Ready to Read: Cradling Literacy Study

Social Innovation Fund
Final Report

Submitted to Clayton Early Learning and Mile High Early Learning

Prepared by
The Butler Institute for Families
Graduate School of Social Work
University of Denver

UNIVERSITY of DENVER
BUTLER INSTITUTE FOR FAMILIES
Graduate School of Social Work

ACKNOWLEDGMENTS

This publication represents a collaborative effort by the Butler Institute for Families at the University of Denver, Graduate School of Social Work, Clayton Early Learning, and Mile High Early Learning. This work was made possible by the Corporation for National and Community Service (CNCS) and Mile High United Way. We would like to thank Julie McCrae, PhD, for her contributions as the Principal Investigator of this study from 2014–2016.

Principal authors:

Ann Deaton Wacker, MA
Butler Institute for Families, University of Denver
Ann.Wacker@du.edu
303.871.4873

Shauna Rienks, PhD
Butler Institute for Families, University of Denver
Shauna.Rienks@du.edu
303.871.3886

Miriam Estrada
Butler Institute for Families, University of Denver
Miriam.Estrada@du.edu
303.871.4003

For further information about the Ready to Read program contact:

Shelly Anderson, MA
Clayton Early Learning
SAnderson@claytonearlylearning.org
303.398.8563

Pamela Harris, PhD
Mile High Early Learning
PamelaH@milehighearlylearning.org
303.962.2623

Recommended citation:
Butler Institute for Families, University of Denver. (2017). *Ready to Read: Cradling Literacy Study, Social Innovation Fund Final Report.* Denver, Colorado.

For more information:
Please visit the Butler Institute for Families website at www.thebutlerinstitute.org

CONTENTS

EXECUTIVE SUMMARY

Clayton Early Learning and Mile High Early Learning are implementing *Ready to Read*, an innovative program designed to increase early literacy skills among low-income children age birth to three in Denver. The *Ready to Read* study in center-based sites began in 2012 after being selected by Mile High United Way (a Social Innovation Fund recipient) to build the evidence base of a promising literacy program, *Cradling Literacy* (*CL*). This is the final evaluation report of the *Ready to Read* center-based five-year study.

Cradling Literacy is a 12-module training curriculum intended for teachers of young children (from birth to age 5) to increase teachers' knowledge and skills related to early childhood development, literacy and learning, and family engagement (Im, Osborn, Sanchez, & Thorp, 2007). *CL* provides a framework for early language and literacy by focusing on the importance of storytelling, culture, and relationships, as well as how children develop language skills, and ways teachers can support emergent literacy and social-emotional literacy. Topics include language, books, text, and stories, with concrete teaching behaviors that can be implemented to increase children's skills. Practical tips and reflections are used throughout the curriculum to demonstrate important topics and reinforce applied learning in classroom settings.

In the *Ready to Read* study, teachers at some early learning centers received *CL* training, along with coaching to reinforce the concepts provided in training. Teachers in other centers did not receive *CL*. Both groups received training and coaching in *Dialogic Reading* (*DR*), an evidence-based early literacy program that focuses on parents and caregivers reading interactively with their children.

This impact study includes four confirmatory research questions, examining whether teachers and families at centers that were implementing *CL* in addition to *DR* provided greater support for literacy activities or experienced greater gains in children's oral and communication skills relative to those at centers that did not use *CL*. The two exploratory research questions address the extent to which dosage (attendance in a *CL* classroom) influences those same outcomes. Based on the strong level of evidence for *DR* and the preliminary level of evidence for *CL*, a moderate level of evidence was targeted. To achieve that level of evidence, this study used a quasi-experimental study design, randomly assigning the six centers to either the experimental (*CL*) or comparison ("programming as usual" or "*DR* only") conditions.

Center-Based Study

Intended Outcomes

- Increased language-rich interactions between children and teachers or parents/caregivers
- Increased quality and frequency of book reading
- Increase in children's oral language and communication skills

Measures

Child developmental assessments and parent surveys were completed at children's entry to a Cradling Literacy classroom and, again, three and six months later.

- Parent Survey of Home Literacy (use of interactive book reading techniques)
- Baby FACES, Reading Books and Telling Stories (frequency of book reading and storytelling)
- Teaching Strategies™ GOLD® (TSG)
- Ages and Stages Questionnaires, Communication Subscale (ASQ-C)
- MacArthur-Bates Communicative Development Inventories (CDI)

Teacher knowledge, behavior, and training satisfaction were assessed through observation and surveys.

- Classroom Assessment Scoring System, infant and toddler version (CLASS® I/T is an observational tool used to gauge teachers' *Emotional and Behavioral Support* of students, as well as their *Engaged Support for Learning*
- Teacher survey (gauged teachers' pre- and post-knowledge of practices that support early literacy and their satisfaction with the training)

Analysis

To examine change over time among intervention and comparison families and teachers, we conducted paired samples *t*-tests and Analysis of Variance (ANOVA).

Implementation and Impact Findings

Ready to Read has been delivered in six early care and education (ECE) centers in the Denver metropolitan area since September 2012. Programming is offered year-round, five days a week, and up to 10.5 hours per day in a classroom setting that has up to eight children per class. Teachers and staff are diverse in terms of age and race/ethnicity and range in qualifications from *Child Development Associate* to holding a bachelor's degree. Classrooms have a minimum adult: child ratio of 1:4.

During the time of this *Ready to Read* study, 817 children ranging in age from 1 month to 30 months were enrolled in one of the six centers. Study eligibility criteria included the child being 30 months or younger at the time of enrollment in a center, and the parent/caregiver consenting to participate in the study within five weeks of center enrollment. A total of 333 families enrolled in the research study, for an enrollment rate of 41%:

- Intervention condition: 186 (42% of eligible participants from 3 sites)
- Comparison condition: 147 (40% of eligible participants from 3 sites)

Study data indicate that *Ready to Read* centers serve diverse families: many children are African American (34%), and one-third are Latino/a (32%). Nearly all qualify for free or reduced lunch (89%), and 4% had an Individualized Family Service Plan (IFSP), a marker of special education services. Rates of IFSPs among children in this study are slightly higher than the 3% found among children in the general population (US Department of Education, 2015). Follow-up data in some form (parent survey, Teaching Strategies GOLD [TSG], Ages and Stages Questionnaire [ASQ-C], or Communicative Development Inventory [CDI]) are available for 64% of the children who completed baseline assessments.

Implementation Evaluation

Originally, the study included five implementation questions related to the *DR* and *CL* programming:

How many individuals were trained on DR*?*

- A total of 1,272 center-based staff, parents and caregivers, community members, and volunteers were trained on *DR*.

How many modules of CL *training did center-based staff receive?*

- Teachers received an average of about nine *CL* modules per year.

Did center-based teachers demonstrate increased knowledge after attending the CL *training sessions? To what extent did teachers in the* DR*-only group demonstrate knowledge of the concepts covered in the* CL *curriculum?*

- Teacher knowledge was assessed via a survey, with results indicating that across all *Ready to Read* teachers (intervention and comparison), there was a statistically significant increase in knowledge of teacher practices to support early literacy between the start of the study and the first follow-up. Further results are not included in this report. For more information, see the *Ready to Read: Cradling Literacy Study, Social Innovation Fund Year 4 Annual Report.*

Did center-based teachers find the CL *training to be useful?*

- Based on training evaluations, teachers had positive views of the *CL* modules and the training delivery. Further results gleaned from training evaluations are not included in this report. For more information, see the *Ready to Read: Cradling Literacy Study, Social Innovation Fund Year 4 Annual Report*.

During the 2015–2016 school year, three additional implementation questions were added to the study:
What are the characteristics of the CL *intervention as it is being implemented?*
What is the consistency of CL *program delivery?*
What is the quality of CL *program delivery?*

These questions were only partially addressed through the CL implementation observational study conducted during the 2015–2016 school year:

- Results of one *CL* training observation indicated that the session included activities with the full group of teachers, as well as a breakout into smaller reflective practice groups to allow for deeper discussion before coming back together at the end of the training.
- Overall, coaches were consistent in how they delivered the training, but flexible in how they adapted to the group of teachers they were working with. Coaches also offered frequent encouragement and a variety of activities to keep teachers engaged. Further results are not included in this report. For more information, see the *Ready to Read: Cradling Literacy Study, Social Innovation Fund Year 4 Annual Report.*

During the 2016–2017 school year, three additional implementation questions were added to the study as a way to gather contextual information about *CL* teachers' and coaches' views of the program at the conclusion of the study:

What are teachers' and coaches' perceptions of the impacts of CL *training and coaching?*

- According to *CL* coaches and teachers, *CL* makes an impact at multiple levels, ranging from the classroom environment and teaching practices to family engagement and parents'/caregivers' and children's skills and behaviors. For example, one teacher credited *CL* for creating a culture that focuses on language and literacy: *"I think that's why our children are so interested in books and excited about them."*

What are teachers' and coaches' views of the content and delivery of CL*? What aspects were most/least successful and helpful to them?*

- Teachers and coaches reported that the *CL* content was useful and that they received helpful tools, resources, and materials. The extent to which the content was "new" or "helpful" to teachers seemed to depend on how experienced each teacher was. The most favorably received modules included: Behavioral Understanding, Development, Relationships, and Stories.
- In terms of format of training delivery, there was a general preference for the monthly training sessions provided in Year 1 and the peer-to-peer training from Year 3. Challenges related to implementing *CL* included finding convenient times for training and coaching, keeping the content from becoming "repetitive" over the course of the project, and balancing other work responsibilities with *CL*.

Program Impact

This study addressed six impact questions (four confirmatory, two exploratory):

Impact (Confirmatory)

Do teachers in the CL *group demonstrate greater short-term improvements in their support of children's language and literacy development than do teachers in the* DR*-only group?*

- *Engaged Support for Learning* (as measured by the CLASS I/T observations) increased significantly across all teachers between baseline and the first follow-up.
- *Emotional and Behavioral Support* (CLASS I/T) increased significantly across all teachers between baseline and the first follow-up.
- Based on data collected during Years 1–3, use of *DR* techniques increased significantly across all teachers between baseline and the first follow-up. For more information, see the *Ready to Read: Cradling Literacy Study, Social Innovation Fund Year 3 Annual Report.*

Thus, on average, teachers demonstrated short-term improvements in support of children's language and literacy, regardless of whether they worked at an intervention or comparison center.

Do teachers in the CL *group demonstrate greater long-term improvements in their support of children's language and literacy development than do teachers in the* DR*-only group?*

- *Engaged Support for Learning* (CLASS I/T) decreased after the first follow-up and fluctuated from time point to time point but remained in the middle range throughout the study.
- *Emotional and Behavioral Support* (CLASS I/T) decreased after the first follow-up and fluctuated from time point to time point but remained in the middle-to-high range throughout the study.
- Teachers' use of *DR* techniques remained consistent between the first and second follow-ups, suggesting that gains were sustained over time. For more information, see the *Ready to Read: Cradling Literacy Study, Social Innovation Fund Year 3 Annual Report.*

Do parents/caregivers in the CL *group show greater gains in their support of their child's language and literacy development than do the parents/caregivers in the* DR-*only group?*

- Overall, parents/caregivers showed a marginally significant increase in storytelling frequency between baseline and the 6-month follow-up, with intervention parents/caregivers increasing significantly over time while comparison parents/caregivers did not (when examined separately by group).
- Parent/caregiver-reported interactive reading increased significantly across all study participants between baseline and the 6-month follow-up.
- Parent/caregiver-reported reading frequency increased significantly across all study participants between baseline and the 6-month follow-up.

Do children in the CL *group demonstrate greater increases in their oral language and communication skills than do children in the* DR-*only group?*

- Parent/caregiver-reported words produced increased significantly across all children between baseline and the 6-month follow-up.
- Children's language and literacy scores on teacher-reported assessments showed statistically significant increases between the first and third Teaching Strategies GOLD (TSG) checkpoints after enrolling in the center.
- For intervention group children, there was a statistically significant increase in the proportion of children classified as "on schedule" in terms of verbal and nonverbal communication between baseline (84%) and the 6-month follow-up (93%). Comparison group children's classifications remained similar over time, with around 90% being "on schedule."

Impact (Exploratory)

Do parents/caregivers whose children attend a center more frequently show greater gains in their support of their child's language and literacy development than do those with a lower level of attendance?

- Parents'/caregivers' gains in reading frequency and quality and storytelling frequency were not related to how frequently their child attended a center.

Do children with higher levels of attendance in a center demonstrate greater gains in their oral language and communication skills than do those with lower levels of attendance?

- Intervention group children who attended a center more frequently made significantly greater gains in language skills and words produced than did intervention group children who attended less frequently.

Summary

Although this five-year study did not establish a moderate level of evidence for the *Cradling Literacy* intervention, there is preliminary evidence that receiving *CL* training and coaching made an impact at the parent and child levels.

In addition, most *Ready to Read* families experienced an increase in parent/caregiver support of language and early literacy, as well as gains in children's vocabularies. Finally, results showed a boost in classroom-level support for learning and teachers' knowledge of practices to support early literacy.

Contribution of the Study

Results of the *Ready to Read* center-based study show a number of positive outcomes. Children's oral language and communication skills increased over time, according to both parent-reported measures (ASQ-C; CDI) and teacher-reported assessments (TSG). This includes increases in children's vocabularies and fewer children having communication scores in the "concern range" at follow-up. Additionally, parents/caregivers made gains in the frequency with which they read and told stories with their child and also increased the quality of reading with their child. Finally, teachers showed short-term increases in their support for learning and emotional/behavioral support in the classroom.

Regarding the impact of the *CL* intervention, results showed that intervention families made significant gains in the frequency of parent-child storytelling and in the proportion of children classified as "on schedule" in terms of verbal and nonverbal communication. By contrast, families in the comparison group did not experience significant growth in these areas. Additionally, intervention group children who attended the center more frequently made greater gains in language skills and words produced than did intervention group children who attended less frequently. Therefore, there is some evidence that assignment to and/or more frequent attendance in a *CL* classroom translates to greater gains in parent support of early literacy and child outcomes.

This five-year study demonstrated that:

- In general, teachers showed short-term gains related to their support of early literacy.
- In general, parents/caregivers significantly increased their support of early language and literacy.
- In general, children's language and literacy skills increased.
- Parents/caregivers in the intervention group made significant gains over time in parent-child storytelling, while comparison parents/caregivers remained stable or showed more modest gains.
- Children who attended an intervention center more frequently made greater gains than did intervention-group children who attended less frequently.

However, we cannot confidently attribute improvements in scores over time solely to the *CL* intervention as *CL* teachers and families did not make significantly *greater* gains than did the *DR*-only group on most indicators of literacy when groups were compared directly to each other (the gains were relative to their own starting point). As such, the goal of increasing the level of evidence for the *CL* program to "moderate" was not met.

Strengths and Limitations

This study shows improved outcomes for both children and parents/caregivers during their participation in *Ready to Read* and seems to provide preliminary support for the effectiveness of the *Cradling Literacy* intervention (when paired with *Dialogic Reading*). The existence of a comparison group was a clear strength of the study design as was the effort to collect multiple types of data from multiple sources. On the other hand, assessing child-level impact was complicated by issues such as lack of group equivalence in terms of race/ethnicity and the possibility of recall bias on parent-reported data. Although the sample size was adequate to detect medium to large effects, interventions such as this are more likely to result in small effect sizes, leaving limited statistical power for some analyses. Results generally did not indicate that the intervention had an impact above-and-beyond "programming as usual" (*DR* only). It is possible that broad measures such as the Classroom Assessment Scoring System, infant and toddler version (CLASS I/T), and TSG are not sensitive to the subtle impacts of the *CL* intervention. Another potential explanation for these findings is that receipt of any sort of early learning programming may be the key factor in promoting early literacy, and adding *CL* does not substantially alter teacher- and family-level outcomes.

Connection to Future Research

Previous research provides a strong level of evidence for *DR*. However, because that intervention focuses on a very small amount of time in a child's life (shared book-reading experiences), the purpose of this *Ready to Read* study was to layer another intervention on top of *DR* in hopes of enhancing a child's exposure to rich language at other times in the

day. *CL* provides ECE teachers with a theoretical foundation for implementing *DR* related to multiple areas of child development. This study provided preliminary evidence that parents/caregivers whose child attended a *CL* center experienced greater gains in storytelling frequency, and children who attended more often made greater language and literacy gains than did their peers with poorer attendance. Future research should replicate this work with larger and more diverse samples to more closely examine for whom the intervention is most effective, with a particular emphasis on differences between newer and more experienced teachers.

INTRODUCTION

Clayton Early Learning has a history of leadership in early care and education for vulnerable children through training, practice, and research. Founded more than 100 years ago as an orphanage and school for needy boys, Clayton offers prenatal through preschool programs (primarily Head Start or Early Head Start) at two school locations—Educare Denver School and Clayton Early Learning School. Clayton provides a range of programming and emphasizes effective teacher-child interaction, evidence-based practices, advocacy for vulnerable children, and collaboration with policy makers and funders.

Mile High Early Learning, Denver's largest and oldest provider of quality subsidized early care and education, has been serving vulnerable children in the Denver area since 1970. Through its year-round Montessori-inspired education programming provided in seven centers (four of which serve infants and toddlers), Mile High Early Learning serves more than 500 children daily and also reaches several thousand children through its drop-in centers annually, in addition to community education and professional learning programs.

In 2012, Clayton Early Learning and Mile High Early Learning centers were selected by Mile High United Way to be part of the Social Innovation Fund initiative for their promising early literacy program, *Ready to Read*. *Ready to Read* addresses the challenges faced by infants and toddlers growing up in low-income households by supporting key adults in their lives to engage more frequently and effectively with them in language-based activities. *Ready to Read* builds on the base of *Dialogic Reading* (*DR*), a book-reading and story sharing intervention for which there is a strong level of evidence. To bolster adult-child interaction throughout the day, *DR* was supplemented by *Cradling Literacy* (*CL*), a training curriculum that is designed to provide teachers with a solid foundation in children's language and literacy development. This study tests whether the addition of *CL* results in even greater gains in children's oral language and communication skills than *DR* alone.

Since 2014, the Butler Institute for Families has served as the evaluator for this initiative, conducting a quasi-experimental impact evaluation to determine the effectiveness of *CL* teacher training and coaching for: 1) increasing the frequency and quality with which parents/caregivers and teachers engage with children in language-based activities, and 2) increasing children's growth in oral language and communication skills. This is the final evaluation report of the five-year study of *Ready to Read*'s *CL* intervention. This report is primarily intended for stakeholders and funders, though it may be of interest to others in the field of early childhood education.

Program Background and Problem Definition

Research has shown that when adults speak more with children, starting from an early age, those children have better vocabularies, which translates to greater reading and writing proficiency. Research also shows that children from low-income families tend to receive less support for language and literacy development than do children from middle- and upper-income families. Specifically:

- On average, low-income parents talk with their children much less than do higher-income parents. By the age of four, the average low-income child has heard 30 million fewer words than have his or her higher-income peers. One key study demonstrated that the vocabulary gap at age three predicted language scores in third grade (Hart & Risley, 2003)
- Another study identified differences in children's vocabulary knowledge based on socioeconomic status. The estimated disparity in vocabulary size between socioeconomic groups was about 15,000 words, with linguistically disadvantaged children knowing about 5,000 words compared to the more advantaged who knew 20,000 words (Moats, 1999).
- A child from a middle-income family typically enters first grade with about 1,000 hours of one-on-one picture book reading time with parents and other adults, compared with a child from a low-income family, who averages fewer than 100 hours (Adams, 1990).

Before children can become proficient in reading and writing, they need to experience the wonder of storytelling and books within environments rich in vocabulary and oral language. As part of their journey, children need opportunities to build conceptual knowledge about the world around them. A critical pathway for developing these fundamental early literacy skills is adult-child interaction with books and storytelling (Shickendanz, 1999).

In an attempt to identify effective language and literacy interventions to alleviate these disparities, Clayton Early Learning and Mile High Early Learning partnered to conduct a quasi-experimental study of a promising early language and literacy intervention in their center-based programs that primarily serve low-income families. *Ready to Read* addresses the early literacy challenges for infants and toddlers growing up in low-income households by supporting key adults in their lives to engage more frequently and effectively with them in language-based activities. As part of standard programming, families receive training in *DR*, an evidence-based literacy intervention that can be easily implemented by parents/caregivers in the home and has been shown to have a positive impact on children's language development. However, because book reading is only a small part of the time that parents/caregivers spend interacting with their children, the current study tested the added value of another intervention that has a preliminary level of evidence, *CL* training and coaching, which provides ECE teachers with a theoretical foundation for implementing *DR*.

Overview of Prior Research

Dialogic Reading. *DR* is an interactive method of sharing picture books with young children ages birth to five. Rather than adults reading and children listening, in *DR*, children learn to become storytellers. According to the What Works Clearinghouse (WWC) maintained by the US Department of Education's Institute of Education Sciences, there is "strong" evidence that *DR* improves oral language skills, based on four randomized controlled studies that met its evidence standards and one randomized controlled study that met evidence standards with reservations. The studies reviewed by WWC focused on children 25 years old, which is older than the children in the *CL* study. However, a synthesis of practice-based evidence supported the use of *DR* with children under age three (Cutspec, 2007).

Cradling Literacy. The *CL* training curriculum is based upon a professional development intervention, *Literacy, Learning, and Life* (*LLL*), developed by Zero to Three. *LLL* was intended to build the capacity of early childhood educators to improve language and literacy outcomes of at-risk children birth to five years (Im, Osborn, Sanchez, & Thorp, 2007). Results of a quasi-experimental study of *LLL* conducted by independent evaluators showed that children enrolled in *LLL* classrooms experienced a gain in their standard Peabody Picture Vocabulary Test-III score of 4.5 points, while children in the comparison group gained only 1.3 points (Im, Osborn, Sanchez, & Thorp, 2007). In addition, there were positive gains from pre- to post-enrollment for *LLL* participants in terms of the quality of the classroom environment, teachers' knowledge of early childhood practices, and families' home literacy environment.

Theory of Change

The theory of change for the *Cradling Literacy* intervention is that when teachers receive *CL* training and coaching, they will gain a theoretical foundation and culturally responsive framework for understanding children's language and literacy development and will engage in more frequent and rich verbal interactions with children throughout the day (both during book-reading sessions and at other times). It is also posited that teachers will be more motivated to implement *DR* more frequently and with greater fidelity, and that they will sustain those efforts over a longer period of time if they have a theoretical foundation for understanding the development of children's oral language and communication skills. Further, after receiving the *DR* and *CL* trainings, these teachers will provide clearer messages and better support to families to engage in language-rich interactions in the home. Richer and more frequent interactions during book reading and at other times will support children's oral language and communication skills. These skills are, in turn, associated with school readiness and later reading success, leading to third-grade reading proficiency.

Cradling Literacy Program Model

This center-based study included three key program components for serving infants and toddlers and their families:

- Year-round full-day programming in classroom settings led by qualified teachers
- Training in *DR* and *CL* for teachers
- Coaching for teachers

Description of Program Components

Center-based programming. All classroom teachers are experienced early childhood educators, providing year-round instruction to infants and toddlers in a classroom setting. During typical programming, teachers infuse literacy-based activities throughout the day and encourage parents' use of literacy activities at home. For example, parents record their reading time with children and turn in reading logs monthly.

***Dialogic Reading* (*DR*) training and coaching.** Parents/caregivers are trained in using interactive reading techniques to support their children's language and literacy development. During the 2016–2017 school year, parents/caregivers were offered *DR* training in October, February, and May; volunteers and community partners received *DR* training in April. Parents/caregivers also receive *DR* coaching in which a coach observes the parent's reading strategies then engages in a goal-setting conversation on the use of the key *DR* techniques, including those that align with the CAR acronym: 1) **C**omment and wait, 2) **A**sk questions and wait, and 3) **R**espond by adding a little more.

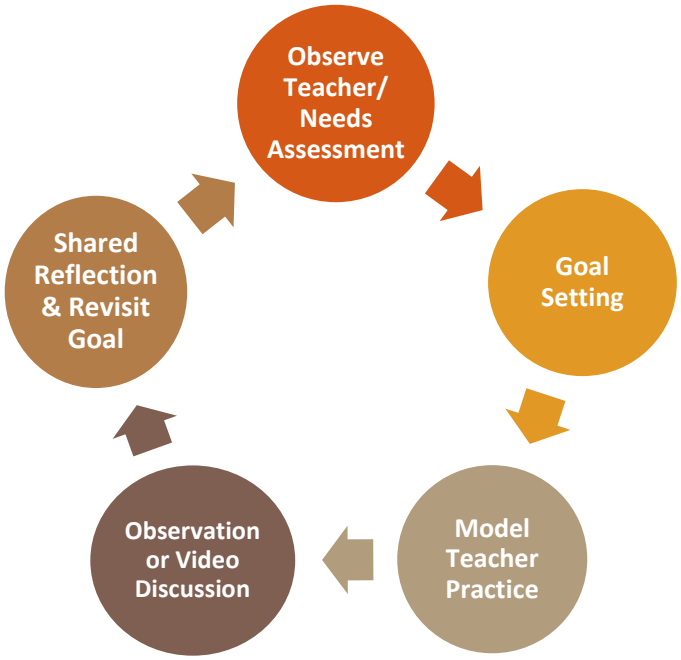
***Cradling Literacy* (*CL*) training.** *CL* training includes 12 modules related to supporting early language and literacy development (see "Curriculum Topics" box) and has been delivered to intervention group teachers since Year 1 of this study (the 2012–2013 school year). Because many teachers were trained in Year 1, delivery was adapted over time to offer pathways for application, reflection, and professional feedback around the content delivered during the initial trainings. In Year 2, trainings were delivered in a "booster session" format, which was structured to support depth of application and praxis around each of the 12 modules. During Year 3, modules were delivered through a peer learning community and peer-to-peer training format, and during Year 4, modules were delivered through reflective practice groups, which involved teams of teachers engaging in activities to help them reflect on the content in a new way. During Year 5, *CL* training was provided for new teachers only, due both to budget cuts and to feedback that previously trained teachers had adequate exposure to the content and did not require a refresher course.

Curriculum Topics

- ☐ Stories
- ☐ Meaningful Experiences
- ☐ Culture
- ☐ Careful Observation
- ☐ Relationships
- ☐ Purposeful Interactions
- ☐ Development
- ☐ Supportive Environments
- ☐ Home Language
- ☐ Behavioral Understanding
- ☐ Family Engagement
- ☐ Social-Emotional Competence

Coaching. Coaching for teachers was provided monthly in groups typically composed of two teachers and one coach. Coaching sessions were guided by the Coaching Cycle, with adaptations made according to teacher needs. In each coaching session, the coach and teachers identified topics to address; goals and activities in progress, completed, or not started; and next steps for both the teacher and coach. Figure 1 displays the coaching components.

Figure 1. *Cradling Literacy* coaching cycle



Program Beneficiaries

The beneficiaries of this program are teachers, children ages 1–30 months, and parents/caregivers at the six ECE sites in metro Denver. Beneficiaries also include community agencies, businesses, and ECE staff members.

Program Outputs

Table 1 depicts the program outputs tracked for the center-based *CL* intervention.

Table 1. *Cradling Literacy* program components and outputs

Program Component	Output
***Cradling Literacy* training and coaching**	Number of teachers receiving training and coaching Frequency of training and coaching received Number of coaching sessions received Duration of coaching sessions Topics in coaching sessions
***Dialogic Reading* training**	Number of teachers, parents/caregivers, volunteers, and partners trained Frequency of trainings
Child attendance	Number of days present at ECE Days attended by quarter

Overview of Impact Study

The overarching goal of the *Ready to Read* program in centers is to increase school readiness among children at high risk for poor educational outcomes due to socioeconomic and other challenges. The program aims to:

- Increase the frequency with which teachers and parents/caregivers engage in language-rich, complex interactions with children, and
- Increase children's oral language and communication skills as a result of more frequent and language-rich interactions with key adults.

This study employed a quasi-experimental design to assess whether *DR* plus *CL* in center-based programming is more effective than *DR* alone. Three centers were randomly assigned to the intervention condition (*CL*), and three centers were randomly assigned to the comparison condition (*DR* only). Study participants included children and their parent or caregiver who were attending the centers in separate neighborhoods of Denver at the beginning of the study or who enrolled in the centers during the course of the study. All families with a child aged 1–30 months were recruited into the study from January 2013 to December 2016. Of the 817 eligible families, 333 participated in this study (186 from intervention centers, 147 from comparison centers).

Targeted Level of Evidence

The evaluation targeted a moderate level of evidence. Although previous research on *DR* demonstrated that this approach is effective for improving children's language and literacy skills, some research has suggested that the degree of implementation of *DR* influences the amount of impact it has. The proposed study tested the added value of *CL*, which gives early childhood educators a theoretical foundation and culturally responsive framework for developing children's oral language and communication skills. It was expected that *DR* and *CL* would complement each other; however, the current level of evidence for *CL* is preliminary, and implementing a professional development intervention is costly. Therefore, it was of both practical and theoretical significance to learn if *CL* has a measurable benefit above and beyond *DR*, which can be implemented rather inexpensively. This study makes a contribution to the literature on this topic by testing the hypothesis that teachers will do a better job of implementing *DR* if they have a theoretical foundation for its use. The *CL* training provides this theoretical foundation. A moderate level of evidence was an appropriate level to target due to the preliminary evidence for the effectiveness of *CL* and the lack of research examining the combined effect of *CL* and *DR*.

Program Implementation and Impact Research Questions

Over the course of the five-year study, we sought to address 16 impact and implementation research questions. Data for the impact questions were collected from classroom observations, parents/caregivers, and teachers, while data for implementation evaluation were collected from program data, observations, and teachers (see also Table 2 in "Study Approach and Methods"). The research questions addressed in this study include:

Impact (Confirmatory)

1. Do teachers in the *CL* group demonstrate greater short-term improvements in their support of children's language and literacy development than do teachers in the *DR*-only group?
2. Do teachers in the *CL* group demonstrate greater long-term improvements in their support of children's language and literacy development than do teachers in the *DR*-only group?
3. Do parents/caregivers in the *CL* group show greater gains in their support of their child's language and literacy development than do the parents/caregivers in the *DR*-only group?
4. Do children in the *CL* group demonstrate greater increases in their oral language and communication skills than do children in the *DR*-only group?

Impact (Exploratory)

5. Do parents/caregivers whose children attend a center more frequently show greater gains in their support of their child's language and literacy development than do those with a lower level of attendance?
6. Do children with higher levels of attendance in a center demonstrate greater gains in their oral language and communication skills than do those with lower levels of attendance?

Implementation

7. How many individuals were trained on *DR*?
8. How many modules of *CL* training did center-based staff receive?
9. Did center-based teachers demonstrate increased knowledge after attending the *CL* training sessions?
10. To what extent did teachers in the *DR*-only group demonstrate knowledge of the concepts covered in the *CL* curriculum?
11. Did center-based teachers find the *CL* training to be useful?
12. What are the characteristics of the *CL* intervention as it is being implemented?
13. What is the consistency of *CL* program delivery?
14. What is the quality of *CL* program delivery?
15. What are teachers' and coaches' perceptions of the impacts of *CL* training and coaching?
16. What are teachers' and coaches' views of the content and delivery of *CL*? What aspects were most/least successful and helpful to them?

Changes to Subgrantee Evaluation Plan

Over the five-year course of this study, there have been several changes from what was proposed in the Subgrantee Evaluation Plan (SEP), particularly related to the evaluation approach. Many factors contributed to the necessity of these changes, including turnover of evaluation teams (twice) during the five-year study as well as decreases in available funding for both program and evaluation efforts. Changes to data collection or evaluation approaches are noted in the "Study Approach and Methods" section of the report. Key program changes include the previously noted changes to training delivery over the course of the study (full training in Y1, booster sessions in Y2, peer learning and training in Y3, reflective practice groups in Y4, and training for new teachers only in Y5) and the addition of coaching for Mile High Early Learning staff in Y4.

STUDY APPROACH AND METHODS

In the *Ready to Read CL* study, the program model involves *CL* training and coaching provided to teachers of children ages 0–3. This study was conducted in Denver, Colorado, from 2012 to 2017 with data collected from two ECE programs that operate six early learning centers. The evaluation plan was developed collaboratively by Clayton Early Learning, Mile High Early Learning, and evaluators from the Buechner Institute for Governance at the University of Colorado, Denver (previously the lead evaluators for this study). Multiple data collection strategies were employed to address the implementation and impact research questions.

Implementation Study Design

CL focuses on teacher training and coaching; therefore, key implementation data include the frequency, duration, and content of *CL* training and coaching; these data, along with information about *DR* training and coaching, were obtained from program records. Estimates of attendance from program records were used to examine families' exposure to teachers who have been trained in *CL*. During years 1–4 of the study, teachers completed a *CL* training satisfaction survey after each training (administered by coaches; 244 surveys completed) and a knowledge survey at baseline (administered by evaluators in fall 2012, or upon starting work at the centers for new teachers; $n = 77$) as well as semi-annually.

Implementation markers included: the number of days children attended *Ready to Read* centers; the number of hours teachers received *CL* training and coaching; the number of *CL* training modules delivered, and the number of families, teachers, volunteers, and community members trained in *DR*. During the 2015–2016 school year, evaluators also observed a *CL* training to capture information about the content, consistency, and quality of these sessions. In the final year of the study, we also collected qualitative data from *CL* coaches (via phone interview) and teachers (via focus group) who participated in *CL* to obtain in-depth information about their views of the content, delivery, and impact of *CL*. Table 2 lists the data collected for each implementation research question.

Table 2. Data Collection by Implementation Research Question

Research Question	Implementation Data	Source
How many individuals were trained on DR*?*	Number trained	Program records
How many modules of CL *training did staff receive?*	Number of modules	Program records
Did teachers demonstrate increased knowledge after attending the CL *training sessions?*	Total score	Teacher knowledge survey
To what extent did teachers in the DR*-only group demonstrate knowledge of the concepts covered in the* CL *curriculum?*	Total score	Teacher knowledge survey
Did teachers find the CL *training useful?*	Teacher-reported satisfaction	Evaluation forms
What are the characteristics of CL *as it is implemented?*	Qualitative themes	Observation
What is the consistency of CL *program delivery?*	Qualitative themes	Observation
What is the quality of CL *program delivery?*	Qualitative themes	Observation
What are teachers' and coaches' perceptions of the impacts of CL *training and coaching?*	Qualitative themes	Interviews and focus groups
What are teachers' and coaches' views of the content and delivery of CL*? What aspects were most/least successful and helpful to them?*	Qualitative themes	Interviews and focus groups

Impact Study Design

To address the impact research questions that explore teacher, parent, and child language and literacy development outcomes, a quasi-experimental longitudinal design was employed to explore differences in outcomes between treatment (*CL*) and comparison families (*DR* only). A key strength of this design is the availability of a comparison group, which provides the potential to reach a moderate level of evidence. The primary limitation is the threat to validity caused by assigning center to treatment condition, which may result in families in intervention centers having different characteristics than those in comparison centers because families are likely to attend the center that is nearest where they live and each neighborhood tends to have unique characteristics.

Sampling, Measures, and Data Collection

Sampling

Child inclusion/exclusion criteria in study. Parents/caregivers of all children enrolled in one of the center-based programs between the age of 1 month and 2.5 years at the start of *Ready to Read* or at program entry to a classroom were recruited for participation. Families who did not enroll in the study within 30 days of program enrollment (plus a one-week grace period) were no longer eligible. Thus, families were only excluded based on age or delay in consenting to join the study.

Table 3 shows families' program and study enrollment rates by school year and condition in the study (intervention and comparison). Since the start of *Ready to Read* in September 2012, a total of 882 children ages 0–30 months enrolled in a center, with roughly equal proportions enrolled in intervention and comparison-assigned centers. Please note that program enrollments for Year 1 look higher than enrollments for Years 2–5 because Year 1 included already-enrolled families, and Years 2–5 only capture the number of new families that enrolled that year. Additionally, the 2016–2017 school year includes children who enrolled only through May 2017 (versus other years with enrollments through August).

Table 3. Enrollments in center-based *Ready to Read* by school year and study condition

	2012–13 school year	2013–14 school year	2014–15 school year	2015–16 school year	2016–17 school year[3,4]	Total
Enrolled in a center[1]	314	197	132	137	102 (37)	882 (817)
Intervention sites	179	104	73	74	49 (17)	479 (447)
Comparison sites	135	93	59	63	53 (20)	403 (370)
Enrolled in the study[2]	131	36	75	75	16	333
Intervention	78	23	40	37	8	186
Comparison	53	13	35	38	8	147
Study enrollment rate (% of eligible)	42%	18%	57%	55%	43%	41%
Intervention (%)	44%	22%	55%	50%	47%	42%
Comparison (%)	39%	14%	59%	60%	40%	40%

[1]Includes only children eligible for the study; children older than 30 months at baseline (n = 91) are not included
[2]Excludes children who enrolled in the study but did not complete a baseline assessment
[3]Numbers for the 2016–17 school year are from September 2016–May 2017; other program years are from September–August
[4]During the 2016–17 school year, study recruitment ended in November 2016; this table reports all program enrollments, with the number eligible for the study in parentheses, as these were used to compute study enrollment rates

Since September 2012, parents/caregivers of 333 children enrolled in the research study, representing 41% of those eligible. Study enrollment rates vary by year: during the first school year, enrollments were slightly below 50%. In

2013–2014, rates were much lower, around 20%, which is primarily due to study recruitment being placed on hold for about four months while study protocols were revised. After recruitment resumed at the start of the 2014–2015 school year, study enrollment rates increased to about 50–60% in Years 3 and 4, and were between 40–50% in Year 5. By study condition, intervention centers included 186 study participants, while comparison sites had 147 study participants.

Measures and Instruments

Teacher-level outcomes. The Classroom Assessment Scoring System, infant and toddler version (CLASS I/T; LoCasale-Crouch, Hamre, La Paro, & Pianta, 2011; La Paro, Hamre, & Pianta, 2012), teacher knowledge survey and training satisfaction survey were used to assess teacher-level outcomes.

CLASS I/T. This assessment is conducted by trained observers and involves 20-minute cycles of classroom observation (typically at least 4 cycles), each followed by a 10-minute period during which the observer reviews his/her notes and assigns scores. Observers rate several "dimensions" of the classroom environment on a scale ranging from 1–7. These scores are then aggregated into two "domains," *Emotional and Behavioral Support* and *Engaged Support for Learning*, that are used in the analyses for this study.

Teacher knowledge. The teacher survey included 37 multiple choice and true-false items designed to assess teachers' growth in knowledge of early language and literacy, adapted from the Supporting Language and Literacy Development in the Classroom assessment (Neuman & Cunningham, 2009). The tool emphasized understanding of child development principles and grounded "real-life" activities in the classroom that specifically support language and literacy.

CL satisfaction. After each *CL* training, teachers rated how well the modules were addressed and how useful the training activities and materials were.

Child-level outcomes. The communication subscale of the Ages and Stages Questionnaire (ASQ-C; Squires et al., 2009), the language production measure of the Mac Arthur-Bates Communicative Development Inventory (CDI; Fenson et al., 2007), and the language and literacy dimensions of Teaching Strategies GOLD (TSG; Heroman, Burts, Berke & Bickart, 2010) were used to assess children's language and literacy development.

ASQ-C. This is a 6-item questionnaire for which parents/caregivers rate the child's development in various age-related skills on a 3-point scale (*yes* = 10, *sometimes* = 5, or *not yet* = 0). Scores correspond with a clinical cutoff to indicate children whose development is on schedule. A score that is close to the cutoff suggests that learning activities and monitoring should be provided. Scores below the cutoff indicate that further developmental assessment is needed. We recoded scores into a dichotomous variable—*on schedule* or *close to/below the cutoff*—and also computed a score indicating whether a child's score: 1) remained the same over time (or decreased), or 2) increased. (Only 10 children's scores decreased over time; therefore, they were collapsed into the "remained the same over time" category.) The authors of the measure report adequate internal consistency estimates (between .82 and .88) and high test-retest reliability (.91) and inter-observer reliability (.92) estimates.

CDI. The parent or caregiver was asked to indicate, on a list of 89 to 100 words (depending on the age of the child), the number of words the child understands and says. The CDI is only administered to children ages 8 months and older. The CDI was normed on more than 1,800 children in three locations, and numerous studies have documented the reliability and validity of these measures.

TSG. TSG is a system of authentic, observation-based assessments for children ages birth through kindergarten. TSG is used in early education settings throughout Colorado to assess children's learning, abilities, strengths, needs, and interests. Children are assessed by teachers at 3-month checkpoints throughout the year. Analyses use language and literacy scaled scores (standardized to a mean of 500 and standard deviation of 50).

Parent-level outcomes. This study used two measures to assess the impact of the intervention on parents'/caregivers' support for their children's language and literacy development.

Parent Survey of Home Literacy (Smith and Dixon, 1995). Parents/caregivers reported (*yes/no*) whether they frequently engage in a series of seven interactive reading techniques. Behaviors include pointing out pictures and letters, asking the child what will happen next, re-reading a story, and encouraging the child to read along when the book uses repeated phrases.

Parent frequency of book reading and storytelling. The frequency with which parents/caregivers read books and tell stories with their child was measured using two items from the Early Head Start Family and Child Experiences Study (Baby FACES; Mathematica Policy Research, n.d.). Parents/caregivers rated these items on a 5-point scale ranging from: *less than once per week* = 1, *once per week* = 2, *a few times per week* = 3, *about once per day* = 4, and *more than once per day* = 5.

Data Collection Activities

Data collection procedures changed during the course of the study. For an overview of these changes, see "Changes to the Subgrantee Evaluation Plan" later in this section. Generally, after a child enrolled in a center, an evaluator called the parent/guardian to explain the study, obtain verbal consent, and let the parent know that the study materials, including a consent form, survey, and self-addressed stamped envelope, would be in the child's "cubby" or left with the classroom teacher. Parents/caregivers either returned the envelope with materials via mail or returned them to the child's cubby/teacher. The evaluator visited each center twice a week to collect surveys. Parents/caregivers who did not respond to the consent and survey received a reminder call the following week and a second reminder call or note and copy of new surveys if necessary the final week the family was eligible to enroll. Teachers assisted by reminding parents/caregivers when there were materials for them. Program staff also assisted by providing updated contact information for parents/caregivers. Child-level TSG assessments were secondary data obtained from center staff.

The study research questions and associated measures and timing are summarized in Table 4.

Table 4. *Ready to Read* center-based study research questions

Research Question	Type of Question	Measures	Timing
1. Do teachers in the *CL* group demonstrate greater *short-term* improvements in their support of children's language and literacy development than do teachers in the *DR*-only group?	Impact: Confirmatory	Classroom Assessment Scoring System, infant and toddler versions (CLASS I/T)	Baseline, semi-annual, or annual follow-up based on availability of data
2. Do teachers in the *CL* group demonstrate greater *long-term* improvements in their support of children's language and literacy development than do teachers in the *DR*-only group?	Impact: Confirmatory	CLASS I/T	Baseline, semi-annual, or annual follow-up based on availability of data
3. Do parents/caregivers in the *CL* group show greater gains in their support of their child's language and literacy development than do the parents/caregivers in the *DR*-only group?	Impact: Confirmatory	Parent Survey of Home Literacy; Baby FACES, Reading Books and Telling Stories	Baseline, 6-month birthday follow-up (enrollments before fall 2014); Baseline, 3- & 6-month follow-up (enrollments since fall 2014).
4. Do children in the *CL* group demonstrate greater increases in their oral language and communication skills than children in the *DR*-only group?	Impact: Confirmatory	Ages and Stages Questionnaires, Communication subscale (ASQ-C); MacArthur-	Baseline, 6-month birthday follow-up (enrollments before fall 2014); Baseline, 3- & 6-

Research Question	Type of Question	Measures	Timing
		Bates Communicative Development Inventories (CDI)	month follow-up (enrollments since fall 2014).
		Teaching Strategies™ GOLD® (TSG): Language and Literacy dimensions	Checkpoints in fall, winter, and spring
5. Do parents/caregivers whose children attend a center more often show greater gains in their support of their child's language and literacy development than do those with a lower level of attendance?	Impact: Exploratory	Parent Survey of Home Literacy; Baby FACES, Reading Books and Telling Stories	Baseline, 6-month birthday follow-up (enrollments before fall 2014); Baseline, 3- & 6-month follow-up (enrollments since fall 2014).
6. Do children with higher levels of attendance in a center demonstrate greater gains in their oral language and communication skills than do those with lower levels of attendance?	Impact: Exploratory	ASQ-C; CDI	Baseline, 6-month birthday follow-up (enrollments before fall 2014); Baseline, 3- & 6-month follow-up (enrollments since Fall 2014).
		TSG Language and Literacy dimensions	Checkpoints in fall, winter, and spring
7. How many individuals were trained on *DR*?	Implementation	Program records	As occurs
8. How many modules of *CL* training did center-based staff receive?	Implementation	Program records	As occurs
9. Did center-based teachers demonstrate increased knowledge after attending the *CL* training sessions?	Implementation	*CL* knowledge survey	Baseline and semi-annual follow-up (discontinued after fall 2016)
10. To what extent did teachers in the *DR*-only group demonstrate knowledge of the concepts covered in the *CL* curriculum?	Implementation	*CL* knowledge survey	Baseline and semi-annual follow-up (discontinued after fall 2016)
11. Did center-based teachers find the *CL* training to be useful?	Implementation	Training evaluation forms	As occurs
12. What are the characteristics of the *CL* intervention as it is being implemented?	Implementation	Qualitative observation of *CL* training	As occurs
13. What is the consistency of *CL* program delivery?	Implementation	Qualitative observation of *CL* training	2015–16 school year
14. What is the quality of *CL* program delivery?	Implementation	Qualitative observation of *CL* training	2015–16 school year
15. What are teachers' and coaches' perceptions of the impacts of *CL* training and coaching?	Implementation	Focus groups and interviews	Spring 2017
16. What are teachers' and coaches' views of the content and delivery of *CL*? What aspects were most/least successful and helpful to them?	Implementation	Focus groups and interviews	Spring 2017

Sample sizes by measure are provided in Table 5. As shown, the center-based sample includes 333 children with baseline data. TSG and parent survey measures provided the largest samples of families with complete baseline and follow-up data. Other measures (ASQ, CDI) had lower rates of completion because they must be completed within a specific age range. Finally, some families left the center before a follow-up could be completed. Missing data and study attrition are discussed later in this section.

Table 5. Sample sizes for parent- and child-level data over time[1]

	Baselines	*Baseline with 3-month follow-up*	*Baseline with 6-month follow-up*
Any measure	333	212	199
Parent Survey	308	126	154
ASQ[2]	279	107	124
CDI[2]	208	73	88
TSG	229	164	133

[1] LENA and video data were discontinued for new participants in 2014 and therefore are not included.
[2]ASQ and CDI data do not include assessments completed if the child was outside the allowable age range.

Teachers

Table 6 shows the demographic characteristics of the ECE teachers who enrolled in the *Ready to Read* study since September 2012. Almost all teachers are female (98%). Fewer than half of the teachers identified as white (43%), 24% as another race/ethnicity, 17% as Latina/o, and 16% as Black or African American. The overall sample included teachers of varying ages, levels of education, and tenure in ECE.

Table 6. Center-based teacher demographic characteristics (*n* = 86)[1]

Characteristic	***N***	**%**
Gender: Female	81	98%
Male	2	2%
Race/Ethnicity: Black / African American	13	16%
Latina/o	14	17%
Other	18	24%
White/Caucasian	37	43%
Age: 18–24	7	9%
25–34	34	42%
35–44	15	19%
45+	24	30%
Education: High School Diploma or GED	10	13%
Child Development Associate (CDA)	12	17%
Associate's Degree	12	16%
Bachelor's Degree	26	35%
Master's Degree	7	9%
Other	8	10%
ECE tenure: < 5 years	18	23%
5–14 years	41	53%
15 years or more	19	24%

[1]This number does not include two teachers who enrolled in the study but did not complete a baseline assessment and four teachers who enrolled in the study after the survey was discontinued.

Families

Table 7 shows the overall demographic characteristics of children in the study. There were roughly equal proportions of male (52%) and female children (48%), and most identified as either white (45%) or Black / African American (34%). Many were eligible for free or reduced lunch (F/RL; 89%). Most children were English-speaking and did not have an Individual Family Service Plan (IFSP)—a plan that indicates early intervention services related to developmental disability or delay.

Table 7. Center-based child demographic characteristics (*n* = 333)

	N	%
Gender: Female	159	48%
Male	172	52%
Race: American Indian or Alaska Native	7	2%
Asian / Pacific Islander	2	1%
Black / African American	109	34%
Multiracial	58	18%
White/Caucasian	145	45%
Ethnicity: Hispanic or Latino	106	32%
Not Hispanic or Latino	226	68%
Free/Reduced Lunch: Eligible	134	89%
Not eligible	16	11%
Language: English	303	92%
Other	28	8%
Individual Family Service Plan: Yes	12	4%
No	289	96%

Baseline Equivalence Analysis

Random assignment. Randomization occurred at the center level rather than at the individual level to minimize the possibility that comparison families would be exposed to the effects of the *CL* intervention via interaction with other teachers in their center (contamination). To be sure each organization had at least one intervention site and one comparison site, the names of Clayton Early Learning's two centers were written on individual pieces of paper, and the original external evaluator drew one to be assigned to the intervention condition. Likewise, the names of Mile High Early Learning's four centers were written on individual pieces of paper, and the evaluator selected two for intervention. The others were assigned to the comparison condition.

To assess the equivalence of intervention and comparison groups at baseline on demographic characteristics and study measures, we conducted chi-square tests and *t*-tests. Because sample sizes for families vary by measure (i.e., participants have complete data for some measures but not for others), analyses were done separately for each measure, as well as for the overall sample.

Teacher characteristics at baseline. In Years 1–4 of the study, we collected teachers' demographic characteristics as part of a survey of teachers' knowledge and beliefs. For this sample ($n = 86$), there were no statistically significant differences between intervention and comparison teachers on the basis of teacher age, sex, race, ethnicity, level of education, or years worked in early childhood education. In addition, the groups did not differ significantly at baseline on measures of early literacy knowledge or views of teacher efficacy. Therefore, based on available data, it appears that teachers in the intervention and comparison groups were fairly similar.

Child and family characteristics at baseline. Characteristics for which statistically significant differences ($p \leq .05$) were found are summarized in Table 8 for the overall sample and for families with baseline and follow-up data for each

measure. As shown, there were differences between the groups by ethnicity for the overall sample and for families with baseline + 6-month follow-up TSG data, with the comparison group having a larger proportion of Hispanic or Latino children than did the intervention group (38% versus 27%, respectively for the overall sample). Notably, however, groups did not differ on the basis of child's language, which might have a more direct influence on language and literacy outcomes than would ethnicity. Furthermore, groups did not differ significantly on measures related to children's skills and parent support of language and literacy.

Table 8. Baseline differences between children in the intervention and comparison groups

Sample	Demographic Characteristics							Child-Level Measures			Parent-Level Measures		
	Age	*Gender*	*Race*	*Ethnicity*	*Language*	*F/RL*	*IFSP*	*ASQ-C*	*CDI*	*TSG*	*Home Lit.*	*Reading*	*Storytelling*
Overall sample (N = 333)[1]				x									
Survey: baseline + 6-month follow-up (n = 154)													
ASQ-C: baseline + 6-month follow-up (n = 124)													
CDI: baseline + 6-month follow-up (n = 88)													
TSG[2]: baseline + 6-month follow-up (n = 133)				x									

An "X" means . . .
Among children in this particular subsample (see rows, left), there was a **significant difference between intervention and comparison groups** on a demographic characteristic or outcome measure (see columns).

[1] $\chi^2(1, N = 332) = 3.95$, $p = .05$; [2] $\chi^2(1, N = 133) = 4.12$, $p = .04$

In summary, although the intervention and comparison groups are similar in many respects, they differ on the basis of child's ethnicity, which could influence the extent to which the intervention impacts parent and child outcomes and the ability to interpret findings.

Differential Attrition Analysis

Study attrition generally did not vary according to children's demographic characteristics, with two exceptions: 1) among the sample of children with ASQ-C data, those who were white were more likely to complete a 6-month follow-up compared to children who identified as races other than white, $\chi^2(1, N = 268) = 4.46$, $p = .04$; and 2) among children with survey data, those who were not eligible for F/RL were more likely to have follow-up data than those who were eligible for F/RL, $\chi^2(1, N = 145) = 4.94$, $p = .03$. Thus, low-income children and those identified as a race other than white may be underrepresented in the sample due to attrition. There were no statistically significant differences in study attrition by treatment condition, indicating that intervention and comparison group families were equally likely to have complete data.

Changes to Subgrantee Evaluation Plan

Several changes to data collection and to analytic strategy were made over the course of the study, including:

Implementation evaluation (started/ended in Year 4). In Year 4 (2015–2016), we received supplemental funding to examine the content, quality, and consistency of the *CL* intervention. Unfortunately, we were only able to conduct one observation before overall funding for the study was cut, necessitating a reallocation of the implementation evaluation funds in order to continue conducting the rest of the study. Results are not included in this report but were included in the Year 4 report, *Ready to Read: Cradling Literacy Study, Social Innovation Fund Year 4 Annual Report*.

Parent and teacher *DR* observations. These observations were discontinued after Years 2 and 3, respectively. During the first two years of the study, parents/caregivers were asked to participate in a video-recorded reading observation while picking up or dropping off their child; however, parents/caregivers often did not have time to do the observation,

and it was challenging (and expensive) for staff to meet families at the centers for these observations. Therefore, after Year 2 of the study, we discontinued the *DR* observations and relied upon the "Parent Survey of Home Literacy" to measure parents'/caregivers' use of interactive reading techniques. Teacher-level *DR* observations were discontinued after Year 3 due to funding cuts. For previous years' *DR* observation results, see the *Ready to Read Cradling Literacy Study: Social Innovation Fund Year 3 Annual Report.*

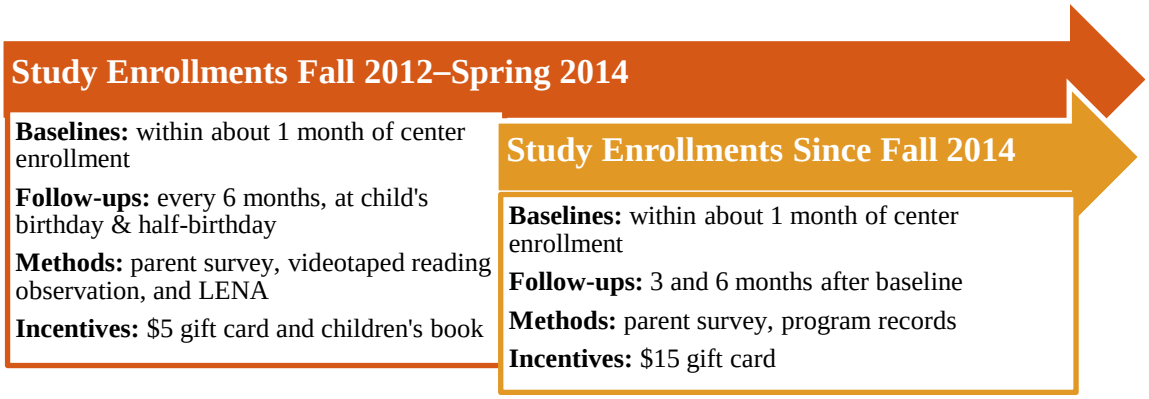
Language Environment Analysis system (LENA). LENA is an audio-recording device that children wear for 10 to 16 hours to record the home literacy environment. Facilitators then present the LENA results to parents/caregivers and discuss ways to increase meaningful adult-child interactions. After Year 2, we only collected LENA data for a subsample of families who had already completed a LENA assessment. These data are not used in outcomes analysis due to the small sample size, but descriptive results of new data collected during Year 4 were reported in *Ready to Read Cradling Literacy Study: Social Innovation Fund Year 4 Annual Report.*

Teacher survey. Administration of the teacher survey was discontinued for Year 5 because according to previous years' results, this measure did not differentiate between intervention and comparison group teachers in terms of teachers' knowledge and beliefs. Instead, we conducted teacher focus groups and interviews with *CL* coaches to gather more in-depth information about staff's views of the implementation and impact of *CL*. For previous years' *DR* observation results, see the *Ready to Read Cradling Literacy Study: Social Innovation Fund Year 4 Annual Report.*

Teaching Strategies™ GOLD® (TSG). To address a problem of missing child-level data encountered in the first two years of the study, we obtained permission to access child-level language and literacy development data that were routinely being collected by the early learning centers using the TSG assessment tool.

Data collection. Figure 2 provides an overview of changes to family data collection procedures. In fall 2014, we shortened the time frame for follow-ups, discontinued *DR* observations and LENA, added TSG data, and increased study incentives. These changes were approved by the University of Denver's Institutional Review Board (IRB).

Figure 2. Family data collection procedures, fall 2012–present



Infeasibility of Propensity Score Matching. Although Propensity Score Matching (PSM) is an increasingly popular method for facilitating treatment and control group comparisons, certain limitations to this study prevented us from being able to conduct PSM as proposed in the SEP. Critical to successful calculation of propensity scores is using a set of characteristics that can predict the probability of being assigned to the treatment group (Rosenbaum & Rubin, 1983). There also needs to be enough overlap of members of the treatment and control groups on those selected characteristics to permit creating matches between treatment group individuals and similar nontreatment individuals. Successful matching also requires a relatively large sample (typically at least 200), with a particularly large pool of control group members to increase likelihood of finding a good match for each member of the treatment group.

Key factors that contributed to our inability to use PSM included: 1) a smaller-than-anticipated sample size (in the SEP, the study enrollment estimate was 90% of available children, but the reality was 41%), and 2) the lack of a sufficient control group from which to create matches (a ratio of at least 2:1 control to treatment is recommended). In addition, we lacked a strong set of demographic data for creating propensity scores and would not have adequate overlap between groups as those characteristics tended to be confounded by site (see Steiner, Cook, Shadish, & Clark, 2010, for a discussion of the importance of covariates). In summary, the combination of a smaller-than-anticipated sample size, the lack of a sufficient control group, and theoretically inadequate covariates precluded us from using PSM. This inability to establish baseline equivalence interfered with the ability to isolate the effects of the *CL* intervention.

Did not use Maximum Likelihood estimation for handling missing data. In the SEP, the original evaluator proposed using MPlus software to conduct Maximum Likelihood (ML) estimation to impute missing data; however, it was not appropriate to use ML due to the nature of this study's missing data. ML requires that the assumption of data missing at random (MAR) is tenable and that covariates associated with missingness are included in the estimation model. We did not use ML in this study because: 1) the missing data are not likely MAR—that is when the probability of missing data on Y is unrelated to the value of Y after controlling for other variables in the analysis. In other words, as demonstrated in the differential attrition analyses, it is likely that the missingness in this study is related to other observed variables; 2) several of the measures in this study cannot have item-level missingness that can be imputed; and 3) the delivery of an intervention with unknown impact makes post-test parameter estimates based on imputed data suspect. In other words, we can't know what post-intervention values would be, and there are insufficient available covariates to serve as auxiliary variables for conducting imputation using Full Information Maximum Likelihood. Furthermore, although the original budget included $600 for MPlus, the current evaluation team did not receive the software or funding necessary to purchase MPlus.

ANALYSIS AND RESULTS

In general, statistical analyses were conducted using the more conservative intent to treat approach with teacher, child, or parent data as the unit of analysis and intervention/control as the grouping variable. Analytic techniques included repeated measures analysis of variance (ANOVA), descriptive statistics, and bivariate analyses (see Table 9). Power analysis, missing data, and attrition were also examined.

Implementation Evaluation Analysis

CL and *DR* training and coaching records, as well as child attendance in *Ready to Read* classrooms, were analyzed descriptively, using counts, percentages, and averages. For teacher focus groups and interviews with coaches, evaluators reviewed qualitative data to identify themes related to the implementation of *CL*.

Statistical Analysis of Impacts

Continuous outcome measures were analyzed using repeated measures analysis of variance (ANOVA), which examined differences between groups and how they changed over time. For categorical outcomes, chi-square tests were used. Parent- and child-level analyses were first conducted using study condition (intervention or comparison) as the independent variable; next, analyses were re-run using dosage (child attendance) as the independent variable. For the analysis of CLASS I/T ratings, we used a factorial ANOVA with study condition and time point as factors. Due to teacher turnover during the course of the study, we could not treat CLASS I/T ratings over time as true "repeated measures" (as though it measured the same individuals). Instead, we examined overall differences in mean ratings by study condition and time point.

Table 9. Statistical approach by research question

Research Question	Type of Question	Analytic Strategy	Unit of Analysis
Do teachers in the *CL* group demonstrate greater *short-term* improvements in their support of children's language and literacy development than do teachers in the *DR*-only group?	Impact: Confirmatory	Descriptives *ANOVA*	Teachers
Do teachers in the *CL* group demonstrate greater *long-term* improvements in their support of children's language and literacy development than do teachers in the *DR*-only group?	Impact: Confirmatory	Descriptives *ANOVA*	Teachers
Do parents/caregivers in the *CL* group show greater gains in their support of their child's language and literacy development than do the parents/caregivers in the *DR*-only group?	Impact: Confirmatory	Descriptives *ANOVA* *t*-tests	Parents/caregivers
Do children in the *CL* group demonstrate greater increases in their oral language and communication skills than do children in the *DR*-only group?	Impact: Confirmatory	Descriptives *ANOVA* Chi-square	Children
Do parents/caregivers whose children attend a center more often show greater gains in their support of their child's language and literacy development than do those with a lower level of attendance?	Impact: Exploratory	Descriptives *ANOVA* *t*-tests	Parents/caregivers
Do children with higher levels of attendance in a center demonstrate greater gains in their oral language and communication skills than do those with lower levels of attendance?	Impact: Exploratory	Descriptives *ANOVA* Chi-square	Children

Research Question	Type of Question	Analytic Strategy	Unit of Analysis
How many individuals were trained on *DR*?	Implementation	Descriptives	Staff, parents/caregivers, volunteers, and community partners
How many modules of *CL* training did center-based staff receive?	Implementation	Descriptives	Center-based staff
Did center-based teachers demonstrate increased knowledge after attending the *CL* training sessions?	Implementation	Discontinued	Teachers
To what extent did teachers in the *DR*-only group demonstrate knowledge of the concepts covered in the *CL* curriculum?	Implementation	Discontinued	Teachers
What are teachers' and coaches' perceptions of the impacts of *CL* training and coaching?	Implementation	Qualitative theme analysis	Teachers and coaches
What are teachers' and coaches' views of the content and delivery of *CL*? What aspects were most/least successful and helpful to them?	Implementation	Qualitative theme analysis	Teachers and coaches

Power Analysis

Total sample sizes for parent and child impact measures ranged from approximately 75 to 150. According to post hoc power analyses conducted in G*Power, there is adequate power to detect a large- or medium-sized effect in repeated measures ANOVA. There is a lower probability of finding small effects (see Table 10). Power estimates for paired samples *t*-tests are lower still, ranging from .52 (for $N = 75$ and a small effect) to 1.00 (for $N = 150$ and a large effect).

Table 10. Post hoc power analysis for repeated measures ANOVA

N	Time Points	Groups	Power to Detect Large Effect (.40)	Power to Detect Medium Effect (.25)	Power to Detect Small Effect (.15)
75	2	2	.99	.98	.72
100	2	2	.99	.99	.84
125	2	2	.99	.99	.91
150	2	2	.99	.99	.95

G*Power Post hoc *F* test for ANOVA: Repeated measures, within-between interaction; Groups = 2, Measurements = 2, Correlation among rep measures = .5

Measure- and Item-Specific Missing Data Analysis Findings

Program attrition and noncompletion of study measures were the primary reasons for missing data and have been described previously. Measure- and item-specific missingness is described below.

CLASS I/T. Of 21 classrooms, 19 had complete data for the five time points included in the analyses. Because this is an observational measure, item-level missingness does not occur.

ASQ-C. As briefly described earlier, there are 17 age-specific versions of the ASQ-C for children between the ages of 0 and 3, with children aging into a different ASQ version approximately every one to two months. If the assessment is not completed within the intended age range, it is not valid and is considered missing for the purposes of the study. During the course of this study, 56 assessments were deemed missing. With regard to item-level missingness, ASQ-C

assessment instructions specify that if there are one or two unanswered items, a score is imputed based on the average of the respondent's answers to the other items (Squires & Bricker, 2009).

CDI. There are three age-specific levels of the CDI for children between the ages of 0 and 3 (Level 1: 8–15.99 months; Level: 2: 16–30 months; Level 3: Greater than 30 months). If this assessment is completed outside the intended age range, it is not valid and considered to be missing. A total of 36 assessments collected during the study were out of range. Because the CDI *words produced* measure is obtained by totaling the number of words the parent has indicated that the child understands and says, item-level missingness is not applicable.

TSG. TSG data were obtained from program records. For TSG checkpoints used in these analyses, fewer than 5% of language or literacy outcome measures were missing.

In contrast to primary data that were collected within about one month of the child's enrollment (and again three and six months later), TSG data were collected by teachers at the early learning centers at fall, winter, and spring "checkpoints." Thus, there was some variation in how long after enrolling in the program a child received a TSG assessment. Of the 133 children with TSG data at three consecutive checkpoints, most had their first assessment within 3 months of when they enrolled in the program (n = 79), but some had their first assessment 3–6 months after program enrollment (n = 43), or more than 6 months after enrollment (n = 11). Children with their first TSG checkpoint 3–6 months after program enrollment had similar initial scores and rates of growth over time as those with their first TSG checkpoint closer to the time of enrollment. However, those with a baseline that occurred more than 6 months after program enrollment had higher initial scores and slower rates of growth than did children with their first TSG checkpoint closer to enrollment. Therefore, children whose first checkpoint was more than 6 months after program enrollment were excluded from the analyses because they do not appear to have a true "baseline" score (n = 11).

Parent survey. Most items included in the analyses had very little missing data (< 1% of responses). An exception was the parent survey of home literacy, which is scored by summing responses to seven *yes/no* items. Some participants provided no response to one or more items (n = 16). After examining those cases, it was determined that nonresponse to those items should be treated as a "no," and previously summed scores were retained.

Implementation Findings

To contextualize our impact analyses, we describe here program implementation, including the delivery of *CL* and *DR* training and coaching and child and family exposure to early learning centers, as well as teachers' satisfaction with *CL* training and their knowledge and beliefs about early learning.

Program Exposure and Dosage

Training and coaching. Training and coaching provided during Year 5 is shown in Table 11. Intervention group teachers received *CL* coaching during most months of the 2016–2017 school year before it concluded in March. *CL* training was provided for new teachers in August. Either *DR* training or coaching was provided to all teachers (intervention and comparison) approximately every other month. For parents/caregivers, *DR* training was provided monthly in the fall and winter, while an additional training for volunteers and community partners was provided in March and April.

During the 2016–2017 school year, coaching related to Cradling Literacy and/or Dialogic Reading was offered for intervention teachers in nearly all months.

For results regarding the implementation of *CL* and *DR* training and coaching in Years 1–3, please see the *Ready to Read Cradling Literacy Study: Social Innovation Fund Year 3 Annual Report*, and for Year 4 results, see the *Ready to Read Cradling Literacy Study: Social Innovation Fund Year 4 Annual Report*. The following is a summary of *CL* and *DR* offerings throughout the five-year grant:

- ***CL* training:** This training was conducted most intensively in Years 1 and 2 of the study, when it was provided in most months (except in the summer). In Years 3 and 4, *CL* training was offered in three to four months, and in Year 5, new teachers received training at the start of the school year only.
- ***CL* coaching:** Coaching was offered nearly every month throughout the study.
- ***DR* training:** This training occurred more frequently during Years 1–3 (training was offered to teachers, community members, and parents/caregivers about every other month, except in summer). In Years 4–5, training was generally provided less often (one to three months of the year), with the exception of community members during Year 4 (five months of the year).
- ***DR* coaching:** This coaching was offered to teachers three to four times a year during the study.

Table 11. Training and coaching from August 2016 to May 2017

Participant	2016–2017 School Year									
	Aug.	**Sep.**	**Oct.**	**Nov.**	**Dec.**	**Jan.**	**Feb.**	**Mar.**	**Apr.**	**May**
Teachers (all)										
DR training		X	X							
DR coaching				X		X				X
Teachers (intervention)										
CL training (new teachers)	X									
CL coaching		X	X	X	X	X	X	X		
Families										
DR training		X	X	X		X	X			
Volunteers & community partners										
DR training*								X	X	

*Includes trainings conducted in PLG and center-based *Ready to Read* sites

***Number trained in* DR.** Since the start of the *Ready to Read* study, a total of 1,272 center-based staff, parents/caregivers, and community members have been trained in *DR*, according to program records (Table 12).

Implementation Research Question

How many parents, facilitators, community partners, and volunteers were trained on *Dialogic Reading*?

Table 12. Number of individuals trained in *Dialogic Reading* September 2012–May 2017

Trainee	*N*
Staff	284
Parents/caregivers	376
Volunteers*	252
Community partners*	360
Total	**1,272**

*Includes trainings conducted in PLG and center-based *Ready to Read* sites

Teacher CL training and coaching exposure. In August 2016, seven new teachers received the *CL* training, which covered the 12 *CL* modules. In addition to formal *CL* training sessions, teachers received monthly coaching. During the 2016–2017 school year, teachers participated in an average of 18 hours of coaching (range: 5–40); this is lower than previous years (approximately 30–40 hours of coaching were provided each year during Years 1–4 of the study) and is likely because coaching ended earlier in the school year during the last year of the grant. Each coaching session covered about two *CL* modules on average. Overall, the most frequently discussed modules were: *Relationships* (16% of all coaching sessions), *Supportive Environments* (11%),

Implementation Research Question

How many modules of *Cradling Literacy* training did center-based staff receive?

Culture (11%), and *Meaningful Experiences* (11%); while the modules least frequently discussed were *Careful Observation* (5%), *Purposeful Interactions* (5%), *Development* (5%), and *Home Language* (3%). Coaching content was based on individual teachers' needs, so some variation in the frequency with which modules were addressed is to be expected.

Child attendance. Data on child attendance at the ECE center (the number of days attended per quarter) came from program records. For intervention sites, attendance was used as a proxy for families' *CL* dosage; for comparison sites, attendance was exposure to "programming as usual." Table 13 displays the average number of days children attended intervention and comparison centers per year.

Table 13. Child attendance (in days) by study year[1]

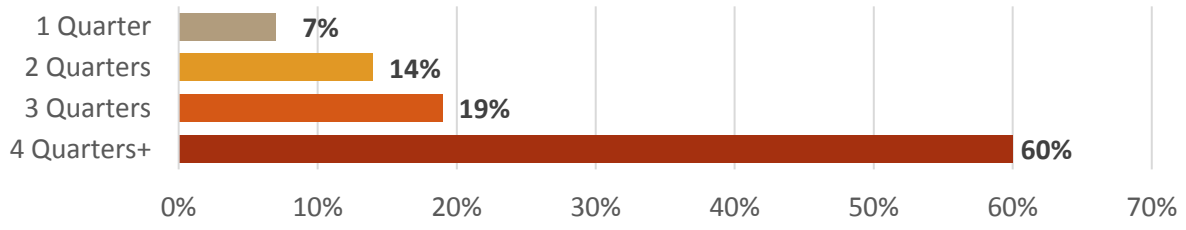
	Intervention Children Average number of days (range)	Comparison Children Average number of days (range)
Child's first year in center (*n* = 327)	99 (2–260) *n* = 183	117 (1–265) *n* = 144
Child's second year in center (*n* = 213)	135 (4–268) *n* = 118	150 (7–263) *n* = 95
Child's third year in center (*n* = 54)	138 (2–256) *n* = 27	143 (12–256) *n* = 27
Child's fourth year in center (*n* = 7)	96 (13–192) *n* = 6	24[2] *n* = 1

[1]Attendance information was not available for six children; [2]No range is available as there is only one child in this category.

On average, children attended about 100 days during their first (or only) year at a center. We also examined attendance by quarter and found that children attended an average of about 45 days per quarter. This suggests a tendency of children to be enrolled in a center for relatively short durations but to attend frequently during those periods.

Child enrollment duration. To better understand families' exposure to center-based programming, we also examined duration of enrollment, defined as the number of quarters in which a child attended at least one day (Figure 3). As shown, more than half of children attended for four or more quarters (60%) while 33% attended for two or three quarters. Seven percent of children attended for only one quarter. Notably, the proportion of children who attended for 4+ quarters (60%) has increased compared to Y4 data (47%). This could be because study enrollment ended in November (and therefore did not include most recent enrollments in the program), but it could also indicate that families were being retained in the program longer than in past years of the study.

Figure 3. Attendance by number of quarters, fall 2012 to spring 2017



As is to be expected, children who attended for a year or longer were younger, on average, when they enrolled, versus those who attended for less than a year (see box, right), which suggests that one factor influencing enrollment duration is children's "aging out" of the program.

Quarters attended during study	Average enrollment age (months)
1 Quarter	17.44
2 Quarters	16.94
3 Quarters	16.68
4 Quarters or more	13.02

Attendance at the time of study assessments. Table 14 includes estimates of child attendance at the time of the 6-month follow-up survey. Please note that this is based on a subsample of children who have complete baseline and 6-month follow-up data. Estimates were obtained by summing attendance data across quarters. Because most assessments were done midway through a quarter, we estimated the proportion of the quarter that had passed and adjusted quarterly attendance totals accordingly. As shown, comparison group participants attended an average of 12 days more than the intervention group; however, this difference between groups was not statistically significant.

Table 14. Attendance (in days) by study condition: Children with baseline and 6-month follow-up survey data

	Intervention	Comparison
Attendance at time of assessment	135.20 (21–331) *n* = 82	147.72 (55–253) *n* = 68

[1]Based on dates of 6-month follow-up parent surveys; [2]Children without a 6-month follow-up survey are not included

Teacher and Coach Views of the Cradling Literacy Model

Through interviews with coaches and focus groups with teachers, we sought to understand their perceptions of the impact of *CL* training and coaching, as well as their views of the utility of the content and delivery.

Implementation Research Question 15: *What are teachers' and coaches' perceptions of the impacts of* CL *training and coaching?*
Implementation Research Question 16: *What are teachers' and coaches' views of the content and delivery of* CL*? What aspects were most/least successful and helpful to them?*

Overall, teachers and coaches perceived the *CL* training to be useful and reported that it brought organization to foundational skills. However, some felt that the booster sessions were repetitive and the delivery of the curriculum was time consuming and challenging to fit in with teaching/personal schedules. Teachers recommended that the content only be delivered once and that the peer-to-peer training was the preferred format for a refresher.

Both teachers and coaches did observe an *impact* on teaching practices, children, and parent behavior and the classroom environment:

- **Teachers**
  - Teachers reported more intentionality and focus on developing children's literacy skills.
  - Teachers integrated literacy into more parts of the day such as nap time, lunch, and on the playground.
- **Children**
  - Teachers saw significant gains in literacy and language development for many children.
- **Parents**
  - Parents became more engaged in literacy development through the *DR* training.
  - Teachers perceived parents to read more to their children at younger ages.
- **Classroom Environment**
  - *CL* funds allowed teachers to enrich their classroom environments with more books and materials to encourage literacy development.

CL Content

Focus group and interview participants generally viewed the content of the *CL* intervention as helpful for providing organization to foundational information learned previously, while incorporating a focus on literacy and language.

- *"It's all pretty foundational. And also concepts that are really, really important I mean are foundational. So, basic, and yet really important concepts."*—Coach
- *"I think it's made us all stronger professionals. This is all we know about and internalize and value every part of this curriculum before introducing Cradling Literacy, but it takes it and builds on it."*—Teacher
- *"I feel like it was useful in that it brought just more definition, if you will, to what we already do and how we respect and honor the families that we work with."*—Teacher

However, while *CL* was viewed as helpful overall, some of the modules' content seemed like a review for more experienced teachers. As the curriculum did not account for differences in teachers' skill levels, this made it difficult for some teachers to engage with the material and for coaches to engage some teachers.

As one coach stated, *"I think some of them felt that they were ready to dive deeper. It's hard with a group of such varied individuals to meet everyone where they were."* Another coach added the additional challenge of engaging teachers with more experience, *"You have those teachers that are lifelong learners, and even if the content is not new, they are going to enter ready to learn. And then we have teachers who have been doing this for three years already. The content isn't new and it was harder to engage them."*

Teachers shared similar concerns and felt that parts of the *CL* intervention could be most beneficial to teachers who were newer to early childhood. One teacher mentioned, *"For somebody who is coming new to early childhood, all of these would be really helpful. But for somebody with early childhood experience, it does kind of feel like we're reviewing things that we've discussed before."* A teacher who described herself as more *"seasoned"* said, *"Some of the content was great and we didn't know it. But the things that we might have already known, it was like, 'Ok, we do know that.'"*

Additionally, since the information was foundational and built upon previous trainings, some individuals did view the *CL* content to have overlap with other professional development trainings already in place. For example, one teacher described how the content was similar to the Touchpoints training that is required by their school, *"I've taken Touchpoints training and it kind of is related to it in a way. Like a lot of the things that they talk about there, they talk about in here . . . it's not the same, but it's alike."*

Module-specific feedback. The differences in experience among teachers were further highlighted when individuals were asked to share which of the modules they found most and least helpful. Most modules received mixed reviews, with a few standing out as definitely helpful or not helpful. Individuals generally found modules to be useful when they thought they included new, relevant, and applicable information. Modules were viewed as not helpful when were found to be repetitive or simplistic. A couple of individuals did comment that they found all of the *CL* modules to be helpful. One coach remarked, *"I think it was very useful, extremely useful. . . . The content was amazing. I don't know that any of them weren't useful."* Table 15 provides a summary of individual views for why modules were seen as helpful or not helpful.

Table 15. *CL* module-specific feedback

Modules	Helpful	Not Helpful
Behavioral Understanding	• Informative and applicable to current school approach.	
Development	• Informative, important, and applicable information. Very helpful to how to set up classroom	

Modules	Helpful	Not Helpful
	environment for aiding development.	
Relationships	• Important foundational information and skills.	
Stories	• Informative and applicable information. Not many trainings discuss child identity development. Encompasses culture and teacher interaction.	
Careful Observation	• Applicable to current school approaches and defined the importance of doing observations. Foundational information.	• Some overlap to work already being done elsewhere. Coaches had to work to make this information *"new."*
Culture	• Resulted in useful discussion.	• Some overlap to previous professional development.
Family Engagement	• Resulted in useful discussion and problem solving.	• Simplistic and repetitive compared to other trainings and initiatives already in place.
Meaningful Experiences	• Foundational and applicable information.	• Overlap with Purposeful Interactions.
Purposeful Interactions	• Applicable information.	• Overlap with Meaningful Experiences.
Social-Emotional Competence	• Important information; impacted teacher behaviors.	• Some overlap to previous professional development.
Supportive Environments	• Important information about developmentally appropriate environments.	• Simplistic and repetitive, did not add much to teachers' learning. This topic needed to be expanded.
Home Language	• Interesting and important information.	• Repetitive and similar to information from Development & Culture. Could be embedded in Family Engagement. • Did not seem to be applicable to current ELL teaching structures—*"more abstract."* • Some schools did not have many ELL students and did not think this was applicable to their classrooms.

Dialogic Reading Content

Part of the *CL* intervention initiative was to embed *DR* into the content. This training was very well received among teachers and coaches and viewed as helpful while individuals were able to draw the similarities to *CL*. *"Dialogic Reading is a concrete strategy for reading with young children. So it reinforces a lot of the theoretical concepts of Cradling Literacy,"* said a coach. One teacher discussed how *DR* was related to several specific *CL* modules, *"Dialogic Reading is creating a meaningful experience. It's a really intentional way of doing it. Taking that individual time with the child. Giving them the opportunity to discuss the book, what they're looking at. And maybe creating a story for the book on their own. It plays to that relationships piece, too."*

One coach did mention that it was difficult for teachers to have frequent one-to-ones with children every day and that this piece of the model might need to be reworked to better fit with class schedules and responsibilities, *"I think the biggest barrier that I saw was expecting them to use Dialogic Reading techniques individually and that was unrealistic. Yes, they are with [the children] for many hours each day, but a lot of those hours are taken up with routines, being outside, eating meals, and changing diapers and all kinds of stuff that takes away from their time."*

CL Training and Coaching

Delivery of the *CL* curriculum represented one of the greatest challenge areas for participants. Overall, teachers felt most engaged during the first year when the content felt newer. They also found the peer-to-peer trainings, discussions, and handouts especially helpful and engaging. Challenges included the repetitiveness of the booster sessions, the training schedules, logistics of coaching, and the required time commitments. Feedback is organized below by successes and challenges.

Peer-to-Peer Teaching

*"You feel really **empowered** with knowing your module really well. And it's nice to hear your peer, like you're more likely to learn from a peer, I think, than to be talked to. I felt like we were talking to one another."*

—Teacher

Successes

- First year of *CL* implementation**:** *"I think that initial training was good. I remember them being much more fully implemented."*—Coach
- Peer-to-peer teaching: *"I got a lot from it and so many different ideas from your peers. Things that I did take back to the classroom."*—Teacher
- Handouts*:* *"What I thought was helpful also was them giving us copies of the PowerPoints that we were talking about because going back to it at work, you could go back to the PowerPoints if you wanted to understand more."*—Teacher

Challenges

- Repetitiveness of booster sessions: Repetition of the *CL* content was a recurring theme noted by teachers and coaches.
    - *"It was like the repetitiveness of it. The booster sessions made me feel not as engaged with it. That's when it felt like a waste [of time]."*—Teacher
    - *"We got into the second year and the third year, and the teachers were like 'this is so repetitive.'"*—Coach
- Training schedules: While individuals voted for the trainings to be scheduled on Saturdays, some teachers were frustrated that the trainings were scheduled outside of regular work hours. The perceived repetitiveness of the booster sessions also came up within this challenge: *"And then that second year, the booster sessions, it was a shorter time that we were here. Which was sort of annoying. It was our Saturday. I don't live in Denver, so traveling all the way to Denver for two hours to hear something that I already know, was really frustrating."*—Teacher
- Logistics of coaching: Coaching support was generally provided by pulling teachers out from their classrooms. Logistical difficulties resulting from this included the need to find extra help for the class during the teacher's absence and challenges with continuing on with daily activities. *"Probably the most challenging aspect is the complexity of being in the classroom and being in a child care center with like logistics. So, meeting times. Meeting spaces. Meeting with the whole team. Those pieces are really challenging . . . with call-out, staffing, and ratios."*—Coach
- Time commitment. Teachers and coaches alike felt strained with the time they needed to devote to *CL* activities given their regular job duties and other professional development initiatives.
    - *"I think that the most challenging part in my role is the actual time that we have to invest in coaching, because I'm also a manager and supervisor of these classrooms."*—Coach
    - *"The coaches would give us assignments to do, but then we would have a crazy week with different people coming and going and we wouldn't get the assignment done and it was kind of hard to go back to the coaching."*—Teacher

Recommendations. Along with providing feedback on *CL*, individuals also recommended some improvements. This included only doing *CL* for one year along with the peer-to-peer training or discussion-based teaching, having the opportunity to build relationships with different coaches, and having a *"cheat-sheet"* for all modules during coaching sessions.

CL Impact

While teachers and coaches did experience some challenges with the *CL* intervention, they were able to observe several positive impacts on their classroom environment, teaching, children, and in parent behavior and engagement.

Classroom environment. Teachers appreciated that *CL* funds provided the opportunity to purchase materials and enrich classroom environments. As one teacher put it, *"I thought the resources that you got to make our environments supportive, more aesthetically pleasing, they definitely had a great impact and made it worth it."* A coach mentioned how much being able to purchase materials and add to classroom libraries helped add to the overall impact of *CL*: *"universally the library became a much more important and well-used part of the classroom because of the fact that we did have some funding to help them purchase materials, books, and other things for the library areas. It just made a huge difference in the way the classrooms were put together. You know the money really helped, the training really helped. I did see a pretty significant change in what was happening."*

> *"In the summertime in the nicer weather on our smaller playground, we have a little bookshelf, well we did last year. We had it out there underneath the trees and like a little blanket so that they could read the books they liked out there."*
>
> —Teacher

Impact on teachers. Individuals felt that the biggest impact on teaching practices was in creating intentionality around developing literacy and language skills by providing tools. Teachers described how learning and developing their skills through a *"literacy lens"* really helped in several aspects of their work and how they sought opportunities for literacy development throughout class schedules including playground time, during lunch, and nap time. One teacher said, *"It just gave me more understanding and different ways to facilitate an interaction with a child. When you're reading, or doing an experience, or trying to understand behavior and like just being more intentional about what I'm doing because I have these new tools and I'm trying to use them."* Coaches also observed an impact and change in teacher practices, *"I think teachers became a great deal more intentional about what they were doing and I think that they began to see the big picture a lot better after going through the Cradling Literacy training and coaching and observations."*

Success Story

A coach described how much a child has grown since initially entering school with some developmental challenges: *"This little guy created a whole page full of scribbles that looked a lot like handwriting. He was not even three years old yet and yet he had developed the ability to hold a pencil properly. He really had good strong fine motor skills at that point. His language was amazing. He used 4-, 5-, 6-, 7-word sentences. Very, very fluently. You know, for this little guy, Cradling Literacy was a godsend."*

Impact on children. Children whose teachers received the intervention were observed to have an increased interest in books and resulting positive developmental outcomes. Teachers and coaches credited *CL* for creating a culture that facilitated this development; *"I think that's why our children are so interested in books and excited about them because we set up our environment and our attitudes. I think that's kind of what draws that out of them,"* described a teacher. Individuals told several success stories where young children would engage with books, ask questions, tell their own stories, and grow their developmental skills. Another teacher described a child in her class who *"can hold up a book and she can basically read the book to the kids in the classroom. I think that has a lot to do with us reading to her and doing Dialogic Reading with her since she was like 3 months old."*

Impact on parents/caregivers and family engagement. The impact of *CL* went from teachers to the children and their families. Families had the opportunities to attend *DR* trainings but also learned how to best support their children's literacy development from teachers. The information provided during the *DR* trainings was viewed as very helpful, *"We had done a Dialogic Reading training for the parents here and one of the parents had said that she didn't know that she was supposed to read to kids younger. She didn't know that they were supposed to be read to before five*."

Teachers sought ways to pass on their knowledge to families and involve them more in their classroom environment. One teacher stated that the most important impact was *"definitely teaching the parents, as well, how important it is to*

read with your child. Whether it's just like 15 minutes before they go to bed. But trying to get the family involved as well." Another teacher described how she would play recordings of parents/caregivers reading books for their children, *"we have the parents record or we recorded them reading the book. So it also allowed the children to listen to it whenever throughout the day. It was therapeutic for them to listen to it."*

Impact Findings

Classroom-Level Impact

Impact Research Question 1 (Confirmatory): *Do teachers in the* CL *intervention demonstrate greater short-term increases in their use of language-rich interactions with children than do teachers in the* DR*-only (comparison) group?*

Impact Research Question 2 (Confirmatory): *Do teachers in the* CL *intervention demonstrate greater long-term increases in their use of language-rich interactions with children than do teachers in the* DR*-only (comparison) group?*

At the classroom level:

- **Both *Support for Learning* and *Emotional and Behavioral Support* fluctuated over time**
- **The greatest gains occurred one year into the study**

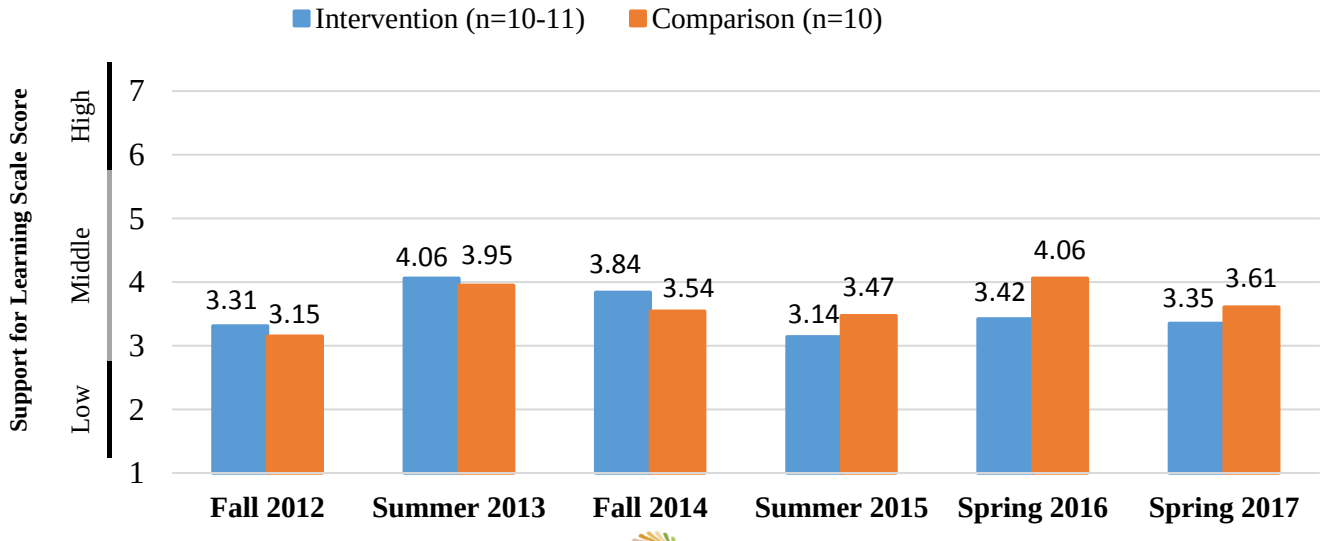
Analysis approach. To answer these research questions, we conducted ANOVAs, which assessed changes over time and between groups on two domains of the CLASS I/T: *Emotional and Behavioral Support* and *Engaged Support for Learning*. Results of these analyses are summarized below.

Engaged Support for Learning. We conducted a 2 (group) x 6 (time) ANOVA and found no time x group interaction, but there was a statistically significant main effect of time across all classrooms between fall 2012 and summer 2013, (p = .01, for the ANOVA contrast between these time points). Figure 4 shows mean scores by group and time point. Overall, results suggest that across study conditions, Support for Learning peaked in summer 2013, and increased again slightly in spring 2016, with scores remaining in the "middle" range of this measure during most of the *Ready to Read* study.

Takeaway: Engaged Support for Learning

- Scores on *Engaged Support for Learning* fluctuated over time.
- About one year into the study, both intervention and comparison classrooms increased significantly.

Figure 4. CLASS I/T *Engaged Support for Learning:* Baseline and 6-month follow-up (n = 21)



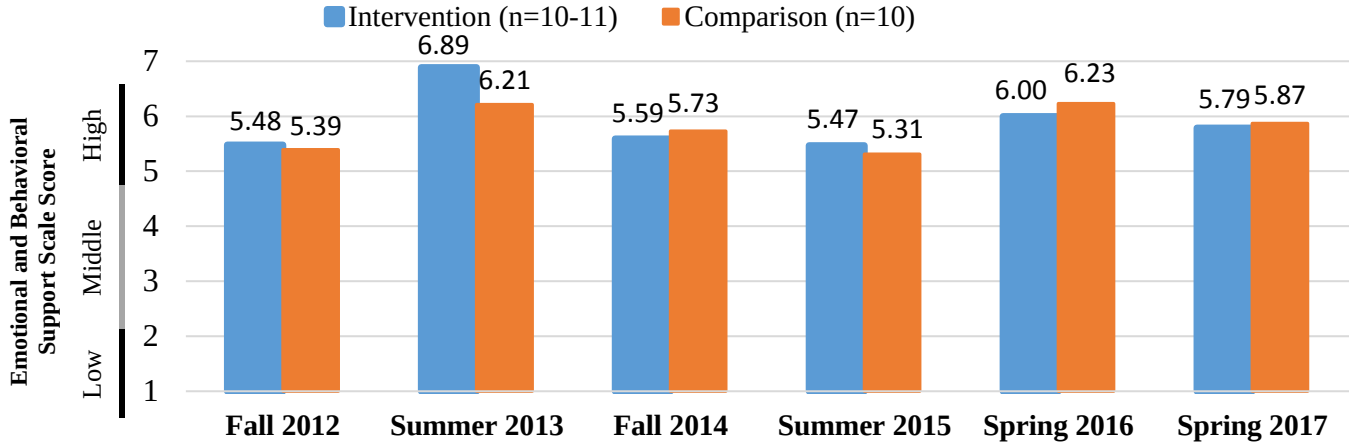
Emotional and Behavioral Support. We conducted a 2 (group) x 6 (time) ANOVA and found no time x group interaction, but there was a significant main effect of time. ANOVA contrasts between individual time points showed a statistically significant increase in scores between fall 2012 and summer 2013 ($p < .001$) and a decrease from summer 2013 to fall 2014 ($p < .01$). Scores then increased again from summer 2015 to spring 2016 ($p = .02$) and remained consistent through the end of the study.

As shown in Figure 5, scores were generally similar by group at each time point. Although the intervention group appeared to make greater gains than did the comparison group between fall 2012 and summer 2013, though this difference was not statistically significant (possibly due to low statistical power). Like the *Engaged Support for Learning* results, *Emotional and Behavioral Support* scores peaked in summer 2013 and spring 2016. Overall, *Emotional and Behavioral* Support scores remained in the middle-to-high range during the *Ready to Read* study.

Takeaway: Emotional and Behavioral Support

- Scores on *Emotional and Behavioral Support* fluctuated over time.
- About one year and four years into the study, scores for the intervention and comparison classrooms increased significantly.

Figure 5. CLASS I/T *Emotional and Behavioral Support*: Baseline and 6-month follow-up (*n* = 21)



Parent-Level Impact

Research Question 3 (Confirmatory): *Do parents/caregivers in the* CL *intervention show greater gains in their support of their child's language and literacy development than do parents/caregivers in the* DR*-only (comparison) group?*

Impact Research Question 5 (Exploratory): *Do families with more program exposure show greater gains in their support of their child's language and literacy development than do those with less program exposure?*

Analysis approach. To answer the parent-level research questions, we conducted 2 (group) x 2 (time) repeated measures ANOVAs to analyze the following outcomes: Parent Survey of Home Literacy total scores (parent-reported use of seven interactive reading techniques) and parent-reported frequency of reading and storytelling. If descriptive statistics suggested that the intervention group made greater gains than did the comparison group, we ran paired-samples *t*-tests separately by group as a follow-up analysis.

Analyses assessed impact of the intervention between baseline and 6 months by condition (intervention vs. comparison). Please note that we did not include 3-month assessments because these were not completed by families who enrolled in the study prior to September 2014, and therefore the 6-month analyses provided the largest sample sizes while also allowing for a longer period of time for the *CL* programming to make an impact. We also conducted 6-month analyses

on level of attendance at an early learning center. The purpose of this is to supplement the intent-to-treat analyses (assignment to intervention or comparison group) with analyses by level of program exposure.

Overall, *Ready to Read* parents increased their support of children's language and literacy development.

- Frequency and quality of parent-child reading increased significantly between baseline and 6 months.

Findings are mixed in terms of their support for the *CL* intervention:

- **Results for parent-child storytelling suggest the intervention has an impact.**
  - Intervention parents made significant gains over time.
  - Comparison group scores did not change over time.
- **Results for frequency and quality of parent-child reading suggest the *CL* intervention does not have an impact.**
  - There were no discernable differences between the intervention and comparison groups in terms of these outcomes.
- **It does not appear that more frequent attendance in a *CL* classroom has an effect on parents' support of language and literacy development.**
  - All parents made gains over time, but this was not related to how frequently their children attended a center.

Intervention impact on parent support of literacy at 6-month follow-up. We conducted 2 (group) x 2 (time) repeated measures ANOVAs to analyze each parent-level outcome from baseline to the time of the 6-month follow-ups. Analyses revealed no statistically significant time x group interaction, which suggests that the *CL* intervention did not differentially impact intervention and comparison groups on these outcomes. However, across the full sample of *Ready to Read* participants, there were statistically significant main effects for time on reading frequency, $F(1, 151) = 15.07$, $p < .01$, $\eta^2 = .09$, and use of interactive reading techniques $F(1, 150) = 9.41$, $p < .01$, $\eta_p^2 = .06$. We also found a marginally significant main effect of time for storytelling frequency, $F(1, 151) = 3.29$, $p = .07$, $\eta_p^2 = .02$. This means that there was an overall increase in parent-reported reading, storytelling, and use of interactive reading techniques between enrollment in the *Ready to Read* program and the 6-month follow-up assessment. The effect sizes, however, are "small."

Takeaway: Parent Impact

Parent-reported reading, storytelling, and interactive reading increased across all participants between baseline and the 6-month follow-up.

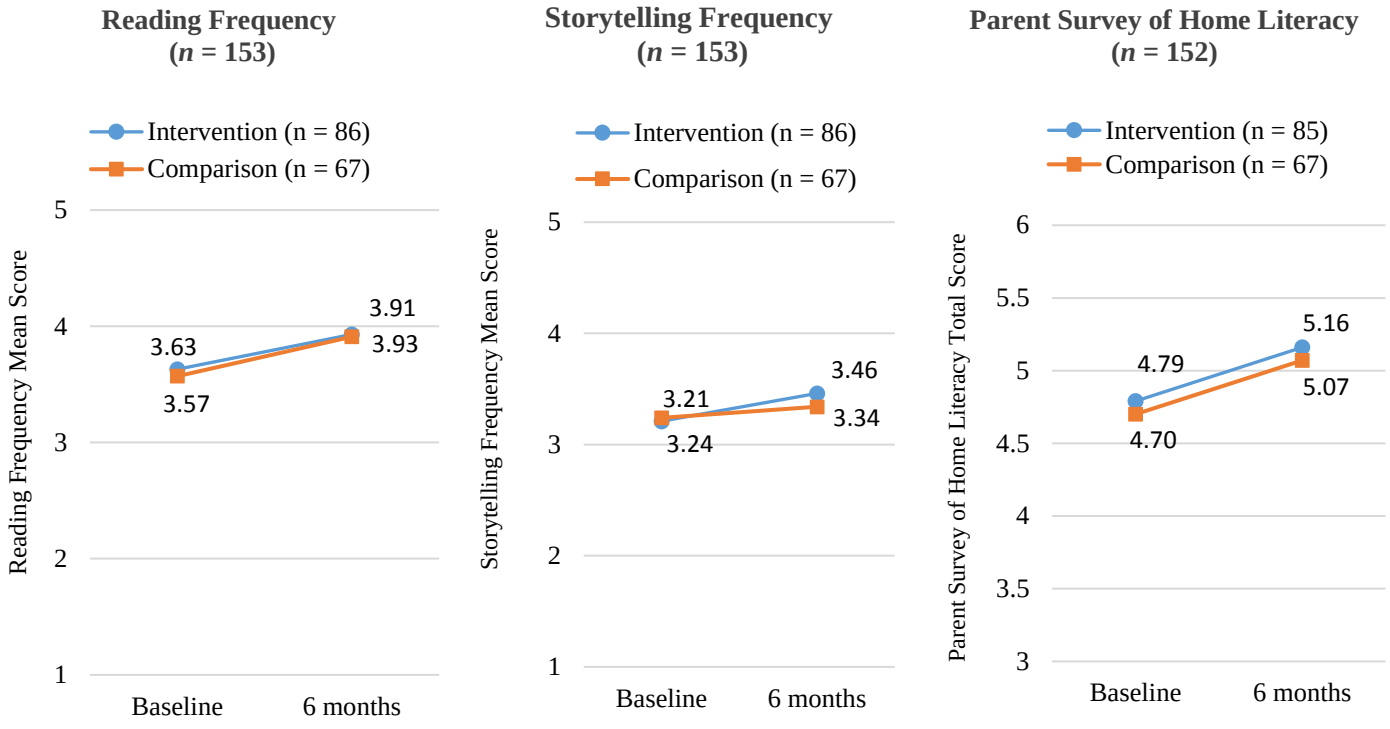
Figure 6 shows baseline and 6-month reading frequency, storytelling frequency, and use of interactive reading techniques. As shown, both the intervention and comparison groups increased over time. For storytelling frequency, it appears that the intervention group made slightly greater gains than did the comparison group.

To further examine change in storytelling frequency over time by group, we conducted paired samples *t*-tests for each group separately. The intervention group made a marginally significant increase in storytelling frequency, $t(85) = -1.80$, $p = .07$; this represents a "small" effect ($d = .20$). The comparison group, however, did not have a statically significant change in scores over time, $t(66) = 0.78$, $p = .43$, $d = .08$. Thus, *t*-test results suggest that intervention group parents/caregivers engaged in storytelling more frequently over time, while parents/caregivers in the comparison group did not increase storytelling significantly. This suggests that the *CL* intervention encourages families to tell stories together more frequently.

Takeaway: Storytelling Frequency

- Intervention group: Increased significantly
- Comparison group: No significant increase

Figure 6. Parent support of language and literacy at baseline and 6-month follow-up



Impact of program exposure on parent support of literacy at 6-month follow-up. As exploratory analyses, we examined changes in parent-level outcome measures over time by days in attendance at an early learning center. For intervention children, attendance constitutes their *CL* intervention "dosage," while for comparison site children, attendance represents their exposure to "programming as usual." Within each study condition, we determined the median number of days children attended and assigned each child at or above the median to a "Higher Attendance" category and those below the median to a "Lower Attendance" group. Results were nonsignificant and were similar to those conducted by group and are provided in Appendix A.

Child-Level Impact

Impact Research Question 4 (Confirmatory): *Do children in the* CL *intervention demonstrate greater increases in their oral language and communication skills than do children in the* DR*-only group?*

Impact Research Question 6 (Exploratory): *Do children with more exposure to* CL *demonstrate greater increases in their oral language and communication skills than do children with a lower level of exposure?*

Overall, children's language and literacy skills increased during their participation in *Ready to Read* (regardless of condition):

Children's *language* skills increased significantly across their first three checkpoints (TSG).
Children's *literacy* skills increased significantly across their first three checkpoints (TSG).
Children's vocabularies increased significantly between baseline and follow-up (CDI).

Results also provide some support for program impact:

Intervention group children who attended a center more frequently made significantly greater gains in language skills and vocabulary than did intervention group children who attended less frequently.
A statistically significant proportion of intervention group children moved from "at or below" the ASQ-C clinical cutoff to "above" the cutoff during the study. There was no significant change over time in classification for the comparison group.

Analysis approach. To answer these research questions, we assessed changes over time and by group (intervention vs. comparison) for study measures related to children's language and literacy development. Outcome measures included teacher-reported TSG language and literacy scores, as well as the parent-reported CDI words produced and ASQ-C scores. Results of these analyses are summarized below. We also conducted 6-month analyses by days in attendance at an early learning center. Additionally, we analyzed TSG outcomes by child age at baseline to determine whether *CL* has a differential impact for infants and toddlers.

Teaching Strategies™ GOLD®

TSG is a system of authentic, observation-based assessments for children ages birth through kindergarten. TSG is used in early education settings throughout Colorado to assess children's learning, abilities, strengths, needs, and interests. Children are assessed by teachers at 3-month checkpoints (in fall, winter, and spring), and results are shared with families through a Family Conference Form.

TSG Language. We analyzed TSG Language scores for the group of children who had three complete checkpoints (n = 122). To contextualize findings, TSG scores are standardized to a mean of 500, providing what can be thought of as "widely held expectations" for children's development based on their age. At the time of each child's first TSG checkpoint, 72% of children in the study "met" or "exceeded" the widely held expectations for language. The average language score at each child's third checkpoint was 470.24, with average scores among intervention and comparison children of 472.55 and 467.87, respectively (see Table 16).

Table 16. TSG Language scores (n = 122)

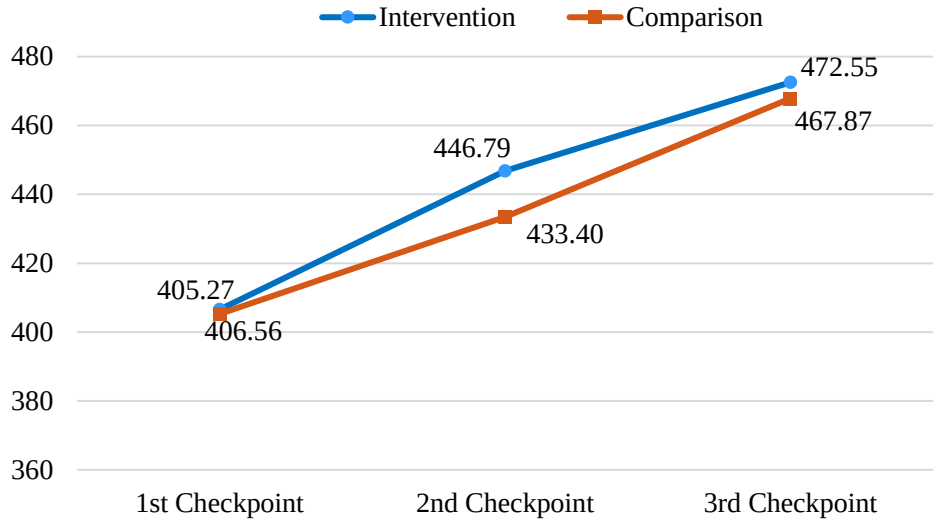
Study Condition	1st Checkpoint		2nd Checkpoint		3rd Checkpoint	
	Mean	*SD*	Mean	*SD*	Mean	*SD*
Intervention (n = 62)	406.56	72.51	446.79	74.68	472.55	74.48
Comparison (n = 60)	405.27	70.07	433.40	75.92	467.87	62.56
Total (n = 122)	405.93	71.03	440.20	75.28	470.24	68.63

We conducted a 2 (group) by 3 (time) repeated measures ANOVA and found a significant main effect of time $F(2, 240) = 129.06$, $p < .001$, $\eta_p^2 = .52$, which means that the full sample of children experienced growth in their language skills over time. The time x group interaction, however, was not statistically significant, indicating that the intervention does not appear to have differentially impacted intervention and comparison children over time, $F(2, 240) = 1.21$, $p = .30$, $\eta_p^2 = .01$ on TSG Language. Figure 7 shows mean scores by checkpoint for the intervention and comparison groups. Notably, the intervention group makes a slightly greater gain than the comparison group between the first and second checkpoints, which could suggest that children in *CL* classrooms experienced a short-term "boost" in language development.

Takeaway: TSG Language

TSG Language scores increased significantly across all study participants over children's first three checkpoints.

Figure 7. TSG Language over time (*n* = 122)



TSG Literacy. The development of literacy skills among young children includes demonstrating knowledge of print material, such as using and appreciating books, responding to books and other texts, and demonstrating phonological awareness, such as rhymes and alliteration or the similarity of words that sound alike. We assessed emergent literacy skills as captured in TSG using repeated measures ANOVA. Table 17 shows results for the group of children who had three complete checkpoints (*n* = 112). At the time of each child's first TSG checkpoint, 83% of children in the study "met" or "exceeded" the widely held expectations for literacy. The average literacy score at each child's third checkpoint was 469.26, with average scores among intervention and comparison children of 476.22 and 461.78, respectively.

Takeaway: TSG Literacy

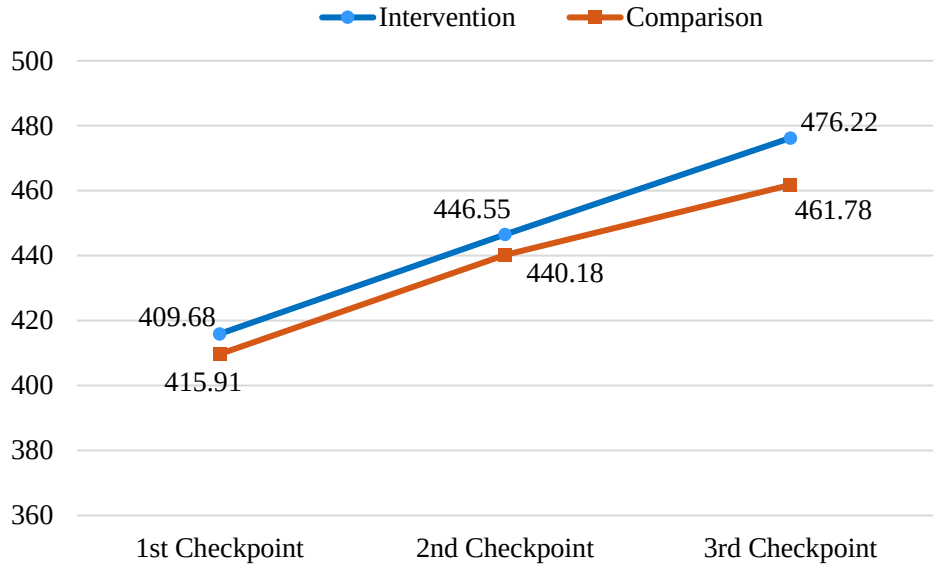
TSG Literacy scores increased significantly across all study participants over children's first three checkpoints.

Table 17. TSG Literacy scores (*n* = 112)

Study Condition	1st Checkpoint		2nd Checkpoint		3rd Checkpoint	
	Mean	*SD*	Mean	*SD*	Mean	*SD*
Intervention (*n* = 58)	415.91	69.57	446.55	71.45	476.22	63.60
Comparison (*n* = 54)	409.68	60.50	440.18	64.62	461.78	58.47
Total (*n* = 112)	412.91	65.13	443.48	68.01	469.26	61.34

We conducted a 2 (group) by 3 (time) repeated measures ANOVA for children with three check points and found a significant main effect of time, $F(2, 220) = 49.09$, $p < .001$, $\eta_p^2 = .31$, but no significant time x group interaction, $F(2, 220) = 0.61$, $p = .54$, $\eta_p^2 = .01$. This indicates that average literacy scores for all children increased significantly over time, but that rates of increase were similar for children in both the intervention and comparison conditions. Figure 8 shows mean scores by checkpoint for the intervention and comparison groups.

Figure 8. TSG Literacy scores over time (*n* = 112)



Program Dosage

We conducted analyses to determine the influence of *Cradling Literacy* "dosage" on children's TSG language and literacy outcomes. Dosage for children is defined as the number of days the child attended the ECE center study classrooms during the time period of their TSG checkpoints. Table 18 indicates the average dosage among intervention and comparison children at the time of each child's 6-month checkpoint.

Table 18. Average *Ready to Read* attendance

	Average Days Attended at Third TSG Checkpoint (*n* = 130)
Intervention	184.30
Comparison	196.30
Total	189.93

We analyzed TSG language and literacy scores related to program dosage using repeated measures ANOVA, grouping children into "lower" and "higher" dosage indicated by the number of days in the classroom. Results for the literacy measure were similar to the intent-to-treat analysis described previously and are provided in Appendix B. However, analyses of the language measure were different from intent-to-treat analyses and are provided below.

TSG language by dosage. We conducted 2 (dosage level) x 3 (time) repeated measures ANOVA for the intervention and comparison groups separately.

For the intervention group, we found a statistically significant dosage by time interaction, indicating that children who attended a *CL* center more frequently had a greater rate of growth with respect to language development than did intervention children who attended a *CL* center less frequently, $F(2, 116) = 4.39$, $p = .01$, $\eta_p^2 = .07$. In particular, ANOVA contrasts showed a statistically significant difference in scores by group between the child's first checkpoint and second checkpoint, $F(1, 58) = 6.72$, $p = .01$, $\eta_p^2 = .10$. Thus, it appears that frequent

Takeaway: Intervention Group Dosage

Intervention group children who attended the center more frequently made greater gains than did children who attended less often.

attendance at a center makes an especially positive impact on children's language growth between a child's first two TSG assessments. The size of this effect, however, is small.

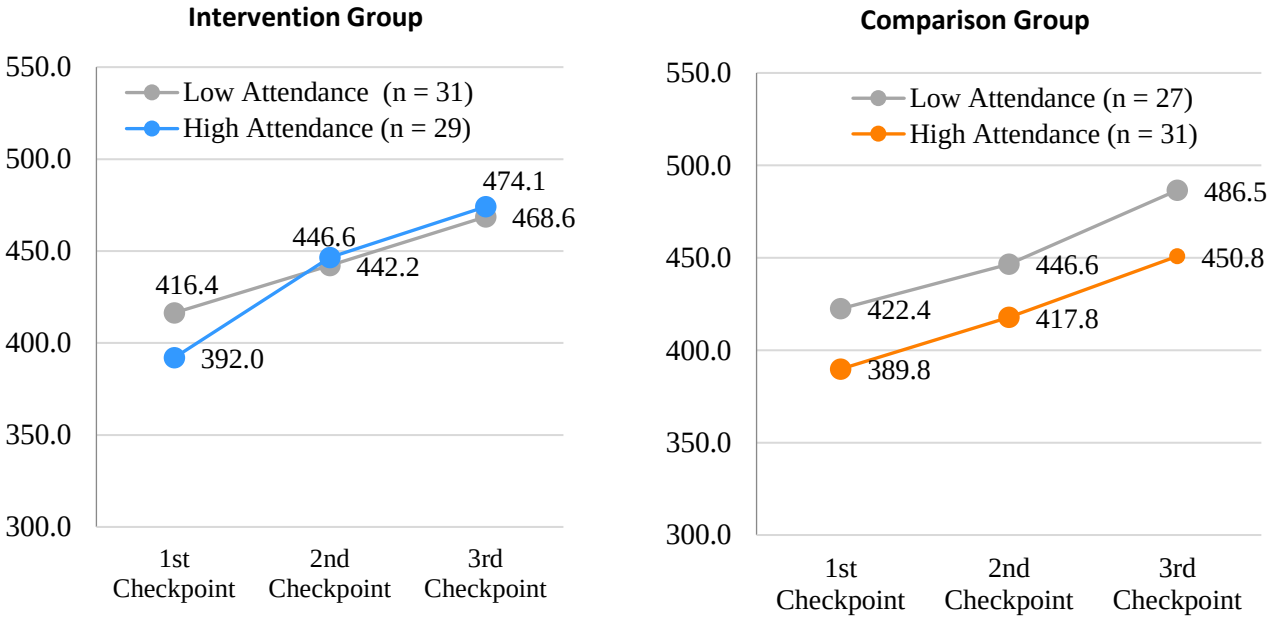
For the comparison group, there was not a statistically significant dosage by time interaction; there was, however, a statistically significant main effect of time, $F(2, 112) = 60.98$, $p < .001$, $\eta_p^2 = .52$, and a marginally significant main effect of dosage, $F(1, 56) = 3.87$, $p = .05$, $\eta_p^2 = .06$. This indicates that the entire group of comparison children made significant gains over time and there are differences in scores by how frequently children attended the center, but attendance does not appear to influence children's rate of growth over time.

Takeaway: Comparison Group Dosage

Scores of comparison group children who attended the center more frequently remained below those of less frequent attendees over time.

Figure 9 shows average TSG Language scores at each the child's first three checkpoints for the intervention and comparison groups at "high" and "low" attendance levels. Notably, children who attended the center less often had higher scores at the first checkpoint in both intervention and comparison centers. In comparison centers (which did not receive *CL* training and coaching), all children made gains, but the children who attended more frequently remained below those who attended less often over time. By contrast, in intervention centers, children who attended more frequently were able to "catch up" to the scores of the child who attended less often by the second TSG checkpoint. This is a positive finding, as it suggests that more *CL* programming translates to greater gains in language development. However, it also suggests that children who do not attend the center as frequently may not benefit from *CL*.

Figure 9. TSG Language scores by condition and attendance



Parent-Reported Child Outcomes

Takeaway: CDI Words Produced

Words produced increased significantly across all study participants.

CDI words produced at 6 months. To compare the number of words produced by children in the *CL* intervention and comparison groups over time, we conducted a 2 (group) x 2 (time) repeated measures ANOVA and found no time x group interaction. There was, however, a statistically significant main effect of time, $F(1, 86) = 36.15$, $p < .001$, $\eta_p^2 = .30$. This means that the number of words children produced (as reported by the parent) increased at similar rates for the intervention and comparison groups between baseline and 6 months. Descriptive statistics are shown in Table 19. Overall, these results suggest that increases in scores are the result of children's development over time and/or due to attendance at an early learning center in general rather than being due to an effect of *CL*.

Table 19. CDI words produced (*n* = 88)

	Baseline		6 Months	
	Mean	*SD*	Mean	*SD*
Intervention (*n* = 53)	33.25	32.12	49.64	33.67
Comparison (*n* = 35)	24.20	28.98	45.60	31.70

Takeaway: ASQ Communication

Intervention group: Increased significantly

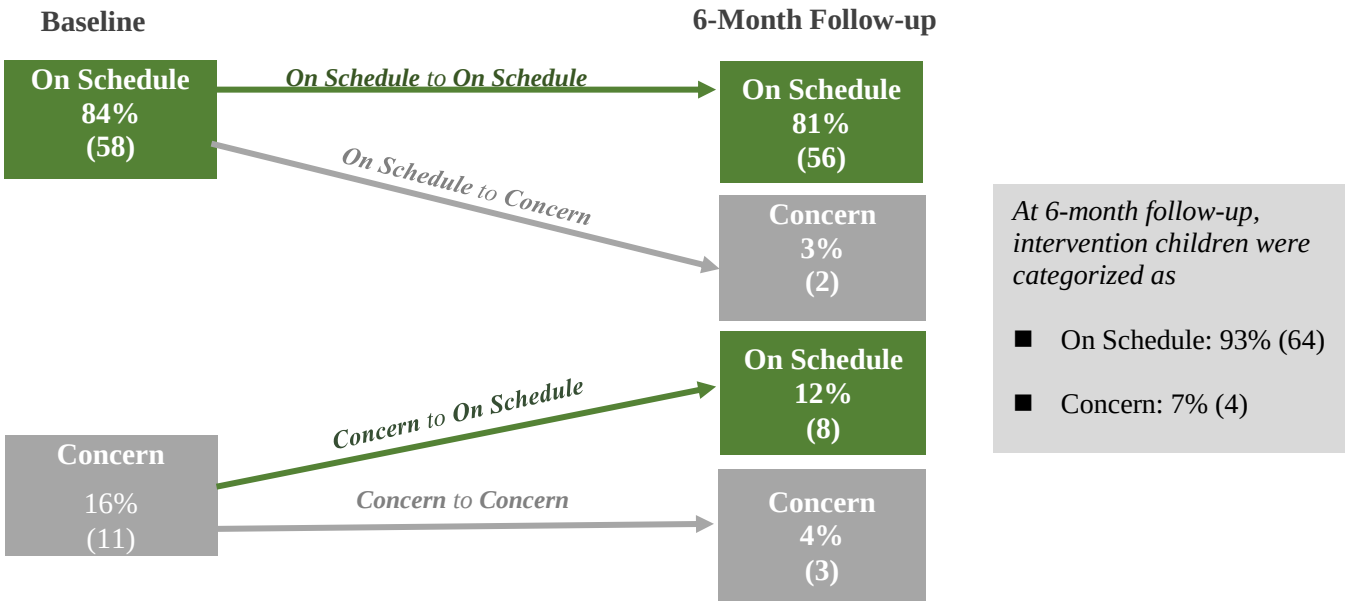
Comparison group: No significant increase

ASQ-C at 6 months. The ASQ-C measure assigns a child to one of three categories related to development in verbal and nonverbal communication skills as "on schedule," "close to the cutoff," or "below the cutoff." To compare changes in children's categorizations over time and to determine whether there are differences in change over time by study condition, we recoded ASQ-C scores to determine whether a child's score 1) remained the same (or decreased), or 2) increased. (Only 10 children's scores decreased over time; therefore, they were collapsed into the "remained the same" category for analysis.)

Results of the chi-square test showed no statistically significant difference between groups regarding changes in ASQ-C scores, $\chi^2(1, N = 124)$, $p = .65$. This means that there is no more variation in classification of children between the groups than we would expect due to chance. However, analyses conducted separately by group showed statistically significant changes over time. These are described below.

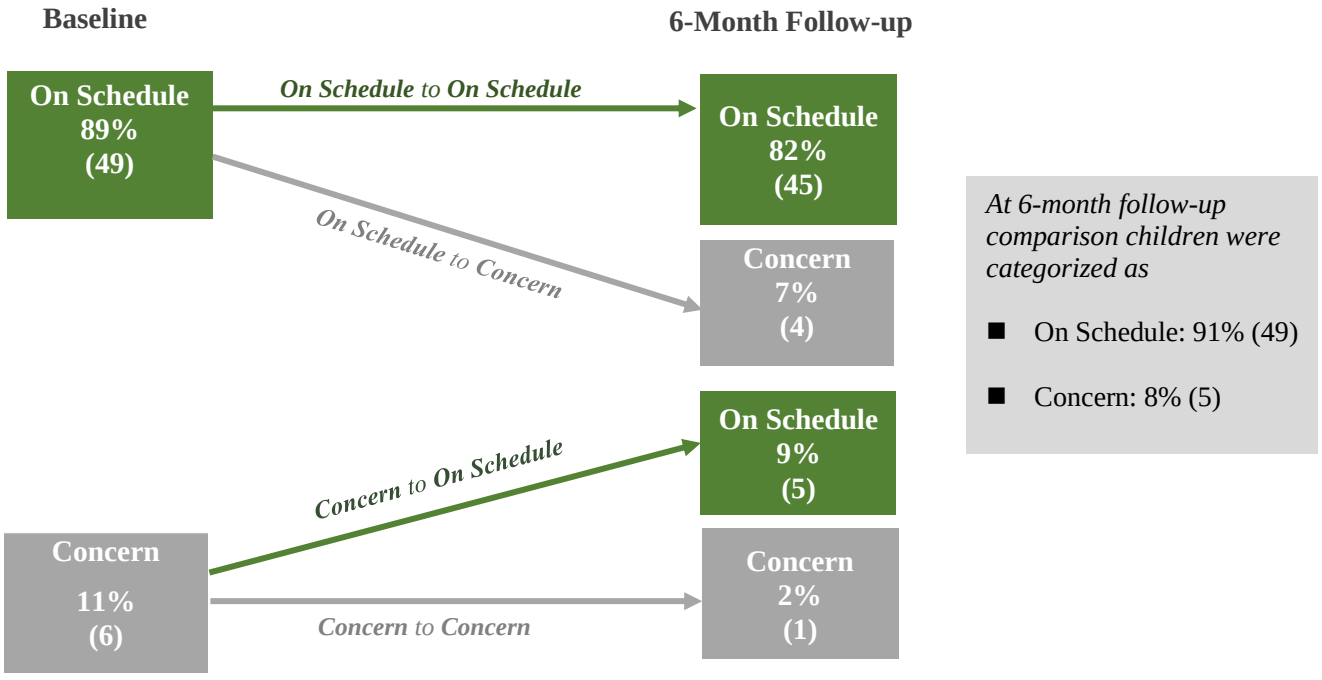
Intervention group. We conducted a chi-square test to assess change over time in ASQ-C classification among intervention group children specifically. Results of a Fisher's Exact Test showed statistically significant change over time, $\chi^2(1, N = 69)$, $p = .03$, whereby the proportion of intervention children who were "on schedule" increased from 84% to 93% (see Figure 10

Figure 10. Intervention children by ASQ categorization (concern *vs* on-schedule) over time (*n* = 69)



Comparison group. A chi-square test was also conducted to assess change over time in ASQ-C classification among comparison group children. Although some children changed classifications over time, this was not statistically significant based on a Fisher's Exact Test, $\chi^2(1, N = 55)$, $p = .45$. Descriptive information is show in Figure 11.

Figure 11. Comparison children by ASQ-C categorization (concern *vs* on schedule) over time (*n* = 55)

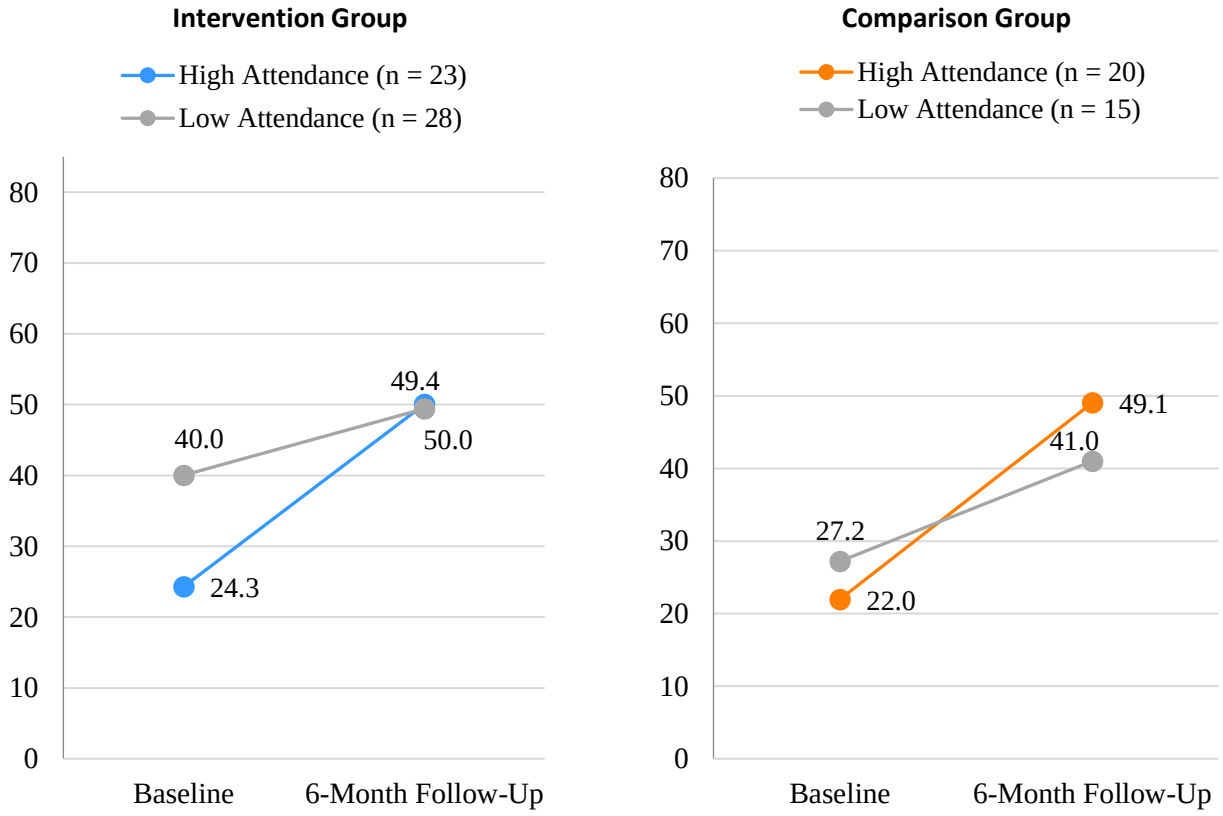


Parent-Reported Child Outcomes by Program Dosage

We conducted exploratory analyses to examine changes in 6-month outcome measures by dosage. Chi-square tests were not performed for ASQ-C due to lower-than-optimal cell sizes (< 5 per group) when cross-tabulated by condition and dosage level. Descriptive results are provided in Appendix A. For the CDI, we conducted the ANOVAs described previously for each group with dosage (child attendance) as the independent variable. Results are presented below.

CDI words produced at 6 months. To compare the number of words produced by children over time and by attendance level, we conducted a 2 (group) x 2 (time) repeated measures ANOVA separately for the intervention and comparison groups. For the intervention group, we found a marginally significant time x group interaction, $F(1, 49) = 3.49$, $p = .07$, $\eta_p^2 = .07$, in which children who attended the center more frequently made greater gains in words produced than did children who attended less frequently. However, we did not find a significant difference in the growth rates of comparison group children based on program attendance, $F(1, 33) = 2.61$, $p = .12$, $\eta_p^2 = .07$. This provides some evidence that *CL* made an impact on children who attended the center more frequently. Results are shown in Figure 12.

Figure 12. CDI words produced: Baseline and 6-month follow-up by attendance



CONCLUSION: FINDINGS, LESSONS LEARNED, AND NEXT STEPS

Summary of Implementation Findings

More than one-half of children in the study (60%) attended a center for a full school year or longer, though the average number of days children attended during their first (or only) year in a center was about 100. Relative to results from past years, this suggests that consistency of attendance may be a challenge, but also that enrollment duration has increased in *Ready to Read* centers over time.

More than 1,000 teachers, parents/caregivers, community members, and/or volunteers have received *DR* training since 2012. During the course of the five-year study, *DR* training was offered most frequently during Years 1–3 of the grant, while *DR* coaching was provided at consistent intervals across years.

The focus of this study was *CL*, for which training was provided approximately monthly early in the study, a few times a year in the middle, and only for new teachers at the end. *CL* coaching was offered during almost all months of the project. Teachers and coaches reported that *CL* had a multifaceted impact on their schools, making a difference in the classroom environment, teaching practices, parent and child behavior, and family engagement. Overall, the *CL* modules were perceived as being useful, though some teachers (especially those who were more experienced) felt that some content was more of a review. In terms of program delivery, teachers and coaches reported that *CL* provided helpful resources, tools, and materials. Preferred modes of training delivery included the monthly training provided during Year 1 and the peer-to-peer training offered in Year 3. Booster sessions provided in Year 2 were generally seen as repetitive (and therefore less helpful). Participants noted that classroom pullouts for the coaching presented logistical challenges in setting up meeting spaces, finding substitute teachers, and maintaining classroom schedules. Additionally, coaches and teachers struggled at times with maintaining their regular work duties in addition to responsibilities resulting from *CL*.

Based on teachers' and coaches' feedback, it is recommended to:

- Emphasize the content in modules that seem to provide the most informative, important, and applicable information (e.g., Behavioral Understanding, Development, Relationships, and Stories) while focusing less on those that duplicate other modules or seem less applicable to teachers' practice.
- Keep the engaging peer-to-peer learning and coaching that promote discussion-based learning, which make *CL* more adaptable and engaging for teachers at all levels of experience. Consider dropping the "booster sessions."
- Schedule the training during regular work hours and find convenient times for the coaching.
- Adapt assignments or expectations and time commitments to teacher and coach needs while keeping in mind their other job duties and professional development requirements.

Summary of Teacher-Level Impact

CLASS I/T observations conducted between 2012–2017 showed increases in teachers' Support for Learning and Emotional/Behavioral Support scores during about the first 10 months of *Ready to Read*. This was followed by a decline in scores, approximately two years into the study. Interestingly, about four years into the study, scores increased significantly again. One possible explanation is that involvement in the *Ready to Read* program—whether as an intervention or comparison site—boosted scores at certain points during the study; for example, perhaps

Evidence of *CL* Effectiveness at the Teacher Level?	
Support for Learning	No
Emotional/Behavioral Support	No
***DR* techniques**	No

excitement about the project increased scores initially. Another possibility is that scores are influenced by the time of year assessments were conducted. For example, it is possible that teachers have more time to attend to domains measured in the CLASS I/T during the spring compared to other times of the year. Turnover of coaches during the study may have also influenced results. Overall, results indicate that the ECE centers provided highly rated *Emotional and Behavioral Support* in the classroom over time, with slightly lower scores (though still in the "middle" range) on *Engaged Support for Learning*. Results do not, however, indicate that *CL* made an impact on these outcomes.

DR observations of teachers conducted during Years 1–3 of the study indicated that all teachers experience gains in their use of interactive reading techniques between the baseline assessment (fall 2012) and the first follow-up (spring 2013). Moreover, scores remained consistent between the first and second follow-up (spring 2014), suggesting that the initial gain was maintained over time. However, it did not appear that intervention teachers made greater gains than did comparison group teachers (as might have been expected with the additional developmental context provided by *CL*).

Summary of Parent-Level Impact

Data collected during the five-year *Ready to Read* study indicate that, on average, parents/caregivers increased their support of children's language and literacy development. Parents/caregivers showed statistically significant increases in reading frequency and use of interactive reading techniques over time across all centers. In addition, parents/caregivers showed a marginally significant increase in storytelling frequency. Effect sizes, however, were small.

Group x time comparisons. Analyses (2 x 2 ANOVAs) examining whether intervention group parents/caregivers showed *greater* gains over time in support of their children's language and literacy skills than did the comparison group parents/caregivers revealed no differences in rates of growth over time.

Pre- and post-analyses by group. For one parent-level outcome, storytelling frequency (*How often do you tell a story with your child?*), intervention group parents/caregivers increased significantly over time (a small effect), while parents/caregivers in the comparison group remained more consistent in their storytelling over time. This shows some preliminary evidence that *CL* may be linked to parent storytelling frequency.

Evidence of *CL* Effectiveness at the Parent Level?	
Reading frequency	No
Interactive reading	No
Storytelling frequency	Yes—preliminary

Program dosage. Analyses examining parent-level outcomes that were conducted based on level of child attendance ("high" versus "low") showed no discernable differences in parent support of children's language and literacy development by dosage.

Overall, results show that *Ready to Read* parents/caregivers made significant gains in the quality and frequency with which they read with their children, regardless of whether they attend an intervention or comparison center. It is possible that the *CL* intervention made a positive impact on parent-child storytelling frequency; however, there is not adequate evidence to demonstrate that the *CL* intervention made an overall impact on parents'/caregivers' support of their children's general language and literacy skills.

Summary of Child-Level Impact

Overall, children across intervention and comparison conditions made statistically significant gains in their language and literacy skills across their first three TSG checkpoints. In addition, parents/caregivers reported that their children's vocabularies increased significantly between baseline and six months. However, results did not suggest that intervention group children showed *greater* gains over time in their language and literacy skills (as measured by TSG) than did children in the comparison group. Utilizing the communication subscale of the ASQ-C, we did, however, find

preliminary evidence that *CL* made an impact at the child level: Children in the intervention group increased significantly over time while children in the comparison group remained more consistent in their communication skills over time.

Evidence of *CL* Effectiveness at the Child Level?	
Language (TSG)	Yes—preliminary
Literacy (TSG)	No
Communication (ASQ-C)	Yes—preliminary
Vocabulary (CDI)	Yes—preliminary

Analyses by dosage. According to the TSG Language and CDI words produced measures, children who attended intervention centers more frequently made significantly greater gains in their language skills and vocabularies than did children who attended intervention centers less frequently. Notably, the children who attended centers more frequently had lower baseline scores than did children who attended less often, which could mean that *CL* helps to "close the gap" among children who enter early learning centers with different levels of language and vocabulary skills. Children in the comparison group did not differ in growth based on attendance.

Lessons Learned

Conducting intervention research in community settings such as early childhood centers is challenging, especially when the goal is to increase the level of evidence for a program. Factors such as low recruitment success, teacher turnover, and child program and study attrition limit the available data, making it difficult to establish cause-effect relationships, especially when the utilization of statistical matching procedures that could account for treatment bias between intervention and control groups are not feasible.

It can also be difficult to identify measures that are sensitive to the impacts of an intervention. For example, perhaps, as teachers and coaches suggested, there were impacts at the center or organizational level, such as increased emphasis on early literacy, which could not be detected by the measures used in this study. Furthermore, changes at the organizational level could explain why we observed gains across all study participants for several measures. Finally, because all participants received support through center-based "programming as usual," as well as *DR*, this made it difficult to detect an effect for the relatively small program component of *CL*.

A factor that we believe facilitated the data collection process was recruiting and retaining a data collector who was able to build trust and rapport with both families and teachers. In addition, the teachers proved to be an invaluable source of support for on-site data collection.

Next Steps

Next steps for the early learning centers are expected to include continuing with *Cradling Literacy* training and coaching and considering ways of modifying the delivery based on feedback from teachers and coaches. For example, centers could explore the possibility of conducting more peer-to-peer trainings.

Recommendations for doing a quasi-experimental study in community settings such as this include: strategies for handling recruitment challenges and attrition (due to its impact on sample size), determining data requirements when establishing a comparison group through matching, evaluating program implementation as part of the study, and collecting multiple types of data from multiple sources.

REFERENCES

Adams, M. (1990). *Beginning to read: Thinking and learning about print*. Cambridge, MA: MIT Press.

Bond, T. G., & Fox, C. M. (2006). *Applying the Rasch model* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Cutspec, P. (2007). Influences of dialogic reading on the language development of toddlers *Winterberry Research Syntheses, 1*(18).

Fenson, L., Marchman, V.A., Thal, D. J., Dale, P. S., Reznick, J. S., & Bates, E. (2007). *Mac-Arthur Bates communication development inventories: Second edition.* Baltimore, MD: Paul H. Brooks Publishing Co., Inc.

Hart, B. & Risley, T. R. (2003). The early catastrophe: The 30 million word gap by age 3. *American Educator, 27*(1). Retrieved from https://www.aft.org/sites/default/files/periodicals/TheEarlyCatastrophe.pdf

Heroman, C., Burts, D. C., Berke, K., & Bickart, T.S. (2010). *Teaching Strategies GOLD: Objectives for development & learning: Birth through kindergarten.* Washington, DC: Teaching Strategies, Inc.

Huebner, C. E., & Joye, E. W. (2004). *Dialogic Reading Observation Form.* University of Washington: Unpublished.

Hoy, W. K., & Woolfolk, A. E. (1993). Teachers' sense of efficacy and the organizational health of schools. *The Elementary School Journal, 93*, 355–372.

Joye, E. (2007). *A psychometric examination of the Dialogic Reading Observation Form in a sample of English and Spanish speaking caregivers* (Doctoral dissertation). Retrieved from http://search.proquest.com/docview/304861168?accountid=14608 (304861168)

Im, J. H., Osborn, C. A., Sanchez, S. Y., & Thorp, E. K. (2007). *Cradling literacy: Building teachers' skills to nurture early language and literacy from birth to five.* Washington, DC: ZERO TO THREE.

La Paro, K. M., Hamre, B. K., & Pianta, R. C. (2012). *Classroom Assessment Scoring System: Manual, Toddler*. Baltimore, MD: Paul H. Brookes Publishing Co.

LoCasale-Crouch, J., Hamre, B., La Paro, K. M., & Pianta, R. C. (2011). *Classroom Assessment Scoring System: Infant Manual.* Unpublished manual.

Mathematica Policy Research. (n.d.) *Baby FACES: Parent Interview 2010–2012 Versions*. [Measurement instrument].

Moats, L. (1999). *Teaching reading IS rocket science*. Washington, DC: American Federation of Teachers.

Neuman, S. B., & Cunningham, L. (2009). The impact of professional development and coaching on early language and literacy instructional practices. *American Educational Research Journal, 46*(2), 532–566.

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. Biometrika, 70(1), 41-55.

Shickendanz, J. (1999). *Much more than the ABCs: The early stages of reading and writing*. Washington DC: NAEYC.

Smith, S. S., & Dixon, R. G. (1995). Literacy concepts of low- and middle-class four-year-olds entering preschool. *The Journal of Educational Research, 88*, 243–253.

Squires, J., & Bricker, D. (2009). *Score adjustment chart for ASQ-3: For use when item responses are missing.* Retrieved from http://agesandstages.com/wp-content/uploads/2015/02/ASQ-3_Score_Adjustment_Chart.pdf

Squires, J., Bricker, D., Twombly, E., Nickel, R., Clifford, J., Murphy, K., Hoselton, R., Potter, L., Mounts, L., & Farrell, J. (2009). *Ages and stages questionnaire: Third edition.* Baltimore, MD: Paul H. Brooks Publishing Co., Inc.

Steiner, P. M., Cook, T. D., Shadish, W. R., & Clark, M. H. (2010). The importance of covariate selection in controlling for selection bias in observational studies. *Psychological Methods*, *15*(3), 250.

US Department of Education. (2015). Number and percent of infants and toddlers receiving early intervention services under IDEA, Part C, by age and state and race/ethnicity. Retrieved August 11, 2016, from http://www2.ed.gov/programs/osepidea/618-data/static-tables/index.html

APPENDIX A. PARENT-LEVEL IMPACT ANALYSES BY DOSAGE

Measure	Condition	ANOVA Results	Mean and Standard Deviation by Time Point and Group
Reading Frequency	Intervention (n = 82)	Time: $F(1, 80) = 7.56$, $p = .01$, $\eta_p^2 = .09$ Time x Dosage: $F(1, 80) = 0.09$, $p = .76$, $\eta_p^2 = .001$	Time point / Attendance / *M* / *SD* / *n* Baseline / Low / 3.61 / 1.16 / 46 Baseline / High / 3.61 / 0.99 / 36 Post 1 / Low / 3.96 / 1.11 / 46 Post 1 / High / 3.89 / 0.78 / 36
	Comparison (n = 67)	Time: $F(1, 65) = 7.90$, $p = .01$, $\eta_p^2 = .11$ Time x Dosage: $F(1, 65) = 0.92$, $p = .34$, $\eta_p^2 = .01$	Time point / Attendance / *M* / *SD* / *n* Baseline / Low / 3.62 / 1.24 / 29 Baseline / High / 3.53 / 1.06 / 38 Post 1 / Low / 4.10 / .77 / 29 Post 1 / High / 3.76 / .97 / 38
Storytelling Frequency	Intervention (n = 82)	Time: $F(1, 80) = 2.76$, $p = .10$, $\eta_p^2 = .03$ Time x Dosage: $F(1, 80) = 0.04$, $p = .84$, $\eta_p^2 = .001$	Time point / Attendance / *M* / *SD* / *n* Baseline / Low / 3.28 / 1.41 / 46 Baseline / High / 3.08 / 1.13 / 36 Post 1 / Low / 3.50 / 1.24 / 46 Post 1 / High / 3.36 / 1.27 / 36
	Comparison (n = 67)	Time: $F(1, 66) = 0.33$, $p = .57$, $\eta_p^2 = .01$ Time x Dosage: $F(1, 66) = 2.59$, $p = .11$, $\eta_p^2 = .04$	Time point / Attendance / *M* / *SD* / *n* Baseline / Low / 3.31 / 1.39 / 29 Baseline / High / 3.18 / 1.39 / 38 Post 1 / Low / 3.17 / 1.34 / 29 Post 1 / High / 3.47 / 1.01 / 38
Home Literacy	Intervention (n = 81)	Time: $F(1, 79) = 6.76$, $p = .01$, $\eta_p^2 = .08$ Time x Dosage: $F(1, 79) = 2.19$, $p = .14$, $\eta_p^2 = .03$	Time point / Attendance / *M* / *SD* / *n* Baseline / Low / 5.09 / 1.72 / 46 Baseline / High / 4.40 / 1.54 / 35 Post 1 / Low / 5.28 / 1.48 / 46 Post 1 / High / 5.11 / 1.39 / 35
	Comparison (n = 67)	Time: $F(1, 65) = 5.33$, $p = .02$, $\eta_p^2 = .08$ Time x Dosage: $F(1, 65) = 1.18$, $p = .28$, $\eta_p^2 = .02$	Time point / Attendance / *M* / *SD* / *n* Baseline / Low / 4.45 / 1.38 / 29 Baseline / High / 4.89 / 1.39 / 38 Post 1 / Low / 5.03 / 1.32 / 29 Post 1 / High / 5.10 / 1.31 / 38

APPENDIX B. CHILD-LEVEL ANALYSES BY DOSAGE

Measure	Condition	ANOVA Results	Mean and Standard Deviation by Time Point and Group
TSG Literacy	Intervention (n = 57)	Time: $F(2, 110) = 55.44$ $p < .001$, $\eta_p^2 = .50$ Time x Dosage: $F(2, 110) = 1.34$, $p = .27$, $\eta_p^2 = .02$	
	Comparison (n = 53)	Time: $F(2, 102) = 35.97$ $p < .001$, $\eta_p^2 = .41$ Time x Dosage: $F(2, 102) = 1.79$, $p = .17$, $\eta_p^2 = .03$	

Time point	Attendance	*M*	*SD*	*n*
Checkpoint 1	Low	412.73	76.40	30
	High	419.00	63.81	27
Checkpoint 2	Low	434.53	82.82	30
	High	457.67	55.72	27
Checkpoint 3	Low	465.23	66.40	30
	High	486.96	60.25	27

Time point	Attendance	*M*	*SD*	*n*
Checkpoint 1	Low	412.08	68.66	25
	High	405.82	53.73	28
Checkpoint 2	Low	447.88	71.62	25
	High	429.93	56.03	28
Checkpoint 3	Low	477.84	65.03	25
	High	447.57	50.05	28

ASQ-C categorization (Concern/On Schedule) by Time Point and Attendance: Intervention Group

	Concern at Baseline		On Schedule at Baseline	
	Concern at Follow-Up 1	On Schedule at Follow-Up 1	Concern at Follow-Up 1	On Schedule at Follow-Up 1
Low attendance (intervention group) (n = 37)	5% (2)	11% (4)	--	84% (31)
High attendance (intervention group) (n = 29)	3% (1)	14% (4)	7% (2)	76% (22)

ASQ-C categorization (concern/on schedule) by Time Point and Attendance: Comparison Group

	Concern at Baseline		On Schedule at Baseline	
	Concern at Follow-Up 1	On Schedule at Follow-Up 1	Concern at Follow-Up 1	On Schedule at Follow-Up 1
Low attendance (comparison group) (n = 24)	--	4% (1)	8% (2)	88% (21)
High attendance (comparison group) (n = 31)	3% (1)	13% (4)	7% (2)	77% (24)