1.0 (a perfect positive relationship), with 0 meaning no relationship.
>
> **Item-total correlation (ITC).** A measure of the reliability of a multi-item scale and a tool for improving such scales. It determines the correlation between an individual item and the sum of the values of all other items. With a set of 10 items, there would be 10 item-total correlations. For item 1, it is the correlation between it and the sum of the other 9 items, and so on.
>
> **Scale.** Grouping or categorization of items together to measure a construct.

---

[4] One center completed all sections of the implementation interview except one; one center completed all implementation data collection but not cost, and is not included in the total sample for Phase 2.

Recode items. We began preparing the data for analysis by recoding individual implementation items (interview questions) so that a higher score on a single item meant more of a presence of what we were measuring. The recoded items were either a hierarchy (captured in an ordinal variable) or a sum (examples shown in Exhibit IV.1). There are a few binary items (for example, yes/no), such as whether a center provides financial or other assistance for teachers to support additional training or education. We assessed the variation in the recoded items by looking at descriptive statistics, including each item's minimum and maximum values and its distribution of values across all the centers.

Exhibit IV.1. Examples of recoded items based on a hierarchy and a sum

Item	Recode
Frequency of curriculum training for teaching staff	0 = No curriculum or no training is provided 1 = Just occasionally (once a year or as needed) 2 = Just initially (upon initial curriculum implementation or during new staff orientation) 3 = Occasionally and initially 4 = Monthly 5 = Monthly and initially
Individual(s) that provide input into teaching staff performance	0 = None of the categories are marked 1-4 = Sum of contributors with value of 4 representing four or more categories with input • Center leadership (director/assistant director/supervisor) • Other teaching staff • Other center staff (e.g., family support worker, specialists) • Parents • Coach (if other than supervisor) • Other (specify) that represents a distinct and relevant category

Examine pairwise correlations. After we computed descriptive statistics for each recoded item (means, range, standard deviation), we explored the statistical associations between the recoded items within a key function by computing pairwise correlations—the association between each set of two items. From each pair of items that were highly associated (with a correlation of .90 or greater), we excluded from further analysis one item that we deemed less substantively important. For example, if two similar questions about hiring lead and assistant teachers were highly correlated, we removed the question about assistant teachers. We took this step to ensure that each item that remained captured something about the construct we wanted to measure that was unique in some way. Removing highly related items reduces item redundancy to help get to a smaller item set for efficient measurement and reduces burden for respondents.

Examine item-total correlations. To further refine the item set for a key function before conducting a factor analysis, we computed an item-total correlation for each remaining item (after removing highly associated items in the prior step). We first added up the total score—the sum of the values—of all remaining items we wanted to explore for each key function. We then computed the correlation between each individual item and a total score (excluding the value for the item itself).[5] Items with high item-total

---

[5] Note that item-total correlation is substantively different from a factor loading. No latent factor is created in computing item-total correlations. In contrast, a factor loading comes out of a statistical procedure that computes a

correlations account for more variation in a scale than items with low item-total correlations. Items with high item-total correlations are the items that might primarily define the meaning of the scale and are effective at making distinctions between centers in the construct being measured (different levels of implementation).

We maintained items that had a minimum item-total correlation of 0.2 (absolute value). We selected this minimum that reflects a moderate ability of an item to contribute to making a distinction in the construct of interest—for example, between centers about the level of implementation of a key function—and could be considered for potential inclusion in a scale (Pope n.d.).

Conduct confirmatory factor analysis (CFA).
After we identified the conceptual grouping of items for each key function using the steps outlined above, we turned to confirming the statistical associations among the items within each key function using confirmatory factor analysis. Confirmatory factor analysis tests whether hypothesized relations among the items are supported by the data as a step toward assessing construct validity. Confirmatory factor analysis can also further reduce an item set that forms a scale to those items that work together to measure a construct most reliably. Some items drop out as the analysis determines a factor solution that has the best fit to the data and produces the highest reliability possible among the items.[6] The confirmatory factor analysis produces factor loadings for each item that reflect how much the item contributes to measuring the factor solution; the higher the factor loading, the more information the item contributes to the solution.

> **Steps to develop the implementation measures**
>
> 1. **Recode items** so that a higher score aligns with contributing to stronger implementation.
>
> 2. **Examine pairwise correlations** to reduce item redundancy and identify items that contribute unique information.
>
> 3. **Examine item-total correlations** to develop an item set that can make distinctions in the level of implementation between centers.
>
> 4. **Conduct confirmatory factor analysis** to assess construct validity and reliability.
>
> 5. **Standardize the score for each item** in the final set on a scale from 0 to 1.
>
> 6. **Compute the key function score** as the average of the individual items.

> **Construct validity.** An estimate of how well a measure assesses the theoretical construct it claims to measure

We set a benchmark to retain items that had factor loadings of 0.32 or greater in the confirmatory factor analysis, representing a moderate contribution to measuring the factor.[7] (See Appendix A for further information.)

---

latent factor and measures the proportion of variation in each item accounted for by the latent factor. Item-total correlations and factor loadings likely differ for this reason, and we use them in different steps to develop the measures.

[6] Reliability from a confirmatory factor analysis is often referred to as internal consistency reliability. This assesses how well an item set is measuring what you want it to measure.

[7] Loadings below 0.30 are typically considered nuisance or non-salient (Brown 2015). We included items with factor loadings ≥ 0.32, corresponding to at least 10 percent ($0.32^2 = 0.10$) of the variance in the item shared with a factor.

> **Good fit to the data** suggests that a statistical model (in this case, a confirmatory factor solution) is a good representation of the data. Good fit suggests that the scales would have similar reliability when computed with data from another sample of centers, assuming similar values and distributions of the items.
>
> **Observed variables** are directly measured (captured in our items).
>
> **Latent factors** are variables that are not directly observed but contribute to information about a construct.
>
> **Factor loading** is the correlation between the item and the factor that it measures (as represented by the correlation coefficient between the observed variable and the latent factor). The higher the factor loading for an observable variable, the more that variable contributes to the solution, or answer, produced by the analysis.

Exhibit IV.2 presents the item sets for each key function that meet the minimum value of 0.2 for item-total correlations. The items shown in green italic font are those that contributed to the factor solution (with factor loadings of at least 0.32) produced by the confirmatory factor analysis.

Exhibit IV.2. Theoretical groupings of items that meet the item-total correlation threshold (0.2 or greater) for each key function

Key function (sample size)	Theoretical groupings	Item total correlation[a]	Factor loadings
Structural Supports for Instruction and Caregiving (N = 31)	**Ratio and group size:**		
	M1. Ability to meet age-based group size targets	0.941	0.716
	M2. Ability to meet age-based teacher/child ratio targets	0.906	0.641
	Materials and equipment:		
	M3. Availability of equipment to support caregiving	0.474	0.644
	M4. Availability of materials to support instruction	0.593	0.668
	M5. Access to outdoor play area	0.364	-0.076
	Staff qualifications:		
	N2a, item 1. Education level with highest proportion of lead teachers	0.503	0.307
	N2a, item 2. Proportion of lead teachers who have at least 18 hours of ECE coursework	0.394	0.810
	N2c/d, item 1. Education level of education specialists/program directors	0.325	-0.129
	N2c/d, item 2. Proportion of education specialists/program directors who have at least 18 hours of ECE coursework	0.246	0.922
	N2c/d, item 3. Proportion of education specialists/program directors who are certified	0.252	0.988
	N2b, item 1. Education level with highest proportion of assistant teachers	0.258	-0.270
	N2b, item 2. Proportion of assistant teachers who have at least 18 hours of ECE coursework	0.265	0.568

Key function (sample size)	Theoretical groupings	Item total correlation[a]	Factor loadings
	Ability to meet teacher qualifications and retain teachers:		
	N4a. Ability to meet targeted qualifications for lead teachers	0.293	0.266
	O4a. Factors that influence lead teachers to stay at the center	0.398	0.560
	O1. What is the approximate proportion of teacher turnover?	0.394	-0.281
Instructional Planning, Coordination, and Child Assessment (N = 31)	**Selection, length of use, and training on curriculum:**		
	E4, item 1. Criteria for selecting or developing infant/toddler curriculum	0.337	-0.033
	E4, item 2. Criteria for selecting or developing preschool curriculum	0.350	0.065
	E3, item 2. Duration of main curriculum used for preschoolers	0.771	0.420
	F2. Method for training center leadership on the curriculum	0.700	0.852
	F4. Method for training teaching staff on the curriculum	0.791	0.850
	F6. Frequency of curriculum training for teaching staff	0.808	0.397
	Planning time for teachers:		
	G1. Existence of paid planning time for various types of staff	0.200	0.138
	G2. Frequency and duration of paid planning time for staff	0.804	0.542
	Screening and child assessment:		
	H2. Level and qualification of staff who administer developmental screening/assessment	0.360	-0.235
	H3. Duration of main screening/diagnostic tool use[b]	0.826	0.628
	J3. Duration of use of main measurement tool for formative assessment[b]	0.666	0.725
	J4. Criteria for selecting formative assessment tool	0.641	0.746
	J5. Teacher use of information gathered through formative assessments to guide instruction[b]	-0.444	0.470
	J6. Process for tracking child's instructional assessments	0.353	0.601
	J10. Criteria for selecting summative assessment tool	0.208	-0.406
	K4. Method of training teaching staff on the assessment tool(s)	-0.741	0.175
	Parent communication about children's learning:		
	L1a and L1b. Frequency and format with which the center provides descriptive information formally to parents	0.314	0.089
	L4a and L4b. Frequency and format with which the center provides information to parents about how their child compares to other children	0.279	-0.023
	L5. Length and frequency of parent-teacher conferences	0.831	0.506

Key function (sample size)	Theoretical groupings	Item total correlation[a]	Factor loadings
Center Administration and Planning (N = 30)	**Structural and procedural supports:**		
	M6. Facility improvement needs	0.424	0.676
	N2d. Certification of center director	0.222	0.023
	O1. Turnover in center staff other than teaching staff	0.375	0.205
	AC1. Process for informing teaching staff of licensing requirements or standards the center must follow	0.298	0.417
	AC2. Process for monitoring compliance with requirements and standards	0.552	0.449
	Z3. Facilitators in ability to change policies and procedures or make improvements	0.432	0.214
	Staff engagement in center planning and improvement:		
	Y4. Staff discussion of quality assurance/need for quality improvement[c]	0.262	0.318
	X1, X2 SP. Existence of strategic planning activities and level of staff involvement in strategic planning efforts	0.674	0.828
	X1, X2 QI. Existence of quality improvement activities and level of staff involvement in quality improvement efforts	0.764	0.961
	X1, X2 PE. Existence of program evaluation activities and level of staff involvement in program evaluation efforts	0.676	0.881
	Setting and tracking center-level goals:		
	Y2. Concepts emphasized in center mission	0.248	0.305
	X3. Types of goals set for the center	0.477	0.529
	X4. Frequency of reviewing progress toward center goals	0.364	0.394
	X5. Use of data in tracking center goals	0.470	0.632
	X6. Intentionality in creating action plans to achieve goals and communicating progress to center staff	0.537	0.360
	Collaborative center oversight (boards):		
	AA2. Existence of board and focus of input to center operations	0.583	0.741
	AA3. Level of board oversight of center operations	0.504	0.792
	AA5. Existence of a parent policy council and frequency of their meetings	0.598	0.957
	AA6. Focus of parent group's input to center operations	0.711	0.975
	AA8. Center director participation on community boards/councils	0.429	0.141

Key function (sample size)	Theoretical groupings	Item total correlation[a]	Factor loadings
Workforce Development (N = 30)	**Interviewing and evaluating teaching staff:**		
	P3. Involvement of teaching staff in interview processes for teacher candidates	0.697	0.323
	P4a. Interview activities for lead teacher candidates	0.619	0.511
	W2. Typical individual performance goals set for teaching staff	0.495	0.432
	W3. Process for determining individual performance goals for teaching staff	0.748	0.659
	W4. Center processes for monitoring and evaluating performance of teaching staff	0.526	0.368
	W7. Individual(s) who provide input into teaching staff performance	0.634	0.563
	Setting priorities for and assessing training and professional development for teaching staff:		
	R1. Process for determining training and professional development priorities for teaching staff	0.555	0.508
	R2. Center's priorities for training and professional development of teaching staff	0.583	0.582
	R4. Staff level of formal involvement in determining individual training and professional development needs	0.608	0.493
	R6. Formality of the process for determining if training meets center goals and staff needs	0.547	0.648
	S6. Type of activities used to assess knowledge acquisition or skill development following training	0.547	0.571
	Supports for training and professional development of teaching staff:		
	S2. Number of training and professional development days provided by the center [c]	0.386	0.316
	T2, item 1. Center provides financial or other assistance for teachers to support external training or education	0.831	0.676
	T2, item 2. Formality of supplemental assistance for teachers to support external training or education	0.352	0.179
	T3. Center's approach to compensating staff for time spent attending external training	0.232	0.328
	T4. Center's approach to providing classroom coverage when teaching staff attend training	0.242	0.140
	Classroom observations and coaching to support teaching staff:		
	U1. Goals of classroom observations	0.429	0.265
	U7. Extent to which teaching staff visit other classrooms to observe practice and where	0.779	0.594
	U8. Activities following teachers' observations of other classrooms	0.693	0.619
	V1. Occurrence of individualized coaching for teaching staff	0.246	0.427
	V5. Activities involved in individualized coaching	0.612	0.570

Key function (sample size)	Theoretical groupings	Item total correlation[a]	Factor loadings
Child and Family Support (N = 29)	**Total services (in center or referral):**		
	Tot_serv. Number of services for which the center provides access	0.894	0.644
	Tot_cat. Count of categories of services for which the center provides access	0.910	0.923
	Service provision (in center):		
	Tot_incenter1. Count of all services explicitly provided by the center with a cost	0.591	0.736
	Tot_incenter2. Count of all services provided in the center, but by another entity and at no cost	0.529	0.506
	Cat_incenter. Count of all categories of services provided in the center (cost or no cost)	0.797	0.835
	Process for deciding on and tracking service receipt:		
	C3_Health. Collaboration in deciding if health services are needed (who decides)	0.324	0.218
	C5_Health. Method for tracking health service receipt	0.658	0.677
	C8_Therapy. Process for deciding if therapeutic services are needed	0.300	0.275
	C11_Therapy. Frequency of tracking therapeutic service receipt	0.475	0.415
	C12_Therapy. Method for tracking therapeutic service receipt	0.535	0.506
	C15_Coun. Process for deciding if counseling services are needed	0.710	0.839
	C16_Coun. Collaboration in deciding if counseling services are needed (who decides)	0.321	0.428
	C18_Coun. Frequency of tracking counseling service receipt	0.734	0.775
	C19_Coun. Method for tracking counseling service receipt	0.620	0.744
	C23_SS. Collaboration in deciding if social services are needed (who decides)	0.703	0.746
	C25_SS. Frequency of tracking social service receipt	0.898	0.899
	C30_Ed. Collaboration in deciding if adult or parent education or services are needed (who decides)	0.830	0.914
	C33_Ed. Method for tracking adult or parent education or services receipt	0.771	0.880

Source: Data collected for the Assessing Implementation and Cost of High Quality Early Care and Education (ECE-ICHQ) Project, Phase 2.

Note: Items in green font contributed to the factor solution (with factor loadings of at least 0.32) produced by the confirmatory factor analysis.

[a] We used the absolute value of the ITC. Two items (J10 and K4) have negative values, likely due to highly skewed distributions. We opted to include these items in the scale because the distribution of these items in different samples may vary.

[b] When more than one screening or assessment tool was used in a center, the item focuses on the main tool.

[c] Item with a factor loading that rounds to 0.32.

C. Findings about the properties of the implementation measures

Once we had identified item sets for potential use as scales to measure implementation of each function, we looked at the psychometric properties, particularly reliability. We also explored how well the resulting measures (or scores) for each function could distinguish levels of implementation across centers (based on variation in the scores), how the scores are associated among the functions, and how the scores vary based on center characteristics. Together, this information sheds light on the feasibility of the measures to help us understand implementation of the key functions and whether we have solid draft measures that are ready for more rigorous testing.

1. Reliability of measures

With the steps described, we used two approaches to identify items to measure implementation for each key function: (1) defining the item set based on item-total correlations and (2) using confirmatory factor analysis.

We ran two confirmatory factor analyses for each of the item sets. In the first column of Exhibit IV.3, we show the total number of items included in the scale (or factor) based on each approach. The number of items included in the confirmatory factor analysis based on item-total correlations is shown first. Shown in parentheses is the reduced number of items that met the minimum value for factor loading (of 0.32) in the confirmatory factor analysis that included all the items based on item-total correlations. These item counts match to those presented in Exhibit IV.2, where all the items included based on item-total correlations are listed for each key function; items that had the minimum value of 0.32 for a factor loading are in green font.

Exhibit IV.3. Model fit statistics based on confirmatory factor analysis that includes all items that meet the item-total correlation threshold (0.2 or greater) for each key function

Key function	Items in the full CFA model (items with loadings >= 0.32)	McDonald's Omega (ω)[c] for full model (for items with loadings >+ 0.32)	Posterior p-value (PPP)[a]	95 percent confidence interval for difference f-statistic[b]
Structural Supports for Instruction and Caregiving	15 (9)	0.804 (0.912)	0.081	-16.650–101.238
Instructional Planning, Coordination, and Child Assessment	19 (11)	0.698 (0.873)	0.155	-42.156–118.531
Center Administration and Planning	20 (15)	0.903 (0.928)	0.159	-34.551–126.809
Workforce Development	21 (18)	0.857 (0.866)	0.46	-84.868–99.362
Child and Family Support	18 (16)	0.939 (0.947)	0.087[d]	-21.496–136.165

Source: Data collected for the Assessing Implementation and Cost of High Quality Early Care and Education (ECE-ICHQ) Project, Phase 2.

[a] PPP values ≥ 0.05 typically indicate good fit of the model to the data. The PPP statistic refers to the CFA model, which includes all items with ITC > .20. An excellent-fitting model is expected to have a PPP around 0.5, whereas a low value, approaching 0.0, indicates poor fit (Muthén and Asparouhov 2011).

[b] A 95 percent confidence interval for the difference in the f-statistic for the real and replicated data should contain zero. The PPP confidence interval refers to the CFA model, which includes all items with ITC >= .2

[c] An omega of close to or above 0.70 is considered sufficient for the early stages of scale development.

[d] In the CFA model for Child and Family Support, we additionally allowed for the residual correlations between items C11_Therapy with C12_Therapy (r = 0.690), C15_Count with C16_Count (r = 0.780), and Tot_incenter2 and Cat_incenter (r = 0.905). Because of the similarity of their content, these items appear to share common variance after controlling for the latent factor.

We found that both approaches identified item sets that produced reliable scales to measure each key function. We measured reliability by McDonald's Omega (ω) and used reliability that reaches 0.70 as a marker that is sufficient for the early stages of development (Nunnally 1978). We found that reliability of the scales produced for each key function using the items that met the item-total correlation threshold of an absolute value of .20 or greater ranged from $\omega = 0.698$ to 0.939 (shown in the second column of Exhibit IV.3). Reliability of the scales for each key function produced with items from the confirmatory factor analysis (with factor loadings of at least 0.32) was even higher, ranging from $\omega = 0.866$ to 0.947 across the functions (shown in parentheses in the second column of Exhibit IV.3). This suggests that the items in the scales are internally consistent—they work well together to measure implementation of each key function with a relatively small amount of measurement error.

> **Item sets for each key function produced reliable scales.**

The confirmatory factor analysis models for each key function provided a good fit to the data—as measured by the predictive posterior p-values (PPP, a measure of goodness of fit for Bayesian CFA models) and the confidence interval, shown in the next two columns of Exhibit IV.3. A good fit of the data suggests that the scales would have similar reliability (internal consistency) when computed with data from another sample of centers, assuming similar values and distributions of the items.

2.   Variation in implementation scores

For each key function, we created summary scores using the item sets determined through the item-total correlation method. We were reluctant to be restrictive in limiting the items to those indicated by the factor models based on the data from just 30 centers. We wanted to balance the need for data reduction (to reduce the burden of data collection on centers) with the need for further testing with a larger sample of centers in a planned field test. Because both methods produced reliable scales, we decided to use the larger item set we created from the item-total correlations to create and examine implementation scores for each key function. We want to include those items we see as promising for measuring each key function in a further test with more settings. This approach will improve our ability to create scales with strong psychometric properties, based on the larger field test sample.

Before computing scores on a key function, we standardized each item to be on a scale from 0 to 1. We rescaled each item to fit within a 0 to 1 value using the minimum and maximum possible score for each item. After rescaling the items, we computed the key function score as the average of all individual item

values within a key function. Each item contributed equally to the score; we did not weight any items to carry more value than others.[8]

The scores for each key function have substantial variation; we are able to discriminate among centers in the middle of the distribution as well as the centers along the tails (Exhibit IV.4). The mean scores lean somewhat to the higher end, ranging from a mean score of 0.505 for the Child and Family Support function to 0.629 for the Instructional Planning, Coordination, and Child Assessment. We might expect this outcome, given that the sample centers all participated in a quality rating and improvement (QRIS) system. We had fewer centers with a low QRIS rating level (10 centers) in the sample, relative to those with a high QRIS rating (20 centers). The scores are generally well distributed, although the range and standard deviations for the functions of Instruction and Caregiving and Instructional Planning, Coordination, and Child Assessment reflect a tighter range of scores than those of the other functions. The distribution of the scores for each function fare well when compared to the ideal of a normal distribution, or bell curve. Skewness is an indication of the symmetry of the data around the mean; the skewness of the scores for each key function suggests approximate symmetry (based on values between -.5 and .5). Kurtosis tells us about the flatness of a distribution, relative to a normal distribution; flatter data cannot make distinctions between scores as well as data that form a normal distribution. The kurtosis of the scores of each key function (with values above 2) suggest moderately normal distributions (a normal distribution has a kurtosis of 3). Exhibit IV.5 shows graphic presentations of the scores for each key function.

> **Implementation scores for each key function discriminate among centers.**

Exhibit IV.4. Descriptive statistics for each key function's summary scores based on items that meet the ITC threshold (0.2 or greater) for each key function

Key function	Mean	Median	Standard deviation	Range	Skewness	Kurtosis
Structural Supports for Instruction and Caregiving	0.591	0.572	0.119	0.356–0.872	0.144	2.80
Instructional Planning, Coordination, and Child Assessment	0.629	0.629	0.123	0.360–0.816	-0.520	2.45
Center Administration and Planning	0.468	0.449	0.142	0.234–0.750	0.112	2.05
Workforce Development	0.593	0.605	0.148	0.196–0.832	-0.787	3.52
Child and Family Support	0.505	0.470	0.184	0.071–0.812	-0.302	2.63

Source: Data collected for the Assessing Implementation and Cost of High Quality Early Care and Education (ECE-ICHQ) Project, Phase 2.

---

[8] Scores are sometimes created by giving more weight to some items over others based on factor loadings from the confirmatory factor analysis. Given the small sample size and the early stage of development, we decided to create unweighted scores for each key function.

Exhibit IV.5. Distributions of implementation scores for each key function

Instruction and caregiving scores
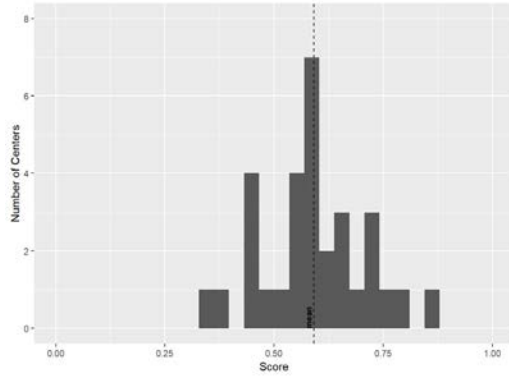
Instructional planning, coordination, and child assessment scores
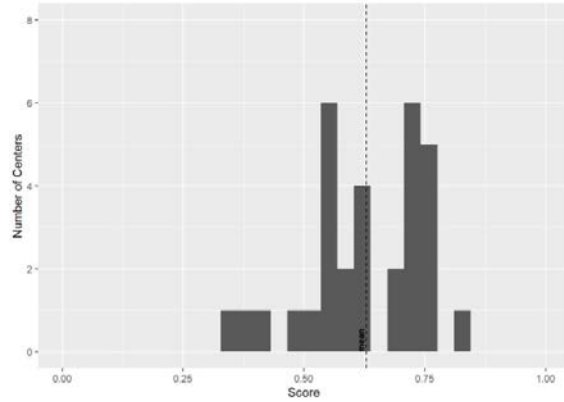
Center administration and planning scores
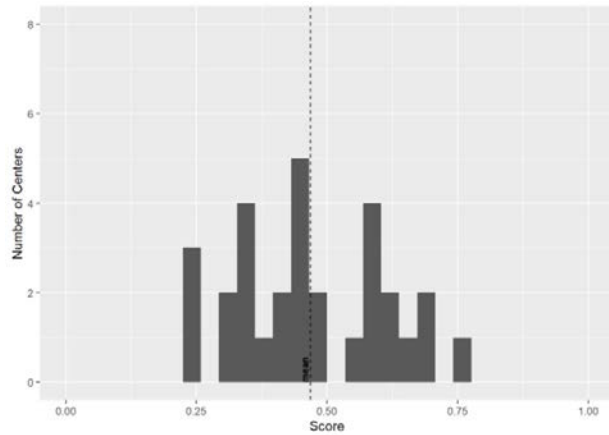
Workforce development scores
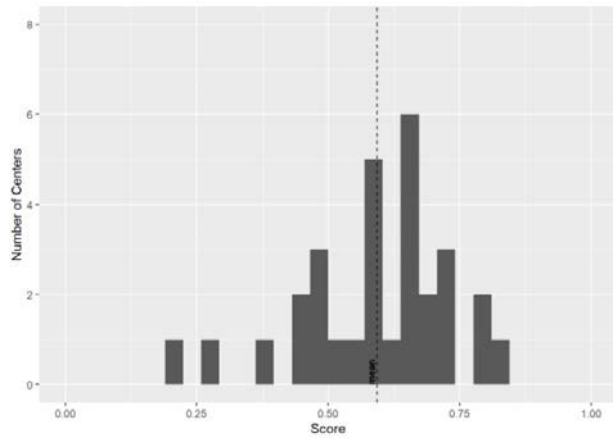
Child and family support scores
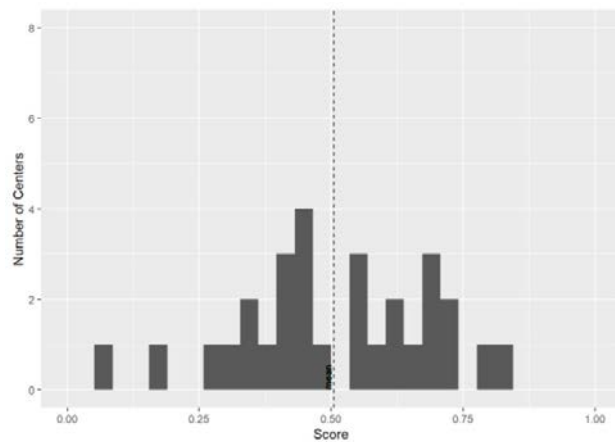
Source: Data collected for the Assessing Implementation and Cost of High-Quality Early Care and Education (ECE-ICHQ) Project, Phase 2.

3. Associations among measures (dimensionality)

We explored the correlations between the implementation scores across the five key functions and found moderate, significant associations among them in ways we might expect. Exhibit IV.6 shows the significant associations between the implementation scores for each of the five key functions. The scores for Center Administration and Planning are positively and significantly correlated with the scores of the other four functions, ranging from .41 to .55, which suggests the central nature of this function in driving implementation. The

> **Implementation scores for the key functions are correlated, but each also provides distinct information.**

Instructional Planning, Coordination, and Child Assessment implementation score is associated with those of Instruction and Caregiving and Workforce Development, in addition to the score for Center Administration and Planning. The implementation score for Child and Family Support is distinct from three functions; it is not significantly correlated with Instruction and Caregiving, Instructional Planning, Coordination, and Child Assessment, or Workforce Development. The score for Child and Family Support is associated with the score for Center Administration and Planning.

Exhibit IV.6. Significant correlations among implementation measures of the five key functions
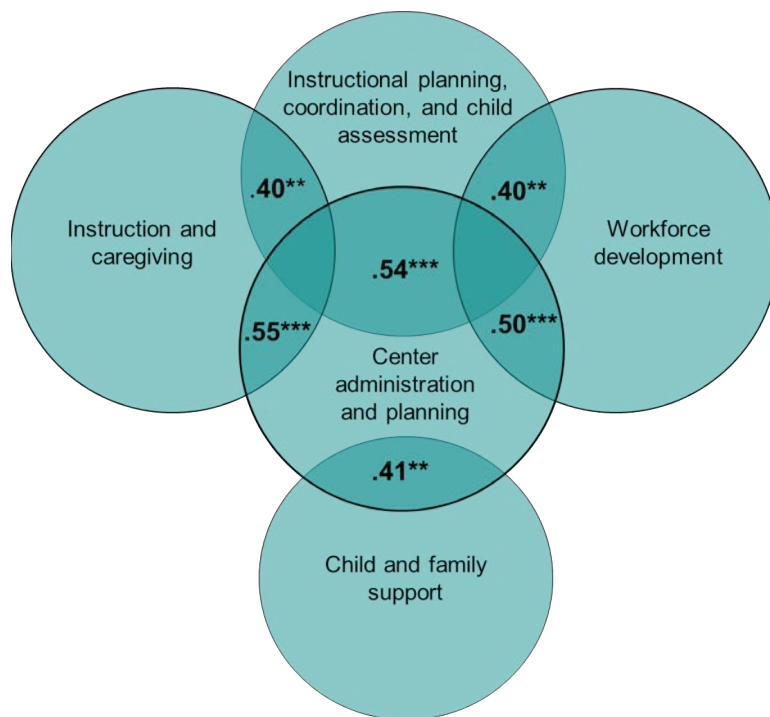
Source: Data collected for the ECE-ICHQ Project, Phase 2.

Note: The figure depicts which scores are correlated but is not reflecting the relative size of the overlap between functions.

** Significant at the 0.05 level

*** Significant at the 0.01 level

4. Variation in implementation scores with center characteristics

We examined the patterns in the summary scores for each of the key functions when compared with center characteristics. We could not conduct statistical tests for significance between the mean scores of centers with different characteristics, given the small number of centers within each characteristic subgroup.

The summary scores of the functions varied by center characteristics in ways we would expect, offering some preliminary evidence of validity for the measures (Exhibit IV.7). For example, in our sample, centers with high QRIS ratings have higher implementation scores, on average, on each key function than centers with low QRIS ratings. Similarly, not-for-profit centers have higher mean scores than for-profit centers. Centers that were

> **Implementation scores for the key functions vary with center characteristics in logical and systematic patterns.**

funded with a majority of Head Start or state pre-K funding and mixed funding centers have higher mean scores than centers that rely largely on private tuition or those that rely on subsidies. This might reflect the presence of standards that come with participation in Head Start and state pre-K programs. Centers that are more specialized in serving specific age groups also tend to have higher scores than those that serve children that range in age. For example, centers that serve only children ages 0 to 3 or those ages 3 to 5 have higher mean scores than centers that serve children from birth through school age. The patterns in scores held across each key function.

Exhibit IV.7. Implementation scores for the five key functions, by center characteristics

Center characteristics	Instruction and Caregiving mean (S.D.)	Instructional Planning, Coordination, and Assessment mean (S.D.)	Center Admin and Planning mean (S.D.)	Workforce Devel. mean (S.D.)	Child and Family Support mean (S.D.)	Sample size
QRIS rating[a]						
Low	0.507 (0.075)	0.559 (0.133)	0.341 (0.095)	0.500 (0.167)	0.388 (0.182)	10
High	0.626 (0.117)	0.669 (0.104)	0.532 (0.118)	0.640 (0.115)	0.567 (0.156)	20
Age groups of children served						
Ages 0 to 5 and school age	0.566 (0.085)	0.642 (0.100)	0.411 (0.146)	0.553 (0.172)	0.438 (0.179)	11
Ages 0 to 5 years only	0.567 (0.107)	0.600 (0.157)	0.462 (0.132)	0.591 (0.135)	0.517 (0.192)	12
Ages 0 to 3 or 3 to 5 only	0.652 (0.167)	0.671 (0.093)	0.569 (0.113)	0.659 (0.121)	0.593 (0.159)	7
Program size (based on licensed capacity)						
Small center	0.581 (0.150)	0.575 (0.118)	0.467 (0.147)	0.599 (0.121)	0.530 (0.147)	12
Large center (serving 75 children or more)	0.590 (0.096)	0.670 (0.116)	0.469 (0.144)	0.589 (0.167)	0.488 (0.208)	18

Center characteristics	Instruction and Caregiving mean (S.D.)	Instructional Planning, Coordination, and Assessment mean (S.D.)	Center Admin and Planning mean (S.D.)	Workforce Devel. mean (S.D.)	Child and Family Support mean (S.D.)	Sample size
Profit status						
For profit	0.592 (0.141)	0.586 (0.112)	0.395 (0.160)	0.508 (0.183)	0.378 (0.187)	11
Not for profit	0.583 (0.107)	0.659 (0.126)	0.511 (0.114)	0.642 (0.098)	0.583 (0.135)	19
Embedded in a larger organization or part of a chain[b]						
No	0.576 (0.132)	0.634 (0.122)	0.447 (0.163)	0.556 (0.180)	0.439 (0.196)	16
Yes	0.598 (0.104)	0.629 (0.131)	0.492 (0.115)	0.635 (0.087)	0.577 (0.144)	14
Funding						
High subsidy[c]	0.550 (0.111)	0.570 (0.100)	0.392 (0.148)	0.552 (0.171)	0.495 (0.202)	8
Majority private tuition[d]	0.553 (0.145)	0.550 (0.125)	0.468 (0.160)	0.548 (0.166)	0.384 (0.132)	8
Mixed public or mixed public/private	0.631 (0.088)	0.705 (0.064)	0.499 (0.180)	0.634 (0.171)	0.559 (0.169)	4
Majority Head Start or state pre-K[e]	0.625 (0.109)	0.718 (0.092)	0.518 (0.096)	0.646 (0.096)	0.599 (0.174)	10
Total						
All centers	0.591 (0.119)	0.629 (0.123)	0.468 (0.142)	0.593 (0.148)	0.505 (0.184)	30

Source: Data collected for the Assessing Implementation and Cost of High Quality Early Care and Education (ECE-ICHQ) Project, Phase 2.

[a] The high QRIS category includes Arkansas centers with a rating of 3, Arizona centers with a rating of 4 or 5, Pennsylvania centers with a rating of 4, and all Head Start centers. The low QRIS category includes Arkansas centers with a rating of 1, Arizona centers with a rating of 2 or 3, and Pennsylvania centers with a rating of 2.

[b] Part of a larger organization means the center is part of, and shares resources/staff with, a larger organization, such as a YMCA, university, or other social services agency.

[c] Includes centers in which 50 percent or more of the children are supported by funding from CCDF and less than 30 percent of children are supported by funding from other federal, state, or local government sources.

[d] Includes centers in which less than 30 percent of the children are supported with funding from CCDF.

[e] Includes centers in which 90 percent or more of the children are supported through private tuition paid by their parents or guardians without any public funding.

CCDF = Child Care and Development Fund.

D. Summary and next steps

With the Phase 2 analysis, we aimed to create a streamlined implementation data collection tool that will provide useful summary information on the implementation of each key function—what a center does and how—without heavily burdening respondents. The item sets in the measures for each key function demonstrate initial evidence of reliability and validity. Our analytic approach helped us to identify items that have a high value in measuring each construct, with sound statistical properties given the sample size.

We identified a smaller set of items that we recommend be the focus of further testing in a field test with a larger sample of centers. Overall, we reduced the number of items in the implementation interview by approximately 48 percent.

We will further test the psychometric properties of the implementation measures for the key functions in a field test, as described in Chapter VII. Through the field test, we will retest the measures using the processes described to determine a final, potentially smaller item set that produces reliable scales for each function. Using data from the field test, we might revise item coding based on the results of testing alternative coding approaches that could improve the item distribution or substantive meaning. In the field test, we also plan to use externally validated measures to validate the implementation measures for each key function.

V. Developing the Cost Measures

We set out to create cost measures that could delve deeper into explaining variations in quality by aligning with implementation measures for the five key functions of center-based ECE settings. Past studies have found that higher quality centers tend to have higher costs, but the association between cost and quality does not have a consistent pattern or strength across studies (Caronongan et al. 2016). Measures that go beyond total costs and work together with information about implementation could help us clarify different cost structures or trade-offs that might lead to quality.

In this chapter, we describe the cost measures we developed with data collected from centers participating in Phase 2 of the multi-case study. Not all the centers in the Phase 2 data collection could complete the cost data collection or provide high quality data to support the analysis. As a result, we built the measures using data from 25 of the 30 centers included in Phase 2. We first define the cost measures and describe our method for estimating each of them. We then present findings about the measures, such as average costs across functions and by center characteristics. We end the chapter with a brief discussion of next steps.

A. Defining the cost measures

We designed the cost measures to describe (1) the cost of providing ECE to children ages 0 to 5 in a center and (2) how a center uses resources to deliver ECE services. We specified four cost measures in the multi-case study:

> **Absolute** cost measures capture how much ECE services cost (in dollars).
>
> **Relative** cost measures capture how centers use their resources to deliver ECE services (as percentages of total costs).

1. **Cost per child care hour.** This measure reflects the cost to provide care for one child for one hour. It is the total cost (in dollars) for a center to provide services to children ages 0 to 5 during a one-year period divided by the total hours of care the center provides to children ages 0 to 5 during that year.

2. **Cost per child care hour by function.** This measure reflects the cost to provide care and carry out other activities related to each key function for one child for one hour (in dollars).

3. **Cost allocation by resource category.** This measure captures how a center's costs are allocated to resource categories. This measure is the percentage of a center's annual costs allocated to each of eight resource categories.

4. **Cost allocation by function.** This measure captures how a center's costs are allocated to each key function. This measure is the percentage of a center's annual costs allocated to each of the key functions specified in the conceptual framework.

The first two measures are absolute measures that reflect cost in dollars that support the center's activities as a whole or for a specific function. These absolute measures convey the magnitude of investments.

The second two measures—as percentages—are relative measures that show how much centers spend on one resource or function relative to another. They can capture trade-offs that centers make to balance costs within available resources, but they do not tell us about the size of the investments in dollars. Exhibit V.1 provides a description of the cost measures.

Exhibit V.1. Cost measures

Measure		Description of the measure	How the measure is calculated
Cost per child care hour	$	The cost to provide care for one child for one hour	Total cost for a center to provide services during a one-year period divided by the total hours of care the center provided during that period
Cost per child care hour by function	$	The cost to provide care and carry out activities related to each key function for one child for one hour	Total cost for a center to provide services related to each key function during a one-year period divided by the total hours of care the center provided during that period
Cost allocations by resource category	%	The proportion of a center's total cost spent on each of the resource categories	The total annual cost for each resource category divided by total annual cost
Cost allocations by function	%	The proportion of a center's total cost spent on each key function	Total annual cost for each key function divided by total annual cost

B. Estimating the cost measures

We estimated measures using data collected from centers participating in the multi-phase case study. In this section, we describe the steps we took to develop the measures.

We collected three essential data elements through the cost workbook and time-use survey to produce the cost measures for each center. The elements included cost, enrollment, and time-use data:

- **Cost data.** Respondents provided information in the cost workbook on all the resources they used to provide services to children 0 to 5 over a one-year period.

- **Enrollment data.** Respondents provided information in the cost workbook on the number of children they served and the number of hours they provided care during the one-year period.

- **Time-use data.** Staff at each center completed the time-use survey to provide information on how they spent their time.

We estimated center-level costs by using a bottom-up method that calculates total costs based on the value of the ingredients or resources necessary to provide services. The first step involved identifying the type and quantity of resources that a center uses to deliver its services over a certain period of time—resources such as personnel, facilities, supplies and materials, and so on—as collected in different tabs of the cost workbook (see Exhibit III.5 in Chapter III).

The second step was to assign a dollar value to each resource identified, either based on the amount reported in the cost workbook or by estimating the value. For some categories, we adjusted reported costs or estimated the value of resources received in kind. For example, we

> **Estimating center-level costs**
>
> Calculate costs using the "ingredients" or resource cost method:
>
> 1. Identify the type and quantity of resources used to deliver a program or service
>
> 2. Determine the monetary value of these resources
>
> 3. Sum values to estimate total cost
>
> Allocate costs to key functions based on:
>
> 1. Coding of line items or resources
>
> 2. Staff time use

calculated the annual *value of equipment* purchased and facilities improvements completed during the cost period using their estimated useful life. We also estimated the *cost of facilities received in kind* using

information on the size of the facility in square feet and information we gathered on the average cost per square foot to lease a similar space in the local area. We *adjusted local resource prices* to national equivalents to account for geographic differences in wages and the cost of other resources. We used an adjustment factor created from state-level and national Bureau of Labor Statistics data by geographic area and occupation.

The third and last step was to sum all the values for resources in the cost workbook to produce an estimate of total costs.

Total cost per child care hour. To estimate the cost per child care hour, we combined the estimates of total annual cost with data on the number of hours of care the center provided during the year. We calculated the average number of child care hours the center provided each week by multiplying the total number of children enrolled in the center by the average number of hours of care children received per week. We calculated the total number of child care hours the center provided during the one-year period by multiplying the average number of hours of care the center provided during a week by the number of weeks the center was open. To calculate the cost per child care hour, we divided a center's total annual cost by the total number of child care hours the center provided during the year.

Cost allocations by resource category. To calculate allocations by resource category, we estimated the proportion of costs used for each resource category. The cost workbooks, which were organized by resource category (that is, there was a separate worksheet for each resource category as shown in Exhibit III.5 in Chapter III), contained all the information we needed to calculate resource category-specific cost estimates. We summed the costs of the individual resources reported for each resource category to calculate eight separate amounts, one for each resource category. (Like the total cost estimates, these estimates reflect adjusted costs and the estimated value of resources received in kind). We then calculated the percentage of a center's annual costs allocated to each category by dividing the total cost for each resource category by total annual cost.

1.   Cost measures by function

For the ICHQ cost measures, we took the additional step of allocating costs to the key functions based on how staff use their time in delivering services and how resources are used.

Allocating personnel costs to key functions based on time use. Personnel costs comprise the majority of costs in center-based ECE settings. We wanted to be able to allocate personnel costs across the five key functions to produce finer-grained measures of costs than total costs or costs by resource category. We used staff time-use surveys to collect the data. To calculate time-use allocations, we grouped activities in the time-use survey related to each key function (presented in Exhibit III.6 in Chapter III). We provide detailed descriptions of the activities included in the time-use survey for each key function in Appendix B.

For each time-use respondent, we calculated the total number of hours spent on each activity in a year. For activities we expected staff to engage in on a regular basis, such as lesson planning, we asked staff to report their time use in terms of weekly hours and converted those weekly hours to a total for the year. For activities we expected respondents to participate in less frequently, such as attending workshops, we asked staff to report their time use in terms of yearly hours. To determine the total hours across activities for each key function for each respondent, we combined weekly and yearly hours across relevant

activities. We used the sum across all key functions as the denominator.[9] We then calculated a time allocation for each key function, based on the proportion of hours spent on each function and the sum of weekly and yearly hours across all functions for an individual respondent.

$$Allocation\ for\ key\ function\ A = \frac{Hours\ across\ activities\ for\ key\ function\ A}{Total\ hours\ across\ all\ key\ functions}$$

The sum of weekly and yearly total hours across all key functions may exceed the total number of hours a staff member actually works in a year. The time-use survey includes instructions that the total number of hours a respondent reports for weekly activities should equal the respondent's scheduled hours in a typical week. Even so, in Phase 2, the median total number of hours reported across all key functions was 2,989, and the mean was 3,692. For comparison, the total number of working hours for a full-time staff member (working 40 hours per week and 52 weeks per year) would be 2,080. We took the approach of summing all reported hours because the relative proportions of hours for each key function are likely to be a reasonable estimate of staff members' time-use allocations, even when the total number of working hours reported might not be precise.

We averaged time-use allocations across all teaching staff in a center (head or lead teachers, teachers, and assistant teachers). We took this approach because it was difficult in some centers to match time-use data and cost data for teaching positions, due to differences between time-use survey respondents' self-reported titles and job positions and the titles and job positions reported in the cost workbook (which was completed by a different respondent). To test the sensitivity of the cost measures to using combined or separate time-use allocations for teaching positions in each center, we examined whether and how cost allocations by function changed under each approach. The estimates of average allocations to functions for larger groups of centers were not sensitive to the choice to combine or separate time-use percentages across teaching positions. However, the different approaches may affect function allocations or rankings at the center level. To improve accuracy at the center level, we plan to refine the data collection approach to make it feasible to produce and use separate time-use allocations for head or lead teachers and assistant teachers in the future. Specifically, in addition to entering center-specific titles or positions in the cost workbook, we are asking respondents to indicate a job category for each staff member that aligns with the job categories in the time-use survey.

> **Allocating personnel costs to key functions**
>
> Assigned based on reported time use for specific positions such as directors, assistant directors, and education coordinators
>
> Used average time allocations reported across all teaching positions to assign costs to key functions for all teaching staff
>
> Used titles and position descriptions to map other professional, administrative, and operations staff to key functions

We also administered the time-use survey to key administrators in the center, including center directors, assistant directors, and education program leads. We did not collect time use for other professional, administrative, or operations staff. For these staff positions, we allocated personnel costs to key functions

---

[9] We explored an alternative approach in which we subtracted out a proportion of hours to account for intermittent work displacing regular work during parts of the year. This approach did not make a substantial difference in estimates because, overall, hours spent on intermittent work only account for a small proportion of total hours in a year.

based on their titles and position descriptions in the cost workbook. For example, we assigned the costs of family support specialists to the Child and Family Support function.

Allocating nonpersonnel costs to key functions. We allocated nonpersonnel resources to key functions based on the purpose or description of the service or item purchased or received. Our method for allocating nonpersonnel costs to key functions varied by resource category (see Exhibit V.2). We allocated all costs in one resource category—staff training and education—to one function, the Workforce Development function. We tried in earlier stages of the multi-case study to break down training by topic, but it was too burdensome on the centers to provide the level of detail needed. More commonly, we allocated line items to specific functions based on the item descriptions. For example, we allocated classroom supplies to Instruction and Caregiving and licensing fees, insurance, and taxes to Center Administration and Planning. Last, we allocated resources that support multiple functions, such as facilities and overhead, across functions using the same proportions as those we applied to personnel costs based on the staff time-use survey.

Exhibit V.2. Methods for allocating nonpersonnel costs to key functions

Resource category	Allocation method
Staff training and education	Allocate all costs to Workforce Development function
Contracted services	Allocate line items to functions based on descriptions
Facilities	Allocate across functions in same proportions as salaries and fringe benefits
Supplies and materials	Allocate line items to functions based on descriptions
Equipment	Allocate line items to functions based on descriptions
Miscellaneous	Allocate line items to functions based on descriptions
Overhead or resources from a larger organization	Allocate across functions in same proportions as salaries and fringe benefits OR Allocate to key functions based on descriptions of individual resources provided by the larger organization

Cost allocations by key function. To calculate allocations by key function, we estimated the proportion of costs used for each key function using the methods described above. We summed the costs of the individual resources reported for each key function and calculated the percentage of a center's annual costs allocated to each key function by dividing the total cost for each key function by total annual cost.

Cost per child care hour by key function. We used the percentage of costs allocated to each key function to calculate a dollar value of the cost of child care hour by key function. We multiplied the total cost per child care hour by the percentage allocation for each key function to determine the cost per child care hour for each of the five key functions.

C. Findings about staff time use

In this section, we summarize the findings from our analysis of the time-use data. The sample for the time-use analysis includes selected staff from all 30 centers that participated in Phase 2. The average number of respondents per center was 14 but ranged from 3 to 45 based on the size of the center. We included a total of 421 surveys in the analysis; 359 were completed by teaching staff. Appendix B includes tables that present the mean and median time-use allocations for each key function, by staff role.

Time-use allocations by staff position. On average, center staff reported spending their time in expected ways. Center directors and teacher-directors (who perform both teaching and administrative duties in a location without a center director) spent most of their time (44 percent) on the Center Administration and Planning function (Exhibit V.3). Educational directors or coordinators spent about one-third of their time on the Instructional Planning, Coordination, and Child Assessment function, and another third on the Workforce Development function. They also spent about a quarter of their time, on average, on the Center Administration and Planning function. Teacher-directors spent the majority of their time between the Instruction and Caregiving function (23 percent, on average) and the Instructional Planning, Coordination, and Child Assessment function (31 percent, on average). They also spent about one-third of their time, on average, on the Center Administration and Planning function.

> **Staff time use and key functions**
>
> Directors reported spending more of their time on Center Administration and Planning than other staff.
>
> Lead teachers, head teachers, or co-teachers reported spending more of their time on Instructional Planning, Coordination, and Child Assessment relative to assistant teachers, who reported spending more of their time on Instruction and Caregiving.

Exhibit V.3. Percentage of total working hours spent on each key function by center administrators
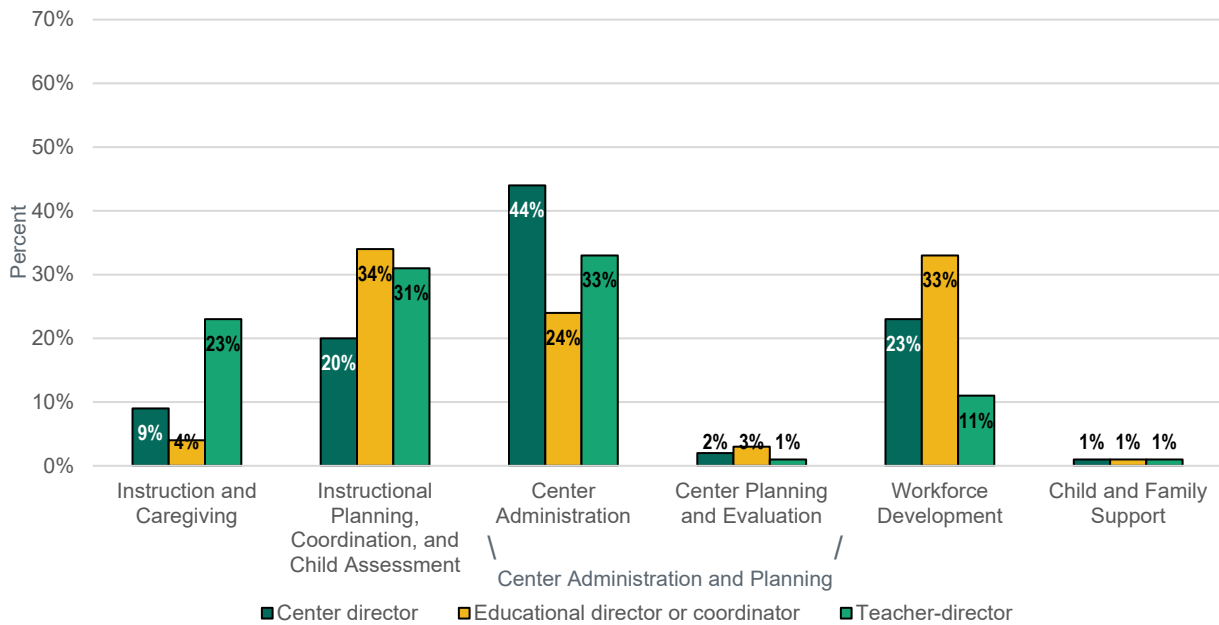
Source: Data collected for the Assessing Implementation and Cost of High Quality Early Care and Education (ECE-ICHQ) Project, Phase 2.

Note: Center Administration and Center Planning and Evaluation were later combined into the Center Administration and Planning function.

Teaching staff reported spending most of their time on the Instruction and Caregiving and Instructional Planning, Coordination, and Child Assessment functions. There were minimal differences in how different teaching staff—head or lead teachers, teachers, or assistant teachers—spent their time (Exhibit

V.4). Of note, however, is that head or lead teachers and teachers spent more time on the Instructional Planning, Coordination, and Child Assessment function and less on the Instruction and Caregiving function, compared to assistant teachers for whom the pattern is opposite.

Exhibit V.4. Percentage of total working hours spent on each key function by teaching staff

Source: Data collected for the Assessing Implementation and Cost of High Quality Early Care and Education (ECE-ICHQ) Project, Phase 2.

Note: Center Administration and Center Planning and Evaluation were later combined into the Center Administration and Planning function.

Time-use allocations by center characteristics. We see some variation in patterns of staff time use among subgroups of centers. For example, teaching staff in centers funded primarily by Head Start or state pre-K funds tend to spend less time on Instruction and Caregiving and more time on Instructional Planning, Coordination, and Child Assessment than teaching staff in centers that fall into the other funding categories (high subsidy, private tuition, or mixed public or mixed public/private funding) (Exhibit V.5). We did not see much variation in reported time use for key administrators or teaching staff between centers that had a high QRIS rating and those that had a low QRIS rating.

Exhibit V.5. Percentage of total working hours spent by all teaching staff on each key function, by funding source

Source: Data collected for the Assessing Implementation and Cost of High Quality Early Care and Education (ECE-ICHQ) Project, Phase 2.

Notes: This figure presents time combined for all teaching staff including head, lead, or co-teachers, teachers, and assistant teachers. Center Administration and Center Planning and Evaluation were later combined into the Center Administration and Planning function. High subsidy includes centers in which 50 percent or more of children are supported through funding from CCDF and less than 30 percent of children are supported from other federal, state, or local government sources. Mixed funding includes centers in which the majority of the children are not supported by any of the other sources and less than 30 percent of children are supported through child care subsidies.

D. Findings about cost measures

Patterns we observe in the cost measures suggest that they function as anticipated. However, we observed more variation among subgroups of centers in absolute cost measures than in relative cost measures. The absolute cost measures of total cost per child care hour and cost per child care hour by key function (in dollars) are able to make distinctions between centers with different characteristics. The relative measures (that report allocations as percentages) do not show noticeable differences between most subgroups of centers.

We built the cost measures using data from 25 centers that provided complete data in the cost workbook to cover the essential categories of costs.[10]

> **Findings from the cost measures**
>
> - The mean total cost per child care hour was $5.47, with a range of $0.92 to $10.77.
> - The total cost per child care hour varied between centers based on QRIS rating, ages of children served, profit status, and funding, but not program size.
> - Staff compensation and facilities costs made up nearly 80 percent of total costs, on average.
> - Centers spent the largest portion of costs (about one-third, on average) on the Instruction and Caregiving function.
> - The way costs are spread across the key functions varied by funding mix but not by other center characteristics.

Total cost per child care hour. We observed notable differences in the total cost per child care hour among key subgroups of centers. The mean cost per child care hour among all centers was $5.47, with a range of $0.92 to $10.77. The mean cost per child care hour was higher, on average, among centers with high QRIS ratings than among centers with low QRIS ratings (Exhibit V.6). This finding is in line with previous studies that observed higher costs in centers of higher quality (Caronongan et al. 2016). The average cost per child care hour was lower for centers that served school-age children in addition to children ages 0 to 5. Costs per child care hour did not differ markedly by program size.

Exhibit V.6. Costs per child care hour vary by quality rating level and center characteristics
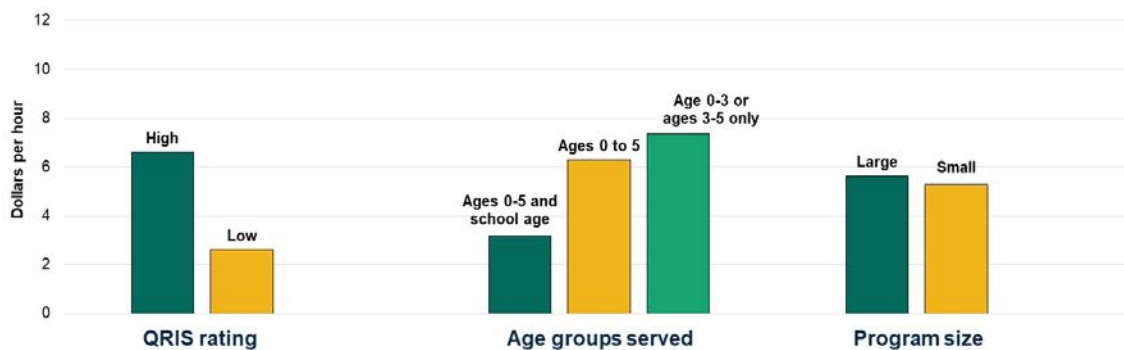
Source: Data collected for the Assessing Implementation and Cost of High Quality Early Care and Education (ECE-ICHQ) Project, Phase 2.

a We categorized centers into high and low QRIS ratings based on the requirements for the different levels in each state in the multi-case study.

b Program size is based on licensed capacity. Large centers are those that can serve 75 children or more; small centers can serve less than 75 children.

The mean cost per child care hour was higher among not-for-profit centers, compared to for-profit centers; among centers embedded in a larger organization, compared to centers that were not; and among

---

[10] We collected complete cost data from 26 centers, but we excluded from analysis one center that had a cost per child care hour that was more than three standard deviations above the mean. Appendix B includes detailed tables of results for the cost measures that both include and exclude this center outlier.

centers with majority Head Start or state pre-K funding, compared to centers funded through all other sources or funding mixes (see Exhibit V.7).

Exhibit V.7. Costs per child care hour vary by profit status and funding
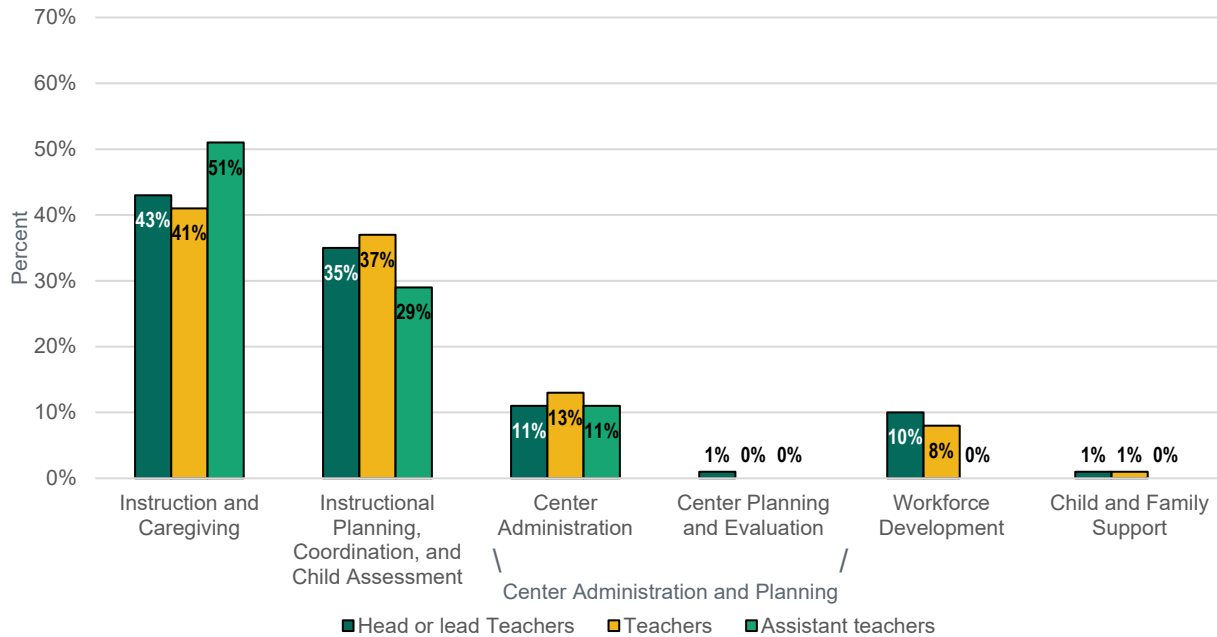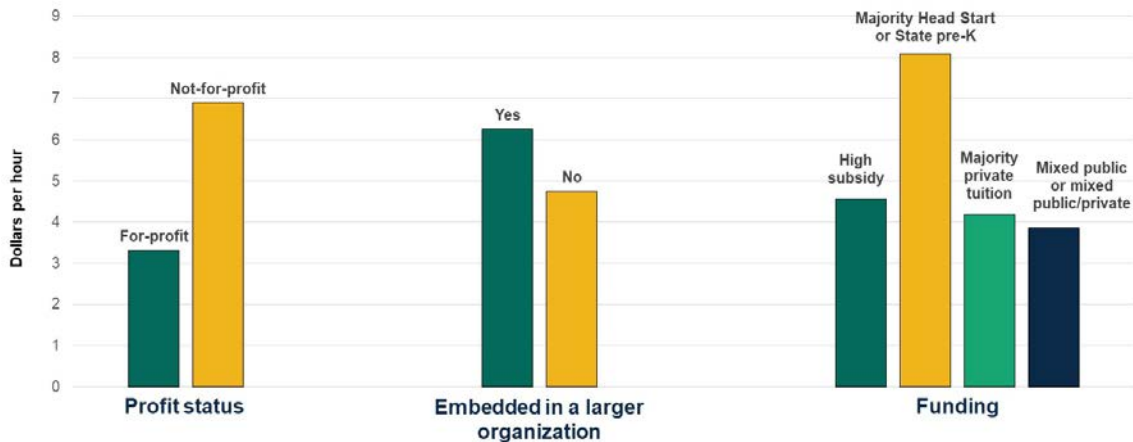

Source: Data collected for the Assessing Implementation and Cost of High Quality Early Care and Education (ECE-ICHQ) Project, Phase 2.

Notes: High subsidy includes centers in which 50 percent or more of children are supported through funding from CCDF and less than 30 percent of children are supported from other federal, state, or local government sources. Mixed funding includes centers in which the majority of the children are not supported by any of the other sources and less than 30 percent of children are supported through child care subsidies.

Cost allocations by resource category. As expected, centers invest in staff compensation (salary and fringe benefits) as the largest proportion of costs (62 percent, on average) (Exhibit V.8). Facilities and supplies and materials account for the next largest proportions of costs (16 percent and 9 percent, respectively, on average). The remaining resource categories comprise small proportions of center costs such as contracted services (4 percent), overhead costs (3 percent), and other direct/miscellaneous costs (3 percent). Just one percent of costs, on average, goes to staff training and education and to equipment.
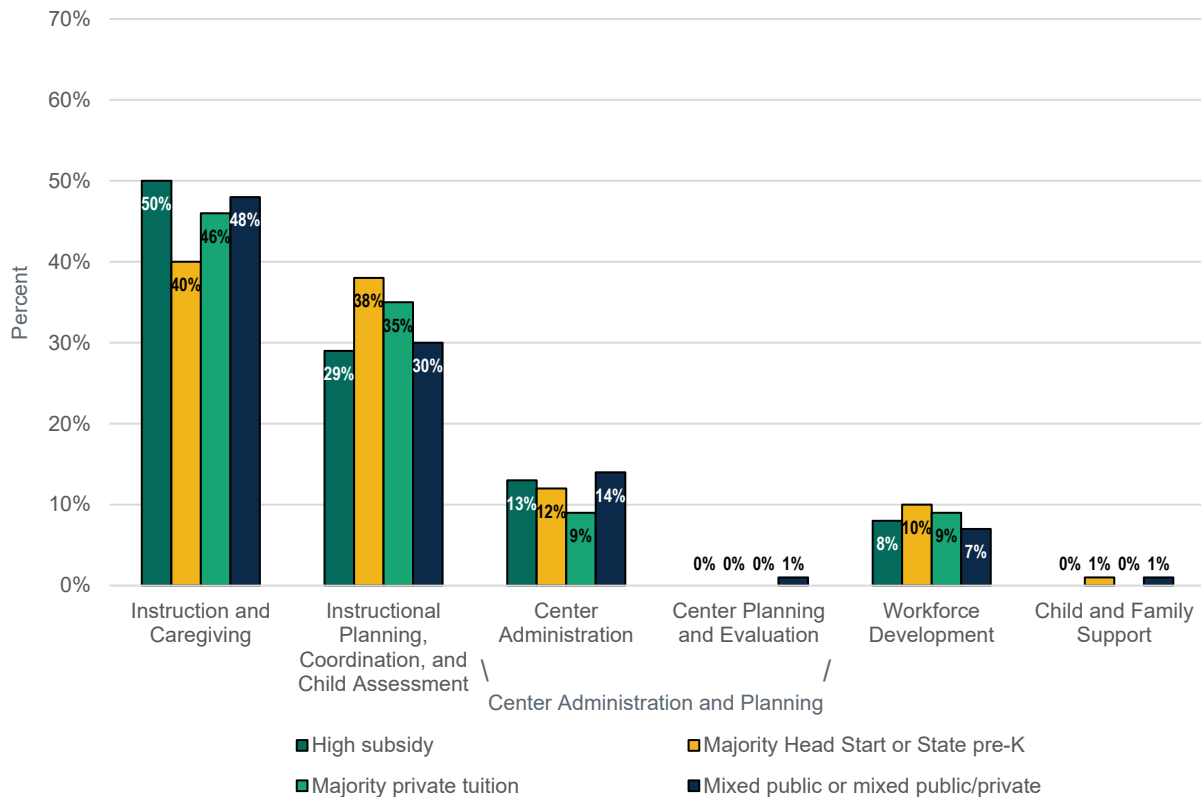
Exhibit V.8. Cost allocations by resource category

Source: Data collected for the Assessing Implementation and Cost of High Quality Early Care and Education (ECE-ICHQ) Project, Phase 2.

We see some differences in the cost allocations among resource categories by center characteristics. For example, high QRIS-rated centers invested more in staff compensation, on average, than low QRIS-rated centers, suggesting that high-rated centers may pay their staff more than low-rated centers (Exhibit V.9). Centers with different funding mix also have different patterns in their cost allocations across the main resource categories, reflecting the trade-offs centers make in how they use available resources (Exhibit V.10).

Exhibit V.9. Cost allocations by resource category, by QRIS rating
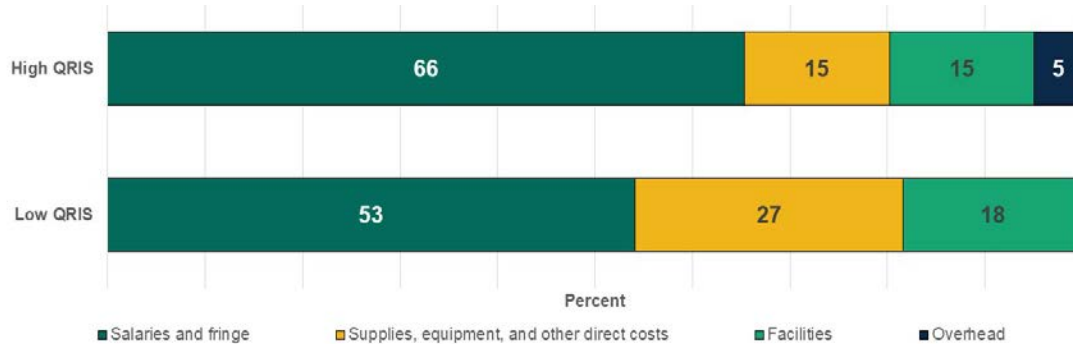
Source: Data collected for the ECE-ICHQ Project, Phase 2.

Notes: Percentages may not sum to 100 due to rounding. Figures for subgroups are means. Supplies, equipment and other direct costs includes the following resource categories: supplies and materials; contracted services; staff training and education; equipment; and miscellaneous. We categorized centers into high and low QRIS ratings based on the requirements for the different levels in each state in the multi-case study.

Exhibit V.10. Cost allocations by resource category, by funding source
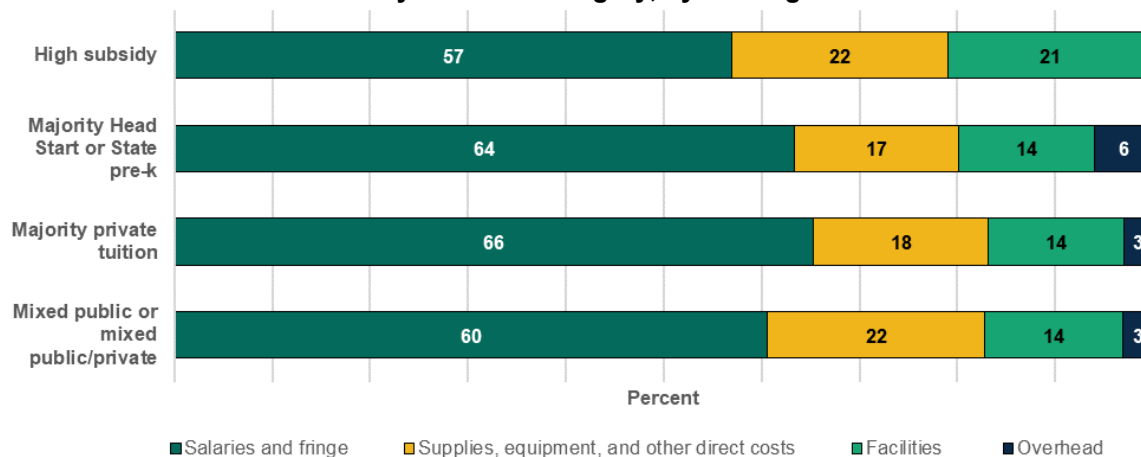
Source: Data collected for the ECE-ICHQ Project, Phase 2.

Notes: Percentages may not sum to 100 due to rounding. Figures for subgroups are means. Supplies, equipment and other direct costs includes the following resource categories: supplies and materials; contracted services; staff training and education; equipment; and miscellaneous. We categorized centers into high and low QRIS ratings based on the requirements for the different levels in each state in the multi-case study.

Cost allocations by key function. On average, the centers allocated 33 percent of costs to the Instruction and Caregiving function and 22 percent of costs to the Instructional Planning, Coordination, and Child Assessment function (Exhibit V.11). The centers spent just over one-quarter of their total costs (26 percent, on average) on the Center Administration and Planning function. Spending on the Workforce Development function accounted for 8 percent of total costs, on average, and spending on the Child and Family Support function accounted for 10 percent, on average.

Exhibit V.11. Cost allocations by key function
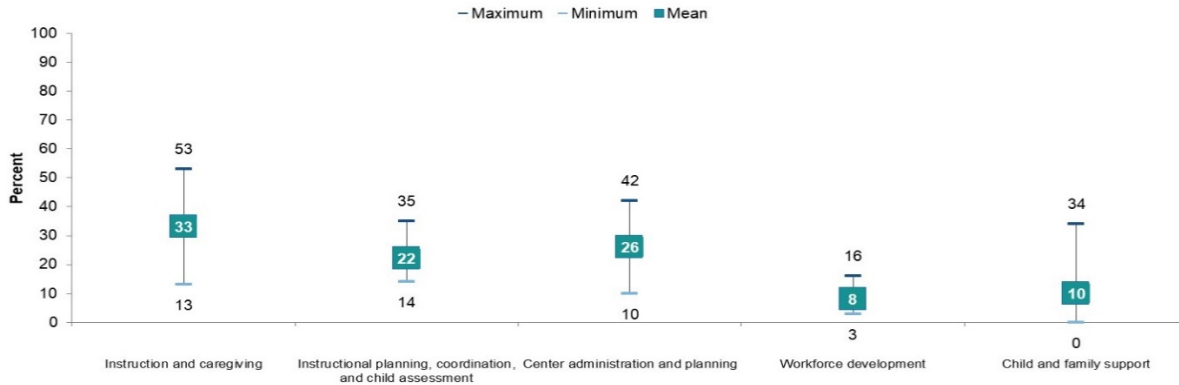
Source: Data collected for the Assessing Implementation and Cost of High Quality Early Care and Education (ECE-ICHQ) Project, Phase 2.

We hypothesized that there would be differences between centers in how they spent their resources on different functions. We found, however, that the way that centers spent resources across the key functions did not differ much by most center characteristics. For example, we do not see much difference in the cost allocations across the functions between high-rated QRIS centers and low-rated QRIS centers (Exhibit V.12).

Exhibit V.12. Percentage of costs allocated to each of the key functions, by QRIS rating level
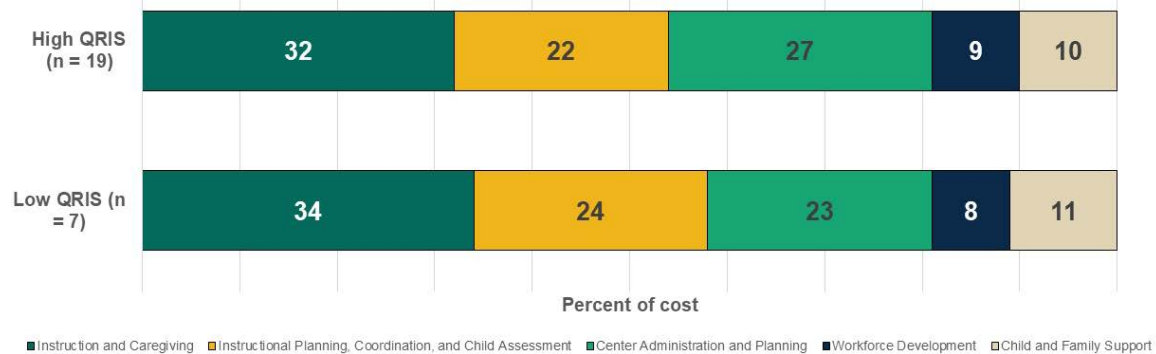
Source: Data collected for the Assessing Implementation and Cost of High Quality Early Care and Education (ECE-ICHQ) Project, Phase 2.

Note: We categorized centers by high and low QRIS ratings based on the requirements for the different levels in each state in the multi-case study.

We did see some differences in spending across the key functions between centers by funding mix. Centers funded primarily through Head Start and state pre-K tended to spend relatively equal proportions of their total costs on the Instruction and Caregiving function and on the Instructional Planning, Coordination, and Child Assessment function (Exhibit V.13). In other types of centers, a larger proportion of costs was spent on the Instruction and Caregiving function compared to the Instructional Planning Coordination, and Child Assessment function. Also, we see that centers that relied on private tuition spent 5 percent of their costs, on average, on the Child and Family Support function compared to the minimum of 10 percent spent in the centers with other funding. This finding might reflect the differences in the characteristics and needs of children served across the centers.

Exhibit V.13. Percentage of costs allocated to each of the key functions, by funding

Funding	Instruction and Caregiving	Instructional Planning, Coordination, and Child Assessment	Center Administration and Planning	Workforce Development	Child and Family Support
High subsidy (n = 6)	35	18	27	7	13
Majority Head Start or state pre-K (n = 8)	28	24	28	8	12
Majority private tuition (n = 7)	38	27	20	10	5
Mixed public or public/private (n = 4)	32	19	31	8	10

Percent of cost

■Instruction and Caregiving ▫Instructional Planning, Coordination, and Child Assessment ■Center Administration and Planning ■Workforce Development ▫Child and Family Support
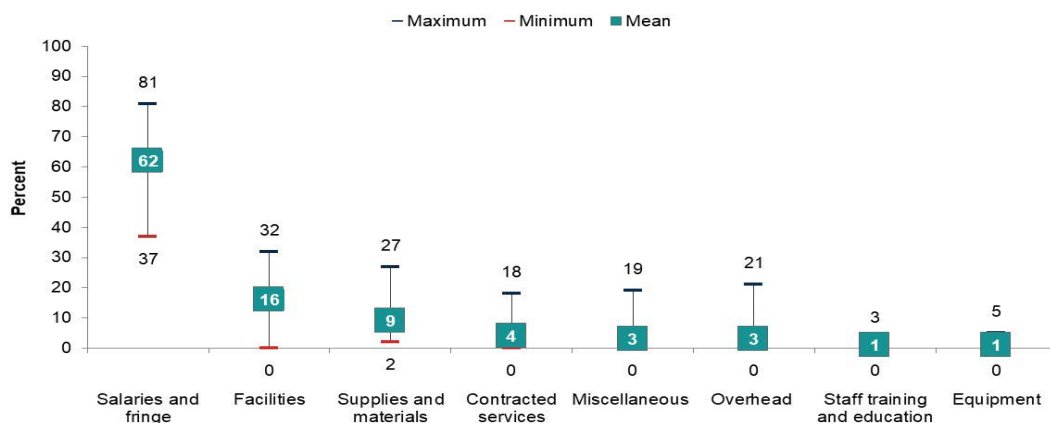
Source: Data collected for the Assessing Implementation and Cost of High Quality Early Care and Education (ECE-ICHQ) Project, Phase 2.

Note: High subsidy includes centers in which 50 percent or more of children are supported through funding from CCDF and less than 30 percent of children are supported from other federal, state, or local government sources. Mixed funding includes centers in which the majority of the children are not supported by any of the other sources and less than 30 percent of children are supported through child care subsidies.

Cost per child care hour by function. This measure captures the use of resources across the key functions and provides information about the size of the investment in each function. Although we saw little variation in the cost allocations by key function, the total cost per child care hour varied substantially across centers. This variation in the total costs—or the investments that centers make—carries through into the cost per child care hour by function. As with the total cost per child care hour, we see variation in the cost per child care hour by function between centers that vary by the ages of children served, profit status, funding mix, and QRIS rating. For example, centers with a high QRIS rating spend more than twice as much per hour on each function than centers with a low QRIS rating (Exhibit V.14). Similarly, centers that serve children ages 0 to 5 spend about twice as much per hour on each function than centers that serve children ages 0 to 5 and school-age children; not-for-profit centers spend about twice as much per hour on each function than for-profit centers; and centers with majority Head Start or state pre-K funding spend about twice as much per hour on each function than centers funded through all other sources or funding mixes. As with the total cost per child care hour, there is less variation in the cost per child care by function between centers that vary by state, program size, and whether the center is embedded in a larger organization.

**Exhibit V.14. Cost per child care hour by function, by QRIS rating level**
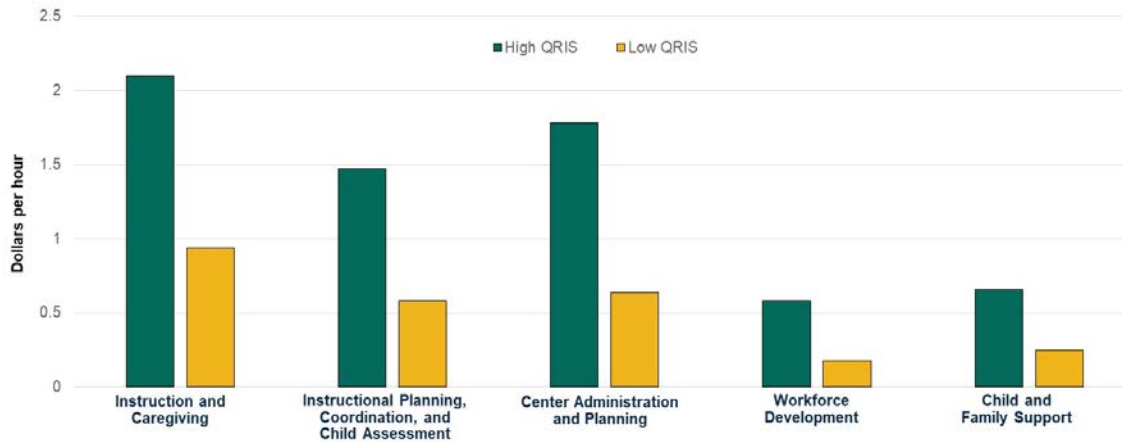
Source: Data collected for the Assessing Implementation and Cost of High Quality Early Care and Education (ECE-ICHQ) Project, Phase 2.

Note: We categorized centers by high and low QRIS ratings based on the requirements for the different levels in each state in the multi-case study.

E. Summary and next steps

Based on the analysis and our findings about the cost measures, we are focusing on three refinements to cost data collection and measures development for a field test.

Calculating facilities costs. To ensure the cost measures are consistent and comparable across all centers, we included the estimated value of facilities that centers received in kind. A concern with this approach is that costs for centers that receive facilities in kind might differ systematically from the costs for centers that do not. For example, the costs for centers that receive facilities in kind might be higher if these centers tend to use larger facilities than they would if they paid for space. We looked at the value of the total cost per child care hour and the cost allocations for centers that received a subsidy for facilities costs through in-kind donations or other supports (9 centers) and centers that did not receive a subsidy for facilities (17 centers). We did not find differences and concluded that our approach to estimating the value of facilities received in kind is not likely to drive differences in the cost measures between centers. Because facilities costs represent a relatively large proportion of center costs, on average, we plan to continue with this approach that supports consistency and comparability in estimating costs across centers.

In this analysis, we could not distinguish between spaces used only for children ages 0 to 5 and spaces used only for, or shared with, school-age children (for example, space used for before- and after-school care). Estimates of facilities costs may include the value of space used for the care of school-age children. We have revised data collection for a field test to collect information about the space used to care exclusively for school-age children, so that we can adjust to focus on the cost of facilities used to care for children ages 0 to 5.

Allocating teaching staff time to key functions. We are taking steps to support assigning personnel costs more precisely to key functions so the cost measures can better reflect different teaching staff structures and pay scales across centers. For example, we revised the job titles and position descriptions included in the time-use survey and cost workbook to improve how they fit together so we can calculate

time use by specific teaching positions. Specifically, we will estimate time-use allocations separately for (1) lead teachers, head teachers, and co-teachers, (2) assistant teachers, aides, and teaching assistants, and (3) floaters and substitute teachers.

Producing costs per child care hour by age groups of children served. To better understand cost patterns across centers, and potential cost shifting across the ages of children served, we will estimate a new measure to capture the cost per child care hour by age group. We will estimate the cost per child care hour for all children ages 0 to 5, and we will develop separate estimates for infants (ages 0 to under 18 months), toddlers (ages 18 months to under 36 months), and preschoolers (ages 3 to 5). These new estimates may highlight differences in costs by age group served. We will test five measures in the field test: (1) cost per child care hour, (2) cost per child care hour by age group, (3) cost per child care hour by key function, (4) cost allocations to resource categories, and (5) cost allocations to key functions.

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

VI. Exploring the Relationship between Implementation, Cost, and Quality

A key goal of the ECE-ICHQ project is to produce measures that can clarify the relationship between implementation and cost, to help us better understand how centers can achieve high quality. We examined associations between implementation scores for each key function and each of our cost measures to explore whether they are related in ways we would expect. We also explored whether the implementation and cost measures provide unique information beyond what one might expect based on a center's QRIS rating or other center characteristics. We end this chapter with a discussion of refinements we might need to make to the implementation and cost measures that could improve the way they work together and inform pathways to quality.

A.   The relationship between implementation and cost measures

We built the implementation and cost measures around the five key functions to capture how centers work to achieve quality. If we specified measurement well for each of the five functions, we would expect to see a relationship between the implementation measures that capture what a center does and how, and the cost measures that capture total costs and the use of resources. We examined the relationship between the implementation measures and each of the different cost measures through correlations that assess the degree to which they vary together. In the sections that follow, we discuss, in turn, the relationship between the implementation score for each of the five key functions and (1) the total cost per child care hour, (2) cost allocations by key function, and (3) cost per child care hour by key function.
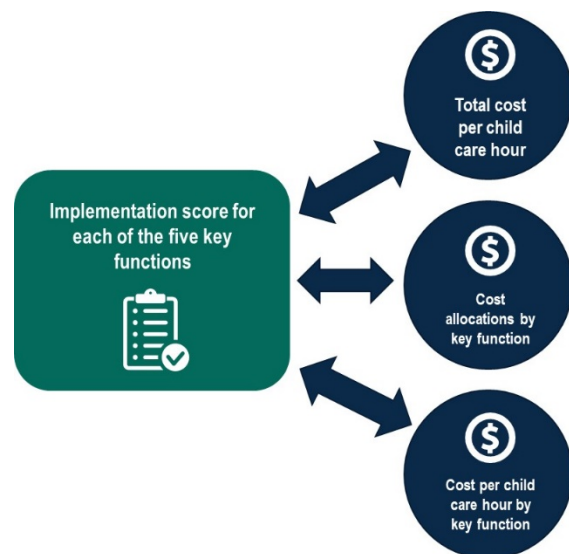
1.   Implementation scores and total cost per child care hour

Centers with higher implementation scores tend to have higher total costs per child care hour. We found positive, significant associations between four of the five implementation scores and total cost per child care hour (Exhibit VI.1). These findings suggest that centers with higher levels of implementation for each key function (based on the practices they pursue—what they do; and the intentionality with which they pursue them—how they do it) tend to have higher costs. We would also expect to see a positive, significant association between the implementation score and the total cost per child care hour for the Instruction and Caregiving function. The lack of association suggests we need to improve implementation measurement of structural features of care—like staff-to-child ratios and staff qualifications—to capture differences in costs across centers. In addition, the costs of structural features of care cannot solely be captured in the Instruction and Caregiving function; these costs are spread across the five functions through labor costs associated with staff time in each function. This might also contribute to a weaker association between the implementation and cost measures for this function.

Exhibit VI.1. Correlations between implementation scores and total costs per child care hour

Implementation scores	Total cost per child care hour
Instruction and Caregiving score	0.24
Instructional Planning, Coordination, and Child Assessment score	0.43**
Center Administration and Planning score	0.47**
Workforce Development score	0.52***
Child and Family Support score	0.46**

Source: Data collected for the ECE-ICHQ Project, Phase 2.

Note: Sample size is 25 centers; cost data collection was completed with 26 centers in the total sample, one outlier is not included.

 ** Significant at the 0.05 level.

*** Significant at the 0.01 level.

2.   Implementation scores and cost allocations by key function

We found few associations between the implementation scores for each of the five key functions and the cost allocations by key function (meaning the percentage of total costs that centers spend on each function). Implementation scores for four of the five functions are positively associated with the cost allocation to the Center Administration and Planning function but generally not to the cost allocations of other key functions (Exhibit VI.2). This suggests that centers with higher implementation scores on each function, except Workforce Development, tend to have a substantial portion of their costs allocated to the Center Administration and Planning function. Only for the Center Administration and Planning function are the implementation score and the cost allocation *to the same function* positively, and significantly, associated; we did not find relationships between the implementation score and cost allocations for the other four functions.

Exhibit VI.2. Correlations between implementation scores and cost allocation by key function

Implementation scores	Cost allocation by key function (percent of total costs)				
	Instruction and Caregiving	Instructional Planning, Coordination, and Child Assessment	Center Administration and Planning	Workforce Development	Child and Family Support
Instruction and Caregiving	-0.16	-0.21	**0.36***	0.11	-0.07
Instructional Planning, Coordination, and Child Assessment	**-0.48***	-0.02	**0.55***	0.10	0.005
Center Administration and Planning	-0.16	-0.15	**0.46***	0.01	-0.17
Workforce Development	-0.20	-0.16	0.24	0.33	-0.006
Child and Family Support	-0.03	-0.34	**0.44***	-0.12	-0.13

Source: Data collected for the ECE-ICHQ Project, Phase 2.

Note: Sample size is 25 centers; cost data collection was completed with 26 centers in the total sample, one outlier is not included. Green, bold font denotes statistically significant correlations.

  * Significant at the 0.10 level.

 ** Significant at the 0.05 level.

*** Significant at the 0.01 level.

We do see some negative correlations between the implementation scores and cost allocations in Exhibit IV.2. Most of the negative correlations are not statistically significant, but a few are sizable in magnitude. One negative correlation is statistically significant. Centers with higher scores on the Instructional Planning, Coordination, and Child Assessment function tend to have lower percentages of costs allocated to the Instruction and Caregiving function.

The cost allocations are relative measures, reflecting the idea that centers face trade-offs in how they make use of available resources. Therefore, the correlations between cost allocations and implementation should not be interpreted as the relationship between what a center spends and the level of implementation. Rather, they show how the percentage of costs that centers allocate to certain key functions *relative to others* is related to the level of implementation. For example, the negative correlation between the Instructional Planning, Coordination, and Child Assessment implementation score and the percentage of costs allocated to the Instruction and Caregiving function might suggest shifts in how teachers use their time and how other staff support them to improve practice in the classroom between the two functions. The positive correlations between the implementation score and the cost allocation to the Center Administration and Planning function suggest that centers that are able to allocate a substantial percentage of costs to Center Administration and Planning may be those centers that have more resources to devote to implementation supports in general.

3.   Implementation scores and cost per child care hour by key function

To further explore the relationship between implementation and cost, we looked at correlations between the implementation score for each key function and the cost per child care hour for each key function. The cost per child care hour by key function is a measure of how centers allocate their costs in dollar terms; it is an absolute measure of cost by function rather than a relative measure.

We found positive associations between the implementation scores and cost per child care hour by key function (shown in green font in Exhibit VI.3).

Associations between implementation scores and cost per child care hour

- Higher implementation scores in the Instructional Planning, Coordination, and Child Assessment function are positively and significantly associated with higher costs per child care hour in (1) Instructional Planning, Coordination, and Child Assessment, (2) Center Administration and Planning, and (3) Workforce Development.

- Higher implementation scores in the Center Administration and Planning function are positively and significantly associated with higher costs per child care hour in (1) Instructional Planning, Coordination, and Child Assessment and (2) Workforce Development.

- Higher implementation scores in the Workforce Development function are positively and significantly associated with higher costs per child care hour in (1) Instruction and Caregiving, (2) Instructional Planning, Coordination, and Child Assessment, (3) Center Administration and Planning, and (4) Workforce Development.

- Higher implementation scores in the Child and Family Support function are positively and significantly associated with higher costs per child care hour in Center Administration and Planning.

This set of correlations shows that centers that spend more (in absolute dollar terms) on each key function also tend to show higher levels of implementation. We see positive correlations between implementation scores and the corresponding cost per child care hour for the same function (shown in Exhibit VI.3 in the gray shaded cells). These correlations are significant for three of the five functions. This finding indicates a relationship between the resources used for a key function and the level of implementation of that key function. (Appendix C lists the items included in each measure, by key function.)

Costs for structural supports are spread across functions because the cost measures are constructed based on staff time use. Structural features are large cost drivers because they are associated with staffing patterns and compensation levels, but teacher time and compensation are spread across functions (particularly across Instruction and Caregiving and Instructional Planning, Coordination, and Child Assessment) in our cost measures. As a result, the implementation score for the Instruction and Caregiving function might not be highly associated with the cost per child care hour for the function.

Exhibit VI.3. Correlations between implementation scores and cost per child care hour, by center function

Implementation scores	Cost per child care hour by function				
	Instruction and Caregiving	Instructional Planning, Coordination, and Child Assessment	Center Administration and Planning	Workforce Development	Child and Family Support
Instruction and Caregiving	0.13	0.11	0.32	0.23	0.14
Instructional Planning, Coordination, and Child Assessment	0.08	**0.40****	**0.58*****	0.33	0.27
Center Administration and Planning	0.29	0.34	**0.60*****	**0.35***	0.19
Workforce Development	**0.42****	**0.42****	**0.46****	**0.61*****	0.19
Child and Family Support	0.34	0.26	**0.51****	0.24	0.34

Source: Data collected for the Assessing Implementation and Cost of High Quality Early Care and Education (ECE-ICHQ) Project, Phase 2.

Note: Sample size is 25 centers; cost data collection was completed with 26 centers in the total sample, one outlier is not included. Green, bold font denotes statistically significant correlations. The gray shaded cells denote the correlations between the cost per child care hour and implementation score of the same center function.

* Significant at the 0.10 level.

** Significant at the 0.05 level.

*** Significant at the 0.01 level.

B. The relationship between implementation, costs, and quality

To examine how implementation and costs vary by center quality, we created figures to show where centers fall in terms of cost per child care hour, implementation score, and quality level. Exhibit VI.4 shows this figure for the Center Administration and Planning function, plotting the implementation score and total cost per child care hour for each center. We use QRIS ratings as a proxy for quality, with blue dots representing centers with high quality ratings and red dots representing centers with low quality ratings. We categorized centers into high and low QRIS ratings based on the requirements for the different levels in each state in the multi-case study (discussed in Chapter III).[11] The dashed lines in the figure indicate the median cost per child care hour and implementation score for the function, so that each panel on the figure shows where centers fall compared to the rest of the centers in our sample. Figures for the other key functions are included in Appendix D.

We see variation on the cost per child care hour and implementation scores even among centers with similar levels of quality. For example, not all centers with a high QRIS level have high implementation scores or high costs. The lower left panel has some blue dots, representing centers that have high QRIS ratings but have implementation scores at or lower than the score of half of our sample. Similarly, a couple centers with low QRIS ratings have at median cost per child care hour or an above median implementation score. These data suggest that our cost and implementation measures are capturing information that is not fully captured by QRIS ratings, and that our measures can contribute to a more complete understanding of the relationships among cost, quality, and implementation. The ECE-ICHQ measures appear to be capturing additional, unique information that might clarify pathways to quality through the combination of implementation and cost information.

---

[11] The high QRIS category includes Arkansas centers with a rating of 3, Arizona centers with a rating of 4 or 5, Pennsylvania centers with a rating of 4, and all Head Start centers in each state. The low QRIS category includes Arkansas centers with a rating of 1, Arizona centers with a rating of 2 or 3, and Pennsylvania centers with a rating of 2.

Exhibit VI.4. Distribution of implementation scores and cost per child care hour for the Center Administration and Planning function, by quality level
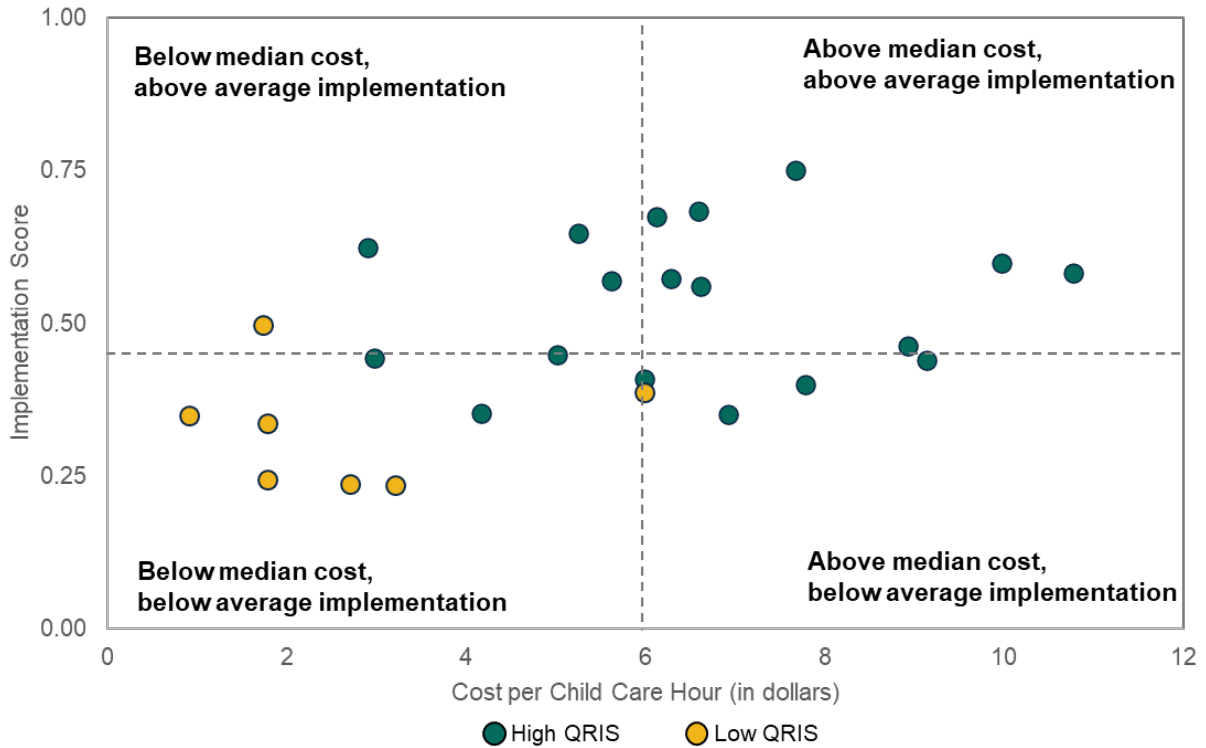
Source: Data collected for the Assessing Implementation and Cost of High Quality Early Care and Education (ECE-ICHQ) Project, Phase 2.

Note: Each dot represents one of the 25 centers in Phase 2. We categorized centers into high and low QRIS ratings based on the requirements for the different levels in each state in the multi-case study.

To better understand the variation in our measures, we looked more closely at the measures for two centers in our sample. We wanted to look at a scenario in which two centers have similar levels of quality but different costs, to see if there are differences in the way they use resources or in their implementation that enable one center to achieve the same quality as the other but at lower cost.

Exhibit VI.5 shows the costs per child care hour by function for two centers in the same state with high QRIS ratings but differences in the total cost per child care hour. The higher-cost center spends $3.70 more per child care hour than the lower-cost center. Each center receives a mix of public and private funding and serves children ages 0 to 5 as well as school age. We see that the higher-cost center spends more, in absolute dollar terms, on all the functions, compared to the lower-cost center. We might think that the package of services each center offers—or what each center does—may differ.

Exhibit VI.5. Cost per child care hour, by key function: Centers with high QRIS ratings, different total cost per child care hour

Source: Data collected for the Assessing Implementation and Cost of High Quality Early Care and Education (ECE-ICHQ) Project, Phase 2.

Note: Both centers have mixed funding and serve children ages 0 to 5 as well as school age. Total cost per child care hour is $2.89 for the lower-cost center and $6.59 for the higher-cost center.

Exhibit VI.6 shows differences in implementation scores between the two centers. The largest difference in implementation scores is in the Workforce Development function (a difference of about .14). The difference of .14 is sizable, given that the standard deviation for this function is .15. This indicates that the higher cost center is doing considerably more in terms of workforce development than the lower cost center.

Exhibit VI.6. Implementation scores: Centers with high QRIS ratings, different total cost per child care hour
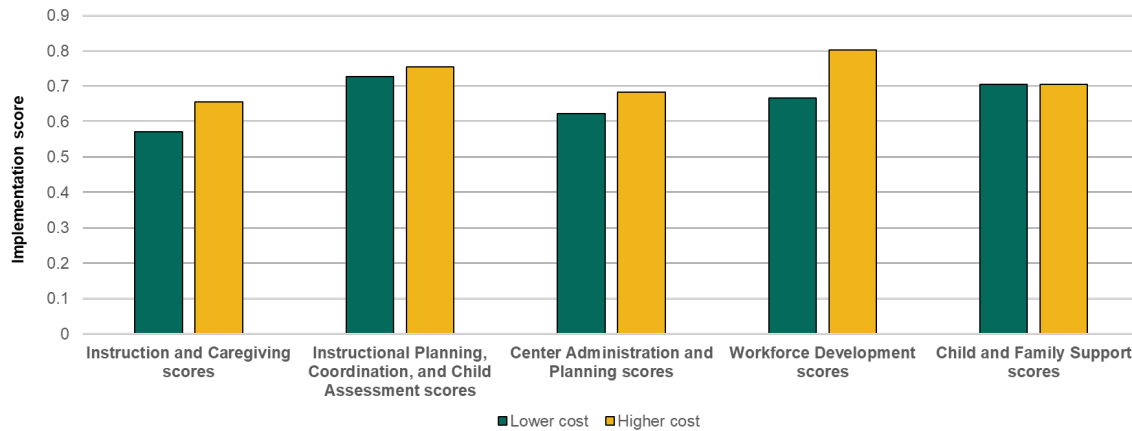
Source: Data collected for the Assessing Implementation and Cost of High Quality Early Care and Education (ECE-ICHQ) Project, Phase 2.

Note: Both centers have mixed funding and serve children ages 0 to 5 as well as school age. Total cost per child care hour is $2.89 for the lower-cost center and $6.59 for the higher-cost center.

We can also look at the way each center uses its resources across the key functions. Exhibit VI.7 shows the cost allocations of the two centers across the key functions. The higher-cost center allocates more of its resources to Instruction and Caregiving and Child and Family Support, compared to the lower-cost center, which allocates more of its resources to Instructional Planning, Coordination, and Child Assessment and Center Administration and Planning. As costs are bounded by a center's resources, it could be that the lower-cost center must devote enough resources to accomplish what is needed or required in Center Administration and Planning (for example, to track and meet requirements or standards), forcing trade-offs to adjust costs in other functions.

Exhibit VI.7. Cost allocations of two centers with high QRIS ratings, but different total cost per child care hour
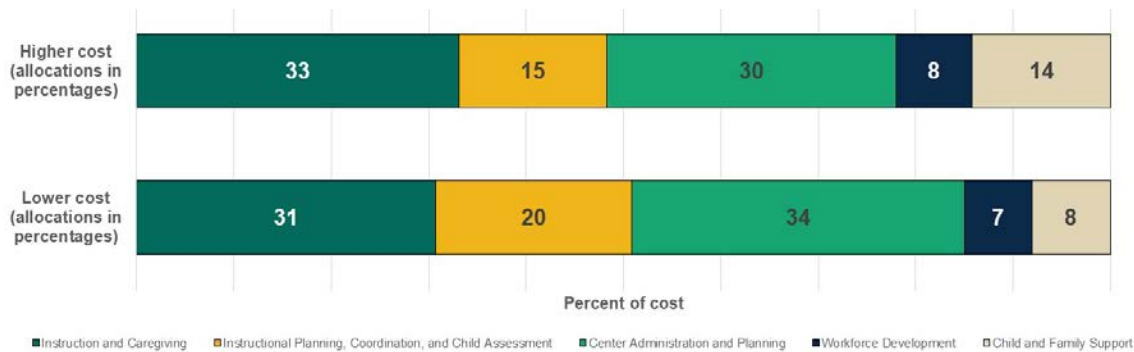
Higher cost (allocations in percentages)	33	15	30	8	14
Lower cost (allocations in percentages)	31	20	34	7	8

Percent of cost

■ Instruction and Caregiving ■ Instructional Planning, Coordination, and Child Assessment ■ Center Administration and Planning ■ Workforce Development □ Child and Family Support

Source: Data collected for the Assessing Implementation and Cost of High Quality Early Care and Education (ECE-ICHQ) Project, Phase 2.
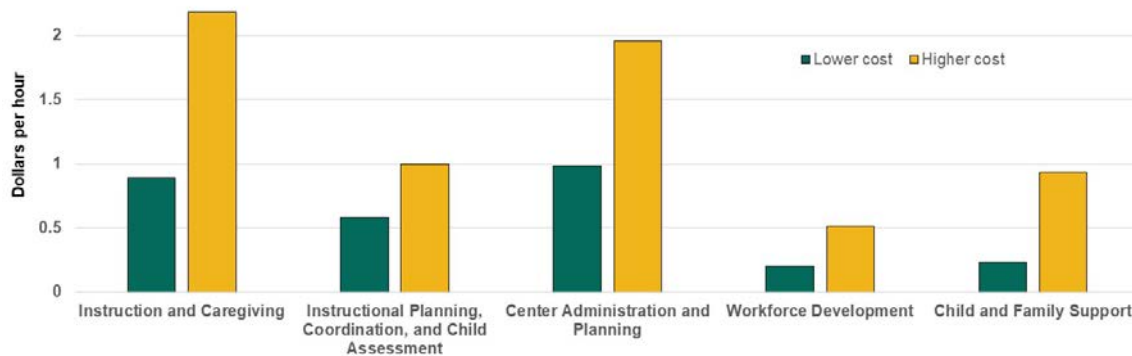
Note: Both centers have mixed funding and serve children ages 0 to 5 as well as school age. Total cost per child care hour is $2.89 for the lower-cost center and $6.59 for the higher-cost center.

The cost differences between these centers might be driven by differences in wages for staff and facility costs—the two large cost drivers. Staff and facilities costs are spread across the key functions based on the time-use allocations from the staff time-use survey. Because staff and facilities costs are not captured in just one function—as would be a function-specific cost such as paying for training—differences in these costs between centers might not be fully reflected in the cost measures by function.

C. Implications and next steps

We know from past cost work that there is not a clear linear relationship between costs and quality, or a specific cost at which we can say a center can achieve quality. We need to understand more about what a center does—the package of services it provides and how—and the decisions or trade-offs it makes in its resource use.

The correlations between the ECE-ICHQ implementation and cost measures suggest that they can help us better understand how centers support quality. Cost per child care hour by key function shows how additional dollars spent can support implementation. Cost allocation measures shed light on the resource trade-offs centers make to support implementation and quality.

Centrality of the Center Administration and Planning function. Correlations between and among the cost and implementation measures highlight the importance of the Center Administration and Planning function. Both the cost allocation measures and the cost per child care hour measures of the Center Administration and Planning function are strongly correlated with the implementation scores of the other functions. This suggests that activities related to Center Administration and Planning can support quality through strong implementation.

Refinements to measurement. The implementation score for the Instruction and Caregiving function captures decisions and activities related to the structural features of care. We would expect to see a positive association between the implementation score for this function and the total cost per child care hour for the center, but we did not find a significant association. This score might lack the specificity needed to accurately capture differences in the costs of structural features across centers. We refined the implementation instrument to collect detail on teacher–child ratios and classroom group sizes by

classroom and to improve the quality of the data on teacher qualifications. We expect this refinement to improve measurement of the structural features captured in the Instruction and Caregiving function implementation score and to result in a positive correlation with the total cost per child care hour.

We must also explore a way to reflect differences in staff and facilities costs in the measures by function, possibly by including specifics about staff compensation levels and paid versus unpaid facilities in the implementation measures.

VII. Next Steps for Testing and Use of the ECE-ICHQ Measures

The data from the multi-case study supported development of a new set of measures that might shed light on effective and efficient pathways to high quality in early care and education center-based settings. The next step in the measures development process is to test the measures in a large sample of centers. In this chapter, we discuss plans for a field test and close the report with ideas about how the measures can help inform important research, policy, and practice decisions for the field.

A. Plans for a field test of the draft measures

We will use information collected in a field test in 2021 to produce valid measures that assess how centers implement activities and use resources to support high quality early care and education. Specifically, the goals of the field test are to (1) refine the implementation measures to further improve their psychometric properties, (2) refine the cost measures to identify costs associated with each center function and across ages of children, and (3) test the associations among implementation, cost, and quality measures. Using data collected in the field test, we will first create the implementation and cost measures, assess their functionality and validity, and then examine associations between implementation, cost, and quality. Research questions for the field test are presented in Exhibit VII.1.

Exhibit VII.1. Research questions for the ECE-ICHQ field test, by goal

Goal 1: Refine the implementation measures to further test and improve their psychometric properties

- Can the implementation data be collected with strong inter-rater reliability?
- Are the implementation scores valid measures of the activities in each function? Are the items for each center function holding together to form implementation scores? (construct validity)
- Are implementation scores reliable measures? Do they consistently measure the activities in each function?
- Do the implementation scores capture multiple and distinct dimensions of implementation?
  - Are the implementation scores for each function associated with the scores for other functions in expected ways?
  - Do the implementation scores for each function measure distinct information?
- Are the scores associated with other measures of quality or program functioning in expected ways? (concurrent validity)
- Is there enough variation in implementation scores so that we can identify centers with high, middle, and low levels of implementation?
- Do the implementation scores vary in expected ways based on different center characteristics?

Goal 2: Refine the cost measures to identify costs associated with each center function and across ages of children

- Can the cost data be feasibly collected in a range of center-based settings?
- Do time-use allocations for different center functions vary in expected ways among staff in different job categories?
- Is there enough variation in cost measures so that we can identify centers with high, middle, and low levels of cost?
- Do costs vary in expected ways based on different center characteristics and for different age groups of children?

Goal 3: Test the associations between implementation, cost, and quality measures
• What is the relationship between implementation and cost measures? Does the strength of the relationship between implementation and cost measures vary by center function?
• What is the relationship between cost and quality? Does the strength of the relationship between cost measures and quality vary by center function?
• Do the implementation and cost measures explain variation in quality beyond what is explained by center characteristics?
• Do each of the measures (implementation and cost) relate to quality independent of the other?
• Does the relationship between cost and quality vary by levels of implementation?
• What are the different profiles of implementation and cost measures that exist among centers?
– Do the profiles of high quality centers look different from those of centers with low quality?
– Do there appear to be efficient paths to high quality (at lower cost) based on patterns in implementation scores?

The plans for a field test include collecting data from a purposive sample of 80 centers in up to four states. To reflect the variety of geographic and policy contexts in which early care and education centers operate, we will select centers from varied locations within each state. The sampling and recruitment strategy for the field test will also ensure that centers vary in characteristics such as whether the center serves infants and toddlers, size, profit status, funding mix, and whether they are embedded in a larger organization.

We will collect different types of data in the field test to construct the ECE-ICHQ measures and to help test and interpret them (Exhibit VII.2). We plan to administer the revised implementation interview by telephone with the center director or person most knowledgeable about the early childhood education program to collect the data to construct the implementation measures. To collect data needed to construct the cost measures, we will support completion of the Excel-based cost workbook remotely and administer web-based time-use surveys to all teaching staff and select administrators . We will collect data on center characteristics to ensure variation of centers within the field test sample and to examine whether the measures vary in expected ways based on center characteristics.

Exhibit VII.2. Data collection in the ECE-ICHQ field test to support measures development

Implementation data collection to develop implementation measures
• Implementation score for Structural Supports for Instruction and Caregiving
• Implementation score for Instructional Planning, Coordination, and Child Assessment
• Implementation score for Center Administration and Planning
• Implementation score for Workforce Development
• Implementation score for Child and Family Support

Cost data collection to develop cost measures

- **Cost per child care hour.** This measure is the total annual cost (in dollars) for a center to provide services to children ages 0 to 5 divided by the total hours of care provided to children ages 0 to 5 during the year.

- **Cost per child care hour by age.** This measure is the cost of providing one hour of child care to children ages 0 to <18 months, 18 to <36 months, and 3 to 5 years separately. It represents the cost of personnel who serve each age group plus a share of other center costs, based on the proportion of total child care hours provided to children in each age group.

- **Cost allocation by resource category.** This measure is the percentage of a center's annual costs allocated to each of eight categories of resources or uses: (1) personnel (salaries and fringe benefits), (2) staff training and education, (3) contracted services, (4) facilities (including in-kind), (5) supplies and materials, (6) equipment, (7) miscellaneous or other costs, and (8) overhead or costs of services provided by a larger organization.

- **Cost allocation by function.** This measure is the percentage of a center's annual costs allocated to each of the center functions specified in the conceptual framework.

- **Cost per child care hour by function.** This measure is the cost per child care hour allocated to each function (in dollars)—that is, the cost per child care hour multiplied by the proportion of annual costs allocated to each center function.

Center characteristics to assess variation in measures across centers

- Size (based on capacity)
- Ages of children served
- Profit status
- Embedded in a larger organization or chain
- Funding mix

Quality measures to test validity

- Center rating level from state Quality Rating and Improvement System (QRIS)
- Center accreditation status based on the standards of the National Association for the Education of Young Children (NAEYC)

Staff survey of center environment, supports, and conditions to test validity

- Center-level scores on the Supportive Environmental Quality Underlying Adult Learning (SEQUAL)

Quality measures will support identification of the sample and analyses of relationships between the ECE-ICHQ measures and quality. We will use the Quality Rating and Improvement System (QRIS) rating level as a sampling characteristic to engage centers of varying quality in the field test. We will also use the QRIS rating as a center characteristic in the first sets of analyses that examine whether the measures vary in expected ways with QRIS level. We had planned to conduct classroom observations in each center using the Early Childhood Environment Rating Scale-Revised (ECERS-R) and the Infant/Toddler Environment Rating Scale-Revised (ITERS-R) to directly assess the quality of each center during the study period to support the analyses of relationships between implementation, cost, and quality. Given the timing of the field test during the COVID-19 pandemic, we will not pursue classroom observations.

B.   Examining concurrent validity

A key objective of the field test is to assess the validity of the ECE-ICHQ measures—that they measure what they intend to measure. We will explore the concurrent validity of the ECE-ICHQ implementation measures by examining associations between implementation scores and center characteristics, QRIS ratings and accreditation status, and a validated measure of the center work environment, supports, and conditions based on staff surveys.

There are several types of concurrent validity, and we propose to measure both criterion and convergent (or construct) validity.

Examining criterion validity. We plan to use various measures of quality to examine criterion validity of the ECE-ICHQ implementation measures. For example, we can test for significant mean differences in implementation scores of centers that are or are not NAEYC-accredited or have higher or lower QRIS ratings. This approach does not necessarily indicate that the measures are capturing the same constructs, but that they are focused on the same objectives. We would expect that high implementation scores would be associated with higher quality—as measured through NAEYC accreditation or QRIS ratings.

Examining convergent validity. We plan to explore convergent validity using scales of a validated staff survey measure. Because the implementation measures are intended to fill a measurement gap, they are not well aligned with existing, validated measures as a whole. However, where similarities exist between the implementation measures and specific scales of a validated survey, we can assess whether the implementation measures are measuring what they intend, as evidenced by positive correlations with similar scales.

> **Concurrent validity** demonstrates the association (usually measured as a correlation) between a score on a given measure and performance on another measure of the same or similar construct obtained at approximately the same time.
>
> Following are three types of concurrent validity:
>
> - **Criterion validity** tests which scores on a measure are statistically related to a criterion (such as accreditation or quality rating level) or to scores on some other measure (preferably an established measure) of the same objectives or criteria.
>
> - **Convergent validity** is a type of construct validity that shows a positive relationship with measures of the same or similar constructs.
>
> - **Discriminant validity** demonstrates there is not an association between measures that capture different constructs. It can provide evidence that a measure is unique in capturing constructs not covered in other measures.

We will explore convergent validity by administering to all teaching staff a validated survey that shares similarities in constructs with the implementation measures. A measure based on staff surveys can provide a different perspective on similar topics as the implementation measures that are constructed from an interview with center directors and educational program managers. We identified the Supportive Environmental Quality Underlying Adult Learning (SEQUAL; Center for the Study of Child Care Employment) as a measure that provides an opportunity to test convergent validity of the implementation scores for all five functions. The SEQUAL consists of a self-administered survey of teachers who work with children from birth to age 5 in center-based settings to assess their perceptions of their work environment and the supports and conditions that affect their practice. We will explore convergent

validity by looking at whether implementation scores for certain functions relate closely to SEQUAL domains with similar constructs.

C. Examining relationships between implementation, cost, and quality

A primary goal of the ECE-ICHQ project is to understand how the interplay between implementation and cost explains variation in center quality. Our analysis of the Phase 2 data suggests that the ECE-ICHQ measures are capturing variation among centers that might shed light on pathways to quality through the combination of implementation and cost information. A field test sample of 80 centers will allow us to further explore, through multivariate analysis, the unique (and combined) contributions of the implementation and cost measures in explaining variations in quality. Multivariate analyses will allow us to hold some variables constant to examine the unique influence of selected variables on quality. Through multivariate analyses, we can explore the following:

- Examine whether the cost and implementation measures explain variation in quality beyond what is explained by center characteristics

- Examine the degree to which each set of measures (cost and implementation) is related to quality, independent of the other

- Test whether the relationship between cost and quality varies by levels of implementation (that is, whether implementation moderates the relationship between cost and quality)

D. Potential uses of the ECE-ICHQ measures

Because the ECE-ICHQ measures have been designed to support an integrated assessment of the implementation and costs of center-based child care, researchers, policymakers, and practitioners might use them in a variety of ways. The measures could be used independently or together to support descriptive analysis of center-based ECE settings or systems that inform program management, quality improvement, or funding. Examples of potential uses include the following:

- At the program level, the measures could complement or inform QRIS ratings. Center leaders might use enhanced information linking implementation and quality to make decisions about actions for technical assistance or program improvement.

- State leaders might use measures of the cost of high quality care to inform subsidy payments and policymaking that promotes cost-efficient paths to enhance program quality. If observations vary by geographic location or center characteristics, policymakers and administrators might target funding and quality improvement efforts appropriately.

- Researchers and program developers might use insights from the ECE-ICHQ measures to compare cost and implementation patterns among different early care and education program models.

- Federal policymakers could use data collected with the ECE-ICHQ instrument to inform cost modeling for efforts to expand access to high quality early care and education.

Ultimately, the ECE-ICHQ measures have the potential to support the efforts of researchers and others to continue to clarify the relationships between implementation, costs, and quality. These insights, in turn, can inform decisions about how best to use resources at the center level to deliver high quality early care and education more efficiently.

Our intent has always been to build measures of implementation and cost for use with observational measures of quality to examine the variation in center capacities and resources that make a difference for children. Better understanding the association between the ECE-ICHQ implementation and cost measures and quality is essential to help the field (1) dig deeper than price, market rates, and total labor costs to inform center decision making and policy and (2) contribute to equity considerations by identifying the costs of implementation that can help achieve quality by looking at the package of services a center provides and the characteristics of the children it serves.
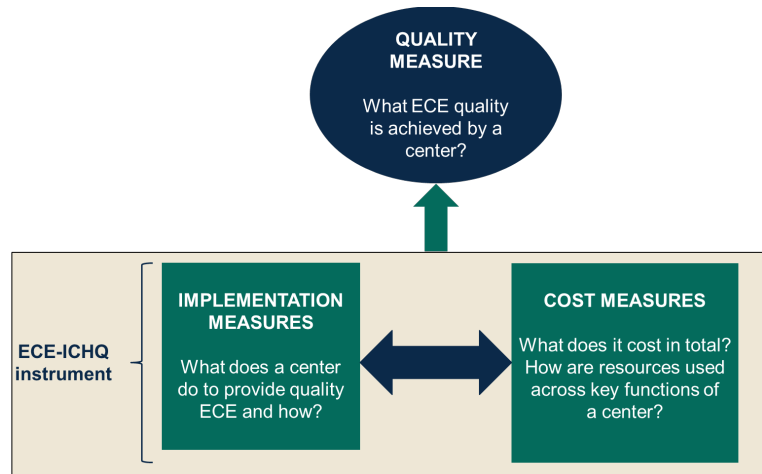
References

Asparouhov, T., and B. Muthén. "Prior-Posterior Predictive P-values." *Mplus Web Notes*, no. 22. 2017. Available at https://www.statmodel.com/download/PPPP.pdf. Accessed August 26, 2020.

Belfield, C., and H. Schwartz. "The Cost of High-Quality Pre-School Education in New Jersey." Newark, NJ: Education Law Center, 2007.

Berkel, C., A. M. Mauricio, E. Schoenfelder, and I. N. Sandler. "Putting the Pieces Together: An Integrated Model of Program Implementation." *Prevention Science,* vol. 12, no. 1, 2011, pp. 23–33.

Blau, D. M., and H. N. Mocan. "The Supply of Quality in Child Care Centers." *The Review of Economics and Statistics,* vol. 84, no. 3, 2002, pp. 483–496.

Brown, T. A. *Confirmatory Factor Analysis for Applied Research*. New York, London: The Guilford Press, 2015.

Burwick, A., H. Zaveri, L. Shang, K. Boller, D. Daro, and D. Strong. "Costs of Early Childhood Home Visiting: An Analysis of Programs Implemented in the Supporting Evidence-Based Home Visiting to Prevent Child Maltreatment Initiative." Princeton, NJ: Mathematica Policy Research, 2014.

Caronongan, P., G. Kirby, K. Boller, E. Modlin, and J. Lyskawa. "Assessing the Implementation and Cost of High Quality Early Care and Education: A Review of Literature." OPRE Report #2016-31. Washington, DC: U.S. Department of Health and Human Services, Administration for Children and Families, Office of Planning, Research and Evaluation, 2016.

Clark, L. A., and D. Watson. "Constructing Validity: Basic Issues in Objective Scale Development." *Psychological Assessment*, vol. 7, no. 3, 1995, pp. 309–319.

Damschroder, L. J., D. C. Aron, R. E. Keith, S. R. Kirsh, J. A. Alexander, and J. C. Lowery. "Fostering Implementation of Health Services Research Findings into Practice: A Consolidated Framework for Advancing Implementation Science." *Implementation Science,* vol. 4, no. 50, 2009. doi:10.1186/1748-5908-4-50

Deng, L., and W. Chan. "Testing the Difference Between Reliability Coefficients Alpha and Omega." *Educational and Psychological Measurement*, vol. 77, no. 2, 2017, pp. 185–203. Available at https://doi.org/10.1177/0013164416658325.

Durlak, J. A., and E. P. Dupre. "Implementation Matters: A Review of Research on the Influence of Implementation on Program Outcomes and the Factors Affecting the Implementation." *American Journal of Community Psychology,* vol. 41, nos. 3–4, 2008, pp. 327–350.

Elsworth, G. R., A. Beauchamp, and R. H. Osborne. "Measuring Health Literacy in Community Agencies: A Bayesian Study of the Factor Structure and Measurement Invariance of the Health Literacy Questionnaire (HLQ)." *BMC Health Services Research*, vol. 16, no. 1, 2016, pp. 1–14. Available at https://doi.org/10.1186/s12913-016-1754-2.

Fixsen, D. L., S. F. Naoom, K. A. Blase, R. M. Friedman, and F. Wallace. *Implementation Research: A Synthesis of the Literature.* Tampa, FL: University of South Florida, Louis de la Parte Florida Mental Health Institute, The National Implementation Research Network (FMHI Publication #231), 2005.

Gold, A. H., A. Malhotra, and A. H. Segars. "Knowledge Management: An Organizational Capabilities Perspective." *Journal of Management Information Systems*, vol. 18, no. 1, 2001, pp. 185–214.

Green, S. B. "Coefficient Alpha: A Reliability Coefficient for the 21st Century?" *Journal of Psychoeducational Assessment*, vol. 29, no. 4, 2011, pp. 377–392. Available at https://doi.org/10.1177/0734282911406668.

Greenhalgh, T., G. Robert, F. Macfarlane, P. Bate, and O. Kyriakidou. "Diffusion of Innovations in Service Organizations: Systematic Review and Recommendations." *Milbank Quarterly*, vol. 82, no. 4, 2004, pp. 581–629.

Harms, T., R. Clifford, and D. Cryer. *The Early Childhood Environment Rating Scale—Revised*. New York, NY: Teachers College Press, 1998.

Heerwegh, D. "Small Sample Bayesian Factor Analysis." PhUSE paper, 2014. Available at http://www.lexjansen.com/phuse/2014/sp/SP03.pdf. Accessed August 26, 2020.

Helburn, S. W. (ed.). *Cost, Quality, and Child Outcomes in Child Care Centers: Technical Report.* Denver, CO: Economics Department, University of Colorado at Denver, 1995.

Henseler, J., C. M. Ringle, and M. Sarstedt. "A New Criterion for Assessing Discriminant Validity in Variance-Based Structural Equation Modeling." *Journal of the Academy of Marketing Science*, vol. 43, no. 1, 2015, pp. 115–135. Available at https://doi.org/10.1007/s11747-014-0403-8.

Kline, R. B. *Principles and Practice of Structural Equation Modeling*. New York: Guilford Press, 2011.

Levin, H. M., and H. L. Schwartz. "What Is the Cost of a Preschool Program?" Paper presented at the Association for Education Finance Analysis Conference, Baltimore, MD, March 2007.

Marshall, N. L., C. L. Creps, N. R. Burstein, F. Glantz, W. W. Robeson, and S. Barnett. "The Cost and Quality of Full Day, Year-Round Early Care and Education in Massachusetts: Preschool Classrooms." Cambridge, MA: Wellesley Centers for Women and Abt Associates, Inc., 2001.

Marshall, N. L., C. L. Creps, N. R. Burstein, J. Roberts, J. Dennehy, W. W. Robeson, and F. Glantz. "The Cost and Quality of Full Day, Year-Round Early Care and Education in Maine: Preschool Classrooms." Cambridge, MA: Wellesley Centers for Women, Muskie Institute of the University of Southern Maine, and Abt Associates, Inc., 2004b.

Marshall, N. L., C. L. Creps, J. Roberts, F. B. Glantz, and W. W. Robeson. "The Cost and Quality of Full-Day Year-Round Early Care and Education in Massachusetts: Infant and Toddler Classrooms." Cambridge, MA: Wellesley Centers for Women and Abt Associates, Inc., 2004a.

Metz, A., T. Halle, L. Bartley, and A. Blasberg. "The Key Components of Successful Implementation." In *Applying Implementation Science in Early Childhood Programs and Systems*, edited by T. Halle, A. Metz, and I. Martinez-Beck. Baltimore, MD: Paul H. Brookes Publishing Co., 2013.

Meyers, D. C., J. A. Durlak, and A. Wandersman. "The Quality Implementation Framework: A Synthesis of Critical Steps in the Implementation Process." *American Journal of Community Psychology,* vol. 50, nos. 3–4, 2012, pp. 462–480.

Mitchell, A. W. "Modeling the Cost of Quality in Early Achievers Centers and Family Child Care." Washington Department of Early Learning, QRIS National Learning Network, 2014.

Muthén, B., and T. Asparouhov. "Bayesian SEM: A More Flexible Representation of Substantive Theory Web Tables." *Psychological Methods*, vol. 17, no. 3, 2011, pp. 313–335. https://doi.org/10.1037/a0026802

NORC. National Survey of Early Care & Education's Revised Center-Based Provider Questionnaire, November 28, 2011. Available at https://www.acf.hhs.gov/opre/resource/center-based-provider-questionnaire-revised.

Nunnally, J. C. *Psychometric Theory*. New York: McGraw-Hill, 2nd ed, 1978.

Nunnally, J. C., and I. H. Bernstein. *Psychometric Theory*. New York, NY: McGraw Hill, 3rd ed, 1994.

Pianta, Robert C., Karen M. LaParo, and Bridget K. Hamre. *Classroom Assessment Scoring System Manual, Pre-K*. Baltimore, MD: Paul H. Brookes, 2006.

Pope, G. Blog post, n.d. Available at https://blog.questionmark.com/item-analysis-analytics-part-6-determining-whether-a-question-makes-the-grade. Accessed November 5, 2018.

Raykov, T., and G. A. Marcoulides. *Introduction to Psychometric Theory*. New York, NY: Taylor & Francis, 2011.

Schwartz, H. L., and L. A. Karoly. "Cost Study of the Saint Paul Early Childhood Scholarship Program." Technical report. Santa Monica, CA: RAND Corporation, 2011.

Spillane, James P., Richard Halverson, and John B. Diamond. "Towards a Theory of Leadership Practice: A Distributed Perspective." *Journal of Curriculum Studies*, vol. 36, no. 1, 2004, pp. 3–34.

Strauss, M. E., and G. T. Smith. "Construct Validity: Advances in Theory and Methodology." *Annual Review of Clinical Psychology*, vol. 5, no. 1, 2009, pp. 1–25. Available at https://doi.org/10.1146/annurev.clinpsy.032408.153639.

Talan, T. N., and P. J. Bloom. *Program Administration Scale: Measuring Early Childhood Leadership.* New York, NY: Teachers College Press, 2004.

Talan, T. N., and P. J. Bloom. *Program Administration Scale: Measuring Early Childhood Leadership and Management.* New York, NY: Teachers College Press, 2011.

Tout, K., R. Starr, M. Soli, S. Moodie, G. Kirby, and K. Boller. "Compendium of Quality Rating Systems and Evaluations." OPRE Report. Washington, DC: Office of Planning, Research, and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services, 2010.

Trizano-Hermosilla, I., and J. M. Alvarado. "Best Alternatives to Cronbach's Alpha Reliability in Realistic Conditions: Congeneric and Asymmetrical Measurements." *Frontiers in Psychology*, vol. 7, May 2016, pp. 1–8. Available at https://doi.org/10.3389/fpsyg.2016.00769.

van Erp, S., J. Mulder, and D.L. Oberski. "Prior Sensitivity Analysis in Default Bayesian Structural Equation Modeling." *Psychological Methods*, 2017. Available at https://doi.org/10.1037/met0000162. Accessed August 26, 2020.

Wandersman, A., V. H. Chien, and J. Katz. "Toward an Evidence-Based System for Innovation Support for Implementing Innovations with Quality: Tools, Training, Technical Assistance, and Quality Assurance/Quality Improvement." *American Journal of Community Psychology,* vol. 50, nos. 3–4, 2012, pp. 445–459.

Wandersman, A., J. Duffy, P. Flaspohler, R. Noonan, K. Lubell, L. Stillman, and J. Saul. "Bridging the Gap Between Prevention Research and Practice: The Interactive Systems Framework for Dissemination and Implementation." *American Journal of Community Psychology,* vol. 41, nos. 3–4, 2008, pp. 171–181.

Yoshikawa, H., C. Weiland, J. Brooks-Gunn, M. Burchinal, L. Espinosa, W. Gormley, and J. Ludwig, et al. "Investing in Our Future: The Evidence Base on Preschool Education." New York: Foundation for Child Development and Ann Arbor, MI: Society for Research in Child Development, 2013.

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

APPENDIX A

Technical Information to Support
Implementation Measures Development

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

A.   Supporting information for the creation of summary measures

Examining pairwise correlations. For each pairwise combination of items, we produced an appropriate correlation coefficient depending on the item type. The items are largely ordinal with a few binary items, and the Mplus software is capable of producing appropriate correlations for categorical outcomes—tetrachoric (for binary data) or polychoric (for ordered-categorical data) pairwise correlations, depending on the item type. Both tetrachoric and polychoric correlations assume categorical response variables are underlain by normal response distributions. Therefore, they estimate what the correlation between variables would be if responses were made on a continuous scale. Resulting correlation coefficients are also theoretically invariant over changes in the number or "width" of rating categories, which is particularly important because the items may have different value ranges.

Once we obtained the correlation matrix for the items within a key function, we identified correlations above 0.90. There is no universal threshold for collinearity/discriminant validity, but a threshold of 0.85 (Clark and Watson 1995; Kline 2011) and 0.90 (Gold et al. 2001) are most commonly used in practice (see, for example, Henseler et al. 2015). Excluding items that lack discriminant validity is important for two reasons. First, including items that are highly redundant reduces the breadth of the scale, and the resulting high reliability would be associated with a reduction of the scale validity (Strauss and Smith 2009). Second, in the extreme cases (that is, in cases when correlations are approaching 1.0), such items could prevent model estimation because of multicollinearity.

Examining item-total correlations to refine the item set. We calculated the item-total correlation in Mplus using the procedure described by Raykov and Marcoulides (2011, pp. 108–109). To calculate this statistic, we first computed the total score (the sum of all items provisionally included in the scale) and then computed correlations between each individual item and a total score.[12] Items with higher item-total correlation statistics account for more shared variation in the scale and therefore can be thought of as the items that may primarily define or "capture" the meaning of the scale. Item-total correlation is also referred to as an index of discrimination and corresponds to the degree to which centers with a higher overall score on a key function also got high scores on an individual item. Item-total correlations above 0.3 suggest that an item is "strongly discriminating" (that is, an item is effective at discriminating centers with different levels of implementation) and should remain in a scale (Nunnally and Bernstein 1994, p. 304). We included items with item-total correlations of an absolute value above 0.2, which corresponds to moderate outcome discrimination and could potentially be considered for including in the scale (Pope n/d).

Confirmatory factor analysis (CFA). A CFA tests a particular model rather than all possible models and so can be estimated with fewer cases. Researchers can refine the model based on results and then test again using an iterative approach.

We tested whether grouping the items together in a single scale is consistent with the available data using a Bayesian CFA model. With this small sample, we need to limit the number of parameters to be estimated, so we conducted factor analysis at the key function level.[13] The Bayesian estimation approach is generally recommended when the sample size is small (Heerweg 2014; ven Erp et al. 2017). Because CFA is a theory-driven procedure, we expect that the key function constructs we initially develop may

---

[12] Biserial and polyserial correlations were computed between binary and continuous and ordinal and continuous variables, respectively.

[13] Our ability to conduct the analysis across multiple functions was limited. Because the sample size is small, the statistical analyses would be underpowered, meaning that confidence intervals around point estimates are large.

need subsequent iterative refinements that are informed by the results. For example, some items may end up having weak factor loadings and need to be removed from the model.

Model fit. We relied on the global fit statistics, such as predictive posterior p-value (PPP)[14], which measures whether correlations reproduced by the model are significantly different from the correlations observed in the data; and local fit statistics, such as loadings in factor analysis, which show the strength of the relationship of each item and an underlying factor. A PPP statistic is obtained via a fit statistic f based on the usual chi-square test of the null hypothesis against an alternative hypothesis.

> Predictive posterior p-value (PPP) is a Bayesian measure of statistical fit of the model to the data; values ≥ 0.05 typically indicate good fit of the model to the data, and values ≥ 0.50 indicate excellent fit to the data (Muthén and Asparouhov 2011).
>
> Omega statistics for internal consistency reliability; looking for close to or above 0.70 as sufficient for the early stages of scale development (Nunnally and Bernstein 1994).
>
> Factor loading is the measure of correlation between the item and the factor that it measures. Loadings below 0.30 are typically considered non-essential (Brown 2015). We included items with factor loadings ≥ .32, corresponding to at least 10 percent ($0.32^2=0.10$) shared variance between the item and a factor.

An excellent-fitting model is expected to have a PPP value around 0.5 and an f-statistic difference of zero, falling close to the middle of the confidence interval. A positive lower limit for the 95 percent confidence interval of the f-statistic is in line with a low PPP and indicates poor fit (Muthén and Asparouhov 2011). There is no theory for how low PPP can be before the model is significantly ill-fitting at a certain level. A PPP value of > 0.05 is generally used as a cutoff, indicating that the discrepancy between the suggested factor model and the data is not statistically significant and thus the model is an acceptable fit to the data (Muthén and Asparouhov 2011; Elsworth et al. 2016). We should stress, however, that in our case, low PPP values together with small sample size may show limited power of the data to reject the incorrect model.[15] We also inspected Bayesian posterior distribution graphs to ensure convergence of model parameters.[16]

Model reliability. Based on the analyses of the factor structure of the scales within each key function, we computed McDonald's Omega (ω) reliability of the functions (Green 2011). McDonald's Omega relies on the estimates from a factor analytic model—in particular, the commonalities and either the variance-covariance matrix or the correlation matrix. The observed score variance includes both the variance of the expected (true) scores and the variance of error scores.

---

[14] It tests the fit to the data of the model with the given priors and is based on comparing the model with the unrestricted mean and variance-covariance model (Asparouhov and Muthén 2017)

[15] The power of PPP to reject a misspecified model increases with the increasing sample size (Hoofs et al. 2018).

[16] Root Mean Squared Error of Approximation (RMSEA), Tucker-Lewis Index (TLI), Comparative Fit Index (CFI), Standardized Root Mean Residual (SRMR), and other statistics commonly used for model evaluation estimated via the most frequently used estimation methods (e.g., Maximum Likelihood) do not exist for the Bayesian estimation method.

Reliability is the ratio of true variance of the scores (signal) and the total variance of the scores (signal + noise):

$$\omega = \frac{\sigma^2_{true\ variance}}{\sigma^2_{true\ variance} + \sigma^2_{error\ variance}} \; ;$$

Calculations of the Omega coefficient account for the differential strength of the relationship between each item and a factor, and therefore provide a more accurate estimate of reliability than coefficient alpha (Trizano-Hermosilla and Alvarado 2016; Deng and Chan 2017). To evaluate the reliabilities of the scales, we follow the guidance of Nunnally (1978), who considered reliabilities of 0.70 as sufficient for the early stages of research.

B.   Supporting information for the creation of implementation scores

We added all of the items together to create the key function score.[17]

The items within the instrument were primarily ordinal with some binary items. Before computing centers' scores on a key function, we normalized each item to be on a scale from 0 to 1 using a theoretical minimum and maximum possible score for each item. We used the theoretical (rather than observed) minimum and maximum score for each item to ensure that in the process of normalization, the centers with the lowest and highest observed scores would not get the key function scores of 0.00 and 1.00 unless they had the lowest theoretically possible combination of scores on each individual item. Otherwise, for example, a center with the lowest combination of scores in our sample (not necessarily the lowest theoretically possible combination of scores) would receive a score of 0.00 due to rescaling.[18] This approach has two important implications. First, it allows us to compute a score on the key function even for just one center, as normalization is based on the theoretical distribution of the item scores rather than observed distribution, for which we would need more than one center. Additionally, the scores computed via this method are not referenced to the observed distribution of the scores in a sample, so computed scores would be comparable across multiple samples.

After normalizing the items, we computed the key function score as the average of all individual item scores within a key function.

---

[17] An alternative approach would be to additionally weight each observed item score by a factor loading of that item, thus giving more weight to the items that measure a key function more reliably. However, because of a small sample size, our estimates for factor loadings were less precise. In our simulations, centers' scores that were obtained based on the weighted and unweighted items were very similar and almost perfectly correlated.

[18] Normalization is based on the following formula: normal_x1 = (x1 – x1(min)) / (x1(max) – x1(min)), where x1 is the variable to be normalized, and x1(min) and x1(max) are the lowest and highest *observed values* in the sample for the variable x1, respectively. In this formula, we substituted x1(min) and x1(max) with the lowest and highest *theoretically possible* values.

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

APPENDIX B

Technical Information to Support
Cost Measures Development

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

Exhibit B.1. Time use survey activities included in key functions

Key function	Activity in time-use survey	Description
Instruction and Caregiving	Providing instruction or care	This includes one-on-one or group interaction with children during play or planned activities; leading a lesson; helping children during mealtimes; naptimes; transition times; and field trips.
Instructional Planning, Coordination, and Child Assessment	Planning curriculum, activities, and lessons	This includes developing or reviewing lesson plans and other forms of instructional planning, and preparing and setting up classroom materials, either individually or with other staff. This may include time during the school day when children are napping or otherwise engaged.
	Conducting child assessments during or outside of classroom time	This includes assessments to determine whether a child needs further developmental assessment and/or services, to assess growth and development, or to determine children's needs and plan instruction. Include time spent coordinating assessments and administering, scoring, and reviewing results. Also includes time spent documenting children's progress such as recording anecdotes and observations in a log.
	Parent communication	This includes oral (face-to-face or by phone) or written communication (such as preparing a daily report or preparing a newsletter) with parents about the care, activities, or progress of their child(ren). This may include preparing daily reports for parents when children are napping or otherwise engaged.
	Periodic child assessments and parent conferences	This includes time spent preparing for, conducting, and reporting periodic assessments of children's progress, growth, or development. Also includes time spent preparing for and conducting periodic parent conferences to discuss the care, activities, and progress of children.
	Curriculum planning and development	This includes periodic activities related to curriculum review, selection, and development, including time spent ensuring alignment with standards or program requirements.
Center Administration	Ongoing regulatory compliance and reporting	This includes daily activities to ensure and maintain compliance with federal, state, and city agencies or programs (for example, monitoring child-staff ratios, or completing injury report forms).
	Managing center finances	This includes collecting payments, budgeting, payroll, accounting, and tracking expenditures.
	Managing and maintaining center operations	This includes overseeing staff schedules, the work of facilities and maintenance staff, working with vendors, and organizing space.
	Periodic regulatory compliance and reporting	This includes periodic activities to prepare for and complete annual monitoring requirements of federal, state, and city agencies or programs.
Leadership, Planning, and Evaluation	Planning and goal-setting	This includes assessing center performance, developing written plans, conducting center self-assessments of quality, pursuing quality improvement grants, participating in planning or board meetings, and marketing and fundraising activities.

Key function	Activity in time-use survey	Description
Workforce Development	Staff supervision, coaching, or mentoring	This includes scheduled or unscheduled consultation with a supervisor or supervisees and providing or receiving coaching or mentoring either one-on-one or in groups. Also include time spent conducting or discussing classroom observations. Also includes time spent providing coaching, mentoring, or technical assistance (TA) to other center staff.
	Staff meetings	This includes time for general staff meetings on topics not already covered in other rows.
	Recruiting and hiring teaching staff	This includes preparing job listings, reviewing applications, screening, and interviewing.
	Evaluating staff performance	This includes time spent preparing for and participating in periodic reviews to evaluate performance or set goals for the coming year, including developing or reviewing training or professional development plans.
	Training and professional development activities	This includes time spent preparing, conducting, or participating (in-person or online) in training or professional development activities that occur during scheduled work hours, including planned professional development days, if applicable. Does not include time spent pursuing a degree or credential outside of work hours.
	Periodic coaching, mentoring, or technical assistance	This includes time spent receiving targeted or periodic coaching, mentoring, or technical assistance (TA) that occurs irregularly and is not part of ongoing supervision or program support activities within the center. Also includes time spent providing coaching, mentoring, or TA to other center staff.
Child and Family Support	Planning and conducting family engagement or family support activities	This includes events for currently enrolled families (such as back-to-school nights or class performances), planning or providing parent education or other family support services, and conducting home visits.

Note: In the revised conceptual framework, the Center Administration function and the Leadership, Planning, and Evaluation function are combined into a single function, Center Administration and Planning.

Exhibit B.2. Time use allocations, by key function and staff role (proportions of total working hours)

Key functions	Center directors			Teacher-director			Educational director or coordinator			Head or lead teachers			Teachers			Assistant teachers		
	Mean	Median	N	Mean	Median	N	Mean	Median	N	Mean	Median	N	Mean	Median	N	Mean	Median	N
Instruction and Caregiving	0.09 (0.10)	0.08	28	0.23 (0.10)	0.21	5	0.04 (0.05)	0.02	10	0.43 (0.14)	0.43	30	0.41 (0.15)	0.41	23	0.51 (0.19)	0.52	23
Instructional Planning, Coordination, and Child Assessment	0.20 (0.13)	0.21	28	0.31 (0.05)	0.31	5	0.34 (0.23)	0.26	10	0.35 (0.09)	0.37	30	0.37 (0.13)	0.34	23	0.29 (0.13)	0.30	23
Child and Family Support	0.01 (0.01)	0.01	28	0.01 (0.01)	0.00	5	0.01 (0.01)	0.01	10	0.01 (0.01)	0.00	30	0.01 (0.01)	0.00	23	0.00 (0.00)	0.00	23
Workforce Development	0.23 (0.15)	0.19	28	0.11 (0.08)	0.15	5	0.33 (0.17)	0.31	10	0.10 (0.05)	0.09	30	0.08 (0.05)	0.07	23	0.08 (0.04)	0.07	23
Center Administration	0.44 (0.19)	0.46	28	0.33 (0.11)	0.37	5	0.24 (0.21)	0.16	10	0.11 (0.07)	0.11	30	0.13 (0.09)	0.13	23	0.11 (0.10)	0.10	23
Leadership, Planning, and Evaluation	0.02 (0.03)	0.01	28	0.01 (0.01)	0.00	5	0.03 (0.02)	0.02	10	0.01 (0.01)	0.00	30	0.00 (0.01)	0.00	23	0.00 (0.00)	0.00	23
Total N			28			5			10			30			23			23

Source: Data collected for the Assessing Implementation and Cost of High-Quality Early Care and Education (ECE-ICHQ) Project, Phase 2.

Exhibit B.3. Time use allocations for teaching staff, by function (proportions of total working hours)

	Number of centers	Instruction and Caregiving	Instructional Planning, Coordination, and Child Assessment	Workforce Development	Child and Family Support	Center Administration and Planning	
						Center Administration	Leadership, Planning, and Evaluation
All centers							
Mean	30	0.45	0.34	0.09	0.00	0.12	0.00
Median	30	0.47	0.32	0.08	0.00	0.12	0.00
Maximum	30	0.68	0.54	0.22	0.02	0.25	0.03
Minimum	30	0.15	0.21	0.02	0.00	0.01	0.00
QRIS rating[a]							
High	20	0.47	0.34	0.08	0.01	0.11	0.00
		(0.11)	(0.09)	(0.03)	(0.00)	(0.05)	(0.01)
Low	10	0.42	0.33	0.10	0.00	0.14	0.00
		(0.13)	(0.08)	(0.05)	(0.00)	(0.05)	(0.01)
Age groups served							
Ages 0-5 and school-age	11	0.42	0.33	0.10	0.00	0.15	0.01
		(0.11)	(0.08)	(0.05)	(0.00)	(0.04)	(0.01)
Ages 0-5 years only	12	0.49	0.34	0.08	0.00	0.08	0.00
		(0.11)	(0.07)	(0.03)	(0.00)	(0.04)	(0.00)
Age 0-3 or ages 3-5 only	7	0.44	0.35	0.07	0.01	0.13	0.01
		(0.13)	(0.11)	(0.03)	(0.00)	(0.05)	(0.01)
Program size							
Large center (serving 75 children or more)	18	0.44	0.35	0.09	0.00	0.12	0.00
		(0.11)	(0.08)	(0.03)	(0.00)	(0.05)	(0.00)
Small center	12	0.47	0.32	0.09	0.00	0.11	0.01
		(0.13)	(0.09)	(0.05)	(0.00)	(0.06)	(0.01)
Profit status							
For-profit	11	0.42	0.35	0.09	0.00	0.13	0.01
		(0.12)	(0.08)	(0.06)	(0.00)	(0.04)	(0.01)
Not-for-profit	19	0.47	0.33	0.08	0.01	0.11	0.00
		(0.11)	(0.08)	(0.03)	(0.00)	(0.06)	(0.01)

	Number of centers	Instruction and Caregiving	Instructional Planning, Coordination, and Child Assessment	Workforce Development	Child and Family Support	Center Administration and Planning	
						Center Administration	Leadership, Planning, and Evaluation
Embedded in a larger organization or part of a chain[b]							
Yes	14	0.46	0.35	0.08	0.01	0.10	0.00
		(0.12)	(0.09)	(0.02)	(0.00)	(0.05)	(0.00)
No	16	0.45	0.32	0.09	0.00	0.14	0.01
		(0.12)	(0.07)	(0.05)	(0.00)	(0.05)	(0.01)
Funding							
High subsidy[c]	8	0.50	0.29	0.08	0.00	0.13	0.00
		(0.07)	(0.06)	(0.03)	(0.00)	(0.07)	(0.00)
Majority Head Start or state pre-K[d]	10	0.40	0.38	0.10	0.01	0.12	0.00
		(0.10)	(0.08)	(0.03)	(0.00)	(0.06)	(0.01)
Majority private tuition[e]	8	0.46	0.35	0.09	0.00	0.09	0.00
		(0.18)	(0.10)	(0.06)	(0.00)	(0.04)	(0.01)
Mixed public or mixed public/private	4	0.48	0.30	0.07	0.01	0.14	0.01
		(0.04)	(0.02)	(0.01)	(0.01)	(0.02)	(0.01)

Source: Data collected for the Assessing Implementation and Cost of High-Quality Early Care and Education (ECE-ICHQ) Project, Phase 2.

Note: Means are calculated at the center level among centers in each group.

[a] The high QRIS category includes: Arkansas centers with a rating of 3, Arizona centers with a rating of 4 or 5, Pennsylvania centers with a rating of 4, and all Head Start centers. The low QRIS category includes: Arkansas centers with a rating of 1, Arizona centers with a rating of 2 or 3, and Pennsylvania centers with a rating of 2.

[b] Part of a larger organization means the center is part of and shares resources/staff with a larger organization, such as a YMCA, university, or other social services agency.

[c] Includes centers in which 50 percent or more of the children are supported by funding from CCDF and less than 30 percent of children are supported by funding from other federal, state, or local government sources.

[d] Includes centers in which less than 30 percent of the children are supported with funding from CCDF.

[e] Includes centers in which 90 percent or more of the children are supported through private tuition paid by their parents or guardians without any public funding.

Exhibit B.4. Time use allocations for directors, by function (proportions of total working hours)

	Number of centers	Instruction and Caregiving	Instructional Planning, Coordination, and Child Assessment	Workforce Development	Child and Family Support	Center Administration and Planning	
						Center Administration	Leadership, Planning, and Evaluation
All centers							
Mean	28	0.09	0.20	0.23	0.01	0.44	0.02
Median	28	0.08	0.21	0.19	0.01	0.46	0.01
Maximum	28	0.39	0.50	0.65	0.06	0.77	0.14
Minimum	28	0.00	0.01	0.05	0.00	0.05	0.00
QRIS rating[a]							
High	19	0.06	0.18	0.25	0.02	0.47	0.03
		(0.05)	(0.13)	(0.15)	(0.02)	(0.20)	(0.03)
Low	9	0.17	0.24	0.17	0.01	0.39	0.01
		(0.12)	(0.12)	(0.14)	(0.01)	(0.14)	(0.01)
Age groups served							
Ages 0-5 and school-age	10	0.16	0.24	0.17	0.01	0.41	0.02
		(0.13)	(0.11)	(0.09)	(0.01)	(0.15)	(0.01)
Ages 0-5 years only	11	0.05	0.18	0.28	0.01	0.46	0.02
		(0.05)	(0.15)	(0.18)	(0.01)	(0.19)	(0.03)
Age 0-3 or ages 3-5 only	7	0.07	0.18	0.22	0.02	0.47	0.03
		(0.05)	(0.12)	(0.14)	(0.02)	(0.25)	(0.05)
Program size							
Large center (serving 75 children or more)	18	0.09	0.20	0.22	0.01	0.46	0.02
		(0.10)	(0.12)	(0.15)	(0.01)	(0.19)	(0.02)
Small center	10	0.10	0.20	0.24	0.02	0.42	0.03
		(0.10)	(0.15)	(0.15)	(0.02)	(0.19)	(0.04)
Profit status							
For-profit	10	0.15	0.26	0.19	0.01	0.38	0.01
		(0.13)	(0.10)	(0.11)	(0.01)	(0.18)	(0.01)
Not-for-profit	18	0.06	0.17	0.25	0.02	0.48	0.03
		(0.05)	(0.14)	(0.16)	(0.02)	(0.18)	(0.04)

	Number of centers	Instruction and Caregiving	Instructional Planning, Coordination, and Child Assessment	Workforce Development	Child and Family Support	Center Administration and Planning	
						Center Administration	Leadership, Planning, and Evaluation
Embedded in a larger organization or part of a chain[b]							
Yes	13	0.07	0.18	0.28	0.02	0.44	0.01
		(0.06)	(0.14)	(0.17)	(0.02)	(0.18)	(0.02)
No	15	0.12	0.22	0.18	0.01	0.45	0.03
		(0.12)	(0.12)	(0.11)	(0.01)	(0.19)	(0.04)
Funding							
High subsidy[c]	7	0.10	0.18	0.20	0.02	0.47	0.03
		(0.09)	(0.13)	(0.16)	(0.02)	(0.11)	(0.05)
Majority Head Start or state pre-K[d]	10	0.06	0.18	0.26	0.02	0.46	0.01
		(0.05)	(0.15)	(0.17)	(0.01)	(0.25)	(0.02)
Majority private tuition[e]	7	0.12	0.28	0.20	0.01	0.38	0.02
		(0.09)	(0.10)	(0.13)	(0.01)	(0.16)	(0.03)
Mixed public or mixed public/private	7	0.10	0.18	0.20	0.02	0.47	0.03
		(0.09)	(0.13)	(0.16)	(0.02)	(0.11)	(0.05)

Source: Data collected for the Assessing Implementation and Cost of High-Quality Early Care and Education (ECE-ICHQ) Project, Phase 2.

Note: Figures for subgroups are means. Means are calculated at the center level among centers in each group. Percentages may not sum to 100 due to rounding.

[a] The high QRIS category includes: Arkansas centers with a rating of 3, Arizona centers with a rating of 4 or 5, Pennsylvania centers with a rating of 4, and all Head Start centers. The low QRIS category includes: Arkansas centers with a rating of 1, Arizona centers with a rating of 2 or 3, and Pennsylvania centers with a rating of 2.

[b] Part of a larger organization means the center is part of and shares resources/staff with a larger organization, such as a YMCA, university, or other social services agency.

[c] Includes centers in which 50 percent or more of the children are supported by funding from CCDF and less than 30 percent of children are supported by funding from other federal, state, or local government sources.

[d] Includes centers in which less than 30 percent of the children are supported with funding from CCDF.

[e] Includes centers in which 90 percent or more of the children are supported through private tuition paid by their parents or guardians without any public funding.

.

Exhibit B.5. Percentage of costs by resource category, all centers

	Number of centers	Salaries and fringe	Staff training and education	Contracted services	Facilities	Supplies and materials	Equipment	Miscellaneous	Overhead or cost of services from a larger organization
All centers									
Mean	26	62	1	4	16	9	1	3	4
Median	26	63	0	3	15	8	0	2	0
Maximum	26	81	3	18	32	27	5	19	21
Minimum	26	37	0	0	0	2	0	0	0
QRIS rating[a]									
High	19	65	1	4	15	7	1	2	6
		(7.7)	(0.7)	(3.4)	(7.2)	(4.3)	(0.9)	(2.2)	(7.0)
Low	7	53	0	5	18	14	2	6	0
		(12.1)	(0.7)	(7.3)	(7.5)	(8.4)	(1.9)	(6.6)	(0.0)
Age groups served									
Ages 0-5 and school-age	9	56	1	5	16	13	1	5	2
		(12.6)	(0.6)	(6.5)	(7.4)	(7.7)	(1.6)	(6.3)	(4.5)
Ages 0-5 years only	10	65	1	3	15	9	1	2	5
		(8.3)	(0.9)	(3.5)	(6.6)	(4.5)	(1.7)	(1.9)	(6.7)
Age 0-3 or ages 3-5 only	7	66	0	5	15	5	0	2	6
		(7.6)	(0.4)	(3.0)	(9.2)	(3.7)	(0.3)	(1.6)	(7.9)
Program size									
Large center (serving 75 children or more)	15	63	1	2	15	10	1	3	4
		(10.0)	(0.7)	(2.5)	(5.5)	(5.8)	(1.5)	(4.9)	(6.3)
Small center	11	61	1	7	17	8	1	3	4
		(11.2)	(0.6)	(5.9)	(9.5)	(7.0)	(1.4)	(2.9)	(7.0)
Profit status									
For-profit	10	61	0	3	12	13	1	5	3
		(12.1)	(0.3)	(4.1)	(6.4)	(7.5)	(1.9)	(5.9)	(7.4)
Not-for-profit	16	63	1	5	18	7	1	2	4
		(9.5)	(0.7)	(5.0)	(7.4)	(4.5)	(1.0)	(1.7)	(6.1)

	Number of centers	Salaries and fringe	Staff training and education	Contracted services	Facilities	Supplies and materials	Equipment	Miscellaneous	Overhead or cost of services from a larger organization
Embedded in a larger organization or part of a chain[b]									
Yes	13	61 (9.4)	1 (0.8)	5 (5.0)	19 (7.5)	8 (4.7)	1 (1.0)	2 (1.8)	4 (6.6)
No	13	63 (11.5)	1 (0.5)	4 (4.5)	12 (5.6)	11 (7.5)	1 (1.7)	4 (5.4)	4 (6.6)
Funding									
High subsidy[c]	6	57 (14.3)	1 (0.7)	6 (7.1)	21 (7.0)	10 (9.4)	1 (1.8)	4 (3.6)	0 (0.8)
Majority Head Start or state pre-K[d]	9	64 (5.4)	1 (0.5)	4 (3.0)	14 (5.5)	8 (4.5)	1 (0.6)	2 (1.5)	8 (6.5)
Majority private tuition[e]	7	66 (7.8)	1 (1.0)	4 (4.9)	14 (9.4)	10 (7.7)	1 (1.6)	2 (2.2)	3 (7.7)
Mixed public or mixed public/private	4	60 (16.2)	1 (0.3)	3 (3.6)	14 (5.5)	10 (3.4)	1 (1.9)	7 (9.0)	3 (6.8)

Source: Data collected for the Assessing Implementation and Cost of High-Quality Early Care and Education (ECE-ICHQ) Project, Phase 2.

Note: Costs adjusted to account for state- and metro-level differences in average wages. Figures for subgroups are means. Means are calculated at the center level among centers in each group.

[a] The high QRIS category includes: Arkansas centers with a rating of 3, Arizona centers with a rating of 4 or 5, Pennsylvania centers with a rating of 4, and all Head Start centers. The low QRIS category includes: Arkansas centers with a rating of 1, Arizona centers with a rating of 2 or 3, and Pennsylvania centers with a rating of 2.

[b] Part of a larger organization means the center is part of and shares resources/staff with a larger organization, such as a YMCA, university, or other social services agency.

[c] Includes centers in which 50 percent or more of the children are supported by funding from CCDF and less than 30 percent of children are supported by funding from other federal, state, or local government sources.

[d] Includes centers in which less than 30 percent of the children are supported with funding from CCDF.

[e] Includes centers in which 90 percent or more of the children are supported through private tuition paid by their parents or guardians without any public funding.

.

Exhibit B.6. Percentage of costs by resource category, excluding outlier

	Number of centers	Salaries and fringe	Staff training and education	Contracted services	Facilities	Supplies and materials	Equipment	Miscellaneous	Overhead payments
Mean	25	62	1	4	16	9	1	3	3
Median	25	64	0	3	15	8	0	2	0
Maximum	25	81	3	18	32	27	5	19	21
Minimum	25	37	0	0	0	2	0	0	0
QRIS rating[a]									
High	18	66	1	4	15	7	1	2	5
		(7.8)	(0.7)	(3.4)	(7.4)	(4.3)	(0.9)	(2.2)	(6.1)
Low	7	53	0	5	18	14	2	6	0
		(12.1)	(0.7)	(7.3)	(7.5)	(8.4)	(1.9)	(6.6)	(0.0)
Age groups served									
Ages 0-5 and school-age	9	56	1	5	16	13	1	5	2
		(12.6)	(0.6)	(6.5)	(7.4)	(7.7)	(1.6)	(6.3)	(4.5)
Ages 0-5 years only	9	65	1	3	15	9	1	2	3
		(8.5)	(0.9)	(3.6)	(6.9)	(4.5)	(1.7)	(1.9)	(3.4)
Age 0-3 or ages 3-5 only	7	66	0	5	15	5	0	2	6
		(7.6)	(0.4)	(3.0)	(9.2)	(3.7)	(0.3)	(1.6)	(7.9)
Program size									
Large center (serving 75 children or more)	14	63	1	3	15	11	1	4	3
		(10.3)	(0.7)	(2.5)	(5.6)	(5.7)	(1.5)	(5.0)	(4.2)
Small center	11	61	1	7	17	8	1	3	4
		(11.2)	(0.6)	(5.9)	(9.5)	(7.0)	(1.4)	(2.9)	(7.0)
Profit status									
For-profit	10	61	0	3	12	13	1	5	3
		(12.1)	(0.3)	(4.1)	(6.4)	(7.5)	(1.9)	(5.9)	(7.4)
Not-for-profit	15	63	1	5	18	7	1	2	3
		(9.8)	(0.8)	(5.1)	(7.5)	(4.6)	(1.0)	(1.8)	(4.2)
Embedded in a larger organization or part of a chain[b]									
Yes	12	61	1	5	20	8	1	2	3
		(9.8)	(0.8)	(5.1)	(7.6)	(4.7)	(1.0)	(1.9)	(4.4)
No	13	63	1	4	12	11	1	4	4
		(11.5)	(0.5)	(4.5)	(5.6)	(7.5)	(1.7)	(5.4)	(6.6)

	Number of centers	Salaries and fringe	Staff training and education	Contracted services	Facilities	Supplies and materials	Equipment	Miscellaneous	Overhead payments
Funding									
High subsidy[c]	6	57 (14.3)	1 (0.7)	6 (7.1)	21 (7.0)	10 (9.4)	1 (1.8)	4 (3.6)	0 (0.8)
Majority Head Start or state pre-K[d]	8	64 (5.4)	1 (0.5)	4 (3.0)	14 (5.9)	9 (4.5)	1 (0.7)	2 (1.5)	6 (4.2)
Majority private tuition[e]	7	66 (7.8)	1 (1.0)	4 (4.9)	14 (9.4)	10 (7.7)	1 (1.6)	2 (2.2)	3 (7.7)
Mixed public or mixed public/private	4	60 (16.2)	1 (0.3)	3 (3.6)	14 (5.5)	10 (3.4)	1 (1.9)	7 (9.0)	3 (6.8)

Source: Data collected for the Assessing Implementation and Cost of High-Quality Early Care and Education (ECE-ICHQ) Project, Phase 2.

Note: Costs adjusted to account for state- and metro-level differences in average wages. Percentages may not sum to 100 due to rounding. Figures for subgroups are means. Means are calculated at the center level among centers in each group.

[a] The high QRIS category includes: Arkansas centers with a rating of 3, Arizona centers with a rating of 4 or 5, Pennsylvania centers with a rating of 4, and all Head Start centers. The low QRIS category includes: Arkansas centers with a rating of 1, Arizona centers with a rating of 2 or 3, and Pennsylvania centers with a rating of 2.

[b] Part of a larger organization means the center is part of and shares resources/staff with a larger organization, such as a YMCA, university, or other social services agency.

[c] Includes centers in which 50 percent or more of the children are supported by funding from CCDF and less than 30 percent of children are supported by funding from other federal, state, or local government sources.

[d] Includes centers in which less than 30 percent of the children are supported with funding from CCDF.

[e] Includes centers in which 90 percent or more of the children are supported through private tuition paid by their parents or guardians without any public funding.

.

Exhibit B.7. Cost per childcare hour, all centers (dollars)

	Number of centers	Annual child care hours provided	Cost per childcare hour	Range of cost per childcare hour	Standard deviation of cost per childcare hour
All centers					
Mean	26	177,890	6.36	n.a.	n.a.
Median	26	125,232	5.99	n.a.	n.a.
Maximum	26	665,397	28.57	n.a.	n.a.
Minimum	26	12,771	0.92	n.a.	n.a.
QRIS rating[a]					
High	19	186,583	7.75	2.89 - 28.57	5.48
Low	7	154,296	2.59	0.92 - 5.99	1.67
Age groups served					
Ages 0-5 and school-age	9	253,558	3.17	0.92 - 6.59	2.00
Ages 0-5 years only	10	162,334	8.53	2.98 - 28.57	7.40
Age 0-3 or ages 3-5 only	7	102,825	7.36	5.24 - 10.77	1.93
Program size					
Large center (serving 75 children or more)	15	245,730	7.16	1.78 - 28.57	6.54
Small center	11	85,380	5.27	0.92 - 8.92	2.67
Profit status					
For-profit	10	177,867	3.32	0.92 - 7.67	2.16
Not-for-profit	16	177,905	8.26	2.89 - 28.57	5.78
Embedded in a larger organization or part of a chain[b]					
Yes	13	147,297	7.97	2.89 - 28.57	6.56
No	13	208,484	4.75	0.92 - 9.97	3.01
Funding					
High subsidy[c]	6	93,281	4.56	1.79 - 5.99	1.83
Majority Head Start or state pre-K[d]	9	181,500	10.37	6.12 - 28.57	7.04
Majority private tuition[e]	7	105,465	4.19	0.92 - 7.78	2.73
Mixed public or mixed public/private	4	423,424	3.86	1.78 - 6.59	2.06

Source: Data collected for the Assessing Implementation and Cost of High-Quality Early Care and Education (ECE-ICHQ) Project, Phase 2.

Note: Costs were adjusted to account for state- and metro-level differences in average wages. Figures for subgroups are means. Means are calculated at the center level among centers in each group.

[a] The high QRIS category includes: Arkansas centers with a rating of 3, Arizona centers with a rating of 4 or 5, Pennsylvania centers with a rating of 4, and all Head Start centers. The low QRIS category includes: Arkansas centers with a rating of 1, Arizona centers with a rating of 2 or 3, and Pennsylvania centers with a rating of 2.

[b] Part of a larger organization means the center is part of and shares resources/staff with a larger organization, such as a YMCA, university, or other social services agency.

[c] Includes centers in which 50 percent or more of the children are supported by funding from CCDF and less than 30 percent of children are supported by funding from other federal, state, or local government sources.

[d] Includes centers in which less than 30 percent of the children are supported with funding from CCDF.

[e] Includes centers in which 90 percent or more of the children are supported through private tuition paid by their parents or guardians without any public funding.

n.a. = not applicable.

Exhibit B.8. Cost per childcare hour, excluding outlier (dollars)

	Number of centers	Annual child care hours provided	Cost per childcare hour	Range of cost per childcare hour	Standard deviation of cost per childcare hour
All centers					
Mean	25	183,178	5.47	n.a.	n.a.
Median	25	125,664	5.99	n.a.	n.a.
Maximum	25	665,397	10.77	n.a.	n.a.
Minimum	25	12,771	0.92	n.a.	n.a.
QRIS rating[a]					
High	18	194,410	6.59	2.89 - 10.77	2.20
Low	7	154,296	2.59	0.92 - 5.99	1.67
Age groups served					
Ages 0-5 and school-age	9	253,558	3.17	0.92 - 6.59	2.00
Ages 0-5 years only	9	175,294	6.30	2.98 - 9.97	2.42
Age 0-3 or ages 3-5 only	7	102,825	7.36	5.24 - 10.77	1.93
Program size					
Large center (serving 75 children or more)	14	260,019	5.63	1.78 - 10.77	2.88
Small center	11	85,380	5.27	0.92 - 8.92	2.67
Profit status					
For-profit	10	177,867	3.32	0.92 - 7.67	2.16
Not-for-profit	15	186,718	6.90	2.89 - 10.77	2.08
Embedded in a larger organization or part of a chain[b]					
Yes	12	155,763	6.26	2.89 - 10.77	2.27
No	13	208,484	4.75	0.92 - 9.97	3.01
Funding					
High subsidy[c]	6	93,281	4.56	1.79 - 5.99	1.83
Majority Head Start or state pre-K[d]	8	198,476	8.09	6.12 - 10.77	1.82
Majority private tuition[e]	7	105,465	4.19	0.92 - 7.78	2.73
Mixed public or mixed public/private	4	423,424	3.86	1.78 - 6.59	2.06

Source: Data collected for the Assessing Implementation and Cost of High-Quality Early Care and Education (ECE-ICHQ) Project, Phase 2.

Note: Costs were adjusted to account for state- and metro-level differences in average wages. Figures for subgroups are means. Means are calculated at the center level among centers in each group.

[a] The high QRIS category includes: Arkansas centers with a rating of 3, Arizona centers with a rating of 4 or 5, Pennsylvania centers with a rating of 4, and all Head Start centers. The low QRIS category includes: Arkansas centers with a rating of 1, Arizona centers with a rating of 2 or 3, and Pennsylvania centers with a rating of 2.

[b] Part of a larger organization means the center is part of and shares resources/staff with a larger organization, such as a YMCA, university, or other social services agency.

[c] Includes centers in which 50 percent or more of the children are supported by funding from CCDF and less than 30 percent of children are supported by funding from other federal, state, or local government sources.

[d] Includes centers in which less than 30 percent of the children are supported with funding from CCDF.

[e] Includes centers in which 90 percent or more of the children are supported through private tuition paid by their parents or guardians without any public funding.

Exhibit B.9. Cost per child care hour (dollars) and percentage allocation by key function, all centers

	Number of centers	Total cost per child care hour	Instruction and Caregiving		Instructional Planning, Coordination, and Child Assessment		Workforce Development		Child and Family Support		Center Administration and Planning	
			Cost per child care hour	% Allocation	Cost per child care hour	% Allocation	Cost per child care our	% Allocation	Cost per child care hour	% Allocation	Cost per child care hour	% Allocation
All centers												
Mean	26	6.36	1.90	32	1.43	22	0.58	9	0.60	10	1.84	27
Median	26	5.99	1.82	32	1.00	22	0.41	9	0.32	10	1.29	26
Maximum	26	28.57	5.15	53	6.59	35	3.46	16	2.32	34	11.32	42
Minimum	26	0.92	0.12	13	0.32	14	0.08	3	0.02	0	0.22	10
QRIS rating[a]												
High	19	7.75	2.26	32	1.74	22	0.73	9	0.73	10	2.28	28
		(5.48)	(1.01)	(8.78)	(1.39)	(5.73)	(0.72)	(2.59)	(0.66)	(7.06)	(2.34)	(8.75)
Low	7	2.59	0.94	34	0.58	24	0.18	8	0.25	11	0.64	23
		(1.67)	(0.70)	(13.48)	(0.36)	(5.69)	(0.14)	(4.23)	(0.31)	(11.27)	(0.58)	(7.08)
Age groups served												
Ages 0-5 and school-age	9	3.17	1.06	32	0.65	22	0.25	8	0.32	11	0.90	27
		(2.00)	(0.73)	(10.79)	(0.35)	(5.76)	(0.16)	(3.13)	(0.36)	(9.98)	(0.68)	(7.17)
Ages 0-5 years only	10	8.53	2.52	34	1.91	23	0.90	9	0.88	11	2.32	23
		(7.40)	(1.23)	(11.34)	(1.71)	(4.99)	(0.97)	(3.38)	(0.77)	(6.31)	(3.23)	(8.53)
Age 0-3 or ages 3-5 only	7	7.36	2.11	30	1.75	22	0.57	8	0.57	9	2.36	32
		(1.93)	(0.56)	(7.34)	(1.02)	(7.16)	(0.23)	(2.54)	(0.56)	(9.06)	(0.90)	(7.78)
Program size												
Large center (serving 75 children or more)	15	7.16	1.90	30	1.66	23	0.67	8	0.82	13	2.12	26
		(6.54)	(1.19)	(9.36)	(1.59)	(5.20)	(0.83)	(2.75)	(0.68)	(8.13)	(2.72)	(8.25)
Small center	11	5.27	1.91	36	1.11	22	0.47	9	0.31	7	1.46	27
		(2.67)	(1.03)	(10.41)	(0.70)	(6.45)	(0.32)	(3.48)	(0.41)	(7.04)	(0.89)	(9.06)
Profit status												
For-profit	10	3.32	1.09	32	0.76	24	0.26	8	0.30	11	0.91	25
		(2.16)	(0.78)	(11.60)	(0.47)	(6.10)	(0.22)	(3.91)	(0.29)	(9.27)	(0.80)	(7.78)
Not-for-profit	16	8.26	2.41	32	1.85	21	0.79	9	0.79	10	2.42	27
		(5.78)	(0.97)	(9.30)	(1.48)	(5.31)	(0.76)	(2.48)	(0.71)	(7.70)	(2.51)	(8.93)

	Number of centers	Total cost per child care hour	Instruction and Caregiving		Instructional Planning, Coordination, and Child Assessment		Workforce Development		Child and Family Support		Center Administration and Planning	
			Cost per child care hour	% Alloca-tion	Cost per child care hour	% Alloca-tion	Cost per child care our	% Alloca-tion	Cost per child care hour	% Alloca-tion	Cost per child care hour	% Alloca-tion
Embedded in a larger organization or part of a chain[b]												
Yes	13	7.97 (6.56)	2.27 (1.18)	32 (10.25)	1.88 (1.62)	23 (5.78)	0.79 (0.84)	9 (2.54)	0.64 (0.60)	8 (6.19)	2.39 (2.83)	27 (9.91)
No	13	4.75 (3.01)	1.54 (0.93)	33 (10.17)	0.98 (0.67)	22 (5.64)	0.38 (0.33)	8 (3.47)	0.56 (0.67)	12 (9.71)	1.29 (0.92)	26 (6.97)
Funding												
High subsidy[c]	6	4.56 (1.83)	1.64 (0.84)	35 (7.18)	0.79 (0.33)	18 (3.41)	0.33 (0.18)	7 (2.45)	0.54 (0.51)	13 (12.66)	1.25 (0.70)	27 (9.10)
Majority Head Start or state pre-K[d]	9	10.37 (7.04)	2.53 (1.14)	27 (7.35)	2.50 (1.70)	24 (4.77)	0.99 (0.96)	9 (2.68)	1.08 (0.74)	12 (7.31)	3.27 (3.15)	29 (9.82)
Majority private tuition[e]	7	4.19 (2.73)	1.70 (1.22)	38 (14.58)	1.03 (0.62)	27 (6.02)	0.44 (0.38)	10 (4.38)	0.15 (0.12)	5 (5.27)	0.88 (0.69)	20 (5.10)
Mixed public or mixed public/private	4	3.86 (2.06)	1.25 (0.71)	32 (1.29)	0.70 (0.29)	19 (2.63)	0.30 (0.16)	8 (0.76)	0.42 (0.34)	10 (3.69)	1.19 (0.61)	31 (2.60)

Source: Data collected for the Assessing Implementation and Cost of High-Quality Early Care and Education (ECE-ICHQ) Project, Phase 2.

Note: Costs were adjusted to account for state- and metro-level differences in average wages. Figures for subgroups are means. Means are calculated at the center level among centers in each group. Figures in parentheses are standard deviations.

[a] The high QRIS category includes: Arkansas centers with a rating of 3, Arizona centers with a rating of 4 or 5, Pennsylvania centers with a rating of 4, and all Head Start centers. The low QRIS category includes: Arkansas centers with a rating of 1, Arizona centers with a rating of 2 or 3, and Pennsylvania centers with a rating of 2.

[b] Part of a larger organization means the center is part of and shares resources/staff with a larger organization, such as a YMCA, university, or other social services agency.

[c] Includes centers in which 50 percent or more of the children are supported by funding from CCDF and less than 30 percent of children are supported by funding from other federal, state, or local government sources.

[d] Includes centers in which less than 30 percent of the children are supported with funding from CCDF.

[e] Includes centers in which 90 percent or more of the children are supported through private tuition paid by their parents or guardians without any public funding.

Exhibit B.10. Cost per childcare hour (dollars) and percentage allocation by key function, excluding outlier

	Number of centers	Total cost per child care hour	Instruction and Caregiving		Instructional Planning, Coordination, and Child Assessment		Workforce Development		Child and Family Support		Center Administration and Planning	
			Cost per child care hour	% Allocation	Cost per child care hour	% Allocation	Cost per child care our	% Allocation	Cost per child care hour	% Allocation	Cost per child care hour	% Allocation
All centers												
Mean	25	5.47	1.77	33	1.22	22	0.47	8	0.54	10	1.46	26
Median	25	5.99	1.71	32	1.00	22	0.37	8	0.32	11	1.27	25
Maximum	25	10.77	3.41	53	3.65	35	1.07	16	2.32	34	4.07	42
Minimum	25	0.92	0.12	13	0.32	14	0.08	3	0.02	0	0.22	10
QRIS rating[a]												
High	18	6.59	2.10	32	1.47	22	0.58	9	0.66	10	1.78	27
		(2.20)	(0.75)	(8.38)	(0.76)	(5.89)	(0.29)	(2.54)	(0.60)	(7.24)	(0.85)	(8.50)
Low	7	2.59	0.94	34	0.58	24	0.18	8	0.25	11	0.64	23
		(1.67)	(0.70)	(13.48)	(0.36)	(5.69)	(0.14)	(4.23)	(0.31)	(11.27)	(0.58)	(7.08)
Age groups served												
Ages 0-5 and school-age	9	3.17	1.06	32	0.65	22	0.25	8	0.32	11	0.90	27
		(2.00)	(0.73)	(10.79)	(0.35)	(5.76)	(0.16)	(3.13)	(0.36)	(9.98)	(0.68)	(7.17)
Ages 0-5 years only	9	6.30	2.23	36	1.39	23	0.61	9	0.75	11	1.32	21
		(2.42)	(0.86)	(10.37)	(0.51)	(5.29)	(0.37)	(3.44)	(0.69)	(6.57)	(0.68)	(6.50)
Age 0-3 or ages 3-5 only	7	7.36	2.11	30	1.75	22	0.57	8	0.57	9	2.36	32
		(1.93)	(0.56)	(7.34)	(1.02)	(7.16)	(0.23)	(2.54)	(0.56)	(9.06)	(0.90)	(7.78)
Program size												
Large center (serving 75 children or more)	14	5.63	1.67	31	1.31	23	0.47	8	0.73	13	1.46	25
		(2.88)	(0.81)	(9.09)	(0.85)	(5.39)	(0.31)	(2.62)	(0.61)	(8.29)	(1.00)	(7.66)
Small center	11	5.27	1.91	36	1.11	22	0.47	9	0.31	7	1.46	27
		(2.67)	(1.03)	(10.41)	(0.70)	(6.45)	(0.32)	(3.48)	(0.41)	(7.04)	(0.89)	(9.06)
Profit status												
For-profit	10	3.32	1.09	32	0.76	24	0.26	8	0.30	11	0.91	25
		(2.16)	(0.78)	(11.60)	(0.47)	(6.10)	(0.22)	(3.91)	(0.29)	(9.27)	(0.80)	(7.78)
Not-for-profit	15	6.90	2.23	33	1.53	21	0.61	9	0.70	10	1.83	27
		(2.08)	(0.66)	(8.77)	(0.80)	(5.48)	(0.28)	(2.41)	(0.65)	(7.94)	(0.85)	(8.61)
Embedded in a larger organization or part of a chain[b]												

	Number of centers	Total cost per child care hour	Instruction and Caregiving		Instructional Planning, Coordination, and Child Assessment		Workforce Development		Child and Family Support		Center Administration and Planning	
			Cost per child care hour	% Allocation	Cost per child care hour	% Allocation	Cost per child care our	% Allocation	Cost per child care hour	% Allocation	Cost per child care hour	% Allocation
Yes	12	6.26 (2.27)	2.03 (0.83)	33 (9.79)	1.49 (0.82)	23 (6.03)	0.57 (0.27)	9 (2.50)	0.53 (0.45)	9 (6.45)	1.64 (0.95)	26 (9.60)
No	13	4.75 (3.01)	1.54 (0.93)	33 (10.17)	0.98 (0.67)	22 (5.64)	0.38 (0.33)	8 (3.47)	0.56 (0.67)	12 (9.71)	1.29 (0.92)	26 (6.97)
Funding												
High subsidy[c]	6	4.56 (1.83)	1.64 (0.84)	35 (7.18)	0.79 (0.33)	18 (3.41)	0.33 (0.18)	7 (2.45)	0.54 (0.51)	13 (12.66)	1.25 (0.70)	27 (9.10)
Majority Head Start or state pre-K[d]	8	8.09 (1.82)	2.21 (0.62)	28 (7.08)	1.98 (0.78)	24 (5.08)	0.69 (0.29)	8 (2.51)	0.96 (0.68)	12 (7.61)	2.26 (0.96)	28 (9.62)
Majority private tuition[e]	7	4.19 (2.73)	1.70 (1.22)	38 (14.58)	1.03 (0.62)	27 (6.02)	0.44 (0.38)	10 (4.38)	0.15 (0.12)	5 (5.27)	0.88 (0.69)	20 (5.10)
Mixed public or mixed public/private	4	3.86 (2.06)	1.25 (0.71)	32 (1.29)	0.70 (0.29)	19 (2.63)	0.30 (0.16)	8 (0.76)	0.42 (0.34)	10 (3.69)	1.19 (0.61)	31 (2.60)

Source: Data collected for the Assessing Implementation and Cost of High-Quality Early Care and Education (ECE-ICHQ) Project, Phase 2.

Note: Costs were adjusted to account for state- and metro-level differences in average wages. Figures for subgroups are means. Means are calculated at the center level among centers in each group. Figures in parentheses are standard deviations.

[a] The high QRIS category includes: Arkansas centers with a rating of 3, Arizona centers with a rating of 4 or 5, Pennsylvania centers with a rating of 4, and all Head Start centers. The low QRIS category includes: Arkansas centers with a rating of 1, Arizona centers with a rating of 2 or 3, and Pennsylvania centers with a rating of 2.

[b] Part of a larger organization means the center is part of and shares resources with a larger organization, such as a university or social services agency.

[c] Includes centers in which 50 percent or more of the children are supported by funding from CCDF and less than 30 percent of children are supported by funding from other federal, state, or local government sources.

[d] Includes centers in which less than 30 percent of the children are supported with funding from CCDF.

[e] Includes centers in which 90 percent or more of the children are supported through private tuition paid by their parents or guardians without any public funding.

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

APPENDIX C

Measurement of Each Function
Across Types of Data Collection

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

Exhibit C.1. Measurement across types of data collection: Structural Supports for Instruction and Caregiving

Implementation	Time Use	Cost
Ratio and group size • Ability to meet age-based group size requirements or standards • Ability to meet age-based teacher/child ratio requirements or standards **Materials and equipment** • Availability of equipment to support caregiving • Availability of materials to support instruction • Access to outdoor play area **Staff qualifications** • Education level with highest proportion of lead teachers • Proportion of lead teachers who have at least 18 hours of ECE coursework • Education level of education specialists/program directors • Proportion of education specialists/program directors who have at least 18 hours of ECE coursework • Proportion of education specialists/program directors who are certified • Education level with highest proportion of assistant teachers • Proportion of assistant teachers who have at least 18 hours of ECE coursework **Ability to meet teacher qualifications and retain teachers** • Ability to meet targeted qualifications for lead teachers • Factors that influence lead teachers to stay at the center • Approximate proportion of teacher turnover	• **Providing instruction or care.** Includes one-on-one or group interaction with children during play or planned activities; leading a lesson; helping children during mealtimes; naptimes; or transition times, or field trips. Include naptimes only if you are not also conducting another activity.	• Percentage of salaries of staff who spend time providing instruction or care • Contracted services (such as substitute teachers), classroom materials and supplies, equipment (such as classroom furniture or playground equipment), and other direct costs used for providing instruction or care (such as costs for field trips) • Allocations of facilities, equipment, and overhead costs

Exhibit C.2. Measurement across types of data collection: Instructional Planning, Coordination, and Child Assessment

Implementation	Time Use	Cost
Selection, length of use, and training on curriculum • Criteria for selecting or developing infant/toddler curriculum • Criteria for selecting or developing preschool curriculum • Duration of main curriculum used for preschoolers • Method for training center leadership on the curriculum • Method for training teaching staff on the curriculum • Frequency of curriculum training for teaching staff **Planning time for teachers** • Existence of paid planning time for staff • Frequency and duration of paid planning time for staff **Screening and child assessment** • Type and training of staff who administer developmental screening/assessment • Duration of main screening/diagnostic tool use • Duration of use of main measurement tool for formative assessment • Criteria for selecting formative assessment tool • Teacher use of information gathered through formative assessments to guide instruction • Process for tracking child's instructional assessments • Criteria for selecting summative assessment tool • Method of training teaching staff on the assessment tool(s) **Parent communication about children's learning** • Frequency and format with which the center provides descriptive information formally to parents • Frequency and format with which the center provides information to parents about learning and development of their child compared to national norms • Length and frequency of parent-teacher conferences	• **Curriculum planning and development.** Includes periodic activities related to curriculum review, selection, and development, including ensuring alignment with standards or program requirements. • **Planning curriculum, activities, and lessons.** Includes developing or reviewing lesson plans and other forms of instructional planning; preparing and setting up classroom materials. • **Conducting child assessments during or outside of classroom time.** Includes assessments to determine whether a child needs further developmental assessment and/or services, to assess growth and development, or to determine children's needs and plan instruction; time spent coordinating assessments and administering, scoring, and reviewing results; documenting children's progress such as recording anecdotes and observations in a log. • **Periodic child assessments and parent conferences.** Includes preparing for, conducting, and reporting periodic assessments of children's progress, growth, or development; preparing for and conducting periodic parent conferences to discuss the care, activities, and progress of children.	• Percentage of salaries of staff who spend time on: – Curriculum planning and development – Planning curriculum, activities, and lessons – Conducting child assessments during or outside of classroom time – Periodic child assessments and parent conferences • Curriculum and assessment materials • Other contracted services, materials and supplies, equipment, and other direct costs used for instructional planning, coordination, and child assessment • Allocations of facilities, equipment, and overhead costs

Exhibit C.3. Measurement across types of data collection: Center Administration and Planning

Implementation	Time Use	Cost

Structural and procedural supports
 	
• Facility improvement needs
• Certification of center director
• Turnover in center staff other than teaching staff
• Process for informing teaching staff of licensing requirements or standards the center must follow
• Process for monitoring compliance with requirements and standards
• Facilitators in ability to change policies and procedures or make improvements
 	
Staff engagement in center planning and improvement
 	
• Staff discussion of quality assurance/need for quality improvement
• Existence of strategic planning activities and level of staff involvement in strategic planning efforts
• Existence of quality improvement activities and level of staff involvement in quality improvement efforts
• Existence of program evaluation activities and level of staff involvement in program evaluation efforts
 	
Setting and tracking center-level goals
 	
• Concepts emphasized in center mission
• Types of goals set for the center
• Frequency of reviewing progress toward center goals
• Use of data in tracking center goals
• Intentionality in creating action plans to achieve goals and communicating progress to center staff
 	
Collaborative center oversight (boards)
 	
• Existence of board and focus of input to center operations
• Level of board oversight of center operations
• Existence of a parent policy council and frequency of their meetings
• Focus of parent group's input to center operations
• Center director participation on community boards/councils
	• **Planning and goal-setting.** Includes assessing center performance, developing written plans, conducting center self-assessments of quality, pursuing quality improvement grants, participating in planning or board meetings, and marketing and fundraising activities.
	• **Ongoing regulatory compliance and reporting.** Includes daily activities to ensure and maintain compliance with federal, state, and city agencies or programs (for example, monitoring child-staff ratios, or completely injury report forms); activities related to accreditation or participation in Early Head Start/Head Start or a state pre-K program. Regulations may include those related to licensing, subsidy, quality rating and improvement systems (QRIS), Child and Adult Care Food Program (CACFP), and Individuals with Disabilities Education Act (IDEA).
	• **Managing center finances.** Includes collecting payments, budgeting, payroll, accounting, and tracking expenditures.
	• **Managing and maintaining center operations.** Includes overseeing staff schedules, the work of facilities and maintenance staff, working with vendors, and organizing space.
	• **Periodic regulatory compliance and reporting.** Includes periodic activities to prepare for and complete annual monitoring requirements of federal, state, city agencies or programs; activities related to accreditation or participation in Early Head Start/Head Start or a state pre-K program. Regulations may include those related to licensing, subsidy, quality rating and improvement systems (QRIS), Child and Adult Care Food Program (CACFP), and Individuals with Disabilities Education Act (IDEA).
		• Percentage of salaries of staff who spend time on:
		  – Planning and goal-setting
		  – Ongoing regulatory compliance and reporting
		  – Managing center finances
		  – Managing and maintaining center operations
		  – Periodic regulatory compliance and reporting
		• Contracted services (such as legal, technology, and program evaluation services), materials and supplies (such as office and janitorial supplies), equipment (such as office furniture), and other direct costs used for center administration or planning (such as licensing fees, insurance, and taxes)
		• Allocations of facilities, equipment, and overhead costs

Exhibit C.4. Measurement across types of data collection: Workforce Development

Implementation	Time Use	Cost

Interviewing and evaluating teaching staff
 	
• Involvement of teaching staff in interview processes for teacher candidates
• Interview activities for lead teacher candidates
• Typical individual performance goals set for teaching staff
• Process for determining individual performance goals for teaching staff
• Center processes for monitoring and evaluating performance of teaching staff
• Individual(s) that provide input into teaching staff performance
 	
Setting priorities for and assessing training and professional development for teaching staff
 	
• Process for determining training and professional development priorities for teaching staff
• Center's priorities for training and professional development of teaching staff
• Staff level of formal involvement in determining individual training and professional development needs
• Formality of the process for determining if training meets center goals and staff needs
• Type of activities used to assess knowledge acquisition or skill development following training
 	
Supports for training and professional development of teaching staff
 	
• Amount of training and professional development days provided by the center
• Center provides financial or other assistance for teachers to support external training or education
• Formality of supplemental assistance for teachers to support external training or education
• Center's approach to compensating staff for time spent attending external training
• Center's approach to providing classroom coverage when teaching staff attend training
	• **Staff supervision, coaching, or mentoring.** Includes scheduled or unscheduled consultation with a supervisor or supervisees and providing or receiving coaching or mentoring either one-on-one or in groups; time spent conducting or discussing classroom observations; time spent providing coaching, mentoring, or technical assistance (TA) to other center staff.
	• **Recruiting and hiring teaching staff.** Includes preparing job listings, reviewing applications, screening, and interviewing.
	• **Evaluating staff performance.** Includes time spent preparing for and participating in periodic reviews to evaluate performance or set goals for the coming year, including developing or reviewing training or professional development plans.
	• **Training and professional development activities.** Includes time spent preparing, conducting, or participating (in-person or online) in training or professional development activities that occur during scheduled work hours, including planned professional development days. Excludes time spent pursuing a degree or credential outside of work hours.
	• **Periodic coaching, mentoring, or technical assistance.** Includes time spent receiving targeted or periodic coaching, mentoring, or TA that occurs irregularly and is not part of ongoing supervision or program support activities within the center; time spent providing coaching, mentoring, or TA to other center staff.
	• **Hours spent taking classes.** Includes credit hours in the past 12 months. Excludes time spent fulfilling coursework such as doing homework.
		• Percentage of salaries of staff who spend time on:
		  – Staff supervision, coaching, or mentoring
		  – Recruiting and hiring teaching staff
		  – Evaluating staff performance
		  – Training and professional development activities
		  – Periodic coaching, mentoring, or technical assistance
		  – Hours spent taking classes
		• Contracted services (such as trainers or coaches), materials and supplies (such as supplies for staff training), equipment, and other direct costs used for workforce development (such as tuition reimbursement or travel expenses for professional development)
		• Allocations of facilities, equipment, and overhead costs

Implementation	Time Use	Cost
Classroom observations and coaching to support teaching staff • Goals of classroom observations • Extent to which teaching staff visit other classrooms to observe practice and where • Activities following teachers' observations of other classrooms • Occurrence of individualized coaching for teaching staff • Activities involved in individualized coaching		

Exhibit C.5. Measurement across types of data collection: Child and Family Support

Implementation	Time Use	Cost
Total services (in center or referral) • Number of services for which the center provides access • Count of categories of services for which the center provides access **Service provision (in center)** • Count of all services explicitly provided by the center with a cost • Count of all services provided in the center but by another entity and at no cost • Count of all categories of services provided in the center (cost or no cost) **Process for deciding on and tracking service receipt** • Process for deciding if therapeutic services are needed • Process for deciding if counseling services are needed • Collaboration in deciding if health services are needed (who decides) • Collaboration in deciding if counseling services are needed (who decides) • Collaboration in deciding if social services are needed (who decides) • Collaboration in deciding if adult or parent education or services are needed (who decides) • Frequency of tracking therapeutic service receipt • Frequency of tracking therapeutic service receipt • Frequency of tracking social service receipt • Method for tracking health service receipt • Method for tracking therapeutic service receipt • Method for tracking counseling service receipt • Method for tracking adult or parent education or services receipt	• **Planning and conducting family engagement or family support activities.** Includes events for currently enrolled families (such as back-to-school nights or class performances), planning or providing parent education or other family support services, and conducting home visits.	• Percentage of salaries of staff who spend time planning and conducting family engagement or family support activities • Contracted services (such as health and medical services), materials and supplies (such as health and hygiene items provided to families), equipment (such as health care equipment or vehicles needed to provide family support services), and other direct costs used for child and family support (such as parent/caregiver programs and trainings) • Allocations of facilities, equipment, and overhead costs

APPENDIX D

Distribution of Implementation Scores and Cost per Child Care Hour by Key Function

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

Exhibit D.1. Distribution of implementation scores and cost per child care hour for the Structural Supports for Instruction and Caregiving function, by quality level
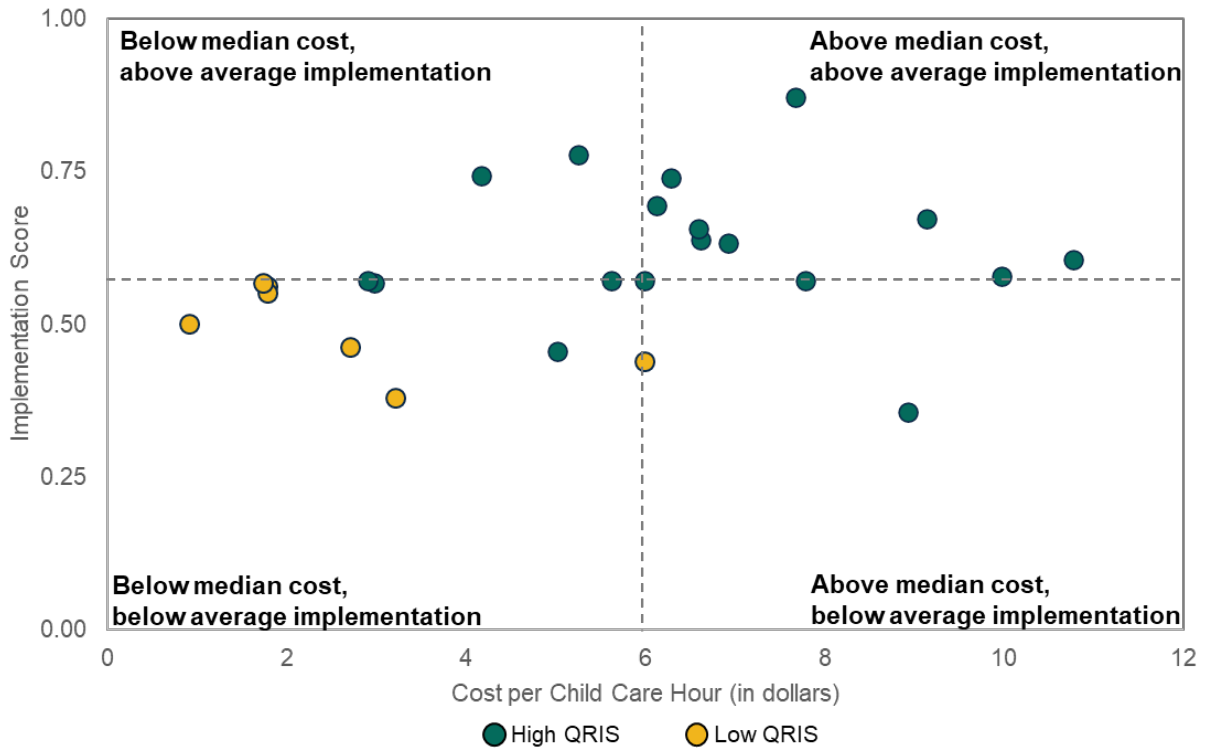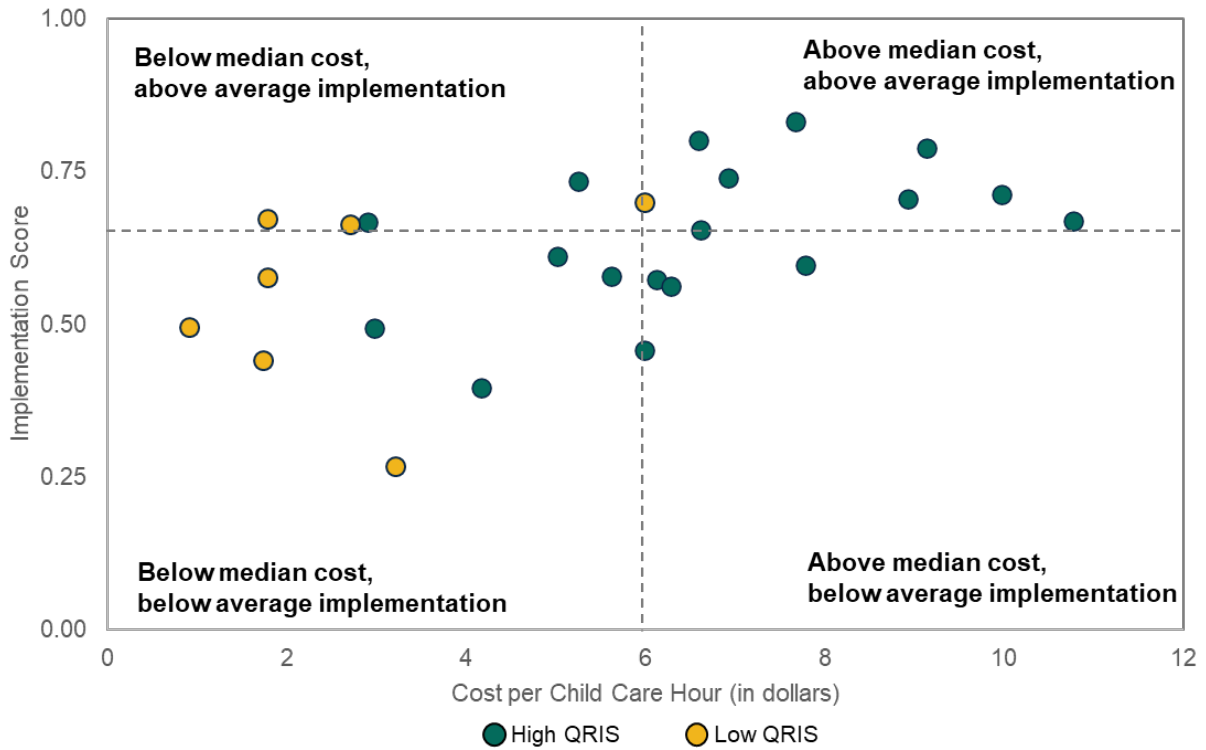

Source: Data collected for the Assessing Implementation and Cost of High Quality Early Care and Education (ECE-ICHQ) Project, Phase 2.

Note: Each dot represents one of the 25 centers in Phase 2 for which we have cost measures. We categorized centers into high and low QRIS ratings based on the requirements for the different levels in each state in the multi-case study.

Exhibit D.2. Distribution of implementation scores and cost per child care hour for the Instructional Planning, Coordination, and Child Assessment function, by quality level
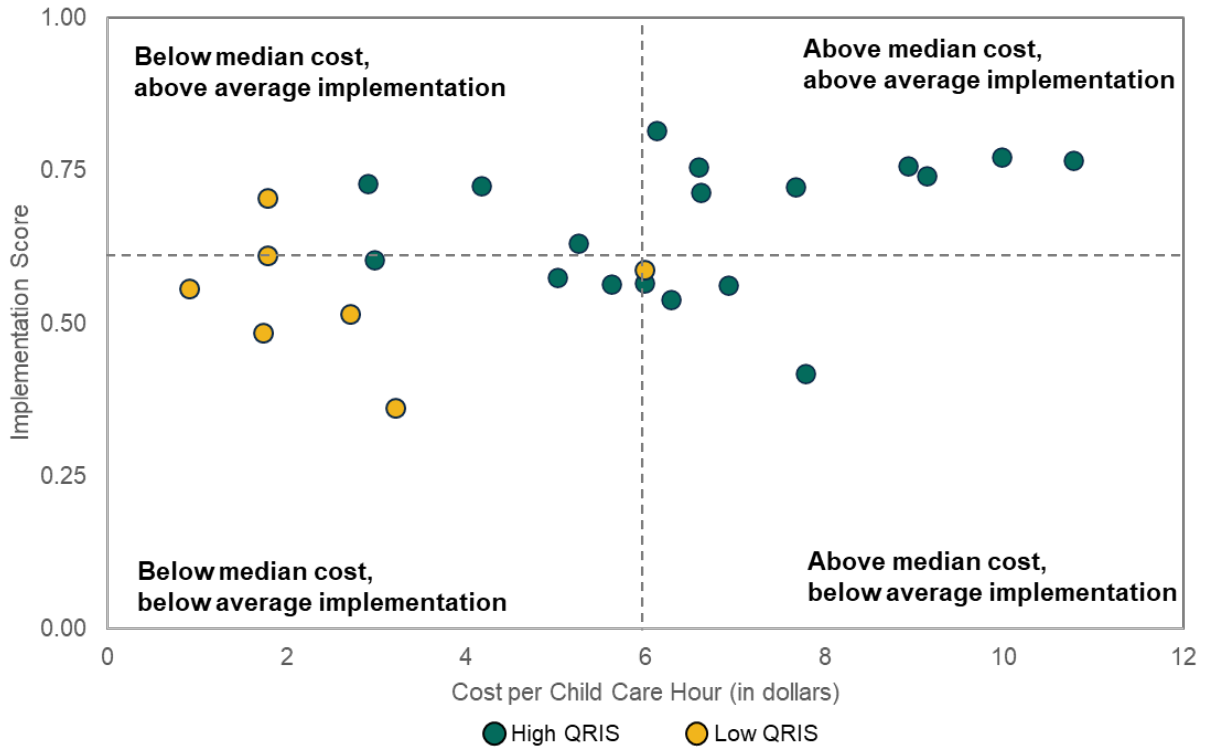
Source: Data collected for the Assessing Implementation and Cost of High Quality Early Care and Education (ECE-ICHQ) Project, Phase 2.

Note: Each dot represents one of the 25 centers in Phase 2 for which we have cost measures. We categorized centers into high and low QRIS ratings based on the requirements for the different levels in each state in the multi-case study.

Exhibit D.3. Distribution of implementation scores and total cost per child care hour for the Workforce Development function, by quality level

Source: Data collected for the Assessing Implementation and Cost of High Quality Early Care and Education (ECE-ICHQ) Project, Phase 2.

Note: Each dot represents one of the 25 centers in Phase 2 for which we have cost measures. We categorized centers into high and low QRIS ratings based on the requirements for the different levels in each state in the multi-case study.

Exhibit D.4. Distribution of implementation scores and total cost per child care hour for the Child and Family Support function, by quality level
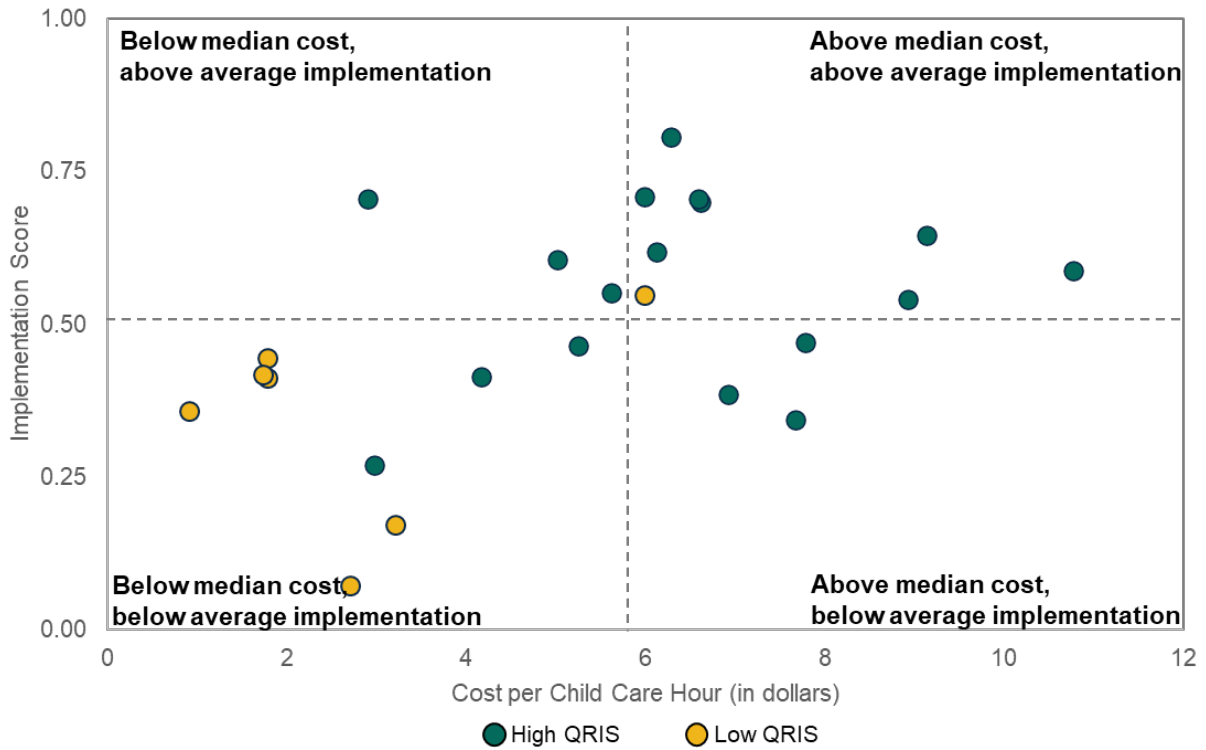
Source: Data collected for the Assessing Implementation and Cost of High Quality Early Care and Education (ECE-ICHQ) Project, Phase 2.

Note: Each dot represents one of the 25 centers in Phase 2 for which we have cost measures. We categorized centers into high and low QRIS ratings based on the requirements for the different levels in each state in the multi-case study.

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

Mathematica
Progress Together

mathematica.org