NCEE 2012-4003

U. S. DEPARTMENT OF EDUCATION

Evaluation of Program for Infant/Toddler Care (PITC): An On-Site Training of Caregivers

ies NATIONAL CENTER FOR EDUCATION EVALUATION AND REGIONAL ASSISTANCE
Institute of Education Sciences

Evaluation of Program for Infant/Toddler Care (PITC): An On-site Training of Caregivers

Final Report

March 2012

Authors:

Phyllis Weinstock
Berkeley Policy Associates

Johannes Bos
Berkeley Policy Associates

Fannie Tseng
Berkeley Policy Associates

Emily Rosenthal
Berkeley Policy Associates

Lorena Ortiz
Berkeley Policy Associates

Chantelle Dowsett
University of Texas, Austin

Aletha Huston
University of Texas, Austin

Alison Bentley
University of Texas, Austin

Project Officer:

Ok-Choon Park
Institute of Education Sciences

NCEE 2012-4003
U.S. Department of Education

U.S. Department of Education

Arne Duncan

Secretary

Institute of Education Sciences

John Q. Easton

Director

National Center for Education Evaluation and Regional Assistance

Rebecca A. Maynard

Commissioner

March 2012

This report was prepared for the National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, under contract ED-06CO-0014 with Regional Educational Laboratory West administered by WestEd.

IES evaluation reports present objective information on the conditions of implementation and impacts of the programs being evaluated. IES evaluation reports do not include conclusions or recommendations or views with regard to actions policymakers or practitioners should take in light of the findings in the report.

This report is available on the Institute of Education Sciences website at http://ncee.ed.gov and the Regional Educational Laboratory Program website at http://edlabs.ed.gov.

Alternate Formats Upon request, this report is available in alternate formats, such as Braille, large print, audiotape, or computer diskette. For more information, please contact the Department's Alternate Format Center at 202-260-9895 or 202-205-8113.

Disclosure of potential conflict of interest

Regional Educational Laboratory West, housed at WestEd, contracted with Berkeley Policy Associates to conduct a third-party evaluation of the Program for Infant/Toddler Care, a WestEd intervention. None of the authors or other staff involved in the study from Berkeley Policy Associates and its subcontractors, the University of Texas at Austin and Survey Research Management, or any members of the Technical Work Group for the study, has financial interests that could be affected by the content of this report. The evaluation was conducted independent of WestEd staff, who developed and implemented the Program for Infant/Toddler Care.

Contractors carrying out research and evaluation projects for IES frequently need to obtain expert advice and technical assistance from individuals and entities whose other professional work may not be entirely independent of or separable from the tasks they are carrying out for the IES contractor. Contractors endeavor not to put such individuals or entities in positions in which they could bias the analysis and reporting of results, and their potential conflicts of interest are disclosed.

Acknowledgments

This study and this report were made possible by a collaborative effort of many child care practitioners, families, and researchers. The authors would like to thank the following individuals for their time, dedication, and assistance.

We would like to thank the many child care providers, teachers, administrators, and parents who took time from their busy schedules to participate in this study. The study would not have been possible without their cooperation and support.

At WestEd, we thank the REL West research directors, Neal Finkelstein and Thomas Hanson, for their assistance and support at all stages of this study. Of the Program for Infant/Toddler Care staff, we thank Ron Lally, Peter Mangione, Janet Poole, Arlene Paxton, Kerry Kriener-Althen, Maria Niggle, Diane Fellows, and the many PITC coordinators who provided invaluable assistance to this study.

We appreciate the technical consultation and advice provided by the REL West Technical Working Group: Jamal Abedi, Lloyd Bond, Geoffrey Borman, Brian Flay, Tom Good, Corinne Herlihy, Joan Herman, Heather Hill, Roger Levine, Juliet Shaffer, and Jason Snipes.

A large team of assessors and analysts supported data collection and analysis for this study. We offer special thanks to Linda Kuhn, Tony Lavender, and Gretchen Stage at Survey Research Management Boulder; to Kristina Lara, Dina De Veer, and Vincent Chan at Berkeley Policy Associates; and to Amy Imes at the University of Texas, Austin.

Contents

Figures

Tables

Abbreviations

BSRA: *Bracken School Readiness Assessment, Third Edition*

BSID: *Bayley Scales of Infant and Toddler Development, Third Edition,* Cognitive Subscale

CBCL 1½ -5: Child Behavior Checklist for children ages 1½-5

FCCERS-R: Family Child Care Environment Rating Scale-Revised

ITERS-R: Infant/Toddler Environment Rating Scale-Revised

ITSEA: Infant-Toddler Social and Emotional Assessment

PITC: Program for Infant/Toddler Care

PITC/PARS: Program for Infant/Toddler Care Program Assessment Rating System

PLS-4: *Preschool Language Scale, Fourth Edition,* Expressive Communication Subscale

Executive Summary

In recent decades, the quality of child care for children younger than three years old has been of concern nationally and within the West Region. Among mothers of these children, 59.4 percent were in the labor force as of March 2008 (U.S. Department of Labor 2009). The Early Childhood Longitudinal Study conducted earlier in the decade found that, of children younger than three with working mothers, 38 percent spent 35 hours or more in child care and 17 percent spent 15 to 34 hours in child care (Flanagan and West 2004). The National Institute of Child Health and Human Development (NICHD) Study of Early Child Care and Youth Development (2003, 2005) found that the quality of child care during a child's first three years was related to their school readiness, expressive language, and receptive language at age three. And 14 percent of child care centers and 12 percent of regulated family child care homes in California were rated good to excellent, based on the environment rating scales quality measures (Kontos et al. 1995; Helburn and Culkin 1995; Peisner-Feinberg 1999).

Advances in research on early brain development in the 1990s underscored the critical importance of children's early learning experiences and the potential benefits of effective early interventions. A White House Conference on Early Childhood Development highlighted this research in 1997 and, in 1999, the U.S. Department of Education created the National Center for Early Development and Learning to conduct further research on interventions in young children's "critical periods" of development. In 2003, the National Infant and Toddler Child Care Initiative, established by the U.S. Department of Health and Human Services, began offering technical assistance to states, territories, and tribes to help them improve the quality and supply of child care for their youngest children.

Child care improvement strategies have focused increasingly on the education, training, and professional development of early childhood teachers and child care providers. In 2001, the Committee on Early Childhood Pedagogy concluded that there is a serious mismatch between the preparation (and compensation) of the average early childhood professional and the growing expectations of parents and policy makers" (National Research Council, p. 261). Responding to this concern, the National Infant and Toddler Child Care Initiative has been helping states design professional development policies and systems that are informed by standards based on core competencies for infant/toddler caregivers (National Infant and Toddler Child Care Initiative 2010).

Little research has been conducted on the effectiveness of training strategies for child care providers. In a recent literature review (Zaslow, Tout, Halle, Whittaker, and Lavelle 2010), the authors commented, "There is a need to expand understanding of the strategies that are most effective for educators working with infants and toddlers." The report concluded that "the research on early childhood professional development is at an early stage" (p. xi).

Preliminary studies, mostly nonexperimental, have identified training strategies that warrant additional research, including the use of specific curricula, on-site consultation, high intensity and/or duration interventions, and focus on practice (Fukkink and Lont 2007; Zaslow et al. 2010). The Child Care Bureau of the U.S. Department of Health and Human Services, in the recent Quality Intervention for Early Care and Education evaluation, sponsored randomized controlled trials measuring the impact of on-site consultation models of caregiver training. One

study found positive effects of an assessment-based consultation model on quality of care in child care centers but not in family child care homes; no significant child effects were found (Bryant et al. 2009). An impact study of Seeds for Success, a child care quality improvement program in Washington state that includes coaching and quality improvement grants, found positive impacts on child care quality in centers and family child care homes (Boller et al. 2010). This study did not measure child effects.

The current study used an experimental intent-to-treat design to measure the impact of an established intervention, the on-site caregiver training component of the Program for Infant/Toddler Care (PITC), on child development and child care program quality. The PITC was developed by WestEd in 1985, in partnership with the California Department of Education. Over the next 25 years, more than 1,500 early childhood trainers across 30 states became PITC-certified trainers. More than 1,000 Early Head Start trainers have also been trained by the PITC. Regional Educational Laboratory West, administered by WestEd, contracted with Berkeley Policy Associates to conduct a third-party evaluation of the PITC. Berkeley Policy Associates and its subcontractors, the University of Texas, Austin, and Survey Research Management, conducted the evaluation independently of the WestEd staff who developed and implemented the intervention.

PITC is informed by brain development research that emphasizes early relationships as the foundation for healthy child development. Its approach incorporates six essential policies: primary care (assignment of a primary caregiver to each child), small groups, continuity of care, individualized schedules and routines, inclusion, and cultural sensitivity. The training imparts information on infant-toddler development. It encourages practices that facilitate healthy development and sensitivity to children's home communities, cultures, and languages. The training includes program policy recommendations and addresses program operation and environmental arrangements. The PITC training has been delivered in California for more than 20 years, with some modifications in format and additions of support components. Trainers work with child care programs to develop a schedule of 64 hours of training and 40 hours of technical assistance and support, to be delivered over a 10 to 18 months, with a minimum of 4 hours of training or technical assistance per month. The course is delivered to individual child care centers (with at least four staff and a director participating) and to small groups of family child care providers. More limited versions of the PITC, often as between one and four half-day or full-day workshops, have been offered in other states.

The intervention under study combines direct caregiver training and on-site coaching or other tailored assistance. For center-based child care programs, the intervention is delivered on-site. Family child care providers participate in groups of 5 to 10 programs, coming together for training sessions in a provider's home or in a convenient community center or school. The trainer also visits individual family child care homes for on-site technical assistance and support. Technical assistance includes observations, director meetings, and "reflective action planning," group meetings in which staff reflect on progress and plan further improvements. Individual participants can earn up to $350 in "professional growth incentives" by participating in at least 56 hours of training and 8 hours of reflective action planning.

This study is the first rigorous effectiveness trial of the on-site caregiver training component of PITC. It was implemented over 2007–2010 in six Southern California counties and four Arizona counties. The study sample of 251 child care programs included 92 child care centers and

159 licensed family child care homes, and the sample of 936 children included an average of eight children per center and between one and two children per family child care home. Child care programs were the unit of random assignment. Programs were enrolled in the study subject to consent to participate by program directors and a minimum number of staff who worked with children under the age of three (at least four staff in child care centers and one in family child care homes). In addition, parental consent was needed for children's participation in child-level data collection activities. Children were eligible to participate if younger than 27 months. Parents of at least five children in each child care center and parents of at least one child in each family child care home would have to consent to the study in order for the program to enroll. Children whose parents did not consent were not part of the child-level data collection activities and children could not be added to the study after their programs' random assignment. Data were collected on programs and children at baseline (before random assignment) and in two follow-up waves. The same data were collected on treatment and control group members, and all members of the original sample were contacted for follow-up regardless of whether children remained in their original child care settings and whether programs remained in PITC training. At each wave of data collection, there were no statistically significant differences in response rates between treatment and control groups of centers, family child care homes, or children.

The primary questions focus on child outcomes:

- What is the impact of the PITC on a composite measure of children's cognitive and language skills, at least 6 months after its full delivery to the children's child care programs (within an average of 23 months after random assignment)?
- What is the impact of the PITC on a composite measure of children's social and behavioral skills, at least 6 months after its full delivery to the children's child care programs (within an average of 23 months after random assignment)?

The secondary questions focus on child care quality:

- What is the impact of the PITC on global child care quality at least 4 months after the PITC ends (within an average of 21 months after random assignment)?
- What is the impact of the PITC on a composite measure of the quality of child care programs' staff-child interactions at least 4 months after the PITC ends (within an average of 21 months after random assignment)?

The design of the impact evaluation combined with the duration of the professional development program meant that children participating in the study would be exposed to fully trained child care workers only if the children remained in child care sites for a minimum of 15 months, on average. Children enrolled in participating child care sites were recruited before providers were randomly assigned to treatment and control groups. PITC implementation in treatment sites began, on average, two months after random assignment, and lasted for an average of 13 to 14 months. Of children in the treatment group, 60.0 percent remained in their original program for 15 months or more.

The confirmatory research questions were addressed using hierarchical linear regression models to account for the effect of clustering observations within programs. Each impact analysis included covariates collected before random assignment to improve the estimates' statistical

precision and reduce the likelihood that random sampling variation would affect the impact estimates.

To reduce the number of statistical comparisons in the study, researchers used or created composite measures of child and program outcomes. For the primary child measures, a cognitive/language composite was formed by averaging the *z*-scores from the *Bracken School Readiness Assessment, Third Edition* (BSRA; Bracken 2007) and the *Preschool Language Scale, Fourth Edition*, Expressive Communication Subscale (PLS-4; Zimmerman, Steiner, and Pond 2002). The BSRA measures pre-academic skills and includes six subtests: colors, letters, numbers/counting, sizes, comparisons, and shapes. The PLS-4 measures expressive communication skills for children from birth through 6 years, 11 months of age. A composite child socioemotional/behavior measure was formed using parent ratings on The Child Behavior Checklist (CBCL 1½-5; Achenbach and Rescorla 2000) and The Positive Behavior Scale (Polit 1996). The CBCL 1½-5 has subscales that measure internalizing problems (emotionally reactive, anxious/depressed, somatic complaints, and withdrawn behaviors) and externalizing problems (attention problems and aggressive behavior). The Positive Behavior Scale has three subscales: compliance/self-control, social competence and sensitivity, and autonomy.

Composites were also used to measure program outcomes. One composite is a global quality measure of the environment rating scales, which includes comparable items of the Infant/Toddler Environment Rating Scale-Revised (ITERS-R; Harms, Clifford, and Cryer 2003) or the Family Child Care Environment Rating Scale-Revised (FCCERS-R; Harms, Cryer, and Clifford 2007). These are widely used observational measures of child care quality. A second composite child care quality measure, focusing on staff-child interactions, was constructed using items from both the ITERS-R/FCCERS-R and the PITC Program Assessment Rating System (PITC-PARS). The PITC-PARS is an observational measure of child care quality designed by PITC staff in accordance with how the PITC measures the quality of care children receive from birth through age three in home-based and center-based settings. For the construction of the staff-child interactions composite, four selected items from the Quality of Caregiver Interaction Subscale of the PITC-PARS were used: facilitation of cognitive development, responsiveness and sensitivity, positive tone and attentiveness, and responsive engagement and intervention.

Other PITC-PARS subscales measuring critical PITC policies were used in the implementation analysis. These subscales measure culturally responsive caregiving, primary caregiving, continuity of care, and group size and ratios.

Other measures for the study were incorporated into questionnaires administered at baseline and at the 15-month follow-up to both treatment and control child care staff. These questionnaires included caregiver knowledge and skills measures and questions about program operations and enrollment, program goals, caregivers' educational and professional backgrounds, and professional development received over the prior 12 months.

The primary findings are:

- The PITC did not have a statistically significant effect on a composite measure of children's cognitive/language scores, measured approximately 6 months (on average) after it ended.

- The PITC did not have a statistically significant effect on children's composite behavior scores, measured at 6 months after it ended. Sensitivity analyses, conducted with two alternative approaches to missing data treatment, had results consistent with these findings.

Secondary research questions addressed the effects of the PITC on child care program quality at, on average, four months after the intervention ended. These estimates also found no significant effects. Findings of this analysis are:

- The PITC did not have a statistically significant effect on global program quality, as measured by trained observers administering the ITERS-R and the FCCERS-R.
- The PITC did not have a statistically significant effect on staff-child interactions, a composite measure incorporating interactions items from the environment rating scales and from the PITC-PARS. Results of sensitivity analyses were consistent these findings.

Analysis of implementation found that, in many child care programs, the intervention was not fully implemented or was not implemented with full participation: Of the 124 child care programs assigned to the treatment group, 11 decided not to participate before receiving any training, and 6 dropped midcourse. In only 59.4 percent of participating family child care homes did at least one caregiver receive the benchmark 56 hours of training, and in 41.9 percent of child care centers, four or more caregivers (the minimum number of participants, plus the director, required for PITC delivery) received at least 56 hours of training. Of children in the treatment sample, 17 percent received no exposure to the PITC, either because they left their original child care programs before start-up or because their programs were among the 11 that declined the intervention after random assignment.

This effectiveness trial was conducted in community child care settings. However, the study was conducted in specific areas of California and Arizona and during an economic recession. Its generalizability is limited by this context and by other features of the study including:

- This study tested a specific implementation model of the PITC, with delivery of 64 hours of training and 40 hours of on-site coaching and support, requiring an average of 14 months for full implementation. The findings should not be generalized to other models of PITC implementation that have different durations and service combinations.
- The study relied on volunteer samples, within the identified geographic regions, of child care providers and families willing to participate in a random assignment study. Recruitment required contacts with many providers who refused to participate or who were unwilling or unable to obtain consent from the number of parents (or, in some centers, from staff) needed to meet the sample requirements. It is possible that providers and families who participated in the study were different than nonparticipants, and results should not be generalized to the larger population.

As an intent-to-treat study, this evaluation measured effects on all children who enrolled in the study and were randomly assigned, including those who left their child care settings well before the PITC was fully implemented. While this design maintained the study's internal validity, it also reduced the treatment-control contrast. Analysis of children's time periods in care found that 25.0 percent of treatment children either left their study programs before start-up, left within 6 months of start-up (implementation required 9–20 months), or attended treatment programs

that declined the intervention. These “treatment” children received minimal or no treatment. Conversely, only 49.5 percent of treatment children remained in their programs for 19 months or more, long enough to experience the potential PITC effects—based on the average study treatment period of 13–14 months—as hypothesized in the conceptual timeline for the study, which posits child effects at 6 months after the intervention.

The PITC incorporates a number of the features that preliminary research and expert opinion in the field suggest are most likely to have a positive effect: focus on relationships, on-site consultation, opportunities for assessment and feedback, and application to practice. However, this study finds no positive main effects and also underscores the difficulties of sustaining participation in an intensive, long-term intervention in a large number of community child care settings across geographically dispersed locations. More research on the PITC and other training interventions is needed for fuller examination of both implementation and impacts. Increased understanding of the “transfer” between training strategies, program quality, and child development would inform improved child care training design and implementation.

1. Introduction and Overview

This study examines the impacts of the Program for Infant/Toddler Care (PITC), a training intervention for child care providers serving children younger than three years old. The goals of the study are to estimate the PITC's effects on child development, particularly the language, cognitive, and social skills closely associated with school readiness; and the quality of child care in family child care homes and centers. The study employed a cluster-based random assignment design using a sample of 251 child care centers and family child care homes in Southern California and Arizona, with 936 children served by these programs.

The PITC was developed by WestEd in 1985 and has been delivered over the past few decades in California and (in more limited formats) in 30 other states. The current study is the first rigorous impact assessment of it and one of few experimental studies focusing on infant/toddler child care. The Regional Educational Laboratory West, housed at WestEd, contracted with Berkeley Policy Associates to conduct a third-party evaluation of the PITC. Berkeley Policy Associates and its subcontractors, the University of Texas, Austin, and Survey Research Management, conducted the evaluation independently of the WestEd staff who developed and implemented the intervention.

Need for the study

The importance and scarcity of quality infant/toddler child care

The quality of child care for very young children is a concern at the national level and within the Western Region. Among mothers of children under age three, 59.4 percent were in the labor force in March 2008 (U.S. Department of Labor 2009). Of children born in 2001, 15 percent were in nonrelative care (either in their own or another family's home) and 9 percent were in center-based care by the age of nine months (Flanagan and West 2004). In 2002, of children through the age of four with working mothers, 38 percent spent 35 hours per week or more in child care, and 17 percent spent 15 to 34 hours per week in child care (Overturf 2002). The most recent large-scale longitudinal study addressing the outcomes of infant/toddler child care, conducted by the NICHD Early Child Care Research Network (2003, 2005), found significant associations between observed quality of early child care—particularly for language interactions between caregivers and children—and children's cognitive, social, and language skills. The study found that the quality of child care during a child's first three years was related to their school readiness, expressive language, and receptive language at three years.

The quality of infant child care garnered much attention in the 1990s. The Carnegie Corporation (1994) highlighted an early childhood "quiet crisis," including the prevalence of infant child care of substandard quality. Advances in neuroscientific research brought greater awareness of what Bailey, Bruer, Symons, and Licktman (2001) refer to as the "windows of opportunity" in early child development. In 1995, the federally funded Head Start preschool program for low-income children and their parents was expanded to include Early Head Start, an infant/toddler component based in both centers and family child care homes. A 1997 White House Conference on Early Childhood Development Learning addressed the new research on brain development in very young children and the critical importance of early learning experiences. In 1999, the

National Center for Early Development and Learning, funded by the U.S. Department of Education, convened to further analyze "critical periods" in early development and discuss implications for effective early interventions (National Center for Early Development and Learning 1999). In 2003, the National Infant and Toddler Child Care Initiative, established by the U.S. Department of Health and Human Services, began offering technical assistance to states, territories, and tribes to help them improve the quality and supply of child care for the youngest children.

Child care quality initiatives have been introduced in California and Arizona in recent years. In California, 14 percent of child care centers and 12 percent of regulated family child care homes were rated good to excellent, based on the Infant/Toddler Environment Rating Scale (ITERS-R) measure of quality (Howes and Brown 2000; Kontos, Howes, Shinn, and Galinsky 1995; Helburn and Culkin 1995; Peisner-Feinberg et al. 1999). The National Association of Child Care Resource and Referral Agencies (2007) ranked California 23rd and Arizona 26th of the 50 states in child care standards and oversight of family child care. Both states, however, fund child care quality initiatives, including most recently, Arizona's child care quality rating system, Quality First. This system, piloted in 2010, includes coaching, quality improvement grants, tuition assistance, and quality assessment and rating. In California, voters in 1998 passed Proposition 10, identifying a special tobacco tax as a source of funding that supports services for children from birth through age five. A California First 5 commission and county-level First 5 commissions administer these funds. Some of the funds raised through this tax support early care and education and caregiver training, including the PITC. The California Department of Education adopted infant/toddler learning and development program guidelines in 2007 (California Department of Education 2007).

The need for research on training of infant/toddler caregivers

As concern about improving infant/toddler care has increased, caregiver training (professional development) has emerged as an effective way to achieve that improvement. A focus on this strategy for infants/toddlers has been aligned with a more general focus on professional development in the early childhood field. In 2000, the Board on Children, Youth and Families of the National Research Council and the Institute of Medicine created the Committee on Integrating the Science of Early Childhood Development to update scientific knowledge of early development and to discuss its implications for policy and practice, including professional development. Among the Committee's findings were that "there is a serious mismatch between the preparation (and compensation) of the average early childhood professional and the growing expectations of parents and policy makers" (National Research Council and Institute of Medicine 2000, p. 261). Responding to this concern, the National Infant and Toddler Child Care Initiative has been helping states design professional development policies and systems that are informed by standards based on core competencies for infant/toddler caregivers (National Infant and Toddler Child Care Initiative 2010).

Preliminary studies, mostly nonexperimental, have identified training strategies for child care providers that warrant more research. These strategies include the use of specific curricula, on-site consultation, high intensity and/or duration interventions, and focus on practice (Fukkink and Lont 2007; Zaslow et al. 2010). A 2005 literature review by Child Care and Early Education Research Connections, a project of the U.S. Department of Health and Human Services, noted the diversity of approaches to training, technical assistance, and mentoring of infant-toddler

caregivers, the fairly limited knowledge on the effectiveness of these approaches, and the need to learn more about their strengths and weaknesses (Kreader, Ferguson, and Lawrence 2005). Zaslow et al. (2010) states that "the literature emphasizes professional development for educators working with preschool-age children: most of the studies covered in this review focused on children in the year or two years before entry into kindergarten. There is a need to expand understanding of the strategies that are most effective for educators working with infants and toddlers" (p. 87) and that "the research on early childhood professional development is at an early stage." (p. xi)

Recent research on caregiver professional development has focused on coaching and consultation models. The Child Care Bureau of the U.S. Department of Health and Human Services, in the recent Quality Interventions for Early Care and Education evaluation, sponsored randomized controlled trials measuring the impact of on-site consultation models of caregiver training (Quality Interventions for Early Care and Education 2009). Final results of an impact study of infant caregiver training by Ramey and Ramey (2008) are forthcoming. Bryant et al. (2009), focusing on child care providers serving three- to four-year-olds tested a "partnerships for children" model of assessment-based, individualized, on-site consultation. The study found significant effects on the quality of teaching and interactions in family child care homes (effect size = .33) but not on quality in centers. Child-level effects, measured in school readiness, language/comprehension, and socioemotional domains, were not significant.

An impact study of a child care quality improvement program in Washington state, Seeds for Success, found positive impacts on child care quality (Boller et al. 2010). Centers and family child care homes, serving a mix of infant-toddlers and preschool-aged children, were randomly assigned to a treatment group that received coaching, quality improvement grants, and funds for professional development and supports; or to a control group that received only professional development opportunities and supports. After six months, positive effects of treatment were statistically significant for family child care providers and for child care centers. The study did not measure child effects.

An earlier quasi-experimental study by Campbell and Milbourne (2005) focused on a training program for infant/toddler caregivers working with low-income children. Program global quality and staff-child interactions were compared before and after a 15-hour training course for a group that received three hours of on-site consultation and a group that received no consultation. Consultation entailed three one-hour sessions that followed a protocol including assessment, planning, and follow-up assessment. No significant effects were reported.

PITC development and design

WestEd developed the PITC in 1985, in partnership with the California Department of Education. Over the next 25 years, more than 1,500 early childhood trainers across 30 states became PITC-certified trainers; in turn, they have trained over 6,000 caregivers in the PITC approach, with training formats ranging from one-day institutes to the full 10–18 month program. The PITC, further developed and fielded since its inauguration, has built strong connections with early childhood stakeholders in the West Region, including Early Head Start, for which it has trained over 1,200 trainers.

The PITC, as with other early childhood professional development models, combines training and on-site consultation or support. It is informed by the neuroscientific research that emphasizes early relationships as the foundation for children's development. *From Neurons to Neighborhoods* noted that "the child's relationship with his or her child care provider seems to play an especially important role with regard to social-emotional development. Children form secure attachments to their child care providers when they are stable and these attachments, in turn, are associated with adaptive social development, just as they are for children and parents" (Shonkoff and Phillips 2000, p. 313).

The PITC promotes relationship-based care responsive to children's needs at each stage of development. The training imparts information on infant/toddler development. It encourages practices that facilitate healthy development and sensitivity to children's home communities, cultures, and languages. The training includes program policy recommendations and addresses program operation and environmental arrangements. The cornerstone of the PITC approach is a primary relationship between each child and a caregiver. It incorporates six essential policies: primary care (assignment of a primary caregiver to each child), small groups, continuity of care, individualized schedules and routines, inclusion, and cultural sensitivity. These policies align closely with eight key aspects of high-quality infant care, identified by the 1998 National Leadership Forum on Quality Care for Infants and Toddlers: health and safety, small groups of three to four infants per caregiver, assigning each baby to a primary caregiver, ensuring continuity of care with the same provider over time, caregiver responsiveness to infant signals, meeting each infant's needs in group care with a focus on individual learning style and temperament, cultural and linguistic sensitivity, and provision of a physical environment with variety, stimulation, and planned activities (Fenichel, Lurie-Hurvitz, and Griffin 1999).

The PITC has been delivered in California for more than 20 years, with some modifications in format and additions of support components. The course is delivered to individual child care centers (with at least five staff[1] participating) and to small groups of family child care providers. The PITC lasts 10–18 months. Trainers work with providers to customize a training schedule that delivers the full curriculum and accommodates staff schedules and other commitments as much as possible. However, at least four hours of workshops or consultation meetings per month is required.

More limited versions of the PITC, often in the form of between one and four half- or full-day workshops, have been offered in other states. In Arizona, four one-day workshops, based on the PITC curriculum have been offered through Child and Family Resources, a child care resource and referral network, at two locations since 1996. Participants may enroll for one workshop or for all four.

The intervention under study combines direct caregiver training and on-site coaching or other tailored assistance. Because few programs can offer staff release time for training, most of the training is in the evening or on weekends. For child care centers, the intervention is delivered on-site. Family child care providers participate in groups of 5 to 10 providers, coming together for

[1] Throughout this report, the terms "staff" and "teachers," as well as "caregivers," may be used when referring to the individuals who care for children in child care settings and are eligible for training from the PITC. Caregivers in family child care homes may also be referred to as "providers." These terms are commonly used in the child care field and are useful in distinguishing caregivers in child care settings from parents and other familial caregivers.

training sessions in a provider's home or in a convenient community center or school. The trainer also visits individual family child care homes for on-site coaching and support.

The intervention is delivered by PITC-certified trainers, experienced practitioners with bachelor's degrees[2] who have completed a trainer certification process. To be certified, trainers must participate in two seven-day trainer institutes and write a 28-page paper presenting a detailed lesson plan for each curriculum topic. In addition to delivering the necessary content, a trainer's lesson plans must demonstrate knowledge of adult learning strategies. To that end, the lesson plans must incorporate multiple instructional strategies, including audiovisual, text, action-based, individual, and group approaches.

The PITC under study delivers 64 hours of training and 40 hours of coaching or other support to each participating child care center or group. The training curriculum is divided into section I (module 1: social emotional growth and socialization and module 2: group care; and section II (module 3: learning and development and module 4: culture, family, and providers. Although participants have some choice of topics and topic sequence within modules, the modules themselves are always delivered in this order. Throughout the course, participants periodically meet with the trainer in a "reflective action planning" process that involves setting goals, reviewing program progress, and revisiting goals.

Child care center staff and family child care providers must complete at least 28 hours of training per section and 4 hours of reflective action planning per section to earn "professional growth incentives," provided in cash, academic credit, or resource grants. Participating in 56 hours of training and 8 hours of reflection action planning is a benchmark for completion and earns $350 in cash incentives. Child care center directors must complete 16 hours of training (8 hours per section), complete 8 hours of reflective action planning, and participate in an initial and concluding meeting with the trainer to earn incentives.

Trainers are taught to tailor their own teaching strategies to the learning styles, preferences, needs, and cultures of the caregivers with whom they work. In assigning trainers, the PITC attempts to match each trainer with child care programs whose caregivers have the same language and cultural background.

Nonexperimental research conducted by WestEd in California statewide and in San Diego County found the PITC to be associated with statistically significant increases in program quality, as measured four months after completion of the intervention. Overall scores on the Infant/Toddler Environment Rating Scale increased from a mean of 4.0 preprogram to 4.38 postprogram ($p < .05$) on a scale of 1 (inadequate) to 7 (excellent) and on the Early Childhood Environment Rating Scale from a preprogram mean of 4.0 to a postprogram mean of 4.48 ($p < .01$; WestEd Child and Family Services 2003). A score of 5 or more is considered "good" on these instruments.

This study is the first random-assignment evaluation of the on-site caregiver training component of PITC and the first to measure its impacts on children.

[2] In special cases, as an alternative to a bachelor's degree, trainers may have an associate's degree plus six years of relevant experience.

Conceptual model

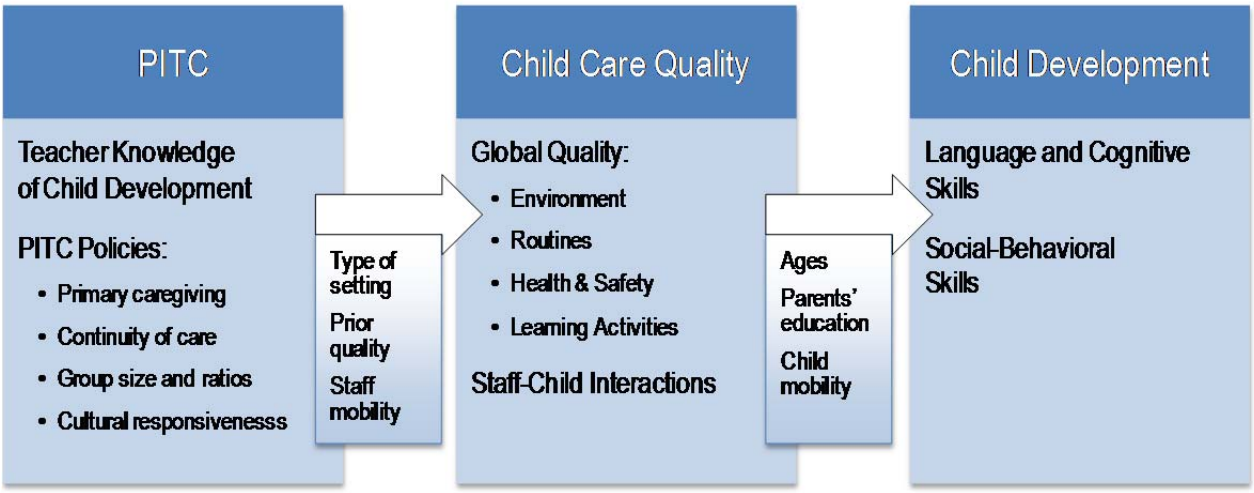
The PITC is informed by a causal model with several stages: caregivers trained in program practices that facilitate children's learning and development will improve in these domains of child care practice; this improved child care, in turn, will foster positive outcomes for children, including school readiness before school entry. The approach to the care and education of children from birth to three years emphasizes two areas key to children's school readiness: cognitive and language development and socioemotional development.

The conceptual model for this study assumes that the effects on children of the PITC caregiver training are the result of improvements in critical dimensions of child care quality (figure 1.1). The PITC delivers curricula on infant development and encourages providers to adopt policies, such as assignment of primary caregivers, that support this development. Caregiver training influences child development via the environment, activities and interactions provided by better trained caregivers. In addition, policies encouraged by the PITC, such as smaller group size and primary caregiving arrangements, foster the circumstances in which these high-quality caregiver-child interactions are more likely to occur. More responsive caregiving and stimulation, in turn, are expected to lead to improved language, cognitive, and social skills for the children.

In practice, these skills overlap as children develop, so individuals caring for infants and toddlers can engage in a single activity and improve both language and cognitive skills. For example, the caregiver might give gentle encouragement so that the infant can explore how to make a toy work and use new descriptive vocabulary words while the infant plays. Similarly, for social development, the caregiver might encourage turn-taking and cooperation while facilitating dramatic play with two or more toddlers and help them learn to regulate their emotions when conflicts arise.

Effects on caregiver quality and, in turn, on child development are also influenced by the characteristics of the child care program (type of setting and prior quality) and the children it serves (ages of children, parents' education; see figure 1.1). Mobility, which affects the duration of children's exposure to particular child care programs, is another factor that moderates the impact of the intervention on children's development.

Figure 1.1. Conceptual model



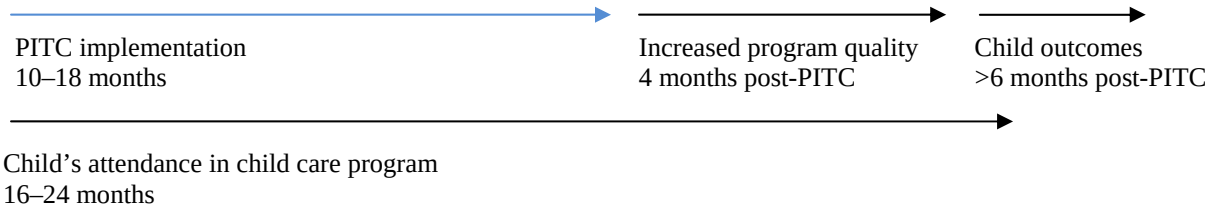
Source: PITC program materials.

Temporal alignment of evaluation design and conceptual model

This study's hypotheses posit that primary outcome measurements are child measurements administered, on average, six months after a program's completion of the PITC (figure 1.2).

As noted, the PITC is designed to be implemented over 10–18 months. Based on the experiences of PITC developers, caregivers need additional time (at least 4 months) after completing the training to fully incorporate new practices and policies. More time is required for children to be potentially affected by PITC-induced changes in care. These expected timeframes for change to occur are in figure 1.2 and are reflected in the study hypotheses and data collection timing.

Figure 1.2. Temporal model



Source: PITC program materials.

Research questions and hypotheses

The confirmatory questions fall in two domains: child development and child care quality. The primary questions focus on child development outcomes:

- What is the impact of the PITC on a composite measure of children's cognitive and language skills, at least 6 months after its full delivery to child care programs (within an average of 23 months after random assignment)?

- What is the impact of the PITC on a composite measure of children's social and behavioral skills, at least 6 months after its full delivery to child care programs (within an average of 23 months after random assignment)?

The secondary questions focus on child care quality:

- What is the impact of the PITC on global child care quality at least 4 months after the intervention ends (within an average of 21 months after random assignment)?
- What is the impact of the PITC on a composite measure of the quality of child care programs' staff-child interactions at least 4 months after the intervention ends (within an average of 21 months after random assignment)?

These four research questions result in tests of four confirmatory hypotheses:

- The PITC increases children's cognitive and language skills 6 months after its full delivery.
- The PITC increases children's social and behavioral skills 6 months after its full delivery.
- The PITC increases global child care quality 4 months after the intervention ends.
- The PITC increases quality of child care staff-child interactions 4 months after the intervention ends.

The analysis tests the hypotheses at the child and program levels. The intervention would be found to have a positive impact on child development if either the first or second primary research question demonstrated a statistically significant positive treatment effect. The intervention would be found to have a positive impact on program quality if either the first or second secondary research question demonstrated a statistically significant positive treatment effect.

Exploratory analyses support the interpretation of the confirmatory impact analyses by describing potential mediating and moderating factors. Prior research consists primarily of experimental tests of high-quality demonstration programs and observational studies of community-based child care that represents a range of quality and settings, but there is little research using random assignment on a sample of community-based child care settings. Hence, exploratory analyses can be particularly valuable for understanding the effects (or lack thereof) of the PITC, guiding future research and informing its targeting of the program to specific settings or children.

Intermediate measures of child care quality and child development were collected when, on average, three-fourths of the full curriculum (typically, the first three of the four curriculum modules) had been completed, and children in the sample were, on average, two and one-half years old. Changes in child care quality and children's skills at this time might mediate later development. These measures are also more proximal to the intervention for children who left the original settings before the full training curriculum was delivered.

Although this report does not include mediational analyses, it does include exploratory analyses of impacts on these proximal outcomes and on subgroups. Exploratory research questions for intermediate impacts include:

- What is the impact of the PITC on global child care quality one year after start-up, at a point when, on average, three-fourths of the intervention has been delivered?
- What is the impact of the PITC on a composite measure of the quality of child care programs' staff-child interactions one year after start-up, at a point when, on average, three-fourths of the intervention has been delivered?
- What is the impact of the PITC on a composite measure of children's cognitive and language skills one year after start-up, at a point when, on average, three-fourths of the intervention has been delivered?
- What is the impact of the PITC on a composite measure of children's social and behavioral skills one year after start-up, at a point when, on average, three-fourths of the intervention has been delivered?

Exploratory questions about variations in subgroup effects address moderators of the PITC's effects. These analyses also contribute to the policy relevance of the study findings. Questions addressed by subgroup impact analyses include:

- How do PITC impacts on children vary across subgroups by type of child care setting (family child care home or child care center), child age (younger than 18 months or 18 months or older), and parents' education level?
- How do impacts on programs vary across subgroups by child care setting and baseline quality?

Given the high levels of mobility and turnover in child care settings, additional exploratory research questions were examined:

- Does the PITC have an impact on child turnover?
- Does the impact of the PITC vary across children with different levels of treatment exposure, as measured by the child's period of attendance in the child care setting?

These exploratory analyses inform the interpretation of the primary impact analysis results and provide the field with potentially useful information on the design of child care quality interventions in ordinary community settings.

Structure of the report

This report is organized in six chapters. Chapter 2 describes the study methods, including detailed information on the sample, data collection, measures, and analytic models. Chapter 3 discusses implementation in detail, as the fidelity and quality of the experimental treatment is crucial in any program evaluation. Two aspects of implementation are important: the extent to which training was completed as prescribed and the extent to which PITC principles were implemented in the settings experienced by the children. Chapter 4 presents the results of the confirmatory analyses testing the primary and secondary hypotheses, and chapter 5 presents the results of the exploratory analyses. Chapter 6 summarizes the key findings and their implications.

2. Study Design and Methodology

This chapter overviews the study design and describes the sample recruitment, random assignment, sample flow and attrition, measures and measurement timeline, data collection procedures and quality oversight, response rates, and sample characteristics, including treatment-control equivalence of the original and final samples.

This study uses cluster-based random assignment to measure its impacts on program quality and on children's development (table 2.1). Child care programs were randomly assigned to treatment and control groups. Treatment group programs were offered the full on-site caregiver training component of Program for Infant/Toddler Care (PITC) intervention. Control group programs were excluded from the intervention but could participate in other professional development programs as usual. Children, nested within programs, were included in the study subject to parental consent. Child care quality outcomes, as well as child language, cognitive, and behavioral outcomes, were measured in two follow-up waves.

Table 2.1. Key study features

Feature	Description			
Study design	Cluster-randomized trial			
Unit of assignment	Child care programs (child care centers and family child care homes)			
Minimum detectable effect sizes	For Type 1, error=0.05, 80 percent or higher power for a minimum detectable effect size of 0.20–0.24 at the child level and 0.37 at the program level, assuming a child intraclass correlation of either 0.10 or 0.20 and R^2 values of 0.30.			
Implementation began	Rolling dates, starting in January 2008			
Measurement timeline				
	10/07–7/08 baseline	1/09–10/09 15 months after random assignment	12/09–5/10 22 months after random assignment	9/09–6/10 23 months after random assignment
Baseline child and program characteristics (covariates)	✓			
Program assessments	✓	✓	✓	
Child assessments		✓		✓

Note: Details of the power calculation, conducted with Optimal Design software (Raudenbush and Liu, 2000), are in table A1 in appendix A. Demographic and background characteristics of children and providers were collected at baseline. Because of the children's ages, brief child care provider assessments of children's language and behavior were collected at baseline; direct child assessments were administered only at follow-up.

The original sample targets were set at 240 child care programs (90 centers and 150 family child care homes) and 1,170 children, allocated equally to treatment and control groups. With an assumption of 10 percent attrition for programs and 20 percent attrition for children, this sample would have a minimum detectable effect size of between 0.20 and 0.23 for children and 0.38 for programs. The effect size for children was deemed reasonable given that an impact study of Early Head Start (U.S. Department of Health and Human Services, Administration for Children and Families 2002) found child effects between 0.1 and 0.2 and the Head Start impact study found child effects between 0.2 and 0.3 (U.S. Department of Health and Human Services 2005).

Program effects are more direct and needed to be larger to transfer to children. The study enrolled a sample of 251 child care programs (including over 300 classrooms were observed) and 1,009 children.

Enrollment occurred on a rolling basis over October 2007–July 2008. Locations of sample recruitment were within six Southern California counties and four Arizona counties. Child care centers, family child care homes with a primary language of English, and family child care homes with a primary language of Spanish were recruited separately, though impact analyses were conducted jointly across program types. Location strata for centers were in Los Angeles, Orange, and Riverside counties in California and in the cities of Tucson and Phoenix and the county of Cochise in Arizona. Location strata for English-speaking family child care homes were local areas in Los Angeles, Orange, and Riverside counties in California and in the cities of Phoenix and Tucson in Arizona. Spanish-speaking family child care homes were also in these locations, as well as in San Diego and Imperial counties in California and Santa Cruz County in Arizona.

After collecting baseline program-level data, researchers randomized providers to condition, blocked by program type and location. Follow-up program- and child-outcome data were collected in two waves. The first was at 14–15 months after random assignment (approximately 12 months after the intervention began). For programs, the second wave was at an average of approximately 22 months after random assignment; for children; it was approximately 23 months after random assignment.

To participate in the study, child care center directors and a minimum of four infant/toddler caregivers per center made a signed commitment to participate in the training if assigned to the treatment group. Family child care providers agreed to participate in training groups, each including 5–10 providers. Each program assigned to the treatment group, including the minimum number of staff specified above, agreed to a specific training plan and a schedule that implemented the full curriculum over 10–18 months.

The counterfactual was “business as usual.” Providers assigned to the control group were not offered participation in the PITC or access to trainers and other PITC resources. In California, trainers flagged all control-group participants in the PITC database to ensure that these providers were excluded until study completion. In Arizona, the PITC was not available outside the study other than in a very limited form, such as a one-day institute. However, providers assigned to the control group could access other services available to child care providers in their states. These services might include those provided by state departments of education, by local institutions of higher education, and by federally funded child care improvement programs administered by state departments of human services. See chapter 3 for details on training participation by both treatment and control programs and staff.

Sample recruitment

The recruitment and selection of providers was guided, first, by the scope of the Regional Educational Laboratory West contract, with its focus on West Region states.

Second, to estimate the net impact of the PITC, it was important that it not significantly penetrate the areas from which the sample was drawn. This excluded certain areas in California already

being served by the PITC. In these areas, providers either would already have been exposed to the PITC or would have previously declined to participate. Neither of these types of provider would make good candidates for an experimental test of the program.

Third, it was essential to choose locations in which certified trainers were available or could be deployed. The areas of California and Arizona from which the sample was drawn met the necessary criteria. These are areas in which the PITC had not been implemented widely, yet they have large numbers of both child care centers and family child care homes. Further, many of these programs serve predominantly Spanish-speaking children whose early education outcomes tend to lag behind those of their English-speaking counterparts (Abedi and Dietel 2004).

The study targeted children who received child care from a participating provider for at least 20 hours per week and were younger than 27 months at the time of enrollment. Older children were excluded because they were likely to transition out of infant/toddler care shortly after random assignment and, therefore, would receive little or no exposure to the intervention.

Programs were recruited through coordinated activities of researchers and PITC staff recruiters. Researchers developed recruitment materials and co-led outreach meetings with PITC staff recruiters. Researchers trained recruiters to explain the study and to administer the brief screening interview. (See appendix B for screening instrument.) The recruiters sometimes worked with provider associations and resource and referral agencies in their areas. These agencies helped recruiters organize outreach meetings and mail study information to providers in their networks.

The steps in sample selection were:

1. PITC staff in California and Arizona obtained lists of licensed child care centers and family child care homes in their designated geographic areas, which were available from state licensing agencies or provider networks. An initial pool of potential study programs was formed from these lists. The pool included all programs identified as serving infants/toddlers, or for which ages served were not identified, and that had not participated in the PITC (those flagged in the PITC database). Recruiters mailed recruitment letters and followed up with phone calls.

2. A total of 570 programs agreed to participate in a screening interview. During the screening, a recruiter collected information on program stability, number of staff members, staff and child turnover, program size, numbers of infants and toddlers, the primary language of staff and children, and current or past participation in the PITC. Programs were eligible if the primary language spoken in the program was English or Spanish, the program had not participated in the intensive form of the PITC, and, for centers, the program served at least five children younger than 24 months or, for family child care homes, the program served at least two children younger than 24 months.

3. Programs that met the screening criteria and were interested in the study were invited to attend a meeting, individually or in groups, to further discuss the study with researchers and recruiters. The agenda for these meetings included a study overview, a PITC overview, and instructions on how to complete forms needed to enroll in the study. Providers were given packets of staff consent forms, parent consent forms, and

information sheets to bring back to their program, along with postage-paid return envelopes.

4. If consent packets were not returned within two weeks, recruiters followed up by phone and then with in-person visits.

Family child care providers were offered a $15 gift card and child care centers were offered a $15 gift card per classroom (for up to two infant toddler classrooms) for collecting the minimum number of forms needed for enrollment in the study. In addition, each enrolled caregiver, as well as the director, was offered a $25 gift card for completing a questionnaire at baseline and again at follow-up. (It was also explained to staff that, if assigned to the treatment group, they would receive professional growth incentives in cash or credits for completing the PITC, but this incentive is integral and not specific to study participation.) Parents were offered a $10 gift card for completing the baseline questionnaire and consent form, and $50 for completing each wave of follow-up assessments. Child care centers were fully enrolled in the study if they met the eligibility criteria and submitted the following:

- Signed consent forms from the director and at least four infant/toddler staff (at least four staff in one classroom or at least two staff in each of two classrooms. Eligible staff were those working directly with children in the targeted classrooms on a regular basis (not floaters). Classrooms needed to include children younger than 24 months.
- Signed consent forms from parents of at least five children in one classroom or three children in each of two classrooms. Children were eligible if younger than 27 months (preferably younger than 24 months) at the time of enrollment and if the primary language in their home was Spanish or English.

Family child care homes were enrolled if they met eligibility criteria and submitted the following:

- A signed consent form from the director/owner. (Enrollment of additional staff or assistants who worked with children was optional.)
- A signed consent form from parents for at least one child younger than 27 months. (Enrollment of more than one child was encouraged.)

Although center directors were asked to speak to all parents of children of eligible ages, were given 10 parent/child consent forms per target classroom, and were given informational flyers about the study, they were not expected to follow up on consent forms beyond the minimum number required for the study. Family child care providers were given consent forms for four parents but again were not expected to follow up on children beyond the minimum required. The research team did not track the program directors' and providers' communications with families. Programs were not asked to document how many families they approached or whether all families in the identified classrooms received the consent materials.

Data on screening participants who did not enroll in the study were destroyed, since these providers did not complete consent forms permitting the study team to maintain or analyze their data.

Enrollment, data collection, and random assignment occurred on a rolling basis. Baseline data collection for each program was completed as soon as possible after the program, including the

required numbers of program staff and children, enrolled in the study. When sufficient programs within at least one stratum (strata were based on type of program and location) had completed baseline data collection, random assignment was carried out for these strata. On completion of random assignment, programs were notified of their treatment or control status and PITC staff were notified that they could contact the treatment programs to begin scheduling treatment. Random assignment was carried out on 11 dates between late 2007 and mid-2008. The process continued until the sample was complete. Follow-up data collection was timed for each program, as well as for study children associated with the program, based on the program's random assignment date. Children could not join the study after their programs' random assignment.

The full sample was recruited as of summer 2008. Final sample numbers slightly exceeded the targets: 92 child care centers, 159 family child care homes, and 1,009 children were enrolled. However, 73 children left their child care settings before random assignment, reducing the child sample to 936. Although children who left their child care settings prior to random assignment were followed throughout data collection, they were excluded from analyses.

Random assignment

The research team developed and carried out random assignment using a routine that produced a 50/50 random assignment ratio across the various random assignment clusters (see appendix C). Researchers conducted random assignment on a rolling basis on 11 dates between November 1, 2007 and August 5, 2008 (see appendix D). Strata were based on location and on program type: centers, English-speaking family child care homes, and Spanish-speaking family child care homes. Locations were identified as local areas that could be served by certified trainers and within which family child care training groups could be formed.

A new cohort was randomly assigned when at least two programs were ready (had completed baseline data collection). For family child care homes, to ensure that groups could be formed, it was necessary to wait until at least four programs could be randomly assigned in a single county or across adjacent areas of two counties. An exception to this was when family child care groups already forming in those areas could accommodate additional members.[3]

Maintaining integrity of random assignment

For all waves of data collection, the study team followed up with all child care programs and with all children who had enrolled and who had been randomly assigned, regardless of whether programs remained in treatment and whether children remained in their original child care arrangements. This follow-up included full attempts to contact and complete data collection with all sample members. Thus, at the program and child levels, the study is a genuine intent-to-treat study. It thus avoided bias that might have resulted from selective attrition or enrollment.

Program crossovers were extremely rare. In California, all control programs were flagged in the PITC training database so that they could not enter training before data collection was complete.

[3] It was important that rounds of random assignment proceed quickly, so that the process of notifying programs of their status and assigning trainers to treatment programs could move ahead as well. There was some concern that programs, having participated in baseline data collection, would become impatient with delay in learning their status and that the risk of attrition would be increased. Further, trainers in some areas were funded by their counties within particular time constraints.

In Arizona, the PITC was not available to programs other than those in the study. In both states, providers assigned to the control group and their staffs might have had informal relationships or conversations with providers assigned to the treatment; however, it is unlikely that such informal relationships or conversations resulted in any serious contamination of control group programs. Adoption of PITC practices requires months of training and a concerted effort by program administrators and staff. To further reduce the likelihood of control group contamination, it was ensured that the recruited providers were not part of corporate entities that implement cross-site training or staff sharing. The participation by sites of Kinder Care and other child care chains were limited to one site per chain in each state.

Child crossovers could have occurred in either direction. At the time of the final child assessment, 48 percent of study children had left their original child care settings. (Many children did not re-enter child care in the same locations and many were no longer younger than three at the time of the change. See chapter 5 for more discussion of child mobility.) Although child care changes and names of some programs to which children moved were collected in parent interviews, program names were too incomplete for them to be checked accurately for child crossovers.

Sample retention and attrition

Figures 2.1 and 2.2 present the eligible and sample numbers at each stage of screening, enrollment, random assignment, and data collection. Numbers of eligible teachers and children within programs are estimated based on program directors' estimates—provided in screening interviews—of numbers of children younger than 24 months and numbers of teachers caring for infants and toddlers up to age three. To reach the program sample target within the study timeline and budget, no more than two classrooms per center were included in the study. In centers that included more than two infant/toddler classrooms, researchers randomly selected one infant classroom and one toddler classroom. Consent from only five children in one classroom—or six children across two classrooms—and five staff (including the director) were required for child care center enrollment, and consent from only one child and the lead provider were required for family child care enrollment. However, additional eligible children and teachers[4] were enrolled if they provided consent before random assignment. Ultimately, an average of about eight children per center and slightly less than two children per family child care home enrolled in the study. An average of nine teachers per center (over 80 percent of eligible teachers) and 1.55 caregivers per family child care home enrolled. Since data could not be collected on children and staff who did not provide consent to participate, their characteristics could not be compared to those of enrollees.

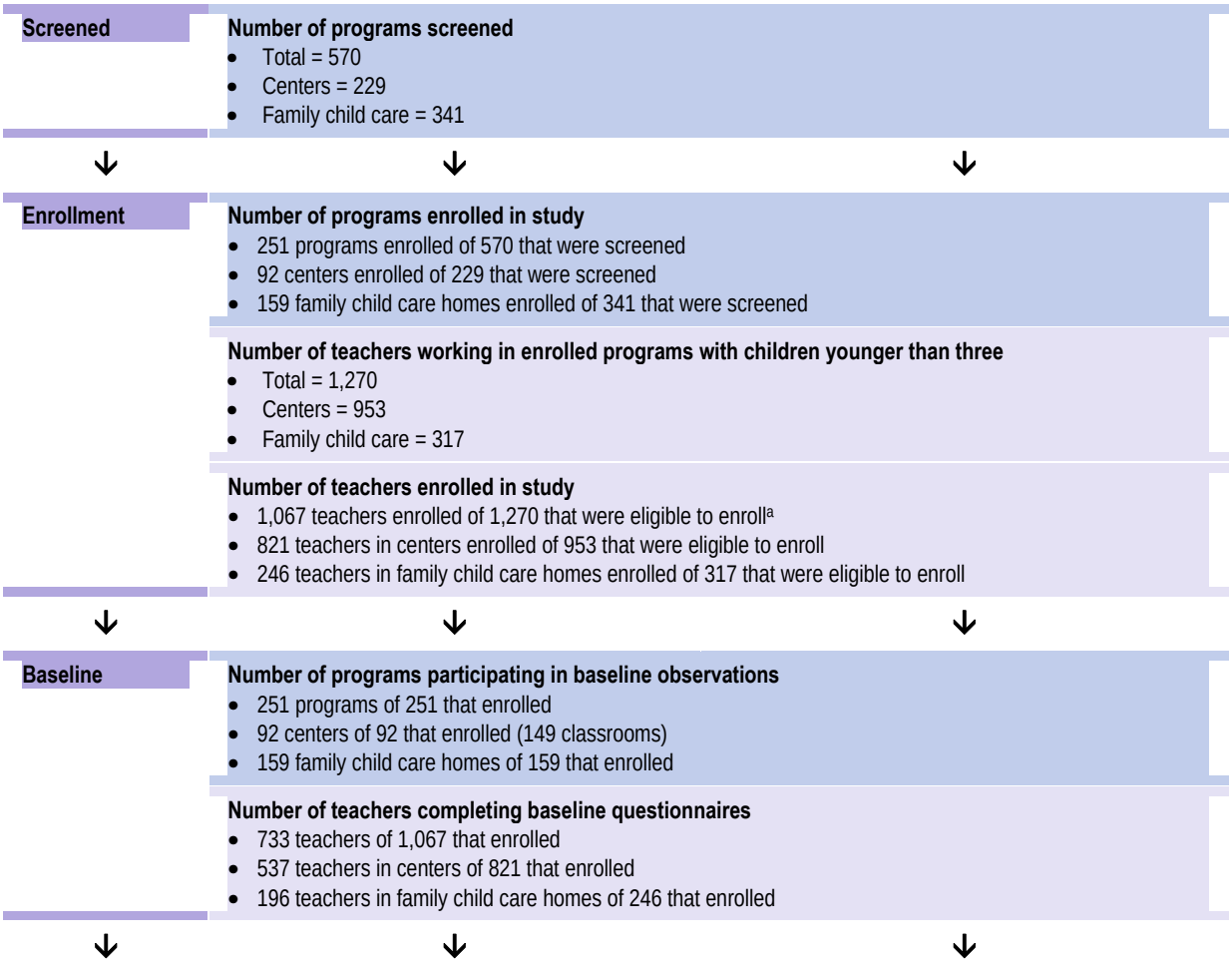
Because procedures for obtaining consent were the same in experimental and control groups, and were carried out before random assignment, the study design minimized or eliminated systematic differences in experimental and control group attrition patterns that might have compromised the study's validity. In addition, researchers followed up with all members of the original child and program samples, using procedures to locate and contact families and providers that had moved, changed contact information, or had not responded to the first several attempts to contact them.

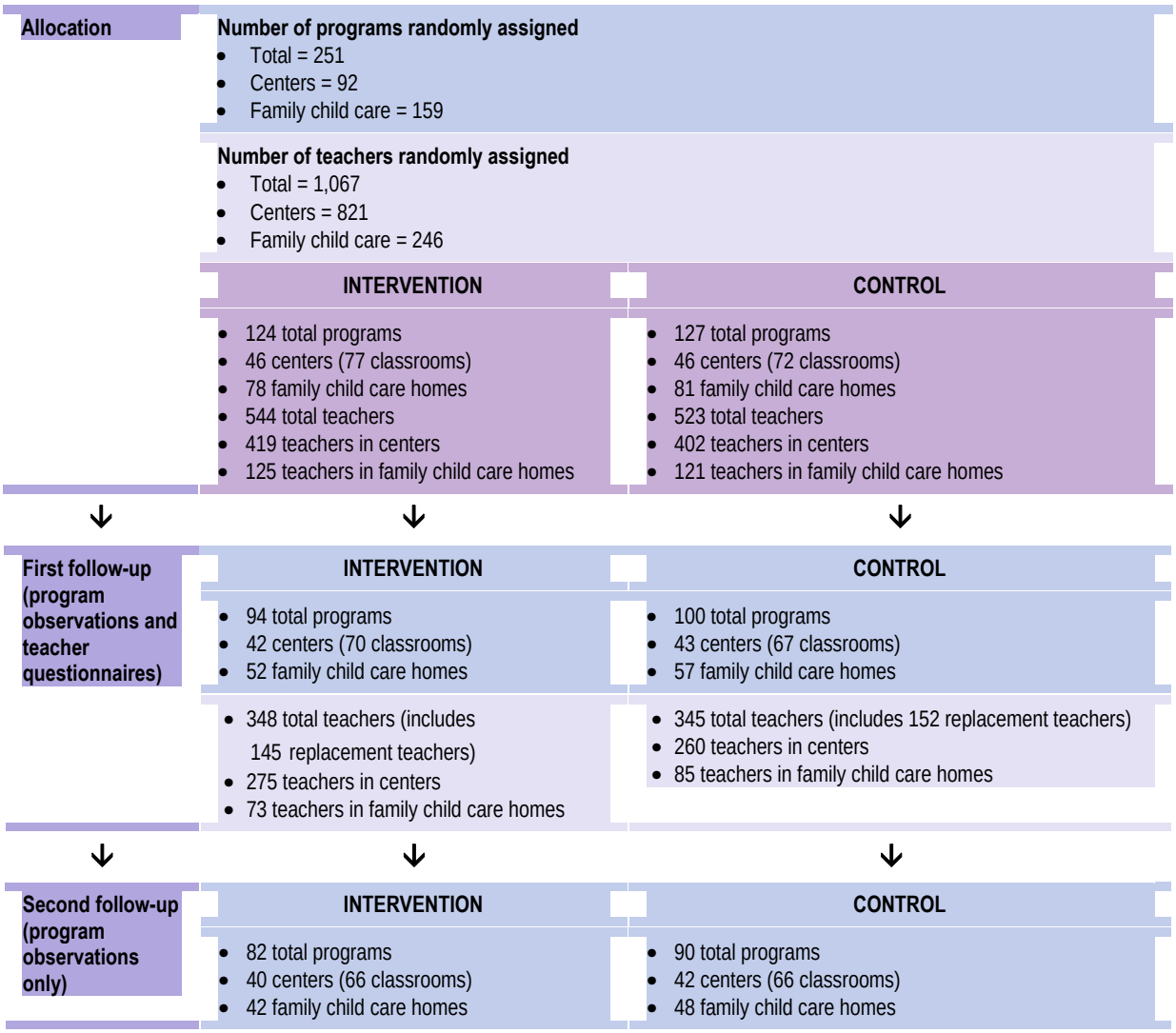
[4] In centers with more than two infant/toddler classrooms, some teachers who worked in classrooms other than the two targeted for the study were enrolled. However, enrolled children were limited to the two target classrooms.

Specific data collection and follow-up procedures are described later in this chapter. Response rates and detailed reasons for nonresponse are provided in table 2.6 and appendix J.

Individual teacher measures were not among confirmatory outcomes for the study. Individual program staff and directors who participated in baseline data collection but who left the programs after random assignment and before follow-up data collection were not included in follow-up. Instead, the study team requested consent to participate from staff and directors who replaced the original study participants. Follow-up questionnaires and observation data, as well as participation data, were collected on these replacement staff if consent was provided.

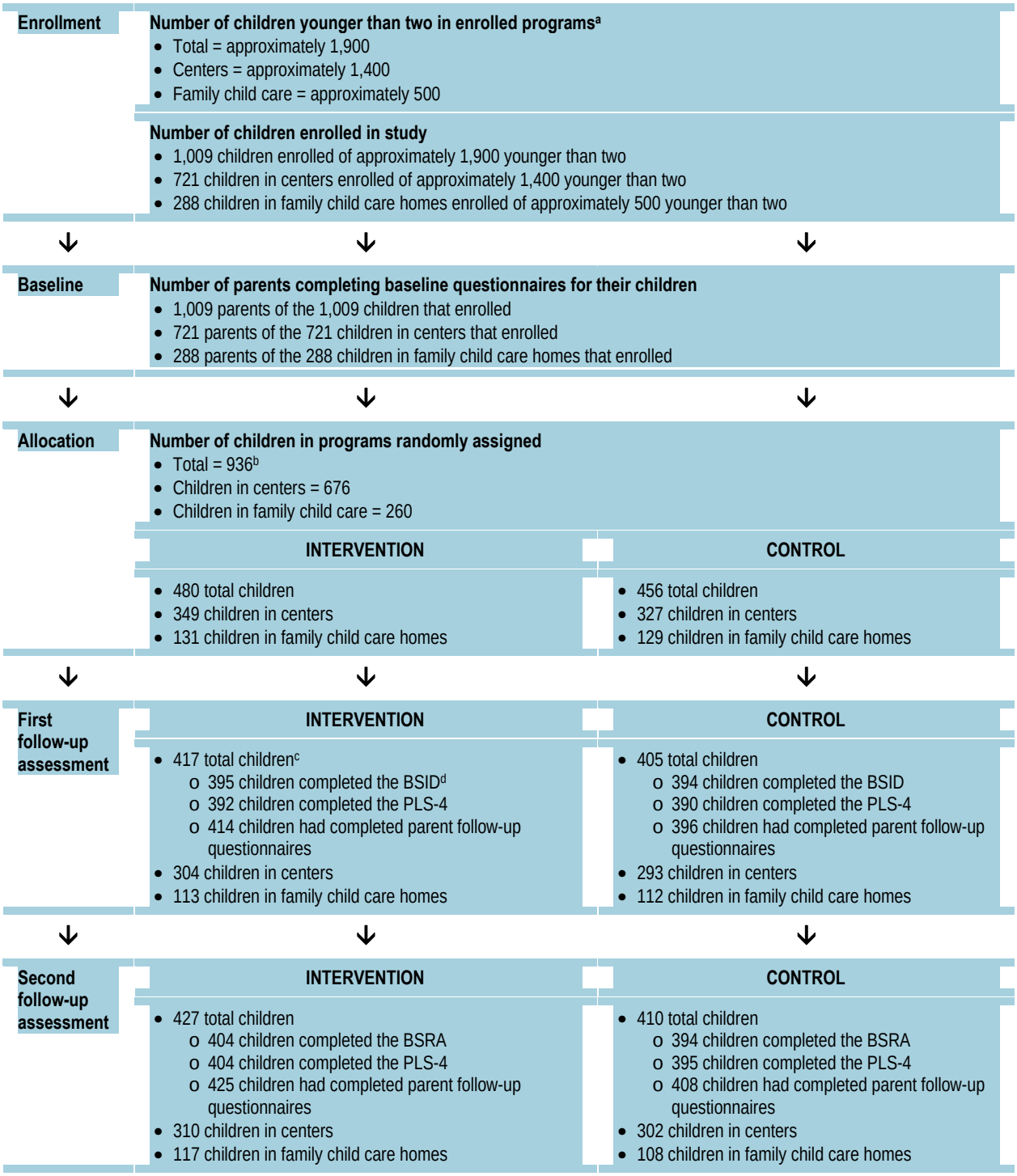
Figure 2.1. Sample flow: programs and teachers





Notes: All teachers working in family child care homes were eligible to participate in the study. Teachers in centers include center directors. Reasons programs were not observed at follow-up include: refusals, program closures, observers were not able to contact the programs, or programs were no longer providing care to infants and toddlers. Of the 92 participating centers, 35 had one participating classroom and 57 had two. At first follow-up of the 85 centers participating, 33 had one classroom and 52 had two. At second follow-up of the 82 centers participating, 32 had one classroom and 50 had two. See table 2.6 and appendix J for response rates and detailed reasons for nonresponse.
a. Includes infant/toddler teachers outside of the two target classrooms.
Source: Sample tracking, screening questionnaire, program observation data files for all three rounds of data collection, and teacher questionnaire data files for baseline and first follow-up.

Figure 2.2. Sample flow: children



a. Total numbers of enrolled children younger than two are program directors' estimates provided in screening interviews. To enroll the program in the study, center directors were asked to collect consent forms from a minimum of five children in one classroom or a minimum of three children in each of two classrooms. Although center directors were asked to speak to all parents of children of eligible ages and were given 10 parent/child consent forms per target classroom, they were not expected to follow up on consent forms beyond the minimum number required. Family child care providers were given

consent forms for four parents but were not expected to follow up on children beyond the minimum of two (later changed to one) per program required for family child care enrollment.
b. 73 children left their study programs after study enrollment and before random assignment. Although these children were followed throughout data collection, they were excluded from the analyses.
c. Total children include those who participated in at least one of the three main components of the child assessment used in the impact analysis: the *Bayley Scales of Infant and Toddler Development, Third Edition*, Cognitive Subscale (BSID)/*Bracken School Readiness Assessment, Third Edition* (BSRA); the *Preschool Language Scale, Fourth Edition*, Expressive Communication Subscale (PLS-4); or the parent follow-up questionnaire (including the behavior assessment). Children do not have complete child assessments at first and/or second follow-up for the following reasons: their parents refused to participate, they were ineligible due to language or severe disability, they moved out of the area, the research office staff were unable to contact or schedule for an assessment, or the data was collected but it was lost or stolen. A child was considered ineligible for the assessment if the parent reported that the child's primary language was not English or Spanish, unless the parent was certain that the child would understand and be able to respond to directions in English or Spanish at least 70 percent of the time. Children with severe developmental delays or other disabilities that could not be accommodated were considered ineligible for the assessments. Parents of children who were considered ineligible were asked to complete a parent follow-up questionnaire and the child care interview. See table 2.6 and appendix J for response rates and details of nonresponse.
d. Completion numbers for the cognitive and language assessments, at both follow-up points, differ slightly (four or fewer children) from the child sample numbers included in the primary and exploratory impact analyses for the cognitive/language composite. These differences are due to documentation of children as "completed" although one or both scores were later determined not to be usable (for language or disability reasons); inability to create composite scores needed for analysis when one component score was missing; and documentation of children as "missing" in sample tracking although scores were used in analysis.
Source: Sample tracking, screening questionnaire data, and child data for all three rounds of data collection.

Data collection

Data were collected at baseline (before random assignment) and in two follow-up rounds. The average follow-up period for program-level data collection was 14.97 months after random assignment for the first follow-up and 21.79 months after random assignment for the second. The average follow-up period for child cognitive and language assessments was 14.77 months after random assignment for the first follow-up and 23.18 months after random assignment for the second. No significant differences were found between treatment and control group follow-up periods for data collection, for either programs or children. See tables E1 and E2 in appendix E for details on follow-up intervals.

Table 2.2 overviews all the measures used in the study, as well as the timing for administration of the measures. See tables F1 and F2 in appendix F for more details of the two rounds of child assessments. Descriptive and psychometric information for the key analytic scales are presented in table F3 in appendix F. Zero-order correlations among the key analytic variables are presented in table G1 in appendix G.

Table 2.2. Data collection overview

Instrument	Respondent group	Content	Mode of administration	Timeline (months after random assignment)
Outcome measures and covariates				
Child care provider screening interview	Child care directors (centers) or owners (family child care homes)	Basic program characteristics: number and ages of children, staffing, languages spoken, program stability	Telephone	Recruitment/ baseline
Parent baseline questionnaire	Parents	Family and child background characteristics: children's ages, parents' education and employment status, ethnicity and languages spoken	Paper and pencil, attached to consent form	Recruitment/ baseline
Center director questionnaire	Child care center directors	Center goals, services, practices, detailed enrollment, and staffing characteristics	Paper and pencil, distributed by mail (after enrollment and at follow-up), collected by observer	Baseline and 15-month program follow-up
Center teacher/ caregiver questionnaire and family child care provider/ caregiver questionnaire	Child care staff working with children younger than three	Caregiver attitudes, beliefs, background, training, education	Paper and pencil, distributed by mail (after enrollment and at follow-up), collected by observer	Baseline and 15-month program follow-up
Individual child form for caregiver	One teacher/ caregiver for each child	Child's relationships, behavior, language in the care setting	Paper and pencil, distributed by mail (after enrollment and at follow-up), collected by observer	Baseline
Center observation instrument: ITERS-R and items from PITC-PARS. Family child care observation instrument: FCCERS-R and items from PITC-PARS	Unit of observation is classroom in centers and whole program in family child care	Program quality indicators	Observation by field researchers	Baseline, 15 months, 22 months
Bayley Scales of Infant and Toddler Development, Third Edition (BSID), Cognitive Subscale	Children	Cognitive skills	In-person assessment/ interview by field researchers	15 months

Instrument	Respondent group	Content	Mode of administration	Timeline (months after random assignment)
Preschool Language Scale, Fourth Edition(PLS-4), Expressive Communication Subscale	Children	Expressive language skills	In-person assessment/ interview by field researchers	15 and 23months
Bracken School Readiness Assessment, Third Edition (BSRA)	Children	Cognitive skills and school readiness	In-person assessment/ interview by field researchers	23 months
Infant-Toddler Social and Emotional Assessment (ITSEA)	Parents	Externalizing, internalizing, dysregulation, and competence	Parent questionnaire (usually self-administered; some in-person interview)	15 months
Child Behavior Checklist 1½-5 (CBCL 1½-5)	Parent	Problem behaviors, internalizing and externalizing	Parent questionnaire	23 months
Polit Positive Behavior Scale	Parent	Compliance, social competence, autonomy	Parent questionnaire	23 months
Parent questionnaire	Parent	Parental employment, child health	Parent questionnaire	15 and 23 months
Parent child care interview	Parent	Child care changes	In-person interview by field researchers	15 and 23 months
Implementation measures				
Staff questionnaire	Child care staff (treatment and control)	Training received over past year	Staff questionnaire (self-administered)	15 months
PITC-PARS measures	Child care programs (treatment and control)	PITC policies	Program observations (supplemented by interviews)	15 and 23 months
PITC training database	Treatment programs and staff	Training duration, schedules, individual hours of participation by topics, activities, and dates	PITC participant forms and sign-in sheets	Throughout treatment
Interviews with trainers and coordinators	PITC trainers and regional coordinators in California and Arizona	Fidelity monitoring, training approaches, factors that affected implementation	Telephone	At conclusion of treatment (December 2009)

Note: ITERS-R is the Infant/Toddler Environment Rating Scale-Revised. PITC-PARS is the Program for Infant/Toddler Care Program Assessment Rating Scale. FCCERS-R is the Family Child Care Environment Rating Scale-Revised.
Source: Authors' study design.

Child measures

Below is a description of each child measure used for the confirmatory and exploratory analyses.[5] All measures are widely used in child development research involving children within this age range. Because some measures have limited standardization age ranges, different measures were used for some cognitive and behavioral outcomes at first and second follow-up waves.

Cognitive development

The *Bayley Scales of Infant and Toddler Development, Third Edition*, Cognitive Subscale (BSID; Bayley 2006) has been used in several national evaluations of young children. A modified short form was used in the Early Childhood Longitudinal Study–Birth Cohort, and an earlier version was used in the Early Head Start Research and Evaluation Project (U.S. Department of Health and Human Services, Administration for Children and Families 2002). The BSID is published in English and is standardized on a sample of 1,700 children (birth to 42 months) stratified by parent education level, race/ethnicity, and geographic region to represent the population based on the 2000 census. The split-half reliability of the cognitive subtest was 0.91 in the standardization sample and the test-retest reliability was 0.81 (Bayley 2006). The BSID, Cognitive subtest correlates 0.79 with the Verbal IQ subscale, 0.72 with the Performance IQ subscale, and 0.79 with the Full Scale IQ score from the *Wechsler Preschool and Primary School Scale of Intelligence, Third Edition* (Wechsler 2003). The BSID Cognitive subtest is correlated 0.57 with the *Preschool Language Scale, Fourth Edition* (PLS-4; Zimmerman, Steiner, and Pond 2002) total language score (Bayley 2006). The BSID includes some school readiness indicators, such as identifying colors and counting. For this study, it was administered in English or Spanish, depending on the predominant language spoken in the child's home.[6]

The *Bracken School Readiness Assessment, Third Edition* (BSRA; Bracken 2007) total score is comprised of six subtests: colors, letters, numbers/counting, sizes, comparisons, and shapes. An earlier version of the BSRA was used in the NICHD Study of Early Child Care and Youth Development (NICHD Early Childcare Research Network 2005). During this 10–15 minute test, the examiner presents pictures and the child is asked to identify concepts within subtest categories by pointing. The BSRA was normed on a sample of 640 children stratified by age, gender, race/ethnicity, geographic region, and parent education level to represent the U.S. population and has both English and Spanish language versions. The publisher reports a test-retest reliability of 0.92 for children younger than five and a split-half reliability of 0.96 or better for children younger than five. The School Readiness Composite has been shown to predict academic achievement (Panter 2000), and the publisher reports that it correlates 0.66 with the total language score from the Preschool Language Scale-4 (Bracken 2007). The administration of

[5] The child measures are copyrighted and cannot be included in the report.

[6] If the child's primary language was Spanish, bilingual assessors used a scoring sheet adapted for this study to include Spanish translations of instructions and prompts for the child to ensure that all the assessors would follow the same administration procedures.

the test is nearly identical in English or Spanish, and it is administered using the predominant language in the child's home.[7]

Language development

The Preschool Language Scale, Fourth Edition, Expressive Communication Subscale (PLS-4; Zimmerman, Steiner, and Pond 2002) measures expressive communication skills for children from birth through 6 years, 11 months of age, using play and structured activities to elicit key indicators of vocal development and social communication. The PLS-4 was selected for use in the Early Head Start Family and Child Experiences Study, and an earlier version of the PLS was used in the National Survey of Child and Adolescent Well-Being (U.S. Department of Health and Human Services, Administration for Children Youth and Families 2005). Trained assessors administered the PLS-4, Expressive Communication Subscale, at the first and second follow-up, for approximately 20 minutes. The scale was normed on a sample of 1,500 children and has both English and Spanish versions that were normed separately. Test-retest reliability of the Expressive Communication Subscale ranged from 0.82 to 0.95 across age groups. The publisher reports a high level of agreement between the PLS-4 and the language strand of the Denver II (Frankenburg, Dodds, Archer, Shapiro and Bresnick 1992), citing evidence that typically developing children score in the normal range on both the PLS-4 and the Denver II. Specifically, all 37 children administered both tests scored within the "normal" range on the Denver II and were within one standard deviation of the mean on the PLS-4 (Zimmerman, Steiner, and Pond 2002). The publisher also provides extensive information regarding clinical validity of the PLS-4, which is able to differentiate those within the normal range from those with language disorders, developmental delay, autism, and hearing.

Behavior/socioemotional development

The *Infant-Toddler Social and Emotional Assessment* (ITSEA; Briggs-Gowan and Carter 2005, Carter and Briggs-Gowan 2006), completed by parents at the first assessment, covers several dimensions of behavior that are hypothesized to be affected by the PITC, both on the positive and problem ends of the spectrum. The ITSEA measures four domains of socioemotional behavior: externalizing, internalizing, dysregulation, and competence. A shortened screening version, the BITSEA, was selected for the Early Head Start Family and Child Experiences Study (U.S. Department of Health and Human Services, Administration for Children and Families 2009). The full version of this scale was chosen for this study because no other single assessment for this age range captures differentiated aspects of both problem behavior and social competence. The ITSEA was normed on a national sample of 600 children, stratified by ethnicity, parent education level, and region, to represent the U.S. population based on the 2000 Census. Publisher-reported test-retest reliability for the domains ranged from 0.76 to 0.91. The publisher reports agreement between the ITSEA domain scores and subscales with the Child Behavior Checklist 1½-5, the Ages and Stages Questionnaire socioemotional scale, and the

[7] The BSRA is standardized for children ages three to six. About 50 children in the sample were younger than three at the time of the second assessment. Members of the research team conducted six practice assessments with children of varied backgrounds between the ages of 27 and 31 months. Assessors established basal scores for all children, and total score ranges were similar to those obtained by the publisher for children 36 months and older. The assessment was therefore administered to all children aged 27 months and older, and a small number of children were delayed in their second assessment until they reached 27 months. Raw scores were used for the analysis.

Adaptive Behavior Assessment. The test has both English and Spanish versions and was written for parents with a grade 4 reading level. A positive behavior score was calculated as the mean of four domains: competence, externalizing (reverse-scored), internalizing (reverse-scored), and dysregulation (reverse-scored). In the PITC sample, internal consistency was 0.63 for the total score.[8]

As part of the second child assessment, parents' ratings of positive behaviors were measured by the *Positive Behavior Scale* (Polit 1996), developed for the New Chance survey (Quint, Bos, and Polit 1997) of over 2,000 low-income mothers and their children. Its 25 items are divided into three subscales: compliance/self-control (thinks before he/she acts, usually does what I tell him/her), social competence and sensitivity (gets along well with other children, shows concern for other people's feelings), and autonomy (tries to do things for him/herself, is self-reliant). The parent responds on a five-point scale, ranging from "never" to "all of the time." In the PITC sample, internal consistency was 0.82 for the total positive behavior score.

At the second child assessment, parents also completed the *Child Behavior Checklist* (CBCL 1½-5; Achenbach and Rescorla 2000), which has 99 problem items that can be used to derive subscale scores for internalizing problems and externalizing problems. The CBCL 1½-5 is a widely used measure in developmental research and was used in the NICHD Study of Early Child Care (NICHD Early Child Care Research Network 2005). CBCL 1½-5 is normed on a national sample of 700 children. The publisher reports an eight-day test-retest reliability of 0.90 for the total problems scale, a 12-month stability coefficient of 0.76, and agreement with the ITSEA and the Toddler Behavior Screening Inventory. For the confirmatory analysis, the sum of the internalizing and externalizing behavior raw scores was used as a measure of behavior problems. In the PITC sample, internal consistency was 0.78 for the total problem behavior score.

Additional child measures used as covariates

At baseline, caregivers (child care program staff) were asked to rate children's behavior problems, including irritability, distractibility, and emotional intensity, using eight items from the Early Childhood Longitudinal Study, Birth Cohort (National Center for Education Statistics 2009). Caregivers were also asked to rate children's current language ability using an eight-item index that begins with early forms of communication, such as vowel-like sounds and gestures, and progresses toward increasingly complex communication, such as two-word sentences, complete sentences, and long and complicated sentences. This measure was also adapted from one used in the Early Childhood Longitudinal Study, Birth Cohort (National Center for Education Statistics 2009) to include items appropriate for children older than two.

Program quality measures

Below is a description of the program quality measures used for the study. Identical measures were used at baseline and during the two follow-up waves. These measures were based on observational scales, with a limited use of interviews for items that cannot be observed. The most

[8] The internal consistency of the total score is lower than the benchmark of 0.70 due to the wide range of behaviors measured. Specifically, the ITSEA includes a mixture of items that are fairly infrequent, such as those in the dysregulation domain, as well as items in the competence domain that are more prevalent in the general population.

widely used measures of child care quality, the environmental rating scales, were used in conjunction with a PITC-aligned measure. Both types of measures were used in a non-experimental self-evaluation conducted by the PITC (WestEd Center for Child & Family Studies, 2003). The Infant/Toddler Environment Rating Scale-Revised (ITERS-R; Harms, Cryer, and Clifford 2006) measures quality experienced for all children (infants to 2½ years of age) in center-based classrooms. The 39 items of the ITERS-R comprise seven subscales. The counterpart of the ITERS-R for home-based settings is the Family Child Care Environment Rating Scale-Revised (FCCERS-R; Harms, Cryer, and Clifford 2007). The publisher reports evidence of concurrent and predictive validity of the original versions of the environment rating scales, citing associations with structural measures of quality as caregiver-child ratios and caregiver education level (Cryer, Tietze, Burchinal, Leal, and Palacios 1999; Phillipsen, Burchinal, Howes, and Cryer 1998) and evidence of predictive validity in relation to child development (Burchinal, Roberts, Nabots, and Bryant 1996; Peisner-Feinberg et al. 1999). The publisher also reports that, because the current revisions maintain the basic properties of the original instruments, the revised scales are expected to maintain validity (Harms et al. 2006, p.2; Harms et al. 2007, p.3).

The family child care observation instrument is comparable to the center observation instrument, with slight adjustments in items to make them appropriate for family child care settings, which have, for example, mixed age groups. A difference between the two instruments is that the FCCERS-R contains two items not present in the ITERS-R: "space for privacy" in the space and furnishings subscale and "math/numbers" in the activities subscale. Because early math concepts are important for school readiness, this indicator was added to the ITERS-R classroom assessments at the first and second follow-up. This was the only deviation from the published scales.

Both environment rating scales provide a global total quality score, a measure used widely in child care research and in child care policy decision making (U.S. Department of Health and Human Services, Administration for Children and Families 2002, 2010). The global score is an average of six subscale scores. The study total scale internal consistency at each assessment ranged from 0.88–0.89 for the ITERS-R and 0.90 to 0.91 for the FCCERS-R. (See table F3 in appendix F for details on the subscales, including sample items and study alpha coefficients.) For the confirmatory analysis, the total score and six items were included the construction of a composite of staff-child interactions. (See the confirmatory factor analysis discussion in appendix H.) Specifically, two items from the listening and talking subscale (helping children use language, helping children understand language) and four items from the interactions subscale (supervision, peer interaction, staff child interaction, and discipline) were converted to *z*-scores and averaged to create the composite.

The PITC Program Assessment Rating System (PITC-PARS) was designed as a more proximal measure of child care quality, developed by PITC staff in accordance with the PITC's philosophy to measure the quality experienced by children from birth through age three in home-based and center-based settings. The full version has 98 items scored either 1 (met) or 0 (not met) that comprise five subscales: quality of caregiver interaction, family partnerships, cultural responsiveness and inclusion of children with special needs, relationship-based care, physical environment, and routines & record keeping. Mangione, Kriener-Althen, Niggle, and Welsh (2006) found a high degree of concurrent validity between the PITC-PARS total score and the

ITERS-R (r = 0.84), ECERS-R (r = 0.88), and FDCRS (r = 0.86). These same researchers reported that concurrent validity of the quality of caregiver interaction with the subscales from the Arnett Caregiver Interaction Scale was moderate (warmth r = 0.60, criticalness, r = –0.70, and distance r = –0.60).

The research team for the current study worked with PITC staff to select items from this measure that provided the least redundancy with the environment rating scales. Full subscales were not administered. For the confirmatory analysis, only four selected items from the quality of caregiver interaction subscale of the PITC-PARS were used, as part of the construction of a composite of staff-child interactions. (See the confirmatory factor analysis discussion in appendix H.) These include facilitation of cognitive development, responsiveness and sensitivity, positive tone and attentiveness, and responsive engagement and intervention. Other PITC-PARS subscales measuring critical PITC policies were used in the implementation analysis. These subscales measure culturally responsive caregiving, primary caregiving, continuity of care, and group size and ratios.

Caregiver beliefs and attitudes measures

The measures for caregiver attitudes and beliefs scales, used in this study as covariates, are identical to those used in the NICHD Study of Early Child Care and Youth Development (NICHD Early Child Care Research Network 2005). These scales were originally used in center-based child care settings in the National Child Care Staffing Study (Whitebook, Howes, and Phillips 1990).

Caregiver attitudes and beliefs about raising young children were assessed using the 30-item Parental Modernity Scale of Child-rearing and Education Beliefs, or Ideas about Raising Children (Schaefer and Edgerton 1985). The 30-items are rated on a five-point Likert scale (1 = strongly disagree to 5 = strongly agree) yielding a total score and two subscores, progressive/democratic beliefs (reflects attitudes favoring self-directed child behavior) and traditional/authoritarian beliefs (reflects attitudes that child behavior should follow adult directives). For the confirmatory analysis, only the total score was used, with higher values indicating more traditional beliefs about raising young children. This scale was chosen because data from the NICHD Study of Early Child Care and Youth Development showed that caregiver attitudes were associated with child care quality (NICHD Early Child Care Research Network 2000). Specifically, caregivers with more progressive/democratic beliefs and less traditional/authoritarian beliefs tended to provide higher quality care that is sensitive and cognitively stimulating (NICHD Early Child Care Research Network 2000). Caregiver support for educational values is likely associated with parental socioeconomic status; thus, this variable may capture a potential source of bias. Internal consistency for the PITC study sample was 0.86.

The Taking Care of Young Children questionnaire was also used in the NICHD Study of Early Child Care and Youth Development (NICHD Early Child Care Research Network 2005). This questionnaire measures caregiver perceptions of concerns and rewards associated with working in a child care environment. The 28 items are rated on a four-point Likert scale (1 = not at all a concern/reward, 4 = extremely a concern/reward), items that were answered not applicable were recoded to 1, because the item cannot be a concern/reward if it does not apply. Items from the "concern" portion of the questionnaire were reverse-scored so that larger composite values

would indicate more positive attitudes about taking care of young children. Internal consistency for the PITC study sample was 0.75.

Teacher/caregiver and director questionnaires

These caregiver beliefs and attitudes measures were incorporated into questionnaires administered at baseline and at the 15-month follow-up to both treatment and control child care staff. Slightly different versions of the questionnaires were administered to teachers/caregivers in centers and in family child care homes and to center directors. Questions addressed program operations and enrollment, program goals, caregivers' educational and professional backgrounds, and professional development received over the prior 12 months.

Implementation measures

Detailed treatment data in the PITC database, gathered through trainers' documentation and participant sign-in sheets, were transferred to the research team for analysis of participation and implementation fidelity. In addition, research staff interviewed PITC coordinators and reviewed a sample of monthly trainer reports to document the context of implementation, challenges to implementation, oversight of trainers, and schedules of training delivery.

Composite measures for confirmatory and exploratory analyses

Child composite measures

To reduce the number of statistical comparisons, researchers used or created composite measures for confirmatory and exploratory analyses (tables 2.3 and 2.4). Child cognitive, language, and behavioral measures administered at the second follow-up wave (an average of 6 months after PITC completion and 23 months after random assignment) are the primary outcomes for the confirmatory analysis.

For the primary child measures, a cognitive/language composite was formed by converting both the BSRA and the PLS-4 expressive language raw scores into z-scores and taking their average. Raw scores, rather than standard scores were used because some of the children were younger than three at the time of assessment and BSRA standard scores were derived from a sample of children ages three to six. The cognitive/language composite was calculated. A total positive behavior composite score was calculated by converting the CBCL 1½-5 and the Polit Positive Behavior Scale raw scores into z-scores and taking their average. Because high scores on this composite represent high levels of positive behavior, the CBCL problem behavior scale was reverse-scored so that high scores would represent high levels of positive behavior.

For the exploratory analysis using the first follow-up child measures, a cognitive/language composite was formed by converting the raw scores from the BSID cognitive and PLS-4 expressive language tests into z-scores and taking their average. Raw scores were used rather than standard scores because some of the children were older than three at the time of assessment and BSID standard scores were derived from a sample of children ages zero to three. For the exploratory analysis using the first follow-up behavior measures, a total behavior problems composite was calculated by converting the ITSEA domain raw scores into z-scores and taking their average. Because high scores on this composite represent high levels of competence, the

ITSEA problem behavior scales were reverse-scored so that high scores would represent high levels of competence.

Table 2.3. Child outcomes: child assessment at 23 months post random assignment (primary outcomes; age range, 25–53 months)

Domain	Instrument
Cognitive/language composite	
Cognitive	*Bracken School Readiness Assessment, Third Edition* (BSRA; Bracken 2007)
Language	*Preschool Language Scale, Fourth Edition* (PLS-4), Expressive Language Subscale (Zimmerman, Steiner, and Pond 2002)
Behavior/socioemotional composite	
	Child Behavior Checklist 1½-5 (CBCL 1½-5): externalizing and internalizing problems (Achenbach and Rescorla 2000)
	Polit Positive Behavior Scale: social competence, compliance, and autonomy (Polit 1996)

Source: Study design.

Table 2.4. Child outcomes: child assessment at 15 months post random assignment (exploratory outcomes; age range 15–45 months)

Domain	Instrument
Cognitive/language composite	
Cognitive	*Bayley Scales of Infant and Toddler Development, Third Edition* (BSID), Cognitive Subscale (Bayley 2006)
Language	*Preschool Language Scale, Fourth Edition* (PLS-4), Expressive Language Subscale (Zimmerman, Steiner, and Pond 2002)
Behavior/socioemotional composite	Infant-Toddler Social and Emotional Assessment (ITSEA): externalizing, internalizing, dysregulation, and competence (Briggs-Gowan and Carter 2005; Carter and Briggs-Gowan 2006)

Source: Study design.

Program composite measures

Program quality measures, collected at 22 months after random assignment of programs (about 4 months after intervention completion), are also post-intervention measures and are analyzed as secondary confirmatory measures. Program quality measures and child measures collected at first follow-up are included in the exploratory analysis.

Outcomes in the domain of program quality include the global quality measure of the ITERS-R/FCCERS-R and a composite measure of staff-child interactions constructed using items from both the ITERS-R/FCCERS-R and the PITC-PARS (table 2.5). Items for the composite were selected to capture interactions between caregivers and children that are "emotionally supportive" and "cognitively stimulating" based on theory (Dowsett, Huston, Imes, and Gennetian 2008; Thomason 2009) and evidence from nonexperimental studies of preschool settings (Howes 2008; Mashburn et al. 2008; Sylva et al. 2006) that these types of interactions may promote children's social behavior and preacademic skills. A confirmatory factor analysis

of the selected staff-child interaction items, using baseline measures, indicated that all these items had loadings above 0.66 on the composite and the model has adequate fit (see appendix H). To address possible concerns that the PITC-PARS measures are overaligned with the intervention, confirmatory impacts on the staff-child interaction composite measure were conducted both with the PITC-PARS measures (internal consistencies of the composites at each time point ranged from 0.90to 0.92) and without the PITC-PARS measures (internal consistencies of the composites at each time point ranged from 0.88 to 0.89). *Z*-scores of selected items were averaged to form the composites.

Table 2.5. Program outcomes: program quality assessments at 22 months (secondary outcome) and 15 months (exploratory outcome)

Domain	Instrument
Global child care quality	
Total quality score	ITERS-R/FCCERS-R global measure
Quality of interactions composite	
Helping children understand language	Infant-Toddler/Family Child Care Environment Rating Scale (ITERS-R/FCCERS-R)
Helping children use language	Infant-Toddler/Family Child Care Environment Rating Scale (ITERS-R/FCCERS-R)
Supervision	Infant-Toddler/Family Child Care Environment Rating Scale (ITERS-R/FCCERS-R)
Peer interaction	Infant-Toddler/Family Child Care Environment Rating Scale (ITERS-R/FCCERS-R)
Staff-child interaction	Infant-Toddler/Family Child Care Environment Rating Scale (ITERS-R/FCCERS-R)
Discipline	Infant-Toddler/Family Child Care Environment Rating Scale (ITERS-R/FCCERS-R)
Responsiveness and sensitivity	Program Assessment Rating Scale (PITC-PARS)
Positive tone and attentiveness	Program Assessment Rating Scale (PITC-PARS)
Responsive engagement and intervention	Program Assessment Rating Scale (PITC-PARS)
Facilitating cognitive development	Program Assessment Rating Scale (PITC-PARS)

Source: Study design.

Data collection

Data collection was carried out by highly trained program observers and assessors. All observers and assessors participated in an initial classroom training delivered by members of the research team, followed by practice observations or assessments. Their performance in administering and scoring the practice assessments determined whether the assessors were then certified to proceed with data collection. Observer and assessor rechecks occurred at established intervals. Along with the training, all observers and assessors received manuals with detailed instructions for administration and scoring of each measure; instructions and scripts for contacting, scheduling, greeting, and concluding the site visits; guidelines for general behavior; information about ethical and consent issues; instructions on data handling; specific requirements for certification and recertification; and information on other administrative issues. See appendix I for more

details on training and reliability-checking for program observers and child assessors are in appendix I.

Program-level data collection

The research team hired, trained, and oversaw field observers who were located, to the extent possible, near the study sites. Observers were hired to conduct either family child care observations or center observations; all family child care observers spoke English and Spanish.

For each of the three waves of program observations (baseline, 15-month follow-up, and 22-month follow-up), an observer spent a half-day collecting on-site data at each family child care home and up to two half-days of data collection at each center. The amount of time for data collection at centers depended on whether one or two classrooms needed to be observed. The number of classrooms that needed to be observed per center depended on how many classrooms enrolled in the study. Each classroom required one half-day of data collection. Of centers enrolled in the study, 35 enrolled one classroom and 57 enrolled two classrooms. For baseline and 15-month follow-up, questionnaires were sent to programs two weeks before the visit, with the request that caregivers complete and submit them to the observers at the visit. Office-based researchers conducted additional phone follow-up to collect outstanding questionnaires. Child care staff and directors who submitted questionnaires received $25 at baseline and another $25 at follow-up. The second program follow-up, at 22 months after random assignment, included observations only.

Observers were trained how to schedule observations and, during biweekly calls with the observers, the field manager discussed strategies and hints for ensuring that as many programs as possible scheduled and completed observations. Sometimes, observers encountered programs where the contact information was no longer correct. In these cases, observers contacted the field manager; office staff conducted Internet yellow and white pages searches or contacted resource and referral agencies to try to obtain current contact information. Programs that noted they were no longer providing care to infants/toddlers when they were contacted for the first follow-up program observation were contacted again for the second program follow-up observation. An observation was scheduled if they were caring for infants/toddlers at that time.

Observers conducting follow-up observations were blind to the programs' random assignment status when they visited the sites. The programs received a letter before the visit explaining that the observers were not associated directly with the PITC and could not discuss the intervention or the study (other than the observation procedures) with the program staff. The letter also provided contact information for reaching both PITC and Berkeley Policy Associates research staff, with the recommendation that program staff contact these individuals with any questions about the PITC or the study. Despite these procedures to maintain blindness of the program observers, some site staff did reveal the treatment/control group status of their programs to the observers at the visit.[9]

Family child care homes were more likely than centers to go out of business before the study was completed or to refuse to participate at follow-up. Some family child care homes could not be observed because they were no longer serving infants or toddlers at the time of follow-up. At the

[9] Researchers did not track the number of unblinded observations.

second follow-up, of 69 family child care homes that did not complete observations, 17 had closed, 15 were no longer serving infants and toddlers, 8 could not be contacted, and 29 refused. There were no differences between treatment and control child care program response rates. See table 2.6 for response rates. Additional details, including reasons for nonresponse at each wave, are in table J1 in appendix J.

Individual program staff and directors who participated in baseline data collection but who left the programs after random assignment and before follow-up data collection were not included in follow-up. Instead, the study team requested consent to participate from staff and directors who replaced the original study participants. Follow-up questionnaires and observation data, as well as participation data, were collected on these replacement staff if consent was provided. Since the program director scheduled the observer's time on-site, replacement teachers usually provided consent for observations. (See appendix K for details of the teacher sample turnover and teacher questionnaire response rates.) Response rates were higher for family child care staff and child care center directors than for other child care center staff.

Child data collection

The research team maintained contact with participating families, facilitated by gathering several alternative phone numbers, as well as making multiple mail and email contacts during the study. Families received birthday cards, as well as quarterly reminder postcards. In addition, families received $50 at the completion of each of the two major rounds of data collection. The research team maintained contact with families who moved out of the area and collected parent follow-up questionnaires even if direct child assessments were not possible.

Office staff contacted parents to meet with the child and a parent at a convenient time and location, most frequently in the child's home. Parents were mailed questionnaires before the child visits and were asked to complete them by the time of the visit if possible; if not, the assessor allowed the parent to complete the questionnaire during the visit.

Child assessors in all cases were blind to the random assignment status of the children. Initial contacts with families were made by research staff who addressed any questions about the study before assessors' calling to schedule a visit with the family. Parents did not discuss the PITC or their random assignment status during the visits.

The child assessment completion rate was 84 percent for the first follow-up and 85 percent for the second (table 2.6). The differences in response rates for the treatment and control groups were not significant at the .05 level. A completed child assessment is one where the child and/or parent completed at least four of the six components of the child assessment: the BSID/BSRA, the PLS-4, the HOME Inventory, the child care interview (with parent), the parent follow-up questionnaire, and the Behavior Observation Inventory.

Reasons for nonresponse to the child assessments varied. Detailed assessment completion rates, including response rates for the specific components of the assessment and reasons for nonresponse at each wave, are in appendix table J1. Some parents refused to participate in the assessments. Some families moved out of the area and could not be assessed. (Among families that had moved out of the area, some parents completed questionnaires although children could not be assessed.) The assessment team was unable to contact some families, and in some cases,

though the assessment team was able to make an initial contact, families were nonresponsive even after multiple attempts to schedule the assessment.

Language and disability were among the causes of noncompletion of child assessments. The parent was asked to verify that the child's primary language was English or Spanish before scheduling the home visit. In cases where Spanish was spoken in the home, a bilingual assessor was assigned to the family. If the parent reported that the child spoke Spanish 50 percent of the time or more, the bilingual assessor was instructed to administer the Spanish version of the PLS-4 and use the appropriate score sheet and stimulus booklet. Children were considered ineligible for the assessments if the parent reported that the child's primary language was neither English nor Spanish, unless the parent was certain that the child would understand and be able to respond to directions in English or Spanish at least 70 percent of the time. Children with severe developmental delays or other disabilities that could not be accommodated were considered ineligible for the assessments.

Table 2.6. Response rates for outcome measures

	Overall		Intervention		Control		Percentage difference between groups	*t*-statistic	*p*-value
Outcome measure	**Number**	**Percent**	**Number**	**Percent**	**Number**	**Percent**			
Number of children enrolled in study	936	100	480	100	456	100			
Completed BSID at first follow-up	789	84	395	82	394	86	4	1.73	.084
Completed PLS-4 at first follow-up	782	84	392	82	390	85	3	1.59	.111
Completed parent follow-up questionnaire at first follow-up	810	86	414	86	396	87	1	0.27	.791
Completed BSRA at second follow-up	798	85	404	84	394	86	2	0.96	.335
Completed PLS-4 at second follow-up	799	85	404	84	395	87	3	1.06	.288
Completed parent follow-up questionnaire at second follow-up	839	90	429	89	410	90	1	0.27	.787
Number of centers enrolled in study	92	100	46	100	46	100			
Completed center program observations at first follow-up	85	92	42	91	43	93	2	0.39	.698
Completed center program observations at second follow-up	82	89	40	87	42	91	4	0.66	.508
Number of center classrooms enrolled in study	149	100	77	100	72	100			
Completed center classroom observations at first follow-up	137	92	70	91	67	93	2	0.18	.143
Completed center classroom observations at second follow-up	132	89	66	86	66	92	6	0.89	.624
Total number of family child care homes enrolled in study	159	100	78	100	81	100			
Completed family child care home observations at first follow-up	109	69	52	67	57	70	3	0.50	.618
Completed family child care home observations at second follow-up	90	57	42	54	48	59	5	0.69	.494

Note: BSID is the *Bayley Scales of Infant and Toddler Development, Third Edition,* Cognitive Subscale.

PLS-4 is the *Preschool Language Scale, Fourth Edition*, Expressive Communication Subscale.
BSRA is the *Bracken School Readiness Assessment, Third Edition.*
Note: Completion numbers for the cognitive and language assessments, at both follow-up points, differ slightly (four or fewer children) from the child sample numbers included in the primary and exploratory impact analyses for the cognitive/language composite. These differences are due to: documentation of children as "completed" although one or both scores were later determined unusable (for language or disability reasons), inability to create composites when one or both component scores were missing, and documentation of children as "missing" in sample tracking although scores were used in analysis.
Source: Sample tracking; child outcome data files for two follow-up rounds of data collection; program observation data files for two follow-up rounds of data collection.

Sample characteristics

To describe the sample, assess the equivalence of the treatment and control groups, and examine any possible bias caused by attrition, the characteristics of the groups were compared at both baseline and follow-up. Variables were selected based on their potential correlation with study outcomes. In addition, these data capture important demographic and organizational features of the program and child samples. Some program and staff characteristics are presented separately for child care centers and family child care homes to allow for comparability with other child care studies and to provide data that are interpretable given different sizes, structures, and staff backgrounds in the two types of settings.

No group differences were statistically significant at baseline (tables 2.7–2.9).

Table 2.7. Program and staff characteristics at baseline, by experimental condition

Measure	Intervention (standard deviation)	Control (standard deviation)	Difference between groups	*t*-statistic or *z*-score	*p*-value
Program characteristics					
Percent of all programs where English is primary language spoken	42.5	42.1	0.4	–0.07	.945
Percent of all programs receiving government funding	70.7	76.8	6.1	1.08	.279
Mean number of children enrolled younger than 24 months	7.2 (9.99)	7.6 (9.92)	0.4	0.30	.763
Rating of program goals[a]					
Religious instruction	1.6 (0.79)	1.5 (0.76)	0.1	0.71	.478
School readiness	2.7 (0.55)	2.7 (0.50)	0.0	0.75	.455
Assisting disadvantaged children	2.7 (0.53)	2.6 (0.61)	0.1	0.76	.450
Staff characteristics					
Percent who speak English as primary language with the children	61.1	59.6	1.5	0.41	.685
Percent who have a bachelor's degree or more education	23.1	20.2	2.9	0.92	.356
Percent who received early childhood education or training in the last 12 months	84.6	86.1	–1.5	–0.58	.559
Mean Taking Care of Young Children[b] total sum score	81.1 (11.34)	81.4 (12.90)	–0.3	–0.30	.766
Mean Modernity Scale[c] total sum score	76.8 (19.27)	78.0 (18.69)	–1.2	–0.85	.398

Note: Total sample sizes range from 213 to 251 for program characteristics and from 712 to 725 for staff characteristics.
a. Center directors and family child care providers were asked how important (not at all = 1, a little important = 2, or very important = 3) each goal is for their child care program: to provide religious instruction, to prepare children for school with a strong academic curriculum, and to provide compensatory education for disadvantaged children.
b. The Taking Care of Young Children questionnaire total sum score has a possible range of 28 to 112, with higher values indicating more positive attitudes about taking care of young children.
c. The Modernity Scale of Child-rearing and Education Beliefs total sum score has a possible range of 30 to 150, with higher values indicating more traditional beliefs about raising young children.
Source: Screening questionnaire, Baseline family child care, center director, and teacher questionnaires.

Comparisons of treatment and control groups on baseline program quality observational measures are presented for the full sample as well as separately for child care centers and family child care homes in table 2.8. No treatment-control group differences were significant at the $p < .05$ level, but for child care centers, differences in global child care quality as measured by the ITERS-R are different at p <.07. The goal of testing differences at baseline is to determine whether to accept the null hypothesis (no differences), not to reject it, so p values less than .10 suggest caution in assuming that the two groups did not differ at baseline. As described below, baseline global child care quality will be included as a covariate in the impact analysis model.

Table 2.8. Program quality at baseline, by experimental condition

Quality measure	Mean	Standard deviation	Mean	Standard deviation	*t*-statistic or *z*-score	*p*-value
	Intervention *n = 124 programs*		Control *n = 127 programs*			
All programs						
Global quality (ITERS-R/FCCERS-R)[a]	3.66	1.01	3.51	1.12	1.05	.293
PITC-PARS structure[b]						
Culturally responsive care summary score	2.27	1.09	2.24	1.17	0.21	.833
Primary caregiving summary score	2.40	1.32	2.27	1.26	0.65	.515
Continuity of care summary score	2.67	1.08	2.84	1.00	–1.26	.209
	Intervention *n = 77 classrooms*		Control *n = 72 classrooms*			
Child care centers						
Center global quality (ITERS-R total score)	3.66	0.85	3.39	0.97	–1.83	.070
Center PITC-PARS program structure						
Culturally responsive care summary score	2.60	1.09	2.76	1.03	0.96	.340
Primary caregiving summary score	2.58	1.26	2.33	1.21	–1.24	.217
Continuity of care summary score	2.27	1.05	2.40	0.93	0.8	.425
	Intervention *n = 78 family child care home*		Control *n = 81 family child care homes*			
Family child care programs						
Family child care global quality (FCCERS-R total score)	3.66	1.17	3.60	1.24	-0.33	.740
Family child care PITC-PARS program structure						
Culturally responsive care summary score	2.08	1.08	1.93	1.13	-0.86	.390
Primary caregiving summary score	2.21	1.49	2.05	1.30	-0.51	.610
Continuity of care summary score	2.90	1.06	3.07	0.97	1.09	.280

Note: A total of 251 programs had completed observations at baseline (92 centers and 159 family child care homes). In the 92 centers, 149 classrooms were observed. (35 centers with one classroom and 57 centers with two classrooms). In centers where there were two classrooms, the scores were averaged across the two classrooms to create the program level scores.
a. The ITERS-R and FCCERS-R global quality scores range from 1 to 7, where 1 is considered inadequate and 7 is considered excellent.
b. The PITC-PARS structure summary scores range from 0 to 4, where 0 indicates that no indicators for the item were met and 4 indicates that all indicators for the item were met.
Source: Observation data from the baseline observations.

Table 2.9. Child characteristics at baseline, by experimental condition

Measure	Intervention (standard deviation)	Control (standard deviation)	Difference between groups	*t*-statistic or *z*-score	*p*-value
Percent of children who are male	52.2	46.5	5.7	–1.72	.085
Child's race/ethnicity					
Percent of children who are White, non-Hispanic	30.0	28.5	1.5	–0.50	.617
Percent of children who are Black, non-Hispanic	7.5	6.4	1.1	–0.69	.493
Percent of children who are Hispanic	51.0	54.1	3.1	0.94	.349
Percent of children living with two parents	64.9	65.1	0.2	0.09	.932
Percent of parents with a bachelor's degree or more education	36.3	34.3	2.0	–0.63	.531
Percent of children where the primary language at home is English	75.3	73.7	1.6	–0.54	.590
Mean child age in months at random assignment	17.8 (6.91)	17.3 (6.95)	0.5	–1.01	.313
Mean hours per week child attends child care program	37.0 (9.91)	36.2 (10.10)	0.8	–1.24	.216
Mean number of siblings	0.9 (1.04)	1.0 (1.14)	0.1	0.92	.359
Mean parent age	29.9 (7.52)	29.6 (7.21)	0.3	–0.62	.538
Mean number of hours parent is employed or in school	36.3 (11.09)	36.6 (10.92)	0.3	0.40	.688
Teacher rated child behavior and language scales					
Mean score on behavior scale[a]	1.1 (0.64)	1.1 (0.64)	0.0	–0.65	.514
Mean score on language scale[b]	2.8 (1.78)	2.6 (1.79)	0.2	–1.49	.136

Note: Total sample sizes range from 893 to 936 for child/family characteristics and from 788 to 812 for the teacher rate child behavior and language scales.

a. Baseline child behavior scale. As part of a self-administered questionnaire, caregivers rated children's problem behaviors, including fearfulness, irritability, focused attention, and adaptability to change. This eight-item scale is based on scored items, using a 0–3 scale, with 0 indicating that the child "is never like this" and 3 indicating that the child "is like this most of the time." Scores were calculated by taking the average of available items; higher scores indicate higher levels of problem behavior.

b. Baseline child language scale. As part of a self-administered questionnaire, caregivers identified the item that best characterized children's language and communication abilities using an eight-item list (scored 0–7). Lower scores indicate more rudimentary language abilities (the child repeats sounds; the child's sounds have a speech-like babble to them) and higher scores indicate more complex language abilities (the child is talking in fairly complete short sentences; the child is talking in long and complicated sentences).

Source: Parent baseline questionnaire and individual child forms completed at baseline by child's primary teacher.

Tables 2.10 and 2.11 compare the baseline characteristics of the treatment and control groups within the retained sample at the second follow-up only. This was the final analytic sample for the confirmatory impact analysis. Within this retained sample, there were no statistically significant differences between treatment and control groups.

Parallel sample characteristics at the first follow-up are in table L1 in appendix L. The sample equivalency is similar to that at the second follow-up, with no significant differences. Retained and nonretained samples are compared in tables L2 and L3 in appendix L. The same characteristics are compared separately for family child care and centers in tables L4 and L5 in appendix L. At the second follow-up, children not retained were less likely to be from English-speaking homes. Both children's language and family child care providers' language were included as covariates in the impact analysis.

Table 2.10. Baseline characteristics of the retained program and staff sample at second follow-up, by experimental condition

Measure	Overall	Intervention (standard deviation)	Control (standard deviation)	Difference between groups	*t*-statistic	*p*-value
Program characteristics (all programs)						
Percent of programs where English is primary language spoken	59.5	60.3	58.9	1.4	0.18	.858
Percent of programs receiving government funding	74.1	67.9	79.8	–11.9	–1.77	.078
Mean number of children enrolled younger than 24 months	9.5	9.5 (11.68)	9.6 (11.10)	–0.1	0.05	.957
Mean baseline global quality scale score (ITERS-R and FCCERS-R)	3.6	3.7 (3.46)	3.5 (1.11)	0.2	1.56	.120
Mean Concerns and Rewards[a] total sum score (program average)	79.7	80.4 (10.10)	79.0 (12.54)	1.4	0.77	.445
Mean Belief[b] Scale total sum score (program average)	79.3	79.5 (16.53)	79.0 (14.80)	0.5	0.20	.844
Mean Goals[c] (program average):						
Religious instruction	1.5	1.6 (0.81)	1.4 (0.72)	0.2	1.45	.149
School readiness	2.7	2.7 (0.52)	2.7 (0.51)	0.0	0.07	.947
Assisting disadvantaged children	2.6	2.7 (0.53)	2.6 (0.61)	0.1	1.24	.218
Staff/provider characteristics						
Percent who speak English as primary language with the children	60.5	61.1	60.0	1.1	0.27	.787
Percent who have a bachelor's degree or more education	22.5	24.7	20.4	4.3	1.23	.221
Percent who received early childhood education or training in the last 12 months	86.4	86.8	86.1	0.7	0.25	.804

Note: Total sample sizes range from 139 to 172 for program characteristics and from 588 to 582 for staff characteristics.
a. The Taking Care of Young Children questionnaire total sum score has a possible range of 28 to 112, with higher values indicating more positive attitudes about taking care of young children.
b. The Parental Modernity Scale of Child-rearing and Education Beliefs total sum score has a possible range of 30 to 150, with higher values indicating more traditional beliefs about raising young children.

c. Center directors and family child care providers were asked how important (not at all=1, a little important=2, or very important=3) each of the following goals is for their child care program: to provide religious instruction, to prepare children for school with a strong academic curriculum, and to provide compensatory education for disadvantaged children.
Source: Screening questionnaire; baseline center director, center staff/teacher, and family child care home questionnaires.

Table 2.11. Baseline characteristics of retained child sample at second follow-up, by experimental condition

Measure	Overall	Intervention (standard deviation)	Control (standard deviation)	Difference between groups	*t*-statistic or *z*-score	*p*-value
Percent of children who are male	49.4	52.5	46.1	6.4	1.79	.074
Child's race/ethnicity						
Percent of children who are White, non-Hispanic	30.2	30.7	29.6	1.1	0.34	.735
Percent of children who are Black, non-Hispanic	6.4	7.2	5.6	1.6	0.90	.367
Percent of children who are Hispanic	52.9	52.0	53.8	–1.8	–0.52	.602
Percent of children living with two parents	65.3	64.9	65.8	–0.9	–0.28	.777
Percent of parents with a bachelor's degree or more education	34.7	34.8	34.7	0.1	0.03	.976
Percent of children where the primary language at home is English	75.2	76.9	73.5	3.4	1.08	.279
Mean child age in months at random assignment	17.4	17.7 (6.95)	17.1 (6.79)	0.6	1.21	.227
Mean hours per week child attends child care program	36.5	37.0 (9.92)	36.0 (10.13)	1.0	1.44	.150
Mean number of siblings	1.0	1.0 (1.03)	1.0 (1.12)	0.0	–0.22	.824
Mean parent age	29.7	29.8 (7.35)	29.5 (7.19)	0.3	0.52	.604
Mean number of hours parent is employed or in school	37.6	37.9 (11.21)	37.4 (11.11)	0.5	0.59	.558
Teacher-rated child behavior and language scales						
Mean score on behavior scale[a]	1.1	1.1 (0.36)	1.0 (0.38)	0.1	1.51	.130
Mean score on language scale[b]	2.7	2.8 (1.14)	2.6 (1.20)	0.2	1.83	.067

Note: Total sample sizes range from 739 to 796.
a. Baseline child behavior scale. As part of a self-administered questionnaire, caregivers rated children's problem behaviors, including fearfulness, irritability, focused attention, and adaptability to change. This eight-item scale is scored items, using a 0–3 scale, with 0 indicating that the child "is never like this" and 3 indicating that the child "is like this most of the time." Scores were calculated by taking the average of available items; higher scores indicate higher levels of problem behavior.
b. Baseline child language scale. As part of a self-administered questionnaire, caregivers identified the item that best characterized children's language and communication abilities using an eight-item list (scored 0–7). Lower scores indicate more rudimentary language abilities (the child repeats sounds; the child's sounds have a speech-like babble to them) and higher scores indicate more complex language abilities (the child is talking in fairly complete short sentences; the child is talking in long and complicated sentences).
Source: Parent baseline questionnaire and individual child forms completed at baseline by child's primary teacher.

Impact analysis

Multilevel models

The confirmatory research questions were addressed using hierarchical linear regression models to account for the effect of clustering of observations within programs. Each impact analysis included covariates collected before random assignment to improve the estimates' statistical precision and to reduce the likelihood of random sampling variation affecting the impact estimates (table 2.12). Note that only covariates at the same or higher levels of aggregation as the outcome variable were used. (For example, child gender was not used as a covariate at the program level.)

Table 2.12. Baseline covariates used in impact analysis

Level and covariate	Baseline measure
Program level	
Program receives government funding or child subsidies	Screening questionnaire
Staff knowledge and beliefs	Staff baseline questionnaire (aggregated to program level)
Staff education level	Staff baseline questionnaire (aggregated to program level)
Staff early childhood training in year prior to baseline	Staff baseline questionnaire (aggregated to program level)
Randomization strata	Dichotomous indicators for N-1 strata (based on program type, program language, location)
Childcare global quality	ITERS-R/FCCERS-R global quality
Child level	
Child demographics	Primary language spoken in the home Gender Age Parent educational level
Caregiver assessment	Language measure Behavior measure
Follow-up interval	Months between random assignment and first follow-up Months between random assignment and second follow-up

The data analysis included a child level and a program level. Within child care centers there was some additional clustering of outcome data within classrooms (each center could have multiple infant/toddler classrooms), but because the data for centers and family child care homes were analyzed together, it was not possible to include a separate hierarchical linear modeling level in the analysis for these classrooms. A separate Huber-White standard error correction for the classroom-level clustering within child care centers was applied using the robust cluster option in Stata.

The hierarchical linear model used for the analysis of a child-level outcome is illustrated in equations 1 and 2.

$$Y_{ik} = \beta_{0k} + \sum_{1}^{X} \alpha_{Xk} X_{Xik} + \varepsilon_{ik} \quad (1)$$

$$\beta_{0k} = \gamma_0 + \gamma_1 E_k + \sum_{2}^{P} \gamma_P P_{Pk} + \phi_k \quad (2)$$

In these equations, Y_{ik} is the assessment outcome for child i in program k. X_{Xik} and P_{Pk} are vectors of child and program-level background variables (covariates). The child-level equation 1 is a function of a random intercept, β_{0k}, a vector of child-level covariates, X_{Xik}, and an error term, ε_{ik}. β_{0k} is a function of the experimental dummy variable E_k, whose coefficient γ_1 is the main PITC treatment effect in this system of equations, a vector of program-level covariates, P_{Pk}, and an error term, ϕ_k. Estimating these two equations simultaneously ensured that the statistical results were appropriate for nested data. Note that larger strata/sites, particularly child care centers rather than family child care homes, were given more weight than smaller strata/sites in the estimation of the aggregate treatment effect.

At the program level, the impact analysis was structured as a simple single-level model:

$$Y_k = \gamma_0 + \gamma_1 E_k + \sum_{2}^{S} \gamma_S P_{Sk} + \phi_k \quad (3)$$

In this equation, Y_k is the outcome for program k, and P_{Sk} is a vector of program-level background variables. γ_1 captures the main PITC treatment effect. In child care centers, Y_k is a simple average of the outcomes for different classrooms observed in that center. The analysis did not include a separate level of analysis for the subset of programs with multiple classroom observations at each observation wave; it was decided that the statistical benefits of pooling across provider type (centers and family child care homes) outweighed the benefits of having more data points for capturing program outcomes for child care centers, which were highly correlated. The averaging of multiple contemporaneous classroom observations per center made those observations more stable and reliable.

Multiple comparison procedures

As described above, the study team conducted two statistical tests in estimating primary child impacts and two tests in estimating secondary program-level impacts. The study team used the Benjamini and Hochberg (1995) method to adjust *p*-values for multiple comparisons within the primary and secondary outcome domains. This procedure involves ordering from smallest to largest *p*-values obtained for each outcome variable across domains at each level (student and teacher) and multiplying each unadjusted *p*-value by $N/(N - j + 1)$, where N is the number of primary outcome variables within a domain and j represents the order of the test. The null hypothesis is rejected if an adjusted *p*-value is less than 5 percent.

Sensitivity analysis

The study team conducted several types of sensitivity analyses to test the robustness of the results across different models and samples. The following types of sensitivity analyses were conducted:

Impact models with fewer baseline covariates

Two sets of alternative models were estimated which included fewer covariates than the benchmark model described above—the first set includes randomization strata and treatment status as covariates and a second set includes randomization strata, treatment status, and a pretest measure of the outcome variable as a covariate (program-level outcomes only).

Impact models using different missing data handling methods

As shown in Puma, Olson, Bell, and Price (2009), listwise deletion, dummy variable imputation, and multiple imputation may be valid methods of handling missing data in group randomized controlled trials, depending on the type of missing data (posttests, covariates, or pretests). The benchmark model used listwise deletion for missing posttests and dummy variable imputation methods for missing covariates. To examine the possible consequences of using different techniques for handling missing data for the impact estimates, researchers obtained estimates from three models/analytic samples: listwise deletion (multilevel ANCOVA impact estimates based on cases with complete data on posttests, pretests, and covariates); multilevel ANCOVA impact estimates based on listwise deletion of posttests and imputing missing pretests and covariates with dummy variable imputation where necessary; and multilevel ANCOVA impact estimates based on listwise deletion of posttests and imputing missing pretests and covariates with multiple imputation where necessary. These techniques are detailed in appendix O.

Program staff-child interaction outcome impact models

The program quality child care staff-child interaction composite outcome measure comprises constructs assessed via the ITERS-R/FCCERS-R and the PITC-PARS. The PITC-PARS constructs are assessed with sets of questions tapping responsive caregiving practices applicable to all child care settings. However, because these measures were developed by the PITC, there might be concern that the PITC-PARS measures overalign with the intervention. To assess the sensitivity of the program staff-child interaction results, impacts on a staff-child interaction composite measure that excludes the PITC-PARS measures were estimated.

No-show adjustments

In this study, an average of approximately one month elapsed between programs' random assignment to PITC and the date when PITC staff first approached them. Eleven programs assigned to treatment decided against participating before they received any program services. Therefore, a no-show correction was included in the sensitivity analysis, using the method developed by Bloom (1984). This correction is based on the assumption that the overall net impact of an intervention, divided by the percentage of individuals/programs that received any services, is an unbiased estimate of the average impact per service recipient. Once a program received any services, it was counted as a "show," and the no-show correction was not applied even if the program stopped participating in the PITC at a later point.

In addition to adjustments for program-level no-shows, similar adjustments were made for child-level no-shows. Children whose parents volunteered for the study and were in programs randomly assigned to the treatment group—but who left the program within one month of random assignment (or later if before the first PITC service date)—would not have received PITC services and are considered no-shows. Again, such adjustments to the intent-to-treat estimates are included in the sensitivity analyses.

3. Implementing the Program for Infant/Toddler Care

This chapter provides an overview of the Program for Infant/Toddler Care (PITC) philosophy and training activities. It then describes several aspects of its implementation in the study programs: the quality, dosage, and types of services received by caregivers in the treatment group at both the individual level and the program level; the treatment/control differences in training receipt and in implementation of PITC principles as of the first and second follow-up data collections; and children's exposure to settings in which caregivers had received PITC training.

Core PITC policies

The PITC philosophy encourages programs to apply six essential policies in caregiving for children younger than three: primary caregiving; continuity of care; small groups; inclusion of children with special needs; individualized schedules and routines; and cultural responsiveness. The PITC imparts knowledge and adoption of these policies through staff training, technical assistance, and support. Descriptions of the policies are presented below, followed by details of the curriculum and modes of service delivery.[10]

Primary caregiving

A primary caregiver should be assigned to cover each child's day in care. If a child's day is longer than the primary caregiver's, a second primary caregiver is assigned. Primary caregiving includes assignment, routines, team caregiving, and communications with parents. Primary caregivers are assigned so that they have a specific relationship with a fraction of the children that is different than their relationships to other children in the group. Primary caregivers care for the same children each day and carry out caregiving routines with these same children most of the time they are in care. Team caregiving involves two people working in the same room coordinating responsibility for the children. The purpose is to smooth the transition if a primary caregiver is absent or if a caregiver leaves the program or room. Formal or informal communication occurs regularly between caregivers. The primary caregivers regularly communicate with the parent(s) of their assigned children and keep records for each child's development and care.

Continuity of care

When possible, infants and toddlers should remain with the same primary caregiver from the time of enrollment until three years of age. Children can experience a stable, long-term relationship not only with their caregiver but also with each other. Although staff turnover inevitably occurs, the program makes efforts to reduce it and to support continuity of care. Either same-age groupings or mixed-age groupings may be used. With same-age groupings, the environment should be modified and made developmentally appropriate as the children grow older or the whole group moves together to a new space. Children who leave this type of arrangement are usually replaced by newly enrolled children of the same age. In mixed-age groupings, newly enrolled infants and toddlers of any age can be added to the group. The

[10] These descriptions are based on a draft of the *PITC-PARS User's Guide* produced by PITC's parent organization, WestEd (WestEd Center for Child & Family Studies 2007).

environment should be interesting and flexible enough to serve children at varying levels of development.

Group size

PITC group size guidelines are presented in table 3.1. Barriers at least three feet high should separate the groups. Appendix M compares these guidelines with state licensing requirements for group sizes and ratios in California and Arizona. Licensing requirements for staff qualifications are presented as well.

Table 3.1. PITC group size guidelines

Group type	Age	Total number of children in group
Same-age groups	Birth–8 months	6
	8–18 months	9
	18–36+ months	12
Mixed-age groups (at least one child is 24 months)	Birth–36+ months	8

Source: WestEd Center for Child and Family Studies 2007.

Inclusion of children with special needs

The PITC has a written policy of compliance with the Americans with Disabilities Act, and caregivers have been trained in caring for children with disabilities or other special needs. If the program serves children with disabilities or special needs, it has appropriate accommodations. The program does outreach to and encourages enrollment of children with disabilities or other special needs.

Routines

Feeding, diapering, washing, and dressing are carried out as appropriate for children's developmental level and individual needs and, in carrying out these routines, caregivers follow guidelines for health and safety. The primary caregiver carries out most routines for the child and uses them as an opportunity to develop a personal relationship with the child and enhance the child's self-esteem, security, and cognitive and language skills.

Cultural responsiveness

The child care program's philosophy statement or handbook should recognize the importance of connecting the children's cultural and linguistic experiences at home to their experiences in child care. An effort is made to have the child's experiences in child care, including routines and environment, mirror those at home. A caregiver should interview the family about practices and routines in the home. The child care environment reflects the children's cultural experiences. Children are allowed to bring in transition objects, blankets, pictures, toys, or food from home. The language spoken by the children is represented in books in the classroom. Routines followed at home, such as holding children until they fall asleep at naptime, are mirrored in the child care setting.

The PITC curriculum

PITC training topics are embedded in four curriculum modules, forming two course sections with two modules each (table 3.2). Emphasis and topic order may vary somewhat within modules, but the modules are consistently delivered in the order presented below. Training offered includes a minimum of two hours on each of the topics below. In addition to these topics, the PITC offers sessions on understanding the essential policies and philosophy and on understanding the need for quality in infant care. Other optional topics are offered at participant request. Typically, a minimum of two training sessions are scheduled per month.

Table 3.2. The PITC curriculum: sections, modules, and training topics

Section I	Section II
Module 1—Social emotional growth and socialization • Temperaments • Social emotional milestones • Understanding children's behavior • Socialization & guidance • Responsive caregiving	***Module 3—Learning and development*** • Facilitating early learning • Discoveries of infancy • Early brain development • Language, communication, & culture • Special needs
Module 2—Group care • Caregiving routines as curriculum • Primary care • Continuity of care • Environments • Individualized care • Small groups	***Module 4—Culture, family, and providers*** • Culture, family, and providers (ask, acknowledge, and adapt) • Harmonizing cultural diversity • Responding to families in culturally sensitive ways • Protective urges of parents and caregivers

Source: WestEd Center for Child & Family Studies, PITC Partners for Quality brochure n.d.

The PITC training includes presentation of videos demonstrating model practices, a series of readings on each topic area, and guided discussion on each program's current practice and how it could be improved. Trainers provide feedback based on program observations and model effective practices in the classroom. The PITC encourages participants to understand, appreciate, and implement the essential policies, but it does not implement them directly. For example, trainers lead participants in the following activities:

- Discussing and writing about how they might make primary caregiving work, or work better, in their program.
- Discussing how primary caregiving and continuity of care work together.
- Discussing and writing about the six PITC policies and how they support attachment and relationship-based care (if seeking academic credit).
- Rating the current cultural sensitivity of their programs on five specific dimensions.

Delivery of the PITC

Each center or group of 5 to 10 family child care providers works initially with the certified trainer to develop their own plan and schedule for being trained, while adhering to the above guidelines. The full implementation is designed to include delivery of 64 hours of training plus additional hours of technical assistance and support over 10–18 months.

Group training

Child care centers receive all training and support on site. At least five infant/toddler staff per center, including an administrator, are asked to commit to participate. Family child care providers participate in groups of 5 to 10 programs within a 30-mile radius of the assigned trainer. Providers come together for group training sessions, meeting at a provider's home or at a convenient community center or school. Trainers encourage ongoing peer support among group members.

Technical assistance and support

In addition to providing group training, trainers visit each program to model techniques, observe programs, and provide feedback to guide program improvement. Trainers also conduct reflective action planning sessions in which providers reflect on program progress and set program goals. Training assistance hours differ for child care centers and family child care homes as follows:

- Each child care center is offered 40 hours of technical assistance and support (in addition to the 64 hours of training), including observations, meetings with the director, and reflective action planning sessions.
- Each individual family child care home is offered 18 hours of technical assistance and support (6 hours of observation and 12 hours of reflective action planning). Again this is in addition to the 64 hours of training delivered to the family child care group.

Written feedback on papers is provided for all individual participants receiving academic units.

Professional growth incentives

Professional growth incentives are an integral part of the PITC and were not added or modified for the study. Participants who meet specific participation targets for each section receive two academic units or $175; four units or $350 may be earned for completing both sections. The cash incentive is available to participants who complete 75 percent of training hours during unpaid, nonwork hours. Programs may also opt to receive a resource grant rather than the individual cash or academic credit incentives.

For center staff and family child care providers to qualify for professional growth incentives, they must complete 28 hours of training and 4 hours of reflective action planning for each section. Thus, completing 56 hours of training and 8 hours of reflective action planning for the whole intervention is a benchmark for complete dosage for most participants. The exception is center directors, who may earn professional growth incentives for each section by participating in 8 hours of training, 4 hours of reflective action planning and one additional meeting with the trainer; thus the benchmark for full completion for directors is 16 hours of training, 8 hours of

reflective action planning, and two meetings (an initial meeting and a concluding meeting). Details of all professional growth incentive requirements are provided in appendix N.

PITC implementation

The PITC was implemented for the treatment settings on a rolling basis as programs were randomly assigned. On average, the implementation start date (first in-person meeting between trainer and staff) was about two months after random assignment. The average delay before start-up was longer in Arizona than in California, and the average duration of implementation was 13 months in Arizona and 13.9 months in California. Implementation for the first wave of treatment programs began in January 2008. All treatment programs had begun by November 2008 and were completed by December 2009.

Because family child care providers participate in the PITC in groups, family child care homes could not be randomly assigned until a sufficient number of programs were enrolled within a local area. Typically, 5 family child care providers is the minimum required for a group. This would have required that 10 or more family child care homes be enrolled and randomly assigned within a radius convenient for group meetings. For the study, the minimum group size for a family child care group was lowered in Arizona so that groups in rural areas could be formed with as few as 2 or 3 providers. Ultimately, 5 family child care groups were formed in Arizona, with the number of providers per group ranging from 2 to 12. In California, some study providers participated in groups with nonstudy providers.[11] Treatment family child care providers in California participated in a total of 16 groups, with the number of study providers per group ranging from 1 to 6.

In California, the PITC is administered through PITC Partners for Quality, based at WestEd's Center for Child and Family Studies. Experienced certified trainers throughout the state are available, either as staff or on a contract, to implement the PITC in their respective service areas. In Arizona, as in many other states, certified trainers are available but have less experience than those in California, and most have no experience with the site-based form of the PITC (the focus of the study).

The primary study hypotheses applied to both centers and family child care providers, but the two settings differed in structure, size, and caregiver characteristics. Much of the research on child care is either limited to one type (usually centers) or includes separate descriptions of centers and family child care. To allow comparison with other studies, it is important to examine implementation for each setting as well as for the total sample. Some measures have different meanings in the two settings. For example, describing "dosage" in a center involves combining the amounts of participation for several caregivers, but dosage for a family provider typically involves participation of one person. The descriptive information in this chapter is also useful to determine whether it is reasonable to assume comparability of training across child care types, to inform understanding of the experimental tests, and to provide information about implementing training for caregivers from a range of settings.

[11] Nonstudy providers were offered the PITC only after random assignment had been completed in the area. Control group providers were flagged in the PITC database and were excluded.

Qualifications and oversight of certified trainers

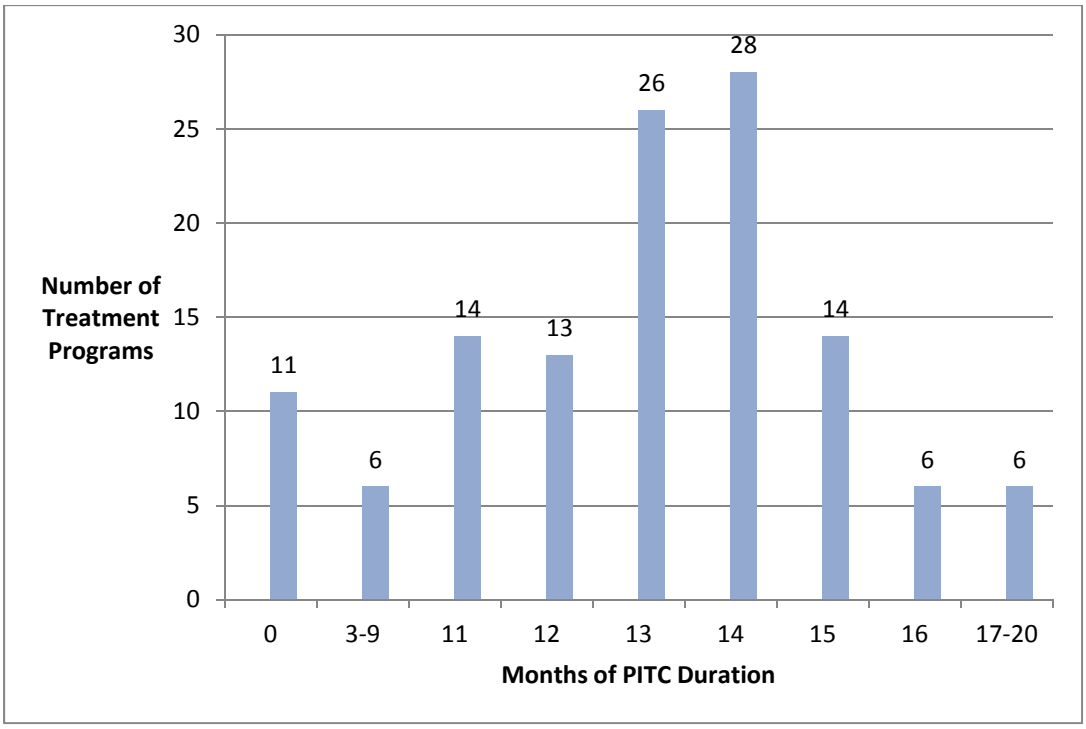
Before the study was implemented, managers of Partners for Quality provided additional training for Arizona trainers who would be working with the study program. This training familiarized the trainers with the intensive, site-based form of the PITC and fostered consistency in procedures across the sites. In addition, Partners for Quality staff observed all trainers in both states twice during implementation to monitor the quality and fidelity of training and provide feedback.

Four trainers in Arizona and 27 in California provided PITC services to study programs. All trainers had been certified through the process described in chapter 1, but Arizona trainers had limited experience in delivering PITC training before the study. Their experience was in delivering a more limited version of the PITC (a series of workshops serving all providers in a region, with no on-site training or support) that had been available in Arizona. California trainers had an average of eight years' experience in PITC training before the study. This experience was with the intensive site-based form of the PITC (the focus of the study). All trainers in both states had experience as a child care provider, teacher, or director. Average experience of trainers in Arizona was 9 years and in California was 22.

Duration of intervention

Treatment programs were encouraged to work with the trainers to develop plans for implementing the full PITC curriculum within 14 months if possible. However, the intervention schedules varied. (As noted in chapter 1, the PITC schedule for each program was customized to accommodate other staff and program commitments. Typically, the timeline ranged 10–18 months.) Average duration of the intervention for programs that completed it was 13 months for family child care homes and 14.4 months for centers. However, duration ranged from 9 months to 20 months for treatment programs that completed the intervention and from 3 months to 16 months for those that participated but did not complete it (figure 3.1).

Figure 3.1. Duration of PITC



Note: n = 124 programs. Eleven programs declined all treatment; six dropped treatment before completion and participated for the following durations: 3 months, 4 months, 6 months, 8 months, 9 months, and 16 months; 107 programs completed treatment and duration ranged from 9 months to 20 months, with an average of 14 months.
Source: PITC information on treatment dates and dropouts

Treatment program and caregiver participation

Eleven programs decided immediately after random assignment that they could not participate in the treatment as planned and declined all services (table 3.3). Six programs began the training but were unable to complete the full course. Reasons included program closures, director and staff turnover, and declining enrollment of infants/toddlers.

Table 3.3. Treatment group participation in PITC

Participation status	All treatment programs	Family child care homes	Centers
Original treatment programs	124 (100 percent)	78	46
Declined to participate	11 (8.9 percent)	#	#
Dropped before completion	6 (4.8 percent)	#	#
Completed	107 (86.3 percent)	66	41

# indicates that the number is withheld because of disclosure risk.
Source: PITC information on treatment dates and dropouts.

Table 3.4 presents sizes of original and final treatment samples for both programs and staff. Individual staff and directors and their programs are defined as "participants," based on records in the PITC participant-tracking database showing one or more participation hours. These records are based on sign-in sheets at all PITC events. (Additional staff may have participated in PITC services provided to these programs; the study includes data only on individuals who consented to participate in the study.) Replacement teachers were those hired or reassigned from other classrooms to replace original sample members who left the programs between study enrollment and follow-up; replacement teachers provided consent to participate at the time of follow-up data collection. Newly hired teachers could enter the PITC at any point; however, they would need to enter midcourse and were less likely to do so later in the process.

Table 3.4. Participation in PITC treatment

Participation status	Total	Family child care	Center child care
Programs			
Assigned to treatment group	124	78	46
Participating in the PITC	113	70	43
Treatment programs receiving the PITC (percent)	91.1	89.7	93.5
Staff			
Assigned to treatment group[a]	544	125	419
Original treatment group staff participating in the PITC	338	83	255
Original treatment staff participating (percent)	62.1	66.4	60.9
Replacement teachers participating[b]	100	6	94
Total sample participating[c]	438	89	349

a. Includes directors, teachers/caregivers, and assistants.
b. These staff provided consent to participate in the study at first follow-up (15 months after random assignment). Program directors were asked to collect consent from new staff hired or were reassigned from other classrooms to replace study participants who had left the programs.
c. Participant data are included for 112 treatment programs. Data from one participating program are missing.
Source: PITC participation database and study sample staff tracking information.

Table 3.5 presents numbers of participants and hours of participation by section for all treatment sample participants, as well as separately for original and replacement participants. Average hours of participation across all three groups of participants were similar for sections 1 and 2 as well as for the full program. However, there were 76 fewer total participants in section 2 than in section 1, reflecting a drop in the number of original participants. Replacement sample members participated almost equally in sections 1 and 2, indicating that most entered early in training. Across family child care and center staff, as shown in table 3.6, average hours of participation again were similar, and the number of participants dropped in section 2 for all categories of participants.

Table 3.5. PITC participation by section for total, original, and replacement sample members

Participation status	All participants		Original participants		Replacement participants	
	Number of participants	Mean hours[a]	Number of participants	Mean hours	Number of participants	Mean hours
All treatment group participants	438	52	338	53	100	54
Section 1	369	30	285	30	84	30
Section 2	293	32	204	33	89	32

a. Includes some meeting hours not allocated to section 1 or section 2.
Source: PITC participation database.

Table 3.6. PITC participation, by section and type of child care

Participation status	Family child care		Centers		Center participants excluding directors		Center directors	
	Number of participants	Mean[a] hours	Number of participants	Mean hours	Number of participants	Mean hours	Number of participants	Mean hours
All treatment group participants	89	52	349	52	299	52	50	55
Section 1	84	29	330	30	281	30	#	32
Section 2	65	32	256	32	215	32	41	29

# indicates number was suppressed due to disclosure risk.
Note: Participants include directors, teachers/caregivers, and assistants.
a. Includes some meeting hours not allocated to section 1 or section 2.
Source: PITC participation database.

Individual completion of the PITC

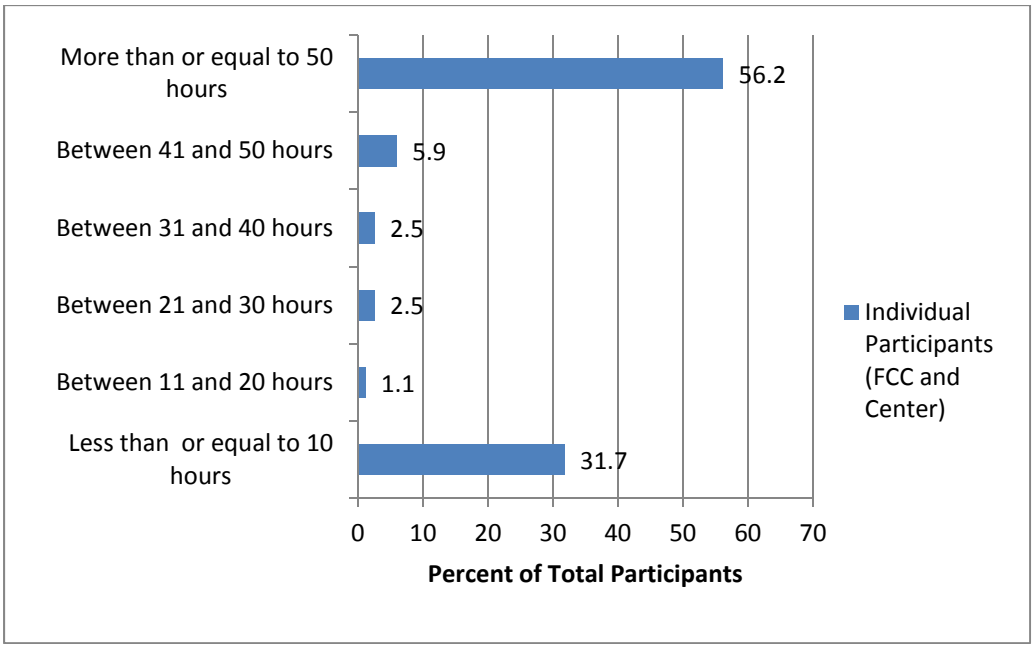
Earning professional growth incentives for a PITC section is a recognition of the individual's completion of that section. Fewer participants earned professional growth incentives for section 2 than for section 1 (table 3.7). Participants who dropped out of the program usually did so after completing 10 or fewer hours (figure 3.2).

Table 3.7. Participants receiving professional growth incentives

Section	Total participants[a] (*n* = 438)		Family child care participants (*n* = 89)		Center participants (excluding directors) (*n* = 299)		Center directors (*n* = 50)	
	Participants receiving incentives	Percent	Participants receiving incentives	Percent	Participants receiving incentives	Percent	Participants receiving incentives	Percent
Section 1	296	67.58	54	60.67	206	68.90	36	72.00
Section 2	181	41.33	41	46.07	120	40.13	20	40.00
Both sections	181	41.33	41	46.07	120	40.13	20	40.00

a. For family child care providers and center staff to earn cash incentives, they must complete 28 hours of training and 4 hours of reflective action planning per section. Center directors are required to participate in 8 hours of training and 4 hours of reflective action planning per section and to attend several additional meetings with the trainer. (See appendix J for more details on professional growth incentives.)
Source: PITC participation database.

Figure 3.2. PITC hours of service received (percent)



Note: n = 345 for child care center staff and n = 89 for family child care staff.
Source: PITC participation database.

Between 60 percent and 70 percent of participants attended training in most topics (table 3.8). A noticeable dropoff occurred after the first training session, on temperaments, which had a 78 percent participation rate. Topics in the last module (culture, family, and providers) had lower levels of participation than earlier modules. Throughout other training topics, participation levels were generally consistent. Average attendance hours per topic were between two and four.

Table 3.8. Participant receipt of PITC training by topic

Training topic[a]	Total participants (*n* = 438)		Family child care staff (*n* = 89)		Center staff (*n* = 349)	
	Number of participants	Percent	Number of participants	Percent	Number of participants	Percent
Temperaments	341	77.85	70	78.65	271	77.65
Social emotional milestones	270	61.64	57	64.04	213	61.03
Understanding children's behavior	285	65.07	57	64.04	228	65.33
Socialization and guidance	275	62.79	54	60.67	221	63.32
Responsive caregiving and identity	303	69.18	62	69.66	241	69.05
Caregiving routines as curriculum	274	62.56	59	66.29	215	61.60
Primary caregiving	311	71.00	63	70.79	248	71.06
Continuity of care	294	67.12	60	67.42	234	67.05
Environments	304	69.41	62	69.66	242	69.34
Individualized care	267	60.96	53	59.55	214	61.32
Small groups	302	68.95	58	65.17	244	69.91
Facilitating early learning	280	63.93	59	66.29	221	63.32
Discoveries of infancy	280	63.93	55	61.80	225	64.47
Early brain development	268	61.19	51	57.30	217	62.18
Language, communication, and culture	272	62.10	58	65.17	214	61.32
Special needs	273	62.33	57	64.04	216	61.89
Culture, family, and providers	272	62.10	54	60.67	218	62.46
Harmonizing cultural diversity	258	58.90	59	66.29	221	63.32
Responding to families in culturally sensitive ways	242	55.25	51	57.30	191	54.73
Protective urges	264	60.27	52	58.43	212	60.74

a. Topics listed in order of delivery.
Source: PITC participation database.

Participation at the program level is an important measure of implementation fidelity. An average of eight staff per center participated in the PITC (table 3.9). All participating centers except for three exceeded the minimum of five staff participants. In most family child care homes, only the lead provider participated. (Additional caregivers participated in only five homes.) Average service hours per participant were 51 for centers and 54 for family child care homes.

However, four or more participants achieved a full 56 hours of training in 41.9 percent of centers (table 3.10). This dosage of training is a benchmark for earning professional growth incentives for both sections (except for center directors). In 86.1 percent of centers and 59.4 percent of homes, at least one participant completed 56 hours of training. In 60.5 percent of centers, three or more participants completed this amount of training.

Table 3.9. Program-level participation in the PITC

Participation	Family child care (*n* = 69 programs)	Centers (*n* = 43 programs)
Average participants per program	1.28	8.12
Minimum	1	2
Maximum	4	18
Program-level hours per participant[a]	54	51
Minimum	1.5	17
Maximum	83	72

Note: Participants of 112 treatment programs are represented in the database. Data are missing from one program that received treatment.
a. Program-level hours per participant equal the sum of all service hours of participants in the center or home divided by the number of participants in the center or home.
Source: PITC participation database.

Table 3.10. Programs that received 56 hours or training or more, by number of participants

Program	Number	Percent
Child care centers (n = 43)		
One or more participant	37	86.1
Two or more participants	28	65.1
Three or more participants	26	60.5
Four or more participants	18	41.9
Five or more participants	13	30.2
Family child care (n = 69 programs)		
One or more participant	41	59.4

Source: PITC participation database.

Treatment/control group training comparison

The "treatment" was the offer of PITC training to the treatment group and not to the control group, but members of either group might also have received other types of training. To determine whether the treatment was associated with actual differences in the overall training experiences of child care staff, follow-up data collection at 15 months after random assignment included a staff survey with a subset of questions on receipt of training, including but not limited to the PITC training, over the past year. (Staff questionnaires were not administered at the second follow-up observation because of response burden issues.)

Treatment center directors (table 3.11), staff (table 3.12), and family child care providers (table 3.13) were more likely than control group members to have received training specifically in infant care over the past 12 months; to have received over 15 hours of training in the past 12 months; and to have received training in cognitive or language development and in physical care. Treatment center directors and staff were more likely than their control group counterparts to have received training in social emotional and physical development. Center staff and family child care providers in the treatment group were also more likely than their control group counterparts to have received training in working with parents.

Table 3.11. Center director education and training at first follow-up, by experimental condition

Director education and training at first follow-up	Intervention (n = 45)		Control (n = 48)				
	Number	Percent	Number	Percent	*t*-statistic	*p*-value	
Degree in early childhood education or related field	24	53.3	26	54.2	–0.08	.937	
Early childhood education or care training in last 12 months	42	97.7	41	85.4	2.17	.034	*
Coursework in early childhood education or related field	19	82.6	18	69.2	1.08	.287	
Training specific to infant care	35	85.4	20	54.1	3.18	.002	**
15 or more hours of training in last 12 months	37	94.9	18	58.1	3.80	.001	**
Member of national state or local professional association for early childhood education	29	65.9	16	33.3	3.27	.002	**
Specific training received in last 12 months							
Early childhood education conference	27	60.0	18	58.3	0.16	.872	
Early childhood education workshop	28	62.2	31	64.6	-0.23	.816	
Early childhood education Child Development Associate training	7	15.6	3	6.3	1.45	.151	
Early childhood education other in-service training	22	48.9	16	33.3	1.53	.130	
Early childhood education college course	11	24.4	14	29.2	–0.51	.612	
Early childhood education adult education	#	#	#	#	#	#	
Early childhood education correspondence course	0	0.0	0	0.0			
Early childhood education no training	#	#	#	#	#	#	
Early childhood education other training	12	26.7	#	#	#	.001	**
Training topic							
Cognitive or language development	36	80.0	26	54.2	2.72	.008	**
Social emotional development	36	80.0	26	54.2	2.72	.008	**
Physical growth and motor skills	32	71.1	24	50.0	2.11	.038	*
Curriculum planning	23	51.1	21	43.8	0.70	.483	
Working with parents	27	60.0	25	52.1	0.76	.448	
Child abuse prevention	15	33.3	14	29.2	0.43	.669	
Health and safety	26	57.8	22	45.8	1.15	.254	
Physical care of children	21	46.7	10	20.8	2.72	.008	**
Discipline practices	21	46.7	17	35.4	1.10	.275	
Other topic	11	24.4	5	10.4	1.78	.078	

# indicates number was removed due to disclosure risk
*p < .05, statistically different from zero at the .05 level, two-tailed test
**p < .01, statistically different from zero at the .01 level, two-tailed test
Source: Center director follow-up questionnaire.

Table 3.12. Center staff/teacher education and training at first follow-up, by experimental condition

Center staff/teacher education and training at first follow-up	Intervention (n = 230)		Control (n = 212)		t-statistic	p-value	
	Number	Percent	Number	Percent			
Degree in early childhood education or related field	49	22.4	44	21.3	0.28	.781	
Child Development Associate credential	58	26.4	39	18.8	1.86	.060	
State-awarded certificates or credentials for early childhood education or related field	50	22.7	45	21.7	0.24	.807	
Early childhood education or care training in last 12 months	209	92.5	187	89.5	1.09	.277	
Coursework in early childhood education or related field	120	65.2	105	57.1	1.61	.109	
Training specific to infant care	193	90.6	131	67.9	5.80	.001	**
15 or more hours of training in last 12 months	144	84.7	70	45.5	8.03	.001	**
Member of national, state, or local professional association for early childhood education	46	20.9	26	12.8	2.24	.026	*
Specific training received in last 12 months							
Early childhood education conference	53	23.0	69	32.6	–2.24	.026	*
Early childhood education workshop	114	49.6	117	55.2	–1.18	.238	
Early childhood education Child Development Associate training	52	22.6	33	15.6	1.89	.059	
Early childhood education other in-service training	86	37.4	57	26.9	2.37	.018	*
Early childhood education college course	73	31.7	56	26.4	1.23	.220	
Early childhood education adult education course	4	1.7	4	1.9	–0.12	.908	
Early childhood education correspondence course	#	#	#	#	#	#	
Early childhood education no training	#	#	#	#	#	#	
Early childhood education other training	43	18.7	15	7.1	3.72	.001	**
Training topic							
Cognitive or language development	181	78.7	120	56.6	5.07	.001	**
Social emotional development	176	76.5	122	57.6	4.31	.001	**
Physical growth and motor skills	166	72.2	109	51.4	4.59	.001	**
Curriculum planning	116	50.4	92	43.40	1.48	.139	
Working with parents	119	51.7	78	36.8	3.19	.002	**
Child abuse prevention	86	37.4	69	32.6	1.07	.287	
Health and safety	129	56.1	101	47.6	1.78	.076	
Physical care of children	107	46.5	67	31.6	3.24	.001	**
Discipline practices	90	39.1	71	33.5	1.23	.219	
Other topic	28	12.2	24	11.3	0.28	.781	

# indicates number was removed due to disclosure risk
*p < .05, statistically different from zero at the .05 level, two-tailed test
**p < .01, statistically different from zero at the .01 level, two-tailed test
Source: Center staff/teacher follow-up questionnaire.

Table 3.13. Family child care provider education and training at first follow-up, by experimental condition

Family child care provider education and training at first follow-up	Intervention (n = 73)		Control (n = 85)		t-statistic	p-value	
	Number	Percent	Number	Percent			
Degree in early childhood education or related field	16	22.9	19	23.5	–0.09	.931	
Child Development Associate credential	19	27.5	21	26.9	0.08	.934	
State awarded credential in early childhood education or related field	15	22.1	17	21.5	0.08	.938	
Early childhood training in last 12 months	59	81.9	64	76.2	0.87	.384	
Coursework in early childhood education or related field	36	59.0	36	52.9	0.69	.492	
Training specific to infant care	52	86.7	47	72.3	2.01	.046	*
15 or more hours of training in last 12 months	45	90.0	41	70.7	2.61	.010	**
Member of national, state, or local professional association for early childhood education	20	29.0	23	28.8	0.03	.975	
Specific training received in last12 months							
Early childhood education conference	26	35.6	40	47.1	–1.45	.148	
Early childhood education workshop	46	63.0	48	56.5	0.83	.407	
Early childhood education Child Development Associate training	21	28.8	19	22.4	0.92	.358	
Early childhood education other in-service training	10	13.7	5	5.9	1.63	.106	
Early childhood education college course	23	31.5	19	22.4	1.30	.197	
Early childhood education adult education course	4	5.5	7	8.2	–0.68	.501	
Early childhood education correspondence course	0	0.0	0	0.0			
Early childhood education no training	#	#	#	#	#	#	
Early childhood education other training	5	6.9	5	5.9	0.25	.805	
Training topic							
Cognitive or language development	54	74.0	48	56.5	2.32	.022	*
Social emotional development	49	67.1	45	52.9	1.82	.071	
Physical growth and motor skills	47	64.4	44	51.8	1.60	.111	
Curriculum planning	35	47.9	28	32.9	1.93	.055	
Working with parents	34	46.6	24	28.2	2.41	.017	*
Child abuse prevention	28	38.4	27	31.8	0.87	.389	
Health and safety	44	60.3	39	45.9	1.81	.072	
Physical care of children	33	45.2	24	28.2	2.24	.027	*
Discipline practices	40	54.8	26	30.6	3.15	.002	**
Other topic	8	11.0	10	11.8	0.16	.875	

# indicates number was suppressed due to disclosure risk.
*p < .05, statistically different from zero at the .05 level, two-tailed test
**p < .01, statistically different from zero at the .01 level, two-tailed test
Source: Family child care provider follow-up questionnaire.

Implementation of PITC policies

The PITC-PARS observational[12] measure was used in the study to measure both staff-child interactions and structural or policy indicators of the PITC essential policies in primary care, continuity of care, cultural responsiveness, and group size. Tables 3.14 and 3.15 compare treatment and control programs, at baseline and for both follow-up periods, on these PITC-PARS measures. These comparisons are shown separately for centers and family child care settings because most are meaningful at the level of individuals within programs. In centers, there are typically several individuals and classrooms as well as a director; in family settings, there is often one caregiver. Some measures have different meanings in each setting. For example, one principle of PITC is assigning a primary caregiver to an infant, but that principle is relevant primarily in centers because most family homes have only one caregiver. Combining the two settings could obscure information about the fidelity and comparability of program implementation.

[12] The PITC-PARS observation is supplemented with review of documents provided by the director (primarily the parent handbook) and clarification questions posed to the director on practices that cannot be observed during the visit (such as those for maintaining continuity of care).

Table 3.14. PITC-PARS measures of program structure for child care center classrooms at baseline, first follow-up, and second follow-up, by experimental condition

	Baseline				First follow-up				Second follow-up			
Program structure	**Treatment group mean** (*n* = 77)	**Control group mean** (*n* = 72)	***t*-statistic**	***p*-value**	**Treatment group mean** (*n* = 70)	**Control group mean** (*n* = 68)	***t*-statistic**	***p*-value**	**Treatment group mean** (*n* = 66)	**Control group mean** (*n* = 67)	***t*-statistic**	***p*-value**
Culturally responsive care	2.60	2.76	0.96	.340	2.11	2.19	0.40	.692	1.97	1.78	–1.06	.292
Philosophy statement/handbook connects home to child care	0.38	0.44	0.84	.404	0.31	0.28	–0.45	.657	0.24	0.09	–2.41	.018 *
Specific practices for culturally consistent care	0.78	0.79	0.18	.855	0.70	0.74	0.46	.648	0.56	0.48	–0.95	.342
Environment reflects children's cultural experiences	0.60	0.67	0.87	.385	0.31	0.43	1.36	.175	0.38	0.37	–0.07	.947
Supportive of cultural preferences and expressions of cultural identity	0.84	0.86	0.29	.773	0.79	0.75	–0.49	.622	0.79	0.84	0.7	.483
Primary caregiving	2.58	2.33	–1.24	.217	2.26	1.82	–1.72	.087	2.44	2.24	–0.85	.397
Primary caregivers assigned	0.44	0.40	–0.48	.635	0.57	0.35	–2.62	.010 **	0.52	0.36	–1.83	.069
Primary caregivers carry out routines for same group of children	0.44	0.40	–0.48	.635	0.51	0.34	–2.11	.037 *	0.49	0.39	–1.12	.264
Familiar caregiver when primary caregiver is absent	0.87	0.88	0.09	.930	0.79	0.81	0.34	.738	0.91	0.99	1.98	.050
Primary caregiver communicates with parents and keeps records	0.83	0.65	–2.53	.012*	0.39	0.32	–0.76	.449	0.53	0.51	–0.26	.794
Continuity of care	2.27	2.40	0.8	.425	2.00	1.77	–1.39	.167	2.12	2.22	0.58	.566
Same-age or mixed-age continuity	0.43	0.51	1.04	.300	0.24	0.16	–1.18	.239	0.14	0.19	0.89	.375
Appropriate procedures for transitioning children to new caregiver	0.52	0.49	–0.4	.686	0.34	0.29	–0.61	.543	0.53	0.60	0.77	.442
Appropriate procedures for	0.51	0.54	0.43	.670	0.44	0.49	0.50	.620	0.55	0.46	-0.95	.344

Program structure	Baseline				First follow-up				Second follow-up			
	Treatment group mean (n = 77)	**Control group mean** (n = 72)	***t*-statistic**	***p*-value**	**Treatment group mean** (n = 70)	**Control group mean** (n = 68)	***t*-statistic**	***p*-value**	**Treatment group mean** (n = 66)	**Control group mean** (n = 67)	***t*-statistic**	***p*-value**
introducing a new child to the group												
Staff professional development activities focus on infant/toddler care	0.82	0.86	0.71	.480	0.97	0.82	–2.95	.004 **	0.91	0.97	1.48	.141
Program structure												
Caregiver: child ratio meets PITC standards	0.39	0.33	–0.71	.478	0.43	0.43	–0.02	.980	0.32	0.22	–1.22	.224
Small groups according to PITC standards	0.18	0.19	0.20	.845	0.34	0.34	–0.06	.955	0.27	0.33	0.70	.488

Note: Items (for example, culturally responsive care) were rated 0 to 4, where 0 indicates that no indicators were met for this item and 4 indicates that all indicators were met for this item. Indicators (for example, philosophy statement/handbook connects home to child care) were rated 0 or 1, where 0 indicates that the indicator was not met, and 1 indicates that the indicator was met. At baseline, 149 classrooms were nested in 92 centers. At first follow-up, 138 classrooms were nested in 85 centers. At second follow-up, 133 classrooms were nested in 82 centers. All observations are treated as independent.

*$p < .05$, two-tailed test.

**$p<.01$, two-tailed test.

Source: Observation data from the child care center baseline, first follow-up, and second follow-up observations.

Table 3.15. PITC Program Assessment Rating Scale measures of program structure for family child care homes at baseline, first follow-up, and second follow-up, by experimental condition

	Baseline				First follow-up				Second follow-up			
Program structure	**Treatment group mean** (n = 78)	**Control group mean** (n = 81)	***t*-statistic**	***p*-value**	**Treatment group mean** (n = 51)	**Control group mean** (n = 51)	***t*-statistic**	***p*-value**	**Treatment group mean** (n = 42)	**Control group mean** (n = 47)	***t*-statistic**	***p*-value**
Culturally responsive care	2.08	1.93	–0.86	.390	2.49	2.33	–0.69	.489	2.67	2.49	–0.80	.427
Philosophy statement/handbook connects home to child care	0.14	0.15	0.13	.899	0.26	0.22	–0.46	.645	0.24	0.28	0.41	.683
Specific practices for culturally consistent care	0.33	0.27	-0.84	.400	0.57	0.49	–0.79	.433	0.69	0.64	–0.51	.608
Environment reflects children's cultural experiences	0.80	0.74	-0.80	.423	0.80	0.78	–0.24	.809	0.81		–0.50	.621
Supportive of cultural preferences and expressions of cultural identity	0.81	0.77	–0.65	.519	0.86	0.84	–0.28	.782	0.93	0.81	–1.66	.100
Primary caregiving	2.21	2.05	–0.51	.609	1.27	1.68	0.93	.358	0.95	1.48	1.02	.315
Primary caregivers assigned	0.61	0.54	–0.59	.560	0.33	0.43	0.74	.464	0.26	0.38	0.78	.440
Primary caregivers carry out routines for same group of children	0.50	0.40	–0.88	.381	0.27	0.32	0.45	.654	0.21	0.29	0.54	.595
Familiar caregiver when primary caregiver is absent	0.51	0.49	–0.23	.819	0.33	0.39	0.46	.644	0.21	0.38	1.17	.251
Primary caregiver communicates with parents and keeps records	0.61	0.73	1.12	.268	0.33	0.54	1.56	.124	0.26	0.43	1.08	.285

Program structure	Baseline				First follow-up				Second follow-up			
	Treatment group mean (*n* = 78)	**Control group mean** (*n* = 81)	***t*-statistic**	***p*-value**	**Treatment group mean** (*n* = 51)	**Control group mean** (*n* = 51)	***t*-statistic**	***p*-value**	**Treatment group mean** (*n* = 42)	**Control group mean** (*n* = 47)	***t*-statistic**	***p*-value**
Continuity of care	2.90	3.07	1.09	.276	3.45	3.33	–0.70	.485	3.57	3.40	–1.09	.280
Same-age or mixed-age continuity	1.00	0.99	–0.98	.328	0.98	0.92	–1.38	.172	0.98	0.89	–1.55	.124
Appropriate procedures for transitioning children to new caregiver	0.58	0.57	–0.11	.909	0.80	0.78	–0.24	.809	0.86	0.83	–0.35	.727
Appropriate procedures for introducing a new child to the group	0.49	0.58	1.17	.242	0.71	0.67	–0.42	.673	0.81	0.75	–0.73	.470
Staff professional development activities focus on infant/toddler care	0.83	0.94	2.10	.037	0.96	0.96	0.00	1.000	0.93	0.94	0.14	.888
Program structure												
Caregiver: child ratio meets PITC standards	0.56	0.62	0.68	.498	0.65	0.67	–0.81	.442	0.38	0.49	1.02	.309
Small groups according to PITC standards	0.71	0.72	0.15	.880	0.67	0.71	0.42	.673	0.52	0.49	–0.32	.749

*$p < .05$.
**$p < .01$.
Note: Items (for example, culturally responsive care) were rated 0 to 4, where 0 indicates that no indicators were met for this item and 4 indicates that all indicators were met for this item. Indicators (for example, philosophy statement/handbook connects home to child care) were rated 0 or 1, where 0 indicates that the indicator was not met, and 1 indicates that the indicator was met. At baseline, *n* = 159 family child care homes (78 assigned to the treatment group and 81 assigned to the control group). At first follow-up, *n* = 102 family child care homes (51 assigned to the treatment group and 51 assigned to the control group). At second follow-up, *n* = 89 family child care home (42 assigned to the treatment group and 47 are assigned to the control group).
Source: Observation data from the family child care home baseline, first follow-up, and second follow-up observations.

At first follow-up, treatment group centers were more likely than control group centers to be implementing primary care assignments, to have primary caregivers carrying out routines for the same group of children, and to have staff development activities focusing on infants/toddlers. At second follow-up, about four months after the PITC was complete, most of these differences had disappeared. Ratings declined for both treatment and control programs on some measures.

Child exposure to the PITC

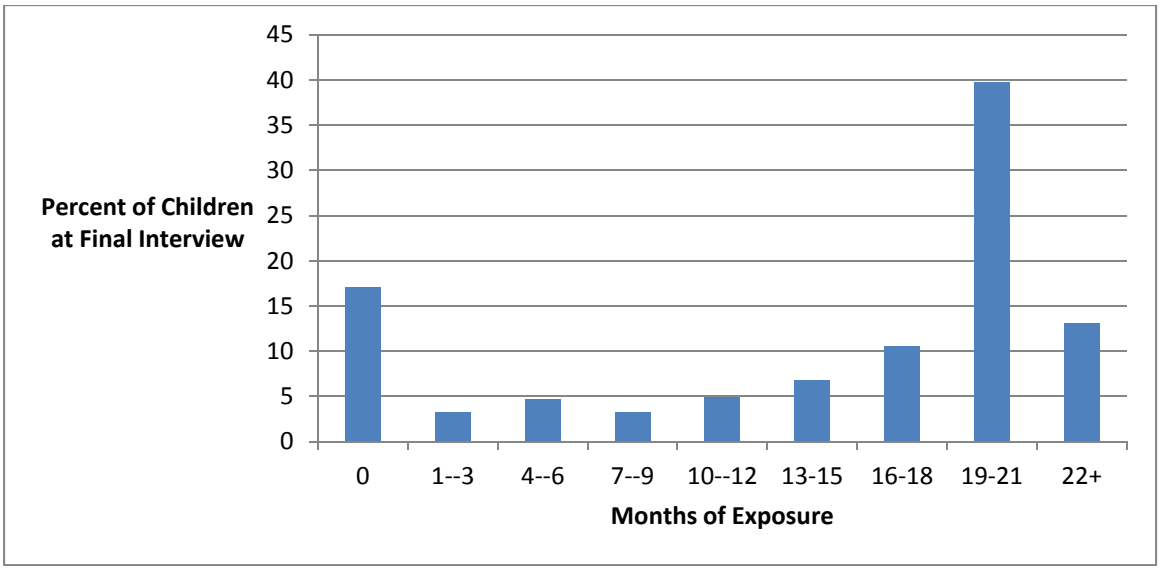
According to the conceptual framework for the study, the potential for PITC-induced child care effects on children occurs six months after completing the full PITC in the child's program. Seventeen percent of treatment children received no exposure to PITC-treated child care, either because their child care programs declined to participate in treatment or because they left the programs before treatment started (table 3.16; figure 3.3). Over 25 percent of children attended the treatment programs for 6 months or less after the PITC training started, and 33 percent attended for 12 months or less. At the time of final child assessment and parent interview, at 23 months after random assignment (on average 19–20 months after the start of treatment), 49.5 percent of treatment children (excluding children whose programs received no treatment) remained in their original child care programs.

Table 3.16. Treatment children's exposure to the PITC

Months between PITC start date and date of final interview or date child left program, if earlier	Total		Children in treatment group family child care homes		Children in treatment group centers	
	Number	Percent	Number	Percent	Number	Percent
0[a]	73	17.1	27	23.1	46	14.8
1–3	14	3.3	#	#	#	#
4–6	20	4.7	#	#	#	#
7–9	14	3.3	4	3.4	10	3.2
10–12	21	4.9	5	4.3	16	5.2
13–15	29	6.8	7	6.0	22	7.1
16–18	45	10.5	9	7.7	36	11.6
19–21	170	39.7	40	34.2	130	41.8
22+	42	9.8	20	17.1	22	7.1
Total treatment group children with child care change data[b]	428		117		311	

# indicates that the number is withheld because of disclosure risk.
Note: "Exposure" is not adjusted for level of program participation or for program dropout before completion
a. Thirty children had no exposure because their programs declined treatment and never participated in the PITC and 43 children had no exposure because they left their original child care programs before the PITC training started.
b. Includes treatment child sample at second follow-up and additional children whose parents reported child care changes at first follow-up and who were not included in second follow-up.
Source: Parent interview at second follow-up and first follow-up on child care changes; PITC staff-provided treatment dates.

Figure 3.3. Illustration of treatment children's exposure to the PITC



Note: n = 428 treatment children (117 in family child care and 311 in centers).
Source: Parent interviews at second follow-up and first follow-up on child care changes; PITC staff-provided treatment dates.

Conclusions

The full PITC curriculum was delivered in treatment programs that followed through with participation. Eleven treatment programs declined to participate, and another six discontinued participation midcourse. An average of eight staff per child care center and one per family child care home participated in some level of services. At the 15-month follow-up, both directors and staff in treatment centers and family programs, compared to control programs, had more training specific to infants, were more likely to have 15 or more hours of training, and had more training in cognitive intellectual development. In centers, they also had more training in socioemotional development and physical growth. Fifteen-month follow-up measures on core PITC policies also showed greater adoption of primary caregiving arrangements in treatment child care centers than control centers, but most differences had disappeared by the second follow-up at four months after the PITC was complete. For children, attrition from their child care programs, as well as the programs' level of participation in the treatment, influenced their exposure to the PITC. Of children in the treatment sample, 49.5 percent remained in their original treatment programs at the time of the second follow-up assessment.

4. Impact Results

This chapter describes the primary impacts of the Program for Infant/Toddler Care (PITC) on the cognitive and behavioral outcomes of children who received child care in the programs assigned to the treatment group. The chapter also presents secondary impacts on the child care program environment—measured with a global program quality score and a staff-child interactions composite score, as described in chapter 2.

The primary cognitive/language and behavioral child outcomes are based on the second follow-up child assessments, which were administered in person to children approximately 23 months after random assignment. Children at second follow-up were between 24 and 58 months of age. The two measures of program quality reported in this chapter are based on the second follow-up observations of child care classrooms by trained observers, at approximately 21 months after random assignment. In centers, 132 of the 149 participating classrooms were observed during the second follow-up for the program observations. Of the 92 centers randomly assigned, 35 had one participating classroom and 57 had two. At the second follow-up, of the 82 centers that participated in the observations, 32 had one classroom that participated in the observations and 50 had two. Of the 159 family child care homes randomly assigned, 90 were observed during the second follow-up.

All child-level impacts (primary impacts) presented in this chapter are based on two-level models that account for the multilevel structure of the data and the potential correlation of within-center error terms that may result from this structure. Center-level impacts (secondary impacts) are based on a linear regression model. The chapter tables display the standard error, a *p*-value, and an adjusted *p*-value for each impact estimate. The standard error indicates the magnitude of the uncertainty of the true mean of each impact, given the number of programs and children in the analysis. The *p*-value indicates the chance of obtaining an impact as large as the estimated impact, if there was no true impact, and the adjusted *p*-value uses the Benjamini-Hochberg (1995) method to account for the number of statistical tests being reported. The resulting adjusted benchmarks are in table 4.1. Impact estimates in this table are considered statistically significant if the adjusted *p*-value is lower than 5 percent, indicating that, accounting for the number of statistical tests, there would be less than a 5-percent chance of obtaining such an impact if there were no true program effect.

Impacts on children's cognitive and language development (primary)

The composite measure of children's language and cognitive development was created by converting the raw scores on the *Preschool Language Scale, Fourth Edition*, Expressive Communication Subscale (PLS-4) and the *Bracken School Readiness Assessment, Third Edition* (BSRA) measures into z-scores and then taking their average (see chapter 2). This composite measure was used to address one of the two primary research questions about the effect of the PITC on child development: What is the impact of the PITC on a composite measure of children's cognitive and language skills, at least 6 months after its full delivery to the children's child care programs (within an average of 23 months after random assignment)?

Table 4.1 displays, by experimental condition, the regression-adjusted group means, difference between treatment and control groups, adjusted *p*-values, and effect sizes for the child cognitive

and language and positive behavior outcomes at second follow-up. The regression-adjusted group means are adjusted using the two-level regression models discussed in chapter 2. The group-mean differences reflect the regression-adjusted program impacts on child outcomes. The *p*-values are adjusted using the Benjamini-Hochberg (1995) correction for multiple comparisons. Finally, the effect sizes are calculated by dividing the group-mean differences by the control group standard deviations.

Children in the treatment group obtained cognitive and language outcomes that were not statistically significantly different from those of children in the control group, as indicated by the adjusted *p*-value of .241 (see table 4.1). The difference in the cognitive and language composite scores between treatment and control group children was –0.083 standard deviations. Thus, there is no evidence that the PITC improved children's development, as measured with this composite outcome. To support interpretation of the composite results, table 4.1 also presents the separate component impacts and adjusted means and effect sizes for the PLS-4 and BSRA measures used to create the language and cognitive development composite, presented as raw scores.

Table 4.1. PITC impacts on child outcomes (primary outcomes)

Child outcome	Adjusted means: Treatment (standard deviation)	Adjusted means: Control (standard deviation)	Difference (standard error)	Adjusted *p*-value	Effect size	Treatment sample size	Control sample size
Language and cognitive development							
Language and cognitive development (composite of PLS-4 and BSRA)	–0.184 (0.729)	–0.116 (0.816)	–0.068 (0.058)	.241	–0.083	402	394
PLS-4 score	44.520 (8.129)	44.966 (8.878)	–0.445 (0.497)	.370	–0.050		
BSRA score	23.474 (17.760)	25.312 (18.861)	–1.838 (1.343)	.171	–0.097		
Positive behavior							
Positive behavior (composite of Polit Positive Behavior Scale and CBCL 1½-5)	–0.050 (0.899)	0.070 (0.751)	–0.121 (0.062)	.106	–0.161	425	408
Polit Positive Behavior Scale	2.954 (0.425)	2.998 (0.374)	–0.032 (0.031)	.295	–0.086		
Externalizing Problems Domain Score from CBCL 1½-5	10.909 (7.571)	9.702 (7.069)	1.207 (0.543)	.026	0.171		
Internalizing Problems Domain Score from CBCL 1½-5	7.200 (5.958)	6.419 (5.352)	0.782 (0.414)	.059	0.146		

Note: PLS-4 is the *Preschool Language Scale, Fourth Edition,* Expressive Communication Subscale.
BSRA is the *Bracken School Readiness Assessment, Third Edition.*
CBCL 1½-5 is the Child Behavior Checklist for children ages 1½-5.
Composite scores are expressed as z-scores, and component measures are expressed as raw scores.
Note: Data were regression-adjusted using a two-level regression model to account for differences in baseline characteristics. Randomization strata were also included in the regression models to account for blocking by geographic region and program type. Effect sizes were calculated by dividing impact estimates by the control group standard deviations of the outcome variables. Missing covariates were imputed using dummy variable missing data imputation methods. To create a positive behavior composite, the externalizing and internalizing domain component z-scores were multiplied by (–1) before taking their average. The child sample for the language and cognitive development impacts include the 798 children with data for the BSRA and the 799 children with data for the PLS-4, less children who had incomplete data for either the PLS-4 or the BSRA. The child sample for the positive behavior impacts include 833 children whose parents completed the parent follow-up questionnaire and therefore had data for the Child Behavior Checklist 1½–5 and the Polit Positive Behavior Scale.
Source: Child assessment data from second follow-up, including the PLS-4 and the BSRA, administered to the children; and the Polit Positive Behavior Scale and the CBCL 1½-5 from the parent questionnaire at second follow-up.

Impacts on children's behavior (primary)

The composite measure of child behavior was created by combining the Positive Behavior Scale and Child Behavior Checklist (CBCL 1½-5) measures, as described in chapter 2. For this analysis, items measuring positive and negative behavior scores were scaled so that the resulting outcome measure represents a progression from negative behavior (more behavior problems, fewer positive behaviors) to positive behavior (fewer behavior problems, more positive behaviors). The child positive behavior outcome results address the second of the two primary research questions about the effect of the PITC on child development: What is the impact of the PITC on a composite measure of parent-reported social and behavioral skills, at least 6 months after its full delivery to the children's child care programs (within an average of 23 months after random assignment)?

Table 4.1 displays, by experimental condition, the regression-adjusted group means, impact and effect size for the child positive behavior composite and the regression-adjusted group means, impacts, and effect sizes for the Positive Behavior Scale and Child Behavior Checklist (CBCL 1½-5) measures used to create the composite.

The table shows that behavior of children in the treatment group was rated less favorably by parents than was behavior of children in the control group. However, the difference between the two groups was not statistically significant. The difference between the treatment and control group children was –0.161 standard deviations. To further elucidate this finding, the regression-adjusted group means and impacts for each of the three behavior components are presented. The p-value for the externalizing behavior component was less than the benchmark p-value of .05, but because these data are presented for clarification and not as formal statistical tests of the hypotheses, they should be interpreted with caution. In any case, there is no evidence that the PITC positively impacted parent-rated social and behavioral development, as measured with this composite outcome.

Impacts on child care program quality (secondary)

Trained observers visited each program to observe the quality of the programs and the quality of interactions between staff members and children, as rated with standardized program assessments for infant and toddler care (see chapter 2). The results of these observations address the two secondary research questions about child care quality:

- What is the impact of the PITC on global child care quality at least 4 months after the intervention ends (within an average of 21 months after random assignment)?
- What is the impact of the PITC on a composite measure of the quality of child care programs' staff-child interactions at least 4 months after the intervention ends (within an average of 21 months after random assignment)?

These questions were addressed by measuring the impact of the PITC on overall program quality, as well as the impact on a subset of quality measures capturing the quality of interactions between program staff and the infants and toddlers in the program. Table 4.2 shows the regression-adjusted group means, groupmean differences, adjusted *p*-values and effect sizes for the program quality outcomes at second follow-up. Programs in the treatment group recorded overall program quality scores that were not statistically significantly different from those of

programs in the control group, as indicated by the adjusted *p*-value of .254, greater than the benchmark *p*-value of .05. The difference between the program quality assessment scores of treatment and control group programs was 0.169 standard deviations. There is no evidence that the PITC improved the overall quality of the infant and toddler child care settings in this study.

Programs in the treatment group recorded staff interaction quality scores that were not statistically significantly different from those of programs in the control group, as indicated by the adjusted *p*-value of .220, greater than .05. The difference between the program quality assessment scores of treatment and control group programs was 0.236 standard deviations. There is no evidence that the PITC improved the quality of staff-child interactions in the infant and toddler child care settings in this study.

Table 4.2. PITC impacts on program quality (secondary outcomes)

	Adjusted means						
Program outcome	**Treatment (standard deviation)**	**Control (standard deviation)**	**Difference (standard error)**	**Adjusted *p*-value**	**Effect size**	**Treatment sample size**	**Control sample size**
Global quality (ITERS-R/FCCERS-R)	3.351 (0.993)	3.198 (0.914)	0.154 (0.134)	–.254	0.169	82	90
Staff-child interactions composite w/ interaction items from ITERS-R/FCCERS-R and PITC-PARS	0.034 (0.693)	–0.129 (0.692)	0.163 (0.102)	.220	0.236	82	90

Note: Data were regression-adjusted using (single-level) linear regression models to account for differences in baseline characteristics. Randomization strata were also included in the regression models to account for blocking by geographic region and program type. Effect sizes were calculated by dividing impact estimates by the control group standard deviation of the outcome variable. Missing covariates were imputed using dummy variable missing data imputation methods. A total of 172 programs had completed observations for the second follow-up observation (82 centers and 90 family child care homes). In the 82 centers, 132 classrooms were observed (32 centers with one classroom and 50 centers with two). In centers where there were two classrooms, the scores were averaged across the two.
Source: Observation data from the second follow-up observations.

Results of sensitivity analyses

A number of sensitivity analyses were conducted to determine the sensitivity of the impact estimates to different variations of the impact model. First, the sensitivity of the measured impacts to the different methods used to impute missing covariate data was tested. As chapter 2 discussed, missing covariates were handled using three different methods:

- Dummy variable adjustment.
- Listwise deletion (deletion of all observations with missing covariates).
- Multiple imputation.

The benchmark impacts in the tables 4.1 and 4.2 used dummy variable adjustment to impute missing covariates. The impacts using listwise deletion and multiple imputation for missing covariates are presented in appendix tables O1 through O4. Overall, the child and center impacts were similar across the three missing data methods. For example, the signs (positive or negative) of the measured impacts were consistent across the three methods. The largest change in effect

size for child impacts was 0.080 standard deviations and the largest effect size change for center impacts was 0.029 standard deviations.

The study team also conducted an analysis of the sensitivity of the secondary, center-level staff-child interaction composite measure to the inclusion of the PITC-PARS measures in the composite. The staff-child interaction impact that includes the PITC-PARS measures is similar to the impact that uses the composite without the PITC-PARS measures; the impacts are both positive, and the effect size for the composite without the PITC-PARS is larger (0.288 standard deviations) than for the composite with the PITC-PARS (0.236), but neither impact is statistically significant. (See appendix table O5.)

Appendix tables O6 and O7 show child and program quality impacts that include adjustments to account for no-shows in the sample. Approximately 4 percent of the child samples for the cognitive and language and behavior outcomes were no-shows. Two percent of the program sample for the program quality outcomes were no-shows. Because there were few no-shows in the child and program samples, the no-show adjustments did not make a significant difference in child and program quality impacts.

A sensitivity analysis was conducted to determine the sensitivity of child and program quality impacts to the exclusion of a small number of crossover sites. Appendix tables O8 and O9 present the impact results using child and program samples that exclude the crossover sites. Removing the crossover sites from the impact sample did not change the child or program quality impacts significantly. The change in the child language and cognitive development impact was 0.002 (–0.068 to –0.066 for the sample that excluded the crossover sites), while the change in the child positive behavior impact was –0.001 (–0.121 to –0.122 for the sample that excluded the crossover sites). At the program-level, the change in the global quality impact was 0.008 (0.154 to 0.162) and the change in the staff-child interactions impact was 0.003 (0.163 to 0.166).

Finally, a sensitivity analysis of the child and program quality impact models to including fewer covariates was conducted. Appendix table O10 presents the results of the child impact model that include only the treatment indicator variable and randomization strata, while tables O11 and O12 present the results of program quality models that include: the treatment indicator variable and randomization strata; and the treatment indicator variable, randomization strata, and baseline program quality score. Child impacts were not sensitive to the covariates in the model. The global quality and staff-child interactions composite program impacts increased and had adjusted *p*-values significant at the 5-percent level for the model that included the treatment indicator variable and randomization strata. However, the program quality impacts for both outcomes were not statistically significant for the model that included the treatment indicator variable, randomization strata, and baseline program quality score.

5. Exploratory Analysis

The exploratory analyses in this chapter examine potential mediators and subgroup differences in Program for Infant/Toddler Care (PITC) impacts. These analyses inform the interpretation of the primary impact analysis results and generate hypotheses for further research.

Child and program impacts at first follow-up

These impact analyses address questions similar to those addressed for the confirmatory analyses but at an earlier measurement time—approximately 14–15 months after random assignment. At this time, in contrast to the timing of the confirmatory impact analyses, treatment for most of the sample programs was not complete. This measurement took place, on average, two months before completion of the average PITC.

Because of this timing, these analyses do not test the central PITC change model: program quality changes, followed by changes in child development, occur several months after the PITC has been delivered. However, these analyses enable us to assess impact using measures that are more proximal to the treatment for children in the sample who left their programs before treatment was completed and for programs that discontinued treatment or closed before treatment was completed. In addition, these earlier potential changes in program quality and child development were mediators of the primary outcomes, which were measured nine months later.

The models for these analyses are identical to those for the primary and secondary analyses. Program measures are also identical. The child measures (other than the language measure) differ from the primary measures, as appropriate for the children's younger ages at this measurement point. Children at first follow-up were between 15 and 45 months of age. A cognitive and language composite measure was constructed using the *Preschool Language Scale, Fourth Edition,* Expressive Communication Subscale (PLS-4) and the *Bayley Scales of Infant and Toddler Development, Third Edition*, Cognitive Subscale (BSID). A positive behavior composite measure was constructed based on four domains of the Infant-Toddler Social and Emotional Assessment (ITSEA). All measures and the construction of the composites are described fully in chapter 2.

Impacts on children's cognitive and language development at first follow-up

There were no statistically significant differences between the program and control groups for either the cognitive and language or positive behavior composite outcomes (table 5.1). For the cognitive and language composite, there was a 0.004 standard deviation difference between the program and control group means, while the program-control group mean difference for the positive behavior composite was –0.061 standard deviations.

Table 5.1. PITC impacts on child outcomes at first follow-up

Child outcome	Adjusted means: Treatment (standard deviation)	Adjusted means: Control (standard deviation)	Difference (standard error)	Adjusted *p*-value	Effect size	Treatment sample size	Control sample size
Language and cognitive development (composite of PLS-4 and BSID)	0.022 (0.947)	0.02 (1.047)	0.004 (0.051)	.933	0.004	394	394
Positive behavior (composite of four ITSEA domains)	0.026 (0.704)	0.067 (0.667)	–0.041 (0.051)	.415	–0.061	414	396

Note: PLS-4 is the *Preschool Language Scale, Fourth Edition,* Expressive Communication Subscale.
ITSEA is the Infant-Toddler Social and Emotional Assessment.
BSID is the *Bayley Scales of Infant and Toddler Development, Third Edition,* Cognitive Subscale.
Data were regression-adjusted using a two-level regression model to account for differences in baseline. The child sample for the language and cognitive development impacts include the 789 children with data for the BSID and the 782 children with data for the PLS-4, less children who had incomplete data for either the PLS-4 or the BSID.
Source: Child assessment data from first follow-up, including the PLS-4 and the BSID, administered to the children; the ITSEA from the parent questionnaire at first follow-up.

Impacts on child care program quality at first follow-up

The impacts on global quality and staff-child interactions at 14 months after random assignment were not statistically significant (table 5.2). The effect size for the global quality outcome was 0.122; the effect size for the staff-child interactions outcome was –0.119.

Table 5.2. PITC impacts on program quality at first follow-up

Program outcome	Adjusted means: Treatment (standard deviation)	Adjusted means: Control (standard deviation)	Difference (standard error)	Adjusted *p*-value	Effect size	Treatment sample size	Control sample size
Global quality (ITERS-R/FCCERS-R)	3.494 (0.938)	3.379 (0.940)	0.115 (0.123)	.353	0.122	94	100
Staff-child interactions composite w/ interaction items from ITERS-R/ FCCERS-R and PITC-PARS	–0.066 (0.602)	0.018 (0.695)	–0.083 (0.089)	.352	–0.119	94	100

Note: Data were regression-adjusted using (single-level) linear regression models to account for differences in baseline characteristics and study design characteristics. Effect sizes were calculated by dividing impact estimates by the control group standard deviation of the outcome variable. Missing covariates were imputed using dummy variable missing data imputation methods. A total of 194 programs had completed observations for the first follow-up observation (85 centers and 109 family child care homes). In the 85 centers, 137 classrooms were observed (33 centers with one classroom and 52 centers with two classrooms). In centers where there were two classrooms, the scores were averaged across the two.
Source: Observation data from the first follow-up observations.

Subgroup analyses at second follow-up

This section presents subgroup impact results for child and program outcomes at second follow-up. Analyses of impacts for child and program subgroups were designed to reflect child development theory and research suggesting that an early intervention might affect children differently depending on their family backgrounds, socioeconomic characteristics, and ages. In addition, child care centers differ from family child care homes in their level of resources, size, and structure. Therefore, they would be expected to experience the PITC in different ways.

The subgroup impact results are exploratory findings intended to identify possible trends and impact modifiers that could be further tested in future research. Therefore, the subgroup impact results should be interpreted with caution, and any conclusions about the impact of the PITC should be based on the confirmatory findings in chapter 4.

Child-level impacts were estimated at second follow-up for subgroups defined by the following baseline characteristics:

- Age (18 months or older / younger than 18 months)
- Child care setting (family child care home / center)
- Parents' education level (high school degree or less / some college or more)

Program-level impacts at second follow-up were estimated for subgroups defined by the following baseline characteristics:

- Program type (family child care home / center)
- Baseline program quality (global quality score below 3.5 on a scale from 1 to 5 / at or above 3.5)

Child- and program-level impacts were estimated for each separate subgroup (tables 5.3–5.7) using regression models identical to the ones used for the confirmatory analyses. For example, at the child level, a hierarchical linear model was used to measure impacts within each child-age subgroup and, at the program level, a linear regression model was used to estimate impacts for each type of care. Differences in impacts on different subgroups were tested for statistical significance using full-sample models that included subgroup-treatment interaction effects of subgroup status (e.g. child age group) with experimental condition; results are noted in the tables.[13]

[13] Subgroup impacts measured using interaction effects are less likely than those measured using separate subgroup impact models to lead to spurious findings due to multiple comparisons, because the number of impact models estimated for each subgroup category is reduced to one. However, this study was not designed with enough statistical power to detect differences in treatment effects across subgroups. Because of this lack of statistical power, the absence of statistically significant subgroup differences using interaction terms does not necessarily imply that impact differences across subgroups do not exist. It does, however, imply that the subgroup impact results are unreliable until the subgroup hypotheses can be tested with further research.

Subgroup impacts on children's language and cognitive development and positive behavior

Impacts on language and cognitive development and on parent-reported positive behavior for children younger than 18 months at baseline were not statistically significant (see table 5.3). For children 18 months or older at baseline, the language and cognitive development impact was not statistically significant, but approached significance ($p < .06$) with an effect size of –0.197. The positive behavior impact was negative and statistically significant, with an effect size of –0.277. These exploratory findings suggest possible negative treatment impacts concentrated among children who were older at baseline. However, there were no statistically significant positive behavior subgroup differences in impacts across child's age at baseline groups.

Table 5.3. PITC subgroup impacts on child outcomes at second follow-up: baseline child's age (younger than18 months and 18 months or older)

	Adjusted means						
Child outcome	**Treatment (standard deviation)**	**Control (standard deviation)**	**Difference (standard error)**	***p*-value**	**Effect size**	**Treatment sample size**	**Control sample size**
Children younger than18 months							
Language and cognitive development (composite of PLS-4 and BSRA)	–0.485 (0.688)	–0.445 (0.718)	–0.040 (0.073)	.587	–0.056	194	211
Positive behavior (composite of Polit Positive Behavior Scale and CBCL 1½-5)	0.005 (0.840)	0.088 (0.775)	–0.083 (0.091)	.364	–0.107	207	219
Children 18 months or older							
Language and cognitive development (composite of PLS-4 and BSRA)	0.086 (0.662)	0.219 (0.757)	–0.149 (0.077)	.055	–0.197	208	182
Positive behavior (composite of Polit Positive Behavior Scale and CBCL 1½-5)	–0.064 (0.950)	0.136 (0.721)	–0.200 (0.097)	.040[a]	–0.277	218	188

Note: PLS-4 is the *Preschool Language Scale, Fourth Edition,* Expressive Communication Subscale.
BSRA is the *Bracken School Readiness Assessment, Third Edition.*
CBCL 1½-5 is the Child Behavior Checklist for children ages 1½-5.
Data were regression-adjusted using two-level regression models to account for differences in baseline characteristics. Randomization strata were also included in the regression models to account for blocking by geographic region and program type. Effect sizes were calculated by dividing impact estimates by the control group standard deviation of the outcome variable. Missing covariates were imputed using dummy variable missing data imputation methods. The child sample for the age subgroup language and cognitive development impacts include the 796 children in the full sample (confirmatory) language and cognitive development impacts, less children with missing age data. The child sample for the age subgroup positive behavior impacts include the 833 children in the full sample positive behavior impacts, less children with missing age data.
a. The impact is significantly different from zero at the .05 level (two-tailed test), but differences in subgroup impacts estimated using a full-sample model with subgroup-interaction terms were not statistically significant.
Source: Child assessment data from second follow-up, including the PLS-4 and the BSRA conducted with the children and the Polit Positive Behavior Scale and the CBCL 1½-5 from the parent follow-up questionnaire.

Language and cognitive development and positive behavior impacts for children in family child care homes were not statistically significant (table 5.4). For children in centers, as well, neither cognitive and language nor behavior impacts were statistically significant, although the behavior findings approached significance ($p < .06$).

Table 5.4. PITC subgroup impacts on child outcomes at second follow-up: child care setting (family child care homes or centers)

	Adjusted means						
Child outcome	**Treatment (standard deviation)**	**Control (standard deviation)**	**Difference (standard error)**	***p*-value**	**Effect size**	**Treatment sample size**	**Control sample size**
Children in family child care homes							
Language and cognitive development (composite of PLS-4 and BSRA)	0.232 (0.703)	–0.174 (0.715)	–0.057 (0.098)	.558	–0.079	108	105
Positive behavior (composite of Polit Positive Behavior Scale and CBCL 1½-5)	–0.135 (0.820)	–0.113 (0.751)	–0.022 (0.118)	.851	–0.029	117	107
Children in child care centers							
Language and cognitive development (composite of PLS-4 and BSRA)	–0.209 (0.740)	–0.134 (0.851)	–0.075 (0.073)	.302	–0.088	294	289
Positive behavior (composite of Polit Positive Behavior Scale and CBCL 1½-5)	0.025 (0.925)	0.188 (0.752)	–0.163 (0.084)	.051	–0.217	308	301

Note: PLS-4 is the *Preschool Language Scale, Fourth Edition*, Expressive Communication Subscale.
BSRA is the *Bracken School Readiness Assessment, Third Edition.*
CBCL 1½-5 is the Child Behavior Checklist for children ages 1½-5.
Data were regression-adjusted using a two-level regression models to account for differences in baseline characteristics. Randomization strata were also included in the regression models to account for blocking by geographic region and program type. Effect sizes were calculated by dividing impact estimates by the control group standard deviation of the outcome variable. Missing covariates were imputed using dummy variable missing data imputation methods. The child sample for the program type subgroup language and cognitive development impacts include the 796 children in the full sample (confirmatory) language and cognitive development impacts. The child sample for the program type subgroup positive behavior impacts include the 833 children in the full sample positive behavior impacts.
Source: Child assessment data from second follow-up, including the PLS-4 and the BSRA conducted with the children and the Polit Positive Behavior Scale and the CBCL 1½-5 from the parent follow-up questionnaire.

For children with parents in the high school or below subgroup, language and cognitive development and positive behavior impacts were not statistically significant (table 5.5). For children with parents in the above high school subgroup, the language and cognitive development impact was not statistically significant, while the positive behavior impact was negative and statistically significant, with an effect size of –0.265. Results from the model using subgroup-interaction terms indicate that there were no statistically significant positive behavior subgroup differences in impacts across parent's education groups.

Table 5.5. PITC subgroup impacts on child outcomes at second follow-up: parents' education (high school or below or above high school)

	Adjusted means						
Child outcome	**Treatment (standard deviation)**	**Control (standard deviation)**	**Difference (standard error)**	***p*-value**	**Effect size**	**Treatment sample size**	**Control sample size**
Children with parents whose highest education level was high school or below							
Language and cognitive development (composite of PLS-4 and BSRA)	–0.491 (0.683)	–0.418 (–0.777)	-0.073 (0.087)	.401	–0.094	108	122
Positive behavior (composite of Polit Positive Behavior Scale and CBCL 1½-5)	–0.272 (1.014)	–0.335 (0.829)	0.062 (0.142)	.659	0.075	107	122
Children with parents whose highest education level was above high school							
Language and cognitive development (composite of PLS-4 and BSRA)	0.009 (0.705)	0.125 (0.787)	–0.116 (0.074)	.121	–0.147	287	259
Positive behavior (composite of Polit Positive Behavior Scale and CBCL 1½-5)	0.004 (0.843)	0.189 (0.695)	–0.184 (0.083)	.027[a]	–0.265	311	271

Note: PLS-4 is the *Preschool Language Scale, Fourth Edition*, Expressive Communication Subscale.
BSRA is the *Bracken School Readiness Assessment, Third Edition.*
CBCL 1½-5 is the Child Behavior Checklist for children ages 1½-5.
Data were regression-adjusted using two-level regression models to account for differences in baseline characteristics. Randomization strata were also included in the regression models to account for blocking by geographic region and program type. Effect sizes were calculated by dividing impact estimates by the control group standard deviation of the outcome variable. Missing covariates were imputed using dummy variable missing data imputation methods. The child sample for the parent education subgroup language and cognitive development impacts include the 796 children in the full sample (confirmatory) language and cognitive development impacts, less 20 children with missing parent education data. The child sample for the parent education subgroup positive behavior impacts include the 833 children in the full sample positive behavior impacts, less 22 children with missing parent education data.
a. The impact is significantly different from zero at the .05 level (two-tailed test), but differences in subgroup impacts estimated using a full-sample model with subgroup-interaction terms were not statistically significant.
Source: Child assessment data from second follow-up, including the PLS-4 and the BSRA conducted with the children and the Polit Positive Behavior Scale and the CBCL 1½-5 from the parent follow-up questionnaire.

Subgroup impacts on child care program quality

The estimated impact of PITC on global quality in family child care homes was not statistically significant, with an effect size of 0.229 (table 5.6). Similarly, the estimated impact on staff-child interactions in family child care homes was not statistically significant, with an effect size of 0.341. For child care centers, the estimated impacts on global quality and staff-child interactions were not statistically significant, with effect sizes of 0.066 and 0.061, respectively.

Table 5.6. PITC subgroup impacts on program quality at second follow-up: child care setting (family child care homes or centers)

	Adjusted means						
Program outcome	**Treatment (standard deviation)**	**Control (standard deviation)**	**Difference (standard error)**	***p*-value**	**Effect size**	**Treatment sample size**	**Control sample size**
Family child care homes							
Global quality (ITERS-R/FCCERS-R)	3.474 (1.133)	3.238 (1.029)	0.236 (0.231)	.309	0.229	42	48
Staff-child interactions composite with interaction items from ITERS-R/FCCERS-R and PITC-PARS	–0.023 (0.744)	–0.266 (0.713)	0.243 (0.162)	.138	0.341	42	48
Child care centers							
Global quality (ITERS-R/FCCERS-R)	3.208 (0.808)	3.158 (0.768)	0.051 (0.153)	.740	0.066	42	40
Staff-child interactions composite with interaction items from ITERS-R/FCCERS-R and PITC-PARS	0.109 (0.636)	0.070 (0.644)	0.039 (0.137)	.775	0.061	42	40

Note: Data were regression-adjusted using two-level regression models to account for differences in baseline characteristics. Randomization strata were also included in the regression models to account for blocking by geographic region and program type. Effect sizes were calculated by dividing impact estimates by the control group standard deviation of the outcome variable. Missing covariates were imputed using dummy variable missing data imputation methods. The program sample for the program type subgroup impacts for the global quality and staff-child interactions outcomes include the 172 programs used for the full sample (confirmatory) impacts. A total of 172 programs had completed observations for the second follow-up observation (82 centers and 90 family child care homes). In the 82 centers, 132 classrooms were observed (32 centers with one classroom and 50 centers with two classrooms). In centers where there were two classrooms, the scores were averaged across the two classrooms.
Source: Observation data from the second follow-up observations.

Estimated impacts on global quality and staff-child interactions in programs with lower baseline quality were not statistically significant, with an effect size of 0.250 for global quality and 0.334 for staff-child interactions (table 5.7). For programs with higher baseline quality, impacts were also not statistically significant, with a global quality effect size of 0.117 and a staff-child interactions effect size of 0.112.

Table 5.7. PITC subgroup impacts on program quality at second follow-up: baseline program quality (lower and higher)

	Adjusted means						
Program outcomes	**Treatment (standard deviation)**	**Control (standard deviation)**	**Difference (standard error)**	***p*-value**	**Effect size**	**Treatment sample size**	**Control sample size**
Lower baseline program quality							
Global quality (ITERS-R/FCCERS-R)	2.995 (0.942)	2.855 (0.785)	0.203 (0.216)	.352	0.250	35	47
Staff-child interactions composite with interaction items from ITERS-R/FCCERS-R and PITC-PARS	–0.208 (0.742)	–0.344 (0.611)	0.204 (0.174)	.247	0.334	35	47
Higher baseline program quality							
Global quality (ITERS-R/FCCERS-R)	3.703 (0.911)	3.588 (0.906)	0.106 (0.199)	.596	0.117	46	42
Staff-child interactions composite with interaction items from ITERS-R/FCCERS-R and PITC-PARS	0.227 (0.583)	0.150 (0.667)	0.075 (0.143)	.602	0.112	46	42

Note: Data were regression-adjusted using two-level regression models to account for differences in baseline characteristics. Randomization strata were also included in the regression models to account for blocking by geographic region and program type. Effect sizes were calculated by dividing impact estimates by the control group standard deviation of the outcome variable. Missing covariates were imputed using dummy variable missing data imputation methods. The program sample for the baseline program quality subgroup impacts for the global quality and staff-child interactions outcomes include the 172 programs in the full sample (confirmatory) impacts. A total of 172 programs had completed observations for the second follow-up observation (82 centers and 90 family child care homes). In the 82 centers, 132 classrooms were observed (32 centers with one classroom and 50 centers with two classrooms). In centers where there were two classrooms, the scores were averaged across the two classrooms. Lower baseline quality is defined as a score lower than 3.5 on an ITERS-R/FCCERS-R global quality scale from 1 to 5; higher baseline quality is a score of 3.5 or above.
Source: Observation data from the baseline and second follow-up observations.

Child mobility analysis

Given that the PITC model of change assumes children's exposure to a full intervention, its effects might have been influenced by high levels of child mobility and wide variation in children's exposure to the treatment. Exploratory analyses addressed impacts related to child-level mobility and the association of children's exposure (duration in treatment setting) with outcomes. These analyses were based on the sample of 837 children whose parents provided child care information at follow-up, which represents 89.4 percent of the original child sample.

For this analysis, "stayers" and "leavers" from the original child care setting were identified using parent reports of changes in child care arrangements at the first and second follow-up interviews. Both mobility and duration in the original child care setting were calculated using a common timeframe, beginning at random assignment and ending 23 months later, which is the average amount of time until the second follow-up occurred.[14] At the time of the second child follow-up assessment, 49 percent of both treatment and control children remained in their original child care setting (table 5.8). The average length of time in the original setting for the full sample was 16.34 months (table 5.9). Children's average number of months in the original setting did not differ by treatment condition, type of care, or child age. (See appendix tables P1 and P2.)

Table 5.8. Mobility of PITC child sample, by experimental condition

	Treatment (n = 428)	Control (n = 409)	Percent of total
Number of children who stayed in or left original care setting			
Stayed in original care setting	210	200	49
Left original care setting	218	209	51
Number of children with short and long durations in original care setting			
Short duration (less than 16 months)	134	154	34
Long duration (16 months or more)	294	254	66

Note: The sample size for child mobility is 837; the sample size for analysis of duration is 836.
Source: Months in original care setting was determined based on parent reports, at both first and second child assessment, of child care change dates.

Table 5.9. Children's months in original child care setting, by experimental condition

Time in original care setting	Treatment mean (standard deviation)	Control mean (standard deviation)	t-statistic	p-value
Months in original care setting	16.59 (7.59)	16.08 (7.5)	–0.96	.336

Note: n = 836.
Source: Months in original care setting was determined based on parent reports, at both first and second child assessment, of child care change dates.

There are several differences at baseline between families who would stay in the original child care setting for the whole study and those who would move to another arrangement by the time of the final child assessment (table 5.10). Children who were stayers were more likely to be

[14] The number of months in the original setting by the time of the second follow-up was calculated using the number of months between random assignment and either the date the parent reported the child left the original setting or the date of the second follow-up child assessment. Thirty-three parents reported that their children had left their setting by the time of the first follow-up but had returned to the original setting by the second follow-up. These children were coded as leavers. Where the interval between random assignment and the second follow-up interview was longer than 23 months, the number of months in the setting was truncated at 23 months for both stayers and leavers; and if parents reported that their children left their original setting 23 months after random assignment, children were recoded as stayers.

White and leavers were more likely to be Hispanic. Stayers were younger, more likely to live in a two-parent home at baseline, in care for more hours per week, and more likely to have siblings than were leavers. Parents of stayers were older, on average, than those of leavers.

Table 5.10. Baseline characteristics for children who stayed in original care settings and those who left by the time of the final child assessment

Measure	Overall	Stayers (standard deviation)	Leavers (standard deviation)	Difference between groups	*t*-statistic	*p*-value
Percent of children who are male	49.4	47.6	51.0	−3.4	−1.03	.302
Child's race/ethnicity						
Percent of children who are White, non-Hispanic	29.6	34.5	25.5	9.0	2.98	.003*
Percent of children who are Black, non-Hispanic	6.2	5.1	7.1	−2.0	−1.27	.204
Percent of children who are Hispanic	53.1	49.3	56.3	−7.0	−2.11	.035*
Percent of children living with two parents	65.1	70.9	60.3	10.6	3.34	.001**
Percent of parents with a bachelor's degree or more education	35.4	38.4	32.8	5.6	1.75	.080
Percent of children where the primary language at home is English	74.6	76.7	73.2	3.5	1.20	.231
Mean child age in months at random assignment	17.5	16.9 (6.97)	18.0 (6.93)	−1.1	−2.51	.012**
Mean hours per week child attends child care program	36.6	37.4 (9.73)	35.9 (10.53)	1.5	2.17	.031*
Mean number of siblings	1.0	1.1 (1.16)	0.9 (0.99)	0.2	2.75	.006**
Mean parent age	29.6	31.1 (6.67)	28.3 (7.45)	2.8	5.87	<.001**
Mean number of hours parent is employed and/or in school	37.8	38.4 (11.35)	37.2 (11.03)	1.2	1.53	.126
Teacher-rated child language and behavior scales						
Child's mean score on language scale	2.7	2.7 (1.18)	2.8 (1.23)	−0.1	−1.27	.204
Child's mean score on behavior scale	1.1	1.0 (0.35)	1.1 (0.42)	−0.1	−1.43	.153

Note: Sample size ranged from 862–908 for parent-rated child and family characteristics and from 841–850 for the caregiver-rated child language and behavior scales, due to item-level missing data.
*Significantly different from zero at the .05 level, two-tailed test.
**Significantly different from zero at the .01 level, two-tailed test.
Source: Parent baseline questionnaire, caregiver baseline child form, and parents' reports of child care changes at first and second child assessments.

One statistically significant difference between treatment and control groups of families whose children stayed in the original care setting across treatment groups emerged: parents in the treatment group were older, on average, at baseline than were control group parents (table 5.11). When baseline characteristics of children who left the original settings were compared across treatment and control groups, no statistically significant differences were found. (See appendix table P3.)

Table 5.11. Baseline characteristics, by experimental condition, for children who stayed in original care settings 22 months after random assignment

Measure	Overall	Treatment (standard deviation)	Control (standard deviation)	Difference between groups	*t*-statistic	*p*-value
Percent of children who are male	47.6	49.1	46.0	3.1	0.63	.532
Child's race/ethnicity						
Percent of children who are White, non-Hispanic	34.5	35.9	33.2	2.7	0.57	.568
Percent of children who are Black, non-Hispanic	5.1	6.6	3.5	3.1	1.47	.144
Percent of children who are Hispanic	49.3	48.6	50.0	–1.4	–0.29	.774
Percent of children living with two parents	70.9	70.8	71.0	–0.2	–0.05	.957
Percent of parents with a bachelor's degree or more education	38.4	39.7	37.1	2.6	0.55	.583
Percent of children where the primary language at home is English	76.7	79.2	74.1	5.1	1.22	.225
Mean child age in months at random assignment	16.9	17.4 (7.30)	16.3 (6.56)	1.1	1.72	.086
Mean hours per week child attends child care program	37.4	38.3 (9.29)	36.5 (10.11)	1.8	1.93	.055
Mean number of siblings	1.1	1.1 (1.09)	1.1 (1.22)	0.0	0.09	.927
Mean parent age	31.1	32.1 (6.57)	30.1 (6.65)	2.0	2.94	.004*
Mean number of hours parent is employed and/or in school	38.4	38.0 (11.88)	38.9 (10.78)	–0.9	–0.81	.417
Teacher-rated child language and behavior scales						
Child's mean score on language scale	2.7	2.8 (1.19)	2.6 (1.16)	0.2	1.88	.061
Child's mean score on behavior problems scale	1.0	1.1 (0.33)	1.0 (0.38)	0.1	1.39	.166

Note: Sample size ranged from 385–389 on the caregiver-rated child language and behavior scales and from 398–414 on parent report of child and family characteristics, due to item-level missing data.
Source: Parent baseline questionnaire, caregiver baseline child form, and parents' reports at second child assessment.

Table 5.12 displays, by condition, the regression-adjusted group proportions, as well as program impacts, standard errors, *p*-values, and effect sizes. These data are based on a logistic regression model predicting whether children remained in the original child care setting for at least 23 months after random assignment. These impacts address the exploratory research question about whether the PITC might decrease mobility if it led to child care that was higher quality or that was perceived by parents as more satisfactory.

Compared with children in the control group, children in the treatment group were no more likely to stay in the original child care setting by the time of the second follow-up assessment. The difference in adjusted proportions of staying in the original setting for the treatment and control group programs was 0.063 standard deviations.

Table 5.12. PITC impacts on child mobility (stayers)

	Adjusted proportions					
Child outcome	**Treatment (standard deviation)**	**Control (standard deviation)**	**Difference (standard error)**	***p*-value**	**Effect Size**	**Sample size**
Child mobility (0=left original setting, 1=stayed until 22 months past random assignment)	0.444 (0.497)	0.413 (0.492)	0.031 (0.052)	.546	0.063	838

Note: Data were regression-adjusted using a two-level logistic regression model to account for differences in baseline characteristics and study design characteristics. Effect sizes were calculated by dividing impact estimate by the control group standard deviation of the outcome variable. Missing covariates were imputed using dummy variable missing data imputation methods.
Source: Parents' reports of child care changes at second child assessment.

Table 5.13 displays child-level subgroup impacts defined by the time spent in the original child care setting (less than 16 months, or 16 months or more) for the composite measures of children's language and cognitive development and positive social behavior (see chapter 2). These analyses address the exploratory research questions about whether the effect of the PITC on child development is moderated by the amount of time the child spent in the treatment setting.

Among children with less time in the original care setting, the treatment group children had scores on language and cognitive development and on parent-reported positive behavior outcomes that were not statistically significantly different from those of children in the control group. For children in the high duration subgroup, language and cognitive impacts were not statistically significant, but there was a nonsignificant trend for the program group to score lower on behavior than the control group. Therefore, there is no evidence to support the notion that the association between PITC treatment and children's development differed for those children with more time spent in the original care setting, as measured by the composite outcomes in this study.

Table 5.13. PITC subgroup impacts on child outcomes: low duration and high duration in original child care setting

Child outcome	Adjusted means: Treatment (standard deviation)	Adjusted means: Control (standard deviation)	Difference (standard error)	*p*-value	Effect size	Sample size
Low duration in original care setting (less than 16 months)						
Language and cognitive development (composite of PLS-4 and BSRA)	–0.048 (0.683)	0.053 (0.871)	–0.101 (0.092)	.263	0.116	265
Positive behavior (composite of Polit Positive Behavior Scale and CBCL 1½-5)	0.315 (0.989)	0.343 (0.812)	–0.029 (0.122)	.815	0.036	289
High duration in original care setting (16 months or more)						
Language and cognitive development (composite of PLS-4 and BSRA)	–0.050 (0.743)	0.028 (0.784)	–0.078 (0.078)	.315	0.402	531
Positive behavior (composite of Polit Positive Behavior Scale and Child Behavior Checklist)	–0.019 (0.855)	0.118 (0.706)	–0.138 (0.078)	.078	0.195	544

Note: PLS-4 is the *Preschool Language Scale, Fourth Edition,* Expressive Communication Subscale.
BSRA is the *Bracken School Readiness Assessment, Third Edition.*
CBCL 1½-5 is the Child Behavior Checklist for children ages 1½-5.
Note: Data were regression-adjusted using (single-level) linear regression models to account for differences in baseline characteristics and study design characteristics. Effect sizes were calculated by dividing impact estimates by the control group standard deviation of the outcome variable. Missing covariates were imputed using dummy variable missing data imputation methods.
Source: Second child assessment measures and parents' reports of child care changes at first and second child assessments.

6. Summary of Findings and Study Limitations

The primary hypotheses were that the Program for Infant/Toddler Care (PITC) would have a positive effect on children's cognitive and social development, as measured approximately six months after completion. The findings are that the PITC did not have a statistically significant effect on a composite measure of children's cognitive/language scores, as measured approximately six months (on average) after completion, and that the PITC did not have a statistically significant effect on children's composite behavior scores, as measured approximately six months after completion. The results of the sensitivity analyses, conducted with two alternative approaches for dealing with missing data, were consistent with these findings.

Because the effect of the PITC on program quality was the hypothesized mediator of its effects on children, a secondary analysis was designed to estimate the effects of the PITC on child care program quality at, on average, four months after completion. Findings of this analysis are that the PITC did not have a statistically significant effect on global program quality, as measured by trained observers administering the Infant/Toddler Environment Rating Scale-Revised (ITERS-R) and the Family Child Care Environment Rating Scale-Revised (FCCERS-R), and that the PITC did not have a statistically significant effect on staff-child interactions, as measured by a composite measure incorporating interactions items from the environment rating scales and from the PITC-PARS. Sensitivity analyses, using alternative approaches for dealing with missing data, eliminating crossovers, and reducing the number of covariates, had consistent findings. An additional sensitivity analysis to test impacts on an alternative staff-child interactions composite that excluded the PITC-PARS items also had consistent findings.

Exploratory analyses were conducted to test PITC effects on children and on programs at an earlier data collection point, on children of different ages, on centers and family child care homes, on children of parents with different education levels, and on children with shorter or longer durations attending the child care setting. The PITC had no significant impacts on children or programs at an earlier data collection point (one year after PITC start-up, on average two months before PITC completion). The exploratory analyses revealed potential negative effects on children's behavior for children older than 18 months, for children in centers, and for children whose parents had higher education levels. No significant impacts were found on any of the program subgroups. Mobility analysis found that PITC impacts did not differ between children with longer (16 months or more) and shorter (less than 16 months) durations in their original child care settings.

Any expected impacts of the PITC presuppose that the participants received the full intervention and that children remained in the settings that received the training. However, treatment programs had lower than expected levels of staff participation. Children's exposure to the treatment was reduced due to children's mobility and lower levels of program participation. Of the 124 child care programs assigned to the treatment group, 11 decided not to participate before receiving any training, and 6 dropped the training midcourse. In 59.4 percent of participating family child care homes, one caregiver received the benchmark 56 hours of training, and in 41.9 percent of child care centers, four or more caregivers (the identified minimum number of participants, plus the director, required for PITC delivery) received at least 56 hours of training. Of children in the treatment sample, 17 percent received no exposure to the PITC either because

they left their original child care programs before start-up or because their programs were among the 11 that declined the intervention after random assignment. In the exploratory analyses, however, there is no evidence of positive impacts on children who remained in their child care settings for a long enough time to be exposed to the trained staff.

Study limitations

This study was carried out in community child care settings. But it was conducted in specific areas of California and Arizona and during an economic recession. Its generalizability is limited by this context and by other features of the study including:

- This study tested a specific implementation model of the PITC, with delivery of 64 hours of training, plus on-site coaching, designed to require between 10 and 18 months for full implementation. The findings should not be generalized to other models of implementation that have different durations and service combinations.
- The study relied on volunteer samples, within the identified regions, of child care providers and families willing to participate in a random assignment study. Recruitment required contacts with many providers who either refused to participate or were not willing or able to get consent from the number of parents (or, in some centers, staff) needed to meet the sample requirements. It is possible that providers and families who participated were different from nonparticipants, and results should not be generalized to the larger population.

A limitation in interpreting this study stems from the widely varying PITC exposure of the children. At least 25 percent of children in the treatment group received minimal or no treatment. As an intent-to-treat study, this evaluation measured effects on all children who enrolled and were randomly assigned, including those who left their child care settings well before the PITC was fully implemented. While this design maintained the study's internal validity, it also reduced treatment exposure. Analysis of children's time periods in care found that 25.0 percent of treatment children left their study programs before start-up, left within 6 months of start-up (implementation required between 9 and 20 months), or attended treatment programs that declined the intervention. These children received minimal or no treatment. A total of 49.5 percent of treatment children remained in their programs for 19 months or more, long enough to experience the potential PITC effects (based on the average study treatment period of 13–14 months), as hypothesized in the conceptual timeline for the study, which posits child effects at 6 months after completion of the intervention.

Implications for future research

Few rigorous studies have been conducted that measure the impacts of child care provider training. Only one random assignment study (Ramey and Ramey 2008), to our knowledge, has focused on infant/toddler caregiver training. Little is known about the effectiveness of various training designs and models in use. The PITC incorporates a number of the features that preliminary research and expert opinion in the field suggest are most likely to have a positive effect: focus on staff-child relationships, on-site consultation, assessment and feedback, and application to practice. However, this study finds no positive main effects and also underscores the difficulties of sustaining participation in an intensive, long-term intervention in a large number of community child care settings across geographically dispersed locations. Bryant

(2009) also suggests problems with fidelity of implementation of a caregiver training model tested in the QUINCE study. The exploratory findings also suggest that children who left the child programs before the end of the study, and therefore had less exposure to the PITC, were different demographically than those who stayed.

Additional research on the PITC and other training programs should address questions such as:

- What program and caregiver characteristics are associated with more faithful and complete participation and implementation? Are particular trainer strategies associated with more sustained provider participation, and do these differ for family child care homes and centers? Would the PITC have an impact in a more condensed and tightly controlled format, or with targeting to low-income children? Findings of other rigorous studies indicate that a greater emphasis on classroom mentoring might increase the likelihood of impact (Boller 2010; Bryant et al. 2009).
- What are the caregiver behaviors that are most likely to have a positive impact on children's language and socioemotional development, and can the PITC and other training programs be modified or refocused to improve these specific caregiver behaviors? Future research should track changes in caregiver language interactions and sensitivity more specifically and with more nuance than is possible with the global quality indicators and PITC-PARS items. Future research should track how children behave in the setting to determine whether training-induced changes in caregiver language interactions and sensitivity lead to children's language production and comprehension, positive peer interactions, and positive adult-child interactions.

Although the issues of incomplete staff participation and child mobility raise questions about interpretation of the findings, these are features of the real world of child care. In future research it would be useful to conduct tests of training efficacy separately from the effectiveness of training in a community setting, to test whether the PITC leads to changes in caregiver behavior under controlled conditions. Once it is established that caregivers are well-trained and are using PITC techniques, random assignment of children to PITC and non-PITC settings would provide a clearer test of the impact of these techniques on child development. Overall, increased understanding of the transfer between training strategies, program quality, and child development would inform improved child care training design and implementation.

Appendix A. Sample Power Estimates

Introduction

This appendix includes estimates of the statistical power of the impact estimates presented in this report. The statistical power of a study is its ability to reliably detect true impact estimates of a given size. Specifically, in designing the study, the study team, in consultation with program staff at the Institute for Education Sciences, decided that the study should have power of 80 percent to detect a true effect of 0.20 standard deviations in child outcomes with 95 percent confidence. That is, the sample would have to be large enough that at least 80 percent of impact estimates of 0.2 standard deviations or larger would have a p value of 0.05 or smaller. This effect of 0.2 standard deviations is known as the study's minimum detectable effect size (MDES). The study team specified a minimum detectable effect size of 0.4 for provider-level outcomes, such as the various measures of provider quality. Both of these target minimum detectable effect sizes were chosen after review of extant literature on programs to improve the quality of early childhood education and related child outcomes. They represent a careful trade-off of the need to identify small program effects that were meaningful given the expected treatment contrast and service intensity and the need to keep the study affordable and feasible.

The statistical power calculations were conducted with Optimal Design software (Raudenbush and Liu, 2000). In addition to the size of the sample, factors that determine statistical power are the extent to which child-level observations are clustered within providers and the extent to which baseline variables (such as the baseline observation of provider quality) predict subsequent variation in the outcomes (the R^2 of the impact regression). The study team used .3 as a reasonable estimate of R^2. The study team did not have reliable data with which to estimate the intra-class correlation of child outcomes within providers. We therefore calculated statistical power for two scenarios: an intra-class correlation of 0.1 and an intra-class correlation of 0.2.

A priori power estimates

Using these assumptions, the study team in consultation with the Institute of Education Sciences established a target sample size of 240 providers (90 centers and 150 family child care homes). Table A1 shows that such a sample would result in minimum detectable effect sizes at the child level of 0.20 for the full sample, 0.25 for children in child care centers, and 0.29 for children in family child care homes. At the provider level, these minimum detectable effect sizes were 0.37 for the full sample, 0.56 for child care centers, and 0.48 for family child care homes. All of these estimates included a 20 percent expected study attrition rate for children and a 10 percent expected attrition rate for providers. These were estimates based on past experience of the team members with study retention; however, the estimated program attrition rate was based on experience with child care centers and proved to be a low estimate for family child care programs.

Table A1. Estimated minimum detectable effect sizes for original sample

Outcome	*J*	*n*	*N*	MDES 1	MDES 2
Child outcomes (cluster random assignment)					
Full sample	240	4.9	936	0.20	0.24
Child care centers	90	8	576	0.25	0.33
Family child care homes	150	3	360	0.29	0.33
	J		***N***	**MDES**	
Provider outcomes (simple random assignment)					
Full sample	240		216	0.37	
Child care centers	90		81	0.56	
Family child care homes	150		135	0.48	

J = number of clusters; *n* = number of children per provider; *N* = number of follow-up data points; MDES 1 = minimum detectable effects size for an intraclass correlation of 0.1; MDES 2 = minimum detectable effect size for an intraclass correlation of 0.2.
Note: A 20 percent attrition rate was estimated for children and a 10 percent attrition rate for providers. Attrition rates were estimated separately. Researchers followed up with individual children even if observations could not be conducted at the provider level.
Source: Calculations using Optimal Design Software (Raudenbush and Liu 2000)

Realized minimum detectable effects

Tables A2 and A3 show the actual realized statistical power of the analyses presented in this report. At 0.18 and 0.16 for child outcomes, and 0.35 and 0.37 for provider outcomes, the minimum detectable effect sizes were similar to those used to design the study sample and smaller than the target of 0.2 for child outcomes.

Table A2. Child-level minimum detectable effect sizes for final study sample (second follow-up)

Outcome	Child sample size	Number of programs	Actual intraclass correlation	R-squared of covariates on outcomes at program-level	R-Squared of outcomes on blocking variable	Average number of programs per block	Alpha	Power (percent)	Minimum detectable effect size
Language and cognitive development (composite of PLS-4 and BSRA)	833	229	0.344	0.398	0.063	4	0.05	80%	0.163
Positive behavior (composite of Positive Behavior Scale and CBCL 1½-5)	833	229	0.026	0.048	0.036	4	0.05	80%	0.185

Note: PLS-4 is the *Preschool Language Scale, Fourth Edition,* Expressive Communication Subscale.
BSRA is the *Bracken School Readiness Assessment, Third Edition.*
CBCL 1½-5 is the Child Behavior Checklist for children ages 1½-5.
Source: Sample tracking and calculations using Optimal Design Software (Raudenbush and Liu 2000).

Table A3. Program-level minimum detectable effect sizes for final study sample (second follow-up)

Outcome	Program sample size	R-squared of program-level covariates on outcomes	Alpha	Power (percent)	Minimum detectable effect size
Global quality (ITERS-R/FCCERS-R)	172	0.360	0.05	80	0.345
Staff-Child Interactions composite with interaction items from ITERS-R/FCCERS-R and PITC-PARS	172	0.313	0.05	80	0.369

Source: Sample tracking and calculations using Optimal Design Software (Raudenbush and Liu 2000).

Appendix B. Child Care Provider Screening Interview (California)

(This interview is being conducted as a phone interview with child care providers. The script for the interview is embedded in the interview below.)

INSTRUCTIONS FOR THE INTERVIEWER

--All text in brackets [ ] are instructions for you, the interviewer, and are not meant to be read aloud.

--All text that is in mixed case should be read aloud. This includes text in mixed case that appears in response options.

--All text that is in all CAPS should not be read aloud. If the all CAPS are in a response option only mark that option if the provider says that as an answer, do not read it aloud to the provider. If the all CAPS are in brackets [ ] they are instruction about what you should say. If the all CAPS are in brackets, for example, [YOUR NAME], say your name and not the words "your name."

--If the provider answers a question with a response that is not given, in most questions you can write what they do say in the space provided. For example, if the provider says, "I don't know" in response to the question, "How many children are enrolled in your program," you can enter the words "don't know" into the space provided for the number of children. The same is true if a provider refuses to answer a question. Please just enter the word "refused."

--When you are done with each screen please click on NEXT at the bottom of the page.

--If a provider changes his/her mind about a response you may go back to a previous response by clicking on PREV at the bottom of the page.

--Once you click DONE or EXIT at the end of the survey you will not be able to go back and change the responses.

BEFORE CALLING THE PROVIDER

[Interviewer should enter information on this page before making the phone call.]

1. **Please enter your ID number.**

   This is the number that you were assigned when you were given the link to this screening interview.

   PITC staff ID number _____

2. **Is this your first attempt at calling this provider?**

   YES
   NO, Please specify how many times you have called this provider _____

3. **Is this a child care center or a family child care program?**

Child care center
Family child care program →*SKIPS TO INTRODUCTION BEFORE QUESTION 40*

BEGIN CENTER SCREENING PHONE CALL

Hi, this is [NAME], I am calling about the Study of the Program for Infant Toddler Care, may I please speak with the center director or administrator?

[If the director is not available please ask when a better time to call would be and ask to leave a message for her/him. Leave your name and let the person who answered the phone know that you are calling about the Study of the Program for Infant Toddler Care and that you will call back at a more convenient time but leave your number. Click EXIT at the upper right corner of this page to leave the screening interview.]

BEGIN INTERVIEW WITH CENTER DIRECTOR

[If the Director did not answer the phone, start below.]

Hi, this is [NAME], I am calling about the Study of the Program for Infant Toddler Care.

[If Director answers the phone start below.]

I got your child care center's contact information from [STATE WHO OR WHERE YOU GOT THEIR CONTACT INFORMATION FROM].

Your child care program was selected as a potential participant in this important research study because it is located within the study region and is licensed to provide child care to children younger than 2 years of age.

I would like to talk to you about your potential participation in this study. This should take about 15 minutes.

4. **Do you have time to discuss this now?**

NO
YES → *SKIP TO QUESTION 6*

5. **When would be a good time for me to call back to discuss this important research study?**

Record time]

Thank you. I will call back then.

[Exit survey. Click on Exit Survey in the upper right corner of this screen. Do not press Next. End phone call.]

6. **About a week ago you were sent information about this study. Did you receive this information?**

[IF PROVIDER ANSWERS NO, SAY: Okay. I will send you that information again.]

NO
YES

Okay, let me review or briefly tell you a little about the study. Then I would like to ask you a few questions about your child care program.

[Please pause and answer questions that the child care provider may have about anything you say below.]

The Study of the Program for Infant Toddler Care is a research project sponsored by the Regional Education Laboratory West and the U.S. Department of Education. This study will help policy makers understand how the Program for Infant Toddler Care, or the PITC, a highly regarded caregiver training program, helps children grow and learn.

The PITC provides on-site training and technical assistance in caregiving practices.

This valuable PITC training is free for participants in the study. In addition, there are professional growth incentives, including either academic credits, $350 in cash, or resource materials, for completion of the PITC curriculum.

About half of the programs in the study will be selected to receive the PITC training over the next year. Other programs will be given priority to receive the PITC in 2009.

Participation in the Study of the PITC is also free. In addition caregivers will receive a $25 gift card for participating in the study in 2007 and again in 2008.

Your program's participation is important. The more programs that can participate in the study the more accurate our conclusions about the impact of the Program for Infant Toddler Care will be.

Do you have any questions or concerns about this study? [Answer questions and try to relieve concerns]

I work for the Program for Infant Toddler Care and can answer any questions you have about the training program as well. Do you have any questions about the training program?

7. **Would you be interested in participating in this study?**

NO
YES → *SKIP TO INTRODUCTION BEFORE QUESTION 9*

8. Why not? What are your concerns about participating in this study?

_____

_____

END INTERVIEW

Thank you for letting me take the time to discuss this study with you.

INTRODUCTION TO QUESTION 9

Great!

I have a few questions I would like to ask you about the child care you or your program provides. Then, I will call you back sometime next week to invite you to attend a meeting to discuss the PITC and the Study of the PITC. At this meeting you may sign up to have your program participate in the study.

All of the answers to these questions that you provide will be kept confidential to the fullest extent possible allowed by law. We are interested in having your complete contact information so that we may contact you in the future. All of the information that we obtain from you will be used only for the purpose of this study.

First, I would like to ask you about your primary language and then I have some questions about how to contact you in the future.

9. What is your primary language?

English
Spanish
DON'T KNOW
REFUSED
Other (please specify) _____

10. What is your name?

FIRST NAME _____
LAST NAME _____

11. What is your job title?

_____

12. What is the best way for us to contact you in the future?

Direct Phone
Mobile/Cell Phone
Email
Other (please specify)

13. What is that number or address?

_____

14. What is the name of this child care center?

[Note to interviewer: The name here is the name of the individual program. For example, if it is a Kindercare or Bright Horizons Center, we want the site-specific name (e.g., Little Angels Kindercare).]

_____

15. What is the address of this center?

[The interviewer must enter something on each line below. Please enter the county yourself if you know it. If you do not know the country, enter "unknown."]

STREET _____
CITY _____
STATE _____
ZIP _____
COUNTY [if known] _____

16. Would you like to designate one staff person at your center to be the primary contact for this study other than yourself?

NO → SKIP TO QUESTION 21
YES

17. What is that person's name?

FIRST NAME _____
LAST NAME _____

18. What is that person's job title?

_____

19. What is the best way for us to contact that person in the future?

Direct phone
Mobile/cell phone
Email
Other (please specify) _____

20. What is that phone number or address?

_____

21. Is your center non-profit or for-profit?

Non-profit
For-profit
DON'T KNOW
REFUSED

22. Is your program independent or is it sponsored by another organization?

Examples of organizations that might sponsor a child care program are a church or government agency

Independent → *SKIP TO QUESTION 24*
Sponsored
DON'T KNOW

23. What type of organization sponsors your child care center?

Head Start
Early Head Start
Social service organization or agency
Church or religious group
Public school
Private school
College or university
Private company or individual
Non-government community organization
State or local government
Other (please specify) _____

24. What are the funding sources for your center?

[Please mark all that apply or all answers that the provider indicates are funding sources.]

Parents/guardians who pay the full fee or amount for care
State/Federal subsidies
Early Head Start
Head Start
Other (please specify) _____

25. When is the center open? What are your hours and days of operation?

Hours [e.g., 9am-6pm] _____
Days [e.g., Monday-Thursday] _____

26. Now I have some questions about the current enrollment at your center.

How many children are currently enrolled at your center?

_____

27. How many children is your center licensed to provide care for?

_____

28. How many children aged 3-24 months are currently enrolled at your center?

_____

29. Of the children age 3-24 months currently enrolled at your center, how many are enrolled 20 hours per week or more? Your best estimate is fine.

_____

30. How many of the children age 3-24 months who are currently enrolled at your center do you expect will remain enrolled until they are 3-years-old?

_____

31. In total, how many classrooms serve children ages 3-24 months at your center?

_____

32. Now, I have some questions about the staff at your center.

In total, how many paid staff work in classrooms with children younger than 36 months?

Paid staff includes all full-time and part-time staff. Include all caregivers, aides, assistants, directors, and other staff who work directly with the children. Do not include bus drivers, cooks, or other staff who do not work directly with the children.

_____

33. What is the language USUALLY spoken in the infant or toddler (children younger than 36 months) classrooms?

This is the language most often used by caregivers when speaking to the children.

English
Spanish
Other (please specify) _____

34. Are other languages spoken by caregivers in the infant or toddler classrooms?

NO
YES. Please specify the languages. _____

35. Please think about the staff members who work directly with the children. How many have you hired in the last 12 months?

Please include only caregivers, aides, assistants, directors, and other staff who work directly with the children. The person hired does not have to still be employed to be included in the count. What is your best estimate?

_____

36. Think about the staff members who work directly with the children. How many have left the center in the last 12 months?

Please include only caregivers, aides, assistants, directors and other staff who work directly with the children.

_____

37. I just have a few more questions about your center.
About how long has this center been operating?

ENTER YEARS _____

38. Is the center planning any major changes, such as expansions, cutbacks, or reorganizations, within the next year?

NO
YES. Please explain. _____

39. Has the center experienced any major changes, such as expansions, cutbacks, or reorganization, within the past year?

NO
YES. Please explain. _____

END CENTER SCREENING INTERVIEW

This completes this interview. Thank you for taking the time to answer these questions. I will be calling in the next week to let you know about the meeting to learn more about participating in the study.

BEGIN FAMILY CHILD CARE PROGRAM PHONE CALL

Hi, this is [NAME], I am calling about the Study of the Program for infant Toddler Care, may I please speak with the owner or licensee of this child care program?

[If the owner or licensee is not available please ask when a better time to call would be and ask to leave a message for her/him. Leave your name and let the person who answered the phone know that you are calling about the Study of the Program for Infant Toddler Care and that you

will call back at a more convenient time but leave your number. Click EXIT at the top right hand corner of the screen to leave the interview.]

BEGIN FAMILY CHILD CARE PROGRAM INTERVIEW

[If someone other than the owner or licensee answers start below.]

Hi, this is [NAME], I am calling about the Study of the Program for Infant Toddler Care.

[If the owner or licensee answered the phones start below.]

I got your contact information from [STATE WHO OR WHERE YOU GOT THEIR CONTACT INFORMATION FROM].

Your child care program was selected as a potential participant in this important research study because it is located within the study region and is licensed to provide child care to children younger than 2 years of age.

I would like to talk to you about your potential participation in this study. This should take about 15 minutes.

40. Do you have time to discuss this now?

NO
YES → *SKIP TO QUESTION 42*

41. When would be a good time to call back to discuss this important research study?

[Record time]

Thank you. I will call back then.

[Exit survey. Click on Exit Survey in the upper right corner of this screen. Do not press NEXT. End phone call.]

42. About a week ago you were sent information about this study. Did you receive this information?

[IF PROVIDER ANSWERS NO, SAY: OKAY, I will send you that information again.]

NO
YES

Okay, let me review or briefly tell you a little about the study. Then I would like to ask you a few questions about the child care you provide.

[Please pause and answer any questions the child care provider may have about anything you say below.]

The Study of the Program for Infant Toddler Care is a research project sponsored by the Regional Education Laboratory West and the U.S. Department of Education. This study will help

policy makers understand how the Program for Infant Toddler Care, or the PITC, a highly regarded caregiver training program, helps children grow and learn.

The PITC provides training and technical assistance in caregiving practices.

This valuable PITC training is free for participants in the study. In addition, there are professional growth incentives, including either academic credits, $350 in cash, or resource materials, for completion of the PITC curriculum.

About half of the programs in the study will be selected to receive the PITC training over the next year. Other programs will be given priority to receive the PITC in 2009.

Participation in the Study of the PITC is also free. In addition, all caregivers will receive a $25 gift card for participating in the study in 2007 and again in 2008.

Your participation is important. The more child care providers that can participate in the study the more accurate our conclusions about the impact of the Program for Infant Toddler Care will be.

Do you have any questions or concerns about this study? [Answer any questions that the provider may have.]

I work for the Program for Infant Toddler Care and can answer any questions you have about the training program as well. Do you have any questions about the training program?

43. Would you be interested in participating in this study?

NO
YES → *SKIP TO INTRODUCTION BEFORE QUESTION 45*

44. Why not? What are your concerns about participating?

_____

END INTERVIEW

Thank you for letting me take the time to discuss this study with you.

INTRODUCTION BEFORE QUESTION 45

Great!

I have a few questions I would like to ask you about the child care you or your program provides. Then, I will call you back sometime next week to invite you to attend a meeting to learn more about the PITC and the Study of the PITC. At this meeting you may sign up to have your program participate in the study.

All of the answers to these questions that you provide will be kept confidential to the fullest extent possible allowed by law. We are interested in having your complete contact information

so that we may contact you in the future. All of the information that we obtain from you will be used only for the purpose of this study.

First, I would like to ask you about your primary language and then I have some questions about how to contact you in the future.

45. What is your primary language?

English
Spanish
DON'T KNOW
REFUSED
Other (please specify) _____

46. What is the name of your child care program?

[Note to Interviewer: Some programs may not have names. If the program does not have a name just enter the word 'none'.]

_____

47. What is your name?

FIRST NAME _____
LAST NAME _____

48. What is the best way for us to contact you in the future?
Direct phone
Mobile/cell phone
Email
Other (please specify) _____

49. What is that phone number or address?

_____

50. What is the address of your child care program?

Again, we just need this information so that we can contact you in the future about the study or send you more information

STREET _____
CITY_____
STATE _____
ZIP _____
COUNTY [if known] _____

Now I have some more specific questions about your program or the care that you provide.

51. When do you provide care? During what hours and on what days do you provide care?

Hours [e.g., 9am-6pm] _____
Days [e.g., Monday-Thursday] _____

Now, I have some questions about the children in your care.

52. How many children are currently in your care?

_____

53. How many children are you licensed to provide care for?

_____

54. How many children in your care are subsidized?

[If provider does not understand what you mean by subsidized, please explain. By subsidized we mean a government agency pays part or all of the fee that you charge for the care of a child or children.]

[If provider answers 'no children', enter ;0' (zero) and move on to the next question.]

_____

55. How many children in your care have parents who pay the full amount for that care?

[If provider says that she/he does not charge for care please indicate that on the line below.]

_____

56. How many children ages 3-24 months are currently in your care?

_____

57. Of the children age 3-24 months currently in your care, how many do you care for 20 hours per week or more? Your best estimate is fine.

_____

58. How many of the children age 3-24 months who are currently in your care do you expect will remain in your care until they are 3-years-old?

_____

59. Now, I have some questions about you and other caregivers or staff you might have for this child care program. How many total caregivers, including yourself, work in the program?

Caregivers include all people over the age of 16 that work either full-time or part-time directly with the children. Include all caregivers, aides, assistants, and others who work directly with the children.

ONE [Provider or Owner is the only caregiver]
MORE THAN ONE CAREGIVER. ENTER NUMBER. _____

60. What is the language USUALLY spoken with infants or toddlers (children younger than 36 months) in your care?

Please tell me the language you (and other caregivers if there are any) use most often when speaking with these younger children.

English
Spanish
Other (please specify) _____

61. In addition to the language you (and other caregivers) speak in most often with infants or toddlers in your care, do you (and/or other caregivers) speak in other languages with these children?

NO
YES. Please specify the other languages. _____

62. I just have a few more questions about your childcare program. About how long has this program been operating?

ENTER YEARS. _____

63. About how long have you been a licensed childcare provider?

ENTER YEARS. _____

64. About how long have you been licensed to provide care to infants or toddlers (children younger than 36 months)?

ENTER YEARS. _____

65. About how long have you personally worked in the childcare field?

This would be in addition to providing care for or raising your own children.

ENTER YEARS. _____

66. Are you planning any major changes in the way that you provide care or to your child care program, such as expansions, cutbacks, or reorganization, within the next year?

NO
YES. Please Explain. _____

67. Has your program experienced any major changes, such as expansions, cutbacks, or reorganization, within the past year?

NO
YES. Please Explain. _____

END OF FAMILY CHILD CARE PROGRAM INTERVIEW

This completes this interview. Thank you for taking the time to answer these questions. I will be calling in the next week or two to let you know about the meeting to learn more about participating in the study.

Appendix C. Method of Random Assignment

The objective of the random assignment was to achieve a random assignment ratio (the ratio of treatment to control group members) that was as close as possible to 1, both for the sample as a whole and within specific random assignment blocks. Such a 50/50 ratio is beneficial because it maximizes statistical power with a given sample size, it minimizes the correlation between the random assignment variable and the blocking variable, and it results in an even distribution of program services across random assignment blocks.

Because some random assignment blocks have odd numbers of programs within them, it is not possible to achieve a random assignment ratio of exactly 50/50 in each block. To address this issue and still maintain an overall random assignment ratio as close as possible to 50/50 we used a SAS program with the following steps to conduct random assignment:

1. All random assignment blocks with even numbers of programs are temporarily excluded from the first step of the random assignment process.
2. All blocks with odd numbers of programs are sorted randomly.
3. The top half of blocks with odd numbers are identified and flagged.
4. Half the programs in blocks with even numbers of programs are assigned to the treatment group.
5. Up to half the programs in blocks with odd numbers of programs that were not flagged in step 3 are assigned to the treatment group.
6. Up to half the programs plus 1 in blocks with odd numbers of programs that were flagged in step 3 are assigned to the treatment group.

For example:

1. There are three blocks, one with 10 programs and two with 5 programs each.
2. The 10-program block is dropped and the two 5-program blocks are sorted randomly.
3. The first of the 5-program blocks is identified and flagged.
4. Five programs in the 10-program block are assigned to the treatment group.
5. Half of 5 programs is 2 1/2, so 2 programs of the second 5-program block are assigned to the treatment group.
6. Half of 5 programs plus 1 is 3 1/2, so 3 programs of the first 5-program block are assigned to the treatment group.
7. The treatment group now has 5+2+3 = 10 programs and the control group has 5+3+2 = 10 programs as well.

In addition to these procedures, there were two additional programs that were randomized one at a time, using a flip of a coin.

Appendix D. Random Assignment Cohorts and Strata

Table D1. Random assignment cohorts and strata

Number of strata	Date of random assignment	Type of program	County or regional area	Number of programs randomly assigned		
				Total	Control	Intervention
4	11/2007 and 12/2007	Centers	Orange, Los Angeles, and Riverside counties, CA	12	6	6
6	1/24/2008	Centers and family child care homes	Los Angeles, Orange, and Riverside counties, CA; Phoenix, AZ	42	21	21
5	2/19/2008	Centers and family child care homes	Los Angeles and Orange counties, CA; Tucson, AZ	16	8	8
5	3/27/2008	Centers	San Diego and Riverside counties, CA; Phoenix and Tucson, AZ	10	5	5
5	4/17/2008	Centers and family child care homes	San Diego, Los Angeles, and Orange counties, CA	10	5	5
6	5/19/2008	Centers and family child care homes	Los Angeles and Orange counties, CA; Phoenix and Tucson, AZ	13	7	6
11	6/13/2008	Centers and family child care homes	Los Angeles, Orange, and Riverside counties, CA; Phoenix and Tucson, AZ	48	25	23
9	7/14/2008	Centers and family child care homes	San Diego and Riverside counties; Phoenix and Tucson, AZ; Cochise County, AZ	57	29	28
9	7/21/2008 and 8/5/2008	Centers and family child care homes	San Diego County, CA; Tucson and Phoenix, AZ; Cochise and Santa Cruz counties, AZ	43	21	22
			Total	251	127	124

Note: Random assignment cohorts on each date included multiple strata based on language, program type, and location. Strata were collapsed for the table presentation due to disclosure risk.
Source: Sample tracking data.

Appendix E. Follow-Up Data Collection Intervals, by Experimental Condition

Table E1. Mean intervals in months between random assignment and child assessment

	Overall			Control			Treatment			Difference between groups	*p*-value
Elapsed time	**Number**	**Mean**	**Standard deviation**	**Number**	**Mean**	**Standard deviation**	**Number**	**Mean**	**Standard deviation**		
Random assignment and first follow-up for the cognitive and language assessment	847	14.77	0.878	427	14.78	0.831	420	14.76	0.925	0.02	.766
Random assignment and first follow-up for the behavior assessment	877	14.72	0.896	435	14.71	0.796	442	14.73	0.883	0.02	.794
Random assignment and second follow-up for the cognitive and language assessment	863	23.18	1.028	430	23.15	1.043	433	23.21	1.014	0.06	.364
Random assignment and second follow-up for the behavior assessment	901	22.83	2.004	447	22.83	1.965	454	22.83	2.043	0	.963

Source: First and second child assessment, sample tracking data.

Table E2. Mean intervals in months between random assignment and program observations

	Overall			Control			Treatment			Difference between groups	*p*-value
Elapsed time	**Number**	**Mean**	**Standard deviation**	**Number**	**Mean**	**Standard deviation**	**Number**	**Mean**	**Standard deviation**		
Random assignment and baseline	251	–1.72	1.739	127	–1.73	1.778	124	–1.70	1.705	0.03	.888
Random assignment and first follow-up	194	14.97	1.119	100	15.05	1.165	94	14.89	1.069	0.15	.338
Random assignment and second follow-up	172	21.79	1.646	90	21.77	1.640	82	21.82	1.663	0.06	.816

Source: Program observation measures at first follow-up and second follow-up, sample tracking data.

Appendix F. Details of Study Measures

Table F1. Details on first child assessment: approximately 15 months after random assignment (age range, 15–42 months)

Measure	Instrument	Subject	Administration time	Standardization ages
Language development	*Preschool Language Scale, Fourth Edition*, Expressive Language Scale (Zimmerman, Steiner, and Pond 2002)	Child	20 minutes	Birth–6 years
Cognitive development	*Bayley Scales of Infant and Toddler Development, Third Edition* (Bayley 2006), Cognitive Subscale	Child	15-25 minutes	1–42 months
Behavior	Infant-Toddler Social and Emotional Assessment (Briggs-Gowan and Carter 2005): Externalizing, Internalizing, Dysregulation, and Competence	Parent	30 minutes	12–36 months
	Bayley Scales of Infant and Toddler Development, Third Edition (Bayley 2006), Behavior Observation Inventory	Interviewer	(5 minutes observation)	1–42 months
Child health	General health rating	Parent	<1 minute	
Parent/child interaction	Home Observation for Measurement of the Environment (HOME; Caldwell and Bradley 1984): Learning Materials, Responsivity, and Acceptance subscales	Parent/ interviewer	5 minutes (plus 5–10 minutes interviewer observation)	6 months–4.5 years
	Parent/child care provider information exchange	Parent	6 minutes	
Parent socioeconomic status	Income/employment	Parent	3–5 minutes	
PITC exposure	Child care changes over past year	Parent	2–5 minutes	

Source: Study design.

Table F2. Details of second child assessment: approximately 23 months after random assignment (age range, 25–50 months)

Measure	Instrument	Subject	Administration time	Standardization ages
Language	*Preschool Language Scale, Fourth Edition*, Expressive Language Scale (Zimmerman, Steiner, and Pond 2002)	Child	30 minutes	Birth–6 years
Cognitive	*Bracken School Readiness Assessment, Third Edition* (BSRA; Bracken 2007)	Child	10–15 minutes	3–6years, 11 months
Behavior	Child Behavior Checklist 1½-5 (CBCL 1½-5; Achenbach and Rescorla 2000)	Parent	20 minutes	1½–5 years
	Polit Positive Behavior Scale (Polit 1996)	Parent	5 minutes	1–18 years
Child health	General health rating	Parent	<1 minute	
Parent/child interaction	Home Observation for Measurement of the Environment (HOME; Caldwell and Bradley 1984): Learning Materials, Responsivity, and Acceptance subscales	Parent/ Interviewer	5 minutes (plus 5–10 minutes interviewer observation)	
	Parent/Child Care Provider Information Exchange	Parent	6 minutes	
Parent socioeconomic status	Income/employment	Parent	3–5 minutes	
PITC exposure	Child care changes over past year	Parent	2–5 minutes	

Source: Study design.

As indicated in these tables, child outcome measures and program quality measures are based on well-established scales that researchers have used in similar child care studies across a wide range of program and research contexts. Measures are copyrighted and cannot be included in the report.

Table F3. Descriptive and psychometric information on primary, secondary, and covariate measures

Measure	Sample items	Sample size	Possible range		PITC study range		Mean (standard deviation)	Internal consistency reliability
			Minimum	Maximum	Minimum	Maximum		
Child cognitive and language development								
Bayley Scales of Infant and Toddler Development, Third Edition, Cognitive Subscale (BSID) raw score	Picks up blocks, finds hidden object, completes simple puzzles, demonstrates representational play, reproduces simple patterns, classifies objects	788	0	91	44	99	68.89 (7.24)	NA
Preschool Language Scale, Fourth Edition, Expressive Communication Subscale raw score	Makes vocalizations, imitates words, names objects, uses word combinations, produces basic sentences, completes analogies	780[a]	0	68	19	63	37.75 (7.62)	NA
		799[a]	0	68	21	79	45.10 (8.50)	NA
Bracken School Readiness Assessment, Third Edition raw score	Colors, letters, numbers/counting, sizes/comparisons, shapes	796	0	85	0	78	29.79 (18.31)	NA
Child socioemotional development								
Infant-Toddler Social and Emotional Assessment raw score	Competence domain (compliance, attention, mastery motivation, imitation/play, empathy, prosocial peer relations); externalizing domain-reversed (activity/impulsivity, aggression/defiance, peer aggression); internalizing domain-reversed (depression/withdrawal, general anxiety, separation distress); dysregulation domain-reversed (negative emotionality, sleep, eating, sensory sensitivity)	810	0	2	0.47	1.49	1.09 (0.17)	0.66

Measure	Sample items	Sample size	Possible range		PITC study range		Mean (standard deviation)	Internal consistency reliability
			Minimum	Maximum	Minimum	Maximum		
Child Behavior Checklist 1½-5 raw score	Externalizing problems (attention problems, aggressive behavior) internalizing problems (emotionally reactive, anxious/depressed, somatic complaints, withdrawn)	833	0	120	0	79	16.86 (11.91)	11.91 0.78
Polit Positive Behavior Scale raw score	Social competence, autonomy, compliance	833	0	4.0	1.43	4.0	2.98 (0.40)	0.82
Child care program quality measures								
Infant/Toddler Environment Rating Scale-Revised[b]	Space and furnishings, personal care routines, listening and talking, activities, interaction, program structure		1	7	1.21 1.79	5.61 5.67	3.53 (0.92) 3.72 (0.92)	0.83 0.88 0.89
Family Child Care Environment Rating Scale-Revised	Space and furnishings, personal care routines, listening and talking, activities, interaction, program structure	159,109,90[c]	1	7	1.30 1.28	6.43 6.69	3.63 (1.2) 3.81 (1.14)	0.91 0.90 0.91
Analytic control variables								
Caregiver-rated child language	Vowel-like sounds, gestures, two-word sentences, complete sentences	864	0	7	0	7	2.73 (1.22)	NA
Caregiver-rated child problem behavior	Irritability, distractibility, emotional intensity	871	0	3	0	3	1.07 (0.38)	0.77
Caregiver traditional beliefs about raising children	Children should not question the authority of parents, children will be bad unless they are taught what is right	891	30	150	5	107	81.38 (7.98)	0.75

Measure	Sample items	Sample size	Possible range		PITC study range		Mean (standard deviation)	Internal consistency reliability
			Minimum	Maximum	Minimum	Maximum		
Caregiver concerns and rewards	Concerns such as not making much money, exposure to illness or injury; rewards such as affection from the children, seeing children's excitement	893	28	112	19	119	76.80 12.95	0.86
Baseline child care quality total score	Space and furnishings, personal care routines, listening and talking, activities, interaction, program structure	936	1	7	1.21	6.69	3.53 (0.92)	0.83 ITERS-R 0.91 FCCERS-R

Note: Standardized alphas computed using the whole sample. Raw scores were used in the creation of the cognitive/language and positive behavior composite scores.
a. The number of children assessed using the Spanish-language version was 123 at the first follow-up and 99 at the second follow-up.
b. There were 149 classrooms nested in 92 centers at baseline; 138 classrooms nested in 86 centers at first follow-up; 138 classrooms nested in 86 centers at second follow-up.
c. There were 159 family child care homes at baseline; 109 at first follow-up; and 90 at second follow-up.
Source: Program, child, and staff measures at baseline, first follow-up, and second follow-up

Appendix G. Zero-Order Correlations among Variables Used in the Impact Analysis

Table G1. Zero-order correlations among variables used in the impact analysis

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Caregiver variables, aggregated to the setting (baseline)																				
1. Caregiver Beliefs scale score	-																			
2. Caregiver Concerns and Rewards scale score	-.09	-																		
3. Highest level of education AGG	-.27	.11	-																	
4. ITERS-R/FCCERS-R total score	-.08	.05	.01	-																
5. English only language spoken at home	.11	-.06	-.07	.06	-															
Child variables																				
6. Child gender	-.01	.00	-.03	-.01	.00	-														
7. Age in months at random assignment	.03	-.09	-.08	.04	.03	-.02	-													
8. Parent highest level of education	-.07	-.03	.20	-.04	-.30	-.03	.01	-												
9. Caregiver-rated child behavior scale	.02	-.14	.10	-.05	.08	.03	-.02	.04	-											
10. Caregiver-rated child communication	-.01	-.12	-.11	.14	.02	-.04	.32	.03	-.01	-										
11. Cognitive/Language composite–first follow-up	-.05	-.06	.03	.11	-.10	-.16	.70	.22	.01	.27	-									
12. Positive Behavior composite–first follow-up	-.04	-.01	.03	.01	-.06	-.07	.09	.08	.00	.03	.23	-								
13. Cognitive/Language composite–second follow-up	-.07	-.02	.08	.03	-.16	-.08	.36	.30	.02	.19	.47	.11	-							
14. Positive Behavior composite–second follow-up	.02	.03	.04	-.01	-.10	-.11	.03	.14	.02	.01	.15	.55	.10	-						
Child care quality, aggregated to the setting																				
15. ITERS-R/FCCERS-R total score–first follow-up	-.20	.01	.12	.54	.20	-.01	.02	-.03	.12	.12	.11	.02	.00	.02	-					
16. ITERS-R/FCCERS-R total score-W3–second follow-up	-.24	.04	.30	.49	.13	-.04	-.03	.04	.05	.08	.07	-.01	.01	.03	.64	-				
17. Quality Interaction composite–first follow-up	-.25	.19	.23	.36	.16	-.01	-.09	-.06	.05	-.06	.03	.03	-.06	.02	.71	.51	-			
18. Quality Interaction composite–second follow-up	-.29	.15	.33	.38	.14	-.02	-.08	.03	.06	-.02	.04	.00	.01	.06	.51	.80	.63	-		
19. Quality Interaction composite environment rating scale only–W2 first follow-up	-.28	.14	.15	.42	.08	-.01	-.05	.03	.03	.02	.05	.03	-.02	.03	.82	.51	.82	.52	-	
20. Quality Interaction composite environment rating scale only–second follow-up	-.25	.12	.32	.37	.05	-.04	-.04	.10	.04	.04	.07	-.01	.04	.04	.53	.87	.51	.86	.55	-

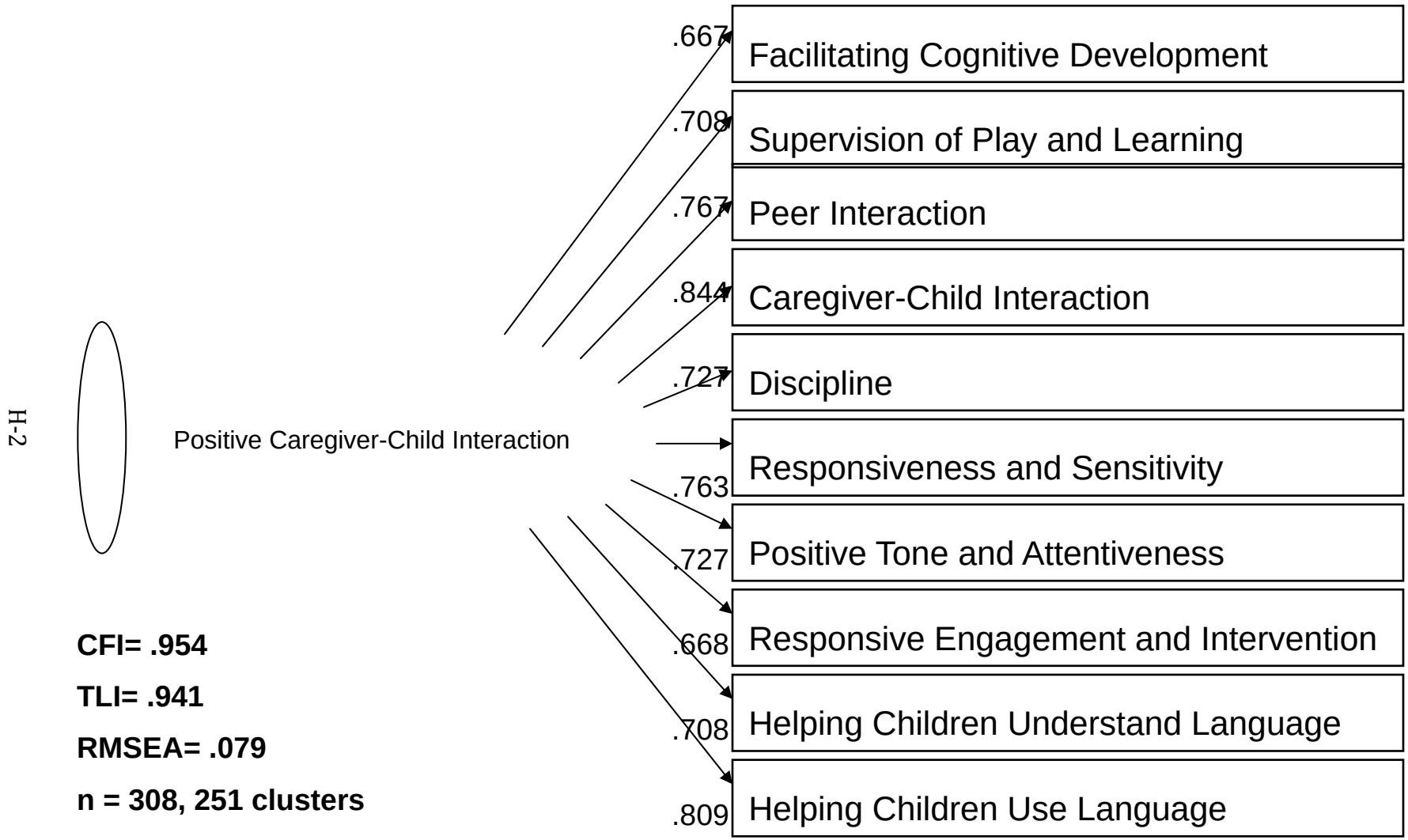
Note: Caregiver self-report n = 893 at baseline; caregiver child rating n = 871 at baseline; child care observations n = 936 at baseline, 819 first follow-up, and 769 second follow-up.
Parent report child behavior n = 936 at first follow-up and 833 at second follow-up.
Child cognitive and language test n = 788 at first follow-up and 799 at second follow-up
Source: Child and program measures (baseline questionnaire, child assessments, and observations) at baseline, first follow-up, and second follow-up; staff measures (questionnaires) baseline and first follow-up.

Appendix H. Caregiver-Child Quality Interaction Composite: Factor Analysis

Guided by developmental theory (Dowsett et al. 2008; Thomason 2009) and published results from nonexperimental data (Howes 2008; Mashburn et al. 2008; Sylva et al. 2006), we selected items from our child care observations that capture the kinds of interactions that occur between caregivers and children that are expected to promote child cognitive, language, and socioemotional development. Six items from the environment rating scales (ITERS-R and FCCERS-R) and four items from the PITC-PARS scale at the baseline assessment were evaluated using confirmatory factor analysis with Mplus software (Version 5.21; Muthén and Muthén 1998–2010; figure H1),and the "cluster" option was used to account for the nesting of classrooms within child care centers. Because the PITC-PARS items were developed by the creators of the PITC and may overly align with the intervention, impacts were run on a modified composite that excludes these items (figure H2).

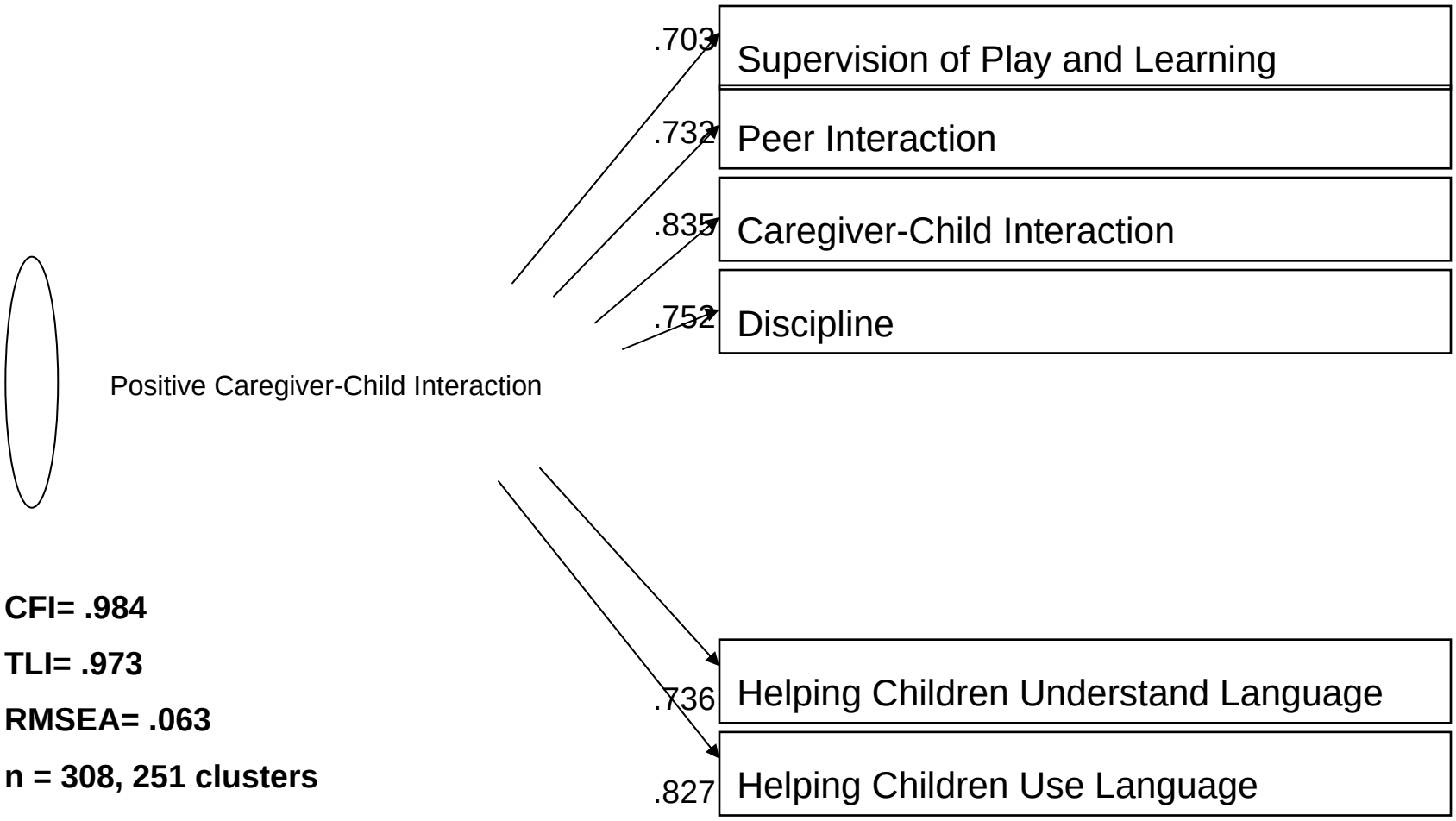
A confirmatory factor analysis of the selected staff-child interaction items, using baseline measures, indicated that all these items had loadings above 0.66 on the composite. The factor models were evaluated using three goodness-of-fit indices: the comparative fit index (Bentler 1990), Tucker-Lewis index (Tucker and Lewis 1973), and root mean square error of approximation (Steiger 1990). A comparative fit index or Tucker-Lewis index ≥ 0.90 indicates acceptable fit, ≥ 0.95 indicates very good fit, and ≥ 0.99 indicates excellent fit. A root mean square error of approximation ≤ 0.05 indicates a close model fit and that values up to 0.08 represent adequate model fit with reasonable errors of approximation in the population (Browne and Cudeck 1993).

Figure H1. Confirmatory factor analysis results: caregiver interaction composite



Source: Study design.

Figure H2. Confirmatory factor analysis results: caregiver interaction composite excluding PITC-PARS items



Source: Study design.

Appendix I. Training and Reliability-Checking of Field Staff

Program observers

Twenty-one observers were trained during this study, seven as anchors. The role of the anchors in data collection was to help train and check interrater reliability of staff hired to conduct program observations. The staff hired to serve as anchors had either previous experience using the observation measures, or a Master's degree in child development or a related field, or both a graduate degree and experience using the observation measures. Anchors were not always based in areas where the observations were conducted, some traveled to participate.

During the study, two anchors left and were replaced with experienced observers who had been reliable at the 90-percent-or-above level for rechecks during the previous round of observations. At the beginning of each round of data collection, an initial center and an initial family child care observation training was held for new observers. Most observers were trained to conduct observations in one type of child care setting (either family child care homes or child care centers). Three of the seven anchors were trained to conduct observations in both types of settings.

Before the training of the full observation team, seven staff completed week-long training sessions to prepare them to serve as anchors for other observers. These sessions were delivered by the authors of the environment rating scales. One week-long training session was devoted to the Family Child Care Environment Rating Scale Revised Edition (FCCERS-R) and the other was devoted to the Infant/Toddler Environment Rating Scale Revised Edition (ITERS-R). Each training involved one day in the classroom, during which trainees watched videos and practiced scoring, as well as four days of practice observations for which trainees were paired with a trainer and conducted interrater reliability checks after each practice observation. The interrater reliability was calculated using adjacent agreement, where scores within one point of each other were considered in agreement. (For example, scores of 1 and 2 were considered in agreement but scores of 1 and 3 were considered in disagreement.) For ratings of *not applicable*, observers were scored as in agreement only if both raters selected *not applicable*. Trainees were considered certified as an anchor observer if their interrater reliability score was 85 percent or higher for the fourth practice observation. All trainees were certified except one, who was not hired to serve as an anchor or to conduct observations.

The authors of the PITC-Program Assessment Rating Scales (PITC-PARS) also provided three days of training to the anchors that involved one classroom day, during which trainees watched videos and practiced scoring, and two days of practice observations for which trainees were paired with a trainer and conducted interrater reliability checks after each practice observation. For these checks, interrater reliability was calculated using percent exact agreement. To be certified as anchors on the PITC-PARS, trainees' interrater reliability score had to be 80 percent or higher for the second practice observation. All but three anchors were certified. The three not certified were required to conduct an additional three days of practice observations with a certified anchor. All were certified after these additional days of training.

After the anchor training, a larger team of trainees participated in an initial one-day classroom training and in a minimum of four practice observations in which they were paired with an anchor and conducted interrater reliability checks after each practice observation. For these checks, for the ITERS-R or the FCCERS-R, interrater reliability was calculated using adjacent agreement. The PITC-PARS interrater reliability was calculated using exact agreement. A combined ITERS-R/FCCERS-R and PITC-PARS observation reliability score was calculated using the total number of items as the denominator and the total number of items with agreement as the numerator. Observers were certified if their interrater reliability score was 85 percent or higher for the fourth practice observation. For observers who did not meet the certification requirement on the fourth practice observation, additional practice observations and training discussions were conducted until they had an 85 percent or higher interrater reliability score.

After every 10 observations, each observer was rechecked for reliability in a paired observation with an anchor. To allow for flexibility in scheduling paired observations with anchors, observers were allowed to complete up to 13 observations before being rechecked. However, after the 13th observation, observers were required to wait until the paired observation to conduct further observations on their own. If an observer did not meet 85 percent reliability during the recheck, he or she was required to complete an additional paired observation before being allowed to conduct observations on his or her own again. Anchors also checked their own interrater reliability with other anchors at least twice during each round of data collection, and after every tenth observation if they were conducting observations in addition to training and checking interrater reliability of other staff.

A total of 71 reliability check observations were conducted during the three rounds of data collection. The overall mean interrater reliability for the three rounds of program observations was 86 percent (standard deviation =6.38). The mean interrater reliability for program observations at baseline was 85 percent (standard deviation =6.51), at the first follow-up was 87 percent (standard deviation =6.77), and at the second follow-up was 88 percent (standard deviation =5.29). For the three rounds of data collection there were10 rechecks in which the interrater reliability score fell below 80 percent; each of these observers completed at least one additional paired observation and then met the 85 percent interrater reliability standard.

Child assessors

The research team hired and trained 13 child assessors living within the study counties, all with experience working with young children. Assessors participated in an initial three-day in-person training that included at least four practice assessments. These practice assessments took place in a laboratory setting with nonstudy children who were in the same age range as the study children. To be certified to collect data in the field, assessors completed at least one practice session with a maximum of two errors in administration and two errors in scoring. Before the second follow-up assessment, assessors conducted videotaped practice assessments, with the videos reviewed by field managers prior to recertification of the assessors to return to the field. Assessors also participated in refresher trainings, trainings in new measures for the second follow-up, and rechecks through in-person sessions and video. Refresher trainings involved practice assessments with nonstudy children in the same age range as the study children. Some refresher practice assessments were conducted in the children’s homes and some were conducted in a laboratory setting. On rechecks, which occurred when assessors restarted assessments after a

break of one month or more, assessors were sent a DVD of an assessment conducted with a child in a home. Assessors were required to view, score, and identify any errors in administration for these videotaped assessments. For example, during a recorded assessment, if an assessor presented the puzzle pieces in the incorrect position or placed the plastic ducks in the wrong position or one was the wrong color or size, the assessor viewing the video would have to note these errors and explain the correct method of administration.

Appendix J. Detailed Response Rates and Reasons for Nonresponse, by Experimental Condition

Table J1. Response rates for outcome measures, including reasons for nonresponse

	Overall		Intervention		Control				
Outcome measure	**Number**	**Percent**	**Number**	**Percent**	**Number**	**Percent**	**Percentage difference between groups**	***t*-statistic**	***p*-value**
Child assessment (first follow-up)									
Number of children enrolled	936	100	480	100	456	100			
Completed	822	88	417	87	405	89	2	0.91	.364
Ineligible	4	<1	#	#	#	#	#	--	--
Refusal	53	6	26	5	27	6	1	0.33	.739
Moved out of area	10	1	#	#	#	#	0	0.72	.473
Unable to contact	37	4	24	5	13	3	2	–1.69	.092
Unable to schedule	10	1	6	1	4	1	0	–0.55	.579
Child assessment components completed at first follow-up									
BSID	789	84	395	82	394	86	4	1.73	.084
Preschool Language Scale	782	84	392	82	390	85	3	1.59	.111
Child care interview (with parent)	823	88	420	87	403	88	1	0.41	.681
Parent follow-up questionnaire	810	86	414	86	396	87	1	0.27	.791
Child assessment (second follow-up)									
Completed	837	89	427	89	410	90	1	0.47	.635
Ineligible	#	#	#	#	#	#	#	--	--
Refusal	54	6	30	6	24	5	1	–0.65	.517
Moved out of area	11	1	6	1	5	1	0	–0.22	.828
Unable to contact	31	3	15	3	16	4	1	0.33	.743
Lost or stolen data	#	#	#	#	#	#	#	--	--

Outcome measure	Overall		Intervention		Control		Percentage difference between groups	*t*-statistic	*p*-value
	Number	Percent	Number	Percent	Number	Percent			
Child assessment components completed at second follow-up									
BSRA	798	85	404	84	394	86	2	0.96	.335
Preschool Language Scale	799	85	404	84	395	87	3	1.06	.288
Child care interview (with parent)	839	90	429	89	410	90	1	0.27	.787
Parent follow-up questionnaire	833	89	425	88	408	89	1	0.46	.649
Child care program observations									
First follow-up									
Centers									
Number of centers enrolled	92	100	46	100	46	100			
Completed	85	92	42	91	43	93	2	0.39	.698
Refusals	#	#	#	#	#	#	#	--	--
Closures	#	#	#	#	#	#	#	--	--
Unable to contact	#	#	#	#	#	#	#	--	--
Family child care									
Total number of family child care enrolled in study	159	100	78	100	81	100			
Completed	109	69	52	67	57	70	7	0.50	.618
Refusals	20	13	11	14	9	11	3	–0.57	.572
Closures	10	6	6	8	4	5	3	–0.71	.478
No longer providing care to infants and toddlers	10	6	4	5	6	7	2	0.59	.557
Unable to contact	10	6	5	6	5	6	0	–0.06	.951

Outcome measure	Overall		Intervention		Control		Percentage difference between groups	*t*-statistic	*p*-value
	Number	Percent	Number	Percent	Number	Percent			
Second follow-up									
Centers									
Completed	82	89	40	87	42	91	4	0.66	.508
Refusals	4	4	#	#	#	#	#	--	--
Closures	6	7	#	#	#	#	#	--	--
Family child care									
Completed	90	57	42	54	48	59	5	0.69	.494
Refusals	29	18	15	19	14	17	2	–0.32	.753
Closures	17	11	9	12	8	10	2	–0.34	.737
No longer providing care to infants and toddlers	15	9	9	12	6	7	5	–0.89	.376
Unable to contact	8	5	#	#	#	#	#	--	--

# indicates that the number is withheld because of disclosure risk.
-- is not applicable.
Note: A completed child assessment is one where the child and/or parent participated in at least one of the three main components of the child assessment used in the impact analysis: the *Bayley Scales of Infant and Toddler Development, Third Edition,* Cognitive Subscale (BSID)/*Bracken School Readiness Assessment, Third Edition* (BSRA); the *Preschool Language Scale, Fourth Edition,* Expressive Communication Subscale (PLS-4); or the parent follow-up questionnaire (the behavior assessment).
Source: Sample tracking database.

Appendix K. Teacher Sample Turnover and Response Rates

Table K1. Details of teacher sample flow and response rates

	Baseline			First follow-up				
	Original enrolled	Questionnaire completions	Response rate (percent)	Total questionnaire completions	Total response rate (percent)	Original teacher questionnaire completions	Original teacher response rate (percent)	Replacement teacher questionnaire completions
Family child care								
Intervention	125	104	83.2	73	58.4	55	44.0	18
Control	121	92	76.0	85	70.2	58	47.9	27
Total	246	196	79.7	158	64.2	113	45.9	45
Center directors								
Intervention	58	50	86.2	45	77.6	34	58.6	11
Control	55	44	80.0	48	87.3	28	50.9	20
Total	113	94	83.2	93	82.3	62	54.9	31
Center staff								
Intervention	361	222	61.5	230	63.7	114	31.6	116
Control	347	221	63.7	212	61.1	107	30.8	105
Total	708	443	62.6	442	62.4	221	31.2	221

Source: Enrollment tracking database; baseline and follow-up one teacher questionnaires.

Table K2. Center staff tenure in program at first follow-up

	Overall		Intervention		Control		Difference	*t*-statistic	*p*-value
	Number	Mean	Number	Mean	Number	Mean			
Center directors									
Mean number of years worked at center	92	7.7	44	7.4	48	7.9	0.30	0.38	.703
Percentage reporting one year or less	93	16.1	45	11.1	48	20.8	9.7	1.27	.207
Center staff									
Mean number of years worked at center	438	4.3	227	4.5	211	4.1	0.4	1.01	.311
Percentage reporting one year or less	442	28.3	230	27.4	212	29.3	1.9	0.43	.666

Source: Center director and center staff follow-up questionnaire.

Appendix L. Additional Sample Equivalence Tables

Table L1. Program, staff, and child baseline characteristics for the retained sample at first follow-up, by experimental condition

Measure	Intervention (standard deviation)	Control (standard deviation)	Difference between groups	*t*-statistic	*p*-value
Program characteristics (all programs)					
Percent of programs where English is primary language spoken	56.7	58.0	–1.3	–0.18	.8537
Percent of programs receiving government funding	67.7	78.8	–11.1	–1.74	.0841
Mean number of children enrolled younger than 24 months	8.6 (11.13)	8.8 (10.67)	–0.2	–0.09	.9263
Mean baseline global quality score on the ITERS-R and FCCERS-R	3.7 (1.05)	3.5 (1.16)	0.2	1.16	.2486
Mean Concerns and Rewards[a] total sum score (program average)	80.7 (9.85)	78.7 (13.14)	1.9	1.13	.2584
Mean Modernity[b] Scale total sum score (program average)	78.0 (17.16)	80.2 (15.60)	–2.2	–0.93	.3554
Mean Goals[c] (program average):					
Religious instruction	1.5 (0.79)	1.5 (0.76)	0.1	0.43	.6702
School readiness	2.7 (0.54)	2.7 (0.53)	0.0	0.22	.8250
Assisting disadvantaged children	2.7 (0.55)	2.6 (0.61)	0.1	0.92	.3590
Staff/teacher characteristics					
Percent who speak English as primary language with the children	60.0	59.4	0.6	0.15	.8830
Percent who have a bachelor's degree or more education	25.2	20.3	4.9	1.46	.1441
Percent who received early childhood education or training in the last 12 months	85.4	85.9	–0.5	–0.18	.8584
Child/family characteristics					
Percent of children who are male	52.8	46.1	6.7	1.88	.0607
Child's race/ethnicity					
Percent of children who are White, non-Hispanic	31.7	29.6	2.1	0.65	.5186
Percent of children who are Black, non-Hispanic	7.0	5.9	1.1	0.59	.5530
Percent of children who are Hispanic	51.3	53.0	–1.7	–0.46	.6423
Percent of children living with two parents	65.5	67.1	–1.6	–0.48	.6311

Measure	Intervention (standard deviation)	Control (standard deviation)	Difference between groups	*t*-statistic	*p*-value
Percent of parents with a bachelor's degree or more education	36.0	35.2	0.8	0.22	.8280
Percent of children where the primary language at home is English	76.7	72.6	4.1	1.30	.1953
Mean child age in months at Random Assignment	17.7 (6.99)	17.2 (6.92)	0.5	0.98	.3295
Mean hours per week child attends child care program	36.9 (9.87)	36.2 (10.29)	0.7	0.99	.3212
Mean number of siblings	1.0 (1.04)	1.0 (1.11)	-0.0	–0.26	.7920
Mean parent age	30.0 (7.35)	29.6 (7.24)	0.3	0.63	.5278
Mean number of hours parent is employed and/or in school	37.7 (11.39)	37.6 (11.27)	0.1	0.11	.9138
Caregiver-rated child behavior and language scales					
Child's mean score on behavior scale[d]	1.1 (0.36)	1.0 (0.38)	0.0	1.57	.1180
Child's mean score on the language scale[e]	2.8 (1.14)	2.7 (1.25)	0.1	1.29	.1983

Note: Total sample sizes range from 160 to 194 for program characteristics, from 625 to 633 for staff/teacher characteristics, and from 722 to 777 for child and family characteristics.

a. The Taking Care of Young Children questionnaire total sum score has a possible range of 28 to 112, with higher values indicating more positive attitudes toward taking care of young children.

b. The Parental Modernity Scale of Child-rearing and Education Beliefs total sum score has a possible range of 30 to 150, with higher values indicating more traditional beliefs about raising young children.

c. Center directors and family child care providers were asked how important (not at all=1, a little important=2, or very important=3) each of the following goals is for their child care program: to provide religious instruction, to prepare children for school with a strong academic curriculum, and to provide compensatory education for disadvantaged children.

d. Baseline child behavior scale. As part of a self-administered questionnaire, caregivers rated children's problem behaviors, including fearfulness, irritability, focused attention, and adaptability to change. This eight-item scale is based on scored items, using a 0–3 scale, with 0 indicating that the child "is never like this" and 3 indicating that the child "is like this most of the time." Scores were calculated by taking the average of available items; higher scores indicate higher levels of problem behavior.

e. Baseline child language scale. As part of a self-administered questionnaire, caregivers identified the item that best characterized children's language and communication abilities using an eight-item list (scored 0–7). Lower scores indicate more rudimentary language abilities (the child repeats sounds; the child's sounds have a speech-like babble to them) and higher scores indicate more complex language abilities (the child is talking in fairly complete short sentences; the child is talking in long and complicated sentences).

Source: Screening questionnaire; baseline center director, center staff/teacher, and family child care home questionnaires; parent baseline questionnaire and individual child forms completed at baseline by child's primary teacher.

Table L2. Program, staff/teacher, and child baseline characteristics, retained and nonretained samples at first follow-up

Measure	Retained (standard deviation)	Not retained (standard deviation)	Difference between groups	*t*-statistic	*p*-value
Program characteristics (all programs)					
Percent of programs where English is primary language spoken	57.4	58.9	–1.5	–0.21	.8363
Percent of programs receiving government funding	73.4	75.0	–1.6	–0.23	.8159
Mean number of children enrolled under 24 months of age	8.7 (10.86)	3.4 (4.46)	5.3	5.05	.0001**
Mean baseline global quality score on ITERS-R and FCCERS-R	3.6 (1.11)	3.4 (0.91)	0.2	1.15	.2516
Mean Concerns and Rewards[a] total sum score (program average)	79.7 (11.63)	81.8 (7.85)	–2.1	–1.5	.1286
Mean Modernity[b] Scale total sum score (program average)	79.1 (16.39)	79.6 (16.77)	–0.5	–0.18	.8594
Mean Goals[c] (program average):					
Religious instruction	1.5 (0.78)	1.5 (0.77)	0.0	0.19	.8458
School readiness	2.7 (0.53)	2.8 (0.51)	–0.1	–0.74	.4583
Assisting disadvantaged children	2.7 (0.58)	2.7 (0.52)	–0.0	–0.18	.8600
Staff/teacher characteristics					
Percent who speak English as primary language with the children	59.7	64.8	–5.1	–0.93	.3512
Percent who have a bachelor's degree or more education	22.7	14.4	8.3	2.03	.0449*
Percent who received early childhood education or training in the last 12 months	85.7	83.2	2.5	0.62	.5334
Child/family characteristics					
Percent of children who are male	49.5	49.4	0.1	0.03	.9779
Child's race/ethnicity					
Percent of children who are White, non-Hispanic	30.6	22.6	8.0	2.02	.0437*
Percent of children who are Black, non-Hispanic	6.4	9.4	–3.0	–1.21	.2292
Percent of children who are Hispanic	52.1	51.6	0.5	0.13	.8993
Percent of children living with two parents	66.3	58.6	7.7	1.84	.0660

Measure	Retained (standard deviation)	Not retained (standard deviation)	Difference between groups	*t*-statistic	*p*-value
Percent of parents with a bachelor's degree or more education	35.6	34.2	1.1	0.33	.7422
Percent of children where the primary language at home is English	74.6	68.8	5.8	1.48	.1391
Mean child age in months at random assignment	17.5 (6.95)	17.9 (6.85)	–0.5	–0.76	.4504
Mean hours per week child attends child care program	36.5 (10.08)	37.3 (9.65)	–0.8	–0.93	.3541
Mean number of siblings	1.0 (1.02)	1.0 (1.15)	0.1	–0.50	.6175
Mean parent age	29.8 (7.30)	29.5 (7.74)	0.3	0.47	.6377
Mean number of hours parent is employed and/or in school	37.6 (11.32)	39.0 (9.18)	–1.4	–1.58	.1155
Teacher rated child behavior and language scales					
Child's mean score on behavior scale[d]	1.1 (0.37)	1.1 (0.44)	0.0	–1.13	.2615
Child's mean score on the language scale[e]	2.7 (1.19)	2.8 (1.33)	–0.1	–0.65	.5175

Note: Total sample sizes range from 218 to 251 for the overall program sample, from 160 to 194 for the retained program sample, and from 47 to 57 for the not retained program sample. Total sample sizes range from 715 to 725 for the overall staff/teacher sample, from 625 to 633 for the retained staff/teacher sample, and from 89 to 91 for the not retained staff/teacher sample. Total sample sizes range from 864 to 936 for the overall child sample, from 722 to 777 for the retained child sample, and from 141 to 159 for the not retained child sample.

*$p < .05$, statistically different from zero at the .05 level, two-tailed test

**$p<.01$, statistically different from zero at the .01 level, two-tailed test

a. The Parental Modernity Scale of Child-rearing and Education Beliefs total sum score has a possible range of 30 to 150, with higher values indicating more traditional beliefs about raising young children.

b. The Taking Care of Young Children questionnaire total sum score has a possible range of 28 to 112, with higher values indicating more positive attitudes about taking care of young children.

c. Center directors and family child care providers were asked how important (not at all=1, a little important=2, or very important=3) each of the following goals is for their child care program: to provide religious instruction, to prepare children for school with a strong academic curriculum, and to provide compensatory education for disadvantaged children.

d. Baseline child behavior scale. As part of a self-administered questionnaire, caregivers rated children's problem behaviors, including fearfulness, irritability, focused attention, and adaptability to change. This eight-item scale is based on scored items, using a 0–3 scale, with 0 indicating that the child "is never like this" and 3 indicating that the child "is like this most of the time." Scores were calculated by taking the average of available items; higher scores indicate higher levels of problem behavior.

e. Baseline child language scale. As part of a self-administered questionnaire, caregivers identified the item that best characterized children's language and communication abilities using an eight-item list (scored 0–7). Lower scores indicate more rudimentary language abilities (the child repeats sounds; the child's sounds have a speech-like babble to them) and higher scores indicate more complex language abilities (the child is talking in fairly complete short sentences; the child is talking in long and complicated sentences).

Source: Screening questionnaire; baseline center director, center staff/teacher, and family child care home questionnaires; parent baseline questionnaire and individual child forms completed at baseline by child's primary teacher.

Table L3. Program, staff/teacher, and child baseline characteristics, retained and nonretained samples at second follow-up

Measure	Retained (standard deviation)	Not retained (standard deviation)	Difference between groups	*t*-statistic	*p*-value
Program characteristics (all programs)					
Percent of programs where English is primary language spoken	59.5	53.9	5.6	0.84	.4036
Percent of programs receiving government funding	74.1	73.1	1.0	0.17	.8633
Mean number of children enrolled under 24 months of age	9.5 (11.34)	3.3 (4.22)	6.2	5.76	.0001*
Mean baseline global quality scale score	3.6 (1.09)	3.6 (1.02)	0.0	0.05	.9584
Mean Concerns and Rewards[a] total sum score (program average)	79.7 (11.38)	81.2 (9.88)	–1.5	–0.97	.3331
Mean Modernity[b] Scale total sum score (program average)	79.3 (15.64)	79.1 (18.18)	0.2	0.08	.9402
Mean Goals[c] (program average)					
Religious instruction	1.5 (0.77)	1.6 (0.79)	–0.1	–1.07	.2839
School readiness	2.7 (0.51)	2.7 (0.56)	0.0	0.25	.8041
Assisting disadvantaged children	2.7 (0.58)	2.7 (0.56)	–0.0	–0.17	.8663
Staff/teacher characteristics					
Percent who speak English as primary language with the children	60.5	59.6	0.9	0.21	.8326
Percent who have a bachelor's degree or more education	22.5	18.1	4.4	1.13	.2603
Percent who received early childhood education or training in the last 12 months	86.5	80.5	6.0	1.61	.1098
Child/family characteristics					
Percent of children who are male	49.4	50.0	-0.6	–0.14	.8908
Child's race/ethnicity					
Percent of children who are White, non-Hispanic	30.2	24.3	5.9	1.41	.1599
Percent of children who are Black, non-Hispanic	6.4	10.0	-3.6	–1.34	.1832
Percent of children who are Hispanic	52.9	47.1	5.8	1.25	.2099
Percent of children living with two parents	65.3	63.0	2.3	0.52	.6051

Measure	Retained (standard deviation)	Not retained (standard deviation)	Difference between groups	*t*-statistic	*p*-value
Percent of parents with a bachelor's degree or more education	34.7	39.0	–4.3	–0.96	.3382
Percent of children where the primary language at home is English	75.2	64.7	10.5	2.57	.0105*
Mean child age in months at random assignment	17.4 (6.88)	18.0 (7.26)	-0.6	–0.89	.3748
Mean hours per week child attends child care program	36.5 (10.03)	37.5 (9.87)	–1.1	–1.14	.2542
Mean number of siblings	1.0 (1.08)	0.9 (1.15)	0.0	0.44	.6631
Mean parent age	29.7 (7.27)	30.1 (7.92)	–0.4	–0.60	.5477
Mean number of hours parent is employed and/or in school	37.6 (11.15)	39.0 (10.04)	–1.3	–1.28	.2004
Teacher rated child behavior and language scales					
Child's mean score on behavior scale[d]	1.1 (0.43)	1.1 (0.37)	0.1	–1.73	.0854
Child's mean score on the language scale[e]	2.7 (1.17)	2.8 (1.48)	–0.1	–0.58	.5633

Note: Total sample sizes range from 218 to 251 for the overall program sample, from 139 to 172 for the retained program sample, and from 68 to 79 for the nonretained program sample. Total sample sizes range from 715 to 725 for the overall staff/teacher sample, from 582 to 588 for the retained staff/teacher sample, and from 133 to 136 for the nonretained staff/teacher sample. Total sample sizes range from 864 to 936 for the overall child sample, from 739 to 796 for the retained child sample, and from 125 to 140 for the nonretained child sample.

a. The Taking Care of Young Children questionnaire total sum score has a possible range of 28 to 112, with higher values indicating more positive attitudes about taking care of young children.

b. The Parental Modernity Scale of Child-rearing and Education Beliefs total sum score has a possible range of 30 to 150, with higher values indicating more traditional beliefs about raising young children.

c. Center directors and family child care providers were asked how important (not at all=1, a little important=2, or very important=3) each of the following goals is for their child care program: to provide religious instruction, to prepare children for school with a strong academic curriculum, and to provide compensatory education for disadvantaged children.

d. Baseline child behavior scale. As part of a self-administered questionnaire, caregivers rated children's problem behaviors, including fearfulness, irritability, focused attention, and adaptability to change. This eight-item scale is based on scored items, using a 0–3 scale, with 0 indicating that the child "is never like this" and 3 indicating that the child "is like this most of the time." Scores were calculated by taking the average of available items; higher scores indicate higher levels of problem behavior.

e. Baseline child language scale. As part of a self-administered questionnaire, caregivers identified the item that best characterized children's language and communication abilities using an eight-item list (scored 0–7). Lower scores indicate more rudimentary language abilities (the child repeats sounds; the child's sounds have a speech-like babble to them) and higher scores indicate more complex language abilities (the child is talking in fairly complete short sentences; the child is talking in long and complicated sentences).

Source: Screening questionnaire; baseline center director, center staff/teacher, and family child care home questionnaires; parent baseline questionnaire and individual child forms completed at baseline by child's primary teacher.

Table L4. Program and staff characteristics of retained program sample at second follow-up, by type of care and experimental condition

Baseline measure	Intervention (standard deviation)	Control (standard deviation)	Difference between groups	*t*-statistic or *z*-score	*p*-value
Program characteristics (family child care)					
Percent of programs where English is primary language spoken	39.0 *n = 41*	39.6 *n = 48*	–0.6	–0.05	.9577
Percent of programs receiving government funding	73.8 *n = 42*	80.9 *n = 47*	–7.1	–0.79	.4326
Mean number of children enrolled under 24 months of age	2.7 (1.52) *n = 42*	2.8 (1.42) *n = 48*	–0.1	–0.19	.8484
Mean baseline global quality scale score	3.8 (1.33) *n = 42*	3.5 (1.28) *n = 48*	0.3	1.06	.2914
Mean family child care provider Taking Care of Young Children total sum score (program average)	79.7 (13.00) *n = 42*	76.1 (15.88) *n = 43*	3.6	1.16	.2504
Mean family child care provider Modernity Scale[b] total sum score (program average)	82.7 (20.87) *n = 42*	79.2 (17.77) *n = 43*	3.5	0.83	.4091
Mean family child care provider goals[c] (program average):					
Religious instruction	1.8 (0.86) *n = 42*	1.6 (0.79) *n = 40*	0.2	1.13	.2635
School readiness	2.7 (0.48) *n = 41*	2.8 (0.42 *n = 43*	–0.1	–0.39	.6990
Assisting disadvantaged children	2.7 (0.54) *n = 36*	2.6 (0.63) *n = 42*	0.1	0.82	.4131
Program characteristics (centers)					
Percent of programs where English is primary language spoken	83.8 *n = 37*	81.0 *n = 42*	2.8	0.32	.7461
Percent of programs receiving government funding	61.5 *n = 39*	78.6 *n = 42*	–17.1	–1.69	.0957
Mean number of children enrolled under 24 months of age	21.4 (12.23) *n = 24*	22.7 (9.63) *n = 25*	–1.3	–0.43	.6701
Mean baseline global quality scale score	3.6 (0.67) *n = 40*	3.4 (0.88) *n = 42*	0.2	1.33	.1883
Mean center staff Taking Care of Children[a] total sum score (program average)	81.0 (5.63) *n = 39*	82.2 (6.02) *n = 39*	–1.2	–0.88	.3797

Baseline measure	Intervention (standard deviation)	Control (standard deviation)	Difference between groups	*t*-statistic or *z*-score	*p*-value
Mean center staff Modernity Scale[b] total sum score (program average)	76.1 (9.05) *n = 39*	78.8 (10.86) *n = 39*	–2.7	–1.21	.2294
Mean center director goals[c] (program average)					
Religious instruction	1.3 (0.66) *n = 37*	1.2 (0.55) *n = 37*	0.1	0.95	.3435
School readiness	2.7 (0.55) *n = 37*	2.6 (0.60) *n = 37*	0.1	0.44	.6628
Assisting disadvantaged children	2.7 (0.53) *n = 35*	2.5 (0.60) *n = 36*	0.2	0.96	.3381
Child care center staff characteristics					
Percent of center staff who speak English as primary language with the children	68.5 *n = 235*	66.5 *n = 239*	2.0	0.46	.6456
Percent of center staff who have a bachelor's degree or more education	27.6 *n = 232*	21.5 *n = 237*	6.1	1.53	.1273
Percent of center staff who received early childhood education or training in the last 12 months	84.9 *n = 231*	88.2 *n = 238*	–3.3	-1.07	.2832
Family child care staff characteristics					
Percent of family child care providers who speak English as primary language with the children	31.0 *n = 58*	32.1 *n = 56*	–1.1	–0.13	.8998
Percent of family child care providers who have a bachelor's degree or more education	12.5 *n = 56*	15.8 *n = 57*	–3.3	–0.50	.6198
Percent of family child care providers who received early childhood education or training in the last 12 months	94.7 *n = 57*	77.2 *n = 57*	17.5	2.76	.0070**

**$p < .01$, significantly different from zero at the .01 level, two-tailed test.

a. The Taking Care of Young Children questionnaire total sum score has a possible range of 28 to 112, with higher values indicating more positive attitudes about taking care of young children.

b. The Parental Modernity Scale of Child-rearing and Education Beliefs total sum score has a possible range of 30 to 150, with higher values indicating more traditional beliefs about raising young children.

c. Center directors and family child care providers were asked how important (not at all = 1, a little important = 2, or very important = 3) each of the following goals is for their child care program: to provide religious instruction, to prepare children for school with a strong academic curriculum, and to provide compensatory education for disadvantaged children.

Source: Authors' analysis of primary data collected for the study.

Table L5. Program and staff characteristics, retained and nonretained samples at second follow-up, by program type

Baseline measure	Retained at second follow-up	Not retained	Difference between groups	*t*-statistic or *z*-score	*p*-value
Program characteristics (family child care)					
Percent of programs where English is primary language spoken	39.3 *n = 89*	48.5 *n = 68*	–9.2	–1.15	.2516
Percent of programs receiving government funding	77.5 *n = 89*	72.1 *n = 68*	5.4	0.78	.4352
Mean number of children enrolled under 24 months of age	2.7 (1.46) *n = 90*	2.3 (1.34) *n = 69*	0.4	1.73	.0864
Mean baseline global quality scale score	3.7 (1.31) *n = 90*	3.6 (1.05) *n = 69*	0.1	0.42	.6786
Mean family child care provider Taking Care of Young Children[a] total sum score (program average)	77.9 (14.56) *n = 85*	80.9 (10.32) *n = 64*	–3.0	–1.46	.1457
Mean family child care provider Modernity Scale[b] total sum score (program average)	81.0 (19.32) *n = 85*	80.1 (19.24) *n = 65*	0.9	0.27	.7875
Mean family child care provider goals[c] (program average)					
Religious instruction	1.7 (0.83) *n = 82*	1.6 (0.80) *n = 59*	0.1	0.55	.5845
School readiness	2.8 (0.44) *n = 84*	2.8 (0.45) *n = 63*	0.0	–0.45	.6517
Assisting disadvantaged children	2.7 (0.59) *n = 78*	2.6 (0.57) *n = 61*	0.1	0.34	.7316
Program characteristics (centers)					
Percent of programs where English is primary language spoken	82.3 *n = 79*	90 *n = 10*	–7.7	–0.61	.5442
Percent of programs receiving government funding	70.4 *n = 81*	80.0 *n = 10*	–9.6	–0.63	.5301
Mean number of children enrolled under 24 months of age	22.1 (10.89) *n = 49*	17.0 (6.67) *n = 5*	5.1	1.01	.3148
Mean baseline global quality scale score	3.5 (0.79) *n = 82*	3.6 (0.82) *n = 10*	–0.1	–0.53	.5977
Mean center staff Taking Care of Young Children[a] total sum score (program average)	81.6 (5.82) *n = 78*	83.0 (6.51) *n = 10*	–1.4	–0.69	.4893

Baseline measure	Retained at second follow-up	Not retained	Difference between groups	*t*-statistic or *z*-score	*p*-value
Mean center staff Modernity Scale[b] total sum score (program average)	77.4 (10.03) *n = 78*	72.6 (5.41) *n = 10*	4.8	1.48	.1419
Mean center director goals[c] (program average)					
Religious instruction	1.2 (0.61) *n = 74*	1.4 (0.70) *n = 10*	–0.2	–0.82	.4172
School readiness	2.6 (0.57) *n = 74*	2.2 (0.92) *n = 10*	0.4	1.42	.1856
Assisting disadvantaged children	2.6 (0.57) *n = 71*	2.8 (0.46) *n = 8*	–0.2	–0.69	.4898
Child care center staff characteristics					
Percent of center staff who speak English as primary language with the children	67.5 *n = 474*	76.4 *n = 55*	–8.9	–1.34	.1818
Percent of center staff who have a bachelor's degree or more education	24.5 *n = 469*	22.2 *n = 54*	2.3	0.37	.7098
Percent of center staff who received early childhood education or training in the last 12 months	86.6 *n = 469*	83.6 *n = 55*	3.0	0.60	.5512
Family child care staff characteristics					
Percent of family child care providers who speak English as primary language with the children	31.6 *n = 114*	48.2 *n = 81*	–16.6	–2.37	.0190*
Percent of family child care providers who have a bachelor's degree or more education	14.2 *n = 113*	15.2 *n = 79*	–1.0	–0.20	.8432
Percent of family child care providers who received early childhood education or training in the last 12 months	86.0 *n = 114*	78.2 *n = 78*	7.8	1.40	.1633

*$p < .05$, significantly different from zero at the .05 level, two-tailed test.

a. The Taking Care of Young Children questionnaire total sum score has a possible range of 28 to 112, with higher values indicating more positive attitudes about taking care of young children.

b. The Modernity Scale of Childrearing and Education Beliefs total sum score has a possible range of 30 to 150, with higher values indicating more traditional beliefs about raising young children.

c. Center directors and family child care providers were asked how important (not at all = 1, a little important = 2, or very important = 3) each of the following goals is for their child care program: to provide religious instruction, to prepare children for school with a strong academic curriculum, and to provide compensatory education for disadvantaged children.

Source: Authors' analysis of primary data collected for the study.

Appendix M. Child Care Licensing Regulations in the Study States

Table M1. Licensing staff qualifications by state

Arizona center directors	California center directors
Must have **one** of the following: • A high school diploma, two years of teaching experience, and 6+ child care/development units OR 60 hours of instruction • 24 college units, 6 being in child care/development; and 18 months of teaching experience • AA degree in a child care/development related field and six months of teaching experience • Bachelor's degree related to child care/development and three months of teaching experience • Obtain an N.A.C., C.D.A., C.C.P., or C.P.C. credential and 18 months of teaching experience	Must have 15 hours of health and safety training and complete **one** of the following: • A high school diploma, four years of teaching experience, 15 units related to child care/development. • An AA degree in a child care/development related field and two years of teaching experience. • A bachelor's degree related to child care/development and one year of teaching experience. • A child development site supervisor permit or a child development program director permit
Arizona center teachers	**California center teachers**
Must have one of the following: • High school diploma and six months of child care experience • An N.A.C., C.D.A., C.C.P., or C.P.C. credential • An AA or bachelor's degree related to child care/development	Prior to employment, must have **one** of the following: • 6 units related to child care/development, or a child development assistant permit • Possess a regional occupation program certificate of training in child care occupations, complete 95 hours of classroom instruction, and 150 hours in supervised field experience in a day care or group home. **After** employment, they must become a fully qualified teacher by: • Obtaining 12 units related to child care/development • Gaining six months of work experience • Obtain a C.D.A. credential • Obtain a child development permit
Arizona family child care providers	**California family child care providers**
Must meet the following requirements: • Be at least 21 years old • Complete department-provided orientation • Satisfy one of the following (A or B): A. Have a high school diploma and either three units related to child care/development OR 60 hours of training in child care/development B. Be registered as a Level II-B with S*CCEEDS • Every 12 months, complete one of the following: A. 12 hours of Arizona T3 training B. Complete one credit related to child care/development C. Complete 12 hours of training in two or more child care/development related subjects.	Must meet the following requirements: • Be at least 18 years old • Must have 15 hours of training on preventative health practices • If operating a large family day care, someone with a pediatric first aid and pediatric cardiopulmonary resuscitation certificate must be present at all times.

Source: Arizona Department of Health Services, Office of Child Care Licensing 2010a,b; California Department of Social Services, Community Care Licensing Division 2005, 2006.

Table M2. Required adult–child ratios in Arizona and California for centers and family child care homes, compared with the PITC

State and centers/homes	Age group	Adult–child ratio
PITC (family child care and centers)	Birth–18 months	1:3
	18–36 months	1:4
	Mixed	1:3 or 1:4[a]
Arizona		
Centers	Birth–12 months	1:5 or 2:11
	12–24 months	1:6 or 2:13
	24–36 months	1:8
Family child care	None specified	1:5 or 2:6-10
California		
Centers	Birth–18 months	1:3
	18–36 months	1:4
Family child care		
• Small child care home	Birth–24 months	1:4[b]
	Mixed	1:6 or 1:8[c]
• Large family child care home	Mixed	2:12 or 2:14[d]

a. The required ratio is 1:3 if more than two children are younger than 24 months; otherwise, it is 1:4.
b. If all children are younger than 24 months the required ratio is 1:4.
c. The required ratio is 1:6 if no more than three children are younger than 24 months. The required ratio is 1:8 if at least two children are over age 6 and no more than two are younger than 24 months.
d. The required ratio is 2:12 if no more than four children are younger than 24 months, six are preschool age (2–6), and two are school age (6+). The required ratio is 2:14 if at least two children are school age and no more than three children are younger than 24 months.
Source: Arizona Department of Health Services, Office of Child Care Licensing 2010a,b; California Department of Social Services, Community Care Licensing Division 2005, 2006.

Appendix N. Program for Infant/Toddler Care Professional Growth Incentives for Child Care Center Directors, Child Care Staff, and Family Child Care Providers (California)

Program for infant toddler care

Professional Growth Incentives
PITC in Practice:
Center-Based Directors or Site Supervisors
(Defined as those who make policy decisions within a Center-Based Program)
Effective July 1, 2007

Certificates of Participation
Participants who attend a minimum of 10 hours of *PITC in Practice* will be eligible to receive a Certificate of Participation at the completion of the services.

Cash Incentive: Cash incentives will be available when at least 75% of the total training and Reflective Action Planning (RAP) hours are held **during unpaid, non-work hours.**

For Section I of *PITC in Practice*, Directors or Supervisors will be eligible for a $175 cash incentive if he or she participates in the following events:

- Getting Acquainted session.
- 8 hours of introductory training (Establishing the Need, PITC Philosophy and Essential Policies, Temperaments, and Responsive Caregiving).
- 4 hours of meeting with the I/T Specialist to discuss progress.
- 4 hours of Reflection Action Planning.

For Section II of *PITC in Practice*, Directors or Supervisors will be eligible for a $175 cash incentive if he or she participates in at least the following events:

- 8 hours of training.
- 4 hours of meeting with the I/T Specialist to discuss progress.
- 4 hours of Reflection Action Planning.
- How Far We've Come session.

Note: In addition to the required sessions, Directors and Supervisors are strongly encouraged to attend the training sessions on topics that are specifically challenging for the program.

Resource Grant
Resource Grants for pre-approved infant/toddler learning materials will be available at the completion of a section for center-based programs in which more than 25% of the total training and RAP hours are held **during paid time**. The amount of the resource grant will be calculated at $175 per qualifying participant. The eligibility requirements are the same as those for the cash incentive (see above under cash incentive).

California PITC Partners for Quality –Professional Growth Incentive Eligibility Guidelines- November 2007

Incentives were identical in Arizona, except that academic units were not available.

Academic Units

Academic units are available to sites that are assigned to an approved Instructor. At the start of each Section, participants (including Directors or Supervisors) will be given the opportunity to enroll for 2 semester units to be earned during that Section.

- ✓ Center-Based Directors or Site Supervisors may choose to enroll for academic units in one or both Sections of *PITC in Practice*. However, they **will not** qualify for the cash incentive or the resource grant during any Section in which they are enrolled for academic units.
- ✓ Center-Based Directors or Supervisors enrolling for academic units must attend a minimum of **26** hours of training, **4** hours of Director Meetings and **4** hours of reflective action planning (RAP) in the Section in which they are enrolled and successfully complete the academic assignments.
- ✓ Academic units may be earned whether the training is taking place on paid or unpaid time.
- ✓ To enroll, a student must submit a $30 enrollment fee (per section) in the form of a money order or cashier's check, made payable to WestEd, and the required registration forms.

How does a Director or Supervisor qualify for a Professional Growth Incentive?

- ✓ Work with or supervise staff working with children birth-to-three for at least 20 hours a week in the program.
- ✓ Complete and submit a participant profile.
- ✓ Sign-in at every event that the Director and/or Supervisor attend (for example: Getting Acquainted Session, Director Meetings, trainings and RAP sessions).
- ✓ Participate in the required sessions as detailed above.

What will jeopardize my Professional Growth Incentive?

- ✓ Failure to sign-in at a training, Director Meeting, and/or RAP session that I attend.
- ✓ A 30-day gap in training that is not included in the pre-approved *PITC in Practice* schedule. See important notes below.
- ✓ Failure to complete the hours necessary to qualify.

Important Notes:

A 30-day gap in training, other than the gap(s) in the approved *PITC in Practice* schedule, will result in the loss of the cash incentive or resource grant. Participant sign-in sheets will not be reviewed for eligibility more than 30 days after the How Far We've Come.

Professional Growth Incentives
PITC in Practice:
Center-Based Teachers/Caregivers

Effective: July 1, 2007

Certificates of Participation

Participants who attend a minimum of 10 hours of *PITC in Practice* will be eligible to receive a Certificate of Participation at the completion of the services.

Cash Incentive

Participants who attend **28** or more training hours and **4** or more hours of reflective action planning (RAP) per Section in which at least 75% of the total training and RAP hours are held **during unpaid, non-work hours** will be eligible to receive a $175 cash incentive at the completion of the Section.

Resource Grants

Resource Grants for pre-approved infant/toddler learning materials will be available at the completion of a Section in which more than 25% of the total training and RAP hours are held **during paid time**. The amount of the resource grant will be calculated at $175 per participant who attends **28** or more hours of training and **4** or more hours of RAP per Section.

Academic Units

Academic units are available to sites that are assigned to an approved Instructor. At the start of each Section, participants will be given the opportunity to enroll for 2 semester units to be earned during that Section.

- ✓ Participants may choose to enroll for academic units in one or both Sections of *PITC in Practice*. However, they **will not** qualify for the cash incentive or the resource grant during any Section in which they are enrolled for academic units.
- ✓ Participants enrolling for academic units must attend a minimum of **26** hours of training and **4** hours of reflective action planning (RAP) in the Section in which they are enrolled and successfully complete the academic assignments.
- ✓ Academic units may be earned whether the training is taking place on paid or unpaid time.
- ✓ To enroll, a student must submit a $30 enrollment fee (per section) in the form of a money order or cashier's check, made payable to WestEd, and the required registration forms.

Please Note: In order to be eligible for *PITC in Practice* and the Professional Growth Incentives, participants must serve infants and toddlers, birth-to-three years old, and be employed in the program for at least 20 hours a week. This includes center-based teaching staff/caregivers, family child care providers, and license-exempt providers. If participants do not meet these requirements, they may participate if the host organization has approved and they may also enroll for academic credit by paying the full cost of $205 for the course.

How do I qualify for a Professional Growth Incentive?

- ✓ Work with children birth-to-three for at least 20 hours a week in the program.
- ✓ Complete and submit a participant profile.
- ✓ Sign-in at **EVERY** training and/or RAP session that I attend.
- ✓ Participate in **at least 28** hours of training and **4 hours of RAP** per Section under the approved plan for the cash incentive or resource grant.
- ✓ Participate in **at least 26** hours of training and **4 hours of RAP** per Section under the approved plan for academic credit.

What will jeopardize my Professional Growth Incentive?

- ✓ Failure to sign-in at a training or a RAP session that I attend.
- ✓ A 30-day gap in training that was not included in the approved *PITC in Practice* schedule. See important notes below.
- ✓ Failure to complete the hours necessary to qualify.

Important Notes:

A 30-day gap in training, unless included in the approved *PITC in Practice* schedule, will result in the loss of the cash incentive or resource grant. Technical assistance hours do not count toward the professional growth incentive. This includes the Getting Acquainted session and the How Far We've Come. Request for review of participant sign-in sheets will not be granted more than 30 days after the How Far We've Come.

Failure to sign-in at a training or a RAP session that the participant attended will require the participant to notify the I/T Specialist within 45 days of the event, at which time, the I/T Specialist will request a verification of attendance form from the main office. This form must be signed by the Director or host contact, I/T Specialist, and participant, then returned to the main office within ten days of receipt in order for the participant to receive credit for attendance.

**Eligibility requirements for Center-Based Directors and Site-Supervisors differ from the requirements on this handout. See separate handout, and/or PQ Application, Program Director Services Agreement for details.

Professional Growth Incentives
PITC in Practice:
Family Child Care Providers
(Including License-Exempt Providers)

Effective: October 1, 2009

Certificates of Participation
Participants who attend a minimum of 10 hours of *PITC in Practice* will be eligible to receive a Certificate of Participation at the completion of the services.

Cash Incentive
Participants who attend **28** or more training hours and **4** or more hours of reflective action planning (RAP) per Section, plus allow 1 observation and 1 hour of individual reflection in their family child care home, per Section, will be eligible to receive a $175 cash incentive at the completion of the Section.

Academic Units
Academic units are available to sites that are assigned to an approved Instructor. At the start of each Section, participants will be given the opportunity to enroll for 2 semester units to be earned during that Section.

- ✓ Participants may choose to enroll for academic units in one or both Sections of *PITC in Practice*. However, they **will not** qualify for the cash incentive during any Section in which they are enrolled for academic units.
- ✓ Participants enrolling for academic units must attend a minimum of **26** hours of training and **4** hours of reflective action planning (RAP) in the Section in which they are enrolled and successfully complete the academic assignments. In addition, participants must allow 1 observation and 1 hour of individual reflection in their family child care home, per Section.
- ✓ Academic units may be earned whether the training is taking place on paid or unpaid time.
- ✓ To enroll, a student must submit a $30 enrollment fee (per section) in the form of a money order or cashier's check, made payable to WestEd, and the required registration forms.

Please Note: In order to be eligible for *PITC in Practice* and the Professional Growth Incentives, participants must serve infants and toddlers, birth-to-three years old, and be employed in the program for at least 20 hours a week. This includes family child care providers, family child care assistants, and license-exempt providers. If participants do not meet these requirements, they may participate if the host organization has approved and they may also enroll for academic credit by paying the full cost of $205 for each course.

How do I qualify for a Professional Growth Incentive?

- ✓ Work with children birth-to-three for at least 20 hours a week in the program.
- ✓ Complete and submit a participant profile.
- ✓ Sign-in at **EVERY** training and/or RAP session that I attend.
- ✓ Fulfill the requirements listed above for either the cash incentive or the academic units.

What will jeopardize my Professional Growth Incentive?

- ✓ Failure to sign-in at a training or a RAP session that I attend.
- ✓ Not allowing the required observations in my family child care home.
- ✓ A 30-day gap in training that was not included in the approved *PITC in Practice* schedule. See important notes below.
- ✓ Failure to complete the hours necessary to qualify.

Important Notes:

A 30-day gap in training, unless included in the approved *PITC in Practice* schedule, will result in the loss of the cash incentive. Technical assistance hours do not count toward the professional growth incentive. This includes the Getting Acquainted session and the How Far We've Come. Request for review of participant sign-in sheets will not be granted more than 30 days after the How Far We've Come.

Failure to sign-in at a training or a RAP session that the participant attended will require the participant to notify the I/T Specialist within 45 days of the event, at which time, the I/T Specialist will request a verification of attendance form from the main office. This form must be signed by the host contact, I/T Specialist, and participant, then returned to the main office within ten days of receipt in order for the participant to receive credit for attendance.

Appendix O. Sensitivity of Impact Estimates to Alternative Model Specifications

Table O1. PITC impacts on child outcomes (primary outcomes) using listwise deletion for missing covariates

	Adjusted means					
Child outcome	**Treatment (standard deviation)**	**Control (standard deviation)**	**Difference (standard error)**	***p*-value (adjusted *p*-value)**	**Effect size**	**Sample size**
Language and cognitive development (composite of PLS-4 and BSRA)	–0.001 (0.737)	0.057 (0.819)	–0.058 (0.060)	.336 (.336)	–0.071	646
Positive Behavior (composite of Positive Behavior Scale and CBCL 1½-5)	–0.078 (0.895)	0.085 (0.734)	–0.163 (0.074)	.027 (.054)	–0.222	646

Note: Data were regression-adjusted using two-level regression models to account for differences in baseline characteristics and study design characteristics. Effect sizes were calculated by dividing impact estimates by the control group standard deviation of the outcome variable. Missing covariates were handled using listwise deletion. Effect sizes were calculated by dividing the impact by the standard deviation of the outcome for the control group.
Source: Child assessment data from follow-up two, including the PLS-4 and the BSRA conducted with the children, the Polit Positive Behavior Scale, and the CBCL 1½-5 from the parent follow-up questionnaire.

Table O2. PITC impacts on child outcomes (primary outcomes) using multiple imputation for missing covariates

	Adjusted means					
Child outcome	**Treatment (standard deviation)**	**Control (standard deviation)**	**Difference (standard error)**	***p*-value (adjusted *p*-value)**	**Effect size**	**Sample size**
Language and cognitive development (composite of PLS-4 and BSRA)	–0.020 (0.733)	0.050 (0.823)	–0.070 (0.056)	.220 (.220)	–0.086	796
Positive behavior (composite of Positive Behavior Scale and CBCL 1½-5)	–0.049 (0.893)	0.069 (0.755)	–0.119 (0.060)	.049 (.098)	–0.157	833

Note: Data were regression-adjusted using two-level regression models to account for differences in baseline characteristics and study design characteristics. Effect sizes were calculated by dividing impact estimates by the control group standard deviation of the outcome variable. Missing covariates were imputed using multiple imputation. Effect sizes were calculated by dividing the impact by the standard deviation of the outcome for the control group.
Source: Child assessment data from second follow-up, including the PLS-4 and the BSRA conducted with the children and the Polit Positive Behavior Scale and the CBCL 1½-5 from the parent follow-up questionnaire.

Table O3. PITC impacts on program quality (secondary outcomes) using listwise deletion for missing covariates

Program outcome	Adjusted means: Treatment (standard deviation)	Adjusted means: Control (standard deviation)	Difference (standard error)	*p*-value (adjusted *p*-value)	Effect size	Sample size
Global quality (ITERS-R/ FCCERS-R)	3.376 (0.993)	3.195 (0.909)	0.181 (0.138)	.190 (.190)	0.199	159
Staff-child interactions composite with interaction items from ITERS-R/ FCCERS-R and PITC-PARS	0.075 (0.691)	–0.104 (0.701)	0.179 (0.105)	.092 (.184)	0.255	159

Note: Data were regression-adjusted using (single-level) linear regression models to account for differences in baseline characteristics and study design characteristics. Effect sizes were calculated by dividing impact estimates by the control group standard deviation of the outcome variable. Missing covariates were handled using listwise deletion. Effect sizes were calculated by dividing the impact by the standard deviation of the outcome for the control group.
Source: Program measures from program observations at second follow-up

Table O4. PITC impacts on program quality (secondary outcomes) using multiple imputation for missing covariates

Program outcome	Adjusted means: Treatment (standard deviation)	Adjusted means: Control (standard deviation)	Difference (standard error)	*p*-value (adjusted *p*-value)	Effect size	Sample size
Global quality (ITERS-R/ FCCERS-R)	3.355 (0.989)	3.190 (0.902)	0.164 (0.131)	.214 (.214)	0.182	172
Staff-child interactions composite with interaction items from ITERS-R/ FCCERS-R and PITC-PARS	0.043 (0.693)	–0.116 (0.675)	0.159 (0.100)	.113 (.226)	0.236	172

Note: Data were regression-adjusted using (single-level) linear regression models to account for differences in baseline characteristics and study design characteristics. Effect sizes were calculated by dividing impact estimates by the control group standard deviation of the outcome variable. Missing covariates were imputed using multiple imputation. Effect sizes were calculated by dividing the impact by the standard deviation of the outcome for the control group.
Source: Program measures from program observations at second follow-up

Table O5. PITC impacts on program quality (secondary outcomes): Staff-child interactions composite excluding PITC-PARS measures

Program outcome	Adjusted means: Treatment (standard deviation)	Adjusted means: Control (standard deviation)	Difference (standard error)	*p*-value	Effect size	Sample size
Staff-child interactions composite without PITC-PARS measures	0.098 (0.779)	–0.129 (0.787)	0.227 (0.120)	.060	0.288	172

Note: Data were regression-adjusted using (single-level) linear regression models to account for differences in baseline characteristics and study design characteristics. Effect sizes were calculated by dividing impact estimates by the control group standard deviation of the outcome variable. Missing covariates were imputed using dummy variable missing data imputation. Effect sizes were calculated by dividing the impact by the standard deviation of the outcome for the control group.
Source: Program measures from program observations at second follow-up

Table O6. PITC impacts on child outcomes (primary outcomes) with adjustments for no-shows

Child outcome	Adjusted means: Treatment (standard deviation)	Adjusted means: Control (standard deviation)	Difference (standard error)	Difference adjusted for no-shows	Treatment sample size	Control sample size
Language and cognitive development						
Language and cognitive development (composite of PLS-4 and BSRA)	–0.184 (0.729)	–0.116 (0.816)	–0.068 (0.058)	–0.071	402	394
Positive behavior						
Positive behavior (composite of Polit Positive Behavior Scale and CBCL 1½-5)	–0.050 (0.899)	0.070 (0.751)	–0.121 (0.062)	–0.126	425	408

Note: PLS-4 is the *Preschool Language Scale, Fourth Edition,* Expressive Communication Subscale.
BSRA is the *Bracken School Readiness Assessment, Third Edition.*
CBCL 1½-5 is the Child Behavior Checklist for children ages 1½-5.
Data were regression-adjusted using two-level regression models to account for differences in baseline characteristics. Randomization strata were also included in the regression models to account for blocking by geographic region and program type. Effect sizes were calculated by dividing impact estimates by the control group standard deviation of the outcome variable. Missing covariates were imputed using dummy variable missing data imputation. The child sample for the language and cognitive development impacts include the 799 children who had complete data for either the PLS-4 or the BSRA, less 3children who had incomplete data for either the PLS-4 or the BSRA. The child sample for the positive behavior impacts include 833 children whose parents completed the parent follow-up questionnaire and therefore had data for the Child Behavior Checklist (CBCL 1½-5) and the Polit Positive Behavior Scale.
Source: Child assessment data from second follow-up, including the PLS-4 and the BSRA conducted with the children and the Polit Positive Behavior Scale and the CBCL 1½-5 from the parent follow-up questionnaire.

Table O7. PITC impacts on program quality (secondary outcomes) with adjustments for no-shows

Program outcome	Adjusted means: Treatment (standard deviation)	Adjusted means: Control (standard deviation)	Difference (standard error)	Difference adjusted for no-shows	Treatment sample size	Control sample size
Global quality (ITERS-R/ FCCERS-R)	3.351 (0.993)	3.198 (0.914)	0.154 (0.134)	0.158	82	90
Staff-child interactions composite w/ interaction items from ITERS-R/ FCCERS-R and PITC-PARS	0.034 (0.693)	–0.129 (0.692)	0.163 (0.102)	0.167	82	90

Note: Data were regression-adjusted using (single-level) linear regression models to account for differences in baseline characteristics. Randomization strata were also included in the regression models to account for blocking by geographic region and program type. Effect sizes were calculated by dividing impact estimates by the control group standard deviation of the outcome variable. Missing covariates were imputed using dummy variable missing data imputation. A total of 172 programs had completed observations for the second follow-up observation (82 centers and 90 family child care homes). In the 82 centers, 133 classrooms were observed (32 centers with one classroom and 50 centers with two classrooms). In centers where there were two classrooms, the scores were averaged across the two classrooms.
Source: Observation data from the second follow-up observations.

Table O8. PITC impacts on child outcomes (primary outcomes) excluding crossover programs

Child outcome	Adjusted means: Treatment (standard deviation)	Adjusted means: Control (standard deviation)	Difference (standard error)	Adjusted *p*-value	Effect size	Treatment sample size	Control sample size
Language and cognitive development							
Language and cognitive development (composite of PLS-4 and BSRA)	–0.043 (0.729)	0.023 (0.819)	–0.066 (0.059)	.265	–0.081	402	391
Positive behavior							
Positive behavior (composite of Polit Positive Behavior Scale and CBCL 1½-5)	–0.050 (0.899)	0.072 (0.750)	–0.122 (0.062)	.100	–0.163	425	405

Note: PLS-4 is the Preschool Language Scale, Fourth Edition Expressive Communication Subscale.
BSRA is the *Bracken School Readiness Assessment, Third Edition.*
CBCL 1½-5 is the Child Behavior Checklist for children ages 1½-5.
Data were regression-adjusted using two-level regression models to account for differences in baseline characteristics. Randomization strata were also included in the regression models to account for blocking by geographic region and program type. Effect sizes were calculated by dividing impact estimates by the control group standard deviation of the outcome variable. Missing covariates were imputed using dummy variable missing data imputation. The child sample for the language and cognitive development impacts includes the 799 children who had complete data for either the PLS-4 or the BSRA, less 3children who had incomplete data for either the PLS-4 or the BSRA. The child sample for the positive behavior impacts includes 833 children whose parents completed the parent follow-up questionnaire and therefore had data for the Child Behavior Checklist (CBCL 1½-5) and the Polit Positive Behavior Scale.
Source: Child assessment data from second follow-up, including the PLS-4 and the BSRA conducted with the children and the Polit Positive Behavior Scale and the CBCL 1½-5 from the parent follow-up questionnaire.

Table O9. PITC impacts on program quality (secondary outcomes) excluding crossover programs

Program outcome	Adjusted means: Treatment (standard deviation)	Adjusted means: Control (standard deviation)	Difference (standard error)	Adjusted *p*-value	Effect size	Treatment sample size	Control sample size
Global quality (ITERS-R/FCCERS-R)	3.353 (0.993)	3.191 (0.916)	0.162 (0.135)	.233	0.177	82	89
Staff-child interactions composite with interaction items from ITERS-R/FCCERS-R and PITC-PARS	0.034 (0.693)	–0.132 (0.695)	0.166 (0.102)	.214	0.239	82	89

Note: Data were regression-adjusted using (single-level) linear regression models to account for differences in baseline characteristics. Randomization strata were also included in the regression models to account for blocking by geographic region and program type. Effect sizes were calculated by dividing impact estimates by the control group standard deviation of the outcome variable. Missing covariates were imputed using dummy variable missing data imputation. A total of 172 programs had completed observations for the second follow-up observation (82 centers and 90 family child care homes). In the 82 centers, 133 classrooms were observed (32 centers with one classroom and 50 centers with two classrooms). In centers where there were two classrooms, the scores were averaged across the two classrooms.
Source: Observation data from the second follow-up.

Table O10. PITC impacts on child outcomes with fewer covariates (treatment indicator variable and randomization strata only)

Child outcome	Adjusted means: Treatment (standard deviation)	Adjusted means: Control (standard deviation)	Difference (standard error)	Adjusted *p*-value	Effect size	Treatment sample size	Control sample size
Language and cognitive development							
Language and cognitive development (composite of PLS-4 and Bayley Scales of Infant Toddler Development, Third Edition, Cognitive Subscale)	0.019 (0.729)	0.013 (0.816)	0.006 (0.077)	.942	0.007	402	394
Positive behavior							
Positive behavior (composite of Positive Behavior Scale and CBCL 1½-5)	–0.036 (0.899)	0.066 (0.751)	–0.102 (0.064)	.218	–0.136	425	408

Note: PLS-4 is the *Preschool Language Scale, Fourth Edition,* Expressive Communication Subscale.
BSRA is the *Bracken School Readiness Assessment, Third Edition.*
CBCL 1½-5 is the Child Behavior Checklist for children ages 1½-5.
Data were regression-adjusted using two-level regression models to account for differences in baseline characteristics. Randomization strata were also included in the regression models to account for blocking by geographic region and program type. Effect sizes were calculated by dividing impact estimates by the control group standard deviation of the outcome variable. Missing covariates were imputed using dummy variable missing data imputation. To create a positive behavior composite, the externalizing and internalizing domain component scores were multiplied by (–1) before creating the composite. The child sample for the language and cognitive development impacts includes the 799 children who had complete data for either the PLS-4 or the BSRA, less 3 children who had incomplete data for either the PLS-4 or the BSRA. The child sample for the positive behavior impacts includes 833 children whose parents completed the parent follow-up questionnaire and therefore had data for the Child Behavior Checklist (CBCL 1½-5) and the Polit Positive Behavior Scale.
Source: Child assessment data from second follow-up, including the PLS-4 and the BSRA conducted with the children and the Polit Positive Behavior Scale and the CBCL 1½-5 from the parent follow-up questionnaire.

Table O11. PITC impacts on program quality with fewer covariates (treatment indicator variable and randomization strata only)

Program outcome	Adjusted means: Treatment (standard deviation)	Adjusted means: Control (standard deviation)	Difference (standard error)	Adjusted *p*-value	Effect size	Treatment sample size	Control sample size
Global quality (ITERS-R/FCCERS-R)	3.483 (0.993)	3.145 (0.914)	0.338 (0.144)	.021	0.369	82	90
Staff-child interactions composite with interaction items from ITERS-R/FCCERS-R and PITC-PARS	0.034 (0.693)	–0.129 (0.692)	0.252 (0.105)	.034	0.364	82	90

Note: Data were regression-adjusted using (single-level) linear regression models to account for differences in baseline characteristics. Randomization strata were also included in the regression models to account for blocking by geographic region and program type. Effect sizes were calculated by dividing impact estimates by the control group standard deviation of the outcome variable. Missing covariates were imputed using dummy variable missing data imputation. A total of 172 programs had completed observations for the second follow-up observation (82 centers and 90 family child care homes). In the 82 centers, 133 classrooms were observed (32 centers with one classroom and 50 centers with two classrooms). In centers where there were two classrooms the scores were averaged across the two classrooms.
Source: Observation data from the second follow-up observations.

Table O12. PITC impacts on program quality with fewer covariates (treatment indicator variable, randomization strata, and baseline program quality score only)

	Adjusted means						
Program outcome	**Treatment (standard deviation)**	**Control (standard deviation)**	**Difference (standard error)**	**Adjusted *p*-value**	**Effect size**	**Treatment sample size**	**Control sample size**
Global quality (ITERS-R/FCCERS-R)	3.393 (0.993)	3.171 (0.914)	0.221 (0.132)	.095	0.241	82	90
Staff-child interactions composite with interaction items from ITERS-R/FCCERS-R and PITC-PARS	0.050 (0.693)	–0.136 (0.692)	0.186 (0.099)	.126	0.268	82	90

Note: Data were regression-adjusted using (single-level) linear regression models to account for differences in baseline characteristics. Randomization strata were also included in the regression models to account for blocking by geographic region and program type. Effect sizes were calculated by dividing impact estimates by the control group standard deviation of the outcome variable. Missing covariates were imputed using dummy variable missing data imputation methods. A total of 172 programs had completed observations for the second follow-up observation (82 centers and 90 family child care homes). In the 82 centers, 133 classrooms were observed (32 centers with one classroom and 50 centers with two classrooms). In centers where there were two classrooms, the scores were averaged across the two classrooms.

Source: Observation data from the second follow-up observations.

Appendix P. Additional Subgroup Analyses of Child Mobility

Table P1. Average months in study child care program: differences by children's ages at time of final assessment

Younger than three (*n* = 266)		Three years or older (*n* = 570)			
Mean	**Standard deviation**	**Mean**	**Standard deviation**	***t*-statistic**	***p*-value**
16.77	7.51	16.14	7.56	1.13	.261

Note: Months between random assignment and date child left program (or between random assignment and date of final assessment, if child remained in program) based on parent interview at final assessment
Source: Parent interview.

Table P2. Average months in study child care program: differences by type of care

Children in family child care homes (*n* = 223)		Children in child care centers (*n* = 613)			
Mean	**Standard deviation**	**Mean**	**Standard deviation**	***t*-statistic**	***p*-value**
15.69	8.05	16.58	7.35	–1.5	.133

Note: Months between random assignment and date child left program, (or between random assignment and date of final assessment, if child remained in program) based on parent interview at final assessment
Source: Parent interview.

Table P3. Baseline characteristics by treatment group for children who left the original care settings 23 months after random assignment

Measure	Overall	Intervention (standard deviation)	Control (standard deviation)	Difference between groups	*t*-statistic	*p*-value
Percent of children who are male	51.0	54.5	47.5	7.0	1.54	.125
Child's race/ethnicity						
Percent of children who are White, non-Hispanic	25.5	25.1	25.9	–0.8	–0.21	.837
Percent of children who are Black, non-Hispanic	7.1	6.5	7.7	–1.2	–0.53	.600
Percent of children who are Hispanic	56.3	56.3	56.3	0.0	0.0	.999
Percent of children living with two parents	60.3	58.9	61.6	–2.7	–0.61	.544
Percent of parents with a bachelor's degree or more education	32.8	32.4	33.2	–0.8	–0.19	.848
Percent of children where the primary language at home is English	73.2	73.3	73.1	0.2	0.05	.962
Mean child age in months at random assignment	18.0	18.0 (6.79)	18.1 (7.09)	–0.1	–0.12	.907
Mean hours per week child attends child care program	35.9	35.7 (10.85)	36.2 (10.22)	–0.5	–0.56	.572
Mean number of siblings	0.9	0.9 (0.94)	0.9 (1.04)	0.0	–0.48	.632
Mean parent age	28.3	27.8 (7.37)	28.8 (7.51)	–1.0	–1.43	.152
Mean number of hours parent is employed and/or in school	37.2	37.6 (10.77)	36.8 (11.29)	0.8	0.80	.427
Teacher-rated child language and behavior scales						
Child's mean score on the language scale	2.8	2.8 (1.18)	2.7 (1.28)	0.1	0.89	.372
Child's mean score on behavior scale	1.1	1.1 (0.44)	1.0 (0.40)	0.1	1.28	.200

Note: Sample size ranged from 456–461 on the caregiver-rated child language and behavior scales and from 488–494 on parent report of child and family characteristics, due to item-level missing data.
Source: Baseline parent questionnaire and baseline caregiver child form.

References

Abedi, J., & Dietel, R. (2004). Challenges in the No Child Left Behind Act for English language learners. *CRESST Policy Brief 7*. UCLA Center for the Study of Evaluation, UCLA Graduate School of Education & Information Studies.

Achenbach, T. M., & Rescorla, L. A. (2000). *The manual for the ASEBA preschool forms & profiles*. Burlington, VT: University of Vermont Department of Psychiatry.

Administration for Children and Families. (2002). *Making a difference in the lives of infants and toddlers: The impacts of Early Head Start.* Washington, DC: U.S. Department of Health and Human Services.

Arizona Department of Health Services, Office of Child Care Licensing. (2010a). Arizona Administrative Code, Substantive Policies, and Arizona Revised Statutes for Child Care Facilities: Interpretation and Clarification of Child Care Licensing Rules. Retrieved May 18, 2010, from http://www.azdhs.gov/als/childcare/ccc_forms/ccc_rules_statutes.pdf

Arizona Department of Health Services, Office of Child Care Licensing. (2010b). Arizona Administrative Rules, Substantive Policies, and Arizona Revised Statutes for Child Care Group Homes. Retrieved May 18, 2010, from http://www.azdhs.gov/als/childcare/ccc_gh_forms/gh_rules_statutes.pdf

Bailey, D., Bruer, J., Symons, F., & Licktman, J. (Eds.). (2001). *Critical thinking about critical periods*. Baltimore, MD: Brookes Publishing.

Bayley, N. (2006). *Scales of infant and toddler development—3rd edition*. San Antonio, TX: Harcourt Assessment.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B, 57*(1), 289–300.

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107,* 238–246.

Bloom, H. S. (1984). Accounting for no-shows in experimental evaluation designs. *Evaluation Review, 8,* 225–246.

Boller, K., Del Grosso, P., Blair, R., Jolly, Y., Fortson, K., Paulsell, D., et al. (2010). *The seeds to success modified field test: Findings from the impact and implementation studies*. Princeton, NJ: Mathematica Policy Research.

Bracken, B. A. (2007). *Examiner's manual: Bracken School Readiness Assessment* (3rd ed.). San Antonio, TX: Harcourt Assessment.

Briggs-Gowan, M. J., & Carter, A. S. (2005). *Brief infant-toddler social-emotional assessment* (Rev. ed.). San Antonio, TX: Harcourt Assessment.

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136–162). Newbury Park, CA: Sage Publications.

Bryant, D., Wesley, P., Burchinall, P., Sideris, J., Taylor, K., Fenson, C., et al. (2009). *The QUINCE-PFI study: An evaluation of a promising model for caregiver training*. Chapel Hill, NC: University of North Carolina, Frank Porter Graham Child Development Center.

Burchinal, M., Roberts, J., Nabors, L., & Bryant, D. (1996). Quality of center child care and infant cognitive and language development. *Child Development, 6,* 606–620.

Caldwell, B. M., & Bradley, R. H. (1984). *Home observation for measurement of the environment.* Little Rock, AR: University of Arkansas at Little Rock.

California Department of Education. (2007). *State schools chief Jack O'Connell introduces innovative infant/toddler publication.* Retrieved February 1, 2011, from http://www.cde.ca.gov/nr/ne/yr07/yr07rel14.asp

California Department of Social Services, Community Care Licensing Division. (2005). *Child care center general licensing requirements.* Retrieved May 18, 2010, from http://www.dss.cahwnet.gov/ord/entres/getinfo/pdf/ccc4.pdf

California Department of Social Services, Community Care Licensing Division. (2006). Manual of policies and procedures, community care licensing division, family child care homes, Title 22, Division 12, Chapter 3. Retrieved May 18, 2010, from http://www.dss.cahwnet.gov/ord/entres/getinfo/pdf/fccman.pdf

Campbell, P. H., & Milbourne, S. A. (2005). Improving the quality of infant-toddler care through professional development. *Topics in Early Childhood Special Education, 25*(1), 3–14.

Carnegie Corporation. (1994). *Starting points: Meeting the needs of our youngest children.* New York: Author.

Carter, A. S., & Briggs-Gowan, M. J. (2006). *Infant–toddler Social and Emotional Assessment (ITSEA): Examiner's manual.* San Antonio, TX: Harcourt Assessment.

Clifford, R. M. (2004, September/October). *Structure and stability of the Early Childhood Environment Rating Scale.* Keynote Address at the Quality in Early Childhood Care and Education International Conference, Dublin, Ireland.

Cryer, D., Tietze, W., Burchinal, M., Leal, D., & Palacios, J. (1999). The prediction of process quality from structural quality in preschool programs: A cross-cultural comparison. *Early Childhood Research Quarterly, 14,* 339–361.

Dearing, E., McCartney, K. M., & Taylor, B. A. (2009). Does higher quality child care promote low-income children's math and reading achievement in middle childhood? *Child Development, 80*(5), 1329–1349.

Dowsett, C. J., Huston, A. C., Imes, A. I., & Gennetian, L. A. (2008). Structural and process features in three types of child care for children from high and low income families. *Early Childhood Research Quarterly, 23,* 69–93.

Duncan, G. J., & National Institute of Child Health and Human Development Early Child Care. (2003). Modeling the impacts of child care quality on children's preschool cognitive development. *Child Development, 74*(5), 1454–1475.

Fenichel, E., Lurie-Hurvitz, E., & Griffin, A. (1999). Seizing the moment to build momentum for quality infant/toddler child care: Highlights of the Child Bureau and Head Start Bureau's

National Leadership Forum on Quality Care for Infants and Toddlers. *Zero to Three, 19*(6), 3–17.

Fiene, R. (2002). Improving child care quality through an infant caregiver mentoring project. *Child & Youth Care Forum, 31*(2), 79–87. Retrieved May 1, 2010, from http://www.childcareresearch.org/childcare/resources/2355

Flanagan, K., & West, J. (2004). *Children born in 2001: First results from the base year of the Early Childhood Longitudinal Study, Birth Cohort (ECLS-B)* (NCES 2005-036). Washington, DC: National Center for Education Statistics.

Frankenburg, W.K., Dodds, J., Archer, P., Shapiro, H., & Bresnick, B. (1992). The Denver II: A major revision and restandardization of the Denver Developmental Screening Test. *Pediatrics, 89*, 91-97.

Fukkink, R. G., & Lont, A. (2007). Does training matter? A meta–analysis and review of caregiver training studies. *Early Childhood Research Quarterly, 22*(3), 294–311.

Gross, D., Fogg, L., Young, M., Ridge, A., Cowell, J. M., Richardson, et al. (2006). The equivalence of the Child Behavior Checklist/1 1/2-5 across parent race/ethnicity, income level, and language. *Psychological Assessment, 18*, 313–323.

Harms, T., Cryer, D., & Clifford, M. (2006). *Infant/toddler environment rating scale-revised*. New York: Teachers College Press.

Harms, T., Cryer, D., & Clifford, M. (2007). *Family child care environment rating scale-revised edition*. New York: Teachers College Press.

Helburn, S. W., & Culkin, M. L. (1995). *Cost, quality, and child outcomes in child care centers.* Denver, CO: University of Colorado at Denver.

Honig, A. S. (2002). *Research on quality in infant toddler programs*. Washington, DC: U.S. Department of Education, National Center for Education Statistics. (ERIC Document Reproduction Service No. ED471910)

Honig, A. S., & Hirallal, A. (1998). Which counts more for excellence in childcare staff: Years in service, education level, or ECE coursework? *Early Child Development and Care, 145*, 31–46.

Howes, C. (2008). Ready to learn? Children's pre-academic achievement in pre-kindergarten programs. *Child & Youth Care Forum, 37*(1), 27–50.

Howes, C., & Brown, J. (2000). *Improving child care quality: A Guide for Proposition 10 commissions.* Los Angeles, CA: UCLA Center for Healthier Children, Families, and Communities.

Kontos, S., Howes, C., & Galinsky, E. (1996). Does training make a difference to quality in family child care? *Early Childhood Research Quarterly, 11,* 427–445

Kontos, S., Howes, C., Shinn, M., & Galinsky, E. (1995). *Quality in family child care and relative care.* New York: Teachers College Press.

Kreader, J. L., Ferguson, D., & Lawrence, S. (2005). Impact of training and education for caregivers of infants and toddlers. *Research-to-Policy Connections3*. Retrieved May 1, 2010, from http://www.childcareresearch.org

Mangione, P., Kriener-Althen, K., Niggle, M. P., & Welsh, K. (2006, June). *Program quality through the PITC lens: Assessing relationship-based care in infant/toddler early care and education programs*. Paper presented at the 15th National Institute for Early Childhood Professional Development. San Antonio, TX.

Mashburn, A. J., Pianta, R. C., Hamre, B., Downer, J. T., Barbarin, O., Bryant, D. M., et al. (2008). Measures of classroom quality in prekindergarten and children's development of academic, language, and social skills. *Child Development, 79*(3), 732–749.

Muthén, L. K., & Muthén, B. O. (1998–2010). *Mplus user's guide* (3rd ed.). Los Angeles, CA: Muthén & Muthén.

National Association of Child Care Resource & Referral Agencies. (2007). *Leaving children to chance: NACCRRA's ranking of state standards and oversight of small family child care homes*. Retrieved October 1, 2010, from http://www.naccrra.org/policy/state_licensing/

National Center for Early Development and Learning. (1999). *Critical Periods Synthesis Conference.* Retrieved October 1, 2010, from http://www.fpg.unc.edu/~ncedl/PAGES/project_summary.cfm?study_id=10

National Center for Education Statistics. (2009). *Early Childhood Longitudinal Study, Birth Cohort (ECLS-B) Longitudinal 9-month-Kindergarten 2007 Restricted-Use Data File and Electronic Codebook DVD* (NCES No. 2010011). Washington, DC: U.S. Department of Education.

National Child Care Information and Technical Assistance Center and the National Association for Regulatory Administration. (2007). *The 2007 child care licensing study.* Retrieved May 1, 2010, from http://naralicensing.org

National Infant and Toddler Child Care Initiative. (2010). *Core knowledge and competencies for infant and toddler caregivers.* Retrieved October 1, 2010, from http://nitcci.nccic.acf.hhs.gov/resources/Corekc.htm

National Research Council and Institute of Medicine. (2000). *From Neurons to Neighborhoods: The Science of Early Childhood Development*. Committee on Integrating the Science of Early Childhood Development. Jack P. Shonkoff and Deborah A. Phillips, eds. Board on Children, Youth, and Families, Commission on Behavioral and Social Sciences and Education. Washington, DC: National Academy Press.

NICHD Early Child Care Research Network. (2000). Characteristics and quality of child care for toddlers and preschoolers, *Applied Developmental Science, 4,* 116–135.

NICHD Early Child Care Research Network. (2005). *Child care and child development: results from the NICHD study of early child care and youth development*. New York: Guilford Press.

NICHD Early Child Care Research Network, and Duncan, G. J. (2003). Does quality of child care affect child outcomes at age 4 1/2? *Developmental Psychology, 39*(3), 451–469.

Overturf Johnson, J. (2005). *Who's minding the kids? Child care arrangements: Winter 2002.* Current Population Reports, P70-101. Washington, DC: U.S. Census Bureau, .

Panter, J. (2000). Validity of the Bracken Basic Concept Scale-Revised for Predictive Performance on the Metropolitan Readiness Test-Sixth Edition. *Journal of Psychoeducational Assessment, 18,* 104–110.

Peisner-Feinberg, E. S., Burchinal, M. R., Clifford, R. M., Culkin, M. L., Howes, C., Kagan, S. L., et al. (1999). *The children of the cost, quality, & outcomes study go to school: Public report.* Chapel Hill, NC: University of North Carolina, Frank Porter Graham Child Development Center.

Phillipsen, L., Burchinal, M., Howes, C., & Cryer, D. (1998). The prediction of process quality from structural features of child care. *Early Childhood Research Quarterly, 12,* 281–303.

Polit, D. F. (1996). *Parenting and child outcome measures in The New Chance 42-month survey*. Unpublished paper. New York: Manpower Demonstration Research Corporation.

Puma, M. J., Olson, R. B., Bell, S. H., & Price, C. (2009). *What to do when data are missing in group randomized controlled trials.* Washington, DC: U.S. Department of Education, National Center for Education Evaluation and Regional Assistance.

Quality Interventions for Early Care and Education. (2009). *Overview.* Retrieved May 1, 2010, from http://www.fpg.unc.edu/~QUINCE/index.cfm

Ramey, S., & Ramey, C. (2008, July). *The "right from birth" study: An evidence-informed training model to improve the quality of early child care and education.* Presentation at the Child Care Policy Research Consortium Meeting, Washington, DC.

Quint, J., Bos, H., & Polit, D. (1997). *New chance: Final report on a comprehensive program for young mothers in poverty and their children.* New York: Manpower Demonstration Research Corporation.

Raudenbush, S. W., & Liu, X. F. (2000). Statistical power and optimal design for multisite randomized trials. *Psychological Methods, 5*(2), 199–213.

Schaefer, E. S., & Edgerton, M. (1985). Parent and child correlates of parental modernity. In I. E. Sigel (Ed.), *Parental belief systems: The psychological consequences for children* (pp. 287–318). Hillsdale, NJ: Erlbaum.

Schafer, J. L. (1997). *Analysis of incomplete multivariate data*. London: Chapman & Hall.

Schochet, P. Z. (2008). Statistical power for random assignment evaluations of education programs. *Journal of Educational and Behavioral Statistics, 33*(1), 62–87.

Steiger, J. H. (1990). Structural model evaluation and modification: An interval estimation approach. *Multivariate Behavioral Research, 25,* 173–180.

Sylva, K., Siraj-Blatchford, I., Taggart, B., Sammons, P., Melhuish, E., Elliott, K., et al. (2006). Capturing quality in early childhood through environmental rating scales. *Early Childhood Research Quarterly, 21*(1), 76–92.

Thomason, A. (2009). Measuring the quality of teacher-child interactions in child care. *Early Education and Development, 20*(2), 285–304

Tucker, L. R., & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika, 38,* 1–10.

U.S. Department of Health and Human Services, Administration for Children and Families. (2002). *Making a difference in the lives of infants and toddlers and their families: The impacts of Early Head Start. Volume I: Final technical report*. Washington, DC: Author.

U.S. Department of Health and Human Services, Administration for Children Youth and Families. (2005). *National Survey of Child and Adolescent Well-Being (NSCAW): CPS Sample Component Wave 1 Data Analysis Report*. Washington, DC: Author.

U.S. Department of Health and Human Services. (2005). *Head Start impact study: first year findings*. Washington, DC: Author.

U.S. Department of Health and Human Services, Administration for Children and Families. (2010). *The Child Care Quality Rating System (QRS) Assessment: Compendium of quality rating systems and evaluations*. Washington, DC: U.S. Administration for Children and Families, Office of Planning, Research and Evaluation.

U.S. Department of Health and Human Services, Administration for Children and Families. (2009). *Early Head Start Child and Family Experiences Study*. Retrieved February 15, 2011, from http://www.acf.hhs.gov/programs/opre/ehs/descriptive_study/index.html#instruments

U.S. Department of Labor. (2009). *Women in the labor force: A databook*. Washington, DC: Author.

U.S. General Accounting Office. (2002). *States have undertaken a variety of quality improvement initiatives, but more evaluations of effectiveness are needed*. GAO 02-897. Washington, DC: Author

Vandell, D. L., & Wolfe, B. (2002). *Child care quality: Does it matter and does it need to be improved?* Madison, WI: University of Wisconsin-Madison, Institute for Research on Poverty. Retrieved May 1, 2010, from http://www.childcareresearch.org/childcare/resources/2144

Wechsler. D. (2003). *Wechsler Preschool and Primary Scale of Intelligence – Third Edition (WPPSI–III)*. San Antonio, TX: Pearson Assessment.

WestEd Center for Child & Family Studies. (2003). *Impact of PITC training on quality of infant/toddler care: Evaluation report.* San Francisco: Author.

WestEd Center for Child & Family Studies. (2007). *PITC-PARS program assessment rating scale: User's Guide for WREL Study.* San Francisco: Author.

WestEd Center for Child & Family Studies. (n.d.). *PITC Partners for Quality brochure.* San Francisco: Author.

Westfall, P. H., & Young, S. S. (1993). *Resampling-based multiple adjustment*. New York: John Wiley and Sons.

Whitebook, M., Howes, C., & Phillips, D. (1990). *The national child care staffing study*. Oakland, CA: Child Care Employee Project.

Zaslow, M., Tout, K., Halle, T., Whittaker, J. V., & Lavelle, B. (2010). *Towards the identification of features of effective professional development for early childhood*

educators. Washington, DC: U.S. Department of Education, Office of Planning, Evaluation and Policy Development.

Zimmerman, I. L., Steiner, V. G., & Pond, E. R. (2002). *Preschool Language Scale–Revised (PLS-R)* (4th ed.). San Antonio, TX: Harcourt Assessment.


DEPARTMENT OF EDUCATION
UNITED STATES OF AMERICA