The Seeds to Success Modified Field Test: Findings from the Impact and Implementation Studies

June 28, 2010

Kimberly Boller
Patricia Del Grosso
Randall Blair
Yumiko Jolly
Ken Fortson
Diane Paulsell
Eric Lundquist
Kristin Hallgren
Martha Kovac

MATHEMATICA
Policy Research, Inc.

Contract Number:
5023

Mathematica Reference Number:
06298.166

Submitted to:

Bill & Melinda Gates Foundation
P.O. Box 23350
Seattle, WA 98102

Project Officer: Abbie Raikes

Submitted by:
Mathematica Policy Research
P.O. Box 2393
Princeton, NJ 08543-2393
Telephone: (609) 799-3535
Facsimile: (609) 799-0005
Project Director: Diane Paulsell

The Seeds to Success Modified Field Test: Findings from the Impact and Implementation Studies

June 28, 2010

Kimberly Boller
Patricia Del Grosso
Randall Blair
Yumiko Jolly
Ken Fortson
Diane Paulsell
Eric Lundquist
Kristin Hallgren
Martha Kovac

ACKNOWLEDGEMENTS

This report would not have been possible without the participation of dedicated staff working at Thrive by Five Washington (Thrive), Educational Service District 105's Ready by Five project in East Yakima, and Puget Sound Educational Service District and its partner in White Center, Child Care Resources. Most importantly, the study would not have happened without the participation of family child providers, child care centers, and coaches. These staff gave generously of their time and energy to answer our questions and let us observe and learn about their work. We gratefully acknowledge their contribution to the evaluation.

We would also like to thank others who contributed to the evaluation design and implementation and to this report. Abbie Raikes, Jodi Haavig, Sara Meyer, and Valisa Smith at the Bill & Melinda Gates Foundation provided guidance, support, and suggestions that shaped all stages of the research and this report. Brenda Blasingame, Danielle Kassow, Rachel Kim, Angel Cantu, and Megan McJennett at Thrive worked closely with us to design and conduct the evaluation. Cynthia Juarez and Angela Abrams at Ready by Five and Caroline Shelton and Adie Fatur at Child Care Resources provided important contextual input and feedback on the evaluation design and report.

Numerous staff at Mathematica played critical roles in the research effort. Margaret Caspe and Yumiko Jolly, aided by Bea Jones, Sumia Ibrahim, and our survey operation center and field staff, led the data collection efforts. Anthippy Petras worked with us to train field staff, and Anthippy and Delores Gonzales ensured the reliability of field staff. Louisa Tarullo and John Deke reviewed drafts and contributed thoughtful comments and suggestions. Cindy George, John Kennedy, and William Garrett oversaw the editing. Cindy McClure, Marjorie Mitchell, and Jill Miller skillfully produced the report.

CONTENTS

TABLES

FIGURES

EXECUTIVE SUMMARY

In 2006, the Bill & Melinda Gates Foundation launched the Early Learning Initiative (ELI) to improve the school readiness of Washington State's children through three main strategies: (1) development of high-quality, community-wide early learning initiatives in two communities; (2) enhancement of statewide systems that support early learning; and (3) support for implementation of promising practices. The foundation joined with other private funders and state officials to form Thrive by Five (Thrive) Washington to energize development and support of high-quality early learning opportunities for all children in the state.

In tandem with the formation of Thrive, the Gates Foundation sought two communities with a high level of need for early learning services and the capacity to develop and implement high-quality, community-wide early learning initiatives. The Gates Foundation selected East Yakima, a neighborhood in the central Washington community of Yakima, and White Center, an unincorporated area just outside Seattle. Thrive has worked with an intermediary agency in each community to develop and implement the initiative. Educational Service District 105, through its Ready by Five project, serves as the intermediary in East Yakima. In White Center, Puget Sound Educational Service District (PSESD) serves as the intermediary for the White Center Early Learning Initiative (WCELI).

Thrive, the two communities, and other stakeholders worked closely with the Washington State Department of Early Learning (DEL) to develop Seeds to Success, a child care quality rating and improvement system (QRIS). In spring 2009, Thrive, Ready by Five, and Child Care Resources (PSESD's partner in White Center) began a six-month field test of a streamlined version of the DEL QRIS, referred to as the Seeds to Success Modified Field Test (Seeds).

Evaluation of Seeds

In 2009, Mathematica Policy Research conducted an impact and implementation evaluation of Seeds. The impact evaluation was designed to determine whether the coaching model and financial incentives implemented as part of Seeds affected the quality of services provided by participating child care businesses (in both family home and center settings), compared with those businesses that did not receive Seeds. Across the two communities, Mathematica randomly assigned the 52 family child care providers and 14 centers that volunteered to participate into treatment and control groups. The treatment group received coaching, quality improvement grants, and funds for professional development opportunities and supports; the control group received funds only for professional development opportunities and supports. The goals of the implementation study were to determine whether Seeds met its goals of providing intensive, high-quality coaching and other supports to participating providers and to capture the lessons learned about implementation during the field test.

Overview of the Seeds to Success Model

For the field test, the Seeds model assessed two quality categories: (1) curriculum and learning environment, and (2) professional development and training. Each participating business received a Seeds rating based on its scores on the two quality categories. Quality ratings ranged from 1 to 5, with 5 being the highest. Businesses could achieve a 5 if they met the requirements in the curriculum and learning environment area (as measured by the Environment Rating Scales [ERS] and the ratio

of children to adults) and met the level 4 Seeds rating requirements in the two professional development and training areas (education and experience).

As part of the Seeds field test, participating providers (those assigned to the treatment group) received the following during the period June through December 2009:

- ***Coaching.*** Each family child care provider, center-based classroom, and center director was eligible to receive eight hours of coaching per month. Informed by their Seeds rating and ERS score, providers and coaches developed quality improvement plans (QIPs) to guide the coaching sessions.
- ***Quality Improvement Grants.*** Programs received quality improvement grants based on their Seeds rating, with higher-rated programs receiving more funding.
- ***Funds for Professional Development Opportunities, Plus Supports.*** Providers were offered professional development opportunities, including funds for training and course work, as well as funds to cover the costs of child care expenses, release time, and books.

Impact Study Main Findings

Data sources for the Seeds impact study included classroom observations, self-administered questionnaires for center directors and educators (lead and assistant teachers), and interviews with family child care providers. Observations conducted by Mathematica at the start (baseline) and at the end (followup) of the field test included the ERS, the Arnett Caregiver Interaction Scale, and counts of the children and adults. The observations and interview data collected at baseline demonstrated that the treatment and control groups were similar; this means that any differences between the two groups at followup were caused by Seeds. This section summarizes the impact study's main findings.

Impacts on Participation in Education and Professional Development Services

Over this relatively short period of time (June through December 2009), family child care providers in the treatment group were not more likely than providers in the control group to be enrolled in an education or training program. However, Seeds did improve center lead and assistant teacher enrollment in an education or training program, and significantly more lead teachers in the treatment group than in the control group attended college courses at least weekly. At followup, family child care providers and center-based lead and assistant teachers in the treatment group were much more likely than those in the control group to report visits from a coach at least weekly.

Impacts on Education and Experience Outcomes

Improvements in staff educational attainment and experience require sustained investments in supports for staff taking courses and in retaining staff in their positions. Because the Seeds evaluation follow-up data collection was only six months after the start of implementation of the Seeds model, the likelihood of observing impacts on educational attainment outcomes was low. Indeed, at followup, we found no consistent pattern of positive impacts of Seeds on family child care providers' educational attainment. More center-based teachers in the treatment group than in the control group earned three credits in the past six months, but Seeds had no impact on completion of a postsecondary degree for center lead teachers and assistants. Although Seeds significantly reduced center-based lead teacher turnover, treatment group lead teachers reported less experience than control group lead teachers at followup.

Impacts on Observed Quality

Child care businesses in the treatment group had significantly higher child care observed quality scores at followup than businesses in the control group. For both family child care providers and child care centers in the treatment group, the ERS total score and most of the ERS subscale scores were significantly higher than control group scores. Most of the differences in observed quality were larger than those reported in other recent studies of child care or preschool quality improvement interventions (Landry et al. 2009; Jackson et al. 2007; Neuman et al. 2009; Powell et al. 2010).

Impacts on Seeds Scores

Despite large differences in observed quality, the overall Seeds scores themselves were not affected by the coaching and quality improvement grants. In the Seeds model, a participating business had to meet all the requirements for curriculum and learning environment and professional development and training to qualify for a rating at that level (referred to as a building block QRIS model). Since no statistically significant impacts on the education and experience components of professional development were found, the very large gains in the ERS did not translate into gains in Seeds scores. For both types of child care businesses, the Seeds scores were not significantly different across treatment and control groups.

Implementation Study Main Findings

Data sources for the Seeds implementation study included (1) executive interviews and focus groups with site coordinators, coaches, and child care staff during site visits conducted in June and November 2009; and (2) service use data collected by coaches and site coordinators from June through December 2009 and analyzed by Mathematica in winter 2010. This section summarizes the implementation study's main findings.

Amount of Coaching Received

The coaches and providers were able to implement the intensive coaching component of the Seeds model. On average, the amount of coaching providers received adhered to the Seeds model; however, participants had a range of experiences. On average, family child care providers, center directors, and lead teachers received between 6 and 11 hours of coaching per month. Assistant teachers in White Center received, on average, less than two hours of coaching per month; in East Yakima, they received, on average, nearly six hours per month.

Uses of Quality Improvement Grants

Across all providers in East Yakima and White Center, quality improvement grants were most commonly used for supplies and materials. During focus groups, providers described using funds to pay for (1) materials, including books and art materials; (2) larger items, such as child-size tables and outdoor play equipment; and (3) safety improvements.

Access to Professional Development

According to administrative data, across communities, one-third to one-half of family child care providers and at least one staff person in each child care center received funding for professional development opportunities for their staff. Fewer providers accessed supports. During focus groups, providers described several barriers that deterred them from using the professional development

opportunities, including (1) limited availability of trainings and classes, particularly near their places of employment or homes; (2) lack of trainings and classes that provided new or relevant information; and (3) low perceived "payoffs" to professional development, because salaries were unlikely to increase as a result of completion of professional development.

Key Lessons

The Seeds impact and implementation evaluation provides important lessons for Washington State and future QRIS activities. The five key lessons are:

1. Implementing the Seeds model is feasible. Coaching and incentives are sufficient to motivate participation and, overall, providers said they found coaching very helpful.
2. Seeds significantly improved observed quality of child care.
3. The short time frame of the field test prohibited substantial gains in educational attainment, but Seeds did increase center-based teacher enrollment in training and educational activities.
4. Due to the building blocks scoring system that requires all standards at a given level to be met before a higher rating can be given, *large* improvements in observed quality did not translate into higher Seeds scores.
5. Rating a small number of areas, coupled with intensive coaching and quality improvement grants, can yield substantial child care quality improvements.

I. INTRODUCTION

In 2006, the Bill & Melinda Gates Foundation launched the Early Learning Initiative (ELI) to improve the school readiness of Washington State's children through three main strategies: (1) development of high-quality, community-wide early learning initiatives in two communities; (2) enhancement of statewide systems that support early learning; and (3) support for implementation of promising practices. The foundation joined with other private funders and state officials to form Thrive by Five (Thrive) Washington to energize development and support of high quality early learning opportunities for all children in the state.

In tandem with the formation of Thrive, the Gates Foundation sought two communities with a high level of need for early learning services and the capacity to develop and implement high-quality, community-wide early learning initiatives. After researching possibilities and consulting with community stakeholders, the Gates Foundation selected White Center, an unincorporated area just outside Seattle, and East Yakima, a neighborhood in the central Washington community of Yakima. Thrive has worked with an intermediary agency in each community to develop and implement the initiative. Educational Service District 105, through its Ready by Five (Rb5) project, serves as intermediary in East Yakima. In White Center, Puget Sound Educational Service District (PSESD) serves as the intermediary for the White Center Early Learning Initiative (WCELI). Three key partners—Child Care Resources (CCR), Open Arms Perinatal Services, and the Seattle/King County Department of Public Health—work with PSESD to manage the initiative and provide services.

Thrive, the two communities, and other stakeholders worked closely with the Washington State Department of Early Learning (DEL) to develop a child care quality rating and improvement system (QRIS). In spring 2009, Thrive and the communities began a field test of a streamlined version of the DEL QRIS system, referred to as the Seeds to Success Modified Field Test (Seeds).[1]

Evaluation of Seeds

In 2009, Mathematica Policy Research conducted an impact evaluation and an implementation study of Seeds. The impact evaluation was designed to determine whether the coaching model and financial incentives implemented as part of Seeds affect the quality of services provided by participating child care businesses (in both family home and center settings), compared with those businesses that did not receive Seeds. Across the two communities, Mathematica randomly assigned the 52 family child care providers and 14 centers that volunteered to participate in the Seeds field test into treatment and control groups.[2] The treatment group received coaching, financial grants, and professional development opportunities and supports; the control group received only professional development opportunities and supports. The goals of the implementation study were to determine whether Seeds implementation successfully met its goals of providing intensive, high-quality

[1] Participating providers initially volunteered to enroll in the DEL QRIS field test. However, due to state budget constraints, the field test was suspended. Thrive then decided to implement a modified version of the state model and invited providers who had volunteered for the state field test to participate.

[2] Given the changes to the field test approach, providers experienced a delay of several months from when they originally volunteered to participate to when the initiative actually began. The providers in the field test are those most motivated to enroll in the field test. They are not a census or a representative sample of all of the providers in the two communities.

coaching and other supports to participating providers and to capture the lessons learned by the participating agencies and providers. Other important study goals included documenting lessons learned about (1) collecting data required for the Seeds rating and (2) using the Seeds rating rubric.

Overview of the Seeds to Success Model

For the modified field test, the Seeds model has two quality categories: (1) curriculum and learning environment and (2) professional development and experience. Each participating business receives a Seeds rating based on its scores on the two quality categories (Table I.1). Quality ratings range from 1 to 5, with 5 being the highest. Businesses can achieve a 5 if they meet the ratio requirements in the curriculum and early learning area and meet the level 4 Seeds rating requirements in all other areas. The Seeds ratings for the modified field test are calculated using the following information:

1. A curriculum and learning environment score was calculated by taking an average of the Environment Rating Scale (ERS) items for family child care providers; for child care centers, a center-level score was calculated by taking the mean of the classroom-level ERS scores.[3,4]

2. A professional development and training score was based on the education and experience data reported by family child care providers and center-based staff (including directors, lead teachers, and assistants) on self-administered questionnaires and interviews.

In the Seeds model, a participating business must meet the requirements in every category of a level to qualify for a rating at that level. The overall Seeds ratings were then calculated by taking the minimum value among the curriculum and learning environment score and the professional development and training score.

As part of the Seeds field test, participating providers (those assigned to the treatment group) received the following:

- ***Coaching.*** Each family child care provider, center-based classroom, and center director was eligible to receive eight hours of coaching per month. Coaching hours for center classrooms were divided between lead teachers and assistants, with more hours intended for lead teachers. Providers and coaches developed quality improvement plans (QIPs) that were used to guide the coaching sessions. Box I.1 includes information about the coaching model used.

- ***Quality Improvement Grants.*** Programs received quality improvement grants based on their Seeds rating, with higher-rated programs receiving more funding (Table I.2).

[3] The version of the scale used during the observations depended on the setting: the Early Childhood Environmental Rating Scales–Revised (ECERS-R) and the Infant/Toddler Environmental Rating Scale–Revised (ITERS-R) were used in child care centers; the Family Child Care Environment Rating Scale–Revised (FCCERS-R) was used in family child care homes (Harms et al. 1998; Harms et al. 2003; Harms et al. 2007).

[4] During the field test, Mathematica conducted the ERS observations.

- ***Professional Development Opportunities and Funds, Plus Supports.*** Providers were offered professional development opportunities, including funds for training and course work. In addition, participating providers had access to funds to cover the costs of child care expenses, release time, and books.

Box I.1. Seeds Coaching Model

In 2008, Thrive developed the Consultative Coaching Program for Early Learning Professionals. This coaching program adapts the Coach Training Institute's Co-Active Coaching[5] model and incorporates Gary Bloom's Blended Coaching[6] techniques and principles of process consultation[7], all within the early learning context. The goal of the Consultative Coaching Program is to train coaches to develop a trusting relationship with early learning professionals so that they can help early learning professionals reflect on their practice (1) in the classroom or in their business and (2) during interactions with the other providers in that setting, with families, and with the children that are in their care. The Seeds coaches were also trained to help the treatment group professionals stay motivated to attain their quality improvement goals and to help establish skills and behaviors that support continuous quality improvement.

Road Map to the Report

This report provides a detailed description of Seeds. Chapter II provides an overview of the impact and implementation studies that comprised the evaluation; specifically, we describe the study designs, data collection processes, and analysis methods. In Chapter III, we present findings from the impact evaluation. In Chapter IV, we discuss findings from the implementation study. Chapter V presents lessons learned about implementing the Seeds model. Appendix A contains additional details about the evaluation methodology.

[5] The Co-Active Coaching model is a relationship-based approach designed to create a relationship between the coach and the staff she works with. When applying the Co-Active Coaching model, coaches are to advise or problem solve together with the coaching recipient (Coaches Training Institute 2010).

[6] The Blended Coaching model is based on more than 15 years of field work at the New Teacher Center, University of California, Santa Cruz. When applying the model, coaches are to move between facilitative and instructional approaches in their practice (Bloom et al. 2005).

[7] Process consultation rests on the belief that the client has the answers to questions about their practice, not the consultant/coach, and that the job of the consultant/coach is to help by facilitating the client's thinking and action rather than imposing their own thinking on the client and her work (Schein 1969).

Table I.1. Seeds Areas Assessed for the Modified Field Test: Requirements by Seeds Level

	Curriculum and Learning Environment
Seed Level 1[a]	Facility that is licensed by DEL or a Tribal Nation and cares for young children (not only school-age children). To be eligible, the facility cannot have a license that is revoked or suspended.
Seed Level 2	Overall average score of 3 or more on the Environment Rating Scale[b]
Seed Level 3	Overall average score of 4 or more on the Environment Rating Scale
Seed Level 4	Overall average score of 5 or more on the Environment Rating Scale
Seed Level 5[c]	Overall average score of 5 or more on the Environment Rating Scale, plus meets ratio requirements:[d] **Family Child Care:** If more than 4 children under 36 months, provider must have an assistant. If more than 8 children total, provider must have an assistant. **Centers**: Infant—1 staff: 3 children Toddler—1 staff: 5 children Preschool—1 staff: 9 children
	Professional Development and Training
	Education[e]
Seed Level 1	Facility that is licensed by DEL or a Tribal Nation and cares for young children (not only school-age children). To be eligible, the facility cannot have a license that is revoked or suspended.
Seed Level 2	***Family Home Provider*** Enrolled in CDA program/ECE credits to meet CDA requirement, or AA-level degree program ***Center Director*** AA-level degree in related field, or 60 college credits ***Center Lead Educator*** 25% have AA-level degree in related field, or 25% have CDA or 15 ECE credits ***Center Assistant Educator*** 50% have high school diploma/GED
Seed Level 3	***Family Home Provider*** CDA or enrolled in AA-level degree program with at least 3 credits earned in last 6 months ***Center Director*** AA-level degree in related field and enrolled in BA program with at least 3 credits earned in last 6 months or has 90 credits or meets the alternative pathway to NAEYC/NAA Director qualifications ***Center Lead Educator*** 50% have CDA or 30 credits in ECE/related field, or 50% have AA-level degree in related field ***Center Assistant Educator*** 50% enrolled in a CDA/AA-level degree program with at least 3 credits earned in last 6 months
Seed Level 4	***Family Home Provider*** AA-level degree or BA in ECE/related field ***Center Director*** BA in ECE/related field or meets the alternative pathway to NAEYC/NAA Director qualifications[f] ***Center Lead Educator*** 100% have AA-level degree in related field, or 50% have BA ***Center Assistant Educator*** 50% have CDA or 15 credits in ECE/related field; the other 50% working toward it
	Experience
Seed Level 1	Facility that is licensed by DEL or a Tribal Nation and cares for young children (not only school-age children). To be eligible, the facility cannot have a license that is revoked or suspended.

Table I.1. *(continued)*

Seed Level 2	***Family Home Provider*** At least 1 year of verifiable experience ***Center Director*** 6 months of verifiable supervisory experience[g] ***Center Lead Educator*** 6 months of verifiable experience ***Center Assistant Educator*** No additional requirement
Seed Level 3	***Family Home Provider*** At least 2 years of verifiable experience ***Center Director*** 6 months of verifiable supervisory experience ***Center Lead Educator*** 6 months of verifiable experience ***Center Assistant Educator*** 3 months of verifiable experience
Seed Level 4	***Family Home Provider*** At least 3 years of verifiable experience ***Center Director*** 1 year of verifiable supervisory experience ***Center Lead Educator*** 1 year of verifiable experience ***Center Assistant Educator*** 6 months of verifiable experience

Source: Seeds to Success: Washington State's Quality Rating and Improvement System, Washington State Department of Early Learning.

[a] For the modified field test, all child care businesses recruited into the study were considered to have met the Seed 1 criteria.

[b] The Environment Rating Scales are designed for use with different age groups and types of care settings: the Infant/Toddler Environment Rating Scale-Revised (ITERS-R; Harms et al. 2003) was used in child care center classrooms with infants and toddlers up to 30 months; the Early Childhood Environment Rating Scale-Revised (ECERS-R; Harms et al. 1998) was used in child care center classrooms with children ages 2-1/2 to 5 years old; and the Family Child Care Environment Rating Scale-Revised (FCCERS-R; Harms et al. 2007) was used in family provider homes.

[c] Five Seeds exists only for the curriculum and learning environment rating area. For the modified field test, a child care business earning 5 Seeds in this area and 4 Seeds in the professional development and training area would have a 5 Seed rating overall.

[d] For the modified field test, child-adult ratios were based on the observed counts of children and adults at the time of the child care observation. The current Seeds model has different ratio requirements for infants, toddlers, and preschool children in centers. For the field test, we computed separate ratios by child age for each classroom (an infant-adult ratio, toddler-adult ratio, and preschool child-adult ratio). A rating of 5 was awarded if the ERS Seeds score was a 4 for the center and the maximum of each of the three ratios computed for any classroom met the ratio requirements.

[e] For the education rating area, a provider/director/educator could obtain a given Seed rating either by explicitly satisfying the associated academic criteria or by achieving a higher-level degree than required. See Chapter V for more details.

[f] In the current Seeds model, a center director who meets the alternative pathway to NAEYC/NAA Director qualifications could earn either 3 or 4 Seeds. For the modified field test, a director meeting these qualifications earned 4 Seeds.

[g] In the current Seeds model, the center director experience requirements are flipped: a director with 1 year of experience would earn 2 or 3 Seeds, while a director with only 6 months of experience would earn 4 Seeds. We corrected this for the modified field test.

AA = Associate's degree; BA = Bachelor's degree; CDA = child development associate; DEL = Department of Early Learning; ECE = early childhood education; NAA = National Afterschool Association; NAEYC = National Association for the Education of Young Children.

Table I.2. Quality Improvement Grants, by Seed Level and Program Type [a]

Seed Level	Family Child Care Providers (Annual Maximum Amount)	Child Care Centers (Annual Maximum Amount)
1	$1,200	$1,800
2	$3,600	$4,800
3	$5,400	$9,000
4	$7,200	$10,800
5	$9,000	$12,600

Source: Seeds to Success: Washington State's Quality Rating and Improvement System, Washington State Department of Early Learning.

[a] For the field test, providers received quality improvement grants that were prorated to account for Seeds lasting less than 12 months.

II. DESIGN OF THE SEEDS IMPACT AND IMPLEMENTATION EVALUATIONS

This chapter describes the impact and implementation studies that comprise the evaluation of Seeds. We begin by describing our research questions. We then provide an overview of the evaluation design and study participants, followed by a summary of the data collection and analytic methods. In each section, we present the summary for the impact study first, followed by the implementation study. Appendix A contains additional details about the evaluation design.

Research Questions

Mathematica's evaluation addressed research questions in two main areas: (1) the impact of Seeds on child care quality, and (2) the implementation of Seeds. By addressing these questions, the evaluation informs program management, program development, and policy decisions. The research questions in each area were:

- ***What was the impact of Seeds on the quality of child care and on provider education and child care experience?*** To answer this question, Mathematica conducted a randomized control trial designed to examine the impacts on child care quality and the child care workforce that can be attributed to Seeds.
  - Did Seeds increase the amount of education, training, and technical assistance services accessed by participating child care businesses?
  - Did Seeds improve the quality of child care available in participating child care businesses?
  - Did Seeds improve the level of education and experience for the workforce employed in participating child care businesses?

- ***Was Seeds implemented as planned?*** To answer this question, Mathematica used qualitative and quantitative methods to explore whether Seeds implementation successfully met its goals of providing intensive, high-quality coaching and other supports for providers.
  - Did staff members receive the targeted number of coaching hours?
  - Did the implementing agencies make coach-provider matches that facilitated cooperation?
  - Did coaches develop supportive relationships with the staff members with whom they worked?
  - How well did the system of providing financial and educational supports work, and how did providers use these supports?

Evaluation Design

To answer the research questions, Mathematica designed complementary impact and implementation studies. The impact study provides a rigorous test of the Seeds model, and the implementation study describes how Thrive and its partners, Ready by Five in East Yakima and

CCR in White Center, recruited participants for the field test, trained the coaches, delivered the coaching model, and administered the financial grants.

Impact Study Design

To test the impacts of the coaching and quality improvement grants provided by Seeds on the quality of care and on Seeds scores, Mathematica conducted a randomized control trial that included random assignment of child care providers to either the treatment group or a control group. The treatment group received coaching, quality improvement grants, and professional development opportunities and supports, while the control group received only professional development opportunities and supports. Because providers were randomly assigned to the control and treatment groups (rather than based on their preferences or some other characteristics of the providers), one can attribute any changes in the treatment group to the Seeds model and not to other characteristics or factors that may influence child care quality and provider education and experience (such as motivation to participate or staff income).

In March 2009, Ready by Five and CCR recruited child care businesses to participate in the modified field test. Fifty-two family child care providers and 14 centers across the two communities agreed to participate.[1] After obtaining informed consent, Mathematica randomly assigned them into the treatment or the control group. Random assignment was conducted separately for family child care providers and centers. For the randomization, family child care providers were stratified by geographic site and language of instruction. This ensured that an equal mix of Spanish-, Vietnamese-, and Somali-speaking providers were assigned to the treatment and control groups. Child care centers were stratified by geographic site, capacity, and ages of children served. This ensured that an equal number of large and small centers were in the treatment and control groups, and that centers in treatment and control groups served a similar range of age groups.

Implementation Study Design

To learn about the implementation of Seeds, the research team collected and analyzed information from two main sources: (1) executive interviews and focus groups with site coordinators, coaches, and staff from participating child care businesses conducted during site visits to East Yakima and White Center at baseline and six months after implementation; and (2) service use data collected by coaches and site coordinators from June 1 through December 31, 2009.

Data Collection

Following random assignment, Mathematica conducted baseline data collection for the impact study, which included observations of all the provider homes and child care centers, as well as surveys of the providers, center directors, and center lead and assistant teachers. Approximately seven months after the baseline data collection period, Mathematica conducted follow-up data collection. Mathematica conducted baseline site visits to East Yakima and White Center for the implementation study in June 2009; follow-up site visits were conducted in November 2009. Mathematica also analyzed service use data collected by Ready by Five and CCR between June 1 and

[1] Those businesses that agreed to participate were a subset of those in the communities. Almost all centers in both communities and the majority of family providers agreed to participate.

December 31, 2009. In this section, we describe our data collection procedures for both the impact and implementation studies. Appendix A contains additional information about data collection.

Impact Study Methodology

The measures used in the evaluation allowed us both to describe the child care businesses participating in the evaluation and to estimate the impact of the Seeds coaching model and financial incentives on the use of education, child care training, and technical assistance services, as well as on the quality of the child care environment and provider education and experience. We used the Seeds framework to identify and organize these measures. For the impact analyses, we used the environment measure from the curriculum and learning environment domain, and the education and experience indicators specified in the professional development and experience domain. Using data collected from classroom observations, we constructed the environment Seeds scores, and using questionnaires and interviews, we constructed scores for providers' education and experience. For each child care business, we computed a Seeds score ranging from 1 to 5 for each of the three primary measures—environment, education, and experience. Higher Seeds scores reflect better quality. In the rest of this section, we describe our data sources, data collector training and certification, and response rates.

Data Sources. The Seeds impact study draws on several data sources, including classroom observations, self-administered questionnaires for center directors and educators (lead and assistant teachers), and interviews with family child care providers.[2] Observations included the ERS, the Arnett Caregiver Interaction Scale (CIS; Arnett 1989), and counts of the children and adults. We used the appropriate ERS to observe each child care setting, and the Arnett CIS for all settings. The ERS include the child care observation tools developed by Harms and colleagues. The ITERS-R consists of 39 items that assess the quality of center-based child care of infants and toddlers up to 30 months (Harms et al. 2003). The 43 items of the ECERS-R assess center-based child care quality provided to children ages 2-1/2 to 5 years old (Harms et al. 1998). The FCCERS-R consists of 38 items that assess the quality of child care provided in family child care homes (Harms et al. 2007). The 30-item Arnett CIS assesses the quality and content of the lead teacher/caregiver's interactions with children. (Box II.1 and Appendix A contain additional information about the ERS and the Arnett CIS.)

Training and Certification. Before the baseline and follow-up rounds of data collection, an expert on the ERS trained Mathematica field interviewers to conduct child care quality observations in child care centers and family child care businesses. Data collection training included classroom instruction and practice observations in child care settings. During practice observations conducted in settings that were not part of the evaluation, the trainer served as the "gold standard" against which field staff members' scores were measured. To be certified to collect study data, a data collector had to have scores within one point of the gold standard's scores on at least 80 percent of the observational items. All field interviewers met this reliability standard and were also found reliable during an in-field quality assurance check conducted by the trainer or a highly experienced interviewer. (Appendix A contains additional information about training, certification, and inter-rater reliability.)

[2] We originally intended to collect transcripts, but the response rates were not high enough to include them in the analysis.

Box II.1. Observation Measures and Their Psychometric Properties

The Environment Rating Scales (ERS) are designed for use with different age groups and types of care settings, but they share the same format and scoring system. Items are rated from 1 to 7, with descriptors provided by the authors for ratings of 1 (inadequate), 3 (minimal), 5 (good), and 7 (excellent).

- The **Infant/Toddler Environment Rating Scale-Revised (ITERS-R) (Harms et al. 2003)** consists of 39 items that assess the quality of center-based child care for infants and toddlers up to 30 months.
- The **Early Childhood Environment Rating Scale-Revised (ECERS-R) (Harms et al. 1998)** consists of 43 items that assess center-based child care quality provided to children ages 2½ to 5.
- The **Family Child Care Environment Rating Scale-Revised (FCCERS-R) (Harms et al. 2007)** consists of 38 items that assess the quality of child care provided in family child care homes.

The **Arnett Caregiver Interaction Scale (CIS) (Arnett 1989)** is a 30-item scale that assesses the quality of the caregiver/teacher's interactions with children. Observers rate the caregiver/teacher on a scale of 1 to 4 in the following areas: Sensitivity, Harshness, Detachment, Permissiveness, and Independence. A score of 1 means the behavior is "not at all" characteristic of the caregiver/teacher, 2 indicates "somewhat" characteristic, 3 "quite a bit," and 4 "very much." All the "negative" items were reverse-coded so that higher scores indicate more positive behavior. For example, a high score on the detachment subscale means providers/teachers are not detached from the children they serve.

All four total scales (ERS and Arnett CIS) demonstrated acceptable internal consistency (Cronbach's alpha) at both baseline and follow-up. Alphas at or above 0.70 are viewed by researchers as adequate and indicate that the items in a scale are measuring the same quality dimension. The ITERS-R total scores had an alpha of 0.93 for baseline and 0.94 for follow-up; the ECERS-R total scores had an alpha of 0.87 for baseline and 0.85 for follow-up; the FCCERS-R total scores had an alpha of 0.83 for baseline and 0.91 for follow-up. The Arnett CIS total scores had alphas ranging from 0.87 to 0.96. There were a few ERS subscales that demonstrated internal consistency below the 0.70 threshold. We did not drop any of these subscales for the impact study, but rather include them because of their wide use by programs and in research.

In addition to these measures, we derived two measures from the counts of children and adults conducted during observations. **Group size** refers to the average number of children observed; the **child-adult ratio** refers to the average ratio of children to adults. (Appendix A, Box A.1 includes a summary of the Washington State licensing standards for child-adult ratio and group size by type of setting, as well as Mathematica's methodology for tabulating child-adult ratio and group size.)

ERS Subscales

SUBSCALE	OBSERVATION ITEMS		
	ITERS-R	ECERS-R	FCCERS-R
Space and Furnishings	Indoor space, room arrangement, furnishings, display for children (5 items)	Indoor space, room arrangement, furnishings, display for children, space for privacy, space and equipment for gross motor play (8 items)	Indoor space used for child care, furnishings, provisions for relaxation and comfort, arrangement of child care space, display for children, space for privacy (6 items)
Personal Care Routines	Greeting and departing, meals and snacks, naps, diapering and toileting, health and safety practices (6 items)	Greeting and departing, meals and snacks, naps, diapering and toileting, health and safety practices (6 items)	Greeting and departing, meals and snacks, naps, diapering and toileting, health and safety practices (6 items)
Listening and Talking / Language-Reasoning[a]	Helping children understand and use language, use of books (3 items)	Books and pictures, encouraging children to communicate, using language to develop reasoning skills, informal use of language (4 items)	Helping children understand and use language, using books (3 items)
Activities	Fine motor; physical play; art; music and movement; blocks; dramatic play; sand and water play; nature and science; use of TV, video, and computers; promoting acceptance of diversity (10 items)	Fine motor; art; music and movement; blocks; dramatic play; sand and water play; nature and science; math and numbers; use of TV, video, and computers; promoting acceptance of diversity (10 items)	Fine motor; art; music and movement; blocks; dramatic play; sand and water play; nature and science; math and numbers; use of TV, video, and computers; promoting acceptance of diversity; active physical play (11 items)
Interaction	Supervision of play and learning, peer interaction, staff-child interaction, discipline (4 items)	Supervision of gross motor activities, general supervision of children, peer interaction, staff-child interaction, discipline (5 items)	Supervision of play and learning, peer interaction, provider-child interaction, discipline (4 items)
Program Structure	Schedule, free play, group play activities, provisions for children with disabilities (4 items)	Schedule, free play, group time, provisions for children with disabilities (4 items)	Schedule, free play, group time, provisions for children with disabilities (4 items)
Parents and Staff	Provision for parents; provision for staff personal and professional needs and growth; staff interaction, cooperation, continuity, supervision, and evaluation (7 items)	Provision for parents; provision for staff personal and professional needs and growth; staff interaction, cooperation, continuity, supervision, and evaluation (6 items)	Provision for parents; balancing personal and caregiving responsibilities, opportunities for professional growth; provision for professional needs (4 items)

Source: ITERS-R: Harms et al. 2003; ECERS-R: Harms et al. 1998; FCCERS-R: Harms et al. 2007.

[a] ITERS-R and FCCERS-R include a Listening and Talking subscale; ECERS-R includes a Language-Reasoning subscale.

Data Collection Response Rates. Baseline data collection began in mid-March and ended in early June 2009. All center classroom observations were completed, and 96 percent of participating center directors and educators completed a questionnaire. We completed baseline observations and interviews with 94 percent of family provider businesses, with three providers declining to participate (Appendix A, Table A.1 summarizes the response rates by data collection period).

Follow-up data collection spanned two months, from January to the end of February 2010. A 100 percent response rate was achieved for child care center observations and questionnaires. The same three family providers refused to participate as at baseline, and one family provider had closed her business. In addition, at follow-up, four providers no longer had children in their care, and one provider was no longer in operation due to a suspended child care license. As a result, we were not able to conduct observations of nine family child care businesses; we were, however, able to conduct interviews with four of the nine providers. The final response rate for family provider interviews was 94 percent, and the response rate for family provider observations was 83 percent.

Implementation Study Methodology

Trained Mathematica site visitors collected data during interviews and focus groups. The same team of site visitors visited each community at baseline and follow-up. In addition, service use data were collected by the site coordinators and coaches. They were trained to enter the data by the contractor that developed the data system; data files were then provided to Mathematica.

Site Visits. Mathematica conducted baseline site visits to East Yakima and White Center in June 2009 and follow-up site visits to both communities in November 2009. During the site visits, the research team conducted executive interviews with site coordinators and group interviews with the coaches (Table II.1). In addition, the research team conducted three focus groups during each visit with child care center directors, child care center lead and assistant teachers, and family child care providers and assistants. During the interviews, site visitors used structured protocols to ensure that consistent information was collected across participants and communities.

Service Use Data. Thrive contracted with Social Solutions, a provider of performance management software, to develop a web-based data system for Seeds. Social Solutions developed a data system using its Efforts to Outcomes (ETO™) software. Social Solutions was responsible for training staff on ETO and providing ongoing technical assistance. Seeds site coordinators and coaches collected service use data and entered these data into the web-based system. Data were collected for each of the following three areas: (1) the goals providers set in their QIPs, (2) the coaching hours staff at participating providers received through Seeds, and (3) the types of

Table II.1. Baseline Site Visit Interview and Focus Group Participants, by Community

	East Yakima		White Center	
Participant Type	Baseline	Follow-Up	Baseline	Follow-Up
Site coordinator	2	2	2	1
Coaches	4	4	6	4
Center directors	3	4	3	3
Center teachers (lead and assistant)	9	6	10	11
Family child care providers, plus assistants	10	8	9	4
Total Number of Participants	**28**	**24**	**30**	**23**

Source: Mathematica implementation study, 2009.

professional development supports staff received through Seeds.[3] The ETO database was designed to track staff and provider progress toward a number of goals defined jointly by staff and coaches in the QIPs. Each record the Seeds coaches entered into ETO tracked the time the coaches spent with staff; the mode of coaching delivery (including one-on-one, group, email, or telephone); and the specific goal within a QIP they worked on. In addition, Seeds site coordinators and coaches entered data on the types of professional development supports staff at participating providers (including providers assigned to the treatment and control groups) received through Seeds. Social Solutions oriented Mathematica staff to the ETO system and provided an analysis data file for the period specified. (Appendix A contains additional information about the variables included in the ETO analysis.)

Sample Characteristics at Baseline and Staff Attrition

At baseline, the child care businesses that agreed to participate in the field test were diverse, both within and across child care settings. The baseline characteristics are summarized in Tables II.2 through II.5 and discussed here. In these tables, businesses in the treatment group are combined with businesses in the control group to present characteristics for the entire study sample.

Family Child Care Providers. A total of 49 family child care providers agreed to participate in baseline data collection. Thirty-eight percent reported having an assistant (Table II.2). The providers enrolled eight children, on average, at the time of the baseline data collection. Family child care providers were 45 years old, on average (Table II.3). Nearly two-thirds of providers (57 percent) identified themselves as Hispanic, 15 percent as white non-Hispanic, and 26 percent as of an "other" race or ethnicity.[4] Nearly half of providers have less than a high school education, and only 14 percent reported having an AA, BA, or completed graduate work. On average, providers reported more than six years of experience at their current job.

Table II.2. Baseline Family Child Care Business Characteristics

Characteristic	Weighted Means or Percentages	SD
Provider employs paid assistants	38	7
Average number of assistants currently employed by provider	0.6	0.1
Number of children enrolled	8	0.7
Percentage of providers with children receiving a child care subsidy	70	7
Number of children with an Individualized Education Plan	0.2	0.1
Provider uses a specific curriculum or combination of curricula	61	8
Sample Size	**49**	

Source: Family Child Care Provider Interview completed in spring 2009.

SD = standard deviation.

[3] Seeds coordinators also used ETO to track the professional development supports received by child care providers in the control group.

[4] Other race or ethnicity includes American Indian, Chinese, Filipino, Korean, Vietnamese, Somali, and Native Hawaiian, among other ethnicities.

Table II.3. Baseline Demographic and Professional Characteristics of Family Child Care Providers

Characteristic	Weighted Means or Percentages	SD
Female	100	0
Age (Years)		
Mean	45	1
Median	46	
Race		
African American	2	2
Hispanic	57	7
White, non-Hispanic	15	5
Other	26	6
Multiple	0	0
Highest Grade or Year of School Completed		
Less than high school	46	7
High school or GED	32	7
Some college	9	4
AA or BA	12	5
Graduate or professional work or degree	2	2
Field of Study		
Child Development, Early Childhood Education, or Elementary Education	7	4
Other	15	5
Not applicable (no advanced study)	78	6
Schooling included five or more college courses in ECD	5	4
Years Worked at Current Job	6.3	0.7
Annual Salary (Dollars)		
Mean	39,630	3,767
Median	38,000	
Sample Size	**49**	

Source: Family Child Care Provider Interview completed in spring 2009.

Note: Categories do not always add up to 100 because of rounding.

ECD = early child development/child development; GED = general equivalency diploma; SD = standard deviation.

Child Care Centers. A total of 14 child care centers agreed to participate in baseline data collection. Sixty-one percent were private, nonprofit organizations (Table II.4). On average, child care centers had four classrooms, five lead teachers, and four assistant teachers. At baseline data collection, the centers enrolled 59 children, on average. Center directors were slightly older, on average, than lead teachers, and lead teachers were slightly older than assistant teachers (43, 35, and 29, respectively; Table II.5). Child care center staff most frequently identified themselves as white non-Hispanic. More than half of directors reported having at least an AA or a BA, and approximately 30 percent of lead teachers reported earning an AA, a BA, or a graduate degree. In contrast, only 17 percent of assistant teachers reported having an AA or a BA. On average, center directors reported six years of experience at their current job, lead teachers reported four years of experience at their job, and assistant teachers reported three years of experience at their job.

Table II.4. Baseline Child Care Center Characteristics

Characteristic	Weighted Means or Percentages	SD
Center Type		
Private, nonprofit	61	14
Private, for-profit	34	14
Other	5	5
Number of Classrooms	4.3	0.6
Number of Lead Teachers	4.9	0.7
Number of Assistant Teachers	4.1	0.9
Number of Children Enrolled in Center	59.2	8.9
Number of Children in the Center That Receive a Child Care Subsidy	33.0	7.6
Number of Children per Classroom with an Individualized Education Plan	0.1	0.1
Center Uses a Specific Curriculum or Combination of Curricula	77	13
Sample Size	**14**	

Source: Baseline Center Director Questionnaire completed in spring 2009.

Note: Categories do not always add up to 100 because of rounding.

SD = standard deviation.

Comparability of the Treatment and Control Groups. To ensure that the treatment and control groups were comparable across a range of characteristics and quality dimensions, in July 2009, Mathematica completed an analysis of baseline data collection undertaken from March to June 2009. The analysis concluded that, across a range of 90 variables, treatment and control groups were very similar overall (see Appendix A, Tables A.1, A.2, and A.3). However, we found statistically significant treatment-control differences in family child care providers' employment of assistants, lead teachers' earned credits, teacher enrollment in training or educational programs, centers' efforts to support teachers' child development associate (CDA) attainment, and teachers' years of experience. More teachers in the control group reported earning credits in the past six months, stated that centers support teachers' CDA attainment, and reported being enrolled in training or education programs. In contrast, more providers in the treatment group reported employing assistants. In addition, teachers in the treatment group reported more years of experience at their jobs. Because of the large number of variables tested, some of these statistical differences were likely due to chance, and thus do not reflect systematic differences across the treatment and control groups.

Staff Attrition

Approximately seven months after baseline data collection, Mathematica conducted follow-up data collection. As at baseline, we conducted observations of all the provider homes and child care centers, as well as surveys of the providers, center directors, and center teachers and assistants. At follow-up, we learned that the participating child care centers had experienced considerable staff turnover between the baseline and follow-up data collection periods (Table II.6).

Table II.5. Baseline Demographic and Professional Characteristics of Child Care Center Directors, Lead Teachers, and Assistant Teachers

	Center Director		Lead Teacher		Assistant Teacher	
Characteristic	Weighted Means or Percentages	SD	Weighted Means or Percentages	SD	Weighted Means or Percentages	SD
Female	89	8	97	2	96	3
Age (Years)						
Mean	43	4	35	2	29	2
Median	41		33		26	
Race						
African American	0	0	7	3	13	6
Hispanic	5	5	20	6	19	6
White, non-Hispanic	73	13	50	9	49	8
Other	5	5	9	6	8	5
Multiple	16	11	14	10	10	6
Highest Grade or Year of School Completed						
Less than high school	0	0	6	3	21	8
High school or GED	5	5	22	6	27	8
Some college	39	14	41	9	35	8
AA or BA	38	14	24	7	17	6
Graduate or professional work or degree	18	10	6	3	0	0
Field of Study						
Child Development, Early Childhood Education, or Elementary Education	71	13	37	9	31	7
Other	23	13	33	8	20	6
Not applicable (no advanced study)	5	5	30	7	50	9
Schooling included five or more college courses in ECD	79	10	38	9	34	8
Years Worked at Current Job	6	2	4	1	3	0
Annual Salary (Dollars)						
Mean[a]	32,237	2,428	19,208	1,258	16,824	700
Median	33,600		20,000		16,416	
Sample Size	**14**		**59**		**51**	

Source: Baseline Center Director Questionnaire and Teacher Questionnaire completed in spring 2009.

Note: Categories do not always add up to 100 because of rounding.

ECD = Early child development/child development; GED = general equivalency diploma; SD = Standard deviation.

[a] One assistant teacher with a very large salary was excluded. This assistant teacher self-identified as a center director as well as assistant teacher.

Table II.6. Staff Turnover at Child Care Centers

		Left Center Between Baseline and Follow-Up			New to Center at Follow-Up	
	Number at Baseline	Number	Percentage	Number at Follow-Up	Number	Percentage
Center Directors	14	0	0	14	0	0
Lead Teachers	59	19	32	56	13	23
Assistant Teachers	56	22	39	55	30	55

Source: Mathematica analysis of staffing lists provided by family child care providers and child care center directors in spring 2009 and winter 2010.

Note: The second column reflects the number of staff who were present at baseline data collection, but were not present at followup. The fifth column reflects the number of staff who were new to the center at followup data collection (they were not present at baseline). These differences do not reflect a one-to-one replacement of staff at baseline by staff at followup. Some staff positions may have been eliminated; as a result the centers did not have the same number of staff at baseline and followup.

All 14 center directors who participated in baseline data collection were still working at the same child care center at follow-up (Table II.6). However, two centers designated a different person as the director at follow-up. In one case, the director at follow-up had been an assistant teacher at baseline; the director at baseline became an assistant teacher at follow-up. In the other case, the center director was a lead teacher at baseline, and the director at baseline continued to serve as the center's owner. Both of these new center directors maintained their teaching positions in addition to being the center director at follow-up.

For lead teachers, we identified 19 teachers who participated in baseline data collection but were no longer working at the same child care center at follow-up (Table II.6). We identified 13 new lead teachers at follow-up who were not present at baseline. We attributed the difference in the number of lead teachers between baseline and follow-up to two main situations: (1) lead teachers may have left soon before follow-up data collection, and centers were unable to hire replacement teachers; and (2) some centers may have decided to reorganize staffing assignments instead of hiring new staff. For example, we identified at least seven cases where a person designated as an assistant teacher at baseline was designated as a lead teacher at follow-up, and four cases where a person designated as a lead teacher at baseline was designated as an assistant teacher at follow-up. Of the 56 assistant teachers present at baseline, we identified 22 who were no longer present at follow-up data collection and 30 new assistant teachers at follow-up data collection.

We did not identify any turnover among the 49 family child care providers observed at baseline. However, four family child care providers no longer had any children in their care at follow-up, and one was no longer in operation due to a suspended child care license.

Analytic Methods

The impact and implementation study analyses rigorously tested the effects of Seeds on child care quality and documented how the coaching and financial incentives were delivered.

Impact Analyses

In preparation for the impact analysis, we cleaned and recoded baseline and follow-up data. Baseline and follow-up questionnaire data were largely complete. However, some teachers, providers, and directors did not provide responses to a small portion of questions. To maximize the amount of data available for analyses, we imputed missing values on teacher questionnaires unrelated to skip patterns with the median value for other teachers at the same center, when available. Similarly, we imputed missing values on director and family child care provider questionnaires unrelated to skip patterns with the median value for other directors and providers in the treatment or control group, depending on the group assignment.[5] In addition, we recoded most missing values from questionnaires related to skip patterns to include all observations in the treatment/control comparisons, thus avoiding the potential problem of differential selection between treatment and control groups.

The impact analyses include only those child care providers and centers that completed both baseline and follow-up questionnaires. Family child care providers that failed to complete either baseline or follow-up provider interviews were excluded from the analysis.[6] However, personnel changes within centers did not result in excluded observations.

With very few exceptions, observational data for the ITERS-R, ECERS-R, FCCERS-R, and Arnett CIS were fully complete. Following established standards for cleaning and coding these instruments, we did not impute missing values for these scales, and we deleted the observation from the analysis when more than 25 percent of the items in a subscale were missing.

Sample Weights. For the impact analyses, all baseline and follow-up data at the provider and center level were pooled across the two communities and weighted to reflect differences in the key characteristics of providers accounted for by the study's design. In addition, we constructed classroom weights by dividing center weights among the number of classroom observations for each center, with the net result that each center has the same total weight within its stratum. We constructed lead teacher classroom weights by dividing center weights among the number of lead teacher classroom observations for each center. We applied different weights to tabulations and tests of significance according to each variable's level of analysis.

Impact Estimation. The impact analysis relies on a regression specification that compares directors, teachers, providers, and classrooms in the treatment group to directors, teachers, providers, and classrooms in the control group at followup, controlling for differences in the two

[5] Nearly all variables used in the baseline descriptive and impact analyses had very low proportions of missing values not related to questionnaire skip logic. Therefore, nearly all variables had less than 10 percent of their values imputed. The exceptions are teacher's salary (52 percent of values imputed), center director's age (21 percent of values imputed), center director's salary (18 percent of values imputed), provider's replacement of assistants (14 percent of values imputed), provider's use of curriculum (13 percent of values imputed), director's report on kindergarten transition planning (11 percent of values imputed), lead teacher's wages (11 percent of values missing), and director's years of experience at his/her current job (11 percent of values missing).

[6] Less than 20 percent of family child care providers failed to complete either baseline or follow-up provider interviews. Furthermore, similar proportions of treatment and control providers did not complete interviews. Under these circumstances, we have confidence that non-completion did not bias the study's findings.

groups.[7] By including the outcomes measured at baseline in the regression, we control for small differences between treatment and control groups at the time of randomization. Thus, the estimated impacts are impacts *above and beyond* baseline differences between the treatment and control groups. Because baseline analysis revealed comparable treatment and control groups, our regression specification produced an unbiased estimate of the impact of coaching and the quality improvement grants on caregivers' education, training, and quality of child care.

In Chapter III, we present regression adjusted treatment and control group means and the difference between the two groups (if the difference is statistically significant we indicate whether it is significant at .10, .05, or .01). We note differences of .10 because the impact evaluation addresses the question of whether the Seeds coaching and quality improvement grants can improve the quality of child care and documenting significant differences at the .10 level helps highlight trends in quality improvement. In fact, most of the differences between the treatment and control groups are at the two higher levels of significance. This was unexpected given the small sample size. We also present the effect size, which is a transformation of the difference between the two groups into a standard unit that allows comparisons across measures with different scale properties. The effect size is defined as the treatment-control difference divided by the standard deviation of the outcome measure among the control group. Because the small sample sizes lead to greater volatility in the impact estimates, the effect sizes will likewise be volatile and should be interpreted with similar caution. Following the procedures used by the What Works Clearinghouse (2008), effect sizes for categorical variables were computed using the Cox (1970) adjustment.

Implementation Study Analyses

To analyze the site visit data, Mathematica created site visit reports using a standard format to ensure consistency. The research team used a qualitative analysis software package, Atlas.ti (Scientific Software Development 1997), to make it easier to organize and synthesize the large amount of data collected during the site visits. This software enabled research team members to use a structured coding scheme for organizing and categorizing data that are linked to the primary research questions (Appendix A includes the coding scheme). After the site visit reports were coded, the research team used Atlas.ti to conduct searches and retrieve data on the research questions and subtopics. The team analyzed these data both within and across sites to identify common themes that emerged across sites, as well as patterns of service delivery, staffing, and other program dimensions.

To examine the types and levels of services providers enrolled in Seeds received during the period from June 1 through December 31, 2009, the research team computed descriptive statistics—such as frequencies, means, and distributions—of variable characteristics included in the ETO data. The team then computed means across all providers and for subgroups of providers (including family child care providers and child care centers) and staff (including directors/owners, lead teachers, and assistant teachers), as appropriate.

[7] Because the regression is a fixed-effects model, it also controls for center and provider characteristics that remained constant throughout the study period. This includes characteristics like geographic location, language(s) spoken, and number of children served.

III. IMPACT STUDY FINDINGS

The primary outcomes Seeds targeted for improvement were child care quality and staff qualifications. To obtain these outcomes, CCR and Rb5 provided the Seeds coaching model and quality improvement grants (discussed in Chapter IV). This chapter presents the differences across the treatment and control groups in access to, and use of, education and training services and coaching services. The chapter also presents the impacts of Seeds on a range of outcomes, including self-reported staff education and experience, observed child care quality, and Seeds scores.[1] For all analyses, we define *impact* as the difference between businesses in treatment and businesses in control at follow-up, controlling for baseline differences. This means that the Seeds coaching and quality improvement grants offered to the treatment group during the six-month field test *caused* any differences between the treatment and control groups at followup. All impacts described in the text are statistically significant, unless noted otherwise.

Key Findings

- More treatment group center teachers than controls reported enrollment in education and training services; no statistically significant differences were found for family child care providers.
- Treatment group family child care providers and center teachers received frequent coaching services.
- Seeds affected a few education and experience outcomes.
- Seeds decreased turnover among lead teachers.
- Seeds significantly improved observed quality.
- Seeds did not improve observed group size or child-adult ratio.
- Despite large impacts on observed quality, coaching and quality improvement grants did not improve Seeds scores.

Impacts on Receipt of Education and Professional Development Services[2]

Seeds did not significantly improve family child care provider enrollment in an education or training program or the receipt of training and technical assistance outside of coaching[3] (Table III.1).

[1] Given the small number of center directors participating in the study and the sample size limitations to conducting many of the analyses of program impacts on their service experiences and outcomes, findings about center directors are not included in this chapter. For activities and outcomes where center director impacts could be computed, they are presented in Appendix A. We caution placing much weight on those analyses.

[2] All impact estimates reported in this chapter were produced using a fixed effects regression model. As a sensitivity check, we also estimated all impacts using a non-fixed effects model that explicitly controlled for stratifying variables and other key center and provider variables. Nearly all estimated impacts were similar in magnitude and significance using this alternate estimation technique.

However, a larger portion of providers in the treatment group reported being enrolled in a Child Development Associate (CDA) program (29 percent versus 10 in control; Table III.1). Because this finding is not statistically significant, it should be interpreted as suggestive evidence regarding how Seeds may have affected progress toward a CDA. In addition, Seeds improved enrollment in courses and other educational programs among center teachers in the treatment group (Table III.2).[4] Lead and assistant teachers in the treatment group were much more likely than those in the control group to report being currently enrolled in a teacher-related training or education program (60 percent of lead teachers in the treatment group versus 15 percent in the control group; 52 percent of assistant teachers in the treatment group versus 15 percent in the control group). There is also some evidence that Seeds may have influenced lead and assistant teachers to enroll in a CDA program.[5] Lead, but not assistant, teachers in the treatment group were more likely than those in the control group to report attending college courses once a month or more (36 and 2 percent for lead teacher treatment versus control group members; Table III.2). More assistant teachers (but not lead teachers) in the treatment group reported participating in training or workshops once a month or more (46 percent of assistant teachers in the treatment group versus 6 percent in the control group).

Table III.1. Seeds Impacts[a] on Family Child Care Provider- Reported Enrollment in Training/Education Programs and Frequency of Coaching (Percentages)

Characteristic	Treatment	Control	Difference	Effect Size[b]
Enrolled in teacher-related training or education program	34	26	7	0.21
Enrolled in Child Development Associate (CDA) Program	29	10	18	3.00
Frequency of training and technical assistance activities				
Once a month or more	41	46	-5	-0.12
Frequency of mentor/coach visits[c]				
Once a week or more	65	0	65***	n.a.
Sample Size	**25**	**23**		

Source: Baseline Family Child Care Provider Interview completed in spring 2009, and Follow-Up Family Child Care Provider Interview completed in winter 2010.

Note: Effect size is not defined when the control group mean equals zero.

[a] Impact is defined as treatment-control differences at follow-up, after controlling for baseline differences.

[b] Effect size is defined as the treatment-control difference divided by the standard deviation of the outcome measure among the control group.

[c] Coaches and mentors could include services provided outside of those provided through Seeds. For the treatment group, we assume these are Seeds coaching/mentoring experiences. Control group members were free to obtain any services other than those provided through Seeds. These questions were designed to assess differences across groups in these types of services.

*** Significant at 0.01.

n.a. = not applicable.

(continued)

[3] Technical assistance refers to training or educational services received from anyone besides coaches during the study period. This phrase was used in teacher questionnaires to capture any services of this type staff received.

[4] Family child care providers and center teachers were asked some questions in slightly different ways and thus the findings are not presented in the same tables. We first present the findings for family child care providers and then for center teachers.

[5] More lead teachers in treatment reported being enrolled in a CDA program. However, this difference was not statistically significant. Similarly, more assistant teachers in treatment reported being enrolled in a CDA program. This difference is not reported in Table III.2 due to the low variability in assistant teachers' responses and small sample sizes. Taken together, however, these findings provide suggestive evidence that Seeds influenced teachers to take courses in pursuit of a CDA. As depicted in Table III.4, there was no impact of Seeds on whether center teachers had CDA.

Table III.2. Seeds Impacts[a] on Center Lead and Assistant Teacher- Reported Enrollment in Training/Education Programs and Frequency of Coaching (Percentages)

	Lead Teachers				Assistant Teachers			
Characteristic	Treatment	Control	Difference	Effect Size[b]	Treatment	Control	Difference	Effect Size[b]
Enrolled in Teacher-Related Training or Education Program	60	15	45***	1.29	52	15	37***	1.11
Enrolled in Child Development Associate (CDA) program[c]	13	4	8	1.73	--	--	--	n.a.
Frequency of Courses Attended at Community or Four-Year College								
Once a month or more	36	2	34***	2.06	29	9	20	0.88
Frequency of Training Workshops Attended								
Once a month or more	35	14	22	0.75	46	6	40**	1.54
Frequency of Ongoing Consultation from a Specialist, Coach, or Mentor								
Once a week or more	70	14	56***	1.61	53	9	44**	1.47
Sample Size	**30- 31**	**26- 28**			**34- 36**	**17- 19**		

Source: Baseline Teacher Questionnaire completed in spring 2009 and Follow-Up Teacher Questionnaire completed in winter 2010.

[a] Impact is defined as treatment-control differences at follow-up, after controlling for baseline differences.

[b] Effect size is defined as the treatment-control difference divided by the standard deviation of the outcome measure among the control group.

[c] Due to a limited amount of variability in teachers' responses and small sample sizes, this figure is not reported for assistant teachers.

** Significant at 0.05; *** Significant at 0.01.

In examining receipt of coaching and mentoring services, we also found several significant differences between the treatment and control groups. Family child care providers in the treatment group were much more likely than providers in the control group to report getting weekly or more frequent visits from a coach (65 percent of treatment group members reported weekly or more frequent visit from coaches, versus none in the control group; Table III.1). Lead and assistant teachers in the treatment group were also much more likely than those in the control group to report receiving weekly ongoing consultation from a specialist, coach, or mentor (70 percent of lead teachers in the treatment group versus 14 percent in the control group and 53 versus 9 percent of assistant teachers; Table III.2). It is important to note that the control group members were free to access any available coaching or professional development services except Seeds coaching. The treatment group reported much higher levels of ongoing consultation from a specialist, coach, or mentor (reflecting the Seeds model) than control group members were able to access on their own.

We found some evidence of an impact of seeds on some education outcomes but the patterns were not consistent across providers. Seeds affected whether staff earned three credits in the past six months (Tables III.3 and III.4). Differences across groups were in the positive direction for providers and lead teachers but were not statistically significant. For assistant teachers the differences were significant. There was no impact overall for center teachers on highest grade or year of school completed or achievement of a CDA. There was a statistically significant negative impact on family provider achievement of a CDA. As described in the limitations section and in Chapter V, lack of a clear pattern of impacts on education is likely a consequence of the short evaluation follow-up period.

One evaluation measure of staff experience providing child care is the number of years staff have worked at their current job (Tables III.3 and III.4).[6] For teachers (both leads and assistants), the number of years working at their current job was consistently lower for the treatment group as compared to the control group (a negative impact, but only marginally significant for assistant teachers; Table III.4). After controlling for baseline differences, lead teachers in the treatment group had around 3 years of experience at their current job, versus 4.5 years of experience among teachers in the control group.[7] For lead teachers, this significant, negative treatment-control difference in experience is likely related to higher turnover and more internal promotions in the control group versus the treatment group. We found that more assistant teachers in the control group were promoted to lead teachers within their centers (29 percent of assistant teachers in control at baseline held other positions in their centers at follow-up—usually lead teacher posts—compared to only 8 percent in the treatment group; Table III.5). Because these newly promoted lead teachers in the control group had more years of experience than the teachers they replaced, lead teachers in the control group "gained" years of experience, on average, relative to lead teachers in treatment during the study period. In essence, the higher level of experience among lead teachers in control versus lead teachers in treatment seems to be a reflection of the higher turnover and more frequent internal promotion within centers in the control group, rather than a direct effect of any Seeds services or supports related to teacher experience. We discuss teacher turnover in more depth in the next section.

Impacts on Teacher Turnover

Lead teachers in the treatment group were much less likely than lead teachers in the control group to leave their centers during the study period (19 percent of treatment group lead teachers who completed baseline questionnaires left by follow-up, versus 45 percent in the control group; Table III.5). In contrast, turnover of assistant teachers was not significantly different across the two

[6] We also asked experience questions of staff members separately by age of children served: How many years have you worked with infants; toddlers; preschool children? Comparisons across treatment and control groups yielded similar patterns to the findings about the number of years staff have worked at their current job. Later in this chapter, we present impacts on the Seeds rating education component, which sets a threshold of the number of months of experience required for each Seeds level across all staff.

[7] Teachers in the treatment group actually had more experience than teachers in the control group at baseline (see Appendix A) and follow-up. However, teachers in the control group reported more experience at follow-up and thus the treatment-control gap in experience observed at baseline decreased during the study period. As a result, regression-adjusted impact estimates show a negative impact of the Seeds field test on teachers' years of experience.

Table III.3. Seeds Impacts[a] on Education and Experience of Family Child Care Providers (Percentages Unless Otherwise Indicated)

	Family Child Care Providers			
Characteristic	Treatment	Control	Difference	Effect Size[b]
Earned at Least 3 Credits in the Past Six Months	21	14	6	0.90
Highest Grade or Year of School Completed				
Some college or higher	38	37	0	0.01
Has Child Development Associate (CDA) Credential	11	24	-13*	-0.54
Average Years Worked at Current Job	7.2	6.6	0.6	0.47
Sample Size	**25**	**23**		

Source: Baseline Family Child Care Provider Interview completed in spring 2009 and Follow-Up Family Child Care Provider Interview completed in winter 2010.

[a] Impact is defined as treatment-control differences at follow-up, after controlling for baseline differences.

[b] Effect size is defined as the treatment-control difference divided by the standard deviation of the outcome measure among the control group.

* Significant at 0.10.

Table III.4. Seeds Impacts[a] on Education and Experience of Center Lead and Assistant Teachers (Percentages Unless Otherwise Indicated)

	Lead Teachers				Assistant Teachers			
Characteristic	Treatment	Control	Difference	Effect Size[b]	Treatment	Control	Difference	Effect Size[b]
Earned at Least 3 Credits in the Past Six Months[c]	--	--	--	n.a.	41	16	25*	2.42
Highest Grade or Year of School Completed								
AA, BA, or higher	28	30	-2	-0.06	20	19	1	0.05
Has Child Development Associate (CDA) Credential	26	13	12	0.48	30	10	20	0.80
Average Years Worked at Current Job	3.1	4.5	-1.4**	-3.90	2.2	3.5	-1.3*	-1.98
Sample Size	**30- 31**	**26- 28**			**34- 36**	**17- 19**		

Source: Baseline Teacher Questionnaire completed in spring 2009, and Follow-Up Teacher Questionnaire completed in winter 2010.

[a] Impact is defined as treatment-control differences at follow-up, after controlling for baseline differences.

[b] Effect size is defined as the treatment-control difference divided by the standard deviation of the outcome measure among the control group.

[c] Due to a limited amount of variability in teachers' responses and small sample sizes, this figure is not reported for lead teachers.

* Significant at 0.10; ** Significant at 0.05.

groups (38 percent of assistant teachers in the treatment group had left their centers by follow-up, compared to 30 percent of control teachers). As mentioned above, assistant teachers in the control group were more likely to switch to a new position within their centers than assistant teachers in the treatment group. This is likely linked to the high turnover of lead teachers in control group centers: as lead teachers in the control group left their centers, it seems they were often replaced with assistant teachers. (As new teachers were hired in the treatment centers, it was expected that they would begin receiving coaching from a Seeds coach.)

Table III.5. Seeds Impacts[a] on Turnover Among Lead and Assistant Teachers (Percentages)

	Lead Teachers				Assistant Teachers			
Characteristic	Treatment	Control	Difference	Effect Size[b]	Treatment	Control	Difference	Effect Size[b]
Changed classroom in the center	5	3	2	0.41	9	20	-11	-0.56
Switched to a new position in the center	8	5	3	0.28	8	29	-21*	-0.94
Left the center by follow-up	19	45	-26*	-0.76	38	30	8	0.21
Switched positions or left the center by follow-up[c]	27	50	-23	-0.60	46	59	-13	-0.32
Sample Size	**32**	**27**			**36**	**20**		

Source: Baseline Teacher Questionnaire completed in spring 2009 and Follow-Up Teacher Questionnaire completed in winter 2010.

[a] Impact is defined as treatment-control differences at follow-up, after controlling for baseline differences.

[b] Effect size is defined as the treatment-control difference divided by the standard deviation of the outcome measure among the control group.

[c] This combines the two categories above to provide a summary of the staff members who either switched positions or left the center by baseline. These are exclusive categories.

* Significant at 0.10.

Impacts on Observed Quality

After controlling for baseline differences, child care businesses in the treatment group had significantly higher child care observed quality scores at follow-up than businesses in the control group.[8] For both family child care providers and child care centers, the ERS total score and most of the ERS subscale scores at follow-up were significantly higher for the treatment group than for the control group (Table III.6 and Figures III.1 and III.2).[9] Also at follow-up, ERS total scores for family child care providers in the treatment group were 0.88 points higher than those of control group providers.[10] ERS total scores for child care centers in the treatment group were 1.52 points higher than those of centers in the control group.[11] These impacts and their effect sizes are very

[8] Total ERS scores among child care centers in control actually decreased from baseline to follow-up by 0.6 points. The largest decreases in quality were found in the Interaction, Program Structure, and Listening and Talking subscales. This substantive decrease in child care quality could be linked to high turnover rates among teachers in control centers.

[9] Treatment-control differences in Table III.6 are not simply differences at follow-up. Rather, they are impacts, defined as the difference between centers and providers in the treatment group and centers and providers in the control group, *after controlling for baseline differences*.

[10] For child care centers, scores were developed at the classroom level, but are reported at the center level. To generate center-level ERS scores, we computed the average of all ITERS-R and ECERS-R classroom scores for each center.

[11] In the 10 centers in the study that had at least one ECERS-R observation and at least one ITERS-R observation, we found that the average ECERS-R total score (N = 26 classrooms) was about one-third of a scale point higher than the average ITERS-R total score (N = 25 classrooms). This discrepancy was present regardless of centers' treatment status (treatment versus control) or the date of the observation (baseline versus follow-up). There are two feasible reasons for this discrepancy in ECERS and ITERS scores within centers: (1) actual differences in quality: it is more difficult to provide high-quality care to children less than 36 months of age versus children from 36 to 60 months of age, or (2) differences unrelated to quality: the ITERS-R instrument produces inherently lower scores than the ECERS-R instrument. These two phenomena could also work in combination to produce comparatively lower ITERS-R scores.

Table III.6. Seeds Impacts[a] on Observed Child Care Quality and Seeds Scores (Means)

	Family Child Care Providers				Child Care Centers			
	Treatment	Control	Difference	Effect Size[b]	Treatment	Control	Difference	Effect Size[b]
Total ERS Score	3.88	3.00	0.88***	5.54	4.80	3.27	1.52***	4.55
ERS: Space and furnishings	3.50	3.09	0.41	2.02	4.60	3.83	0.78**	2.44
ERS: Personal care routines	2.76	2.12	0.65**	3.58	3.66	1.99	1.67***	5.92
ERS: Listening and talking	4.25	3.23	1.02*	4.06	5.64	3.59	2.05***	7.05
ERS: Activities	3.41	2.51	0.90***	6.62	4.18	3.19	0.99**	2.73
ERS: Interaction	4.70	4.20	0.49	1.57	5.89	3.74	2.15**	5.16
ERS: Program structure	5.03	3.27	1.76***	4.40	5.62	3.59	2.03***	4.66
ERS: Parents and providers	5.39	4.06	1.33***	5.23	5.19	3.30	1.90***	3.86
Arnett CIS Total Score	3.28	3.13	0.16	2.23	3.64	3.21	0.42*	3.71
Observed Child/Adult Ratio	2.90	2.46	0.44	1.59	4.63	5.67	-1.03	-1.59
Observed Group Size	4.59	3.56	1.03	2.20	8.16	7.30	0.86	0.85
Final Seeds Score	1.49	1.37	0.12	0.79	2.09	1.45	0.64	2.38
ERS Seeds component	2.49	1.51	0.98***	6.35	3.77	1.52	2.25***	7.39
PD Seeds component	2.00	1.89	0.11	0.52	1.73	1.98	-0.25	-1.15
Education Seeds component	2.00	1.89	0.11	0.50	1.84	1.98	-0.14	-0.48
Experience Seeds component	3.90	3.82	0.07	0.63	2.71	3.25	-0.54	-1.32
Sample Size	**23**	**20**			**7**	**7**		

Source: Baseline Child Care Observation, Family Child Care Provider Interview, and Teacher Questionnaire completed in spring 2009, and Follow-Up Child Care Observation, Family Child Care Provider Interview, and Teacher Questionnaire completed in winter 2010.

[a] Impact is defined as treatment-control differences at follow-up, after controlling for baseline differences.

[b] Effect size is defined as the treatment-control difference divided by the standard deviation of the outcome measure among the control group.

* Significant at 0.10; ** Significant at 0.05; *** Significant at 0.01.

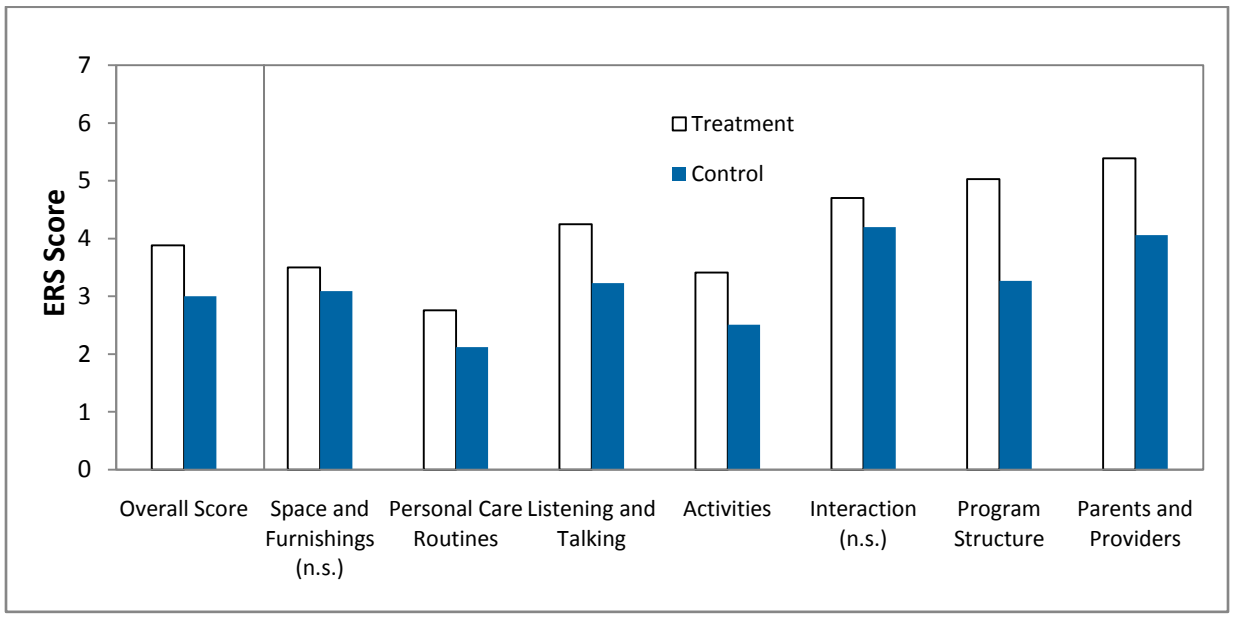
ERS = Environment Rating Scale; CIS = Caregiver Interaction Scale; PD = professional development.

large. The Seeds impacts on quality are much larger than those observed in recent evaluations of quality improvement initiatives (Landry et al. 2009; Jackson et al. 2007; Neuman et al. 2009; Powell et al. 2010).[12] For centers, but not for family child care providers, treatment group scores were significantly higher than control group scores for the CIS total score (Table III.6).

By examining the impacts of Seeds on ERS subscale scores, we have more of a sense of which types of items were more or less challenging to change with this intervention. Treatment group family child care providers had significantly higher ERS subscale scores than the control group on all but two of the subscales: Interaction, and Space and Furnishings (Table III.6). Positive impacts for

[12] For example, the Early Reading First evaluation (Jackson et al. 2007) found coaching impacts of one standard deviation on observed classroom quality, whereas we found Seeds impacts on overall quality as measured by the ERS of 4.6 to 5.6 standard deviations.

Figure III.1. Seeds Impacts[a] on Family Child Care Providers' Observed ERS Scores



Source: Baseline Child Care Observation and Family Child Care Provider Interview completed in spring 2009, and Follow-Up Child Care Observation and Family Child Care Provider Interview completed in winter 2010.

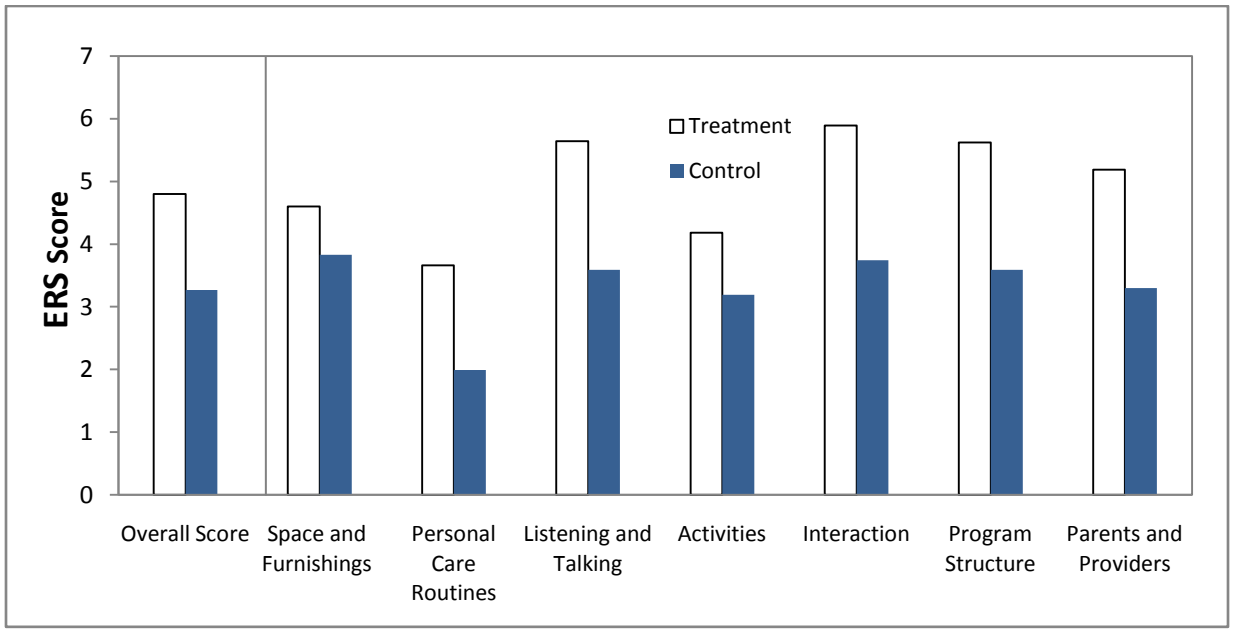
Note: All treatment and control differences were statistically significant except space and furnishings and interactions subscales.

Sample Size: 43.

[a] Impact is defined as treatment-control differences at follow-up, after controlling for baseline differences.

n.s.= not statistically significant.

Figure III.2. Seeds Impacts[a] on Child Care Centers' Observed ERS Scores



Source: Baseline Child Care Observation and Teacher Questionnaire completed in spring 2009, and Follow-Up Child Care Observation and Teacher Questionnaire completed in winter 2010.

Note: All treatment and control differences were statistically significant.

Sample Size: 14.

[a] Impact is defined as treatment-control differences at follow-up, after controlling for baseline differences.

family child care providers on the Personal Care, Listening and Talking, Activities, Program Structure, and Parents and Providers subscale scores ranged from 0.65 points for the Personal Care Routines subscale to 1.76 points for the Program Structure subscale score. The ERS subscale scores for child care centers in the treatment group were all significantly higher than those of the control group, ranging from a difference of 0.78 points for Space and Furnishings to 2.15 for Interaction.

To further explore which aspects of the ERS changed because of the Seeds coaching and quality improvement grants, we analyzed impacts of Seeds at the item level (Appendix A) by type of measure.[13] Statistically significant overall and subscale impacts seem driven by smaller sets of specific items included in each subscale. For example, the nearly one-point impact on the Activities subscale in centers is driven by large impacts on ITERS-R items that assess dramatic play, nature and science, and promotion of diversity in classrooms serving infants and toddlers; of these three items on the ECERS-R, only the impact on dramatic play was statistically significant. The item-level impacts may be useful for providing additional technical assistance for coaches as they work with child care businesses. The items we found impacts on may be those that are most readily changed and may provide guidance for allocating resources in the shorter and longer term. For example, in the first year of their work together, coaches might suggest to child care business staff members that they (1) address the scales and items we found significant impacts on, and (2) simultaneously plan for longer-term activities that may affect specific items and aspects of quality that are harder to change in a short period.

As measured by the CIS, Seeds coaching and quality improvement grants had statistically significant impacts on the quality of teacher-child interactions in child care center classrooms, but not in the family child care homes. The treatment group centers were rated 0.42 points higher on the CIS total score than control group centers (marginally statistically significant difference; Table III.6). The difference between the two groups for family child care providers was not significant. This finding is consistent with the ERS Interaction subscale results described above: Seeds did not improve the observed quality of family child care as measured by the Interactions subscale, but it did improve this aspect of quality of care in centers.[14]

Impacts on Observed Group Size and Child- Adult Ratio

Group size and child-adult ratio are important aspects of child care quality, with demonstrated associations with child outcomes. In addition, child-adult ratio is part of the Seeds rating. For these, there were no statistically significant impacts for family child care providers or for centers, although the means across groups tended to be somewhat higher for the treatment group (Table III.6).[15]

[13] As described in Chapter II, Box II.1, each version of the ERS (FCCERS-R, ITERS-R, and ECERS-R) is made up of somewhat different items. For our overall analyses at the center level, we average the scores across items. For these exploratory item-level analyses, we separately analyzed impacts for each measure at the setting level (all ITERS-R items analyzed separately from all ECERS-R items).

[14] A fundamental difference between Arnett CIS scores and the ERS scores is that CIS scores were not shared with centers and providers to inform improvements and modifications, whereas ERS scores served as a primary input to centers' and providers' quality improvement plans (QIPs).

[15] Because child care center observed quality measures are presented at the center level, group size and child-adult ratio are averaged within a center across the infant/toddler and preschool classrooms and may seem lower than expected for programs serving children through preschool.

Impacts on Seeds Scores

The overall Seeds scores themselves were not affected by the coaching and quality improvement grants. For both types of child care businesses, the Seeds scores were not significantly different across treatment and control groups (for family child care, the scores for treatment and control groups were 1.49 and 1.37; for centers, 2.09 and 1.45, respectively; Table III.6 and Figures III.3 and III.4). To further analyze why the Seeds scores were not significantly different, and to follow up on the findings of few impacts on education and experience, we broke down the Seeds overall rating into its respective ERS and professional development (education and experience) rating standards and examined impacts on these component scores.[16] Not surprisingly, across both types of child care businesses, the ERS was the primary component of the Seeds scores with statistically significant impacts (in family child care homes, the ERS Seeds component score impact was 0.98 points on a scale of 1 to 5; for centers, it was 2.25). There were no statistically significant impacts on the education and experience components of professional development (a few differences were negative but not significant). Given that the scaling of the Seeds rating is based on its three primary components, and progression to the next highest rating level requires that all components are met for each level, the very large gains in the ERS did not translate into gains in Seeds scores. It is important to note that the field test lasted six months and changes in education and experience may be challenging to achieve during this short period. This finding may warrant additional consideration as policymakers broaden Seeds implementation. In Chapter V, we return to this topic.

Limitations

The rigorous random assignment evaluation design yielded robust findings about the impact of Seeds on child care quality. The study and some of its measures have a few limitations that are important to recognize as the findings are interpreted and potentially applied to future quality improvement efforts.

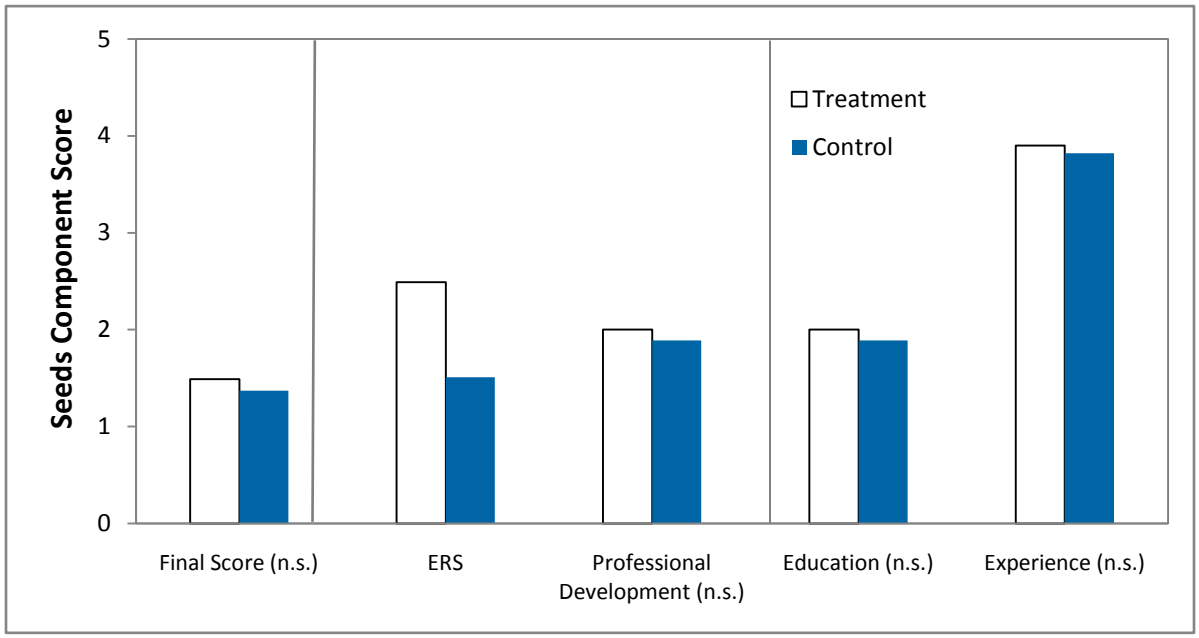
- ***Small Sample Size.*** One limitation is the size of the study sample, especially the number of centers and center directors (N = 14). The small number of center directors limited the analyses we could do with their questionnaire data. The sample size was sufficient for center-level observed quality estimation and despite the small sample size, the impacts were quite large.
- ***Staff Reporting Inconsistencies.*** From baseline to follow-up, staff members sometimes reported differently about their education and experience levels. This supports the need for independent verification of education and experience in QRIS implementation and evaluations. Appendix A presents sensitivity tests we conducted to explore the implications of these reporting issues.
- ***Use of Only Three Observed Child Care Quality Measures as Outcomes.*** The ERS, CIS, and child/adult ratio served as the observed quality outcome measures for the study. If it would have been possible to include another outcome measure, we would

[16] See Chapter I for the education and experience standards within a setting for each rating level.

have another source of evidence about the impact of Seeds on quality. When possible, multiple outcome measures of the same construct are desirable.

- ***Relatively short follow-up period.*** Given the need to provide research results to inform policy and program development in Washington State, the follow-up period for the evaluation was only six months. Seeds developers and program managers were concerned from the outset that it would be challenging to see changes in staff educational status (attainment of college degrees or credentials) during this time. We do not know whether education status may have increased for the treatment group if there was more time between the baseline and follow-up periods. A longer observation period is desirable in QRIS evaluation but must be balanced with policy needs and funding realities.

Figure III.3. Impacts[a] on Family Child Care Providers' Seed Scores



Source: Baseline Child Care Observation and Family Child Care Provider Interview completed in spring 2009, and Follow-Up Child Care Observation and Family Child Care Provider Interview completed in winter 2010.

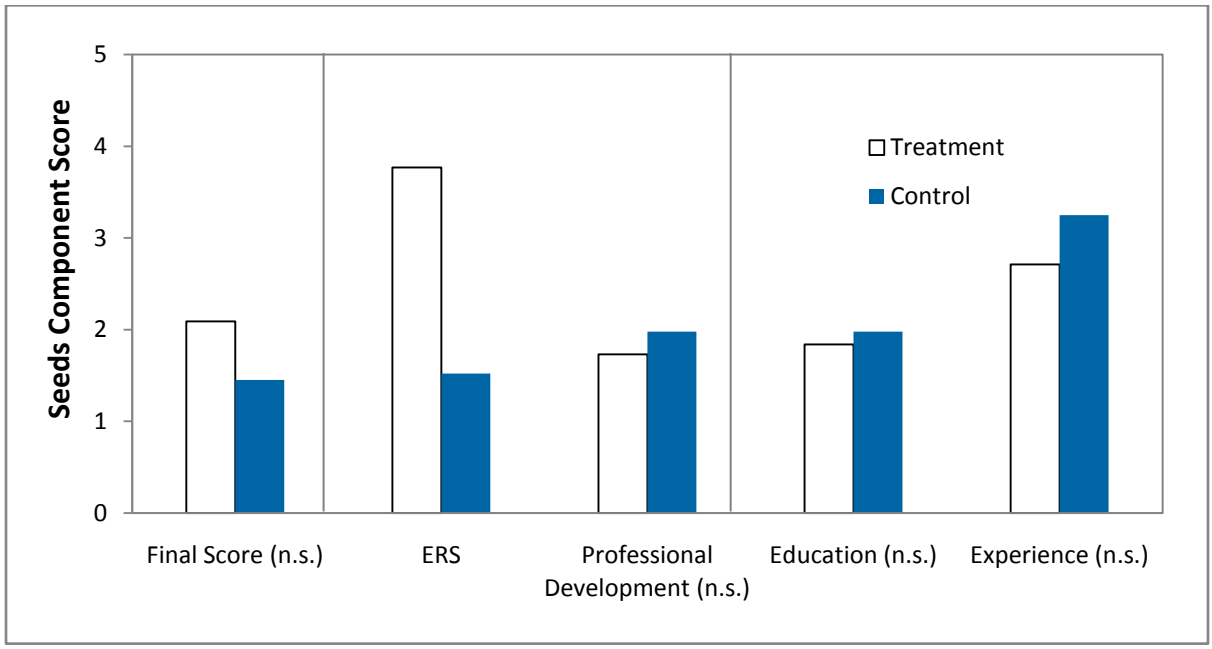
Note: Only ERS had statistically significant differences between treatment and control groups.

Sample Size: 43.

[a] Impact is defined as treatment-control differences at follow-up, after controlling for baseline differences.

n.s. = not statistically significant.

Figure III.4. Impacts[a] on Child Care Centers' Seed Scores



Source: Baseline Child Care Observation and Teacher Questionnaire completed in spring 2009, and Follow-Up Child Care Observation and Teacher Questionnaire completed in winter 2010.

Note: Only ERS had statistically significant differences between treatment and control groups.

Sample Size: 14.

[a] Impact is defined as treatment-control differences at follow-up, after controlling for baseline differences.

n.s. = not statistically significant.

IV. IMPLEMENTATION STUDY FINDINGS

In Chapter IV, we highlight key findings from the implementation study. The findings are based on two types of data: (1) executive interviews and focus groups with site coordinators, coaches, and participating providers conducted during site visits to East Yakima and White Center at baseline and six months after implementation; and (2) ETO data collected by coaches and site coordinators and analyzed by Mathematica six months after implementation. We begin this chapter by discussing the implementation approaches implemented by Rb5 and CCR, including the staffing structure, training and supervision of coaches, process for matching coaches and providers, coach caseloads, and the system for monitoring implementation. We then discuss the implementation experiences of child care providers, including their receipt of coaching, quality improvement grants, and professional development opportunities and supports. The main findings are summarized below.

Key Findings

- Providers saw their coaches as mutual partners and appreciated that coaches wanted them to drive the content of the QIPs and coaching visits.
- ERS scores served as a resource in determining providers' needs.
- Each provider developed a QIP, which served as a basis for the content of coaching.
- On average, the amount of coaching providers received adhered to the Seeds model; however, participants had a range of experiences.
- Coaching was most often provided during one-on-one meetings.
- The most common uses for quality improvement grants were supplies and materials.
- Professional development supports helped some providers access training and educational opportunities, but barriers still exist.

Implementation Approach

In this section, we describe the communities' approaches to employing coaches, assigning coaches to providers, managing caseloads, and tracking service delivery. In East Yakima, Rb5 is the lead agency responsible for the implementation of Seeds. In White Center, CCR is the lead agency. Table IV.1 provides an overview of the implementation approaches, by community, as reported during site visits.

Employing Coaches

At both Rb5 and CCR, a lead staff person provides oversight of Seeds implementation, and a site coordinator directly supervises the people hired or contracted to coach child care providers. Coaches in East Yakima are full-time employees of Rb5; coaches in White Center work as consultants for CCR.

Table IV.1. Organizational Structure, by Community

	East Yakima	White Center
Implementing Agency	Rb5	Child Care Resources (CCR)
Staffing Structure	The early learning professional development coordinator oversees Seeds; the early learning quality coordinator provides direct supervision to the coaches.	The assistant director of provider services and WCELI project lead oversees Seeds; the QRIS coordinator provides direct supervision to the coaches.
Coaches	Rb5 hired 4 coaches to work full-time with providers.	CCR hired 12 coaches as consultants to work with providers.

Source: Mathematica implementation study, 2009.

To identify coaches, both agencies advertised widely for people with experience and expertise in early childhood development, child care, culturally appropriate practice, and adult learning theories. In addition, the agencies sought people fluent in Spanish or Somali to work with the providers in their communities. Rb5 hired four coaches, and CCR contracted with 12 consultants.

Training and Supervision

In preparation for implementation, site coordinators and coaches participated in multiple training sessions, including trainings on coaching, the ERS, and using the ETO data system (Table IV.2). External consultants provided training to the coaches on delivering a relationship-based approach to coaching.[1] An external consultant provided training to the coaches on the ERS. This training was designed to give coaches the information they needed to understand ERS scores so they could explain them to providers and use the scores to guide the coaching sessions.[2] However, coaches were not trained and certified to use the ERS to conduct observations. Social Solutions provided training to the coaches and site coordinators on how to use ETO, including how to record their coaching sessions, input information on professional development, and run reports.

In addition to the three trainings offered by external consultants, Rb5, CCR, and Thrive provided an orientation to the coaches on the Seeds field test and evaluation and site coordinators trained the coaches on how to develop QIPs with providers. Coaches at Rb5 attended training on early childhood curricula and assessments, as well as conferences on early learning.

Table IV.2. Key Trainings for Seeds Coaches

Training	Description
Orientation	Overview of the Seeds to Success modified field test
Consultative Coaching	Instruction on the consultative approach; skill building on developing a strong coaching relationship
Introduction to Assessments	Training on the ERS, including understanding ERS scores
Introduction to ETO	Training on ETO, including how to document visits and complete QIPs in ETO

Source: Mathematica implementation study, 2009.

[1] See Chapter I for a description of the coaching model.

[2] Mathematica shared the scores from the ERS observations, as well as the observation booklets, with the coaches. The coaches, in turn, shared this information with the providers and used the information to inform their QIPs.

Coaches were supervised during team and one-on-one meetings. Coaches at Rb5 met weekly during the field test with the coordinator for supervision. The Rb5 coordinator also conducted on-site observations with the coaches and provided feedback and suggestions based on her observations. In White Center, the coaches met monthly with the site coordinator. In addition, coaches from both communities had the opportunity to meet one-on-one or as a group with a mentor coach hired by Thrive. The mentor coach met with coaches in both communities to provide guidance on the coaching model implemented by the coaches. Site coordinators in both communities provided ongoing support to the coaches on ETO.

During focus groups, the coaches reported that the supervision they received from the site coordinators was helpful as they applied the coaching approaches they were taught in training to their work with the providers. Coaches in East Yakima particularly appreciated the on-site supervision, explaining that the site coordinator was able to provide direct feedback on their interactions with providers. Coaches from both communities described the ongoing supervision provided by the mentor coach as particularly helpful.

Linking Coaches and Providers

The two communities took different approaches to linking coaches and providers. CCR let providers choose their coach, while Rb5 assigned the coaches to providers. CCR prescreened the coaches based on the coaches' availability and whether they spoke the same home language as the provider. The coordinator then gave each provider a list of three possible coaches, along with the coaches' resumes, and allowed providers to rank the coaches in order of preference. Ultimately, the CCR site coordinator made the match between the coaches and providers taking into account the availability of the coaches. Because all Rb5 coaches were fluent in both English and Spanish, the site coordinator assigned coaches to providers at random.

Overall, the site coordinators reported during site visits that the coach and provider assignments worked out well. The site coordinator from Rb5 reported that she made an early adjustment to the assignments to more evenly distribute the caseloads among coaches. In White Center, CCR encountered some challenges during the matching process because there were more Somali-speaking providers than coaches. To address this challenge, the site coordinator assigned an interpreter to work with one provider and two other providers were assigned non-Somali speaking coaches.

Managing Caseloads

The number of providers each coach worked with varied by community. Three of the coaches at Rb5 worked with one child care center and one or two family child care providers (the number of family child care centers depended on the size of the center they were assigned to). The fourth coach worked with eight family child care providers. At CCR, whose coaches were part-time consultants, coaches' caseloads ranged from one to four providers. The number of providers assigned to each coach depended on her availability.

During site visits, coaches from Rb5 described their caseloads as manageable but said they could always use more time to spend with the providers. Some of the CCR coaches reported that they struggled to complete their coaching visits, attend meetings, and record service-tracking information into ETO in the time they allotted to their position as Seeds coaches. Coaches reported that they often struggled to find time to work with assistants, often because the assistants' schedules did not coincide with the times they visited the centers. Coaches from both communities said that

data entry into ETO took up more of their time than expected. One coordinator suggested that in the future coaches carry laptops with them so they could complete data entry while in the field.

Tracking Service Delivery

During site visits, the site coordinators and coaches reported that they used ETO to track service delivery, monitor providers' progress toward their goals, and track their uses of professional development supports. The site coordinators developed Excel spreadsheets to track providers' uses of quality improvement grants. In addition, the site coordinators reported using ETO, plus other documentation (including coaches' calendars), to monitor the amount of time coaches were spending with providers.

The coaches who participated in the site visit interviews reported that recording how they spent their time with providers was useful because it allowed them, during the coaching sessions, to monitor providers' progress toward their goals and review the goals they were working on. However, coaches also described several challenges associated with using ETO. The most frequently mentioned challenge was the amount of time it took them to record details of each coaching session. As a result, some coaches said they recorded information weekly or monthly, rather than daily. Several coaches also expressed frustration that the ETO system does not allow for flexibility in what information can be entered about how time with providers was spent. For example, one coach reported that she had a four-hour meeting with a center director, but a couple of those four hours were spent waiting for the director to deal with issues that arose at the center. The coach said ETO does not allow for an easy way to account for the time she spent waiting. In addition, coaches described several issues with ETO that made it difficult to use, including:

- The named goals listed in providers' QIPs are not listed as options in the goals worked on during coaching sessions; rather, goals are listed by category and goal number (such as Health and Safety #1). Coaches said this design made it difficult for them to correctly account for their time.
- The variable names in ETO are used inconsistently, which can lead to confusion. In particular, goals are sometimes referred to as assessments and sometimes as goals.
- Making changes to information previously entered in ETO is challenging.

Implementation Experiences

Child care providers in the treatment group received coaching, quality improvement grants, and funding and supports for professional development. In this section, we describe the services provided as reported in ETO and described during site visits.

Coaching

Coaches aimed to provide support to family child care providers and their assistants, as well as to child care center directors, teachers, and assistants, as they worked together to improve the quality of care they provided. As described in Chapter I, forming positive relationships between the coaches and the providers was a key component of the Seeds model. The providers were expected to determine the agenda for coaching visits, and the coaches served as a support to advance the providers' goals. During site visits, coaches said they used the following techniques to build and sustain positive relationships with providers: (1) they listened to providers' goals and aspirations, (2)

they respected their knowledge and skills as caregivers, and (3) they spent time with them in the classroom and with the children.

The center directors, teachers, and family child care providers who participated in focus groups described the coaches as mutual partners and reported appreciating that the coaches wanted them to drive the content of the QIPs and the coaching visits. They described how the coaches refrained from giving their opinion unless the providers asked for their input and how the coaches repeatedly stressed that the ultimate decisions about how they attained their goals were theirs. The providers said they appreciated that the coaches listened to their concerns, described what they observed in their homes/classrooms, and made recommendations, then allowed the providers to decide if the recommended changes were the right ones.

Similarly, the coaches and site coordinators described their approach to coaching as being well received by the providers. They explained that, for the most part, this approach was working, and providers were taking the initiative to make changes. However, the coaches and coordinators expressed some tension between their approach of following providers' leads and the need to improve the quality of care. Coaches explained that some providers were reluctant to make more difficult changes. When these tensions arose, coaches described building on their relationship with the providers to encourage them to address the needed changes. At times, they said they had to be more directive with the providers about changes they needed to make to increase their Seeds rating.

Using ERS Scores

As part of the impact evaluation, Mathematica conducted ERS observations at baseline of all child care businesses participating in the Seeds field test. As discussed in Chapters II and III, Mathematica used the ERS scores to assess the process quality of the child care businesses and to calculate the Seeds ratings. In addition, Mathematica shared the scores from these observations, as well as the observation booklets, with the coaches. The coaches, in turn, shared this information with the providers; the providers and coaches used the ERS scores as one source of information to inform their QIPs.[3]

Reaction to ERS Scores. During site visits, coaches described providers' responses to the scores as mixed—some said the scores were what they expected, while others had concerns with the scores. During focus groups, providers had similar reactions. According to child care center directors, the classrooms they thought would score high did not always do so. Family child care providers and center directors reported that they had several questions for coaches about how items were scored and how overall ERS scores were calculated. Several reported purchasing the ERS measures and supporting materials so that they could better understand the standards against which they were being rated.

Concerns About ERS Scores. Some providers and teachers who participated in focus groups expressed concerns about the uniformity and fairness of the ERS observations. For example, some teachers and family child care providers reported that observers spent more time at another provider's home or in another teacher's classroom than in their own, and they were uncertain about

[3] The coaches did not share the ERS booklets with the providers; rather, they used the booklets as a resource when explaining the providers' ERS scores to them.

the reasons for these differences. Another common concern expressed by providers and teachers was related to how scores were assigned for items the observers did not see. For example, if naptime is in the afternoon and the observer was in the classroom in the morning, teachers wanted to know what information the observer used to assign scores to items related to naptime routines. During focus groups, coaches, providers, and teachers reported that, when these issues came up, coaches were not always able to explain why a provider was given a specific score. Coaches explained that it was often challenging for them to interpret the information provided by the observers in the observation booklets, and the level of detail varied by observer and, sometimes, by item. Among providers and teachers who participated in focus groups, those who reported that their coaches were not certain why they received a specific score expressed less confidence in the ERS scores they received.

Developing QIPs

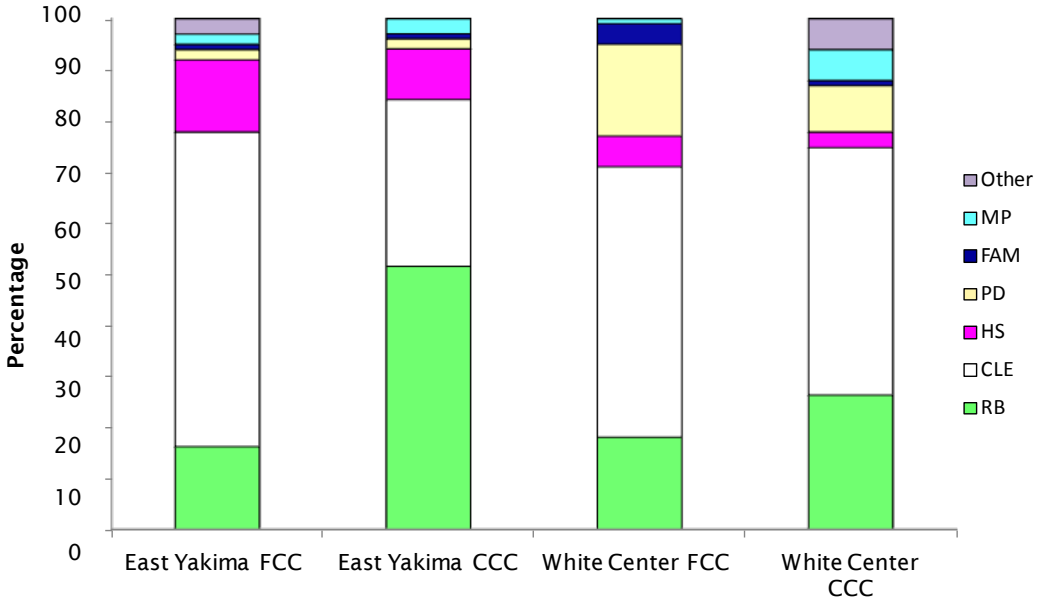
A key activity of the Seeds model was developing QIPs. To develop the QIPs, coaches and providers worked together to identify goals for areas of improvement. As providers achieved the goals in their QIPs, they worked with coaches to develop new goals. In both communities, the coaches developed individual QIPs with family child care providers, center directors, and lead teachers. During site visits, coaches and providers described the main sources of information that they used to inform the plans as comprising provider input, ERS scores, and coaches' observations. QIPs were entered into ETO, and progress toward goals was tracked over time.

During site visits, coaches and providers described the development of QIPs as an ongoing process. New goals were added to the QIPs continuously as initial goals were met or additional needs were identified. Providers explained that coaches often worked with them to break up long-term goals into smaller, more attainable short-term goals. As these short-term goals were met, they would incrementally add new goals that would help them meet their ultimate long-term goal.

Time Spent on QIP Goals. The goals identified in the QIPs drove the content of the coaching sessions. Coaches recorded in ETO the amount of time they spent working on specific goals with each participant during each coaching session. The goals were grouped into six categories: (1) curriculum and learning environment, (2) health and safety, (3) professional development, (4) family relationships and community partnerships, (5) management practices, and (6) other.[4] Table IV.3 lists the types of activities included in each goal type. In addition, coaches reported on times they met with staff but did not focus on a particular goal (we refer to this as general time and relationship building). As Figure IV.1 shows, coaching with family child care providers in both communities focused mostly on topics related to curriculum and learning environment goals. Coaching with staff at East Yakima child care centers focused largely on relationship building and curriculum and learning; in White Center, coaching focused mostly on topics related to curriculum and learning and professional development.

[4] In White Center, professional development goals included time coaches and providers spent identifying goals and professional development opportunities. The site coordinator from East Yakima explained that the coaches originally coded time establishing professional development goals as relationship building in ETO; they later learned from Social Solutions that this was incorrect and should have been coded as professional development. Coaches were not able to identify and correct all cases when this occurred and therefore the amount of time spent on relationship building may be an overestimate.

Figure IV.1. Amount of Coaching Time Spent on Each Goal



Source: ETO, June 1 through December 31, 2009.

Note: CLE = curriculum and learning environment; FAM= family relationships and community partnerships; HS = health and safety; MP = management practices; PD = professional development; RB = general time and relationship building. Number of coaching sessions: for East Yakima FCC = 274; for East Yakima CCC = 516; for White Center FCC = 356; for White Center CCC = 749.

Amount of Coaching Received

During focus groups with providers, they described the process coaches used to schedule visits. Some coaches set up a regular schedule for coaching visits. Other coaches developed schedules weekly, so the providers met with the coaches on different days and at different times, depending on the coaches' and providers' availability.

According to ETO data, on average, family child care providers, center directors, and lead teachers received between four and seven visits per month, and the average amount of coaching was between 6 and 11 hours per month (Table IV.4). Assistant teachers in White Center received, on average, less than two hours of coaching per month; in East Yakima, they received, on average, nearly six hours per month. Although, on average, providers received close to the targeted amount of coaching per month, they had a range of experiences, with some providers receiving, on average, far fewer visits and hours of coaching a month (Table IV.5).

Barriers to Completing Coaching Sessions. During site visits, coaches and coordinators from Rb5 and CCR described several obstacles that impeded their ability to meet the targeted amount of coaching. Some staff, particularly assistant teachers, turned over frequently and so the amount of time they spent in the program was minimal. In addition, there was variability in providers' needs and their openness to the coaching visits. Some family child care providers and child care centers had greater needs (such as administrative turnover or licensing violations) than others and therefore required more frequent visits. In addition, coaches said some family child care providers, center directors, teachers, and assistant teachers were reluctant to allow coaches into their homes or classrooms, explaining that they were too busy to meet with the coach; a few experienced

Table IV.3. Description of Goals Included in Each Goal Category

Description
Curriculum and Learning Environment
Improvements/enhancements to indoor learning environment
Improvements/enhancements to outdoor learning environment
Adding/enhancing a curriculum; adding/enhancing learning activities
Improvements/enhancements to teaching practices
Improvements/enhancements to classroom management and managing difficult behaviors
Adding/adjusting classroom schedules and routines; implementing center-wide schedules and routines
Adding/enhancing assessments and/or observations of children
Health and Safety
Improvements/enhancements to sanitation related to food preparation and meal times
Improvements/enhancements to sanitation related to bathroom and diapering
Improvement/enhancements to other sanitation, including hand washing (if not applicable to food preparation/meal times or bathroom/diapering)
Professional Development
Developing professional development goals
Identifying professional development opportunities
Family Relationships and Community Partnerships
Improvement/enhancement of interactions with parents; providing opportunities for parent engagement (such as parent nights or parent-teacher conferences)
Improvement/enhancement of the information shared with parents (about provider, about child, about curriculum)
Management Practices
Improvements/enhancements to classroom management and managing difficult behaviors
Improvement or development of job descriptions, personnel policies, or other policies related to staff
Improvements to, or development of, a process for holding staff accountable or evaluating staff performance
Improvements/development of systems for supporting professional development for staff (including training, education, membership in professional organizations)
Other
Any other goals that do not fit into categories listed above

Source: ETO, June 1 through December 31, 2009.

Table IV.4. Average Amount of Coaching Received by Staff per Month, by Provider Type

	Number of Staff	Average Number of Coaching Sessions per Month	SD	Average Number of Hours of Coaching Received per Month	SD
East Yakima					
Family child care providers	13	3.9	1.0	7.1	2.5
Child care centers					
Director/owner	7	4.1	1.6	9.8	5.6
Lead teachers	18	3.6	1.6	7.0	4.6
Assistant teachers	11	3.8	1.8	5.8	3.6
White Center					
Family child care providers	13	3.8	0.9	7.0	1.7
Child care centers					
Director/owner	4	7.1	2.7	11.1	2.2
Lead teachers	19	4.4	1.5	6.0	2.6
Assistant teachers	20	2.1	1.7	1.9	2.3

Source: ETO, June 1 through December 31, 2009.

Note: To calculate these data, Mathematica calculated an average for each participant and then averaged across all participants with the same staff type.

SD = standard deviation.

a prolonged illness. In White Center, coaches reported that it was often difficult to schedule visits because of limitations in their own availability and the availability of the providers with whom they worked.

Type of Coaching Sessions

Coaches had a fair amount of flexibility in deciding how to spend their time with providers. They could work one-on-one with specific staff members or offer group sessions on a particular topic to several staff. They also tracked time spent emailing and making telephone calls to providers. According to the ETO data, coaches almost always worked with family child care providers one-on-one (99 percent of the time in East Yakima and 95 percent of the time in White Center; Figure IV.2). Coaches in East Yakima spent 58 percent of their time at child care centers working one-on-one with staff and 38 percent of the time working with groups of providers. In White Center, coaches spent nearly all of their time at child care centers working one-on-one with staff (91 percent of the time). According to the ETO data, coaches spent very little time sending emails and making telephone calls to providers; however, during site visits, coaches reported that they did not always log the time they spent on these activities.

Quality Improvement Grants

Participating providers received quality improvement grants based on their baseline Seeds score (see Table I.2). The grants were designed to help providers make improvements to their child care center or family child care home learning environments, as well as to purchase items to help them improve their instructional or management practice, with an ultimate goal of helping them increase their Seeds score. The Seeds coaches worked with providers to help them identify uses for the grants. All purchases had to be approved by the Rb5 and CCR Seeds coordinators.

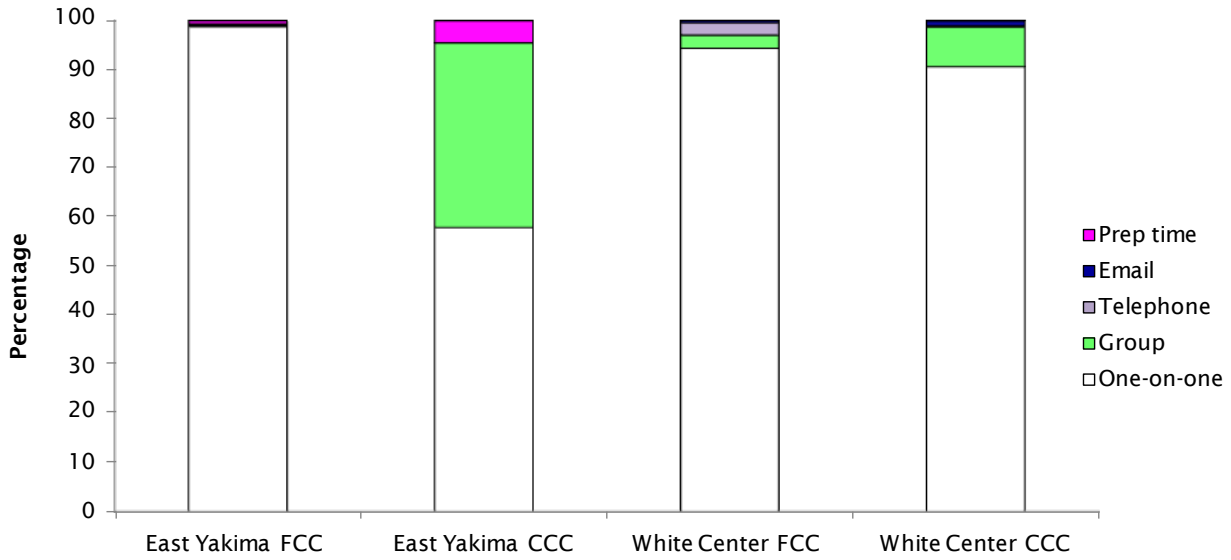
Table IV.5. Average Number of Hours of Coaching Received by Staff per Month, by Provider Type

	Number of Staff			
	East Yakima			
Average Number of Hours of Coaching per Month	Family Child Care Providers	Center Directors	Lead Teachers	Assistant Teachers
Less than 2	1	0	1	0
2 to less than 4	1	1	1	5
4 to less than 6	0	1	7	0
6 to less than 8	6	2	4	2
8 to less than 10	5	0	4	3
10 to less than 12	0	0	0	1
12 or more	0	3	1	0
Number of Staff	**13**	**7**	**18**	**11**
	White Center			
Less than 2	0	0	1	15
2 to less than 4	0	0	1	4
4 to less than 6	3	0	10	0
6 to less than 8	5	0	3	0
8 to less than 10	4	1	2	0
10 to less than 12	1	2	2	1
12 or more	0	1	0	0
Number of Staff	**13**	**4**	**19**	**20**

Source: ETO, June 1 through December 31, 2009.

Note: To calculate these data, Mathematica calculated an average for each participant and then averaged across all participants with the same staff type.

Figure IV.2. Mode of Delivery of Coaching



Source: ETO, June 1 through December 31, 2009.

Note: Number of coaching sessions: for East Yakima FCC = 274; for East Yakima CCC = 516; for White Center FCC = 356; for White Center CCC = 749.

Across all providers in East Yakima and White Center, the most common uses of the quality improvement grants were supplies and materials (Table IV.6). During focus groups, providers described using funds to purchase a wide variety of materials, including books, art materials, dramatic play materials, and manipulatives. Providers and coaches also reported requesting funds to purchase larger items such as bookcases, child-sized tables and chairs, and outdoor play equipment and to make safety improvements. Other uses included the purchase of curriculum materials and additional staff professional development.[5]

Professional Development Supports

Providers enrolled in Seeds were also eligible to receive funding for additional professional development opportunities[6] and support to help them access opportunities. The supports offered were funds for books, release time (or funding to pay for substitutes), and child care for staff while they attended classes and trainings. Coaches worked with providers to help them identify, and apply for, professional development opportunities. The coordinators in both East Yakima and White Center were responsible for working with professional development providers to ensure that opportunities were available to providers. They also reviewed and approved requests for professional development funds and recorded in ETO how funds were used.

[5] Quality improvement grants used for staff professional development included membership in professional associations and other opportunities not paid for by the professional development funds.

[6] Providers assigned to the control group in the Seeds impact study were also eligible to receive funding for professional development. In this chapter, we only detail the experiences of the treatment group members (those providers who were offered quality improvement grants and coaching in addition to professional development funds).

Table IV.6. Uses of Quality Improvement Grants and Funds Distributed, by Provider Type

Uses of Quality Improvement Grants	East Yakima (Percentage of Funds Distributed)		White Center (Percentage of Funds Distributed)	
	Family Child Care Providers	Child Care Centers	Family Child Care Providers	Child Care Centers
Supplies and Materials	87	47	93	94
Curriculum Materials	4	30	5	1
Staff Professional Development	9	23	0	4
Other	1	0	2	1
Number of Providers	13	3	13	4
Total Amount of Funds Distributed	$19,200	$7,800	$18,571	$15,452

Source: Ready by Five and WCELI program records, June 1 through December 31, 2009.

Note: Not all categories add to 100 percent because of rounding.

Types of Professional Development Opportunities. During site visits, the coordinators reported that the types of professional development opportunities available to providers included: (1) course work toward a bachelor's degree; (2) course work towards an associate's degree; and (3) noncredit professional development opportunities, including child development associate (CDA) classes, adult basic education, English as a second language, training offered through local child care resource and referral agencies and other organizations. Rb5 arranged for CDA course work conducted in Spanish and a Spanish literacy program for Spanish speakers called CONEVyT (an online program developed by the Mexican government). The program includes a training component on early childhood education. CCR offered community-based classes for providers in the Somali language.

Enrolling in Professional Development. According to Rb5 and WCELI program records, 4 of the 13 family child care providers in East Yakima and 4 of the 13 family child care providers in White Center received funding for noncredit professional development opportunities between June and December 2009 (Table IV.7). One of the 13 family child care providers in East Yakima and 6 of the 13 providers in White Center received funding for course work towards an associate's degree; no family child care providers received funding for bachelor's level course work. Among child care centers, one of the three centers in East Yakima and all of the centers in White Center received funding noncredit professional development opportunities. In East Yakima three staff members received training using this funding and in White Center 37 staff members received training. Five staff members from two child care centers in East Yakima and 17 staff members from four child care centers in White Center received funding for course work towards an associate's degree. Two center-based staff in East Yakima (in one center) and one in White Center received funding for course work towards a bachelor's degree.

Barriers to Accessing Professional Development. Providers who participated in focus groups described several barriers that deterred them from using the professional development opportunities available through Seeds. Providers reported that courses offered filled up quickly, and they had limited time to make decisions about courses and enroll. Teachers explained that the trainings and courses offered were often repetitive and did not provide them with any new information or ideas. Family child care providers reported that much of the course work was not relevant for their work as home providers. Both teachers and family child care providers expressed frustration that obtaining higher credentials did not guarantee higher pay. Providers said they were more likely to take advantage of the professional development opportunities located close to their neighborhoods and said that, given their schedules, travel even a moderate distance from work and home was a challenge.

Table IV.7. Uses of Professional Development Funds, by Provider Type

	East Yakima		White Center	
	Family Child Care Providers	Child Care Centers	Family Child Care Providers	Child Care Centers
Number of Providers Receiving Funds for Professional Development				
Noncredit	4	1	4	4
Associate's level classes	1	2	6	4
Bachelor's level classes	0	1	0	1
Number of Providers Receiving Professional Development Supports				
Books	1	2	3	3
Child care stipend	1	2	5	3
Release time	2	1	0	1
Total Number of Providers	**13**	**3**	**13**	**4**

Source: Ready by Five and WCELI program records, June 1 through December 31, 2009.

Using Professional Development Supports. In addition to funds for training and course work, staff also accessed supports. As Table IV.7 shows, family child care providers in both communities made use of funding for books and child care stipends; two family child care providers in East Yakima received funding for release time. Although child care centers also made use of these supports, the number of staff using the supports was minimal. In East Yakima, seven center-based staff from two centers received funding for books; in White Center three center-based staff from three centers received funding for books. Program records from East Yakima did not report the use of child care stipends and release at the staff level, but in White Center (which did report this information) the number of staff making use of these supports was also minimal. Eleven staff members received funding for child care stipends and 1 for release time. During site visits, center directors and family child care providers reported that the release time offered through Seeds was not sufficient. Center directors said a shortage of substitute teachers made replacing teachers difficult; instead, they often relied on existing staff for coverage. Family child care providers explained that they were reluctant to leave a substitute in charge of their homes.

Limitations

In this chapter we presented results from our analysis of the ETO data. ETO data were entered by coaches and coordinators responsible for implementing the Seeds model. Coaches and coordinators used these data to assist them with service delivery, not for evaluation purposes. Although Mathematica worked closely with Rb5, CCR, Thrive, and Social Solutions to review and check the data, we anticipate that the data set may include missing or erroneous data. Furthermore, any staff that never received coaching (either because the hours they worked did not coincide with the hours the coach was present at the center or because they were employed by the child care center for a brief period of time) are not accounted for in the ETO data included in this report. Finally, the information from our focus groups with the child care providers can only be generalized to those providers who participated in the groups. Despite these limitations, the ETO and site visit data enhance our understanding of implementation of Seeds.

V. LESSONS LEARNED

The Seeds to Success Modified Field Test impact results and the implementation experiences of Rb5 and CCR provide useful information for Washington State and for other states seeking to improve child care quality. Mathematica also learned important lessons about the process of conducting a QRIS impact evaluation that, taken together with the implementation study findings, provide lessons for QRIS developers and operators, for program managers, and for researchers.

Lessons for Policymakers

Policymakers in Washington State and around the country want to know whether investments in QRIS and other child care quality enhancements improve the quality of care children receive and ultimately improve children's outcomes. In addition, policymakers want to know more about how to design a comprehensive yet relatively low-cost QRIS. Key lessons for policymakers address the impact of the Seeds model as well as the feasibility of implementing it.

Implementing the Seeds Model Is Feasible

The Seeds evaluation found that it was feasible to implement a QRIS using an intensive-coaching model and that the incentives of the system were sufficient to motivate provider participation. All of the expected components of the Seeds model were implemented. Based on information Mathematica analyzed from focus groups, interviews, survey questionnaires, and ETO, we learned that the treatment group received Seeds scores and ratings and detailed information based on their ERS scores and were offered coaching and quality improvement grants based on their Seeds rating. In addition, they were informed that by working toward and achieving a higher score, they would receive larger quality improvement grants in the future. Finally, child care businesses in the treatment group participated in coaching at close to the intended level.

Seeds Significantly Improved the Observed Quality of Child Care

This rigorous impact evaluation demonstrated that over a six-month period Seeds coaching and quality improvement grants had very large, statistically significant impacts on the observed quality of child care in both family child care homes and child care centers. From the impact study, policymakers may conclude that as implemented in the modified field test, Seeds is a promising model for improving observed child care quality, which is a key step toward improving children's school readiness outcomes.

The Short Timeframe of the Field Test Prohibited Substantial Gains in Educational Attainment

As described in Chapter III, this evaluation did not find improvements in staff education or experience; indeed, improvements in education and experience require sustained investments longer than the six-month field test period.[1] We did find, however, that Seeds increased center-based teacher enrollment in training and educational activities.

[1] According to the Seeds model, the education standards measure whether or not staff got a degree in the observation period, as well as some interim measures of progress toward obtaining a degree. "Staff education," as used

Due to the Building Blocks Scoring System That Requires All Standards at a Given Level Be Met Before a Higher Rating Can Be Given, Large Improvements in Observed Quality Did Not Translate into Higher Seeds Scores

The Seeds rating used in the modified field test evaluation is based on its three primary components (learning environment, education, and experience) and progression to the next highest rating requires that all components are met for each level. As a result, because providers did not make comparable gains in the other components, the observed gains in curriculum and learning environment (as measured by the ERS) did not translate into gains in Seeds scores. We recommend that to maintain provider motivation, policymakers should consider interim incentives for improvements in observed quality that do not lead to higher Seeds scores. For example, a provider whose ERS score met the Seed level 3 rating but who did not meet level 3 requirements for all components could be eligible for a partial increase in the quality improvement grant she or he receives.

Formulation of QRIS Standards Has Implications for Costs and Sustainability

The original Seeds model includes four quality areas, with three to six quality standards in each area. To compute the Seeds ratings, the modified field test assessed only two areas and a total of three quality standards (learning environment as measured by the ERS, education, and experience). Creating operational definitions for each standard, collecting the required data, and computing the scores for just three standards was resource intensive.[2] In the Seeds modified field test, a streamlined model yielded large improvements in quality. Developing a model with fewer standards and clear, easily measured scoring criteria may simplify implementation and reduce costs. Unfortunately, there is not a strong research base to guide developers in determining which combinations of standards work best to improve child and family well-being and to promote staff development and retention.

Lessons for State- Level QRIS Developers and Operators

Based on our experience applying the Seeds model, we identified lessons learned about how the Seeds standards were operationalized and implemented.

Modifications to the Rating Criteria Can Simplify Scoring

As specified, the Seeds rating criteria were difficult to operationalize because of their complexity. Both the detailed draft rating system developed by DEL and its partners and the summary rating rubric required additional specification to define clearly the requirements for

(continued)
in this report, therefore, does not refer to any education, but rather education as defined in the Seeds education standards.

[2] In addition to the data collection effort described in Chapter II, Mathematica evaluation team members and QRIS implementation leaders at Thrive met weekly from November 2008 through April 2010 on the QRIS evaluation and addressed issues such as definition of the standards, selecting measures for each standard, and determining the scoring criteria. Because the Seeds scoring rubric had not been used before, we needed to make decisions around how the standards should be operationalized. Thrive staff consulted with their Rb5, CCR, and DEL partners on many of these issues throughout the evaluation period.

meeting each indicator. Modifications to the education, experience, and training standards and the approach to collecting data to rate them could make scoring easier.

- **The education rating standards require clarification to ensure that providers with higher degrees are appropriately recognized.** There were gaps in the education rating standards that made it difficult to categorize those educators with higher degrees or other achievements not specified in the descriptions of the education qualifications that satisfy each Seed level. Defining a linear hierarchy of educational achievements would be helpful for ensuring that educators' various degrees are accounted for consistently and accurately. For example, if a given criterion requires staff members to have a CDA, would that criterion be satisfied by a staff member with an AA, BA, or graduate degree in an ECE-related field? How should degrees in unrelated fields be assessed?[3] To help answer such questions, developers should establish a clear definition of an ECE-related field. It is also possible that some criteria are meant to be satisfied only by the educational achievement specified, and these should be clearly defined.

- **The education rating standards for center teachers do not account for staff with various levels of education.** For example, in order to meet the 2 Seeds requirement for lead educators, 25 percent need to have an AA *or* 25 percent need to have a CDA/15 ECE credits.[4] As currently written, a center that has 15 percent of teachers with an AA and 10 percent with a CDA/15 ECE credits would not meet the 2 Seeds requirement. We recommend that program developers review the education requirements to determine how staff with various levels of education should be categorized. If the intent of the model was for staff to have an AA, a CDA, or 15 ECE credits, this and similar standards should be made more inclusive.

- **Verifying self-reported data is important for rating the education standard, but collecting and using college transcripts for this purpose is time consuming and costly.** For the education standard, the number of college credits earned is an important element and in some cases can determine a child care business's overall Seeds score. From comparing the self-reported data against the transcripts we received, it appeared that credit information was often under- or over-reported on questionnaires which suggested the importance of verifying self-reported data. In order to verify self-reported data, we attempted to collect college transcripts from all participating center directors, teachers, and family providers during both baseline and follow-up periods, but experienced the following challenges that resulted in our inability to use the transcripts:

  - Despite close collaboration with Rb5 and CCR, the response rate was very low for family child care providers (only four transcripts were received during follow-up) and varied widely among child care centers, with some centers providing no transcripts at all and one center submitting transcripts for 100 percent of its staff.

[3] For the purposes of the Seeds impact evaluation, we gave providers and center staff members credit for meeting an indicator if their education was higher than the level indicated.

[4] For the purposes of the Seeds impact evaluation, these requirements were applied as currently stated in the education standards.

- When transcripts were received, various courses across multiple academic institutions needed to be coded in a way that was consistent and fair, but as yet unspecified for Seeds.

As the state proceeds with Seeds, it will be important to determine whether this information could be made available through STARS (the Washington State training and professional development registry) or whether other data sources are available through partnerships with local community colleges and universities. For example, Ready by Five asked providers to sign consent forms and provide their student identification numbers allowing Ready by Five to access their transcripts.

- **The experience rating area is an all-or-nothing area, and one individual's lack of experience could potentially lower the overall Seed score of a child care business.** For the experience rating area, all center lead and assistant teachers must meet the experience threshold required for each Seed level. This means that if a center recently hired a teacher new to providing care, the entire center's Seed score would be downgraded. Such rigid requirements could have unintended consequences on the hiring practices of child care businesses. Given the high rates of staff turnover described in Chapter III, it seems reasonable to revisit this rating area and perhaps model it after the education rating area (which uses percentage requirements) or instead use the average amount of experience to determine the Seed level. This will ensure that Seeds scores are not overly dependent on the experience of any one individual.

- **More guidance is needed to score the training standard that requires fulfillment of QIP-specified training.** At the start of the evaluation, we planned to rate the training standard included in the original Seeds model. The training standard considers completion of training and professional development activities that addresses needs identified in QIPs. However, in reality, the QIPs tended to be broader in terms of the goals that were included. There was also a lack of consistency in the nature of goals in QIPs, with some including short-term goals and some focusing on longer-term goals. For example, a short-term goal might be to research possible AA programs, whereas a long-term goal might be to earn an AA. If child care businesses know that the professional development and training goals in their plans will be examined for rating purposes, staff may be more likely to set short-term goals that can be accomplished easily.

- **Any time periods mentioned in the Seeds scoring rubric should be clearly defined so scores are not unfairly tied to the timing of information collection.** For the Seeds impact evaluation, data collection periods spanned up to 8 to 10 weeks. Answers to questions asking about "the last six months" or about the number of months or years of experience might be different depending on when a child care business was observed and received questionnaires. During follow-up data collection, "the last six months" was defined as the period between July 1 and December 31, 2009. Creating such definitions before each information collection period will help ensure a fair rating process that does not unintentionally penalize some providers and center staff members.

Understanding How Staff at Child Care Centers Are Organized Is Important When Developing Rating Criteria

For child care centers, the Seeds model is currently designed to rate separately center directors, lead teachers, and assistant teachers, but some centers do not make such distinctions among staff. Given that the current Seeds model has separate requirements for center directors, lead teachers,

and assistant teachers in the education and experience rating areas, clear rules need to be established for which educators should be counted in each category.[5] In terms of the overall QRIS model, it might be practical to consider how such definitions coincide with who receives coaching. On the other hand, from the perspective of children and parents, any staff member interacting with children on a regular basis should be qualified to do so regardless of whether he or she directly receives coaching services. Clearly defining how to categorize staff and who is to be included in the ratings of staff education and experience will help to ensure the transparency and accuracy of Seed ratings for child care centers.

Streamlined Approaches for Conducting Observations and Collecting Verification Data May Help Contain Costs

Given the time and training required to collect the observed quality data and information about staff, finding ways to streamline the data collection process will be important for future expansion, but any changes should not compromise the fairness or intent of the ratings. Observing classrooms, collecting questionnaires, and interviewing family child care providers participating in the Seeds to Success Modified Field Test took approximately 10 weeks to complete at baseline and 6 weeks at follow-up. This does not include the time required to train field staff to conduct observations reliably. In addition, only the environment, education, and experience areas were rated. As the Seeds model expands to additional communities, it will be important to simplify some of the rating areas and standards as described earlier. In addition, instead of conducting observations of every classroom in a child care center, perhaps a representative sample of classrooms could be chosen for observation. It may also be useful to explore options for collecting education and experience information centrally from center directors or through a statewide database that documents credentials since distributing and collecting questionnaires from every teacher was also time consuming.

Differences in the Versions of the Measures Used to Assess Quality Raise Issues for Implementing QRIS

QRIS developers and operators often assume that the different versions of the ERS are equivalent in difficulty. However, settings that serve infants and toddlers may find it harder to increase their ERS and Seeds scores. As described in Appendix A, classrooms serving infants and toddlers scored lower on the ITERS-R than classrooms serving preschool-age children scored on the ECERS-R. This pattern is common across studies and some researchers believe it is because providing high quality care for infants and toddlers may be more challenging than it is for older children (Clifford, personal communication March 2010). It is possible that the items themselves are somewhat more difficult on the ITERS-R than on the ECERS-R, but no body of research definitively confirms or disproves this assumption. What this means in the context of a QRIS is that directors, teachers, and coaches may have to work harder to improve the observed scores in infant/toddler classrooms; otherwise, the center's overall Seeds rating might be lower. Some directors might view this as unfair when they compare their center with those that do not serve

[5] For the purposes of the Seeds impact evaluation, we included all staff who work with children in the classroom at least 10 hours per week. If a center did not use lead and assistant teacher titles, we asked center directors to designate a lead teacher for each classroom based on staff education and experience (the teacher with the highest education and most experience should be designated as the lead teacher). Center director designations did not always match self-reported data.

infants and toddlers. An unintended consequence of this pattern could be a disincentive for centers to serve infants and toddlers.

Child Care Businesses That Achieve the Highest Ratings May Require Additional Incentives to Strive for Increased Quality

In the modified field test, child care businesses made very large gains in their ERS scores in just six months, and a few made gains in their Seeds scores as well. It is not clear what will be required to sustain those gains and support these businesses in continuing to improve the quality of their services for children and families. It is also important to note that to obtain the highest Seeds rating in the curriculum and learning environment area, a 5, as part of the modified field test, child care businesses had to have an average ERS score of 5 out of 7 and meet child-adult ratio requirements that are stricter than those required by licensing. A market-based approach to quality such as QRIS rests on the assumption that unless there are incentives for increasing quality, child care businesses will not strive toward obtaining it. When ratings are sustained for a year or more, state QRIS developers might want to consider building in incentives for centers sustaining the highest ratings or providing incentives to these centers for working with other programs to improve the quality of care provided in the community overall.

Lessons for Program Implementers

The implementation of Seeds by Rb5 and CCR can provide useful information for future efforts to refine or expand the Seeds model. In this section, we discuss lessons learned based on the successes and challenges faced by Rb5, CCR, and the participating providers.

When Providing Intensive Coaching, Full-Time Staff with Increased Availability and Flexibility May Reduce Obstacles to Completing Visits

During site visits, coaches and providers described a number of challenges associated with completing coaching visits. Coaches said that holidays, vacations, and other commitments, such as attending trainings and conferences, limited their ability to complete all of the coaching visits they planned for a given month. Providers reported that some days they were too busy to meet with their coaches because of specific activities going on in their classrooms or centers; others described cancellations due to brief illnesses (by either the coach or the provider) and in some cases due to periods of prolonged illness. In White Center, providers and coaches described some difficulties arranging their schedules. For example, providers reported preferring a specific day of the week for the coaching visits but explained that their coaches were not available those days. Providers requested a set schedule for coaching visits so they could plan their week around them. In addition, they said some flexibility within the parameters of the visit is also helpful. For example, a coach may visit a classroom for a period of time in the morning and then schedule to meet with the teacher at naptime to discuss what she observed. With family child care providers, the coach may need to debrief with the provider while children are sleeping or follow-up in the evening by telephone.

Balancing the time dedicated to coaching visits with the time dedicated to administrative tasks can be challenging, especially for part-time staff with other commitments. During site visits, coaches from Rb5 described their caseloads as manageable but said they could always use more time to spend with the providers. Some of the CCR coaches reported that they struggled to complete their coaching visits, attend meetings, and record service tracking information in the time they allotted to their position as a Seeds coach. In addition, coaches from both communities reported spending more time entering data into ETO than originally expected. Given the challenges associated with

using consultants whose time is less flexible than full-time staff, we recommend that programs consider hiring coaches as full-time staff in the future.

Trusting Relationships Facilitate Quality Improvement Efforts

During site visits, teachers and family child care providers described their initial concerns about having someone in their classroom or home commenting on their practice. However, as the field test progressed and they built relationships with their coaches, teachers and family child care providers said they welcomed the coaches' suggestions. In particular, providers said they liked that the coaches did not just come into their classrooms or homes and try to tell them what to do. Rather, the coaches observed the circumstances of each provider and talked through possible ideas for improving practice with the providers. Coaches also stressed the importance of relationship building and of respecting providers as professionals and experts. In addition, coaches said that having the same culture and language as the providers was helpful; when that alignment was not the case, taking the time to learn about the providers' cultural backgrounds was important.

Despite Supports, Barriers Exist That Deter Providers from Participating in Professional Development Opportunities

Providers who participated in focus groups described a number of barriers that deterred them from making use of the professional development opportunities (PD) available through Seeds. Teachers explained that the trainings and coursework were often repetitive and did not provide them with any new information or ideas. Family child care providers reported that much of the coursework was not relevant for their work as home-based providers. Both teachers and family child care providers expressed frustration that obtaining higher credentials did not guarantee higher pay. Providers said they were more likely to take advantage of the PD opportunities located close to their neighborhoods and said traveling even a moderate distance from work and home was a challenge, given their schedules. Center directors and family child care providers reported that the release time offered through Seeds was not sufficient. Center directors said a shortage of substitute teachers made replacing teachers difficult, and instead they often relied on existing staff for coverage. Family child care providers explained that they were reluctant to leave a substitute in charge of their homes.

Child Care Businesses Need Clear Guidelines About Accepted Uses of Quality Improvement Grants

Quality improvement grants have the potential to encourage provider participation and support quality improvement efforts, however some child care directors and family child care providers whose requests for items were not approved expressed frustration about the process. They described having set ideas for how the funds would be used to meet their needs and experiencing disappointment when they were told they could not use the funds in that way. In contrast, coaches and some providers explained that Rb5 and CCR might consider delaying the distribution of the quality improvement grants because the purchases the providers prioritized early in the field test might not align with the purchases they might identify after a few months of coaching. In either case, participating child care providers requested that Rb5 and CCR set clear policies about the uses of quality improvement grants and communicate this guidance to them from the start.

Coaches Could Benefit from Additional Training on ERS; Providers Might also Benefit from Training

Coaches faced challenges interpreting observers' ERS ratings and requested additional training to help them explain the scores to providers and teachers. In future replications of the Seeds model, coaches might benefit from additional training on ERS to ensure that their understanding of the scales is thorough enough to enable them to explain the ratings to providers and teachers. Coaches also need to know how observers were trained and the procedures they were instructed to follow so that they can better explain the process to providers and address misperceptions about unfairness or lack of uniformity in the rating process. Providers, many of whom had bought the ERS materials and books to help them better understand their scores, might also benefit from training that provides an overview of ERS.

Program Operators and Implementers Should Consider Taking Steps to Make the Rating Process More Transparent

To address misperceptions about unfairness or lack of uniformity in the rating process, program implementers in future replications of the Seeds model should convey the established standards for conducting quality observations to the child care businesses so everyone knows what will be assessed and how the assessments will be done. For example, ERS observers trained for the modified field test were instructed to use the scoring booklet to take personal notes about what they saw that would help them derive the scores for each item. It might be important for such notes to be more standardized if they will be shared with coaches and child care businesses. One drawback of such uniformity is that it may seem more like a licensing inspection, but the advantage is that it will appear more uniform to providers. Program implementers should also consider making information about the ERS requirements easily accessible to providers who are considering enrolling in Seeds so that they can understand the standards against which they will be rated. This will increase perceptions of fairness on the part of the providers and help them know what to expect.

ETO's Design Could Be Enhanced to Increase Its Performance Management Capabilities

ETO is an innovative approach to collecting QRIS service data (coaching hours, quality improvement grants, and professional development provided) and although other states do use management information systems to track these data, ETO and software like it have features that allow more flexibility in the design and use of service data. ETO did not adequately link individual center-based teachers and assistants to the classrooms in which they worked. This information was not included because Thrive and Social Solutions anticipated the movement of teachers between classrooms during the pilot. As a result, we could analyze the hours of coaching received by individuals but not by classrooms (which is how the coaching model is defined). To account for the challenges of linking teachers to classrooms, ETO could require coaches to record coaching hours by staff member and classroom. Alternatively, the Seeds model could define coaching requirements by individual rather than by classroom. Furthermore, the unique identification number assigned to child care providers in ETO did not align with the unique identifiers assigned by Mathematica for data collection and analysis. In the future, linking the coaching services and quality improvement grant receipt to the ERS and Seeds scores and other rating components will help program administrators explore how different providers use the services available through Seeds and the implications for their Seeds score. This alignment of information may facilitate further program improvement. For example, program administrators might be able to explore the ratings of classrooms that received varying levels of coaching or they might explore the characteristics of staff who are more and less likely to access professional development. Alignment of different purposes

for the use of service data is important so that all of the potential applications of a QRIS information management system are incorporated in the design phase: (1) tracking services at the individual staff and setting level for accountability purposes, (2) using the data for program improvement, and (3) using the data for research. A design approach that brings all of the potential users of these types of data together from the start will ensure that the data can be used for all of thes purposes.

Lessons for Researchers

Despite the proliferation of QRIS across the country, relatively few resources are available to guide funders, state-level program developers and operators, and researchers in how to conduct rigorous, comprehensive QRIS evaluations. We offer a few lessons that may contribute to QRIS evaluation design and conclude with a recommendation for future research on Seeds.

Coordination Between the Evaluation Team and Program Developers and Managers Is Critical to Success

During the past two years, Mathematica worked closely with the Gates Foundation, Thrive, Rb5, and CCR to plan and implement the evaluation. As we explored the possibility of using a random assignment design, we discussed the level of evidence the funders and program designers wanted to bring to the policy discussion in Washington State. By including both the impact and implementation studies as part of the evaluation, we were able to investigate whether the Seeds model could change child care quality and also how it might do so. Launching QRIS and the evaluation at the same time was challenging because the timing of the start of coaching services was tied to the collection of the baseline data. By taking the time to discuss the evaluation plans with Rb5 and CCR and by frankly sharing concerns and solving problems together, we were able to engage the child care businesses and obtain high response rates during both waves of data collection.

Clearly Mapping Each Child Care Setting Facilitates Tracking Staff and Linking Them to Classrooms

The biggest challenge in QRIS research is that the level of analysis is at the family provider/center level, but careful documentation of the role of each staff member in a family child care home and in centers is crucial for understanding classroom configuration and for tracking individuals and their roles within a center over time. In just six months, we found very high center staff turnover rates. Most impact evaluation designs require tracking the same individual over time. In the QRIS context, it is the function an individual has in the setting that determines the standards he or she is required to meet.

Developing easily administered tracking tools for staff performing different functions and keeping the information updated is important. For example, we know which staff members present at baseline were still there at follow-up, and we know which staff members were new at follow-up. That knowledge enabled us to compute a staff turnover indicator at the individual staff member level that could be aggregated up to the center level. However, we do not know how many different staff members were in those positions between baseline and follow-up. Our turnover indicator therefore is a lower bound; in fact, turnover is probably even higher than what we observed. It is important to understand more about how the staff and children move through a setting and how a study could document exactly how much exposure children got to a lead teacher who received the intended amount of coaching.

The Model Is Worthy of Replication and Continued Study

The evaluation demonstrated that Seeds coaching and quality improvement grants can improve child care quality and that child care businesses will participate. However, the study was only conducted in two locations with a small number of child care businesses. To build the evidence base for this model, broader replication with rigorous evaluation is important. The impacts on quality are promising, but many questions remain, including the following:

- Can the impacts be replicated using the same model and at the same level of service delivery and uptake by child care businesses?
- Would either coaching or the quality improvement grants cause impacts if offered alone?
- Can fewer hours of coaching produce the same impacts?
- Can the impacts of Seeds on observed quality be sustained over time?
- Does the Seeds model improve child outcomes?
- What are the costs of implementing the model? Is it cost-effective?
- What are the costs of going to scale and implementing Seeds and its quality improvement supports?

REFERENCES

Arnett, J. "Caregivers in Day-Care Centers: Does Training Matter?" Journal of Applied Developmental Psychology, vol. 10, 1989, pp. 541–552.

Bloom, G. S., C. L. Castagna, E. Moir, & B. Warren. Blended coaching: Skills and strategies to support principal development. Thousand Oaks, CA: Corwin Press, 2005. Retrieved from http://www.Corwin.Com/Booksproddesc.Nav?Prodid=Book225999#Tabview=Title on June 14, 2010.

Coaches Training Institute. Coach Training. Retrieved from http://www.thecoaches.com/coach-training/ on June 14, 2010.

Cohen, Jacob. "A Coefficient of Agreement for Nominal Scales." Educational and Psychological Measurement vol. 20, no. 1, 1960, pp. 37–46.

Cox, D.R. Analysis of Binary Data. New York: Chapman & Hall/CRC, 1970.

Harms, T., D. Cryer, and R.M. Clifford. Early Childhood Environment Rating Scale - Revised edition (ECERS-R). New York: Columbia University, Teachers College Press, 1998.

Harms, T., D. Cryer, and R.M. Clifford. Family Child Care Environment Rating Scale - Revised edition (FCCRS-R). New York: Columbia University, Teachers College Press, 2007.

Harms, T., D. Cryer, and R.M. Clifford. Infant/Toddler Environment Rating Scale - Revised edition (ITERS-R). New York: Columbia University, Teachers College Press, 2002.

Jackson, R., A. McCoy, C. Pistorino, A. Wilkinson, J. Burghardt, M. Clark, C. Ross, P. Schochet, P. Swank, and S.R. Schmidt. "National Evaluation of Early Reading First." Final Report to Congress. U.S. Department of Education, Institute of Education Sciences, Washington, DC: U.S. Government Printing Office, May 2007.

Landry, Susan H., Jason L. Anthony, Paul R. Swank, and Pauline Monseque-Bailey. "Effectiveness of Comprehensive Professional Development for Teachers of At-Risk Preschoolers." Journal of Educational Psychology, vol. 101, 2009, pp. 448-465.

Neuman, Susan B. and Linda Cunningham. "The Impact of Professional Development and Coaching on Early Language and Literacy Instructional Practices." American Educational Research Journal, vol. 46, 2008, pp. 532-566.

Powell, Douglas R., Karen E. Diamond, and Margaret R. Burchinal. "Effects of an Early Literacy Professional Development Intervention on Head Start Teachers and Children." Journal of Educational Psychology, vol. 102, 2010, pp. 289-312.

Schein, E. Process Consultation: Its Role in Organization Development. Reading, MA: Addison-Wesley Publishing Company, 1969.

Scientific Software Development. "Atlas.ti: Visual Qualitative Data Analysis, Management, and Model Building in Education Research and Business." Berlin, Germany: Scientific Software Development, 1997.

Washington Administrative Code (WAC). Retrieved from http://apps.leg.wa.gov/wac/ on April 27, 2010.

What Works Clearinghouse. "Procedures and Standards Handbook, Version 2.0, 2008." Retrieved from http://ies.ed.gov/ncee/wwc/pdf/wwc_procedures_v2_standards_handbook.pdf on April 29, 2010.

APPENDIX A

METHODOLOGICAL AND TECHNICAL DETAILS

This appendix provides additional technical details about our methodology for collecting and analyzing data for the impact and implementation studies.

Impact Study Design, Data Collection, and Analysis

Design and Sampling

To test the impacts of the coaching and financial incentives provided by Seeds on the quality of care provided and on Seeds scores, Mathematica conducted an experimental study that included random assignment to either the treatment group or a control group. In March 2009, Rb5 and CCR recruited child care businesses to participate in the modified field test.[1] Fifty-two family child care providers and 14 centers across the two communities agreed to participate; after obtaining informed consent, Mathematica randomly assigned them into the treatment or the control group. Random assignment was conducted simultaneously for all participants, although separately for family child care providers and centers. For the randomization, family child care providers were stratified by geographical site and language of instruction. This ensured that equal mixes of East Yakima and White Center child care centers—as well as an equal number of Spanish-speaking family child care providers—were assigned to treatment and control groups. Analogously, child care centers were stratified by the number of children they served. This ensured that equal numbers of large and small centers were in the treatment and control groups.

Following random assignment, Mathematica began the baseline data collection, which included observations of all of the provider homes and child care centers as well as surveys of the providers, center directors, and center teachers and assistants. At the same time, Rb5 and CCR notified treatment group members that they would be offered coaching services and financial grants in addition to the professional development opportunities and funds that were also offered to the control group businesses. Mathematica informed the control group businesses that they were in the control group and would not receive the coaching and financial grants.

To ensure that the treatment and control groups were comparable across a range of characteristics and quality dimensions, in July 2009 Mathematica completed an analysis of baseline data collected from March to June 2009. The analysis concluded that across a range of 90 variables, treatment and control groups were very similar overall (tables A.1, A.2, and A.3). However, we found statistically significant treatment-control differences in family child care providers' employment of assistants, center-based lead teachers' earned credits, center-based teacher enrollment in training or educational programs, and centers' efforts to support teachers' child development associate (CDA) attainment. Given the large number of variables tested, some of these statistical differences were likely due to chance and thus do not reflect systematic differences across the treatment and control groups. Given the small number of statistically significant differences

[1] Participating providers initially volunteered to enroll in the DEL QRIS field test. However, due to state budget constraints, the field test was suspended. Thrive then decided to implement a modified version of the state model and invited providers who had volunteered for the state field test to participate. Given the changes that occurred, providers experienced a delay of several months from when they originally volunteered to participate to when the initiative actually began. The providers in the field test are those most motivated to enroll in the field test. They are not a census or a representative sample of all of the providers in the two communities.

between the two groups, Mathematica concluded that randomization produced treatment and control groups that were comparable at baseline.

Data Sources

The Seeds impact evaluation draws from several data sources, including classroom observations, self-administered questionnaires, and interviews. All participating child care businesses were observed by trained Mathematica field staff. Observations focused on assessing the quality of the child care environment and included the Environment Rating Scales (ERS), the Arnett Caregiver Interaction Scale (CIS) (Arnett 1989), and counts of children and adults (conducted twice during each observation).

The ERS are designed for use with different age groups and types of care settings but share the same format and scoring system. The ERS include the set of child care observation tools developed by Harms and colleagues. The Infant/Toddler Environment Rating Scale-Revised (ITERS-R) (Harms et al. 2003) consists of 39 items that assess the quality of center-based child care of infants and toddlers up to 30 months. The 43 items of the Early Childhood Environment Rating Scale-Revised (ECERS-R) (Harms et al. 1998) assess center-based child care quality provided to children ages 2-1/2 to 5 years old. The Family Child Care Environment Rating Scale-Revised (FCCERS-R) (Harms et al. 2007) consists of 38 items that assess the quality of child care provided in family child care homes. ERS items are rated from 1 to 7, with higher scores reflecting better quality. The 30-item CIS assesses the quality of the lead caregiver/teacher's interactions with children. It can be used in both center- and home-based settings and measures the emotional tone, discipline style, and responsiveness of the caregiver/teacher. The CIS items are rated on a scale of 1 to 4, with higher scores reflecting greater caregiver sensitivity and responsiveness.

For child care centers, center directors and all classroom teachers (lead and assistant teachers) were asked to complete a self-administered questionnaire. The questionnaires asked about enrollment and staffing at the center, classroom activities, curricula and assessments, parent involvement, educational initiatives and training for staff, and employment and educational background. The questionnaires took approximately 20 minutes to complete, and the teacher questionnaire was available in both English and Spanish. For family child care providers, Mathematica field staff conducted one-on-one interviews (in either English or Spanish as preferred by the provider[2]) that covered similar topics. The interviews took approximately 30 minutes to complete. The questionnaires and interview were designed to be as similar as possible. Items were worded differently as needed to ensure that the questions were clear and appropriate to each type of child care business.

Field Staff Training and Certification

For the Seeds impact evaluation, Mathematica field staff were trained to conduct the observations and collect the staff data. Three of eight field staff members had previous experience using the ERS measures. Training and inter-rater reliability testing for the observational measures were conducted in March and December 2009, prior to the baseline and follow-up rounds of data

[2] Eleven providers spoke either Vietnamese or Somali. In these instances, an interpreter was used for the interview and was also present during the observation. Interpreters attended a brief training to ensure all questions were asked accurately and consistently.

collection. Training was conducted by staff members from the University of Washington (UW) and from Mathematica, and the same Mathematica field staff participated in both rounds of data collection.

In March 2009, two separate trainings were held for eight field staff (four based in White Center and four based in East Yakima). The first training focused on collecting data in child care centers and included two days of classroom instruction (one day for ITERS-R and one day for ECERS-R) and two days of practice observations in a child care setting (one day for each measure). The second training focused on collecting data in family provider homes. The four-day family child care training consisted of two days of classroom instruction and two practice FCCERS-R observations in a family provider setting. Staff members lacking previous experience with the ERS measures conducted an extra practice observation during training. In December 2009, field staff attended refresher follow-up training. Refresher training for child care centers included one day of classroom instruction and two days of practice observations (one ITERS-R and one ECERS-R). Refresher training for FCCERS-R included one day of classroom instruction and one day of practice observation.

During the practice observations, the trainers' scores served as the "gold standard" against which the field staff members' scores were measured. To be certified to collect study data for baseline and follow-up, field staff were required to have ERS and Arnett scores within one point of the gold standard rater's scores on at least 80 percent of the observational items (the threshold required by the ERS developers). Similar reliability checks were conducted during both field periods, with four checks in total throughout the study (baseline training, baseline field period, follow-up training, and follow-up field period). All field staff except one passed the ERS certification test during the baseline training. The observer who did not pass the certification test did not participate in the baseline or follow-up data collection.

We assessed inter-rater reliability at baseline and follow-up training to ensure that data would be collected reliably. As shown in Table A.4, at baseline, the average percentage agreement within one point of the gold standard ranged between 89 and 94 percent for the ERS and between 95 and 98 percent for the Arnett CIS. At follow-up, the average percentage agreement ranged between 90 and 95 percent for the ERS and between 96 and 98 percent for the Arnett CIS.

Cohen's kappa coefficient (Cohen 1960) is another statistical test of inter-rater agreement, with coefficients ranging from zero to one. Its computation takes into account random variation in scores across two raters. A value of one means that raters are in total agreement, whereas a value of zero means that any agreement between raters is the result of random chance. For half of the summary scores, kappas averaged across observers during baseline and follow-up training met or exceeded the threshold of 0.60 used by researchers (Table A.4). Given that the inter-rater reliability criteria of the observation measure authors were met, we used these as the outcome measures for the study.

In addition, all data collectors were found reliable during in-field quality assurance checks, with the average percentage agreement with the gold standard ranging between 88 and 98 percent at baseline and 89 and 99 percent at follow-up across all measures (not shown). We can conclude that field staff were adequately trained and demonstrated a high degree of reliability throughout the Seeds impact evaluation data collection.

Data Collection

As described earlier, 52 family child care providers and 14 child care centers consented to participate in the random assignment study. Baseline data collection began after random assignment

and spanned 10 weeks, beginning on March 19 and ending in early June 2009. Table A.5 presents the study response rates by measure and by data collection wave. At baseline, observations and interviews were completed with 49 of the 52 family providers for a final response rate of 94 percent; 3 providers declined study participation. All 58 classroom observations in the 14 participating child care centers were completed, resulting in a 100 percent response rate. Of the 129 center directors, lead teachers, and assistant teachers working in the 14 centers, we received 124 completed questionnaires for an overall response rate of 96 percent. All 14 center directors and all 59 lead teachers completed questionnaires; we were unable to obtain completed questionnaires from 5 of the 56 assistant teachers (for an assistant teacher response rate of 91 percent).

Follow-up data collection began on January 4 and ended on February 26, 2010. Interviews were completed with 48 of the 52 family providers for a final response rate of 92 percent. The same three providers refused to participate as in the baseline data collection, and one provider had closed her business shortly after baseline and left the study. Of the 48 family providers who completed an interview, we were unable to complete observations in five homes. Four family providers did not have any children in their care during the follow-up data collection period (but were still committed to providing care), and one provider had a suspended child care license. The final response rate for family provider observations was thus 83 percent. Classroom observations were again completed in all 57 classrooms[3] operating in the 14 participating child care centers, for a final response rate of 100 percent. We received completed questionnaires from all 125 eligible center directors, lead teachers, and assistant teachers for an overall response rate of 100 percent.

Psychometric Analyses of Constructed Variables

Table A.6 presents key psychometric properties for study variables constructed from the child care observation measures (ITERS-R, ECERS-R, FCCERS-R, and Arnett CIS). We include the full sample mean, standard deviation, range, and internal consistency reliability (Cronbach's alpha) for each scale overall and ERS subscales. The possible range for the three ERS measures is 1 to 7; the possible range for the Arnett CIS is 1 to 4.

Observation data collected during baseline and follow-up demonstrated adequate internal consistency for the four total scales (Cronbach's alpha of 0.83–0.96). Forty-four of the 52 alphas assessed met the reliability standard of 0.70 used by researchers. We did not drop any of these subscales, but rather rely on the long record of these measures and their demonstrated internal consistency in much larger samples to serve as evidence of their reliability.

From the four adult and child counts conducted during the observation, we also derived an average number of observed children and adults. Group size refers to the average number of children observed, and the child-adult ratio refers to the average ratio of children to adults. (Box A.1 includes a summary of the Washington State licensing standards for child-adult ratio and group size by type of setting as well as Mathematica's methodology for tabulating child-adult ratio and group size.)

[3] The number of child care center classrooms and teachers differed slightly between baseline and follow-up due to staff turnover and fluctuations in child enrollment rates.

Box A.1. Washington State Licensing Standards for Child- Adult Ratio and Group Size

The Washington Administrative Code (WAC) as of March 3, 2010 requires that licensed family child care homes and centers meet or exceed minimum thresholds for child-adult ratios and total group size in the home or classroom. The requirements in both settings vary according to the age of the children in care.

Family Child Care Homes. Family child care home ratios and group size requirements are determined by the provider's experience and education/training and by whether there is another adult providing care. There are limits on the number of children younger than 2 years old and the total number of children younger than 12 years old allowed on the premises (including the provider's own children). For example, a licensed provider serving children under age 2 can have a total of six children on the premises but not more than two children under the age of 2. If there are no children under age 2 in care and the provider has one year of experience, the maximum group size increases to eight.

Center- Based Care. In centers, classrooms serving infants (younger than 11 months old) must maintain a child-adult ratio of 4:1 and stay within a maximum group size of 8. Classrooms serving toddlers (12 to 29 months old) must have a ratio of 7:1 and a maximum group size of 14. Classrooms serving preschool children (30 months to 5 years old) must have a ratio of no more than 10:1 and a maximum group size of 20.

For the Seeds analysis, counts of children and adults were conducted twice during each child care, family child care, or child care center observation. Group size is the observed number of children at the time of observation. Child/adult ratios are computed as the number of observed children divided by the number of observed adults (including paid and volunteer staff). Multiple counts for the same child care business were averaged to create each business' mean group size and child-adult ratio. Ratios were not tabulated separately for infants, toddlers, and preschoolers, and thus we cannot determine whether child care businesses met the licensing standards.

Additional Impact Analyses

The impact analyses presented in Chapter III included the key variables that produced stable impact estimates and were required to understand service use and the components of the Seeds scores. Here we present two additional sets of impacts: (1) supplemental impacts on variables reported by providers, center directors, and center teachers that include those presented in Chapter III as well as a few additional variables; and (2) impacts on ERS total scores, subscale scores, and items presented separately for each of the three measures.

Impacts on Additional Self-Reported Variables. As noted in chapters II and III, we did not present impacts in the main report for the 14 center directors because the sample size is small and when few directors endorsed an item or almost all endorsed it, the statistical analyses were unstable and were not valid. In addition, we presented variables that could be constructed in the same way for both family child care providers and centers. Tables A.7, A.8, A.9, and A.10 present the longer set of variables we analyzed for Seeds impacts. In addition, Table A.10 presents the impacts on directors that could be assessed (where the impacts could be computed despite the small sample size). A few variables with significant impacts were not included in Chapter III.

Impacts on ERS Total, Subscale, and Item Scores by Measure. We present these exploratory impact analyses to help understand which subscales and items on the three different ERS measures contribute to the provider- and center-level impacts presented in Chapter III. These analyses are meant to be used for program improvement, there is not a research literature testing the item-level reliability of these items. In addition, it is possible that small changes in the environment on some of the items may lead to large increases in an item score. Given these issues and the small sample sizes, these analyses should be interpreted and used with caution. Tables A.11, A.12, and A.13 present these impacts for the FCCERS-R, ITERS-R, and ECERS-R, respectively. We note that these scores and items are correlated, and we have conducted far more statistical tests than the

number of providers and classrooms observed (these impacts are not adjusted to account for multiple comparisons). This is another reason to interpret these findings with caution and consider them exploratory. There are patterns across subscales and items that might be useful when coaches work with providers and center staff to develop quality improvement plans (QIPs). For example, in the first months of coaching, coaches and the providers and staff they work with might want to focus on improving the items that we found have large impacts while simultaneously planning for how to address the items with smaller or no impacts. In addition, analyses of how the ERS subscales and items contribute to the overall impacts observed also lead to new questions for future research. Next we discuss the item-level impacts observed for each of the ERS, highlighting the subscales and items with treatment-control group differences above 1.5 points.[4]

We conducted FCCERS-R observations in 23 treatment and 20 control group family child care homes at followup. As described in Chapter III, we found significant, positive treatment-control impacts on the FCCERS-R total score (3.88 for treatments and 3.00 for controls) and four FCCERS-R subscales: Personal Care Routines, Activities, Program Structure, and Parents and Staff (Table A.11).

- Within the Personal Care Routines subscale (0.65-point difference between treatment and control groups), Greeting and Departing, and the Meals and Snacks items had higher treatment group scores than the control group. The other four items assess areas that are challenging in all settings and tend to be rated lowest on the ERS, including Nap/Rest and Diapering/Toileting.
- The 0.90-point differential between treatment and control providers in the Activities subscale appears to be driven by significantly higher treatment scores on one item: Art (1.83 points).
- Within the Program Structure subscale (1.76 points), Schedule (1.97 points) has a large group difference.
- Within the Parents and Staff subscale (1.33 points) three items one items had a large group difference: Opportunities for Professional Growth (2.01 points).

We conducted ITERS-R observations in 14 treatment and 11 control group classrooms. Of the three ERS measures, we have the smallest sample size for the ITERS-R and thus the pattern of findings provided here should be interpreted with caution. We found a 1.80 point difference between the treatment and control groups on the ITERS-R Total Score (4.65 versus 2.85) (Table A.12). We found significant treatment-control group differences on all ITERS-R subscales except Space and Furnishings.

- Within the Personal Care Routines subscale (1.78 points), all of the item scores except Greeting and Departing were higher in the treatment group than in the control group. The group difference for the Greeting and Departing Item was large (2.29 points), but it was not statistically significant.
- The 2.15-point differential between treatment and control centers on the Listening and Talking subscale appears to be driven by significantly higher treatment scores on two of

[4] This level of impact is likely to be robust even after controlling for multiple comparisons.

the three subscale items: Helping Children Use Language (3.05 points) and Helping Children Understand Language (2.38 points).

- The 1.54-point differential between treatment and control centers in the Activities subscale appears to be driven by significantly higher treatment scores in 4 of the 10 items: Active Physical Play (1.50 points), Dramatic Play (2.72 points), Nature/Science (2.51), and Promoting Acceptance of Diversity (1.92 points).
- Within the Interaction subscale (2.10-point difference), three of the four items have significant group differences: Supervision of Play and Learning (2.96 points), Peer Interaction (2.06 points), and Discipline (2.09 points).
- Two of the three Program Structure subscale items contribute to the 2.07-point difference we observed: Group Play Activities (2.89 points) and Free Play (2.18 points).
- The 2.50-point differential between treatment and control centers in the Parents and Staff subscale arises from impacts on all seven of its items, with the largest differences observed for three items: Provisions for Parents (3.02 points), Supervision and Evaluation of Staff (3.17 points), and Opportunities for Professional Growth (3.84 points).

For the ECERS-R Total Score, there was a 1.46-point difference between the 17 ECERS-R classrooms in the treatment group centers and the 15 classrooms in the control group. We found significant treatment-control differences on all seven of the ECERS-R subscales (Table A.13).

- Within the Space and Furnishings subscale, the difference of 0.88 appears to be driven by significantly higher treatment scores on two items: Room Arrangement for Play (difference of 3.07 points) and Child Related-Display (difference of 1.99 points).
- The 1.67-point differential between treatment and control centers in the Personal Care Routines subscale appears to be driven by higher treatment scores on three of the six items: Greeting/Departing (3.13 points), Meals and Snacks (1.55 points), and Safety Practices (3.68 points). Treatment and control group centers were not significantly different on the Nap/Rest and Toileting/Diapering items.
- The 2.12-point differential between treatment and control group centers in the Language and Reasoning subscale appears to be driven by significantly higher treatment group scores on three of the four items: Books and Pictures (2.32 points), Encouraging Children to Communicate (4.61 points), and Informal Use of Language (2.43 points).
- Within the Activities subscale, the 0.80 difference between the treatment and control group classrooms was only marginally significant at (0.10) and includes group differences on 2 of the 10 items: Sand and Water Play (1.51 points) and Dramatic Play (1.62 points).
- The 2.27-point differential between treatment and control centers in the Interaction subscale appears to be driven by a higher treatment scores on four of the five items: Supervision of Gross Motor Activities (1.81 points), General Supervision of Children (3.19 points), Discipline (1.95 points), and Interactions among Children (2.22 points).
- The 2.10-point differential between treatment and control centers in the Program Structure subscale appears to be driven by higher treatment scores on one of the three items: Free Play (2.48 points).

- The 1.63-point differential between treatment and control centers in the Parents and Staff subscale appears to be driven by significantly higher treatment scores on three of the six items: Provisions for Parents (2.37 points), Staff Interaction and Cooperation (2.53 points), and Opportunities for Professional Growth (2.15 points).

Taken together, these item-level group differences provide a pattern of impacts on areas that are generally considered quite challenging to improve—areas that require and promote children's language interactions and development. Next steps include working with Thrive, Rb5, and CCR to gather additional information about how the coaching model and quality improvement funds may have affected these particular areas.

Sensitivity Tests to Assess Effects of Staff Reporting Inconsistencies

We found that providers and center staff who participated at both baseline and follow-up did not always report information in interviews and questionnaires consistently. In a number of instances these inconsistencies were not plausible, for example, reporting less education or experience at follow-up than at baseline. To assess the effect of these inconsistencies on Seeds scores and potential bias in reporting across groups, we conducted a number of sensitivity tests using the provider and center director. We conclude that the findings as presented in Chapter III are robust and overall are not affected by the observed respondent reporting errors.

For questions about years of work experience in which we found discrepancies, we used directors' and providers' baseline information, and replaced their follow-up information with their reported years of experience at baseline plus one year (this accounted for the time period between baseline and follow-up). For example, a director that reported 7 years of experience at baseline but only 4 years of experience at follow-up would have the 4 years at follow-up replaced by 8 years (7 years plus 1 year). We found that recomputing the Seeds scores based on these values would decrease the Seeds experience component of three providers, but it would not affect any final Seed scores among these providers. In other words, correcting these discrepancies resulted in no real change to final Seeds scores among providers or directors.

For questions about education levels in which we found discrepancies, we used the highest level of education reported by directors and providers—either their baseline or follow-up report (whichever of the two was higher) for sensitivity tests. We found that recomputing Seeds scores based on these values would increase the Seeds education component of three providers and one director, but it would affect the final Seed scores of just two providers in the study. In other words, correcting these discrepancies resulted in very few changes to final Seeds scores among providers and directors.

In light of these findings, we believe that the Seeds scores would change very little if it were possible to contact all providers, directors, and teachers that gave answers during follow-up that were inconsistent with those provided at baseline. We believe that despite these issues, the study is still methodologically strong, particularly because there does not seem to be a clear pattern of more or different types of discrepancies among treatment centers and providers versus control centers and providers. Even if it were possible to contact all staff (which is unlikely given turnover), it is possible they might report a third answer or that recall would be severely impaired given that baseline was in spring 2009 and follow-up in winter 2010. This is an important lesson about the problems with self-report data even for something as seemingly straightforward as education and experience. These areas would benefit from a verification process if they are included in QRIS and QRIS evaluations.

The only area in which we saw more discrepancies among treatment providers than among control providers was the question about years of experience in child care. At follow-up, family child care providers in treatment tended to overestimate their years of experience relative to what they reported at baseline, whereas providers in control tended to underestimate their years of experience relative to what they reported at baseline. This may point to a possible bias among treatment providers, who may have inflated their years of experience at follow-up given their better understanding of Seeds scores following the intervention. However, because these overestimations do not appear to affect final Seed scores, we maintain that the impact estimates are rigorous and largely bias-free.

Implementation Study Design, Data Collection, and Analysis

Site Visits

We developed site visit protocols, including interview and focus group discussion guides, based on research questions for the Seeds implementation study. We worked closely with Rb5 and CCR to plan the baseline and follow-up site visits, identify and recruit participants for individual interviews and focus groups, and schedule the activities. During the visit, we explored key research questions and topics with multiple informants to triangulate the information we obtained and compare responses across informants with different perspectives.

Analysis of the site visit data was an iterative process. The first step was to develop a coding scheme to apply to the site visit data (Table A.14). We organized the coding scheme according to key research questions. Within each question, we defined codes for key themes and subtopics covered during the interviews and focus groups. The scheme also categorized data by Early Learning Initiative (ELI) community; type of participant (for example, coordinator, coaches, and child care providers); and site visit round (baseline and ELI evaluation years 1, 3, and 7).

The next step was to write up interview and focus group field notes. To facilitate consistent note writing and ensure comparable information across activities and communities, we developed a report template organized according to research questions and key topics. We reviewed write-ups to ensure that field notes were consistent and complete.

We used a qualitative analysis software package, Atlas.ti (Scientific Software Development 1997), to facilitate organizing and synthesizing the qualitative data. We loaded the coding scheme and all field notes into Atlas.ti and two project team members then coded the field notes. To ensure reliability across coders, all coders coded an initial sample of interview reports and then compared and resolved any discrepancies. In addition, the lead coder reviewed a sample of coded reports during the coding process to check reliability.

After all field notes were coded, the research team conducted searches using Atlas.ti to retrieve data on our research questions and subtopics. Data were retrieved on particular codes across all respondents, from individual respondents, and for different categories of respondents (such as coaches and center directors). Finally, we used the system to retrieve the relevant data on specific topics and assess the consistency and quality of information across sources. This approach ensured quality and consistency in our analyses across the project team.

ETO Service Use Data

Thrive by Five contracted with Social Solutions, a provider of performance management software, to develop a web-based data system for the Seeds modified field test. Social Solutions developed a data system using its Efforts to Outcomes (ETO™) software. Social Solutions was responsible for training staff on ETO and providing ongoing technical assistance. The ETO database was designed to track staff and provider progress toward a number of goals defined jointly by staff and coaches in the QIPs. Each record entered into ETO by the Seeds coaches tracked the time the coaches spent with staff; the mode of coaching delivery (including one-on-one, group, email, or telephone); and the specific goal within a QIP they worked on. In addition, Seeds site coordinators and coaches entered data on the types of professional development supports staff at participating providers received through Seeds (including providers assigned to both the treatment and control groups).

Social Solutions gave Mathematica access to the ETO data in fall 2009 in preparation for data analysis. Mathematica worked with Thrive, Social Solutions, Rb5, and CCR to ensure the completeness of the data. A list of the final set of variables included in our analysis is included in Table A.15.

To examine the types and levels of services providers enrolled in Seeds received during the period from June 1, 2009, through December 31, 2009, the research team computed descriptive statistics—such as frequencies, means, and distributions—of variable characteristics. The team then computed means across all providers and for subgroups of providers (including family child care providers and child care centers) and staff (including directors/owners, lead teachers, and assistant teachers) as appropriate.

Table A.1. Baseline Demographic and Professional Characteristics of Family Child Care Providers, by Treatment and Control Status

Characteristic	Weighted Means or Percentages: Full Sample	Standard Deviation: Full Sample	Treatment	Control	Difference
Female	100	0	100	100	0
Age (Years)					
Mean	45.0	1.2	44.5	45.4	-0.9
Median	46		44	47	-3
Race					
African American	2	2	4	0	4
Hispanic	57	7	57	60	-2
White, non-Hispanic	15	5	16	13	4
Other	26	6	22	28	-5
Multiple	0	0	0	0	0
Highest Grade or Year of School Completed					
Less than high school	46	7	52	43	9
High school or GED	32	7	28	34	-6
Some college	9	4	9	9	0
AA or BA	12	5	9	14	-6
Graduate or professional work or degree	2	2	3	0	3
Field of Study					
Child Development, Early Childhood Education, or Elementary Education	7	4	9	6	2
Other	15	5	12	17	-5
No field	78	6	80	77	3
Schooling Included 5 or More College Courses in ECD	5	4	4	6	-2
Years Worked at Current Job	6.3	0.7	5.1	7.5	-2.5*
Annual Salary (Dollars)					
Mean	39,630	3,767	34,270	44,118	-9,847
Median	38,000		30,000	40,000	-10,000
Sample Size	**49**		**26**	**23**	

Source: Family Child Care Provider Interview completed in spring 2009.

Note: Full sample refers to the combined value for both the treatment and control groups.

* Significant at 0.10.

AA = associate of the arts; BA = bachelor of the arts; GED = general equivalency diploma.

Table A.2. Baseline Child Care Center Characteristics, by Treatment and Control Status

Characteristic	Weighted Means or Percentages: Full Sample	Standard Deviation: Full Sample	Treatment	Control	Difference
Center Type					
Private, nonprofit	61	n.a.	57	64	-7
Private, for-profit	34	n.a.	43	25	18
Other	5	n.a.	0	11	-11
Number of Classrooms	4.3	0.6	4.1	4.4	-0.3
Number of Lead Teachers	4.9	0.7	5.0	4.8	0.3
Number of Assistant Teachers	4.1	0.9	5.2	2.9	2.3
Number of Children Enrolled in Center	59.2	8.9	66.8	51.6	15.2
Number of Children That Receive a Child Care Subsidy	33.0	7.6	38.2	27.7	10.5
Number of Children per Classroom with an Individualized Education Plan	0.1	0.1	0.2	0.1	0.1
Center Uses a Specific Curriculum or Combination of Curricula	77	13	89	64	25
Sample Size	**14**		**7**	**7**	

Source: Center Director Questionnaire and Teacher Questionnaire completed in spring 2009.

Note: Full sample percentage and standard deviation refer to the combined value for both the treatment and control groups.

n.a. = not applicable.

Table A.3. Baseline Demographic and Professional Characteristics of Child Care Center Directors, Lead Teachers, and Assistant Teachers, by Treatment and Control Status

	Center Director					Lead Teacher					Assistant Teacher				
Characteristic	Weighted Means or Percentages: Full Sample	SD: Full Sample	T	C	Diff	Weighted Means or Percentages: Full Sample	SD: Full Sample	T	C	Diff	Weighted Means or Percentages: Full Sample	SD: Full Sample	T	C	Diff
Female	89	8	100	79	21	97	2	100	93	7	96	3	93	100	-7
Age (Years)															
Mean	42.9	3.6	43.9	42.2	1.7	35.4	1.7	38.2	32.5	5.7*	29.3	1.6	27.3	31.7	-4.4
Median	40.7		40.8	40.3	0.5	33.0		40.0	28.0	12.0	26.0		24.0	29.0	-5.0
Race															
African American	0	0	0	0	0	7	3	5	8	-3	13	6	2	26	-24*
Hispanic	5	5	0	11	-11	20	6	18	22	-4	19	6	16	22	-6
White, non-Hispanic	73	13	68	79	-11	50	9	54	47	7	49	8	72	24	48***
Other	5	5	0	11	-11	9	6	0	17	-17*	8	5	0	18	-18*
Multiple	16	11	32	0	32	14	10	22	6	17	10	6	11	10	1
Highest Grade or Year of School Completed															
Less than high school	0	0	0	0	0	6	3	2	11	-9	21	8	16	26	-10
High school or GED	5	5	11	0	11	22	6	26	19	7	27	8	30	24	6
Some college	39	14	21	57	-36	41	9	46	36	10	35	8	38	32	6
AA or BA	38	14	43	32	11	24	7	23	26	-3	17	6	16	18	-2
Graduate or professional work or degree	18	10	25	11	14	6	3	4	8	-4	0	0	0	0	0
Field of Study															
Child Development, Early Childhood Education, or Elementary Education	71	13	75	68	7	37	9	36	39	-3	31	7	35	26	9
Other	23	13	14	32	-18	33	8	36	29	6	20	6	16	24	-8
Not Applicable (No Advanced Study)	5	5	11	0	11	30	7	28	32	-4	50	9	49	50	-1
Schooling Included 5 or More College Courses in ECD	79	10	68	89	-21	38	9	42	34	8	34	8	25	44	-19
Years Worked at Current Job	5.9	2.1	9.4	3.5	6.0	3.6	0.9	5.7	1.7	4.0***	2.6	0.4	2.8	2.4	0.4

Table A.3 (continued)

Characteristic	Center Director					Lead Teacher					Assistant Teacher				
	Weighted Means or Percentages: Full Sample	SD: Full Sample	T	C	Diff	Weighted Means or Percentages: Full Sample	SD: Full Sample	T	C	Diff	Weighted Means or Percentages: Full Sample	SD: Full Sample	T	C	Diff
Annual Salary (Dollars)															
Mean	32,237	2,428	33,880	30,075	3,804	19,208	1,258	20,311	18,053	2,257	16,824	700	14,893	21,652	-6,759***
Median	33,600		35,000	33,600	1,400	20,000		20,000	20,000	0	16,416		15,000	23,000	-8,000
Sample Size	**14**		**7**	**7**		**59**		**31**	**28**		**51**		**34**	**17**	

Source: Center Director Questionnaire and Teacher Questionnaire completed in spring 2009.

Note: Full sample percentage and standard deviation (SD) refer to the combined value for both the treatment and control groups.

* Significant at 0.10; *** Significant at 0.01.

AA = associate of the arts; BA = bachelor of the arts; T = treatment group C = control group; Diff. = difference; ECD = early child development/child development; GED = general equivalency diploma; SD = standard deviation.

Table A.4. Inter- Rater Reliability on ERS and Arnett CIS at Baseline and Follow- Up

	Baseline Training		Follow-Up Training	
	Average Percentage Agreement Within 1 Point	Cohen's Kappa	Average Percentage Agreement Within 1 Point	Cohen's Kappa
ITERS-R	89	0.55	95	0.77
ECERS-R	94	0.63	90	0.59
FCCERS-R	92	0.62	93	0.72
Arnett CIS (Infant/Toddler Classrooms)	98	0.54	98	0.51
Arnett CIS (Early Childhood Classrooms)	98	0.68	96	0.58
Arnett CIS (Family Providers)	95	0.46	98	0.68

Source: Practice child care quality observations completed during baseline training (March 2009) and during follow-up training (December 2009).

Note: Inter-rater reliability testing compared field interviewer scores with the scores of the gold standard trainer.

CIS = Caregiver Interaction Scale; ECERS-R = Early Childhood Environment Rating Scale-Revised; ERS = Environment Rating Scales; FCCERS-R = Family Child Care Environment Rating Scale-Revised; ITERS-R = Infant/Toddler Environment Rating Scale-Revised.

Table A.5. Seeds Impact Study Baseline and Follow- Up Response Rates, by Treatment Group

	Full Sample		Treatment Group		Control Group	
	Number	Percentage	Number	Percentage	Number	Percentage
Family Child Care Providers (n = 52)						
Observations						
Baseline	49	94	26	96	23	92
Follow-Up	43	83	23	85	20	80
Interviews						
Baseline	49	94	26	96	23	92
Follow-Up	48	92	25	93	23	92
Child Care Centers (n = 14)						
Classroom Observations						
Baseline (n = 58)	58	100	32	100	26	100
Follow-Up (n = 57)	57	100	31	100	26	100
Questionnaires						
Center Directors						
Baseline (n = 14)	14	100	7	100	7	100
Follow-Up (n = 14)	14	100	7	100	7	100
Lead Teachers						
Baseline (n = 59)	59	100	32	100	27	
Follow-Up (n = 56)	56	100	30	100	26	100
Assistant Teachers						
Baseline (n = 56)	51	91	33	92	18	90
Follow-Up (n = 55)	55	100	33	100	22	100

Source: Mathematica impact study, 2009–2010.

Note: The child care family provider response rates were calculated by using the number of providers who consented to participate in the study as the denominator. For child care centers, the observation denominator is the number of classrooms and for the center director and teacher questionnaires the denominator is the number of center directors and teachers present at the centers during the baseline and follow-up data collection periods. Center directors provided Mathematica a list of all staff working at the center a few weeks before each data collection began.

Table A.6. Descriptive Information for Scales Constructed from Observation Measures, for the Full Sample

Measure	Full Sample Mean	Standard Deviation	Reported Range		Cronbach's Alpha
			Minimum	Maximum	
Total ITERS- R Score					
Baseline	3.42	0.91	1.41	5.03	0.93
Follow-Up	3.81	0.96	1.71	5.30	0.94
ITERS- R Subscale Scores					
Space and Furnishings					
Baseline	3.73	0.72	2.40	5.20	0.34
Follow-Up	4.02	1.07	2.40	6.00	0.72
Personal Care					
Baseline	2.33	0.89	1.00	4.00	0.60
Follow-Up	2.69	1.12	1.00	5.33	0.80
Listening and Talking					
Baseline	3.97	1.59	1.33	7.00	0.72
Follow-Up	4.28	1.44	1.33	7.00	0.80
Activities					
Baseline	2.91	0.88	1.10	4.66	0.77
Follow-Up	3.33	1.01	0.14	5.77	0.84
Interaction					
Baseline	4.37	1.62	1.25	6.25	0.83
Follow-Up	4.71	1.37	2.25	6.75	0.79
Program Structure					
Baseline	4.07	1.55	1.33	6.66	0.69
Follow-Up	4.08	1.37	1.00	7.00	0.60
Parents and Staff					
Baseline	3.70	1.24	1.66	6.29	0.81
Follow-Up	4.32	1.36	1.50	6.57	0.81
Sample Size					
Baseline	**25**	-	-	-	-
Follow- Up	**25**	-	-	-	-
Total ECCERS- R Score					
Baseline	4.06	1.13	1.85	5.48	0.87
Follow-Up	4.20	1.02	2.07	5.98	0.85
ECERS- R Subscale Scores					
Space and Furnishings					
Baseline	4.24	0.86	2.38	5.75	0.62
Follow-Up	4.44	0.98	2.50	6.25	0.74

Table A.6 (continued)

Measure	Full Sample Mean	Standard Deviation	Reported Range		Cronbach's Alpha
			Minimum	Maximum	
Personal Care					
Baseline	2.95	1.23	1.00	5.40	0.70
Follow-Up	2.79	1.02	1.00	5.33	0.69
Language-Reasoning					
Baseline	4.83	1.45	1.75	7.00	0.84
Follow-Up	4.84	1.27	2.00	7.00	0.81
Activities					
Baseline	3.66	1.14	2.00	5.70	0.80
Follow-Up	3.93	1.06	2.00	6.33	0.87
Interaction					
Baseline	4.88	2.04	1.20	7.00	0.95
Follow-Up	4.76	1.64	1.80	6.80	0.91
Program Structure					
Baseline	4.70	1.71	1.33	7.00	0.71
Follow-Up	4.79	1.38	1.66	7.00	0.73
Parents and Staff					
Baseline	3.95	1.07	1.66	5.50	0.66
Follow-Up	4.49	1.40	1.50	7.00	0.82
Sample Size					
Baseline	**35**	-	-	-	-
Follow- Up	**32**	-	-	-	-
Total FCCERS- R Score					
Baseline	2.87	0.72	1.71	5.22	0.83
Follow-Up	3.50	1.04	2.11	6.38	0.91
FCCERS- R Subscale Scores					
Space and Furnishings					
Baseline	3.06	0.88	1.66	5.00	0.62
Follow-Up	3.35	1.19	1.50	6.33	0.75
Personal Care					
Baseline	2.09	0.76	1.00	4.83	0.46
Follow-Up	2.50	1.01	1.16	6.50	0.71
Listening and Talking					
Baseline	3.29	1.27	1.00	7.00	0.65
Follow-Up	3.81	1.46	1.33	6.66	0.76
Activities					
Baseline	2.52	0.72	1.20	5.27	0.73
Follow-Up	3.04	1.02	1.70	5.91	0.87

Table A.6 (continued)

Measure	Full Sample Mean	Standard Deviation	Reported Range		Cronbach's Alpha
			Minimum	Maximum	
Interaction					
Baseline	3.48	1.71	1.00	7.00	0.89
Follow-Up	4.49	1.61	1.25	7.00	0.85
Program Structure					
Baseline	3.14	1.48	1.00	7.00	0.72
Follow-Up	4.26	1.97	1.00	7.00	0.86
Parents and Provider					
Baseline	3.52	1.22	1.66	6.75	0.69
Follow-Up	4.78	1.27	2.00	7.00	0.71
Sample Size					
Baseline	**49**	-	-	-	-
Follow- Up	**43**	-	-	-	-
Total Arnett CIS Score (Infant/Toddler Classrooms)					
Baseline	3.32	0.42	2.53	3.96	0.95
Follow-Up	3.39	0.27	2.83	4.77	0.87
Total Arnett CIS Score (Early Childhood Classrooms)					
Baseline	3.47	0.49	2.23	4.00	0.96
Follow-Up	3.43	0.41	2.06	4.00	0.94
Total Arnett CIS Score (Family Providers)					
Baseline	3.06	0.38	2.20	3.83	0.91
Follow-Up	3.21	0.41	2.03	4.00	0.92

Source: Baseline Child Care Observations completed in spring 2009 and follow-up Child Care Observations completed in winter 2010.

CIS = Caregiver Interaction Scale; ECERS-R = Early Childhood Environment Rating Scale-Revised; FCCERS-R = Family Child Care Environment Rating Scale-Revised; ITERS-R = Infant/Toddler Environment Rating Scale-Revised.

Table A.7. Seeds Impact on Family Child Care Provider- Reported Training/Education, Frequency of Coaching, Business Plan, and Curriculum (Percentages)

Characteristic	Treatment	Control	Difference	Effect Size[a]
Provider is currently enrolled in teacher-related training or education program	34	26	7	0.84
Provider currently working toward CDA	29	10	18	3.00
Frequency of training and technical assistance activities				
Once a month or more	41	46	-5	-0.47
Frequency of mentor/coach visits				
Once a week or more	65	0	65***	n.a.
Mentor/coach came for extended visit	8	0	8	n.a.
Provider has a business plan	70	48	22	2.10
Provider uses a curriculum or combination of curricula	53	51	1	0.13
Sample Size	**25**	**23**		

Source: Baseline Family Child Care Provider Interview completed in spring 2009, and Follow-up Family Child Care Provider Interview completed in winter 2010.

Note: Effect size is not defined when the control group mean equals zero.

[a] Effect size is defined as the treatment-control difference divided by the standard deviation of the outcome measure among the control group.

*** Significant at 0.01.

CDA = child development associate.

n.a. = not applicable.

Table A.8. Seeds Impact on Teacher- Reported Enrollment in Training/Education Programs and Frequency of Coaching (Percentages)

	Lead Teachers				Assistant Teachers			
Characteristic	Treatment	Control	Difference	Effect Size[a]	Treatment	Control	Difference	Effect Size[a]
Center Makes an Effort to Help Teachers Get Their CDA	78	77	1	0.10	76	72	4	0.39
Teacher Currently Enrolled In Teacher-Related Training or Education Program	60	15	45***	4.72	52	15	37***	3.60
Frequency of Courses Attended at Community or Four-Year College								
Once a month or more	36	2	34***	6.37	29	9	20	3.37
Frequency of Training Workshops Attended								
Once a month or more	35	14	22	2.02	46	6	40**	3.33
Frequency of Ongoing Consultation from a Specialist, Coach, or Mentor								
Once a week or more	70	14	56***	13.70	53	9	44**	4.31
Sample Size	**30–31**	**26–28**			**34–36**	**17–19**		

Source: Baseline Teacher Questionnaire completed in spring 2009, and Follow-up Teacher Questionnaire completed in winter 2010.

[a] Effect size is defined as the treatment-control difference divided by the standard deviation of the outcome measure among the control group.

** Significant at 0.05; *** Significant at 0.01.

CDA = child development associate.

Table A.9. Seeds Impacts on Education and Experience of Family Child Care Providers (Percentages Unless Otherwise Indicated)

	Family Care Providers			
Characteristic	Treatment	Control	Difference	Effect Size[a]
Highest Grade or Year of School Completed Some college or higher	38	37	0	0.04
Earned at Least 3 Credits in the Past 6 Months	21	14	6	0.90
Has Child Development Associate (CDA) Credential	11	24	-13*	-1.29
Member of a Professional Association for Early Childhood Education	49	22	27*	2.80
Years Worked at Current Job	7.2	6.6	0.6	0.47
Years of Early Childhood Experience[b]	12.1	8.8	3.3**	2.19
Sample Size	**25**	**23**		

Source: Baseline Family Child Care Provider Interview completed in spring 2009, and Follow-up Family Child Care Provider Interview completed in winter 2010.

[a] Effect size is defined as the treatment-control difference divided by the standard deviation of the outcome measure among the control group.

[b] At follow-up, providers in treatment were more likely than providers in control to report more years of experience than they reported at baseline. This discrepancy is discussed in the summary of sensitivity analyses.

* Significant at 0.10; ** Significant at 0.05.

Table A.10. Seeds Impacts on Education and Experience of Lead and Assistant Teachers (Percentages Unless Otherwise Indicated)

	Lead Teachers				Assistant Teachers			
Characteristic	Treatment	Control	Difference	Effect Size[a]	Treatment	Control	Difference	Effect Size[a]
Highest Grade or Year of School Completed								
AA, BA, or higher	28	30	-2	-0.20	20	19	1	0.10
Has Child Development Associate (CDA) credential	26	13	12	1.48	30	10	20	1.63
Member of a Professional Association for Early Childhood Education	28	24	5	0.53	23	8	14	2.21
Years Worked at Current Job	3.1	4.5	-1.4**	-3.85	2.2	3.5	-1.3*	-1.98
Years of Early Childhood Experience[b]	8.7	8.0	0.7	0.76	5.5	3.2	2.2	2.12
Center Makes an Effort to Help Teachers Get Their CDA	78	77	1	0.10	76	72	4	0.39
Teacher Currently Enrolled In Teacher-Related Training or Education Program	60	15	45***	4.72	52	15	37***	3.60
Frequency of Courses Attended at Community or Four-Year College								
Once a month or more	36	2	34***	6.37	29	9	20	3.37
Frequency of Training Workshops Attended								
Once a month or more	35	14	22	2.02	46	6	40**	3.33
Frequency of Ongoing Consultation from a Specialist, Coach, or Mentor								
Once a week or more	70	14	56***	13.70	53	9	44**	4.31
Sample Size	**30–31**	**26–28**			**34–36**	**17–19**		

Source: Baseline Teacher and Center Director Questionnaire completed in spring 2009, and Follow-up Teacher and Center Director Questionnaire completed in winter 2010.

Note: Effect size is not defined when the control group mean equals zero.

[a] Effect size is defined as the treatment-control difference divided by the standard deviation of the outcome measure among the control group.

[b] This variable is the maximum number of years of experience that teachers reported serving any age group of children. Due to its construction, this variable may not capture the total years of early childhood experience among teachers that have transitioned from serving one age group to a different age group during their career.

* Significant at 0.10; ** Significant at 0.05; *** Significant at 0.01.

AA = associate of the arts; BA = bachelor of the arts.

n.a. = not applicable.

Table A.11. Impacts on FCCERS Total, Subscale, and Item Scores

	Family Child Care Providers (FCCERS-R)			
	Treatment	Control	Difference	Effect Size[a]
Total FCCERS- R Score	3.88	3.00	0.88***	5.54
Space and Furnishings	3.50	3.09	0.41	2.02
Indoor space used for child care	3.90	3.44	0.46	1.16
Furniture for routine care, play, and learning	3.42	3.08	0.34	0.84
Provision for relaxation and comfort	4.52	3.46	1.06**	3.14
Arrangement of indoor space for child care	2.89	2.88	0.00	0.01
Display for children	3.31	2.50	0.81*	3.01
Space for privacy	2.95	3.19	-0.24	-0.62
Personal Care Routines	2.76	2.12	0.65**	3.58
Greeting/departing	6.38	4.75	1.62*	3.40
Meals/snacks	3.35	1.58	1.77**	4.40
Nap/rest	1.40	1.53	-0.13	-0.56
Diapering/toileting	1.75	1.32	0.42	2.44
Health practices	1.75	1.82	-0.06	-0.33
Safety practices	1.93	1.72	0.21	0.96
Listening and Talking	4.25	3.23	1.02*	4.06
Help children understand language	4.54	3.93	0.61	1.74
Help children use language	4.21	4.23	-0.01	-0.03
Use books	4.01	1.55	2.46***	10.39
Activities	3.41	2.51	0.90***	6.62
Fine motor	3.80	3.23	0.58	2.25
Art	4.55	2.72	1.83**	5.40
Music and movement	3.57	3.33	0.24	0.62
Blocks	3.23	2.15	1.08**	5.49
Dramatic play	3.84	2.62	1.21***	5.08
Math/numbers	3.28	2.64	0.64	2.23
Nature/science	3.59	2.48	1.11**	5.14
Sand and water play	3.00	1.70	1.30*	8.83
Promoting acceptance of diversity	3.14	2.71	0.43	1.72
Use of TV, video, and/or computers	3.78	2.36	1.42	4.52
Active physical play	1.79	1.56	0.23	1.95
Interaction	4.70	4.20	0.49	1.57
Supervision of play and learning	4.59	3.83	0.77	1.73
Staff-child interaction	5.63	4.70	0.93	2.23
Discipline	4.23	3.87	0.35	1.34
Interactions among children	4.31	4.37	-0.06	-0.12
Program Structure	5.03	3.27	1.76***	4.40
Schedule	5.07	3.11	1.97***	3.62
Free play	4.39	2.98	1.41**	4.07
Group time	5.21	4.30	0.91	1.28
Provisions for children with disabilities	5.41	5.41	0.00	0.00
Parents and Staff	5.39	4.06	1.33***	5.23
Provisions for parents	5.37	3.97	1.41**	3.71
Balancing personal and caregiver responsibilities	5.39	4.07	1.32**	3.07
Opportunities for professional growth	5.30	3.29	2.01***	5.30
Provisions for professional needs	5.46	5.02	0.44	1.17
Sample Size	**23**	**20**		

Source: Child Care Observation completed in spring 2009 and winter 2010.

n.a. = not applicable.

* Significant at 0.10; ** significant at 0.05; *** significant at 0.01.

Table A.12. Impacts on ITERS- R Total, Subscale, and Item Scores

	Child Care Centers (ITERS-R)			
	Treatment	Control	Difference	Effect Size[a]
Total ITERS- R Score	4.65	2.85	1.80***	6.75
Space and Furnishings	4.19	3.53	0.66	2.75
Indoor space	3.72	2.84	0.88	1.91
Furniture for routine care and play	4.27	3.45	0.82	2.61
Provision for relaxation and comfort	4.49	4.00	0.49	1.12
Room arrangement	4.30	3.67	0.63	1.49
Display for children	4.14	3.67	0.47	1.78
Personal Care Routines	3.60	1.81	1.78***	6.71
Greeting/departing	5.94	3.65	2.29	2.94
Meals/snacks	3.41	0.98	2.44***	9.51
Nap	3.06	1.36	1.70**	6.05
Diapering/toileting	2.72	1.71	1.01*	6.01
Health practices	3.07	1.54	1.53***	5.37
Safety practices	3.28	1.77	1.51**	3.42
Listening and Talking	5.32	3.17	2.15***	5.71
Help children understand language	5.78	3.40	2.38**	4.83
Help children use language	6.35	3.31	3.05***	6.60
Use books	3.81	2.80	1.01	2.60
Activities	4.01	2.47	1.54***	7.39
Fine motor	4.70	3.30	1.39	3.55
Active physical play	3.57	2.08	1.50**	5.75
Art	3.76	2.35	1.42	1.58
Music and movement	3.61	3.55	0.06	0.23
Blocks	3.01	2.91	0.10	0.15
Dramatic play	5.17	2.45	2.72***	6.80
Sand and water play	n.a.	n.a.	n.a.	n.a.
Nature/science	4.04	1.53	2.51***	8.84
Use of TV, video, and/or computers	n.a.	n.a.	n.a.	n.a.
Promoting acceptance of diversity	3.88	1.97	1.92***	7.52
Interaction	5.75	3.66	2.10**	5.03
Supervision of play and learning	6.33	3.36	2.96***	6.05
Peer interaction	6.00	3.93	2.06*	4.49
Staff-child interaction	5.97	4.76	1.22	1.73
Discipline	4.69	2.61	2.09**	5.68
Program Structure	5.22	3.15	2.07***	4.66
Schedule	4.77	3.26	1.51	2.65
Free play	5.12	2.94	2.18**	3.81
Group play activities	5.92	3.03	2.89**	4.28
Provisions for children with disabilities	n.a.	n.a.	n.a.	n.a.
Parents and Staff	5.47	2.97	2.50***	5.54
Provisions for parents	6.21	3.19	3.02***	7.07
Provisions for personal needs of staff	4.14	2.76	1.38*	4.81
Provisions for professional needs of staff	4.25	2.60	1.66*	2.24
Staff interaction and cooperation	5.81	3.94	1.87*	2.51
Staff continuity	5.94	3.50	2.44*	3.80
Supervision and evaluation of staff	6.09	2.92	3.17***	4.00
Opportunities for professional growth	6.18	2.33	3.84***	8.02
Sample Size	**14**	**11**		

Source: Child Care Observation completed in spring 2009 and winter 2010.

[a]Effect size is defined as the treatment-control difference divided by the standard deviation of the outcome measure among the control group.

* Significant at 0.10; ** significant at 0.05; *** significant at 0.01.

ITERS –R = Infant/Toddler Environment Rating Scale-Revised.

n.a. = not applicable.

Table A.13. Impacts on ECERS- R Total, Subscale, and Item Scores

	Child Care Centers			
	Treatment	Control	Difference	Effect Size[a]
Total ECERS- R Score	4.92	3.46	1.46***	4.73
Space and Furnishings	4.85	3.97	0.88**	2.75
Indoor space	4.05	4.05	0.00	-0.01
Furniture for routine care, play and learning	6.36	6.13	0.23	0.74
Furnishings for relaxation and comfort	4.97	4.79	0.18	0.29
Room arrangement for play	5.96	2.90	3.07***	6.10
Space for privacy	4.18	4.11	0.08	0.12
Child-related display	5.29	3.30	1.99***	4.14
Space for gross-motor play	2.01	1.67	0.34	2.92
Gross motor equipment	4.96	4.79	0.17	0.29
Personal Care Routines	3.72	2.05	1.67***	5.79
Greeting/departing	6.44	3.31	3.13***	5.09
Meals/snacks	2.59	1.04	1.55**	10.13
Nap/rest	3.10	3.16	-0.06	-0.11
Toileting/diapering	1.95	1.80	0.16	1.05
Health practices	3.05	1.87	1.18**	3.68
Safety practices	4.98	1.30	3.68***	7.58
Language-Reasoning	5.87	3.75	2.12***	7.31
Books and pictures	4.95	2.63	2.32***	10.55
Encouraging children to communicate	5.91	1.30	4.61***	7.58
Using language to develop reasoning skills	4.77	4.05	0.72	1.78
Informal use of language	6.84	4.42	2.43***	6.25
Activities	4.32	3.52	0.80*	2.55
Fine motor	4.88	3.98	0.90	1.70
Art	5.38	4.14	1.24	1.84
Music and movement	3.44	3.17	0.27	0.67
Blocks	4.89	3.90	0.99	1.89
Sand and water play	4.38	2.87	1.51**	4.47
Dramatic play	5.26	3.64	1.62***	8.91
Nature/science	3.06	3.14	-0.08	-0.14
Math/numbers	4.60	3.51	1.10*	2.62
Use of TV, video, and/or computers	3.47	2.10	1.36	1.31
Promoting acceptance of diversity	3.79	3.13	0.66	3.55
Interaction	6.01	3.74	2.27***	4.07
Supervision of gross motor activities	4.67	2.86	1.81**	3.74
General supervision of children (other than gross motor)	6.12	2.93	3.19***	4.50
Discipline	5.61	3.65	1.95**	2.91
Staff-child interaction	6.06	4.86	1.20	1.88
Interactions among children	6.60	4.38	2.22***	4.65
Program Structure	5.88	3.78	2.10***	4.94
Schedule	3.97	2.98	0.99	3.11
Free play	6.08	3.60	2.48***	4.25
Group time	6.20	5.26	0.94	1.94
Provisions for children with disabilities	n.a.	n.a.	n.a.	n.a.
Parents and Staff	5.08	3.44	1.63***	4.43
Provisions for parents	6.06	3.68	2.37***	4.60
Provisions for personal needs of staff	3.52	2.40	1.12**	3.12
Provisions for professional needs of staff	4.47	3.96	0.51	0.82
Staff interaction and cooperation	6.31	3.78	2.53**	4.72
Supervision and evaluation of staff	5.41	3.98	1.43**	2.94
Opportunities for professional growth	5.11	2.96	2.15***	2.89
Sample Size	**17**	**15**		

Source: Child Care Observation completed in spring 2009 and winter 2010.

* Significant at 0.10; ** significant at 0.05; *** significant at 0.01.

ECERS-R = Early Childhood Environment Rating Scale-Revised.

n.a. = not applicable.

Table A.14. Structured Coding Scheme Used to Organize and Categorize Site Visit Data

Code	Description
Participant Type	
Coordinator	All coordinator/supervisor notes
Coaches	All coaches notes
Center directors	All child care center director notes
Teachers	All child care center-based teacher notes
Family	All family child care provider notes
Community	
East Yakima	All East Yakima notes
White Center	All White Center notes
Data Collection Period	
Baseline	All June 2009 notes
Follow-up	All November 2009 notes
Codes	
Participant info	Name, current position, role and responsibilities, length of time in current position, prior experience, what attracted participant to position, language deliver services in, updates or changes to position
Organization info	Name of organization/child care center, whether work with another teacher
Training hiring	Description of how coaches were hired, interview process, qualifications coordinators were looking for, what attracted coaches to the position, any information about turnover of coaches
Training preservice	Description of training received prior to implementation of Seeds, including topics covered, who provided training, length of training
Training inservice	Description of training received since implementation of Seeds began, including topics covered, who provided training, length of training
Training ERS	Experience using ERS (ECERS-R, ITERS-R, FCCERS-R, SACERS) or another quality measure before pilot began
Training sufficient	Has training been sufficient for work, how helpful training was, most helpful training, most important things learned
Training additional	Additional training needed, topics for additional training
Supervision	Description of supervision including types of supervision, frequency, how information is used
Supervision helpful	How helpful is supervision
Supervision additional	Additional supervision needed
Providers caseload	Number of providers work with, number of center-based versus family providers, if caseload is manageable
Providers description	Description of providers including languages spoken, range of education backgrounds, ages of children in their care, number of children cared for, hours provide care, differences among providers
Providers' strengths	Strengths of providers
Providers' needs	Needs of providers
Recruitment	How providers were recruited into Seeds, including role of coaches and coordinators, what attracted providers, receptivity of providers, providers' expectations about services, reasons providers chose not to participate
Recruitment assigning	How providers were assigned to coaches, if providers were reassigned to other coaches and why
Recruitment retention	Whether providers have dropped out, reasons providers dropped out
Relationship strategies	Strategies used to build relationships with providers, including strategies that work best, additional strategies needed, how providers feel about coaches/what they like best about coaches, whether coaches seem knowledgeable
Relationship challenges	Challenges encountered related to building and maintaining relationships, things providers would like coaches to do differently

Table A.14 *(continued)*

Code	Description
Relationship interpreters	Use of interpreters during coaching visits and whether this effects the relationship
Scores sharing	How were ERS and Seeds scores shared with providers, what role did coaches play in sharing scores, did center directors share scores with teachers and assistant teachers
Scores reactions	How did providers respond to ERS scores and Seeds ratings, what questions or issues arose, what were their concerns
QIP development	Development of QIP, including information used, role of coaches/providers in developing QIP, how long it took to develop QIP
QIP goals	Types of goals in QIPs, differences by provider type
QIP coaching	How QIPs are used during the coaching visits, where QIP is stored and how often it is referenced
QIP monitoring	How progress on QIPs is monitored, role of coaches, role of providers, how disagreements are handled regarding progress, progress made by providers on QIPs (completion of goals)
Grants	How providers are using grants, role of coaching, how grants are monitored
Coaching frequency	How often coaches and providers meet (plan to meet), length of visits, whether providers cancel appointments, barriers to meeting, process used to schedule visits
Coaching visits	What happens during visits, what topics are discussed, whether coaches or providers determine plan for visit, whether coaches follow a curriculum (observations/modeling/resources provided), disagreements between providers and coaches
PD opportunities	Types of professional development opportunities available, types of PD providers enrolled in, uptake of PD, role of coaches in helping providers access PD
PD needs	Other PD topics/opportunities providers would like available, barriers to participation in PD
ETO	Training on ETO, usefulness of ETO, how ETO is used, issues with ETO, whether coaches have changed practice because of ETO, if coordinators use ETO to monitor coaches work
Support	Whether any additional support is needed from CCR/Rb5, Thrive, Gates for implementing Seeds
Coordination	Whether participant coordinates with other demonstration community, whether participant coordinates with Gates, Thrive, Thrive mentor (Holli Bennett), successes and challenges of coordination
Lessons quality	Whether quality or provider practice has improved, whether improvement is expected on followup ERS, types of changes that were easier for providers to make, differences by provider type, whether expect changes to be sustained, improvements expect in the future, is Seeds model working
Lessons Seeds	Whether providers would participate in another quality improvement or professional development initiative, parents' reactions to Seeds, other providers reactions to Seeds, advice for another provider considering enrolling in Seeds
Lessons successes	Successes
Lessons challenges	Challenges
Lessons implementation	On a scale from 1 to 5, how well able to implement Seeds
Lessons changes	Additional services would like offered through Seeds, suggestions for change to the model, advice for model developers

Source: Mathematica implementation study, 2009.

ECERS-R = Early Childhood Environment Rating Scale-Revised; ETO = Efforts to Outcomes; FCCERS-R = Family Child Care Environment Rating Scale-Revised; ITERS-R = Infant/Toddler Environment Rating Scale-Revised; PD = professional development; QIP = Quality Improvement Plan; SACERS = School-Age Care Environment Rating Scale.

Table A.15. ETO Variables Used for Seeds Implementation Study

Variable	Description	Values
Participant Unique Identifier	Unique identification number for child care provider staff	Automatically assigned by ETO
Program	This either lists the coach assigned to a given provider or flags the time with White Center or East Yakima Professional Development	Coach Name – QRIS Coach White Center Professional Development Yakima Professional Development
Provider Agency	The Child Care Center or Family Care Provider at which the participant worked	FCC Name CCC Name
Staff Type	Position of the child care provider	Assistant Teacher Lead Teacher Director/Owner
Contact Location	Mode of contact during coaching sessions between the coach and the child care provider	Email Prep & Documentation Group One-on-One Phone NB: Group
Date of Contact	Date of contact between the coach and provider	June 1, 2009 to December 31, 2009
Outcome Category	The general category of the goal which the provider and coach worked on during a given coaching session	Curriculum and Learning Environment Family Relationships & Community Partnerships Health and Safety Management Practices Other PD Tracking Professional Development
Outcome Name Details	The category and number of the goal which the provider and coach worked on during a given coaching session	Curriculum and Learning Environment # Family Relationships & Community Partnerships # Health and Safety # Management Practices # Other # Professional Development #
Time Spent on Contact	Time spent working on a QRIS goal by the participant and the coach	Minutes

Source: Mathematica implementation study, 2009.

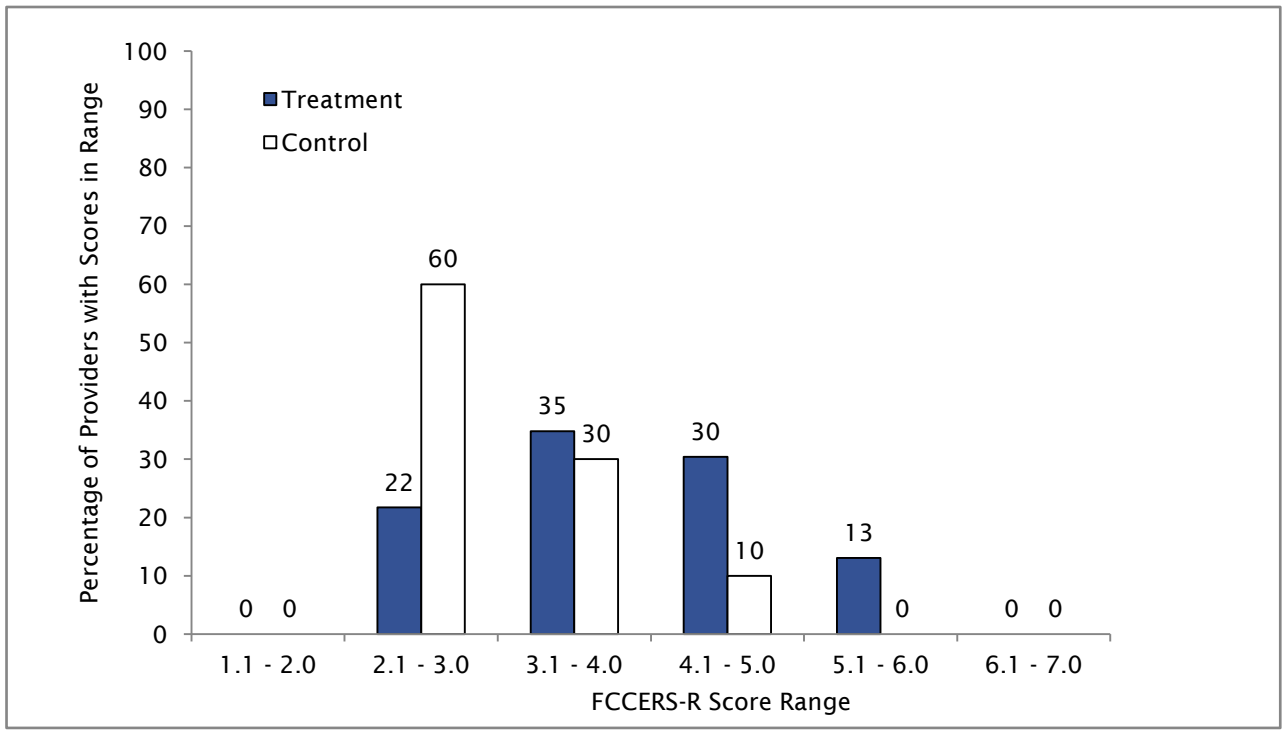
CCC = child care center; ETO = Efforts to Outcomes; FCC = family child care; PD = professional development; QRIS = Quality Rating and Improvement System.

APPENDIX B

DISTRIBUTION OF MEAN ERS AND ARNETT CIS SCORES

In these figures, we provide the distribution of mean ERS and Arnett CIS scores at follow-up by treatment status for family child care providers and centers. These figures demonstrate that none of the family child care homes or centers was rated at the lowest quality levels on any of these measures. The distributions provide information that complements the findings in Chapter III regarding the overall mean scores of treatment and control groups on key measures of child care quality.

Figure B1. Distribution of FCCERS-R Total Scores at Follow-Up, by Treatment and Control for Family Child Care Providers

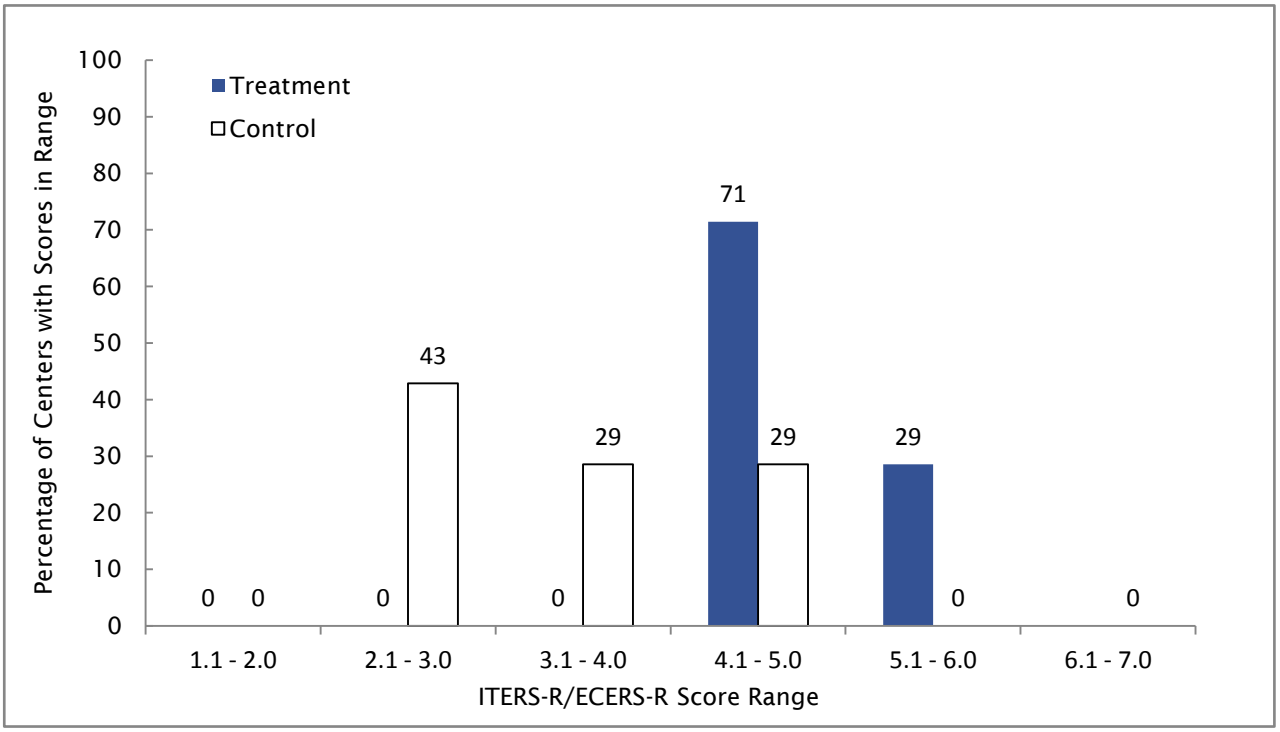


Source: Follow-Up Child Care Observation completed in winter 2010.

Note: Sample size = 43.

FCCERS-R = Family Child Care Environment Rating Scale-Revised.

Figure B2. Distribution of ITERS- R/ECERS- R Total Scores at Follow- Up, by Treatment and Control for Centers

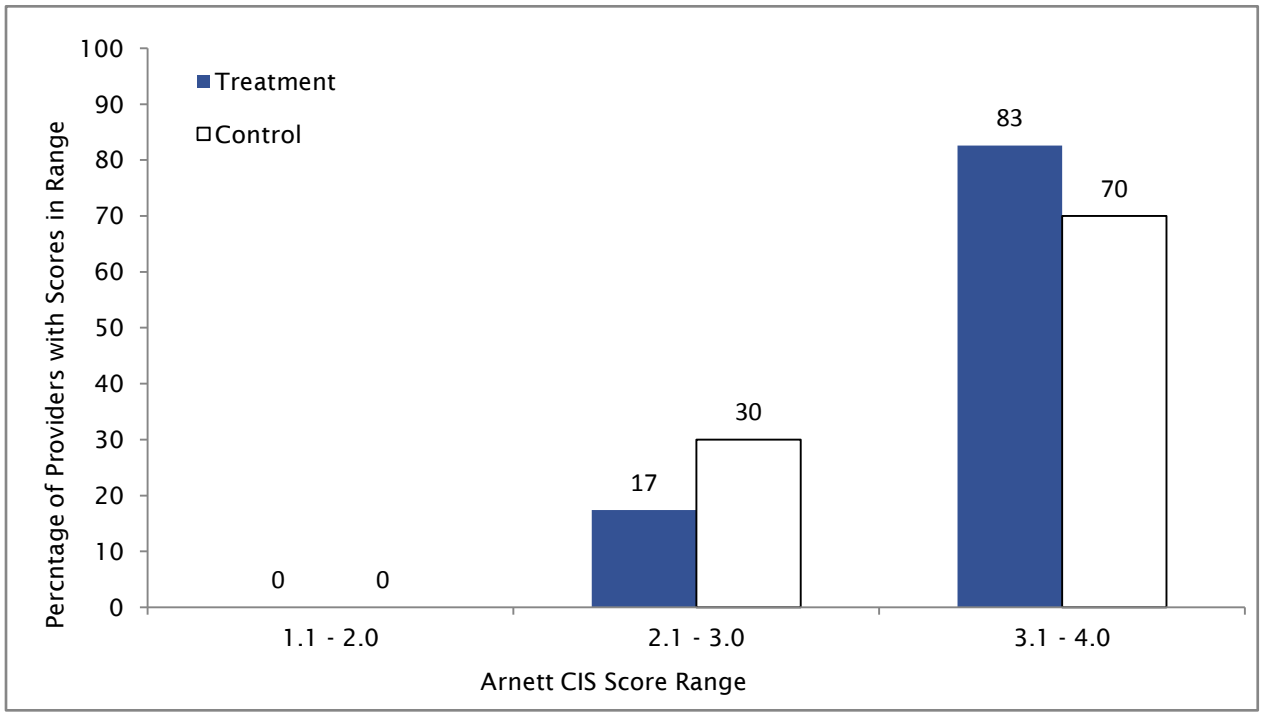


Source: Follow-Up Child Care Observation completed in winter 2010.

Note: Sample size = 14.

ECERS-R = Early Childhood Environment Rating Scale-Revised; ITERS-R = Infant/Toddler Environment Rating Scale-Revised.

Figure B3. Distribution of Arnett CIS Total Scores at Follow- Up, by Treatment and Control for Family Child Care Providers

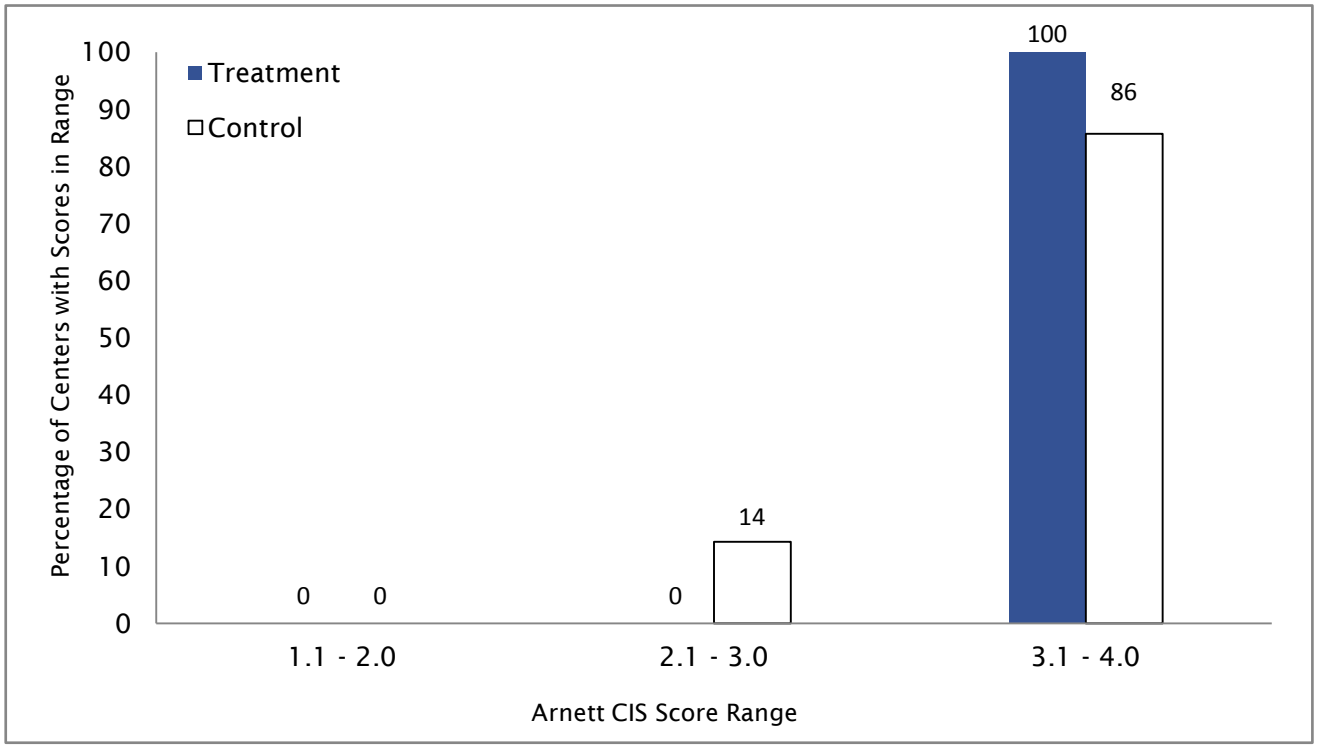


Source: Follow-Up Child Care Observation completed in winter 2010.

Note: Sample size = 43.

CIS = Caregiver Interaction Scale.

Figure B4. Distribution of Arnett CIS Total Scores at Follow- Up, by Treatment and Control for Centers



Source: Follow-Up Child Care Observation completed in winter 2010.

Note: Sample size = 14.

CIS = Caregiver Interaction Scale.

www.mathematica-mpr.com

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

Princeton, NJ ■ Ann Arbor, MI ■ Cambridge, MA ■ Chicago, IL ■ Oakland, CA ■ Washington, DC