Mathematica
Progress Together

Design Option Report for the We Grow Together Process and Implementation Study

June 2023

Tutrang Nguyen, Sally Atkins-Burnett, Shannon Monahan, Louisa Tarullo, and Yange Xue

OPRE Report #2023-181

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

Design Option Report for the We Grow Together Process and Implementation Study

OPRE Report #2023-181
June 2023
Tutrang Nguyen, Sally Atkins-Burnett, Shannon Monahan, Louisa Tarullo, and Yange Xue

Submitted to:

Office of Planning, Research, and Evaluation
Administration for Children and Families
330 C. St., 4th Floor
Washington, DC 20201
Project Officers Ann Rivera and Jenessa Malin

Submitted by:

Mathematica
1100 First St., NE, 12th Floor
Washington, DC 20002
Project Directors Louisa Tarullo and Shannon Monahan

Contract Number: HHSP233201500035I_HHSP23337007T

Acknowledgments

The authors would like to express their appreciation to Judy Cannon, John Deke, Lindsay Glassman, Ann Li, and Cheri Vogel at Mathematica. We thank Juliet Bromer, Margaret Burchinal, April Crawford, Carol Hartman, Toni Porter, Ronna Schaefer, Holly Wilcher, and Martha Zaslow for engaging in conversations with us about the design of the study.

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

Contents

Exhibits

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

Overview

Introduction

High-quality care and education for young children is important, particularly in the first three years of their lives. Early care and education (ECE) outside the home is now a normative experience, with over 50 percent of infants and toddlers regularly receiving nonparental care (Forry et al. 2018; Paschall 2019). Despite high demand, young children may have limited access to responsive and supportive ECE environments because infant and toddler care and education tends to be of low to moderate quality (Schmit and Matthews 2013).

To deliver quality services that promote positive outcomes for young children, effective professional development (PD) for the individuals who care for them is paramount. Effective PD can help infant and toddler caregivers understand quality ECE and how caregiver-child interactions support child development (Zaslow et al. 2010).

One such PD system specifically designed for infant and toddler caregivers is We Grow Together (WGT). WGT strives to improve the quality of care and education by helping infant and toddler caregivers and teachers[1] understand how their interactions with young children support child development and how they can implement best practices for supporting children's positive growth and development. The PD system includes training, strategies, and materials that are aligned with the research-based principles and practices of the Quality of Care for Infants and Toddlers (QCIT)[2] observational measure (Atkins-Burnett et al. 2015).

A 2018–2019 field test of WGT using a pre-post research design provided initial evidence of the potential of WGT to improve caregiver practices (Atkins-Burnett et al. 2020). Mathematica conducted the field test to understand the feasibility of implementing WGT in diverse early care and education (ECE) settings and to collect initial evidence about whether and how WGT could support caregivers in improving their knowledge, beliefs, and practices.

Purpose

This report presents a proposed study design that would enable the field to understand the feasibility and potential benefits of investing time in PD systems for infant and toddler caregivers. The focus of the proposed design is on contextual factors and processes that may influence the adoption, implementation, and potential sustainability of WGT.

To explore the question about the conditions under which WGT can be successfully implemented, the proposed design includes a pilot study and a process and implementation study. The pilot study would be a qualitative study focused on the feasibility of implementing WGT as well as addressing any barriers or challenges to implementation, with four Early Head Start (EHS) centers, four community-based centers, and four family child care homes (FCCs). The process and implementation study would use mixed methods, focusing on the conditions that may influence the use of WGT, its implementation, and the

[1] Throughout the remainder of this document, the term "caregivers" refers to nonparental caregivers and teachers in Early Head Start, community-based child care centers, and family child care homes.

[2] The Quality of Care for Infants and Toddlers (QCIT) measure was formerly called the Quality of Caregiver-Child Interactions for Infants and Toddlers (Q-CCIIT).

resources required. The process and implementation study would involve 40 EHS centers, 60 community-based centers, and 60 FCCs.

Research Questions

1. Pilot study

Three main research questions guide the design of the pilot study:

1. How can WGT be implemented using a team-based approach (professional learning communities [PLCs] with assigned modules) in center-based and FCC settings?
    a. What modifications to the field test approach do participants make in order to implement WGT? How do participants prefer to work with coaches (in person or virtually)?
    b. How do participants prefer to communicate with the implementation team about their use of the materials (phone, email, or text)? How could we improve the delivery of the training and materials to coaches and caregivers?
    c. How do participants prefer to structure their PLCs? What is the optimal frequency and structure of meetings and how does this approach vary across settings?
    d. Does WGT need to be adapted for implementation specifically in FCC settings? That is, might these settings need a slightly different version of WGT compared with center-based settings?
    e. What factors facilitate coaches' and caregivers' access to technology?
2. How can the time challenges associated with participating in PD be addressed?
    a. What conditions facilitate participation in WGT (for example, providing paid PD time, using available substitutes, or using technology and asynchronous coaching activities)?
    b. To what extent do the various approaches to encourage participation of coaches and caregivers increase the amount of time that participants spend using the system?
    c. How does gathering feedback on the use of the WGT PD system itself inform further improvements to WGT?
3. Does WGT meet caregivers' existing PD goals and requirements?

2. Process and implementation study

Eight research questions guide the design of the process and implementation study: two questions related to the conditions that may influence the use of WGT, three questions related to the implementation of WGT, and three questions related to resources.

Conditions

1. How does the classroom-level quality of caregiver-child interactions change between fall and spring?
    a. How do the changes in quality from fall to spring of caregiver-child interactions vary based on initial levels of quality (high, middle, low), experience levels, setting types, and age ranges of children in their group (infants versus toddlers)?
    b. Does the association between level of quality of caregiver-child interactions in the fall and spring vary by coaching experience or size of the PLC?

2. How do caregivers' beliefs and knowledge about ECE and caregiver self-efficacy change between fall and spring?
    a. How do the changes in caregiver beliefs and knowledge about ECE and caregiver self-efficacy from fall to spring vary based on experience levels, setting types, age ranges of children in their group (infants versus toddlers), average PLC beliefs and knowledge, and average PLC self-efficacy?

Implementation

3. What factors are associated with strong implementation of WGT—for example, setting type (EHS, community-based centers, FCCs); size of FCC or classroom and ages of children; level of support from center administration (where applicable); center or PLC climate; physical space; books, toys, and other sensory materials available for use by infants and toddlers; and caregiver beliefs about how to provide care and education to infants and toddlers?
4. What components of implementation are most strongly associated with positive changes in practice—for example, caregivers' level of engagement in the PLC, frequency of PLC meetings, frequency of attendance, level of use of WGT website, frequency of collecting videos of practices and reflecting on their use, and the extent to which participation by caregivers and coaches was sustained throughout the study year?
5. Do the associations between WGT components and positive change in classroom quality differ for classrooms with initial high and low levels of quality or with the PLC's focus?

Resources

6. What is the level of effort needed to implement WGT at the setting and network levels for center leaders, caregivers, coaches, and the study team—for example, time for coaching, for dedicated PD and instructional planning, for substitutes to cover for caregivers when they are meeting with coaches, for the study team to provide technical assistance?
7. What are the costs associated with materials needed to implement WGT—for example, technology costs such as devices, licensing fees, and internet access?
8. Does the level of effort needed to implement WGT vary by setting type and, if so, in what ways?

Highlights

This report describes a study design option that would address the proposed research questions. The research design team proposes exploring several modifications to the PD system, including assigning modules, conducting PD in a team structure, delivering virtual coaching, and making updates to the website's functionalities. The proposed study design seeks to extend findings from the field test by examining implementation of a team-based professional learning community (PLC) that supports caregivers. A PLC is a group of caregivers, or educators, that meet regularly to share their expertise and experience, and work collaboratively to solve common problems of practice (Borko 2004; Buysse et al. 2003; Stoll et al. 2006).

Drawing on what was learned in WGT field test, all caregivers in the proposed pilot study would start with modules and materials in the same domain (for example, Support for Language and Literacy Development). This approach would allow caregivers with the same coach or those in the same PLC to share, discuss, and practice the same content and receive feedback. The modifications to WGT could

allow for testing various approaches to coaching that could lead to cost savings and greater efficiencies for programs if PLCs or online coaching are as effective or more effective than in-person, individualized coaching. Coaches would facilitate the PLCs with groups of caregivers. Additionally, remote coaching could centralize training for coaches in a region or FCC technical assistance network and potentially offer more collaboration and support. PLCs within a setting could motivate and support caregiver change, sustain learning over time, and improve climate and organizational capacities.

Roadmap to the report

Chapter 1 summarizes the development of the WGT system, provides motivation for the proposed study, and presents the research questions and overview of the study design. Chapter 2 describes the design components of the proposed study design, as well as potential modifications of the WGT system. Chapter 3 details the approach to sampling and recruitment for the process and implementation study. Chapters 4 and 5 describe the data collection and analysis plans for the process and implementation study. Finally, this design report is accompanied by a separate appendix that details how WGT aligns with six nationally applicable standards and frameworks related to caregiver-child interactions that support the development of infants and toddlers (Li et al. 2023).

I. Introduction and Goals of the Study

A. Development of the We Grow Together system

Mathematica and its partners developed the We Grow Together Professional Development System (WGT) as part of the Professional Development Tools to Improve the Quality of Infant and Toddler Care project, which was funded in 2015 by the Office of Planning, Research, and Evaluation (OPRE) in the Administration for Children and Families (ACF) at the U.S. Department of Health and Human Services. Mathematica designed WGT to support ACF's vision for helping early care and education (ECE) programs and staff deliver quality services that promote positive outcomes for young children. The proposed study would examine the feasibility of implementing WGT in diverse ECE settings. This report provides an overview of a design option for a pilot study as well as a process and implementation study.

WGT strives to improve the quality of caregiving by helping infant and toddler caregivers and teachers[3] understand how their interactions with young children support child development and how they can implement best practices for supporting children's positive growth and development. The system includes training, strategies, and materials that are aligned with the research-based principles and practices of the Quality of Care for Infants and Toddlers (QCIT)[4] observational measure (Atkins-Burnett et al. 2015). WGT is organized into web-based modules, or sections, that describe caregiving practices to support infants' and toddlers' (1) social-emotional development, (2) language and literacy development, and (3) cognitive development. These modules are further divided into key practices that let caregivers explore the skills and exercise new caregiving techniques. Exhibit I.1 provides an overview of the WGT modules. Training and support materials are available on a user-friendly website that allows flexibility in selection and implementation of professional development (PD) for individualization of approaches. With WGT, coaches[5] support caregivers as they implement practices that promote infant and toddler's social-emotional, cognitive, and language and literacy development. The system is designed for use in center-based classrooms and family child care homes (FCCs), in both Early Head Start (EHS) and community-based settings.[6] In addition to the caregiver PD materials, the system provides a manual and training webinars for coaches as well as a set of materials and supplementary resources for coaches to use in supporting caregivers.

WGT's developers prepared a crosswalk of WGT with six competency frameworks and standards[7] related to caregiver-child interactions that support infant and toddler development to show the alignment in key

[3] Throughout the remainder of this document, the term "caregivers" refers to nonparental caregivers and teachers in Early Head Start, community-based child care centers, and family child care homes.

[4] The Quality of Care for Infants and Toddlers (QCIT) measure was formerly called the Quality of Caregiver-Child Interactions for Infants and Toddlers (Q-CCIIT).

[5] The term "coach" refers to a range of ECE staff who provide professional development, both within programs and employed by outside entities, such as managers and education directors, supervisors, mentors, coaches, employees of technical assistance networks or centers, and master teachers in the ECE setting.

[6] The term "classrooms" refers to both center-based and FCC settings. FCCs can include EHS and/or community-based providers.

[7] These competency frameworks and standards included Child Development Associate® credential, Collaborative for Understanding the Pedagogy of Infant/Toddler Development Competencies, Division for Early Childhood Recommended Practices, Head Start Early Learning Outcomes Framework, Head Start Program Performance Standards, and The ZERO TO THREE (ZTT) Critical Competencies for Infant-Toddler Educators™ (Critical Competencies).

practices. The appendix to this report includes these crosswalks. The research design team found that WGT generally aligned with the competencies and standards related to caregiver-child interactions that support children's development. For example, WGT modules and key practices align with five of the eight subject areas of the Child Development Associate® (CDA®) credential that are related to caregiver-child interactions that support children's development. However, WGT does not address areas such as partnering with families,[8] health and safety, and nutrition. Overall, this work suggests that WGT is aligned with various national competency frameworks and standards that are currently accepted in the field.

B. Context and rationale for the proposed study

High-quality ECE for young children is important, particularly in the first three years of their lives. Child care outside the home is now a normative experience, with over 50 percent of infants and toddlers regularly receiving nonparental care (Forry et al. 2018; Paschall 2019). Moreover, young children's developmental outcomes depend not only on their parents, but also on the care and interactions that occur in ECE settings (National Institute of Child Health and Human Development [NICHD] Early Child Care Research Network 2002). Despite high demand, young children may have limited access to responsive and supportive child care environments because infant and toddler care tends to be of low to moderate quality (Schmit and Matthews 2013).

Interactions between children and their nonparental caregivers are the essential ingredient in the quality of ECE settings (Halle 2011; Sroufe 1988, 2005). Growing knowledge of early brain development (Institute of Medicine [IOM] and National Research Council [NRC] 2015) and the lasting influence of the birth to age 3 period of life (Barnett and Masse 2007; Campbell et al. 2012; Conti et al. 2016) have informed the field's understanding of the critical role that infant and toddler caregivers play in supporting positive outcomes for children. Interactions with caregivers are the mechanism through which relationships form and children's early communication, learning, and competence unfold (Center on the Developing Child 2016; IOM and NRC 2012; Shonkoff 2017). However, on average, infant and toddler caregivers receive the lowest pay and are the least educated in the ECE field (Moreno et al. 2015; Norris 2010). Early childhood programs need cost-effective ways to provide high-quality PD for these critical professionals.

To deliver quality services that promote positive outcomes for young children, effective PD for the individuals who care for them is paramount. Research on PD for caregivers who work with preschool-age children has shown that PD can increase caregivers' knowledge of and skills related to effective interactions in their classroom (Hamre et al. 2012; Mashburn et al. 2010; Pianta et al. 2008; Zaslow et al. 2010). Effective PD can help infant and toddler caregivers understand quality care and how caregiver-child interactions support child development. Such PD can also help caregivers change their practices to better support children's development and learning. However, very few studies have examined the effectiveness of PD strategies in working with nonparental infant and toddler caregivers (Zaslow et al. 2010). Most early childhood research on PD including coaching has focused on the preschool years and involved caregivers who were more educated than the majority of infant and toddler caregivers (National Study of Early Care and Education [NSECE] Project Team 2013). Thus, more research is needed about the potential benefits of different forms of PD for infant and toddler caregivers (IOM and NRC 2015; Milli 2022; Zaslow et al. 2010).

[8] WGT has some materials related to partnering with families, such as family handouts for each practice, but this is not a major component of the system.

Exhibit I.1. We Grow Together modules

Module	Objectives Caregivers will learn and implement the following best practices:
 **Support Social-Emotional Development: Caregiver-Child Relationships**	• Responding to children's social cues • Responding to children's emotional cues • Responding to children in distress • Building a positive relationship • Supervising and joining in play and activities
 **Support Children's Language Use**	• Responding to children's cues • Taking turns in conversation • Asking questions • Extending children's language use • Supporting children's use of new words
 **Support Children's Understanding of Language**	• Using different types of talk • Using lots of specific and new words • Supporting learning about concepts • Engaging children in books • Using themes and projects
 **Support Social-Emotional Development: Regulation of Behavior and Emotions**	• Using responsive routines • Managing behavior and setting limits • Responding to emotional cues • Supporting self-regulation
 **Support Social-Emotional Development: Support Non-Mobile Infants' Peer Interactions**	• Supporting peer interaction and play • Creating a sense of belonging • Supervising and joining in play and activities
 **Support Social-Emotional Development: Support Toddlers' Peer Interactions**	• Supporting peer interaction and play • Extending pretend play • Supporting social problem-solving • Creating a sense of belonging
 **Support Literacy**	• Engaging children in books • Encouraging a positive attitude toward books • Using new words and sentences • Making connections to things not present
 **Support Infants' Cognitive Development**	• Supporting object exploration • Supporting children in making choices • Supporting learning about concepts • Extending knowledge about the world
 **Support Toddlers' Cognitive Development**	• Scaffolding problem-solving • Supporting children in making choices • Extending pretend play • Extending knowledge about the world

C. Field test of the WGT PD system[9]

Mathematica developed the WGT PD system to support nonparental caregivers' interactions with infants and toddlers. The content for caregivers is organized into nine web-based modules that were designed to support infants' and toddlers' social-emotional, language and literacy, and cognitive development. WGT's videos, handouts, and presentations introduce and demonstrate each key practice, which provides an understanding of the PD system's foundational principles and practices. Handouts are available to share with colleagues, supervisors, and families to ensure that the entire caregiving team is involved in supporting and using the key practices. Checklists and self-reflection questions on the WGT website offer support for collaborative progress monitoring, self-reflection, and feedback.

1. Study design and procedures of the WGT field test

A 2018–2019 field test of WGT using a pre-post research design provided initial evidence of the potential of WGT to improve caregiver practices (Atkins-Burnett et al. 2020). Mathematica conducted the field test to understand the feasibility of implementing WGT in diverse ECE settings and to collect initial evidence about whether and how WGT could support caregivers in improving their knowledge, beliefs, and practices.

WGT was implemented in 10 geographic areas in the United States, using local coaches and sampling from a range of ECE settings that served infants and toddlers across multiple localities. In selecting the 10 geographic areas, the research design team used information from public sources on state policies and systems as well as demographic information to achieve diversity by (1) region of the country, (2) state policies on caregiver-child ratios, (3) state requirements or standards for coaches, (4) states with PD registries, (5) states with infant and toddler specialists or mental health specialists, (6) local racial and ethnic composition, (7) children's home languages, and (8) household income. The research design team recruited from localities with ample ECE settings to obtain the coaches and caregivers who would participate in the field test. The team used the Head Start Program Information Report (PIR) to identify EHS programs and to form the geographic areas, some of which contained more than one EHS program. In addition to the PIR, the research design team also sampled from lists of certified center- and family-based child care providers (using the local child care resource and referral system, state quality rating and improvement system databases, and FCC provider networks). For EHS-based centers and classrooms, EHS programs were selected from the PIR. Then, the centers were chosen, and the classrooms were sampled within them. For community-based centers not partnered with EHS grantees, the research design team used the local child care resource and referral system to identify and select centers and classrooms. For community-based FCCs not partnered with EHS grantees, the research design team used FCC provider networks to identify and select coaches and the FCCs to which they were providing their services. The research design team maximized access to FCCs when choosing sites for the field test by recruiting a portion of them from those associated with EHS grantees.[10]

During the field test, caregivers were paired with a trained local coach. They used the WGT resources for up to four months. Before and after WGT implementation, trained observers rated caregiver-child interaction using the QCIT. The QCIT's strong reliability, sensitivity to variation in caregiving, and

[9] More information on the field test of the WGT PD system can be found at: https://acfmain-dev.acf.hhs.gov/opre/report/we-grow-together-professional-development-system-final-report-2019-field-test

[10] The sample included 214 center-based classrooms (89 affiliated with EHS and 125 community-based child care programs) and 57 FCCs, of which 16 were affiliated with EHS.

evidence of validity support its use in providing estimates of quality across and within caregivers (Nguyen et al. 2022). The QCIT offers the opportunity to identify strengths and challenges in caregiving in a variety of settings as well as the potential to test different approaches for improving caregiving for children. When used by trained observers during the WGT field test, the QCIT provided an initial measure of caregivers' strengths and areas for growth. Information from the QCIT shaped the selection of WGT modules for caregivers to work on, with the guidance of coaches. In the field test, the WGT PD system presented caregivers and coaches with different modules. They had the option to start with any one of them. Caregivers' QCIT total and domain scores were examined before and after participating in WGT.

2. Findings from the field test

On average, caregivers participating in WGT made modest improvements in their Support for Social-Emotional Development from fall 2018 to spring 2019. There were no mean differences between the fall 2018 and spring 2019 scores on the overall caregiver-child interaction quality, nor for the other two mean domain scores. EHS caregivers demonstrated improvement in both their Support for Social-Emotional Development and Support for Language and Literacy Development. Center-based, community-based, and FCC classrooms did not differ from fall to spring in average overall quality or on any of the domain scores.

In the field test, the positive associations between time spent on the WGT website[11] and improvements in the quality of caregiver interactions with infants and toddlers[12] supported the hypothesis about the usefulness of WGT in improving the quality of the interactions in ECE settings. Caregivers with greater participation on the WGT website had stronger spring 2019 scores in overall caregiver-child interaction quality, Support for Language and Literacy, and Support for Cognitive Development.[13, 14] However, the use of the WGT website was not associated with Support for Social-Emotional Development in spring 2019 after controlling for other factors.

Using surveys and web analytics, the research design team collected information about the characteristics, background, knowledge, beliefs, and user experiences of the caregivers and coaches who participated in the field test. On average, caregivers spent an average of nine minutes per week and an average of 3.9 total hours across the four months on the WGT website, based on web analytics data. Caregivers did not respond consistently to monthly requests about the time they spent on WGT outside of the website.

3. Implications for a new WGT study

The following are important considerations for a potential future WGT study:

- Whether the measures of dosage are feasible to collect
- The optimal dosage and frequency of coaching contact

[11] Measured by quartile for the number of WGT web pages opened by caregivers.

[12] Measured by QCIT scores in the spring, controlling for the fall QCIT scores and caregiver, classroom, and program characteristics.

[13] Caregivers working only in Support for Social and Emotional Development modules had fewer web pages available to them.

[14] The website did not capture the use of videos or other links to web pages outside of WGT.

- Caregivers' motivations for learning different strategies for supporting children's development

The initial field test of WGT did not include a process and implementation study, but rather was designed as a pre-post study. Therefore, the research design team did not collect reliable information on coaches' frequency of contact with caregivers, the practices that were discussed in coaching sessions, and what content and modules caregivers and coaches spent their time on. In post hoc analyses, the research design team examined the coach and caregiver characteristics that were most linked to using the system. The team found that when the coach attended fewer training webinars, the caregivers spent less time on WGT. EHS coaches, in particular, attended more training webinars than community-based coaches. This finding raises the question of whether caregivers would make greater progress in knowledge and practices if the coaches had greater knowledge of coaching or mentoring infant and toddler caregivers or if the PD was less individualized, which would allow for greater peer support from colleagues in the setting. Establishing additional evidence on the WGT coaching process is an important next step. Head Start Program Performance Standards require EHS programs to use a research-based, coordinated coaching system for caregivers.

D. Modifications to WGT and goals of the study

The field test provided initial evidence of improvements in the quality of caregiver interactions with infants and toddlers through the WGT PD system. Analyses examined observed quality before and after the use of WGT in relation to use of the website, after controlling for caregiver, classroom, and program characteristics. The proposed study design seeks to extend findings from the field test by examining implementation of a team-based professional learning community (PLC) that supports caregivers. A PLC is a group of caregivers, or educators, that meet regularly to share their expertise and experience, and work collaboratively to solve common problems of practice (Borko 2004; Buysse et al. 2003; Stoll et al. 2006).

Proposed modifications to the WGT PD system

- Assigning modules
- Using a team structure
- Offering virtual coaching
- Updating website functionality

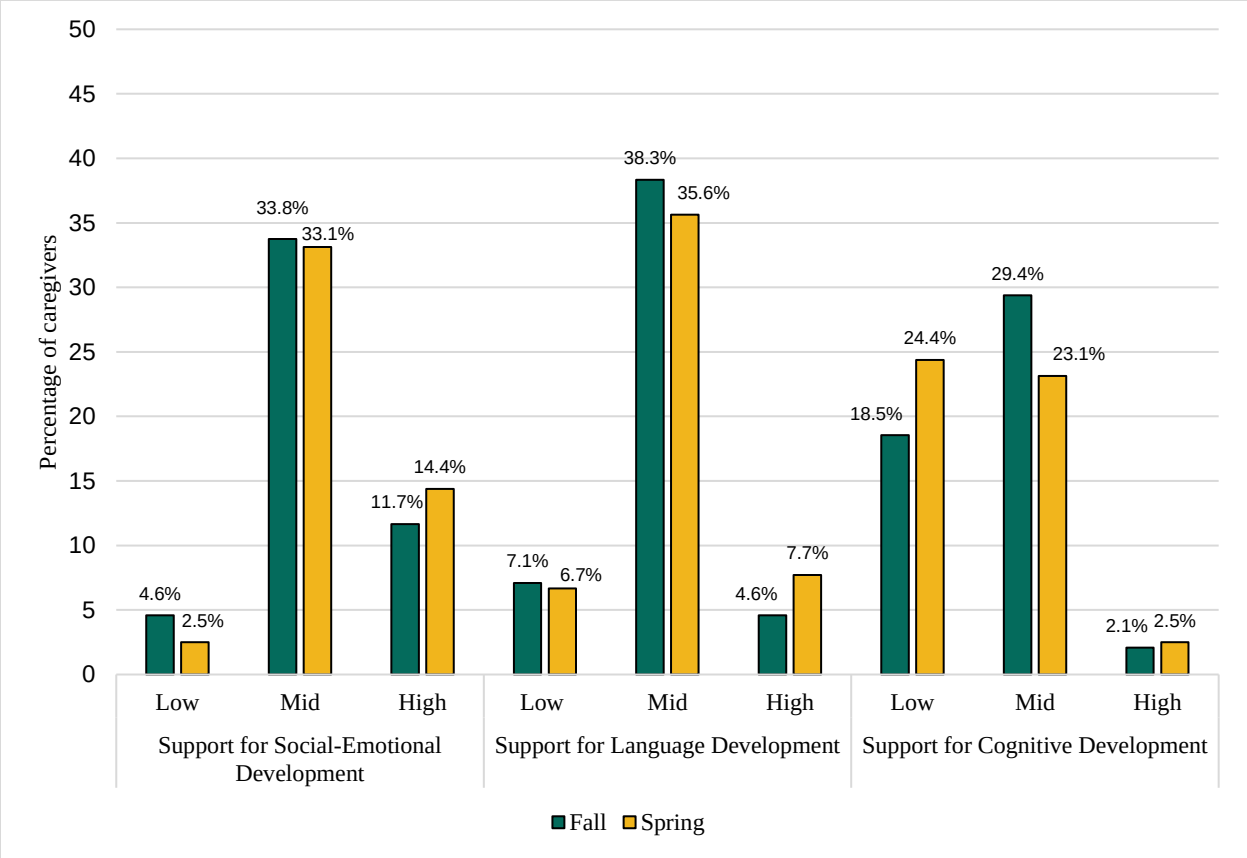
These modifications are discussed in more detail in Chapter II.

As discussed in later chapters of this report, the modifications to WGT would allow for testing various approaches to coaching that could lead to cost savings and greater efficiencies for programs if PLCs or online coaching are as effective or more effective than in-person, individualized coaching. Coaches would facilitate the PLCs with groups of caregivers. For centers, one or more PLCs would be formed for caregivers within each setting based on the center's size. For FCCs, the number of PLCs would vary based on the size of the network. A smaller network could have a single PLC. Remote coaching could centralize training for coaches in a region or FCC technical assistance network and potentially offer more collaboration and support. PLCs within a setting could motivate and support caregiver change, sustain learning over time, and improve climate and organizational capacities.

The focus of the current design option is on contextual factors and processes that may influence the adoption, implementation, and potential sustainability of WGT. Some evidence indicates that WGT supports changes in practice for caregivers with varying initial levels of quality (Atkins-Burnett et al. 2020). In additional exploratory analyses after the field test, the research design team examined factors associated with a change in QCIT scores. The team computed change scores by subtracting the fall QCIT W-score from the spring QCIT W-score. The team examined the descriptive statistics by their fall

baseline quality group (low, middle, and high quality) and found that caregivers who scored the lowest in the fall on the Support for Social-Emotional Development domain[15] made the most growth from fall to spring, suggesting that WGT may be most helpful for caregivers with the lowest initial quality interaction practices. Exhibit I.2 illustrates the distribution of these scores from fall to spring by domain. Between fall and spring, there was a change in the number of caregivers in the different categories of low, middle, and high QCIT scores. Similar to educational interventions where the lowest performing students make greater gains, this phenomenon would be more pronounced when there is a floor or ceiling to the measurement. Caregivers initially in the low-quality group (rating of 1 to 2.99) can demonstrate large positive change, but negative change is restricted. Similarly, caregivers in the high-quality group (above a rating of 5 on a 7-point scale) can only demonstrate limited positive change in practice but have the potential to demonstrate large negative change. In addition to the mean change score differences across groups, the range in the change scores in each domain was large, despite the potential ceiling effects. An important consideration to explore in the next study design would be how and under what circumstances WGT supports positive change in practices for caregivers.

Exhibit I.2. Number of caregivers in QCIT score categories, by domain in fall and spring



Source: WGT Field Test 2019 QCIT observations.

[15] Similar to how the Early Head Start Family and Child Experiences Survey (Baby FACES) team used QCIT scores, the research design team defined raw scores of 1 to 2.99 as the low group, 3 to 4.99 as the middle group, and 5 to 7 as the high group.

Using a team-based PLC approach might support caregivers at multiple levels of initial quality. The PLCs could allow those with the highest levels of quality to improve their practice as they mentor others. For example, prior research from K–12 has shown that professional support through teacher networks and collaboration enhance teachers' effectiveness in their classrooms (Burns et al. 2018; Lomos et al. 2011; Ronfeldt et al. 2015). An important consideration to explore in the next study design is whether to require the assignment of specific modules to groups of caregivers in the same PLCs. This approach would differ from the fully individualized approach that was taken in the prior field test of WGT. PLCs have been used to support PD in a variety of educational settings, including schools and child care programs (Bray et al. 2000; Cherrington 2011). One benefit to a group format such as a PLC—as opposed to individualizing modules based on a caregiver's individual QCIT score—is that it allows group members to ask each other questions, connect and share ideas, and focus on challenges and solutions that emerge from authentic situations in their work (Sheridan et al. 2009). By working on the same modules within a PLC, there would be a starting point for the discussions. Participants could then create and reflect on specific plans and feedback in their caregiving practices. This type of learning community can help create self-sustaining networks of educators and caregivers (Sheridan et al. 2009), which may encourage the implementation of WGT and its practices beyond the proposed study. Interpretation of the findings of the WGT field test as well as anecdotal comments from administrators led the research design team to recommend a team-based approach for delivery of WGT. In addition, other researchers have noted the benefits of collaborative participation in PD (Agosti et al. 2021; Douglass et al. 2015; Douglass et al. 2021).

Policymakers and program administrators will need convincing evidence of the feasibility and potential benefits of investing time in PD systems for infant and toddler caregivers. Program administrators want to know whether WGT will work for their caregivers and settings and what is needed for infant and toddler caregivers to benefit from WGT. Implementation data on the coaching and learning processes would help to establish a research-base for optimal coaching approaches. The field test identified challenges in implementation that should be addressed in order for programs to implement WGT effectively. The most pressing concern (and consistent with reports from other researchers)[16, 17] is ensuring that caregivers and coaches have time to meet and to use the WGT resources. Understanding how programs could address these time challenges may inform other quality improvement efforts. Moreover, required level of effort can be important in understanding the true costs of PD such as WGT. Level of effort is also important in that it might influence participants' involvement in other ways, such as the quality of participation and engagement. In addition, there are costs associated with implementation to support use of technology and materials.

The proposed design includes a pilot study followed by a process and implementation study. The pilot study would identify the feasibility of a team-based PLC approach for infant and toddler caregivers and inform the research team about ways to address potential challenges to this approach. For example, the pilot study would engage participants and center directors in generating solutions for the time challenges identified in the WGT field test, and in other recent early childhood PD efforts (Halle et al. 2021). The research team should also strive to recruit pilot sites that provide a combination of hybrid and in-person coaching approaches. Gathering information about implementation in each of these contexts would help inform the larger implementation and process study.

[16] Personal communication with April Crawford, April 13, 2022.

[17] Personal communication with Ronna Schaffer and Holly Wilcher, April 29, 2022.

E. Research questions

1. Pilot study

Three main research questions guide the design of the proposed pilot study:

1. How can WGT be implemented using a team-based approach (PLCs with assigned modules) in center-based and FCC settings?
    a. What modifications to the field test approach do participants make in order to implement WGT? How do participants prefer to work with coaches (in person or virtually)?
    b. How do participants prefer to communicate with the implementation team about their use of the materials (phone, email, or text)? How could we improve the delivery of the training and materials to coaches and caregivers?
    c. How do participants prefer to structure their PLCs? What is the optimal frequency and structure of meetings and how does this approach vary across settings?
    d. Does WGT need to be adapted for implementation specifically in FCC settings? That is, might these settings need a slightly different version of WGT compared with center-based settings?
    e. What factors facilitate coaches' and caregivers' access to technology?
2. How can the time challenges associated with participating in PD be addressed?
    a. What conditions facilitate participation in WGT (for example, providing paid PD time, using available substitutes, or using technology and asynchronous coaching activities)?
    b. To what extent do the various approaches to encourage participation of coaches and caregivers increase the amount of time that participants spend using the system?
    c. How does gathering feedback on the use of the WGT PD system itself inform further improvements to WGT?
3. Does WGT meet caregivers' existing PD goals and requirements?

2. Process and implementation study

Eight research questions guide the design of the proposed process and implementation study: two questions related to the conditions that may influence the use of WGT, three questions related to the implementation of WGT, and three questions related to resources.

Conditions

1. How does the classroom-level quality of caregiver-child interactions change between fall and spring?
    a. How do the changes in quality from fall to spring of caregiver-child interactions vary based on initial levels of quality (high, middle, low), experience levels, setting types, and age ranges of children in their group (infants versus toddlers)?
    b. Does the association between level of quality of caregiver-child interactions in the fall and spring vary by coaching experience or size of the PLC?
2. How do caregivers' beliefs and knowledge about ECE and caregiver self-efficacy change between fall and spring?

a. How do the changes in caregiver beliefs and knowledge about ECE and caregiver self-efficacy from fall to spring vary based on experience levels, setting types, age ranges of children in their group (infants versus toddlers), average PLC beliefs and knowledge, and average PLC self-efficacy?

Implementation

3. What factors are associated with strong implementation of WGT—for example, setting type (EHS, community-based centers, FCCs); size of FCC or classroom and ages of children; level of support from center administration (where applicable); center or PLC climate; physical space; books, toys, and other sensory materials available for use by infants and toddlers; and caregiver beliefs about how to provide care and education to infants and toddlers?
4. What components of implementation are most strongly associated with positive changes in practice—for example, caregivers' level of engagement in the PLC, frequency of PLC meetings, frequency of attendance, level of use of WGT website, frequency of collecting videos of practices and reflecting on their use, and the extent to which participation by caregivers and coaches was sustained throughout the study year?
5. Do the associations between WGT components and positive change in classroom quality differ for classrooms with initial high and low levels of quality or with the PLC's focus?

Resources

6. What is the level of effort needed to implement WGT at the setting and network levels for center leaders, caregivers, coaches, and the study team—for example, time for coaching, for dedicated PD and instructional planning, for substitutes to cover for caregivers when they are meeting with coaches, for the study team to provide technical assistance?
7. What are the costs associated with materials needed to implement WGT—for example, technology costs such as devices, licensing fees, and internet access?
8. Does the level of effort needed to implement WGT vary by setting type and, if so, in what ways?

F. Overview of the recommended study design

To explore the question about the conditions under which WGT can be successfully implemented, the proposed design includes a pilot study and a process and implementation study. The pilot study would draw on qualitative methods examining the feasibility of implementing WGT PLCs in varied settings, and addressing any barriers or challenges to implementation, with four EHS centers, four community-based centers, and four FCCs. The process and implementation study would use mixed methods, focusing on the conditions that may influence the use of WGT, its implementation, and the resources required. The process and implementation study would involve 40 EHS centers, 60 community-based centers, and 60 FCCs.

1. Pilot study

The research design team proposes that a small implementation and feasibility pilot study be conducted before launching the process and implementation study. This pilot study would focus on implementation and feasibility, using the Learn, Innovate, Improve (LI^2) approach to find solutions to the time challenges identified in the field test. The LI^2 approach draws on the experiences of caregivers, coaches, and

programs in iteratively designing, implementing, and testing solutions to address programmatic challenges. For example, the research team could conduct a series of collaborative sessions with coaches and administrators to understand the supports they would need to spend more time on WGT, such as PD release time; materials and other resources; or more effective communication from coaches and the research team.

Purposive sampling would be used for the pilot study so that the research team can gather information about what works for which caregivers under varying conditions. The research team should look for sites that already have coaches or a coaching infrastructure in place. This condition would ensure that the sites and caregivers are familiar with the coaches, and vice versa.

The research team could also consider selecting centers of different sizes to support understanding of potential differences in challenges. For example, a large center might opt to have one PLC for all caregivers to share strategies and practices for meaningfully interacting with children in their care. Alternatively, the larger centers may decide that implementation is stronger when they divide staff into more than one PLC.

The PLCs for FCCs would likely be formed through existing FCC network connections and coaching relationships. Prior relationships within the networks and with coaches may vary in the number of caregivers per coach, composition, and function. Familiarity among FCC providers in a PLC may differ and could pose challenges (such as trust issues and willingness to engage and share ideas).

It would be important for the pilot sites to participate in the qualitative data collection. This would include participating in person or having sufficient access to technology, if the research team decides to conduct virtual focus groups. Finally, it is crucial that the pilot sites commit to the duration of the study. Retaining and engaging the pilot sites would ensure a successful implementation of WGT for the process and implementation study.

The research team would recruit eight centers and four FCC homes to participate in a four-month pilot study. A four-month pilot study (for example, from September to December) would allow participants time to identify implementation challenges; solve problems; and test solutions (for example, meeting on-site during nap time with support staff coverage in classrooms or meeting on-site before the setting opens). LI^2 encourages participants to brainstorm and test their own ideas for addressing challenges. The research team could leverage this process and develop resources to share among PLC participants by, for example, describing how different PLCs addressed the time challenges they faced.

PLCs would be formed at the center level so that caregivers can work together in a team coaching format toward a common goal. The participating FCCs would form a single PLC. In a case study that examined the implementation of PLCs in FCCs (Gerdes and Jefferson 2015), a PLC was formed among six caregivers within a single county.[18] Qualitative evidence from this case study suggested that caregivers perceived the PLC format to work well. Specifically, caregivers reported that participating in the PLC encouraged their use of developmentally appropriate practices, use of reflection to improve their skills, and awareness of best practices.

In the pilot study, the focus would be on helping caregivers and coaches implement the PLCs and troubleshoot issues and challenges. The QCIT would not be collected. The research team would try out

[18] In this case study, participants were recruited through informational postcards that announced the formation of a PLC to 200 licensed FCCs in one county, using a systematic sampling procedure.

potential modifications to WGT in a subset of classrooms. For example, some centers and FCCs might be asked to participate in virtual coaching to determine its feasibility for the larger process and implementation study.

Completion of the pilot study in December would allow the research team to reflect on what was learned and give the team time to make changes to implementation or create additional supports for WGT to be implemented well. Lessons gained from the pilot study would inform the implementation of WGT in the subsequent larger study and potentially provide a range of solutions to frequently experienced challenges (such as finding time to meet). Importantly, this information would allow the research team to make any necessary adaptations prior to the process and implementation study, such as how the training and materials would be delivered to coaches and caregivers, how to ensure that enough time is spent on WGT, and how to facilitate access to technology.

2. Process and implementation study

The process and implementation study would use a mixed-methods approach, which is important for achieving the aims of this study because both methods would allow for comparing and converging the results for a complete understanding of how to implement WGT in real-world settings. Qualitative methods would provide a greater depth of understanding of the reasons that implementation is or is not successful and would allow for identification of strategies to facilitate implementation. Quantitative methods would provide a better understanding of the predictors of successful implementation (Tashakkori and Teddlie 2003). The research team would administer a number of instruments to caregivers and classrooms in order to understand the conditions that may influence the use of WGT, its implementation, and the resources required.

3. Timeline

The entire study would require two years of planning and preparation time, including applying for Office of Management and Budget (OMB) clearance and conducting a pilot study. This timeline includes an initial year for updating the website, developing materials for a PLC approach, obtaining institutional review board (IRB) clearance and generic OMB approval, selecting sites, and recruiting for the pilot study (Exhibit I.3). The second year would include conducting the four-month pilot study; making necessary adaptations and changes based on findings from the pilot; and finalizing the PD approach (for example, setting the recommended level of dosage) and measures for the process and implementation study (Exhibit I.4). It would also enable the research team to prepare and obtain approval for a full OMB package and IRB clearance and to recruit study participants. The process and implementation study would be conducted the following program year (the third year). This would include conducting a study in a range of ECE settings, including EHS and community-based centers and FCCs. The team would then analyze data and report findings.

Exhibit I.3. Pilot design and preparation for process and implementation study

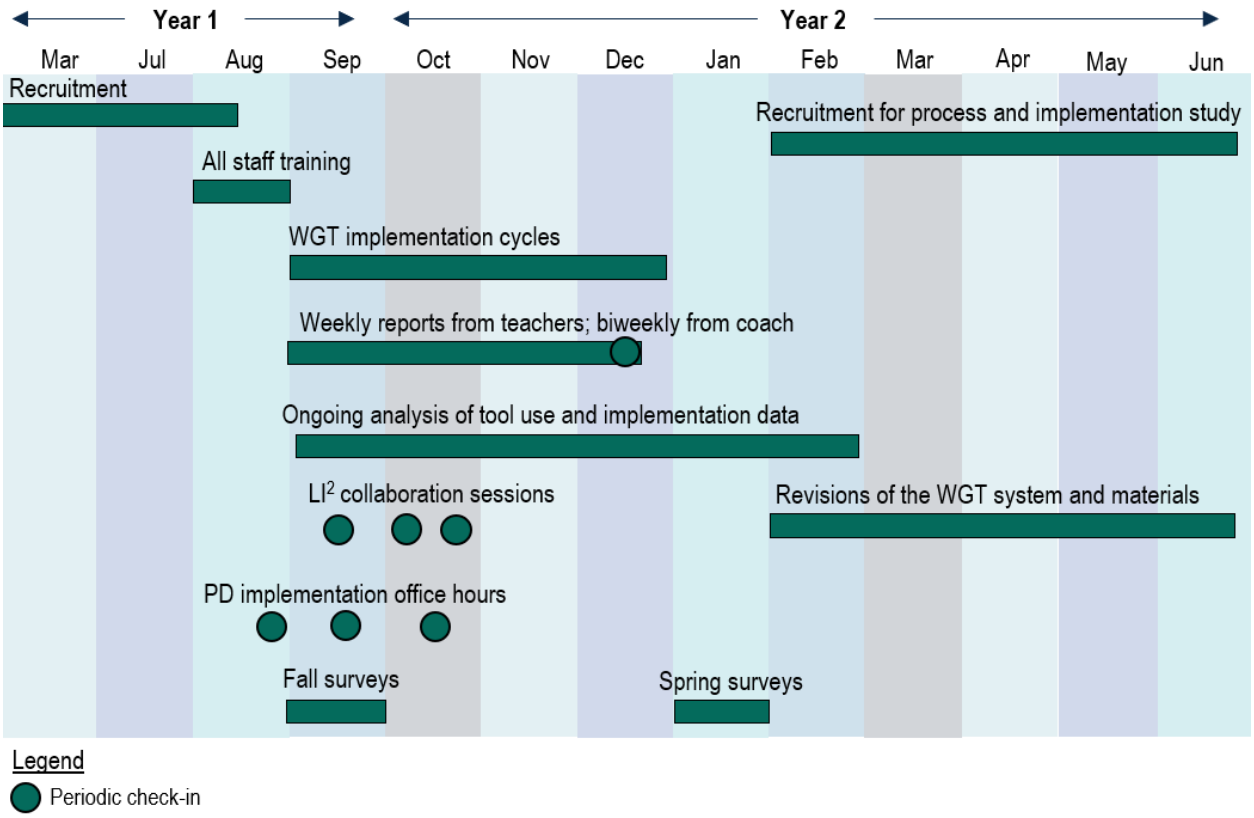
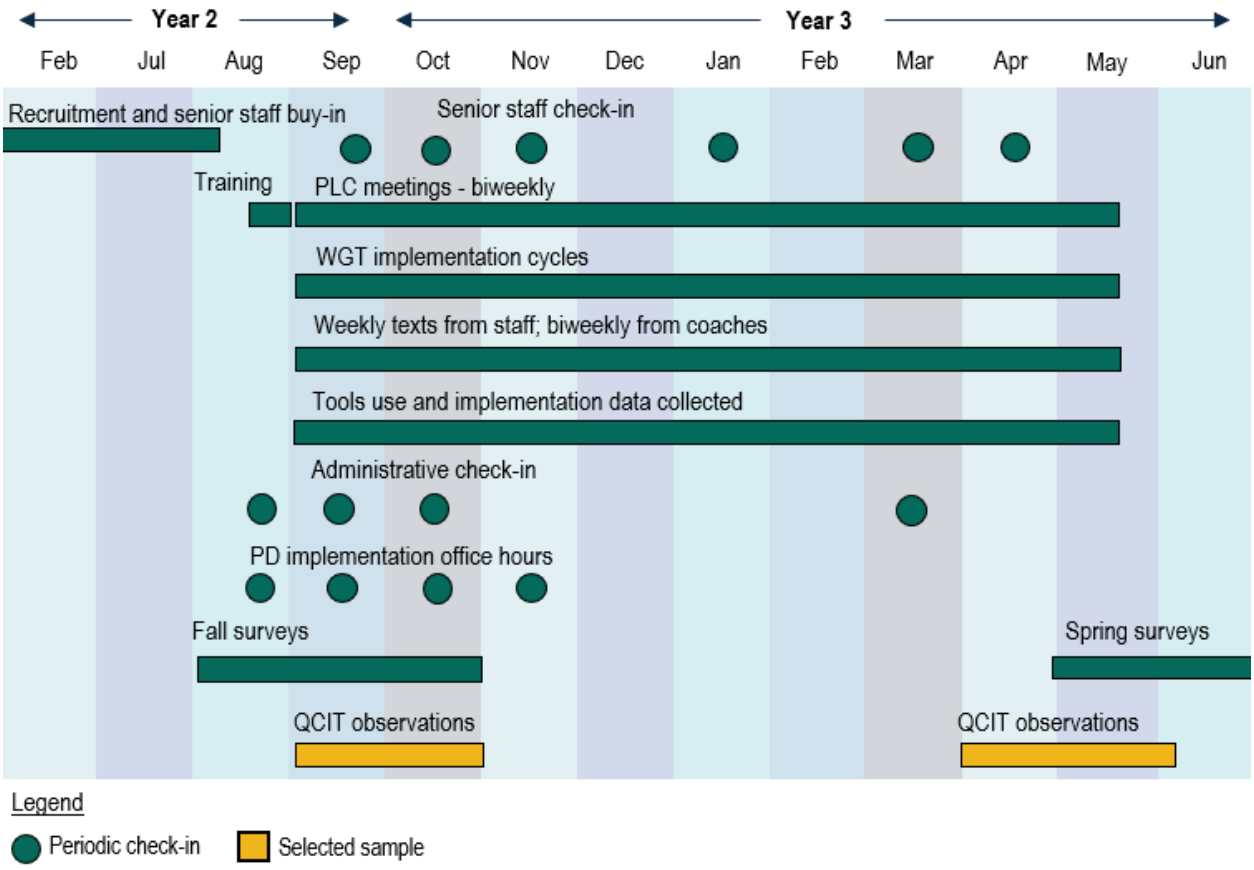


Exhibit I.4. Timeline for process and implementation study



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

II. Design Components

A. Potential modifications to the WGT PD system

Before launching into any potential future study of WGT, the research design team proposes several modifications to the PD system, including assigning modules, conducting PD in a team structure, delivering virtual coaching, and making updates to the website's functionalities.

1. Module assignment

Drawing on what was learned in WGT field test, all caregivers in the proposed pilot study would start with modules and materials in the same domain (for example, Support for Language and Literacy Development). This approach would allow caregivers with the same coach or those in the same PLC to share, discuss, and practice the same content and receive feedback. Qualitative studies in education (Battey and Franke 2008; Schoenfeld 2018) and some in early childhood (Kuh 2012) have shown support for PD where teachers collectively work together on the same content. This approach allows them to discuss both the content and the ways in which they can support children's development with the new practices they learn. After testing in the pilot study, in the proposed implementation study, the majority of the caregivers would start in the same module to maximize opportunities for discussion and to limit variation in implementation. The research team would present modules that were most often selected and showed greater change in practice with one module from each domain from the initial field test, and the group of caregivers would select the one to use—with all working on the same module. However, because the content for infants and toddlers is differentiated, the research team would vary the starting module based on the ages of the children in the classrooms (infant or toddler).

2. Team structure

Because WGT was originally designed for coaches to work one-on-one with a caregiver, the research team would need to modify the system so that it can be used with a larger team of caregivers and their coaches within a PLC. In education, school-level interventions are often more effective than individual approaches—for example, schoolwide Problem Behavior Interventions (Bradshaw et al. 2008; Bradshaw et al. 2009). The PLC approach could foster greater investment of caregiver time and mutual support for improvement in the selected domain. Because caregivers know what they are experiencing the best, PLCs would not be didactic. Rather, the PLCs would be a space for caregivers to share their expertise and experience with others, supported by a coach. It would also be designed as a space for caregivers and coaches to brainstorm strategies and solutions that would lead to equitable change. During the WGT field test, many caregivers asked about sharing materials with their teaching teams. In addition, both coaches and caregivers reported challenges with finding time to implement WGT. Some caregivers received little individual support from their coach. Therefore, the research team could investigate whether creating and piloting PLCs motivates and provides more support for caregivers (Harris and Jones 2010; Stoll et al. 2006; Wesley and Buysse 2001).

The research team would need to update existing materials and create new materials for coaches to provide guidance for using the WGT website with a team of caregivers. The switch from individualized to team coaching in the next phase of WGT would require training coaches on how to work with a team of caregivers or FCC providers. Additionally, the research team would infuse the coaches' training with aspects of the PLC approach that coaches would use with caregivers. For example, the research team can

model PLC techniques at the coaches' training. This can help coaches develop relationships with other coaches to continue to support each other on their own during their training and work with the caregivers. Training for the coaches and PLCs for the caregivers would focus on their assets and successes rather than perceived needs or problems. The caregivers would also need materials that support their learning within a team structure. For example, rather than just individual exploration, these materials may provide more structure for learning by including discussion topics, meeting agendas that are suitable for a group setting, and facilitation guides.

3. Virtual coaching

Virtual coaching could be tested in the pilot study with a small subset of classrooms. The research team would explore the prevalence and feasibility of virtual coaching during the pilot study. Virtual coaching would allow for more coaching support for caregivers who may be faced with geographic or time constraints or concerns with COVID-19. This type of coaching format would allow coaches to work with more caregivers in a shorter amount of time by eliminating travel demands. Caregivers in rural or remote settings could access support from coaches through the use of virtual coaching and other online tools. Research studies have previously shown that positive and supportive coaching relationships can be developed virtually, and therefore facilitate opportunities for teachers to practice new skills (Crawford et al. 2021; Downer et al. 2009; Hirschler and Darcy 2002; Schacter et al. 2019). If WGT were modified to allow for a virtual coaching option, the research team would need to provide technical support to facilitate the sharing of video recordings to observe and reflect on caregiver practices. For example, this approach could be tested with Coaching Companion, an online video sharing and coaching feedback tool used in Head Start.

4. Website functionalities and other updates

Because the last update of WGT was during the 2018–2019 field test, the platform that hosts the WGT website may need to be updated for new functionalities and security. Modifications would be needed to align with the new approach for assigning modules. In addition, URLs in resources would need to be checked. Finally, any new or revised materials and resources would need to be added to the site.

B. Pilot study

The pilot study would help test the feasibility of the design as well as modifications to the system (for example, assignment of modules) and ensure the best test of WGT in real-world settings. Understanding how and what affects adoption, implementation, and sustainability of PD for infant and toddler caregivers is critical for fostering selection and use of WGT.

This pilot study would largely be a qualitative study. However, some quantitative data may also be collected. The pilot study would provide information that would help the research team find solutions to some of the implementation challenges raised in the field test and might generate unique solutions for infant and toddler caregiver PD. For example, the biggest challenge for caregivers was finding the time needed to learn and implement the recommended practices in WGT. The pilot study would help illuminate the supports that are needed, such as PD release time, substitutes, training materials, other resources, or more effective communication from coaches. The research team might also be able to gather information from the pilot study on whether virtual or in-person coaching is better at supporting implementation and for whom. For example, in centers with a larger staff, it may be easier to find time for coaches and caregivers to meet when coaching is provided in person through small, frequent doses.

Alternatively, smaller numbers of staff may mean that there are no substitutes available to staff classrooms and virtual coaching may be needed to add more flexibility for meeting times.

In the field test, the research design team did not receive a high response rate using pop-up surveys to collect data on dosage when the caregivers logged on to the website. In the pilot study, the research team could try other methods for collecting dosage data, such as text messaging. The research team should test the feasibility and use of text messaging to collect data in the process and implementation study. The early childhood literature suggests that text messaging as a form of communication with caregivers is increasingly common. Text messaging has been used to encourage caregivers (Hanno 2022) and parents (Smythe-Leistico and Page 2018; York et al. 2019) to engage in high-quality interactions with children as well as to deliver performance feedback (Barton et al. 2019) and PD to caregivers (Devers and Devers 2018).

Studies in the behavioral sciences—for example, the Building Bridges and Bonds (B3) Study of Responsible Fatherhood Programs—have reported that text messaging can be useful for learning about participants' experiences (B3 Study Team 2022). In health care, researchers have found text messaging to be a reliable and valid method for collecting research data (Shimoni et al. 2020; Whitford et al. 2012). However, we did not find research indicating that this method is a valid and reliable option for infant and toddler caregivers.

An alternative to collecting real-time dosage data, as was used in the initial WGT field test, is using a survey with retrospective questions to caregivers and coaches. Some studies in education have found that data from retrospective, web-based surveys yield similar responses to real-time text message survey data (Olson et al. 2022). However, retrospective data in the WGT field test had some unrealistic responses. For example, some caregivers reported that they worked on WGT for more months than it was available to them. Important considerations for collecting data with text messages are the response rates and respondent burden. The research team would use the pilot study to examine the feasibility of text message data collection before introducing it in the larger study.

The research design team proposes examining implementation in the pilot study using a three-phase approach called Learn, Innovate, Improve (LI^2). In doing so, changes can be made to ensure high fidelity and allocation of the resources necessary for successful implementation. It would also provide a basis for better planning, evaluation, and interpretation of the outcome data in the process and implementation study.

LI^2 is a systematic, evidence-based, human-centered design process for engaging administrators, coaches, and a select group of caregivers in identifying challenges and solutions for "road testing." This process is evidence-based in terms of content (drawing upon the best available evidence and science); process (delivering assistance in accessible, learning-centered formats); and the co-creation of new evidence (through locally driven, decision-focused evaluations). The research team would conduct the pilot study with coaches and caregivers using LI^2 in a virtual format. The coaches would work with caregivers in the format that they typically use, whether that is virtual or in person. This would allow the research team to see what coaches and caregivers are most likely to use, which would inform the format for the process and implementation study. The research team could collect audio recordings from the PLC meetings to develop a checklist of discussion topics that are related to WGT implementation. This checklist would then be used in the process and implementation study.

LI^2 is appropriate for use in this pilot because the steps would guide coaches and caregivers through an incremental change process focused on achieving realistic goals for improving PD in the local context.

The approach emphasizes collaboration and responsive, frequent communication between the research team, coaches, and administrators, which ensures that the solutions to be tested are tailored and responsive to sites' needs and priorities.

LI^2 consists of three phases. – Learn, Innovate, Improve. The Learn phase would focus on listening to and understanding the concerns and challenges that caregivers experience when participating in PD, what motivates them the most in joining and actively participating in a PD program, and how the team approach to learning might support them. The Innovate phase would involve working with caregivers, administrators, and coaches to brainstorm, prioritize, and design solutions that address the identified challenges (for example, availability of time to meet) and leverage the caregiver-reported sources of motivation. The Improve phase would implement the agreed upon solution for two to four weeks and gather both quantitative (for example, number of times the caregivers meet to discuss practice and number of caregivers who collect and share videos of practice) and qualitative data (for example, interviews with coaches and program directors about their perceptions of the implementation of WGT, the success of the innovation, and remaining challenges). Given the cyclical process of the LI^2 steps, the research team would cycle through the phases again as time allows to consider additional changes and improvements that might facilitate the caregivers' ability to benefit from WGT. This ongoing process would continue for about four months.

Example of LI^2: Increasing engagement in the Colorado Works Program

Mathematica worked with Arapahoe County to engage parents with low incomes in the Colorado Works program, the state's version of Temporary Assistance for Needy Families.

Learn: Mathematica and Arapahoe County worked to understand and document the factors that led to low initial engagement in Colorado Works. A series of interviews with staff revealed challenges in communicating with families and completing referrals to employment providers.

Innovate: Arapahoe County Human Services staff designed an intervention comprised of strategies such as decreasing paperwork demands (thus reducing staff and family time spent on this) and clearly communicating to families the value of engagement with the workforce provider. Staff wanted parents to leave the eligibility assessment with a clearer understanding of next steps. The intervention was designed using a road map for change, a tool that helps link the desired outcomes to specific targets (such as shifting parents' attitudes or behaviors) and to evidence-informed strategies.

Improve: In partnership with Mathematica, Arapahoe County launched an experiment of its new approach by randomly assigning about half the team's staff to the intervention while the other half continued what they had been doing. The county tracked parents' immediate sentiments and their attendance rate at orientation. Both groups of parents who received the intervention and the services the usual way reported highly positive experiences and had similar attendance rates. Drawing on the insights from this short experiment, the county has returned to innovating and is focused on addressing the challenges described by eligibility staff. Arapahoe County plans to revise its road map and move toward testing a different approach, using the lessons learned from this experience to improve parents' experiences with the program and in turn increase their engagement.

C. Process and implementation study

The process and implementation study would take a multiple-method, multiple-informant approach to collecting detailed information necessary to address the research questions. The study would explore the

natural variation in conditions that may influence the use of WGT, its implementation, and the resources necessary for a successful implementation. This approach is necessary for describing the relationships among caregivers, coaches, administrators, system and organizational factors, and their roles in the implementation process. Qualitative and quantitative methods would be used for seeking complementarity (Tashakkori et al. 1998; Tashakkori and Teddlie 2003)—that is, using quantitative methods to achieve breadth of understanding (in other words, generalizability) of both content fidelity and outcomes (degree of implementation, experiences with implementation, engaging in PD and PLCs). Qualitative methods would help the research team achieve a depth of understanding of the implementation process and context.

Key components of the process and implementation study would include the following:

- **Modifying WGT based on what was learned from the pilot study.** Based on what the research team learns from the pilot study, the team would evaluate aspects of WGT that need to be modified prior to the process and implementation study. For example, it would be important to learn more about how WGT can be delivered to caregivers in FCC settings and how best to build PLCs to support their implementation of WGT. Although some FCC networks have providers who meet regularly, other providers may be more isolated. It may be that caregivers in FCCs would need to be in PLCs that meet virtually. It would also be important to learn more about the potential options for implementation. For example, within a setting type and location, half of the centers (and FCC PLCs) could be assigned to implement one of two options. These options could include (1) assigning one group to focus on a new domain every two months versus allowing the other group to progress through a selection of modules from each domain as the group decides or (2) varying the frequency of the PLCs. The final decisions about which of these options to test would be based on what is observed in the pilot study.
- **Collecting ongoing implementation data.** This approach would provide information helpful for examining factors associated with changes in quality after use of WGT, understanding how caregivers use WGT, and considering issues concerning sustainability.
- **Examining fidelity to key components of WGT implementation.** The research team would examine the frequency and quality of implementation of components that provide choice. For example, among the key components, the team would look at reports about whether the coaches and caregivers are meeting together and whether they are reflecting on videos. Examples of choice include whether the caregivers focus on a single child or multiple children or on a single time period each day, as they begin to practice a new way of interacting with children or implementing a new strategy. Based on the pilot study, the research team would develop qualitative measures of fidelity to collect during the process and implementation study to understand what is happening in the PLCs and coaching sessions.

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

III. Sampling and Recruitment for the Process and Implementation Study

A. Criteria for site selection

1. Selecting localities

In this section, we suggest criteria for selecting localities (for example, a city, county or counties, urban or rural setting, region, or network) within a state, as well as some of the benefits and trade-offs of recommending these criteria. These suggested criteria emerged from the research design team's calls with experts.

Criteria for site selection

Selecting localities

- System-level infrastructure with in-person or virtual coaching
- Alignment with a compelling motivator or incentive
- Achieving diversity

Selecting centers and FCCs

- Leadership agreement and willingness to support WGT implementation
- Access to and provision of coaches

An important criterion to consider is whether there is system-level infrastructure that provides either in-person or virtual coaching that can incorporate WGT to inform its approach and content. The research team could select a locality with an active Infant Toddler Specialist Network (ITSN) with coaches and active FCC networks with coaches. That is, these networks would be affiliated with an organization that provides coaching. However, the team should exclude localities in states that are currently implementing their own coaching system, such as Texas and Washington, so as not to interfere with their efforts. Some states do not have coaching approaches that are clearly defined. For example, Colorado, Missouri, North Carolina, and Tennessee use practice-based coaching and some form of a coaching model (such as the Program for Infant and Toddler Care, or Pyramid Model). However, these states do not mandate a specific coaching model or curriculum, so localities within the state may make different choices (Child Care State Capacity Building Center 2022). Calls with experts have suggested that such localities within these states may be prime candidates for site selection.[19] Having existing networks in place would make it more feasible for the research team to recruit and train coaches to participate in the intervention, because coaches would already be available in a particular locale. However, selecting localities with this kind of infrastructure would limit the generalizations that could be made and would not be applicable to places without such infrastructure.

In selections of localities, the research team might consider choosing localities in two states that differ on a characteristic of interest, such as one state that provides coaches with training in infant mental health (for example, West Virginia) and one state that does not. This approach would provide a larger group of coaches with easily identified differences in background to help determine whether these factors are associated with implementation and quality. However, there are already a large number of potential contrasts to decide on, so having the locality or localities within one state may be preferable to hold some factors constant.

Given the challenges and pressures on the ECE workforce, alignment with a compelling motivator or incentive (for example, hours, credits, or continuing education units [CEUs] toward a credential or

[19] Personal communication with Ronna Schaffer and Holly Wilcher, April 29, 2022.

degree; paid PD time; alignment with career progression bonuses and salaries; compensation or credential conditional on participation) would be beneficial for adoption of WGT. The availability of this motivator or incentive would make it easier for the research team to recruit caregivers and might support stronger engagement. The potential incentives offered for PD by a center or FCC network may differ in desirability to the caregivers. For example, paid PD time would likely be more attractive than alignment with a career progression. If there is natural variation in the types of incentives offered, then the research team could consider these variables in data collection and multivariate analyses.

The research team could consider ways to achieve diversity if the locality crosses more than one state (for example, Kansas City in both Missouri and Kansas). For example, the research team could achieve diversity by examining state policies on caregiver-child ratios, state requirements or standards for coaches (possibly with lists of coaches who meet requirements), and states that require infant and toddler specialists to obtain training in infant mental health. If the locality is within a single state, the research team could achieve diversity by recruiting both rural and urban settings as well as considering local racial and ethnic composition, children's home languages, and household income. Some of this information would come from the PIR and some from public sources containing state policies and systems. Demographic information would come from the American Community Survey or the ITSN or FCC network.

2. Selecting centers and FCCs

After selecting the localities, the research team would then select centers and FCCs within the selected network or networks. It would be important for network leaders to endorse and support the study. Another important consideration for selecting centers and FCCs is whether the research team can obtain agreement from center and FCC leadership to conduct the study and their willingness to support the implementation of WGT. Past research has highlighted the crucial role that leaders play in promoting buy-in from teachers (Kerr et al. 2006; Spillane 1999; Turnbull 2002). Leadership buy-in is one of the key elements that can shape a culture of agreement and innovation and help strategically plan for the effective implementation of new PD efforts (Park and Datnow 2009; Yoon et al. 2016; Zimmerman 2006). It may be easier for caregivers to embrace new changes and approaches to PD when they see that their center leaders are supportive of these efforts.

Another consideration for selecting centers and FCCs is the availability of coaches. By having access to coaches who already work with caregivers in centers and FCCs, the research team would not need to recruit coaches externally. Caregivers may be familiar with the coaches if the coaches have already been working with them or at their centers. Caregivers may also be more likely to reach out to coaches if they have had a previous working relationship with them. Given that the proposed study may also include some element of virtual coaching and PD, having access to coaches who are familiar to caregivers and the centers and FCCs in general may be particularly important. The coaching relationship is noted as an integral component of some coaching frameworks (Chu 2014; Snyder et al. 2012).

Design option: Select centers and FCCs that serve children from families eligible for subsidies or EHS

With this option, the research team would recruit both community-based centers and FCCs that accept subsidies as well as EHS centers and EHS FCCs. Alternatively, the team could recruit in areas where most families would be eligible for subsidies, even if a given center does not accept subsidy payments.

B. Sampling

1. Sampling

For each locality, the recruiting process would require using various sources, such as the ITSN lists of centers, the PIR, the FCC networks, and lists of licensed center- and family-based child care providers. As an additional resource for sampling, the research team could also use the local child care resource and referral system, state quality rating and improvement system databases, and FCC provider networks. The research team would purposively sample localities to ensure that they represent a range of centers and FCCs serving infants and toddlers. In addition, if a certain locality (such as a county or city) does not have enough settings that are willing to participate, the team may need to extend to an adjoining area (that is still served by the FCC network or ITSN).

2. Sample size considerations

The proposed research questions are aimed at learning what works to improve caregiver-child interactions in infant and toddler child care settings and under what circumstances. Results from the initial WGT field test indicated that not everyone benefited to the same extent. Caregivers who invested more in exploring the WGT website made greater gains than those who did not. In addition, those with weaker observed initial quality on average made greater gains than those who had higher initial quality.

For the proposed study, the research design team expects that there would be stronger effects for caregivers with strong implementation of the WGT process of active learning and a coach who invests time in supporting each caregiver's growth in providing quality care.[20] Caregivers who are less engaged or have more limited communication with coaches would be less likely to show effects. What active learning looks like and what is needed to support that level of implementation would likely differ across settings. The sampling strategy would be focused on generating a sample that is more likely to implement with fidelity so that the research team can understand for whom and under what circumstances caregivers using WGT improve their quality of care.

The research design team proposes a sample size of 60 FCCs (including both EHS and community-based), 60 community-based centers, and 40 EHS centers (see Exhibit III.1). The team assumes that there would be one PLC per coach. If there are multiple PLCs for large centers, then the team would focus on one PLC for the QCIT observation and the caregiver survey. The EHS sample is slightly smaller, given that there is more consistency in programs across the EHS centers than is found in the community settings. FCCs are more diverse than either of the other types of settings (for example, children of different ages in the settings or different environments), so the research team would also draw on qualitative data from both the pilot study and the process and implementation study to understand what is needed to support change in the care of infants and toddlers in these settings.

[20] In the initial WGT field test, caregivers reported how much they felt their coach supported their efficacy. Reports of higher support were associated with more gains in the Support for Cognitive Development domain on the QCIT.

Exhibit III.1. Overview of sampling

	Year 1	Year 2
	Pilot (Four months)	Process and implementation study (One program year)
Early Head Start	4 centers 8 classrooms	40 centers[a] 120 classrooms[b,c] (3 classrooms per center, on average) 240 caregivers (2 caregivers per classroom) 40 coaches (6 caregivers per coach, on average)
Community-based centers	4 centers 8 classrooms	60 centers[a] 180 classrooms[b] (3 classrooms per center, on average) 360 caregivers (2 caregivers per classroom) 60 coaches (6 caregivers per coach, on average)
Family child care	4 FCCs	60 FCCs[d] (1 classroom per FCC) 60 caregivers 10–15 coaches

[a] Centers would be sorted by size (small, medium, large) to ensure variation in size.

[b] "Classrooms" represent both FCC settings and center-based classrooms.

[c] All caregivers in classrooms would receive WGT and be asked to respond to the data collection instruments. However, only two classrooms per center would be observed on the QCIT.

[d] All participating FCCs would be observed on the QCIT.

C. Recruitment

1. Recruitment process

The research team would work with the selected locality to recruit centers to participate in the study. The team would request lists of certified center- and family-based child care providers from the network or networks. If there are not sufficient settings available, the team could obtain these lists using the PIR, local child care resource and referral systems, or state quality rating and improvement system databases. To recruit centers, the research team might send recruitment materials directly to centers, based on the lists that the locality or network shares, or share materials with the locality to send to centers on behalf of the research team.

To recruit sites, the research team should send a clear message via email outreach and telephone follow-up to centers and FCCs about the advantages of participation. For example, the research team could highlight that the system (1) includes research-based content; (2) is practice-based; (3) can be used across different setting types (community-based, EHS, FCCs); and (4) is

Considerations for equity in recruitment and sampling

The research team would make a concerted effort to center equity throughout the study, including during recruitment and sampling. The team would plan to recruit caregivers and coaches representative of the communities they serve. For example, the team would consider demographics (e.g., gender, age, race/ethnicity, ability status) as well as geography (e.g., urbanicity, state, region). The team could also use a "snowball" technique and ask participants with lived experience to recommend additional individuals to recruit. The team could reach out to state and network partners to connect them with individuals with lived experience that they could not have otherwise recruited.

aligned with national frameworks such as ZTT and the CDA®. The research team could also inform potential participants that WGT takes a flexible approach with online and relationship-based aspects and provides several implementation supports, including training, materials, and technical assistance (potentially through a help desk). The team could update recruitment videos that were previously developed for the field test.

2. Administrative support

WGT is more likely to be successfully implemented when there is endorsement from management (such as center directors and education coordinators) as well as organizational support. Caregivers may be more willing to participate in the study if their leaders think WGT is a good choice for their PD. The research team could ask a senior administrative leader of the setting as well as a leader who has direct contact with caregivers (such as an education manager) to sign a letter of commitment before participating in the process and implementation study. This could be one way to create investment in the local leadership to build up a level of commitment in the centers and FCCs. To increase the commitment beyond written endorsement, the research team would also involve local leadership in the implementation of WGT by including them in the initial training session or other introductory meetings, presenting on WGT to the centers and FCCs, and discussing with local leadership about their roles in the study and key supports needed for successful implementation.

3. Potential challenges due to staffing shortages and turnover

Given the large-scale staffing shortages that programs are currently facing, there would likely be high attrition in the study (for example, because of high caregiver turnover). WGT would be implemented with all infant and toddler caregivers within a center. One benefit to implementation at the center level is that even if a caregiver did leave in the middle of the year, caregivers within the same classroom would have been working on the same practices, allowing for an estimate of fall to spring change, although not as precise. Fall surveys would be administered to every infant and toddler caregiver in a center. By doing this, the research team could replace a teacher or caregiver who may leave in the middle of the year, depending on timing. The research team would collect a small set of demographic questions for new caregivers who cycle into a classroom, to be collected in the spring feedback survey.

The research design team is especially concerned about FCC attrition, because this would be a greater threat to the study than attrition of a caregiver in a center. If an FCC provider leaves the study, the research team could attempt to interview the provider about the decision to leave and how WGT was helpful or challenging. To guard against loss of power in the analysis due to attrition, the research team should recruit additional FCCs within the network if feasible. As with the WGT field test and prior PD studies, caregivers would be considered a participant if they have been in the study a minimum of 12 weeks (the full process and implementation study would last one year).

Similar challenges may arise in the four-month pilot, at which point the research team would work with the centers and FCCs to solve any problems. The team should leverage lessons from the pilot study in the process and implementation study—for example, by sharing strategies used by sites during the pilot study.

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

IV. Data Collection in the Process and Implementation Study

This chapter describes the data collection instruments for the process and implementation study. Exhibit IV.1 shows the constructs (including covariates) for the process and implementation study. Some constructs were measured in the prior WGT field test and would be used again; others are new or measured in a different way. The new measures would capture information to examine questions that were not answered in the field test. Exhibit IV.2 presents instruments that the research team could use to collect information on each construct, the target sample, the timing of data collection, and the proposed measures for the process and implementation study. Exhibit IV.2 also indicates the research questions that the construct measures may inform.

A. Classroom roster

For classrooms receiving a QCIT observation, field staff would ask the lead caregivers to complete a brief paper-and-pencil classroom roster in the fall and spring for the children present on the day of the observation. The roster information would be used to confirm whether the classroom qualifies as an infant or toddler classroom. The research team should also collect basic demographic information about all children in the observed infant and toddler classrooms, including their ages, race and ethnicity, languages spoken, and how many children in the classroom have IFSPs, from the classroom rosters. Because the focus is on process and implementation, the team would collect these data to understand how caregivers implement WGT with the populations of children in their care. In addition, caregivers are more likely to sustain a practice if they observe positive change in the children in their setting, so it would be important to understand the characteristics of the children they take care of. This information would be used to understand the populations served by the ECE facilities in the study and allow the research team to control for some potentially unobserved factors.

B. Classroom checklist

The research team should also collect information about the classroom size and ratio, as well as materials and resources. These resources may include books, toys, and sensory materials available to the infants and toddlers. The team could capture this information using a checklist that is completed before or after the QCIT observation for selected classrooms (two per setting) that are observed using the QCIT.

Exhibit IV.1. Constructs from the field test and proposed new constructs for the process and implementation study

Construct	Definition	Collected in the field test	Different measure or approach/mode to collecting construct in the proposed study[a]	New construct, to be collected in the proposed study	Rationale, if different measure or approach/mode to data collection, or new construct
Outcomes					
Quality of caregiver-child interactions	Measure of how well the interactions and relationships between caregivers and children support child development	X			NA
Caregiver self-efficacy	Caregivers' belief in their capacity or ability to provide care for children	X[b]	X		Collect information on caregivers' self-efficacy using different measures from the field test. These measures would include additional items that were not asked in the field test.
Caregiver knowledge and beliefs	What caregivers know and believe about caregiving, child development, and learning in ECE	X[b]			NA
Level of effort	Time that the implementation team, local organizations, setting leadership, coaches, and caregivers spend preparing and implementing WGT as well as any training on WGT	X			Collect reliable information on level of effort to understand the true costs of what it would take to launch and implement WGT.
Covariates					
Classroom characteristics	Classroom size, ratio, and other structural characteristics	X[b]			NA
Classroom materials and resources	Materials and resources available in the classroom such as books, toys, and sensory materials available to the infants and toddlers			X	The initial field test was not an implementation study and therefore did not collect such detailed information to understand the environment in which caregiver-child interactions are taking place.
Caregiver characteristics	Caregiver demographics, education and training background, experience, level of comfort with technology, and past and current PD experiences and/or coaching experiences	X[b]			NA
Coach characteristics	Coach demographics, education and training background, work experience,	X[b]			NA

Exhibit IV.1 (*continued*)

Construct	Definition	Collected in the field test	Different measure or approach/mode to collecting construct in the proposed study[a]	New construct, to be collected in the proposed study	Rationale, if different measure or approach/mode to data collection, or new construct
	demographics, education and training background, experience, level of comfort with technology, past and current coaching experiences				
Organizational climate	Caregivers' perceptions about the environment and culture of their center or FCC			X	Collect information on characteristics such as caregiver burnout, collegiality, relational dynamics, center organizational culture, psychological safety, initial level of trust, and openness or resistance of new ideas and change may help assess the likelihood of caregivers' engagement with WGT in this implementation study.
Readiness to change	Extent to which caregivers are inclined to accept or adopt new caregiving practices	X[b]	X		This construct was measured in the field test but was not predictive. Use a different measure instead to understand readiness to change in this study.
Instructional leadership self-efficacy	Setting leaders' belief in their capacity or ability to lead their settings and caregivers who work with them			X	The initial field test was not an implementation study and therefore did not collect such detailed information from setting leadership.
Coach fidelity of implementation of WGT	Degree to which coaches carried out coaching components of WGT as intended (e.g., sending nudges, conducting meetings)			X	The initial field test was not an implementation study and therefore did not collect such detailed information.
Caregiver Application of WGT practices	Evidence of caregivers using and applying WGT practices with children in their care	X[b]	X		The initial field test was not an implementation study and therefore did not collect such detailed information on caregivers' application of WGT practices. Survey items and text messages would be more detailed for the proposed study.
Caregiver engagement in WGT and PLC (dosage)	Caregiver engagement or involvement with aspects of WGT, including interfacing with the web-based system, WGT	X[b]	X		The initial field test was not an implementation study and therefore did not collect such detailed

Exhibit IV.1 (*continued*)

Construct	Definition	Collected in the field test	Different measure or approach/mode to collecting construct in the proposed study[a]	New construct, to be collected in the proposed study	Rationale, if different measure or approach/mode to data collection, or new construct
	materials, discussions with others in the PLC coaching sessions (e.g., frequency/time on task within the PLCs and coach)				information on caregivers' engagement of WGT aside from the web analytics data. In addition to collecting web analytics data, collect engagement data from text messages.
Coach engagement in WGT	Coach engagement or involvement with aspects of WGT, including preparing WGT materials for the PLCs and preparing and leading discussions with caregivers in the PLC coaching sessions			X	The initial field test did not collect information on coach engagement.
Motivation/burnout	Caregivers' intrinsic values and desire to care for and teach children, and whether they experience emotional job stress			X	The initial field test was not an implementation study and therefore did not collect such detailed information. Information about the degree of burnout caregivers face may help assess the likelihood of their engagement with and implementation of WGT.
Characteristics of PLCs	Frequency of the PLCs, type/mode of interaction, format (subgroups or pairs interacting in the PLCs), how often were videos used, how often providers shared their experiences, size, and attendance			X	The initial field test of WGT was not a team-based format and therefore did not collect information on the characteristics of the PLCs.
Characteristics of coaching sessions	Frequency of video discussion, meeting attendance, mode of delivery (whether virtual or in person), and topics discussed in coaching sessions	X[b]	X		Collect information on characteristics of coaching sessions using a different approach/mode. In the field test, there were low response rates with the pop-up survey. In addition, the pop-up survey only asked caregivers information about the number and length of meetings with their coach. Use a different approach/mode to collect more detailed information using checklists and logs filled out by coaches.

Exhibit IV.1 (*continued*)

Construct	Definition	Collected in the field test	Different measure or approach/mode to collecting construct in the proposed study[a]	New construct, to be collected in the proposed study	Rationale, if different measure or approach/mode to data collection, or new construct
Caregiver and coach feedback on using WGT	Caregivers' and coaches' experiences with implementing WGT, engaging in PD, and supports and barriers to implementing WGT	X[b]	X		Collect information on caregivers' and coaches' experiences with WGT, taking a mixed-methods approach given that the initial field test was not an implementation study and therefore did not collect such detailed information.
Interactions with peers outside of the PLC and how caregivers felt their peers supported their effectiveness	Extent and ways in which caregivers worked together and interacted with their peers (for example, mutual encouragement, sharing information) and felt their peers supported their ability to provide quality care to children			X	The initial field test of WGT was not a team-based format and therefore did not collect information on interactions with peers and on how caregivers felt their peers supported their effectiveness.

[a] The proposed approach is described in more detail in Exhibit IV.2 and the text that follows.

[b] Constructs would be collected in the pilot and process and implementation studies even if they did not change since the field test.

NA = not applicable.

Exhibit IV.2. Constructs, instruments, target sample, timing of data collection, and example measures linked to research questions for the process and implementation study

Construct	Data collection instrument	Target sample[a]	Timing[b]	Example measure(s)	Research question
Outcomes					
Quality of caregiver-child interactions	Observational assessment—QCIT	Two classrooms per center, on average, and all FCCs	Fall and spring of the process and implementation study year	QCIT (Atkins-Burnett et al. 2020; Nguyen et al. 2022)	1, 4, 5
Caregiver self-efficacy	Fall and spring surveys, web- or paper-based	All caregivers in participating classrooms	Fall and spring of the process and implementation study year[a]	Items adapted from the Teacher Self-Efficacy Scale (TSES; Bandura 1997); Teachers' Sense of Efficacy Scale (Tschannen-Moran and Woolfolk Hoy 2001)	2 (and covariate in 3 and 5)
Caregiver knowledge and beliefs	Fall and spring surveys, web- or paper-based	All caregivers in participating classrooms	Fall and spring of the process and implementation study year[a]	Early Head Start Family and Child Experiences Study (Baby FACES) Beliefs About Development Scale (Atkins-Burnett et al. 2017)	2 (and covariate in 3 and 5)
Level of effort	Brief, scale-based text message responses; monthly check-ins between administrative staff and the research team; monthly coaching logs	Implementation team, local organizations, setting leadership, coaches, and caregivers in participating classrooms	Fall and spring of the process and implementation study year[a]	Items developed by research design team	6–8
Covariates					
Classroom characteristics	Classroom roster	All classrooms observed on the QCIT	Fall and spring of the process and implementation study year	WGT Field Test Classroom Roster; items developed by research design team	1, 2, 3, 5, 7, 8
Classroom materials and resources	Checklist	Two classrooms per center and all FCCs observed with the QCIT	Fall and spring of the process and implementation study year	2018 Baby FACES Environmental Checklist (Cannon et al. 2019); items developed by research team	1, 2, 3, 5, 7
Caregiver characteristics[c]	Fall survey, web- or paper-based	All caregivers in participating classrooms	Fall of the process and implementation study year[a]	Items developed by research design team	1, 2, 3, 5, 6
Coach characteristics[d]	Fall survey, web- or paper-based	All coaches in participating classrooms	Fall of the process and implementation study year[a]	Items adapted from the Study of Coaching Practices in Early Care and Education Settings (SCOPE) Coach Survey, 2016–2019 (SCOPE Project Team 2019); items developed by research design team	1, 2, 4–6

Exhibit IV.2 (*continued*)

Construct	Data collection instrument	Target sample[a]	Timing[b]	Example measure(s)	Research question
Organizational climate[e]	Fall survey, web- or paper-based	All caregivers in participating classrooms	Fall of the process and implementation study year[a]	Items adapted from the Essentials 0–5 Survey (Ehrlich et al. 2019); NSECE Teacher Survey 2019; Instructional Activities Scale and Teacher Beliefs Scale (Hart et al. 1990); The Early Childhood Work Environment Survey, 3rd edition. (ECWES; Bloom 2016); Bay Area School Reform Collaborative (BASRC; Porter et al. 2006)	3, 5
Readiness to change	Fall survey, web- or paper-based	All caregivers in participating classrooms	Fall of the process and implementation study year[a]	Items adapted from the Assessment of a Preschool's Readiness for Change (Wanless 2015); University of Rhode Island Change Assessment Scale (URICA; McConnaughy et al. 1983)	3, 5
Instructional leadership self-efficacy	Fall survey, web- or paper-based	All leaders of participating centers	Fall of the process and implementation study year[a]	Items adapted from the Essentials 0–5 Survey (Ehrlich et al. 2019); The Administrator Role Perception Survey (ARPS; McCormick Center for Early Childhood Leadership 2019); The Preschool Instructional Leadership Survey (PILS; Fong and Horsley 2017)	3, 5
Coaching fidelity	Brief, scale-based text message responses; spring surveys, web- or paper-based	All coaches in participating classrooms	During the process and implementation study year[a]	Items developed by research team	3–6, 8
	Fidelity checklist	All coaches in participating classrooms	During the process and implementation study year[a]	Items developed by research team	
Caregiver Application of WGT practices	Brief, scale-based text message responses	All caregivers in participating classrooms	During the process and implementation study year[a]	Items developed by research team	4–6
Caregiver engagement in WGT and PLC (dosage)	Brief, scale-based text message responses; web analytics from the WGT site	All caregivers in participating classrooms	During the process and implementation study year[a]	Items developed by research team	3–6, 8
Coach engagement in WGT	Caregiver report; web analytics from the WGT site	All coaches and caregivers in participating classrooms	During the process and implementation study year[a]	Items developed by research team	3–6, 8

Exhibit IV.2 (*continued*)

Construct	Data collection instrument	Target sample[a]	Timing[b]	Example measure(s)	Research question
Motivation for caregiving	Fall survey, web- or paper-based	All caregivers in participating classrooms	Fall of the process and implementation study year[a]	Items adapted from the NSECE Teacher Survey 2019 (NSECE Project Team 2019)	3, 5
Characteristics of PLCs	Checklists and logs filled out by coaches	All participating coaches	During the process and implementation study year[a]	Items adapted from the Formative Assessment of Collaborative Teams (FACT; Taylor et al. 2014); items developed by research team	4–6
Characteristics of coaching sessions	Checklists and logs filled out by coaches	All participating coaches	During the process and implementation study year[a]	Items adapted from the Coaching Practices Observation Tool (CPOT; Shannon et al. 2021); items developed by research team	3–6
Caregiver and coach experience with WGT	Spring surveys, web- or paper-based	All coaches and caregivers in participating classrooms	Spring of the process and implementation study year[a]	Items adapted from theTeachers' Attitudes about Professional Development (TAP; Torff et al. 2005); items developed by research team	1, 2, 5, 6, 8
	Interviews and focus groups	Subset of participating administrators, coaches, and/or caregivers	Spring of the process and implementation study year[a]	Questions developed by research team	6, 8
Interactions with peers outside of the PLC and how caregivers felt their peers supported their effectiveness	Brief, scale-based text message responses; spring surveys, web- or paper-based	All caregivers in participating classrooms	During the process and implementation study year[a]	Items developed by research team	4, 5, 6

[a] When describing the target sample, "classroom" is used for both FCCs and center-based classrooms. When only a subset of center-based classrooms is involved, this also refers to all FCCs.

[b] The construct and a briefer version of the data collection instrument would be tested in the pilot study.

[c] Caregiver characteristics include demographics, education and training background, experience, level of comfort with technology, and past and current PD experiences and/or coaching experiences.

[d] Coach characteristics include demographics, education and training background, experience, level of comfort with technology, past and current coaching experiences. Additional information may include coaches' knowledge, communication style, relationship skills, and mental health.

[e] Subconstructs of organizational climate would include caregiver burnout, collegiality with one another and/or relational dynamics, center organizational culture and/or climate, psychological safety, initial level of trust.

C. QCIT

Field observers would be trained and certified on the QCIT. Observations conducted by field observers rather than coaches would ensure that the data collection is standardized and independent of the implementation. Teams of certified QCIT observers would be sent to the selected localities over a specific period of time (for example, eight weeks). Details about the size of the QCIT observation team and the duration of data collection would be determined once the localities are selected. To contain costs, the research team would select two infant and/or toddler classrooms per center to conduct the QCIT observations in fall and spring. When centers and FCCs serve both infants and toddlers, the research team would observe one classroom serving primarily younger children (infants) and a second classroom primarily serving older children (toddlers). In the initial WGT field test, infant classrooms were defined as having 50 percent or more of the children under the age of 18 months, and toddler classrooms had more than 51 percent of children who were toddlers. In FCCs, the team would be observing the caregiver's interactions with children within the infant-toddler age range. Before conducting random selection of the two classrooms, the team would need to determine whether or not a caregiver would be able and willing to participate. All participating FCCs would be observed. At least one caregiver in the classroom (preferably a lead caregiver) would need to be comfortable with both reading materials and watching videos in English. The surveys and most of the WGT materials are only available in English, with the exception of handouts for families. Classrooms with at least one caregiver who speaks English would receive an observation, but only the caregivers who are comfortable with English would complete the surveys.

The research team would conduct the QCIT in the fall and spring to ensure that there is a common baseline of observed quality before WGT implementation. Conducting observations once in the fall and once in the spring would allow the research team to make comparisons of the change in quality rather than just reporting on quality at a single point in time. W-scores are continuous measures that account for the differences in the difficulty of the practices. Examining change with W-scores reduces problems with between-subject variability associated with initial raw QCIT scores. Each regular caregiver in a classroom (excluding floaters and temporary substitutes) would be observed using the QCIT. The research team would alternate among observing the caregivers in the classroom. The team would observe a total of six cycles to produce classroom-level scores.

Data collection options for the fall and spring surveys

Because the cost of conducting surveys for every infant and toddler caregiver in the fall and spring may be prohibitive, other options for collecting survey data include the following:

- Collect (1) caregiver surveys from the two selected classrooms in the fall, (2) a spring-only survey of all infant and toddler caregivers, and (3) coach and center director surveys in the fall and spring
- Collect (1) caregiver surveys from the two selected classrooms in the fall and spring and (2) coach and center director surveys in the fall and spring

D. Fall and spring surveys

The research team could ask all caregivers, coaches, and center leaders to report information about their background characteristics, knowledge, and beliefs via a background web survey in the fall. Participants

would receive an email with a link to this web survey to complete online or on their phones before the fall QCIT observation and would receive email reminders if they have not completed the survey.

The research team would collect information from all caregivers in each classroom using a fall background survey, which would help determine caregiver characteristics that may relate to WGT uptake and implementation, help with replacement in case of caregiver turnover, and provide broader information on the PLCs and their participation. The research team would use the background survey to gather information about demographics, education and training backgrounds, professional experience, level of comfort with technology, and past and current PD experiences and/or coaching experiences. In addition, the research team would collect information in both fall and spring regarding caregiver beliefs, knowledge, and self-efficacy. The research team could also measure factors about caregivers that might influence their motivation to learn and improve their practice using WGT. For example, the team could collect information on characteristics such as caregiver burnout, collegiality with one another and/or relational dynamics, center organizational culture and/or climate, mental health, psychological safety,

initial level of trust, and openness or readiness for change. In the spring, all caregivers participating in WGT would also provide feedback about the implementation and perceived benefits of WGT resources and activities, including the PLC meetings.

Design option: Center director survey

Another design option is to include a director survey in the fall and/or spring on self-efficacy, specifically focused on instructional leadership, and another survey in the spring on their observations about the benefits and challenges of WGT PLCs for their infant and toddler classrooms.

To better understand characteristics of coaches that may relate to WGT uptake and implementation, the team could collect information using a fall background survey from all participating coaches. The background survey would be used to gather information about their demographics, education and training backgrounds, professional experience, level of comfort with technology, and past and current coaching experiences. Additional information on coaches' knowledge, communication style, relationship skills, and mental health could also be collected. This information could help explain the relationship between the use of WGT and caregiver changes in interaction quality.

Upon completion of the implementation period, the team would ask all caregivers and coaches to report on their experiences with the WGT via a web survey emailed to them before their spring QCIT observation. The research team would also ask caregivers and coaches about their perceptions of what was most beneficial in supporting changes in practice and in children's development. The team would use email reminders and text messages to encourage participation. On the day of the observation in both the fall and spring, if caregivers have not completed the web survey, the field observer would encourage them to do so (and have paper copies available). The spring survey would also have a section on caregiver characteristics that were asked in the fall survey for any caregivers who stepped into the study as a replacement.

E. Text messages

The team could also ask caregivers and coaches about their WGT experiences in real time, using a limited number of questions in weekly or biweekly text messages. For example, caregivers might report whether they video recorded themselves practicing a given strategy with a child or children, discussed a video with a coach or peers, or whether they collected observational data on children's reactions to the practice

or strategy. The team could circulate up to three questions once per week for caregivers to report on various topics, such as (1) the frequency with which they used checklists and practice summaries and informally shared information with peers, (2) how they felt children responded to the practices, (3) whether they noticed changes in children, (4) whether there was mutual encouragement and motivation in the PLCs and among peers, (5) the frequency with which they received constructive feedback from their coach on the use of the practices, and (6) the time they spend discussing and working with their coach or peers on WGT implementation. Coaches would be asked to rate the overall engagement level in PLC meetings and report the number of additional contacts that they had with caregivers about the module content and implementing the practices in the classroom or FCC. Caregivers and coaches can answer these text message questions quickly and briefly on a 1- to 5-point scale, for example. The number of text message questions would be small (for example, up to three questions once per week).

F. Web analytics

The team could also collect tracking data from the WGT website to capture the amount of time that caregivers spend on it. In addition to measure engagement with the WGT content, web analytics could include information such as the number of pages that caregivers viewed on the website, the number of presentation views, modules that the caregiver accesses, and how often the caregiver works in the module. This information would be gathered periodically throughout the study and at the end of the process and implementation study. The research team could consider the benefits of allowing coaches to view the website log-in information so that they could nudge or try to support caregivers who are not engaging or who seem to be struggling.

G. Coaching checklists and logs

Because PLCs would be introduced for the first time with WGT in this study, the research team could ask the coaches who are leading PLCs to log the frequency of video discussion, mode of delivery, and meeting attendance of each PLC. This log would be quick and electronic, such as a Google form, or a dashboard, or a "quiz" activity in the WGT site. It might be possible to configure these data into graphs for coaches. PLC engagement would be gathered by having the coaches fill out a checklist about topics discussed during each session, which would be used for capturing fidelity. As mentioned earlier, this checklist would be developed in the pilot study to capture the topics discussed in PLCs related to WGT implementation. This approach would allow the research team to document and determine the success of WGT training and to better understand how coaches used the WGT coaching strategies. The research team could also collect action plans and coaching logs and records. The coaching logs filled out at the end of every PLC would provide information about the mode of coaching (that is, whether the caregiver and coach met virtually or in person). In addition, children's responses to practices as well as caregivers' use of visual cues and nudges could be captured in the checklists that the coaches complete during the PLC discussions and video reviews. Coaches would also report on the time they spent leading the PLC as well as the time spent preparing for and reflecting preparing and reflecting on the PLCs so that the research team can understand the level of effort to implement WGT. The logs and checklists would be collected from the coaches at the end of each month.

H. Administrative check-ins

The research team would gather information about the level of effort by collecting information on the time that participating center-based administrators spend as well as another associated cost to implement WGT within centers. During these periodic check-ins with center-based administrators, the research team

would ask about administrative costs, such as hiring extra staff for maintaining classroom ratios while caregivers are in their PLCs, how much were caregivers reimbursed for their time spent on PD, and how much time administrators spent discussing WGT with caregivers. We estimate three to four calls per center-based administrator during this study.

I. Interviews and focus groups

The research team would conduct qualitative interviews with a subset of participating administrators, coaches, and caregivers from the pilot study and in the process and implementation study. The purpose would be to gain insights from study participants on their experiences with implementing WGT, engaging in PD, as well as how they have used WGT outside of the study (for example, how they shared or demonstrated WGT practices with other caregivers not in the study). The interviews would be facilitated and led by a member of the research team using a semi-structured protocol that would be adapted to the roles of the participants and their setting. The research team would ask about how the ITSN, FCC network, state, or region supports implementation of WGT as well as barriers to implementation. The team would interview participants at the site and system level. In addition to interviews, the research team could also conduct focus groups to gather feedback and opinions from caregivers on WGT and its implementation. Interview questions would be tailored accordingly to acknowledge the diversity across and within communities. The qualitative interviews would allow the research team to collect rich data, ensuring that the people involved with, and affected by, implementing WGT —the administrators, coaches, and caregivers—have an opportunity to share their experiences.

V. Analysis Plan for the Process and Implementation Study

This chapter describes the proposed plan for preparing the data for analysis and reporting as well as the analytic approaches for answering the research questions.

A. Preparation of data for analysis

1. Quantitative data

The research team would conduct several preliminary activities to prepare the data for analysis. For each of the data collection instruments (background and feedback surveys for caregivers and coaches and responses to text messages), the team would assess the level of nonresponse overall as well as to specific items. The team would also examine the quality of the data collected to look for outliers, unexpected responses, or inconsistencies. The team would obtain some of the covariates from the recruiting information or the fall observation (for example, whether the classroom serves primarily infants or toddlers) and web analytics data.

The team would construct summary variables and scales and use the appropriate techniques to assess the quality and psychometric properties of the constructed variables. These constructed variables would typically involve combining information from multiple items within a single data collection instrument. The research questions and recommended constructs for the surveys (outlined in the data collection chapter) would guide the set of variables and scales developed. For example, the team would construct scales for organizational climate, caregiving beliefs, and knowledge and beliefs about child development. The team would assess the reliability of those measures using coefficient alpha (Cronbach 1951). The team would examine the correlations among items and/or scales and create composite scores when data reduction is needed (if correlations are high, $r > 0.70$). Interval and continuous variables would be z-scored and used for outcomes in statistical models. This approach would allow for the interpretation of the coefficients in the models as the change in the outcome, in standard deviation units for each 1-point increase in the respective variable. The z-scores on the QCIT would be derived from W-scores (Woodcock 1999), a Rasch-based score. The W-score has several advantages, including an interval-level scale, and is recommended for examining change over time. W-scores also support criterion-referenced interpretation.

2. Qualitative data

The research team would prepare and organize the qualitative data for thematic analysis. The team would gather all of the audio-recorded interview files into one location and, if necessary, convert observational notes to electronic format and scan documents retrieved in paper form. Interview files would be transcribed to support the team in completing the qualitative data analysis. The team would format the data for importing into a qualitative data analysis software package. Importantly, the research team would ensure the anonymity of the data, thus ensuring participant confidentiality. The team would apply best practices, such as replacing the names of individuals and centers in text with pseudonyms as well as obscuring identifying details such as dates, ages, and locations. This would help ensure that the data cannot be linked back to the people, communities, or organizations described within the data. Transcriptions and any notes and documents would be coded by two coders. A codebook would be created to structure how the interviews and focus groups would be coded in the software program. Inter-rater reliability on codes would be estimated.

B. Data analysis

Exhibit V.1 presents the research questions and whether the research team would be conducting quantitative or qualitative data analysis, or both. Exhibit V.2 describes the analysis methods that the research team would use to answer each research question in the process and implementation study. The sections below describe a general approach to analyzing the quantitative and qualitative data.

Exhibit V.1. Research questions to be answered with qualitative data, quantitative data, or both

Research question	Quantitative data	Qualitative data
Natural variation in conditions that may influence the use of WGT		
1. How does the classroom-level quality of caregiver-child interactions change between fall and spring? a. How do the changes in quality from fall to spring of caregiver-child interactions vary based on initial levels of quality (high, middle, low), experience levels, setting types, and age ranges of children in their group (infants versus toddlers)? b. Does the association between level of quality of caregiver-child interactions in the fall and spring vary by coaching experience or size of the PLC?	X	
2. How do caregivers' beliefs and knowledge about ECE and caregiver self-efficacy change between fall and spring? a. How do the changes in caregiver beliefs and knowledge about ECE and caregiver self-efficacy from fall to spring vary based on experience levels, setting types, age ranges of children in their group (infants versus toddlers), average PLC beliefs and knowledge, and average PLC self-efficacy?	X	
Implementation		
3. What factors are associated with strong implementation of WGT—for example, setting type (EHS, community-based centers, FCCs); size of FCC or classroom and ages of children; level of support from center administration (where applicable); center or PLC climate; physical space; books, toys, and other sensory materials available for use by infants and toddlers; and caregiver beliefs about how to provide care and education to infants and toddlers?	X	X
4. What components of implementation are most strongly associated with positive changes in practice—for example, caregivers' level of engagement in the PLC, frequency of PLC meetings, frequency of attendance, level of use of WGT website, frequency of collecting videos of practices and reflecting on their use, and the extent to which participation by caregivers and coaches was sustained throughout the study year?	X	X
5. Do the associations between WGT components and positive change in classroom quality differ for classrooms with initial high and low levels of quality or with the PLC's focus?	X	
Resources		
6. What is the level of effort needed to implement WGT at the setting and network levels for center directors, caregivers, coaches, and the study team—for example, time for coaching, for dedicated PD and instructional planning, for	X	X

Research question	Quantitative data	Qualitative data
substitutes to cover for caregivers when they are meeting with coaches, for the study team to provide technical assistance?		
7. What are the costs associated with materials needed to implement WGT—for example, technology costs such as devices, licensing fees, and internet access?	X	X
8. Does the level of effort needed to implement WGT vary by setting type and, if so, in what ways?	X	X

Exhibit V.2. Description of quantitative and qualitative data analyses for the research questions

Research question	Quantitative		Qualitative	Description of analysis
	Outcome of interest	Example(s) of predictor	Example constructs	
Natural variation in conditions that may influence the use of WGT				
1. How does the classroom-level quality of caregiver-child interactions change between fall and spring? a. How do the changes in quality from fall to spring of caregiver-child interactions vary based on initial levels of quality (high, middle, low), experience levels, setting types, and age ranges of children in their group (infants versus toddlers)? b. Does the association between level of quality of caregiver-child interactions in the fall and spring vary by coaching experience or size of the PLC?	Differences between fall and spring QCIT scores	Initial levels of quality, experience levels, setting types, age ranges of children in their group (infants versus toddlers), coaching experience, PLC size	Conditions that make it easier to implement WGT	• Descriptive analyses of fall-spring change scores (means and standard deviations, ranges) • Bayesian methods • Tests of significance of group differences in change scores based on initial high, middle, and low fall QCIT scores • Multivariate (2L) HLM regression analyses, with spring scores as dependent variables, and experience levels; setting types; age differences as predictors • Moderation analyses
2. How do caregivers' beliefs and knowledge about ECE and caregiver self-efficacy change between fall and spring? a. How do the changes in caregiver beliefs and knowledge about ECE and caregiver self-efficacy from fall to spring vary based on experience levels, setting types, age ranges of	Differences in fall and spring beliefs and knowledge about early care and education and self-efficacy	Initial levels of beliefs and knowledge, experience levels, setting types, age ranges of children in their group (infants versus toddlers), average PLC beliefs and knowledge about ECE, caregiver self-efficacy	NA	• Descriptive analyses of fall-spring change scores (means and standard deviations, ranges) • Bayesian methods • Multivariate (3L) HLM regression analyses, with spring scores as dependent variables, and experience levels; setting types; age differences; average PLC beliefs and knowledge; and

Research question	Quantitative		Qualitative	Description of analysis
	Outcome of interest	Example(s) of predictor	Example constructs	
children in their group (infants versus toddlers), average PLC beliefs and knowledge, and average PLC self-efficacy?				average PLC self-efficacy as predictors
Implementation				
3. What factors are associated with strong implementation of WGT—for example, setting type (EHS, community-based centers, FCCs); size of FCC or classroom and ages of children; level of support from center administration (where applicable); center or PLC climate; physical space; books, toys, and other sensory materials available for use by infants and toddlers; and caregiver beliefs about how to provide care and education to infants and toddlers?	WGT implementation components (the number of participants and level of engagement in the PLC, frequency of attendance, level of use of WGT, frequency of collecting videos of practices and reflecting on their use, the extent to which participation by caregivers and coaches was sustained throughout the study year)	Setting characteristics (setting type, size of FCC or classroom and ages of children, level of support from leaders, center climate, physical space and material resources, caregiver beliefs about how to provide care and education to infants and toddlers)	Barriers and supports to implementation, influence of interactions with coaches, organizational climate, motivation for caregiving, instructional leadership, examples of changes in self-efficacy	**Quantitative** • Descriptive analyses (means and standard deviations, ranges) • Pearson correlations among components and change scores • Data reduction when indicated • Multivariate regression analyses • Multivariate HLM regression analyses **Qualitative** • Thematic analysis to analyze and summarize information • Use standardized templates to organize and document the information and then apply codes and subcodes
4. What components of implementation are most strongly associated with positive changes in practice—for example, caregivers' level of engagement in the PLC, frequency of PLC meetings, frequency of attendance, level of use of WGT website, frequency of collecting videos of practices and reflecting on their	Associations with spring QCIT scores controlling for fall scores	WGT implementation components (the number of participants and level of engagement in the PLC, frequency of attendance, level of use of WGT, frequency of collecting videos of practices and reflecting on their use, the extent to which participation by caregivers	Perceptions about PLC participation, recording and reflecting on practices, interactions with peers, interactions with coaches, reflecting on children's responses and growth	**Quantitative** • Descriptive analyses (means and standard deviations, ranges) • Pearson correlations among components and change scores • Data reduction when indicated • Multivariate HLM (2L) regression analyses **Qualitative**

Research question	Quantitative		Qualitative	Description of analysis
	Outcome of interest	Example(s) of predictor	Example constructs	
use, and the extent to which participation by caregivers and coaches was sustained throughout the study year?		and coaches was sustained throughout the study year)		• Thematic analysis to analyze and summarize information • Use standardized templates to organize and document the information and then apply codes and subcodes
5. Do the associations between WGT components and positive change in classroom quality differ for classrooms with initial high and low levels of quality or with the PLC's focus?	Difference between fall and spring QCIT scores	WGT components (PLCs, virtual or in-person coaching); fall QCIT scores	NA	**Quantitative** • Descriptive analyses (means and standard deviations, ranges) and Pearson correlations with QCIT change scores • Multivariate HLM (2L) regression analyses • Moderation analyses
Resources				
6. What is the level of effort needed to implement WGT at the setting and network levels for center leaders, caregivers, coaches, and the study team—for example, time for coaching, for dedicated PD and instructional planning, for substitutes to cover for caregivers when they are meeting with coaches, for the study team to provide technical assistance?	level of effort	Use of WGT	Cost of implementing WGT; resources needed, level of effort	**Quantitative** • Descriptive analyses (obtain time estimates) **Qualitative** • Analyze qualitative interview data (weekly check-in data) and survey responses from directors and coaches and summarize information on cost and level of effort
7. What are the costs associated with materials needed to implement WGT—for example, technology costs such as devices, licensing fees, and internet access?	Cost	Materials to implement WGT	Cost of implementing WGT; resources needed, time	**Quantitative** • Descriptive analyses (obtain average costs) **Qualitative**

Research question	Quantitative		Qualitative	Description of analysis
	Outcome of interest	Example(s) of predictor	Example constructs	
				• Analyze qualitative interview data and survey responses from directors and coaches and summarize information on level of effort
8. Does the level of effort needed to implement WGT vary by setting type and, if so, in what ways?	Level of effort	Use of WGT by EHS, community-based centers, FCCs	level of effort	**Quantitative** • Descriptive analyses (obtain time estimates) **Qualitative** • Analyze qualitative interview data and survey responses from directors and coaches and summarize information on level of effort

NA = not applicable.

1. Quantitative data

First, the research team would conduct descriptive analyses and examine the means and standard deviations, range of scores, mean fall-spring differences for caregivers overall and with (1) initial low, middle, or high levels of quality, (2) high or low experience levels, (3) different setting types, and (4) primarily working with infants or toddlers.

With a focus on the conditions for optimal growth, the research team would utilize hierarchical linear models (HLM) that examine associations with different factors and account for the nesting within the sample. As evident in the prior field test, the inclusion of covariates also increases the precision of measurement of fall to spring change associated with WGT.

The research team would limit the number of factors that they include in a single analytic model. When appropriate, data reduction techniques would be implemented. Data reduction would be informed by interfactor correlations and exploratory factor analysis of the conditions expected to support change. The research team would examine the interfactor correlations by setting type as well as overall. The research design team hypothesizes that WGT and its components in different combinations for different types of settings (community-based versus EHS centers versus FCCs) would be associated with positive change in caregiver-child interactions and caregiver beliefs, knowledge, and self-efficacy. If these analyses suggest important differences, the team would use the qualitative data to inform understanding of how different conditions do or do not support positive change.

The primary outcome of interest would be the classroom-level QCIT total and domain scores. The team would use the HLMs to examine associations between the outcomes and predictors, controlling for caregiver, classroom, program, and coach characteristics. Given that classrooms are nested within PLCs, the two-level HLMs would account for the non-independence of the observations. Caregiver characteristics would be aggregated to the first level of the models. The use of covariates can increase the precision of estimates and provide added power to detect smaller effect sizes, but care would be taken in the number of covariates included in a single model.

Other outcomes of interest include caregivers' beliefs and knowledge about ECE and their self-efficacy. Because these outcomes are at the individual-level, three-level HLMs (caregivers nested within classrooms, and classrooms nested within PLCs) would account for the non-independence of the observations.

For the moderation analyses that would examine the conditions that may influence the use of WGT, or whether the associations between WGT components and positive change in classroom quality differ for

Bayesian priors and posteriors

The research team would use prior evidence from ECE interventions, and professional development and professional learning community (PLC) studies to help interpret the pre-post change estimates. Using Bayesian priors and posteriors allows for calculating the probability that an intervention had a meaningful effect, given the impact estimate and prior evidence regarding the effects of broadly similar interventions. Priors can be drawn from past evidence to develop an understanding of the probability that programs have effects of various magnitudes. Researchers can incorporate this prior information into the analysis and ultimately assess the probability that a program truly has positive effects and the likely size of those effects. Bayes' Rule calculates the probability that the program of interest truly had a positive effect for study participants given what is observed in the data (the impact estimate and standard error) and how often programs have had positive effects in the past.

classrooms with initial high and low levels of quality, the research team would use HLMs. The team would need to check that assumptions for moderation analyses are met before conducting these analyses.[21] For example, the moderation analyses would examine the cross-level interactions between the initial classroom-level quality of caregiver-child interactions as measured by the QCIT and the coach- or PLC-level variables (such as coaches' coaching experience or caregivers' level of engagement in PLC meetings). As an example, in order to examine the moderation effects of engagement in PLC meetings, level of engagement would be added as a predictor of the slope term for fall to spring classroom quality to determine if the fall to spring classroom quality association depends upon high versus low levels of caregiver engagement in the meetings.

Given that the sample size would be limited for some analyses, Bayesian methods (using priors and posteriors) are recommended to help interpret results—that is, to understand whether effect sizes are meaningful differences when power is weak.

2. Qualitative data

For the qualitative analyses, the research team would develop an initial coding scheme aligned with the construct of interest and from the interview and focus group questions. This would include constructs such as instructional leadership self-efficacy, PLC participation, and organizational climate. In the first stage of coding, the team would code interview and focus group responses related to these key constructs and identify any additional codes that are needed. These additional codes would then be added to other interviews as appropriate. In the second stage, the team would review all data coded within a specific construct to identify broad themes that triangulate across respondents or data sources. For example, "support from peers in the PLCs" and "encouragement from coach" could be emergent themes that fit within "organizational climate." In addition, coders would use a coding scheme to document key information, such as descriptions of the strategies being implemented. As coders read through the interview and focus group transcripts, some additional codes and subcodes may need further specifying. Emergent codes would then be added to the codebook and applied to all relevant data.

[21] Assumptions include continuous dependent and independent variables, a nominal moderator that has at least two groups, a linear relationship among included variables, no multicollinearity, no significant outliers, and the distributions of the variables should be approximately normal.

References

Agosti, J., S. Doyle, A. Douglass, and L. Mendes. "Field Guide for Implementation of a Breakthrough Series Collaborative in Early Care and Education." OPRE Report #2021-210. Office of Planning, Research, and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services, 2021.

Atkins-Burnett, S., S. Monahan, L. Tarullo, Y. Xue, E. Cavadel, L. Malone, and L. Akers. "Measuring the Quality of Caregiver-Child Interactions for Infants and Toddlers (Q-CCIIT)." OPRE Report #2015-13 and Appendices. Washington, DC: U.S. Department of Health and Human Services, Administration for Children and Families, Office of Planning, Research, and Evaluation, 2015.

Atkins-Burnett, Sally, Harshini Shah, Laura Kalb, and Cheri Vogel. "Teacher Beliefs About Infant/Toddler Care and Education." Princeton, NJ: Mathematica Policy Research, 2017.

Atkins-Burnett, S., L. Tarullo, S. Monahan, F. Hurwitz, T. Bruursema, A. Li, E. Blesson, et al. "The We Grow Together Professional Development System Final Report of the 2019 Field Test." OPRE Report #2020-170. U.S. Department of Health and Human Services, Office of Planning, Research, and Evaluation, Administration for Children and Families, 2020.

Bandura, A. *Self–Efficacy: The Exercise of Control*. New York: Freeman, 1997.

Barnett, W. Steven, and Leonard N. Masse. "Comparative Benefit-Cost Analysis of the Abecedarian Program and Its Policy Implications." *Economics of Education Review*, vol. 26, no. 1, 2007, pp. 113–125.

Barton, E. E., M. N. Rigor, E. A. Pokorski, M. Velez, and M. Domingo. "Using Text Messaging to Deliver Performance Feedback to Preservice Early Childhood Teachers." *Topics in Early Childhood Special Education*, vol. 39, no. 2, 2019, pp. 88–102.

Battey, D., and M. L. Franke. "Transforming Identities: Understanding Teachers Across Professional Development and Classroom Practice." *Teacher Education Quarterly*, vol. 35, no. 3, 2008, pp. 127–149.

Bloom, P. J. *Measuring Work Attitudes: Technical Manual for the Early Childhood Job Satisfaction Survey and Early Childhood Work Environment Survey*. New Horizons, 2016.

Borko, H. (2004). Professional Development and Teacher Learning: Mapping the Terrain. *Educational Researcher*, vol. 33, no. 8, pp. 3–15.

Bradshaw, C. P., C. W. Koth, K. B. Bevans, N. Ialongo, and P. J. Leaf. "The Impact of School-Wide Positive Behavioral Interventions and Supports (PBIS) on the Organizational Health of Elementary Schools." *School Psychology Quarterly*, vol. 23, no. 4, 2008, p. 462.

Bradshaw, C. P., C. W. Koth, L. A. Thornton, and P. J. Leaf. "Altering School Climate Through School-Wide Positive Behavioral Interventions and Supports: Findings from a Group-Randomized Effectiveness Trial." *Prevention Science*, vol. 10, no. 2, 2009, pp. 100–115.

Bray, J. N., J. Lee, L. L. Smith, and L. Yorks. *Collaborative Inquiry in Practice: Action, Reflection and Making Meaning*. Thousand Oaks, CA: Sage, 2000.

Building Bridges and Bonds (B3) Study Team. "Using Text Messages to Learn about Participants' Experiences." 2022. Available at https://www.acf.hhs.gov/opre/using-text-messages-learn-about-participants-experiences. Accessed December 1, 2022.

Burns, M. K., M. R. Naughton, J. L. Preast, Z. Wang, R. L. Gordon, V. Robb, and M. L. Smith. (2018). "Factors of Professional Learning Community Implementation and Effect on Student Achievement." *Journal of Educational and Psychological Consultation*, vol. 28, no. 4, pp. 394–412.

Buysse, V., K. L. Sparkman, and P. W. Wesley, P. W. (2003). "Communities of Practice: Connecting What We Know With What We Do." *Exceptional Children*, vol. 69, no. 3, pp. 263–277.

Campbell, F. A., E. P. Pungello, M. Burchinal, K. Kainz, Y.Pan, B. H. Wasik, O. A. Barbarin et al. "Adult Outcomes as a Function of an Early Childhood Educational Program: An Abecedarian Project Follow-Up." *Developmental Psychology*, vol. 48, no. 4, 2012, pp. 1033–1043.

Cannon, J., K. Schellenberger, A. Defnet, A. Bloomenthal, Y. Xue, and C. A. Vogel. "Baby FACES 2018: Data Users' Guide." Washington, DC: Office of Planning, Research, and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services, 2019.

Center on the Developing Child. "From Best Practices to Breakthrough Impacts: A Science-Based Approach to Building a More Promising Future for Young Children and Families." 2016. Retrieved from https://developingchild.harvard.edu/resources/from-best-practices-to-breakthrough-impacts/.

Cherrington, S. "Early Childhood Teachers' Thinking and Reflection Within Their Communities of Practice." Unpublished PhD thesis, Victoria University of Wellington, 2011.

Child Care State Capacity Building Center. "2022 Statewide Infant/Toddler Specialist Networks; A State Profile Compendium." Office of Planning, Research, and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services, 2022. https://childcareta.acf.hhs.gov/sites/default/files/public/itsn_state_profile_compendium_final_9.20.22_508_compliant_edits_lm.pdf.

Chu, M. *Developing Mentoring and Coaching Relationships in Early Care and Education: A Reflective Approach,* 1st ed. Boston, MA: Pearson, 2014.

Conti, G., J. J. Heckman, and R. Pinto. "The Health Effects of Two Influential Early Childhood Interventions on Health and Healthy Behaviours." *Economic Journal,* vol. 126, no. 596, 2016, pp. F28–F65.

Cox, S., R. Parmer, G. Strizek, and T. Thomas. "Documentation for the 2011–12 Schools and Staffing Survey (NCES 2016-817)." Washington, DC: U.S. Department of Education, National Center for Education Statistics, 2016. https://nces.ed.gov/surveys/sass/pdf/1112/SASS4A.pdf.

Crawford, A., K. A. Vaughn, C. L. Guttentag, C. Varghese, Y. Oh, and T. A. Zucker. "'Doing What I Can, But I Got No Magic Wand': A Snapshot of Early Childhood Educator Experiences and Efforts to Ensure Quality During the COVID-19 Pandemic." *Early Childhood Education Journal,* vol. 49, no. 5, 2021, pp. 829–840.

Cronbach, L. J. "Coefficient Alpha and the Internal Structure of Tests." *Psychometrika*, vol. 16, no. 3, 1951, pp. 297–334.

Devers, C., and E. Devers. "Text Messaging to Increase Job Satisfaction Among Early Childhood Educators." In *E-Learn: World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education*. Association for the Advancement of Computing in Education (AACE), October 2018, pp. 1168–1171.

Douglass, A., A. Carter, F. Smith, and S. Killins. "Training Together: State Policy and Collective Participation in Early Educator Professional Development." *New England Journal of Public Policy*, vol. 27, no. 1, 2015, p. 5. http://scholarworks.umb.edu/nejpp/vol27/iss1/5.

Douglass, A., K. Tout, and S. Doyle. "Considerations for Incorporating the Breakthrough Series Collaborative (BSC) as a Quality Improvement Methodology in Early Childhood Systems." OPRE Report #2021-211. Office of Planning, Research, and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services, 2021.

Downer, J. T., M. E. Kraft-Sayre, and R. C. Pianta. "Ongoing, Web-Mediated Professional Development Focused on Teacher-Child Interactions: Early Childhood Educators' Usage Rates and Self-Reported Satisfaction." *Early Education and Development,* vol. 20, no. 2, 2009, pp. 321–345.

Ehrlich, Stacy B., Debra Pacchiano, Amanda G. Stein, Maureen R. Wagner, Sangyoon Park, Elizabeth Frank, Stuart Luppescu et al. "Early Education Essentials: Validation of Surveys Measuring Early Education Organizational Conditions." *Early Education and Development,* vol. 30, no. 4, 2019, pp. 540–567.

Fong, K., and H. L. Horsley. "A Multidimensional Rasch Analysis of the Preschool Instructional Leadership Survey." Chicago, IL: University of Illinois, 2017.

Forry, N., R. Madill, E. Shuey, T. Halle, G. Ugarte, and J. Borton. "Snapshots from the NSECE: How Much Did Households in the United States Pay for Child Care in 2012? An Examination of Differences by Child Age." OPRE Report #2018-110. Office of Planning, Research and Evaluation, 2018.

Gerdes, J., and T. Jefferson. "How a Professional Learning Community Changed a Family Child Care Provider's Beliefs and Practices." *YC Young Children,* vol. 70, no. 5, 2015, pp. 8–13.

Halle, T., J. Cleveland, T. Bamdad, K. Nagle, K. Tout, A. Douglass, J. Agosti et al. "Promoting a Culture of Continuous Learning in Early Care and Education Settings." OPRE Report #2021-208. Office of Planning, Research, and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services, 2021.

Halle, T., R. Anderson, A. Blasberg, A. Chrisler, and S. Simkin. "Quality of Caregiver-Child Interactions for Infants and Toddlers (QCCIIT): A Review of the Literature." OPRE Report #201-25. Washington, DC: Office of Planning, Research, and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services, 2011.

Hamre, B. K., R. C. Pianta, M. Burchinal, S. Field, J. LoCasale-Crouch, J. T. Downer, C. Howes, et al. "A Course on Effective Teacher-Child Interactions: Effects on Teacher Beliefs, Knowledge, and Observed Practice." *American Educational Research Journal,* vol. 49, no. 1, 2012, pp. 88–123.

Hanno, E. C. "Nudging Early Educators' Knowledge, Beliefs, and Practices: An Embedded Randomized Controlled Trial of Text Message Supports." *Early Education and Development,* 2022, pp.1–21.

Harris, A., and M. Jones. "Professional Learning Communities and System Improvement." *Improving Schools,* vol. 13, no. 2, 2010, pp. 172–181.

Hart, C., D. Burts, R. Charlesworth, P. Fleege, M. Ickes, and M. Durland. "Instructional Activities Scale and Teacher Beliefs Scale." 1990.

Hirschler, J., and C. Darcy. "Using Technology to Support Preschool Teachers' Professional Development." In *Society for Information Technology & Teacher Education International Conference*. Association for the Advancement of Computing in Education, 2002, pp. 2550–2551.

Institute of Medicine and National Research Council. "The Early Childhood Care and Education Workforce: Challenges and Opportunities: A Workshop Report." Washington, DC: The National Academies Press, 2012.

Institute of Medicine and National Research Council. "Transforming the Workforce for Children Birth Through Age 8: A Unifying Foundation." Washington, DC: The National Academies Press, 2015.

Kerr, K. A., J. A. Marsh, G. S. Ikemoto, H. Darilek, and H. Barney. "Strategies to Promote Data Use for Instructional Improvement: Actions, Outcomes, and Lessons from Three Urban Districts." *American Journal of Education*, vol. 112, no. 4, 2006, pp. 496–520.

Kuh, Lisa Porter. "Promoting Communities of Practice and Parallel Process in Early Childhood Settings." *Journal of Early Childhood Teacher Education*, vol. 33, no. 1, 2012, pp. 19–37.

Lomos, C., R. H. Hofman, and R. J. Bosker. (2011). "Professional Communities and Student Achievement–A Meta-Analysis." *School Effectiveness and School Improvement*, vol. 22, no. 2, pp. 121–148.

Mashburn, A. J., J. T. Downer, B. K. Hamre, L. M. Justice, and R. C. Pianta. "Consultation for Teachers and Children's Language and Literacy Development During Pre-Kindergarten." *Applied Developmental Science*, vol. 14, no. 4, 2010, pp. 179–196.

McConnaughy, E. A., J. O. Prochaska, and W. F. Velicer. "Stages of Change in Psychotherapy: Measurement and Sample Profiles." *Psychotherapy: Theory, Research and Practice*, vol. 20, no. 3, 1983, pp. 368–375.

McCormick Center for Early Childhood Leadership. "Administrator Role Perception Survey—Center Based." Wheeling, IL: McCormick Center for Early Childhood Leadership, April 2019.

Milli, J. "Why Investing in Child Care Providers Is Essential for Providers, Children, and Families." Center for Law and Social Policy, National Women's Law Center, and the Century Foundation, May 2022.

Moreno, A. J., S. Green, and J. Koehn. "The Effectiveness of Coursework and Onsite Coaching at Improving the Quality of Care in Infant-Toddler Settings." *Early Education and Development*, vol. 26, no. 1, 2015, pp. 66–88.

National Survey of Early Care and Education Project Team. "Number and Characteristics of Early Care and Education (ECE) Teachers and Caregivers: Initial Findings from the National Survey of Early Care and Education (NSECE)." OPRE Report #2013-38. Washington, DC: Office of Planning, Research, and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services, 2013.

National Survey of Early Care and Education Project Team. "National Survey of Early Care and Education (NSECE), [United States], 2019." Inter-university Consortium for Political and Social Research, 2022. https://doi.org/10.3886/ICPSR37941.v5.

National Institute of Child Health and Human Development (NICHD) Early Child Care Research Network. "Child-Care Structure→ Process→ Outcome: Direct and Indirect Effects of Child-Care Quality on Young Children's Development." *Psychological Science*, vol. 13, no. 3, 2002, pp. 199–206.

Nguyen, T., S. Atkins-Burnett, S. Monahan, L. Tarullo, Y. Xue, & J. Cannon. "Supporting Responsive Caregiving for the Youngest Children: Psychometric Evidence of the Quality of Care for Infants and Toddlers (QCIT) Observational Measure." *Early Education and Development*. Published ahead of print, December 5, 2022. https://doi.org/10.1080/10409289.2022.2144685.

Norris, D. J. "Raising the Educational Requirements for Teachers in Infant Toddler Classrooms: Implications for Institutions of Higher Education." *Journal of Early Childhood Teacher Education*, vol. 31, no. 2, 2010, pp. 146–158.

Olson, A. N., S. Cotner, C. Kirkpatrick, S. Thompson, and S. Hebert. "Real-Time Text Message Surveys Reveal Student Perceptions of Personnel Resources Throughout a Course-Based Research Experience." *PloS One*, vol. 17, no. 2, 2022, p. e0264188.

Paschall, K. "Nearly 30 Percent of Infants and Toddlers Attend Home-Based Child Care as Their Primary Arrangement." Child Trends, 2019. https://www.childtrends.org/blog/nearly-30-percent-of-infants-and-toddlers-attend-home-based-child-care-as-their-primary-arrangement.

Pianta, R. C., A. J. Mashburn, J. T. Downer, B. K. Hamre, and L. Justice. "Effects of Web-Mediated Professional Development Resources on Teacher-Child Interactions in Pre-Kindergarten Classrooms." *Early Childhood Research Quarterly*, vol. 23, no. 4, 2008, pp. 431–451.

Porter, K. E., J. C. Snipes, and J. Eisberg. "The Search for Progress; Elementary Student Achievement and the Bay Area School Reform Collaborative's Focal Strategy." MDRC, 2006.

Ronfeldt, M., S. O. Farmer, K. McQueen, and J. A Grissom. (2015). "Teacher Collaboration in Instructional Teams and Student Achievement." *American Educational Research Journal*, vol. 52, no. 3, pp. 475–514.

Schachter, R. E., H. K. Gerde, and H. Hatton-Bowers. "Guidelines for Selecting Professional Development for Early Childhood Teachers." *Early Childhood Education Journal*, vol. 47, no. 4, 2019, pp. 395–408.

Schmit, S., and H. Matthews. "Better for Babies: A Study of State Infant and Toddler Child Care Policies." Washington, DC: Center for Law and Social Policy, 2013.

Schoenfeld, A. H. "Video Analyses for Research and Professional Development: The Teaching for Robust Understanding (TRU) Framework." *ZDM*, vol. 50, no. 3, 2018, pp. 491–506.

Shannon, D. K., P. A. Snyder, M. L. Hemmeter, and M. McLean. "Exploring Coach-Teacher Interactions Within a Practice-Based Coaching Partnership." *Topics in Early Childhood Special Education*, vol. 40, no. 4, 2021, pp. 229–240.

Sheridan, S. M., C. P. Edwards, C. A. Marvin, and L. L. Knoche. "Professional Development in Early Childhood Programs: Process Issues and Research Needs." *Early Education and Development*, vol. 20, 2009, pp. 377–401.

Shimoni, N., S. Nippita, and P. M. Castaño. "Best Practices for Collecting Repeated Measures Data Using Text Messages." *BMC Medical Research Methodology*, vol. 20, no. 1, 2020, pp. 1–7.

Shonkoff, J. P. "Breakthrough Impacts: What Science Tells Us About Supporting Early Childhood Development." *YC Young Children*, vol. 72, no. 2, 2017, pp. 8–16. https://www.jstor.org/stable/90004117.

Smythe-Leistico, K., and L. C. Page. "Connect-Text: Leveraging Text-Message Communication to Mitigate Chronic Absenteeism and Improve Parental Engagement in the Earliest Years of Schooling." *Journal of Education for Students Placed at Risk (JESPAR)*, vol. 23, nos. 1–2, 2018, pp. 139–152.

Snyder, P. A., M. L. Hemmeter, K. Artman, K. Kinder, C. Pasia, and T. McLaughlin. "Characterizing Key Features of the Early Childhood Professional Development Literature." *Infants & Young Children*, vol. 25, 2012, pp. 188–212.

Spillane, J. P. "External Reform Initiatives and Teachers' Efforts to Reconstruct Their Practice: The Mediating Role of Teachers' Zones of Enactment." *Journal of Curriculum Studies,* vol. 31, no. 2, 1999, pp. 143–175.

Sroufe, L. A. "The Role of Infant-Caregiver Attachment in Development." In *Clinical Implications of Attachment,* edited by J. Belsky and T. Nezworski. Hillsdale, NJ: Erlbaum, 1988, pp. 18–38.

Sroufe, L. A. "Attachment and Development: A Prospective, Longitudinal Study from Birth to Adulthood." *Attachment & Human Development,* vol. 7, no. 4, 2005, pp. 349–367. https://doi.org/10.1080/14616 73050 0365928.

Stoll, L., R. Bolam, A. McMahon, M. Wallace, and S. Thomas. "Professional Learning Communities: A Review of the Literature." *Journal of Educational Change,* vol. 7, no. 4, 2006, pp. 221–258.

Study of Coaching Practices in Early Care and Education Settings (SCOPE) Project Team. "Study of Coaching Practices in Early Care and Education Settings (SCOPE) Coach Survey, 2016–2019." Inter-university Consortium for Political and Social Research, 2019. https://doi.org/10.3886/ICPSR38290.v1.

National Survey of Early Care and Education Project Team. "National Survey of Early Care and Education (NSECE), [United States], 2019." Inter-university Consortium for Political and Social Research, 2019. https://doi.org/10.3886/ICPSR37941.v5.

Park, V., & Datnow, A. "Co-Constructing Distributed Leadership: District and School Connections in Data-Driven Decision-Making." *School Leadership and Management*, vol. 29, no. 5, 2009, 477–494.

Tashakkori, A., C. Teddlie, and C. B. Teddlie. *Mixed Methodology: Combining Qualitative and Quantitative Approaches,* vol. 46. Sage, 1998.

Tashakkori, A., and C, Teddlie. "Issues and Dilemmas in Teaching Research Methods Courses in Social and Behavioural Sciences: US Perspective." *International Journal of Social Research Methodology,* vol. 6, no. 1, 2003, pp. 61–77.

Taylor, M. J., P. R. Hallam, C. T. Charlton, and D. G. Wall. "Formative Assessment of Collaborative Teams (FACT): Development of a Grade-Level Instructional Team Checklist." *NASSP Bulletin,* vol. 98, no. 1, 2014, pp. 26–52. https://doi.org/10.1177/0192636513514109.

Torff, B., D. Sessions, and K. Byrnes. "Assessment of Teachers' Attitudes About Professional Development." *Educational and Psychological Measurement,* vol. 65, 2005, pp. 914–924.

Tschannen-Moran, M., and A. Woolfolk Hoy. "Teacher Efficacy: Capturing an Elusive Construct." *Teaching and Teacher Education,* vol. 17, no. 7, 2001, pp. 783–805.

Turnbull, B. "Teacher Participation and Buy-In: Implications for School Reform Initiatives." *Learning Environments Research,* vol. 5, no. 3, 2002, pp. 235–252.

Wanless, S. B., C. Groark, and B. Hatfield. "Assessing Organizational Readiness." In *Handbook of Social Emotional Learning,* edited by J. Durlak, R. Weissburg, and T. Gullotta. New York, NY: Guilford Publications, 2015, pp. 360–376.

Wesley, Patricia W., and Virginia Buysse. "Communities of Practice: Expanding Professional Roles to Promote Reflection and Shared Inquiry." *Topics in Early Childhood Special Education,* vol. 21, no. 2, 2001, pp. 114–123.

Whitford, H. M., P. T. Donnan, A. G. Symon, G. Kellett, E. Monteith-Hodge, P. Rauchhaus, and J. C. Wyatt. "Evaluating the Reliability, Validity, Acceptability, and Practicality of SMS Text Messaging as a Tool to Collect Research Data: Results from the Feeding Your Baby Project." *Journal of the American Medical Informatics Association,* vol. 19, no. 5, 2012, pp. 744–749.

Woodcock R. W. "What Can Rasch-Based Scores Convey About a Person's Test Performance?" In The New Rules of Measurement: What Every Psychologist and Educator Should Know, edited by S. E. Embretson & S. L. Hershberger. Mahwah, NJ: Lawrence Erlbaum, 1999.

Yoon, S. A., Anderson, E., Klopfer, E., Koehler-Yom, J., Sheldon, J., Schoenfeld, I., et al. "Designing Computer-Supported Complex Systems Curricula for the Next Generation Science Standards in High School Science classrooms." *Systems,* vol. 4, no. 4, 2016, 38.

York, B. N., S. Loeb, and C. Doss. "One Step at a Time: The Effects of an Early Literacy Text Messaging Program for Parents of Preschoolers." *Journal of Human Resources,* vol. 54, no. 3, 2019, pp. 537–566. https://doi.org/10.3368/jhr.54.3.0517-8756R.

Zaslow, M., K. Tout, T. Halle, J. V. Whittaker, and B. Lavelle. "Toward the Identification of Features of Effective Professional Development for Early Childhood Educators." Literature review. Office of Planning, Evaluation, and Policy Development, U.S. Department of Education, 2010.

Zimmerman, J. "Why Some Teachers Resist Change and What Principals Can Do About It." *National Association of Secondary School Principals (NASSP) Bulletin*, vol. 90, no. 3, 2006, 238–249.

Mathematica Inc.

Princeton, NJ • Ann Arbor, MI • Cambridge, MA
Chicago, IL • Oakland, CA • Seattle, WA
Tucson, AZ • Woodlawn, MD • Washington, DC

mathematica.org website

EDI Global, a Mathematica Company

Operating in Tanzania, Uganda, Kenya, Mozambique, and the United Kingdom