Ready to Read: Play and Learn Study

Social Innovation Fund
Final Report

Submitted to Clayton Early Learning and Mile High Early Learning
September 2017

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

ACKNOWLEDGEMENTS

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

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@milehighmontessori.org
303.962.2623

Recommended citation:
Butler Institute for Families, University of Denver (2017). *Ready to Read: Play and Learn 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 from birth to age three in Denver. The Ready to Read study in Play and Learn sites began in 2012 after being selected by the Mile High United Way (a Social Innovation Fund recipient) to build the evidence base of a promising literacy program. Since 2014, the Butler Institute for Families has served as the evaluator for this initiative. This is the final evaluation report of the five-year study of Ready to Read's *LENA Feedback* intervention in Play and Learn groups.

Play and Learn Groups (PLGs) are structured play groups for young children and their caregiver(s) led by a trained facilitator. As part of standard programming, families receive: training in *Dialogic Reading* (*DR*), an evidence-based literacy intervention; parent meetings on developmental topics; and coaching in new reading and parenting skills. Facilitators also receive coaching to support their own skill development and application.

Some PLGs include the Language Environment Analysis system (LENA), 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 (this intervention will be referred to as *LENA Feedback*) and discuss ways to increase meaningful adult-child interactions that support the development of oral language and communication skills, which have been associated with greater school readiness and literacy.

This impact study includes two confirmatory research questions, examining whether families who received *LENA Feedback* in addition to *DR* and the other PLG components experienced greater gains in child oral and communication skills and greater frequency and quality of literacy activities in the home relative to those who did not receive the feedback. The exploratory research questions address the extent to which dosage of *LENA Feedback* influences those same outcomes. Based on the strong level of evidence for *DR* and the preliminary level of evidence for the *LENA System*, a moderate level of evidence was targeted. To achieve that level of evidence, this study used a quasi-experimental study design, randomly assigning five PLGs to either the experimental (*LENA Feedback*) or comparison (PLGs "as usual") conditions.

Play and Learn Study

Intended Outcomes

- Increased language-rich interactions between children and parents or caregivers
- Increased quality and frequency of book reading
- Increased oral language and communication skills among children

Measures

Child developmental assessments and parent surveys were completed at PLG entry and at children's birthday and half-birthday every year until turning age 3 or leaving the PLG.

- Parent Survey of Home Literacy (use of interactive book reading techniques)
- Baby FACES, Reading Books and Telling Stories (frequency of book reading and storytelling)
- LENA Conversational Turns (frequency with which child is engaged in conversation)
- Ages and Stages Questionnaires, Communication Subscale (ASQ-C)
- MacArthur-Bates Communicative Development Inventories (CDI)

Analysis

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

Implementation and Impact Findings

Ready to Read has been delivered in five Play and Learn groups (PLGs) in the Denver Metropolitan area since January 2013. During the school year (approximately September to May), PLG sessions are offered twice per week at each of the five sites for 2 hours. Sessions are delivered by a trained facilitator. In the summer, PLGs offer "field trips" for families, such as trips to parks or museums.

From January 2013 to May 2017, 393 children ranging in age from 0 to 30 months were enrolled at one of the five PLGs included in this study. Of the 370 children enrolled in programming who met eligibility criteria, 210 (57%) participated in the PLG research study:

- Intervention condition: 132 children (58% of eligible PLG participants from three sites)
- Comparison condition: 78 children (55% of eligible PLG participants from two sites)

PLGs serve diverse families with unique needs: most of the children enrolled in the study are Latino/a (84%), while 86% qualify for free or reduced lunch, and 8% have an Individualized Family Service Plan (IFSP), a marker of special education services. Notably, rates of IFSPs among children in this study are much higher than the 3% found among children in the general population (U.S. Department of Education, 2015). Follow-up data in some form (parent survey, LENA, ASQ-C, or CDI) are available for 67% of the children who completed baseline assessments.

Implementation of the PLG Intervention

This study originally had four implementation questions related to *DR* training and coaching, *LENA Feedback*, and fidelity of programming:

How many individuals were trained on DR?

- 1,009 PLG staff, parents, and community members have been trained in *DR.*

How many PLG and DR coaching sessions were received by facilitators?

- Coaching frequency varied by implementation year, with facilitators in the most recent school year (2016-2017) receiving coaching as-needed.

How frequently did parents in the DR+LENA Feedback group receive LENA Feedback?

- Parents received the LENA recording device and feedback bi-monthly.
- Parents used the LENA device an average of 3 times, and 75% of intervention parents reported receiving *LENA Feedback* sessions.
- On average, *LENA Feedback* sessions ranged from 16 to 21 minutes.

What were parents' perceptions of the LENA Feedback? Did they find it useful? Did they understand the LENA reports?

- Parents were highly satisfied with the *LENA Feedback* sessions, finding the feedback helpful and easy to understand.
- 87% reported that they were "very likely" to change their behavior as a result of the *LENA Feedback.*

During the 2015-2016 school year, three additional implementation questions were added to the study:

What are the characteristics of the LENA intervention as it is being implemented?
What is the consistency of program delivery?
What is the quality of program delivery?

These questions were only partially addressed through the LENA implementation observational study conducted during the 2015-16 school year:

- *LENA Feedback* sessions between parents and facilitators focused on increasing parent-child conversation time and limiting TV or technology exposure.
- PLG facilitators participated in coaching sessions that generally included discussion of the facilitator's strengths, areas for improvement, teaching techniques, and goals for the future.
- *LENA Feedback* was delivered in a consistent and engaging manner.

Program Impact

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

Do parents in the DR+LENA Feedback group show greater gains in their support of their child's language and literacy development than parents in the DR Only group?

- **Conversational turns** increased significantly across all study participants between baseline and follow-up.
- Intervention parents showed significantly greater growth in **reading frequency** than did comparison parents.
- **Storytelling frequency** increased significantly across all study participants between baseline and follow-up, with intervention parents increasing significantly over time while comparison parents did not (when examined separately by group).
- Parent-reported **interactive reading** increased significantly across all study participants between baseline and follow-up with intervention parents increasing significantly over time while comparison parents did not (when examined separately by group).

Do children in the DR+LENA Feedback group demonstrate greater increases in their oral language and communication skills than children in the DR Only group?

- Words produced increased significantly across all children between baseline and follow-up.
- A slightly larger proportion of intervention children versus comparison children were *on schedule* for **verbal and nonverbal communication** (91% versus 87%, respectively), though there was not a statistically significant difference in change over time by group.

Do parents who receive more LENA Feedback show greater gains in their support of their child's language and literacy development than those who do not receive feedback?

- Parents who received two or more feedback sessions made significantly greater gains in **reading frequency** than did comparison group parents who did not receive feedback.

Do children whose parents receive more LENA Feedback demonstrate greater increases in their oral language and communication skills than children whose parents do not receive feedback?

- There was no discernable difference in child outcomes based on *LENA Feedback* dosage.

Contribution of the Study

Results of the Ready to Read PLG study show a number of positive outcomes. Overall, families appeared to improve in a number of ways related to literacy-based interactions. Children increased in the number of words understood and said, fewer children had communication scores in the "concern-range" at follow up, and parents reported greater use of interactive reading strategies in the home over time. Additionally, Ready to Read scores for conversational turns were similar to the LENA Research Foundation's national norms. This is impressive given that Ready to Read targets low-income families, while LENA data are based upon a nationally representative sample. The PLG program and *LENA Feedback* sessions were viewed very favorably among parents, contributing to their understanding and use of home literacy interactions with children. Facilitators provided parents with concrete skills in interactive reading through *DR* training and coaching, and they offered strategies for reducing TV time and increasing parent-child conversation in the home.

Regarding impact of the *LENA Feedback* intervention, results showed that intervention parents made significant gains in the frequency and quality of reading with their children; by contrast, parents in the comparison group did not experience significant growth in these areas. Additionally, parents who participated in *LENA Feedback* two or more times engaged in significantly more reading over time, while those who did not receive feedback or received fewer feedback sessions showed more modest growth.

This 5-year study demonstrated that:

- In general, parents significantly increased their quality and quantity of book reading;
- In general, parents engaged in more parent-child conversational turns;
- Children's vocabularies and communication skills increased, and;
- Parents in the intervention group (who received *LENA Feedback* and *DR*) made statistically significant gains over time in support of their child(ren)'s language and literacy skills, while comparison parents (in the *DR* only group) remained stable or showed more modest gains.

Summary

Although this 5-year study did not establish a moderate level of evidence for the *LENA Feedback* intervention, there is preliminary evidence that receiving *LENA Feedback* helped parents read more frequently and interactively with their children.

In addition, most families (regardless of whether they received the *LENA Feedback* intervention) increased in parent support of language and early literacy and experienced gains in children's vocabularies.

In sum, by most indicators of literacy, parents who received *LENA Feedback* did not make significantly *greater* gains than those who did not receive feedback (when included in the same analysis); therefore, we cannot confidently attribute improvements in scores over time to the LENA intervention. At the child level, vocabularies of both intervention and comparison children increased significantly, but we did not find evidence that the *LENA Feedback* intervention was responsible. Thus, the goal of increasing the level of evidence for the PLG-LENA program to "moderate" was not met.

Strengths and Limitations

This study shows improved outcomes for both children and parents during their participation in Ready to Read and seems to provide preliminary support for the effectiveness of the *LENA Feedback* intervention. 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 language spoken and race/ethnicity and a possible recall bias on parent-reported data. Although the sample size increased every year, statistical power was limited for some analyses, whereby only large or medium-sized intervention effects could have been detected. The relatively small number of participants in each group combined with a between-groups size differential greater than the recommended 1.5 may also have inflated the Type I error rate, resulting in rejecting the null hypothesis when we should not have (the intervention to comparison group sizes on parent measures ranged from 1.8 to 2.2).

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. Past studies of the LENA System had limited external validity. This study provided preliminary evidence that families receiving *LENA Feedback* may have experienced greater gains in reading frequency, and to a lesser extent, storytelling and interactive reading compared to families that did not receive that intervention. This points to the importance of utilizing personalized feedback when working to increase knowledge and develop skills in this population. Future research should replicate the findings here with larger samples and more diverse programs and seek to more closely examine for whom the different elements of the intervention are most effective.

INTRODUCTION

Clayton Early Learning Center's implementation of Play and Learn groups (PLGs) first began in 2009. PLGs are intended for low-income families of children age birth to 3 years who either cannot access or choose not to enroll their children in formal early childhood education programs and who are often socially isolated. The program focuses on engaging parents early in their child's education and encouraging enrollment in quality preschool prior to kindergarten to improve school readiness and reading. In 2012, Clayton Early Learning and Mile High Early Learning were selected by Mile High United Way to be part of the Social Innovation Fund initiative for their promising early literacy programming. As a result of that funding, Mile High Early Learning Center added two PLGs, that became part of this study, along with three of Clayton's PLGs. 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 *LENA Feedback* for improving family and child literacy outcomes. This is the final evaluation report of the five-year study of Ready to Read's Play and Learn program. 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 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).

A critical opportunity for developing 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 -- two early childhood education providers who serve low-income families -- partnered to conduct a quasi-experimental study of promising early language and literacy programming. Play and Learn groups are structured play groups for young children and their caregiver(s) led by a trained facilitator. As part of standard programming, families receive: training in *Dialogic Reading* (*DR*), an evidence-based literacy intervention; parent meetings on developmental topics; and coaching in new reading and parenting skills. Facilitators also receive coaching to support their own skill development and application. The *DR* program has strong evidence to support it, indicating that *DR* can be easily implemented by parents in the home and that use of *DR* by parents has a positive impact on children's language development. However, because book reading is only a small part of the time that parents spend interacting with their children, the current study tested the added value of another intervention that has a preliminary level of evidence: the LENA System, which can be used to provide tailored and data-informed feedback to parents about the home language environment to encourage parents' language use with children throughout the day.

Overview of Prior Research

Dialogic Reading. *DR* is an interactive method of sharing picture books with young children ages birth to 5. Rather than adults reading and children listening, in *DR*, children learn to become storytellers. According to the What Works Clearinghouse (WWC) maintained by the U.S. 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 2-5 years old, which is older than the children in the PLG study. However, a synthesis of practice-based evidence supported the use of *DR* with children under age 3 (Cutspec, 2007).

LENA System. The LENA System features Digital Language Processors (DLPs) that children wear to capture adult-child conversational interactions, child vocalizations, and the audio environment (additional sounds, such as television). The LENA System software analyzes the data captured by the DLPs and generates a variety of reports about verbal interactions with a given child over time. A few studies suggest that as adults receive feedback through these reports, they increase their language use with the children in their care. However, these studies used small samples and less rigorous designs, resulting in only preliminary evidence for the effectiveness of *LENA Feedback*. For example, in one study, although 102 families who were given *LENA Feedback* initially increased the amount of talk with their children, there was some decrease in this effect over time (Gilkerson & Richards, 2009).

Theory of Change

The theory of change for the *LENA Feedback* intervention is that when parents and caregivers receive feedback on the extent to which they engage in frequent and extended conversations with children, they will engage in more frequent and rich verbal interactions with children. These interactions will, in turn, support the development of children's oral language and communication skills, which will result in greater school readiness and reading success, ultimately leading to 3rd grade reading proficiency.

PLG Program Model

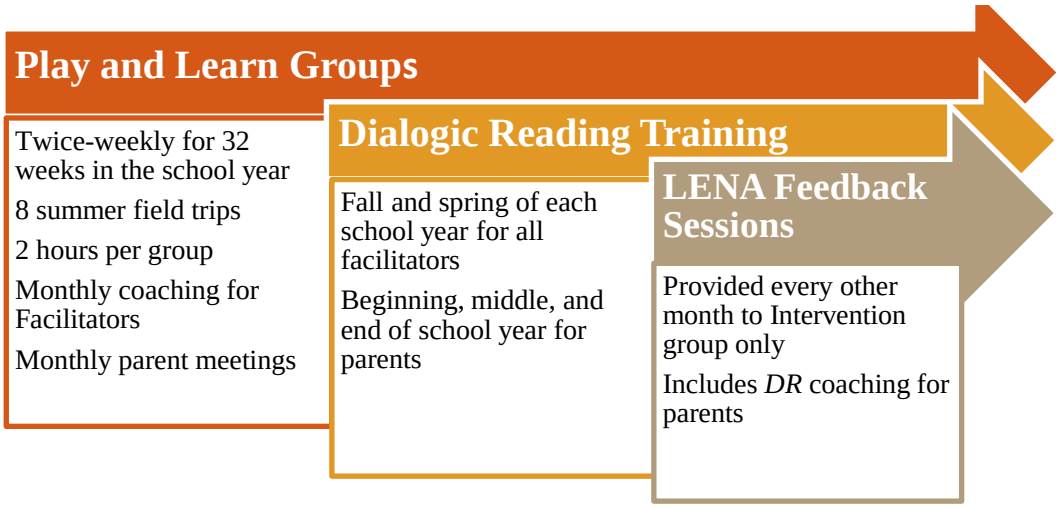
The Play and Learn Groups that operate at three intervention sites and two comparison sites serve families of approximately 100 infants and toddlers per year. The PLG intervention study included three primary components:

- The PLG basic programming of twice-weekly, 2-hour adult-child activity sessions for 32 weeks during the school year and 8 sessions during the summer, led by a facilitator;
- Training in *DR* for PLG facilitators, parents, volunteers, and community partners, and;
- *LENA Feedback* sessions and *DR* coaching provided to families who were in the intervention group.

Facilitators also participated in monthly coaching sessions provided by the Ready to Read project director throughout the school year. Figure 1 depicts the timeline and delivery of PLG program components. It was expected that families who received these program components would demonstrate:

- an increase in quality and frequency of book reading,
- an increase in language-rich interactions between children and parents or caregivers, and
- an increase in oral language and communication skills among children.

Figure 1. Structure of PLG program components



Description of Program Components

Play and Learn Group. All PLG facilitators are experienced early childhood educators and are bi-lingual in English and Spanish. Facilitators receive coaching by the project director. During the PLG session, staff model developmentally-appropriate activities with children (in which literacy goals are embedded), coach parents/caregivers in how to use the activities at home, and provide information on how to support children's development. Interested parents serve as parent leaders to help facilitators with planning activities for children based on their ages and developmental stages. Parent meetings that focus on specific child development and parenting topics are also provided monthly for 8 months, and families are assisted with support services such as ESL classes and child care as needed.

***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 2016-17 school year, parents were offered *DR* training in October, February, and May; volunteers and community partners received *DR* training in April (PLG facilitators were trained during previous school years). Parents 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.

***LENA Feedback* sessions.** Feedback sessions occurred after every LENA administration (bi-monthly). During the feedback session, facilitators and parents engaged in conversation about LENA results and worked together to generate solutions for maximizing the amount of quality child-adult interaction that occurs in the home. The *LENA Feedback* provided to parents included graphs of child-parent interactions that occurred at home over the course of a 10-to-16-hour period.

Program Beneficiaries

One or more parents or caregivers attend the PLG with their child(ren). Parents/caregivers are recruited to PLGs through "word of mouth," and materials and presentations provided to schools, libraries, and local community groups such as churches and child care centers. Beneficiaries include community agencies, businesses, parents, children, volunteers, and early care and education staff.

Program Outputs
Table 1 depicts the program outputs tracked for the PLG *LENA Feedback* intervention.

Table 1. PLG program components and outputs

Program Component	Output
Play & Learn Group	# of families and children enrolled Frequency of attendance # of coaching sessions provided to facilitators by the PLG program director # of parent meetings held
***DR* Training**	# of parents/caregivers, facilitators, volunteers, and partners trained Frequency of trainings
***LENA Feedback* Sessions**	# of sessions received by parents/caregivers Frequency of sessions

Overview of Impact Study

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

- Increase the frequency with which parents and 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 *LENA Feedback* in PLGs is more effective than *DR* alone. Three PLG sites were randomly assigned to the intervention condition (*DR* + *LENA Feedback*), and two PLG sites were randomly assigned to the comparison condition (*DR* Only). Groups included children and their parent or caregiver who were attending the PLGs in separate neighborhoods of Denver at the beginning of the study or who enrolled in the PLG during the course of the study. All families with a child aged 0-30 months were recruited into the study from January 2013 to December 2017. Of the 370 eligible families, 210 participated in this study (132 from intervention sites,78 from comparison sites).

Targeted level of evidence

The evaluation targeted a moderate level of evidence given that *DR* has strong evidence to support it, but that *LENA Feedback* has only preliminary evidence. Although *DR* is easily implemented by parents and has been shown to successfully impact children's language development, it is hypothesized that the *LENA Feedback* may encourage parents' language use with children throughout the day. Although it makes sense that *DR* and *LENA Feedback* would complement each other, the LENA System is expensive. Thus, it is of both practical and theoretical significance to learn if *LENA Feedback* has a measurable benefit above and beyond an intervention like *DR*, which can be implemented rather inexpensively.

Program Implementation and Impact Research Questions

The study included eleven impact and implementation research questions. Data for the impact questions were collected from parents/caregivers, while data for implementation evaluation were collected from parents or PLG staff. Families were assessed at baseline and at children's half-birthday until children aged out of the PLG (see also Table 4 in "Study Approach and Methods"). The research questions addressed in this study include:

Impact (Confirmatory)

1. Do parents in the *DR+LENA Feedback* group show greater gains in their support of their child's language and literacy development than do parents in the DR Only group?

2. Do children in the *DR+LENA Feedback* group demonstrate greater increases in their oral language and communication skills than do children in the DR Only group?

Impact (Exploratory)

3. Do parents who receive more *LENA Feedback* show greater gains in their support of their child's language and literacy development than those who do not receive feedback?
4. Do children whose parents receive more *LENA Feedback* demonstrate greater increases in their oral language and communication skills than children whose parents do not receive feedback?

Implementation

5. How many individuals were trained on *DR*?
6. How frequently did parents in the *DR+LENA Feedback* group receive *LENA Feedback*?
7. What were parents' perceptions of the *LENA Feedback*? Did they find it useful? Did they understand the LENA reports?
8. How many PLG and *DR* coaching sessions were received by facilitators?
9. What are the characteristics of the LENA intervention as it is being implemented?
10. What is the consistency of program delivery?
11. What is the quality of program delivery?

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 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.

In terms of PLG programming, the following notable changes occurred:

- The project experienced turnover among PLG facilitators. It is possible that differences in the linguistic abilities of facilitators who were part of the study over time may have influenced programming and enrollment (e.g., a facilitator who is fluent in Spanish may have been more successful in recruiting Spanish-speaking families).
- There was also a change in the staff person who provided coaching for the PLG facilitators.
- By design, PLGs encourage interested parents to take on a leadership role in the PLGs, and there was likely some variation in the dynamics of each PLG depending on the parent(s) currently in that role.
- There was a change in the physical location of one of the PLGs due to a change in the availability of space.
- Program leaders indicated that the PLG programming became more formalized over the course of the five-year project.

STUDY APPROACH AND METHODS

In the Ready to Read PLG study, the program model involves parents using the LENA recording device and receiving feedback bi-monthly throughout the program. This results in parents receiving feedback sessions fairly frequently, with discussions between the parent(s) and facilitator focused on individualized strategies for enhancing adult-child communication. This study was conducted in Denver, Colorado from 2012 to 2017 with data collected from two early childhood education programs that operate five Play and Learn groups. 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

To evaluate program implementation, we analyzed the frequency with which parents received *LENA Feedback*, and parents' perceptions of the *LENA Feedback*. Surveys were administered by PLG facilitators to collect parents' report of the duration of feedback sessions and the degree to which the feedback was helpful and easy to understand. Parents also reported how likely they were to change their behavior as a result of the *LENA Feedback*. For this study, we utilized data from 138 parents who completed a total of 401 *LENA Feedback* surveys.

Other markers of implementation included: the number of PLG sessions participants attended, the frequency of training and coaching provided to PLG facilitators; and the amount of *DR* training provided to families, facilitators, and community members. 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 frequently did parents in the DR+LENA group receive LENA Feedback?	Number of surveys completed	*LENA Feedback* surveys from parents
What were parents' perceptions of the LENA Feedback? Did they find it useful? Did they understand the LENA reports?	Parent-reported satisfaction	*LENA Feedback* surveys from parents
How many PLG and DR coaching sessions were received by facilitators?	Number of sessions delivered	Program records

Impact Study Design

To address the impact research questions that explore parent and child language and literacy development outcomes, a quasi-experimental longitudinal design was employed to explore differences in outcomes between treatment (*DR + LENA Feedback*) and comparison families (*DR* only, though families contributed LENA data for comparison purposes but did not receive feedback). Strengths of this design include the availability of a comparison group and LENA data from both groups, providing potential to reach a moderate level of evidence. Limitations include threats to validity whereby assignment at the PLG level may result in families in treatment groups having different characteristics than those in comparison groups because membership in PLG is based on the area of town where they live. Another possible limitation is if wearing the LENA device served as its own intervention (regardless of whether or not families received feedback), causing families in both groups to increase their vocalizations with children.

Sampling, Measures, and Data Collection

Sampling

Child inclusion/exclusion criteria in study. All children under the age of 2.5 years who were enrolled in one of the PLG programs at the start of Ready to Read or who entered the PLG during the course of the study were recruited. To participate in the study, families needed to provide consent within 30 days of program enrollment to allow data collected from new families to reflect a true baseline. Thus, families were only excluded based on age or delay in consenting to join the study.

Table 3 shows program and study enrollment rates by school year and condition in the study (intervention and comparison). Since the start of Ready to Read in January 2013, a total of 393 children ages 0-30 months enrolled in a PLG. 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. For the 2016-17 school year, study recruitment ended in November; therefore, children who enrolled in a PLG in December 2016 or later (n = 23) were not eligible for the study. The table below reports the total program enrollments for 2016-17, with the number eligible for the study in parentheses.

Table 3. New enrollments in Play and Learn Ready to Read by school year and study condition[1]

	2012-13 school year	2013-14 school year	2014-15 school year	2015-16 school year	2016-17 school year[4]	Total
Total enrolled in a Play and Learn Group[2]	106	77	62	67	81 (58)	393 (370)
Intervention sites (N = 3)	71	51	34	35	49 (38)	240 (229)
Comparison sites (N = 2)	35	26	28	32	32 (20)	153 (141)
Enrolled in study[3]	47	44	41	38	40	210
Intervention sites (N = 3)	31	29	21	23	28	132
Comparison sites (N = 2)	16	15	20	15	12	78
Study enrollment rate (% of eligible)	44%	58%	64%	58%	69%	57%
Intervention (%)	44%	59%	59%	68%	74%	58%
Comparison (%)	46%	58%	71%	47%	60%	55%

[1]Includes pairs of siblings from the same family (n=32) and one family with three siblings.[2] Includes only children eligible for the study
[3]Excludes children who enrolled in the study but did not complete a baseline assessment
[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 2013, a total of 210 parents consented to participate in the research study, representing 57% of those eligible. By study condition, intervention sites included 132 study participants, while comparison sites had 78 study participants. Study enrollment rates were similar across the intervention and comparison conditions.

Measures and Instruments

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

ASQ-C. This is a six-item questionnaire for which parents or 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 to the child. If the score is below the cutoff, 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 2) increased. 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 Language Production. 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 1800 children in three locations, and numerous studies have documented the reliability and validity of these measures.

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

LENA Conversational Turns. The LENA™ system provides a measure of adult-child conversational turns, which occur when an adult speaks and the child responds, or when the child produces a vocalization and the adult responds. Conversational turn count data are standardized to reflect frequency in a 12-hour period. LENA provides other data such as child vocalizations, adult words, and an estimate of the child's developmental age, but these measures tend to be highly correlated, and conversational turns was deemed the most salient outcome to the PLG study. To reduce the impact of outliers, conversational turn count data were Winsorized, wherein the top 5% of scores were set equal to the value of the 95th percentile, and the bottom 5% of scores were set equal to the value of the 5th percentile. The developers of the LENA System compared "Adult Word Count" estimates obtained from the LENA Pro Software with the results obtained from human transcribers and found that the two were correlated .92, providing evidence for the reliability of the software (Xu, Yapanel, & Gray, 2009).

Parent Survey of Home Literacy (Smith and Dixon, 1995). Parents 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 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 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

A research staff member attended PLG sessions regularly to describe the study to new families and gain their consent to participate. Baseline data were collected from families within about 1 month of their enrollment in the PLG, and again approximately every 6 months, at the time of the child's birthday and half-birthday. At four of the five PLG sites, the program facilitators sent LENA devices home with the family and returned them to the evaluation team for analysis; at one PLG site, the evaluation team distributed/collected LENAs. Videotaped reading observations were conducted at the PLG by evaluation staff (Years 1-4 only). Procedures for collecting parent surveys and child assessments changed for Year 5 of the study; for more information, see "Changes to the Subgrantee Evaluation Plan" section on pages 18-20.

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

Table 4. Ready to Read PLG study research questions

Research Question	Type of Question	Measures	Timing
1. Do parents in the *DR+LENA Feedback* group show greater gains in their support of their child's language and literacy development than do parents in the *DR* Only group?	Impact: Confirmatory	Parent Survey of Home Literacy Baby FACES Reading Books and Telling Stories Dialogic Reading Observation Schedule (DROS; Years 1-4 only)	Baseline Child's half-birthday until aging out
2. Do children in the *DR+LENA Feedback* group demonstrate greater increases in their oral language and communication skills than do children in the *DR* Only group?	Impact: Confirmatory	Ages and Stages Questionnaires, Communication subscale MacArthur-Bates Communicative Development Inventories (CDI)	Baseline Child's half-birthday until aging out

Research Question	Type of Question	Measures	Timing
3. Do parents who receive more *LENA Feedback* show greater gains in their support of their child's language and literacy development than those who do not receive feedback?	Impact: Exploratory	Parent Survey of Home Literacy Baby FACES Reading Books and Telling Stories DROS (Years 1-4 only)	Baseline Child's half-birthday until aging out
4. Do children whose parents receive more *LENA Feedback* demonstrate greater increases in their oral language and communication skills than children whose parents do not receive feedback?	Impact: Exploratory	Ages and Stages Questionnaires, Communication subscale MacArthur-Bates Communicative Development Inventories (CDI)	Baseline Child's half-birthday until aging out
5. How many individuals were trained on DR?	Implementation	Training sign-in sheets	As occurs
6. How frequently did parents in the *DR+LENA Feedback* group receive *LENA Feedback*?	Implementation	Parent *LENA Feedback* Survey	Bi-monthly: Intervention group only
7. What were parents' perceptions of the *LENA Feedback*? Did they find it useful? Did they understand the LENA reports?	Implementation	Parent *LENA Feedback* Survey	Bi-monthly: Intervention group only
8. How many PLG and DR coaching sessions were received by facilitators?	Implementation	Coaching logs	As occurs
9. What are the characteristics of the LENA intervention as it is being implemented?	Implementation	Qualitative observation of *LENA Feedback* & facilitator coaching	2015-16 school year
10. What is the consistency of program delivery?	Implementation	Structured observation of *LENA Feedback* & facilitator coaching	2015-16 school year
11. What is the quality of program delivery?	Implementation	Quality assessment of *LENA Feedback* & facilitator coaching	2015-16 school year

Note: Research questions 9 through 11 (related to implementation) were added for the 2015-2016 school year when supplemental funding was available but were discontinued when funding was discontinued.

Data collection rates at baseline and first follow-up assessment (referred to as "Follow-up 1" throughout the report) are provided in Table 5 by measure and overall (for any measure) as well as by group (intervention and comparison). The PLG study sample included 210 children with baseline data from at least one study measure. For the first follow-up assessment, completion rates ranged from about 40% to 65% of children.

Table 5. Data collection rates by study measure at Baseline and Follow-up 1

Measure	*Number with a Baseline*			*Number of Baselines with Follow-up 1*		
	Intervention	Comparison	Total	Intervention	Comparison	Total
Any data collection method	132	78	210	95	46	141
LENA	117	54	171	80	34	114
Parent survey	116	71	187	67	31	98
ASQ-C	100	64	164	45	23	68
CDI*	72	53	125	38	21	59

*Only includes children who were eligible for the instrument based on age (8 months or older)

Importantly, we did not continue to collect data from families once they left a PLG. Thus, the study completion rates above reflect both program and research study attrition. One key challenge was the difficulty retaining comparison families in the study; this could be because families not receiving the *LENA Feedback* intervention were less invested and engaged in the study. We also experienced difficulty in collecting the ASQ measure, in particular, because the age windows are finite and require close monitoring to ensure that families complete the correct instrument. For example, in some cases, ASQs were "outdated" by the time parents returned the measure, such as a 3-to-5-month assessment completed when the child was 6 months old. Notably, the CDI measure includes only children ages 8 months and older, limiting data on that outcome.

Table 6 shows the demographic characteristics of children enrolled in the study. There were roughly equal proportions of male (51%) and female children (49%), and most children were identified as White (88%) and Hispanic or Latino (84%). Spanish was the primary language for many (62%). Most children were eligible for free or reduced lunch (F/RL; 86%). At 8%, the proportion of children in this study with an Individual Family Service Plan (IFSP) was much higher than the rate in the general population, which is estimated at 3% (U.S. Department of Education, 2015).

Table 6. Play and Learn child demographic characteristics (*n* = 210)

	N	%
Gender: Female	103	49.0
Male	107	51.0
Race: White/Caucasian	184	88.0
Other	25	12.0
Ethnicity: Hispanic or Latino	177	84.3
Not Hispanic or Latino	33	15.7
Free/Reduced Lunch: Eligible	166	85.6
Not eligible	28	14.4
Language: English	52	25.0
Spanish	128	61.5
English and Spanish	28	13.5
Individual Family Service Plan: Yes	14	7.6
No	170	92.4

Similar to the children, mothers (or other caregivers attending PLGs) were mostly White (91%), Hispanic or Latino (80%), and the majority spoke Spanish (53%) or English and Spanish (24%).

Baseline Equivalence Analysis

Random assignment. Randomization occurred at the PLG level rather than at the individual level to minimize the possibility that *DR*-only families would be exposed to the LENA intervention via interaction with other families in their PLGs (contamination). Names of the five sites were written on individual pieces of paper and the original external evaluator drew three of five PLG locations to be assigned to the intervention condition while the other two were assigned to the comparison condition. Despite randomization, analyses revealed significant demographic differences between intervention and comparison children, and, notably, the two randomly selected comparison sites were the PLGs that had not been in operation prior to Ready to Read. 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 vary by measure (i.e., participants have baseline data for some measures but not for others), analyses were done separately for each measure as well as overall, based on data from any source.

Characteristics for which statistically significant differences ($p \leq .05$) were found are summarized in Table 7 for each measure. As shown, intervention and comparison groups generally differed on race, ethnicity, and/or language, with the intervention group having more Hispanic or Latino, bilingual or Spanish-speaking, and white participants than did the comparison group. In addition, for the ASQ-C sample, the intervention group had a significantly higher proportion of boys (56%) than did the comparison group (30%), and in the CDI sample, the intervention group had more families who are eligible for free or reduced lunch (97%) than did the comparison group (70%). These were not included as covariates in analyses because they were confounded with assignment to treatment group. There were no significant differences between groups on child age at baseline, suggesting that results will not be differentially influenced by children's language ability based on their developmental stage.

Groups did not differ on baseline parent and child impact measures, except for book reading frequency. For example, among those who completed a parent survey, comparison group parents reported significantly higher rates of book reading at baseline ($M = 4.42$, $SD = 0.62$) compared with the intervention group ($M = 3.79$, $SD = 1.04$). Interpretation of results take into account this lack of baseline equivalence.

Table 7. Statistically significant differences between intervention and comparison groups at baseline[1]

	Demographic characteristics							Child-level measures		Parent-level measures				
	Age	Gender	Race	Ethnicity	Language	F/RL	IFSP	ASQ-C	CDI	LENA	DROS	Home Literacy	Reading freq.	Story-telling
LENA (n =114)			X	X	X									
Survey (n = 98)			X	X	X								X	
ASQ-C (n = 68)		X	X	X	X								X	
CDI (n = 59)			X		X	X								
Any Measure (N = 210)			X	X	X								X	

[1] $p \leq .05$

In summary, there are differences between the intervention and comparison groups, which include characteristics such as home language and frequency of book reading that may 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 one exception: among the sample of children with survey data, those who spoke Spanish or were bilingual were significantly more likely to complete a follow-up compared with children who spoke English only, $\chi^2(1, N = 152) = 6.62$, $p = .01$. Thus, English-speaking families may be under-represented in the sample due to attrition. There were no statistically significant differences in study attrition by condition, indicating that intervention and comparison group families are equally likely to have complete data.

Changes to Subgrantee Evaluation Plan

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

Changes to data collection procedures. During years 1-4 of the study, an evaluation team member collected baseline and follow-up survey packets (including questions about parent-child reading behavior, the ASQ-C, and the CDI) in-

person at the PLG. During Year 5, budget cuts necessitated that evaluation staff spend less time on-site; therefore, PLG facilitators distributed the parent survey packets and envelopes in which parents could return the surveys confidentially; those were returned to evaluators by the PLG facilitator or by the parent via postal mail. This change to the data collection protocol was approved by the University of Denver's Institutional Review Board (IRB).

Discontinuation of the *Dialogic Reading* videotaped observations of parent-child reading. In Years 1 through 4, videotaped observations of parent-child reading were used to assess changes in parents' use of *DR* strategies. When funding was decreased after Year 4, it was necessary to modify the evaluation strategy. In-person reading observations are resource-intensive and since the "home literacy" measure on the parent survey provides similar data, the parent-child reading observations were discontinued for Year 5. Findings from observational data collected during Years 1-4 indicated that use of interactive reading techniques increased significantly across all study participants between baseline and follow-up; for analyses conducted separately by group, the intervention group showed significant gains while the comparison group did not, providing some evidence that *LENA Feedback* made an impact. For more information, please see the Year 4 report, *Ready to Read: Cradling Literacy Study, Social Innovation Fund Year 4 Annual Report*.

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 *LENA Feedback* intervention and coaching provided to PLG facilitators. Unfortunately, we were only able to conduct a few observations (three *LENA Feedback* sessions and two coaching sessions) 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. Findings indicated that *LENA Feedback* sessions consistently focused on increasing parent-child conversation time and limiting TV or technology exposure and were consistent in terms of duration (about 10-15 minutes long). PLG facilitator coaching sessions generally included discussion of the facilitator's strengths, areas for improvement, teaching techniques, and goals for the future. However, sessions varied in terms of who initiated the discussion, as one session appeared to be more supervisor-led while the other was co-led. For more information, please see the Year 4 report, *Ready to Read: Cradling Literacy Study, Social Innovation Fund Year 4 Annual Report*.

Infeasibility of using Propensity Score Matching. Although Propensity Score Matching (PSM) is an increasingly popular method for facilitating comparisons between treatment and control groups that have important differences at baseline, 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 non-treatment 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 57%), 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, where our ratio was 116:71 on the outcome measure for which we had the most data). In addition, although the original proposal indicated that possible variables to include in PSM were, among others: income; household size; caregiver race/ethnicity, employment status, marital status, education, and age; as well as several other child characteristics, there was not a strategy in place for collecting data on most of those characteristics. The only data available to this evaluation team were child gender, home language, race/ethnicity, presence of individual family service plan (IFSP), and eligibility for free or reduced lunch (for which there was little variability as over 85% of the children in the study qualified). Thus, we lacked a strong set of characteristics 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 *LENA Feedback* intervention.

Did not use Maximum Likelihood (ML) estimation for handling missing data. In the SEP, the original evaluator proposed using MPlus software to conduct 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: a) 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; b) several of the measures in this study cannot have item-level missingness that can be imputed; c) 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 (FIML). 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 used the more conservative intent to treat approach with 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 8). Power analysis, missing data, and attrition were also examined.

Implementation Evaluation Analysis

We analyzed *LENA Feedback* survey data with SPSS software; analyses presented here are descriptive, including means, counts, frequencies, and percentages. PLG attendance and *DR* training data were obtained from programs; analysis includes summary information such as the number of trainings by month and total number of trainings. For *LENA Feedback* and PLG coaching observations (included in the Year 4 report), evaluators reviewed qualitative data to identify themes related to implementation fidelity.

Statistical Analysis of Impacts

Continuous outcome measures were analyzed using repeated measures analysis of variance (ANOVA), which examined differences between groups and change over time. For categorical outcomes, chi-square tests were used. We first conducted analyses using study condition (i.e., child's enrollment in intervention or comparison site) as the independent variable; next, analyses were re-run using dosage (number of *LENA Feedback* sessions received) as the independent variable. If initial analyses showed non-significant results but descriptive statistics suggested scores were moving in a positive direction (i.e., the intervention group's scores increased more than those of the comparison group), we conducted pre/post analyses for the intervention and comparison groups separately as a follow-up (using *t*-tests and chi-square tests).

Table 8. Statistical approach by research question

Research Question	Type of Question	Analytic Strategy	Unit of Analysis
1. Do parents in the *DR+LENA Feedback* group show greater gains in their support of their child's language and literacy development than do parents in the DR Only group?	Impact: Confirmatory	*ANOVA* Descriptives *t*-tests	Parents
2. Do children in the *DR+LENA Feedback* group demonstrate greater increases in their oral language and communication skills than do children in the DR Only group?	Impact: Confirmatory	*ANOVA* Descriptives Chi-square	Children

Research Question	Type of Question	Analytic Strategy	Unit of Analysis
3. Do parents who receive more *LENA Feedback* show greater gains in their support of their child's language and literacy development than those who do not receive feedback?	Impact: Exploratory	*ANOVA* Descriptives	Parents
4. Do children whose parents receive more *LENA Feedback* demonstrate greater increases in their oral language and communication skills than children whose parents do not receive feedback?	Impact: Exploratory	*ANOVA* Descriptives Chi-square	Children
5. How many individuals were trained on *DR*?	Implementation	Descriptives	Staff, Parents, Volunteers, and Community Partners
6. How frequently did parents in the *DR+LENA Feedback* group receive *LENA Feedback*?	Implementation	Descriptives	Parents
7. What were parents' perceptions of the *LENA Feedback*? Did they find it useful? Did they understand the LENA reports?	Implementation	Descriptives	Parents
8. How many PLG and *DR* coaching sessions were received by facilitators?	Implementation	Descriptives	Facilitators

In some families, more than one child is enrolled in the study. The sample of participants (n = 210) included 32 sibling pairs and two families with three siblings. Analyses in previous years were conducted both with and without sibling pairs/groups, and results were similar, except that some significant results became non-significant when siblings were removed. We believe this may be due to decreased power. Therefore, all results presented in this report include siblings.

Power Analysis

Current total sample sizes for each impact measure range from approximately 60 to 120. According to post hoc power analyses conducted in G*Power, there is generally adequate power to detect medium or large effects in ANOVA with all sample sizes, and there is also sufficient power to detect small effects in the larger samples (see Table 9). However, power to detect small effects (as would be expected with a literacy intervention of this nature) is inadequate for measures that have smaller sample sizes. Please note that although some participants completed more than two time points (e.g., they have a baseline + two follow-ups), adding a third assessment did not increase power because few respondents had complete data (all 3 time points). Power estimates for paired samples *t*-tests are lower, ranging from .31 (N = 60 and a small effect) to 1.00 (N = 120 and a large effect).

Table 9. Post hoc power analysis for repeated measures ANOVA

N	Timepoints	Groups	Power to detect large effect (.40)	Power to detect medium effect (.25)	Power to detect small effect (.15)
60	2	2	.99	.97	.62
80	2	2	.99	.99	.75
100	2	2	1.00	.99	.84
120	2	2	1.00	.99	.93

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

Study attrition and non-completion of study measures by method of data collection are the primary reasons for missing data and have been described previously. Measure- and item-specific missingness is described below.

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 the study, this included a total of 20 baseline assessments and 16 Follow-up 1 assessments. 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. If there are more than two unanswered questions, the assessment is not scored (n = 1).

CDI Words Produced. There are 3 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 six baseline assessments and 14 Follow-up 1 assessments collected during the study were out of range. Because this measure is obtained by totaling the number of words the parent indicated the child understands and says, item-level missingness is not applicable.

Parent survey. Items included in the analyses had very little missing data (< 2% of responses). For the parent survey of home literacy, which is scored by summing responses to seven *yes/no* items, some participants (n = 15) provided no response to one or more items. After examining those cases, it was determined that non-response to those items should be treated as a "no," and previously summed scores were retained.

LENA. According to developers, LENA recordings must be at least 10 hours in duration in order to obtain usable data. Between 2012 and 2017, a total of 53 LENAs had less than 10 hours of recording and were excluded from analyses; this includes 14 LENAs from the 2016-2017 school year and 39 from previous years. Evaluators and PLG facilitators review LENA recordings and request that families repeat the LENA recording if it is too short; however, sometimes it is not possible to obtain a usable LENA recording even after multiple attempts. Because LENA measures are automatically generated by LENA software, item-level missingness is not applicable.

Implementation Findings

Program Exposure and Dosage

***LENA Feedback* dosage.** A total of 138 PLG participants completed at least one *LENA Feedback* survey since January 2013. According to the surveys, participants received an average of about 3 feedback sessions (Table 10). Study participants reported that feedback lasted between 16 minutes and 21 minutes, on average, across sites. Notably, earlier in the study (Years 1-3), *LENA Feedback* ranged from 12-24 minutes. Therefore, these results suggest that PLG facilitators "tightened up" the duration of *LENA Feedback* sessions to last about 15-20 minutes, versus much shorter or longer sessions.

Research Question

How frequently did parents in the DR+LENA group receive LENA Feedback?

Table 10. LENA Feedback surveys by site

Site	Total number of respondents	Average number of feedback surveys completed per participant	Average duration of feedback
Clayton	34	3.0	21 minutes
Green Valley Ranch	50	3.0	19 minutes
Southwest	54	2.7	16 minutes

Table 11 presents the number of *LENA Feedback* sessions delivered at the time of study participants' Follow-up 1 assessments. By the first follow-up assessment (approximately six months after the baseline assessment,) 75% of intervention group parents had received one or more *LENA Feedback* sessions. Although these results show some participants not having received *LENA Feedback* at follow-up, it is important to note that *LENA Feedback* dosage is obtained from the feedback satisfaction surveys. Therefore, more parents may have received *LENA Feedback* than is indicated by having a completed survey.

Table 11. LENA Feedback dosage at the time of study assessments for intervention group

	Follow-up 1
Dosage at time **assessment** was completed[1]	
0 *LENA Feedback* sessions	25.4% (17)
1 *LENA Feedback* session	40.3% (27)
2+ *LENA Feedback* sessions	34.3% (23)

[1]Based on dates of follow-up parent surveys

PLG Program Attendance & Enrollment Duration. PLGs are held weekly throughout the school year, and parents can attend PLGs until their children turn age 3. Table 12 presents the average number of PLGs attended by study participants in the intervention and comparison group conditions. Please note that the tracking database for attendance was implemented in the 2013-14 school year; thus, attendance is unavailable for some participants who enrolled in 2012-13, particularly for those attending comparison sites.

Intervention participants attended about eight PLG sessions on average by the time baseline data were collected, approximately one month after program enrollment. Attendance varied considerably in the intervention group, the upper range of which is in part due to some participants having attended the PLG before the start of the Ready to Read study. Follow-up assessments for intervention participants occurred after about 41 PLG sessions. Comparison group participants completed baselines after attending approximately 5 sessions, and follow-ups occurred after about 34 sessions.

Table 12. Average and range of parent attendance at Play and Learn Groups

	Intervention		Comparison	
	Baseline	Follow-up 1	Baseline	Follow-up 1
Attendance at time of assessment	8.04 (0-33) *n* = 108	41.20 (14-83) *n* = 64	5.06 (1-18) *n* = 53	34.41 (13-58) *n* = 29

[1]Based on dates of baseline and follow-up parent surveys

Data for all study participants (shown above) suggest that the intervention group participants generally attended more sessions than did members of the comparison group. Bivariate comparisons among participants with both baseline and follow-up data for each measure indicated there was a statistically significant difference in the number of sessions attended by study condition at the time of the baseline LENA, $t(94.19) = 2.81$, $p = .01$, Follow-up 1 survey, $t(91) = 2.17$, $p = .03$, and Follow-up 1 ASQ-C, $t(56) = 2.04$, $p = .05$. That is, intervention parents attended more sessions than did comparison parents at the time of those assessments. Overall, there is considerable variation in PLG attendance, both within and between the intervention and comparison groups.

We also examined PLG enrollment duration. Table 13 shows the length of participation for those who enrolled during previous school years. These data show that about one-half of children (52%) remained in the program for at least one

year, 33% were enrolled for 6-11 months, and 19% were in the program for less than 6 months. Of the 40 participants who enrolled in the study during the 2016-17 school year, 77% remained enrolled through May 2017 (the end of regular PLG programming).

Table 13. PLG enrollment duration, school years 2012-13 through 2015-16

	n	%
Previous school years (2012-13 through 2015-16)		
Less than 6 months	31	19
6-11 months	55	33
1-2 years	50	30
More than 2 years	31	18

Participant Satisfaction

Research Question
What were parents' perceptions of the LENA Feedback?

As shown in Table 14, respondents reported high levels of satisfaction with *LENA Feedback*, with mean scores of 6 or higher on the 7-point rating scale. Notably, 87% of respondents reported they were "very likely" to change their behavior as a result of the *LENA Feedback*.

Table 14. Satisfaction with LENA Feedback (n =401)

	Mean	Std. Dev.
How helpful is this feedback?	6.89	0.47
How easy to understand is this feedback?	6.80	0.55
How likely are you to change your behavior as a result of this feedback?	6.83	0.52

Ready to Read Training and Coaching

Research Question
How many individuals were trained on DR?

Since the start of the Ready to Read study, a total of 1,009 PLG staff, parents, and community members have been trained in *DR*, according to program records (Table 15).

Table 15. How many individuals have been trained in Dialogic Reading? September 2012-May 2017

Trainee	***N***
Staff	8
Parents	389
Volunteers*	252
Community partners*	360
Total	**1,009**

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

During the 2016-17 school year, PLG facilitators received coaching sessions as-needed during the school year (Table 16). Coaching consists of a cycle with a needs assessment (video or written observation), feedback, and goal-setting. "On the spot" coaching also occurs while the coach is visiting a group. During this year, peer-to-peer coaching also occurred among PLG facilitators. *DR* trainings for facilitators were not held because all facilitators had been trained in *DR* during previous school years.

Families received *DR* coaching and *LENA Feedback* on an ongoing basis during most months of the school year. *DR* training for parents was offered at the beginning, middle, and end of the school year, and *DR* trainings for volunteers

and community partners were offered once during the 2016-17 school year. Most events in which community members and volunteers were trained took place in years 1-4 of the program.

Table 16. Training and coaching provided to PLG staff, families, parents, and volunteers by month (September 2016-May 2017)

Participant	2016-2017 School Year								
	Sep.	Oct.	Nov.	Dec.	Jan.	Feb.	Mar.	Apr.	May
Facilitators									
PLG coaching		x	x		x	x			
***DR* training**									
Families									
***DR* training**		x				x			x
***DR* coaching**		x	x	x	x	x	x	x	x
LENA Feedback		x	x	x	x	x	x	x	
Volunteers & community partners									
DR* training							x	x	

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

Impact Findings

Parent-level Impact

Impact Research Question 1 (Confirmatory): *Do parents in the DR+LENA group show greater gains in their support of their child's language and literacy development than do parents in the DR only group?*

Results indicate that, overall, Ready to Read parents increased their support of children's language and literacy development, regardless of intervention condition:

(1) The frequency and quality of parent-child reading increased significantly between baseline and follow-up.

(2) Parents and children engaged in significantly more conversational turns over time.

Results also provide some support for the effectiveness of the *LENA Feedback* intervention:

(1) Intervention parents showed statistically significantly greater growth in reading frequency than did comparison parents.

(2) Intervention parents showed statistically significant increases in use of interactive reading techniques and storytelling frequency while the comparison group's scores remained consistent over time.

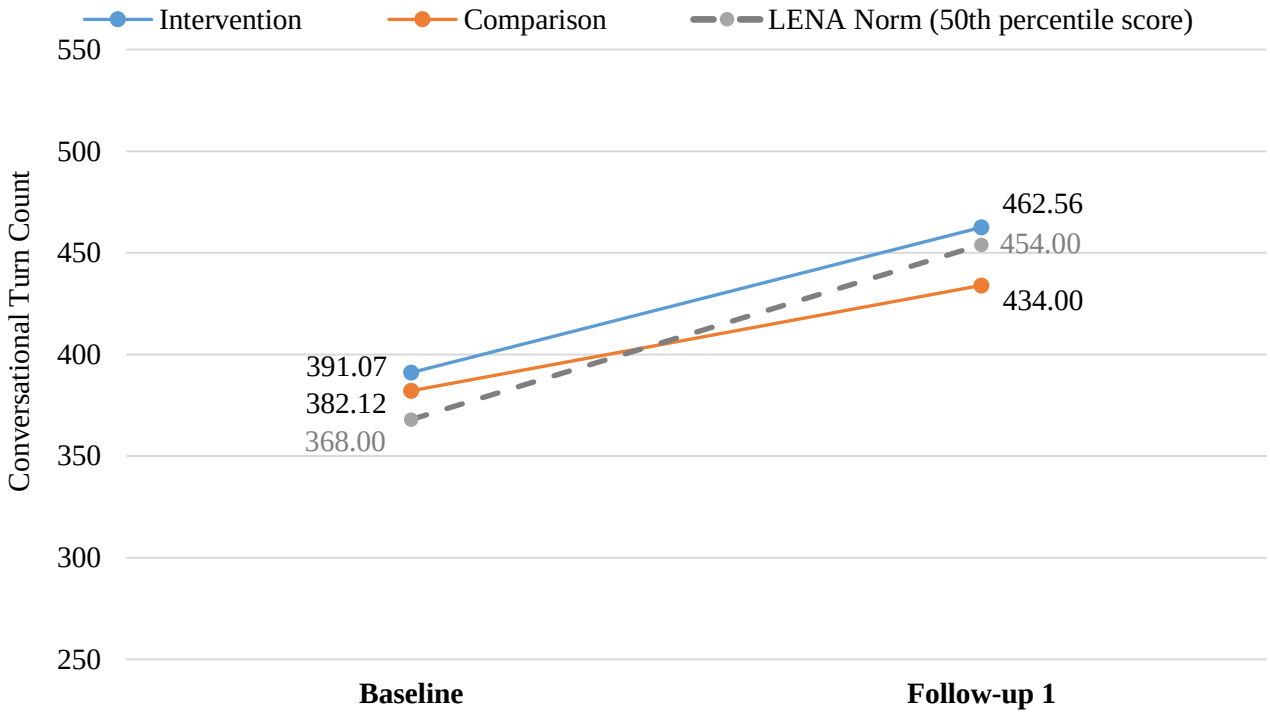
Analysis approach. To answer this research question, we conducted repeated measures ANOVAs to analyze group differences and changes over time for the following outcomes: LENA conversational turns (prevalence of parent-child verbal interactions); Parent Survey of Home Literacy total scores (parent-reported use of 7 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 as a follow-up analysis. Results are summarized below.

LENA Conversational Turns. We conducted a 2 (group) x 2 (time) repeated measures ANOVA and found no statistically significant effect for the time x group interaction, $F(1, 112) = 0.20$, $p = .65$, $\eta p^2 = .002$, indicating similar rates of change over time in use of conversational turns by group. There was, however, a statistically significant main effect of time, $F(1, 112) = 7.98$, $p = .01$, $\eta p^2 = .07$. This means that the number of parent-child conversational turns increased across all study participants (irrespective of group membership); the size of this effect, however, was "small."

Take-away:
Conversational turns increased significantly across all study participants between baseline and follow-up.

Figure 2 displays mean scores for conversational turns by time point and group; this figure also provides scores for the 50th percentile of the LENA Research Foundation's (n.d.) national normative sample.* As shown, these groups engaged in roughly 75 more conversational turns between baseline and follow-up assessments. All Ready to Read families were above the LENA norm at baseline (their scores were above the 50th percentile); however, by the first follow-up assessment, the comparison group fell slightly below the LENA norm, while the intervention group remained above the norm. Overall, this indicates that parent-child vocal interactions are increasing across all study participants and are generally on par with LENA norms.

Figure 2. LENA Conversational Turns: Baseline and Follow-up 1 by Condition (and LENA Norms) (*n*=114)



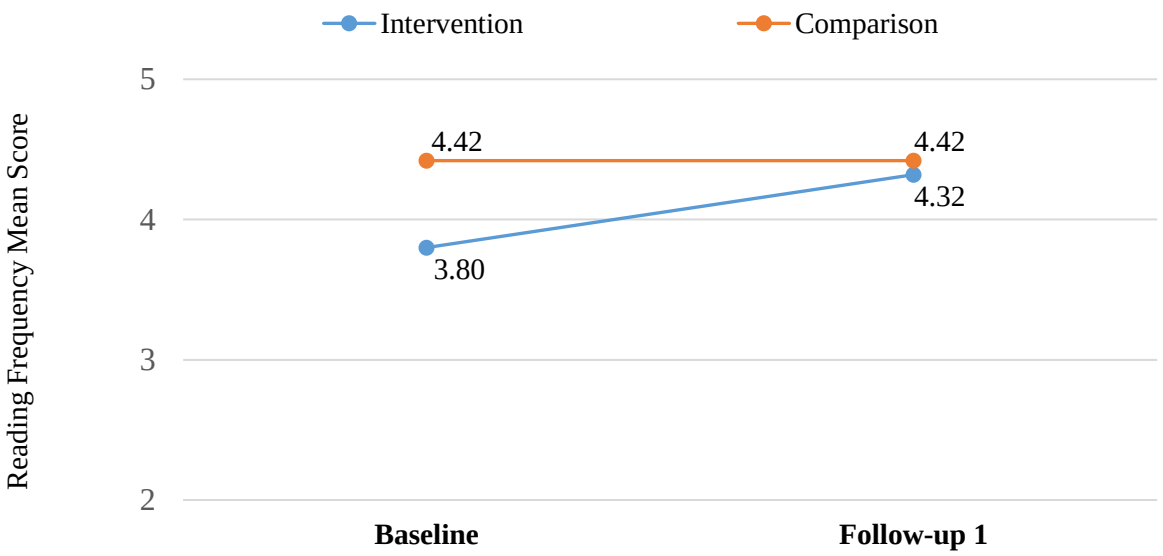
*Baseline: average child age is 16.29 months; LENA norms based on daily conversational turn count for the 50th percentile of 16 month olds. Follow-up 1: average child age is 21.91 months; LENA norms based on daily conversational count for the 50th percentile of 22 month olds.

Reading Frequency. Parent-reported reading frequency data were analyzed using a 2 (group) x 2 (time) repeated measures ANOVA. We found a statistically significant time x group interaction, indicating that intervention and comparison group participants had different rates of progress with respect to reading frequency, $F(1, 94) = 5.50$, $p = .02$, $\eta p^2 = .05$. Parents in the intervention group had a greater increase in their reading frequency scores between the baseline and Follow-up 1 surveys than did the comparison group. The effect size was "small."

Take-away:
Intervention parents showed significantly greater growth in reading frequency than did comparison parents.

Figure 3 shows mean reading frequency scores by time point for the intervention and comparison groups. At baseline, intervention parents had a mean score of 3.80; this roughly corresponds to *daily* reading. At Follow-up 1, intervention parents' mean score increased to 4.42, which indicates that parents were reading *daily* or *more than once a day*. Comparison group parents reported reading *daily* or *more than once a day* at baseline, and their reading frequency scores remained relatively flat over time. Notably, the intervention group, which consists of a large proportion of families who speak Spanish or are bilingual, had statistically significantly lower reading frequency scores at program enrollment than did the comparison group. This suggests that these parents could use some support with reading with their children. Results indicate that the *LENA Feedback* intervention may have had an effect on the frequency with which parents report reading with their children. Although intervention parents' scores were initially lower than those of comparison parents, *LENA Feedback* seems to have helped them "catch up" with the comparison group.

Figure 3. Reading Frequency: Baseline and Follow-up 1 by Condition (*n* = 96)



Storytelling Frequency. Storytelling frequency scores were analyzed using a 2 (group) x 2 (time) repeated measures ANOVA. The time x group interaction was not statistically significant, indicating that the frequency with which parents told stories with their children did not change differentially over time based on assignment to an intervention or comparison site. There was, however, a statistically significant main effect of time, $F(1, 95) = 5.72$, $p = .02$, $\eta_p^2 = .06$, which means that there was overall increase in scores over time.

Take-away: Overall

Storytelling frequency increased significantly across all study participants between baseline and follow-up.

Figure 4 shows mean scores for storytelling frequency by time point and group. As shown, scores for both groups increased over time, and the intervention group made a slightly greater gain than did the comparison group. At baseline, the intervention group's mean score ($M = 3.20$) roughly corresponds with telling stories *a few times a week*, while the follow-up score ($M = 3.79$) indicates parents engaged in storytelling closer to *once a day*, on average. At the time of the follow-up, both groups reported reading an average of about *once a day*.

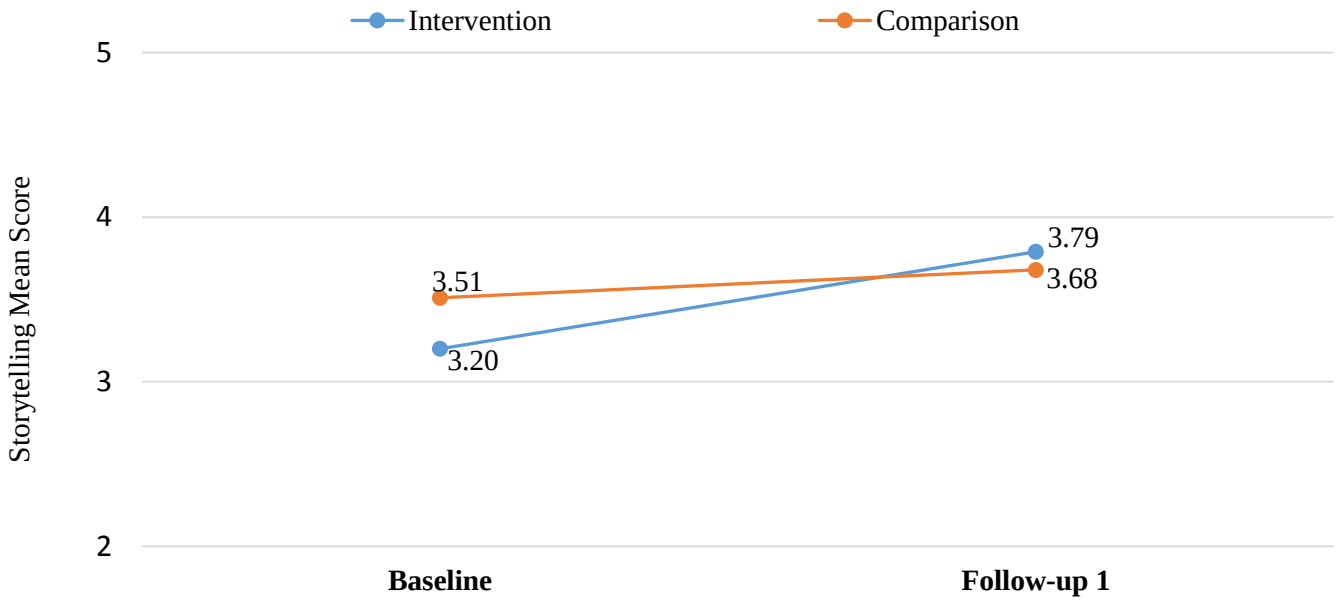
To further examine change over time by group, we conducted paired samples *t*-tests for each group separately. The intervention group made a statistically significant increase in storytelling frequency, $t(64) = -3.22$, $p < .01$; this represents a "medium-sized" effect ($d = .41$). The comparison group, however, did not have a statically significant change in scores over time, $t(30) = -.82$, $p = .42$, $d = .18$. Thus, *t*-test results suggest that intervention group parents engaged in storytelling more frequently after having received *LENA Feedback*, while the scores of parents in the comparison group did not change significantly. It is important to note, however, that the sample size for the comparison group is small ($n = 31$), which limits our ability to detect statistically significant change over time. Overall, results suggest the *LENA Feedback* intervention encourages families to tell stories together more frequently.

Take-away: By Group

Intervention group: Increased significantly

Comparison group: No significant increase

Figure 4. Storytelling Frequency: Baseline and Follow-up 1 by Condition (*n* = 97)



Parent Survey of Home Literacy. Parent Survey of Home Literacy scores (parent-reported number of interactive reading techniques used) were analyzed using a 2 (group) x 2 (time) repeated measures ANOVA. The time x group interaction was not statistically significant, indicating similar rates of change over time in use of reading techniques between study conditions. There was, however, a statistically significant main effect of time, $F(1, 94) = 7.92$, $p = .01$, $\eta_p^2 = .06$, which means that scores increased significantly over time (irrespective of group membership).

Take-away: Overall

Parent-reported interactive reading increased significantly across all study participants between baseline and follow-up.

Figure 5 shows mean scores for use of parent-reported interactive reading by time point and group. As shown, scores for both groups increased over time, but the intervention group appears to have made a greater gain than did the comparison group; that is, comparison group parents reported using about 5.5 of the interactive techniques at both baseline and follow-up, while intervention group parents used about 5.0 techniques at baseline and 5.5 at follow-up.

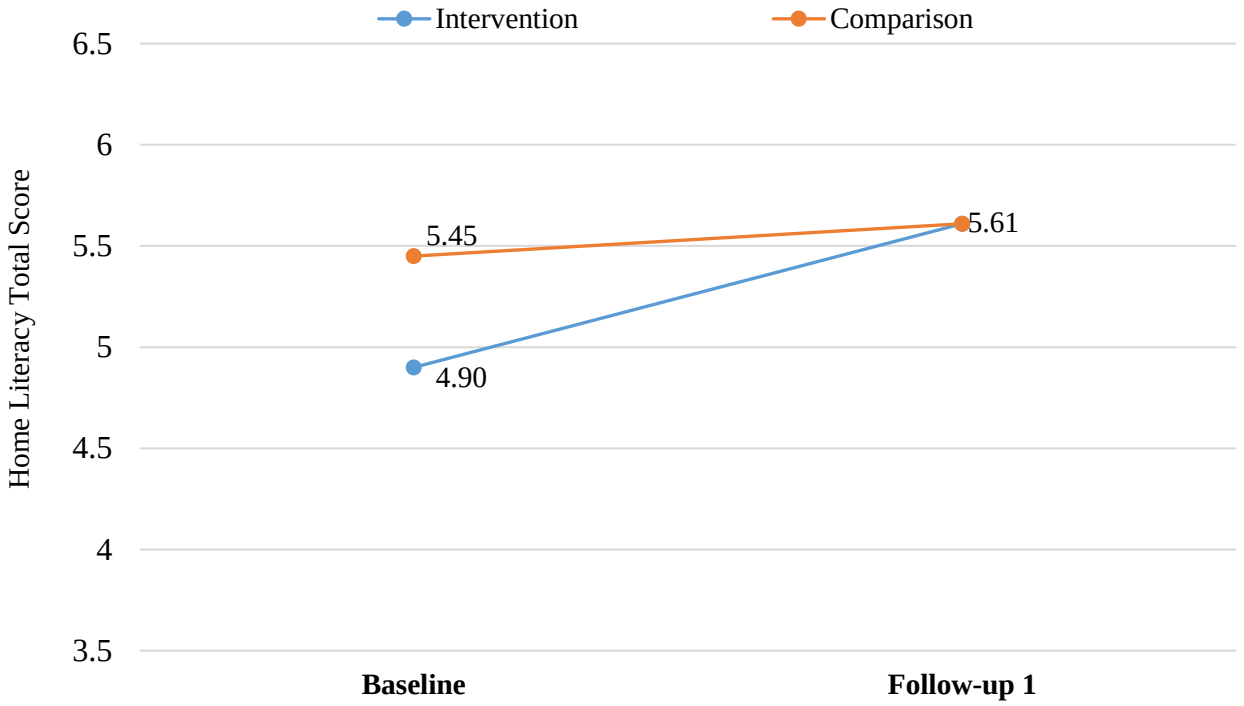
Based on these trends in the descriptive data, we conducted paired samples *t*-tests for each group separately. The intervention group had a statistically significant increase in interactive reading, $t(64) = -3.22$, $p < .01$; this represents a "medium-sized" effect ($d = .40$). The comparison group, however, did not show a statically significant change in scores over time, $t(30) = -.82$, $p = .42$, $d = .15$. Although it is possible that non-significant findings for the comparison group are due to the small sample size ($n = 31$), these results suggest that intervention group parents experienced growth in their use of interactive reading techniques after having received *LENA Feedback*, while the scores of parents in the comparison group remained stable over time.

Take-away: By Group

Intervention group: Increased significantly

Comparison group: No significant increase

Figure 5. Home Literacy (Interactive Reading Techniques): Baseline and Follow-up 1 by Condition (*n* = 96)



Impact Research Question 3 (Exploratory): *Do parents who receive more LENA Feedback show greater gains in their support of their child's language and literacy development than those who do not receive feedback?*

To address this exploratory question, we examined changes in parent-level outcomes over time based on *LENA Feedback* dosage categorized into four levels: 0 feedback sessions (comparison group); 0 feedback sessions (intervention group), 1 feedback session, and 2 or more feedback sessions. The purpose of this is to supplement the intent-to-treat (ITT) analyses (based purely on assignment to intervention or comparison group) with analyses by dosage of the intervention received (if any) at the time of the Follow-up 1 assessment. This involved conducting the ANOVAs as described in the previous section with dosage as the independent variable for each parent-level outcome studied. As expected, results using dosage were similar to those conducted by group and are provided in Appendix 1. However, analyses of reading frequency provided additional information about the impact of *LENA Feedback*, and those results are presented below.

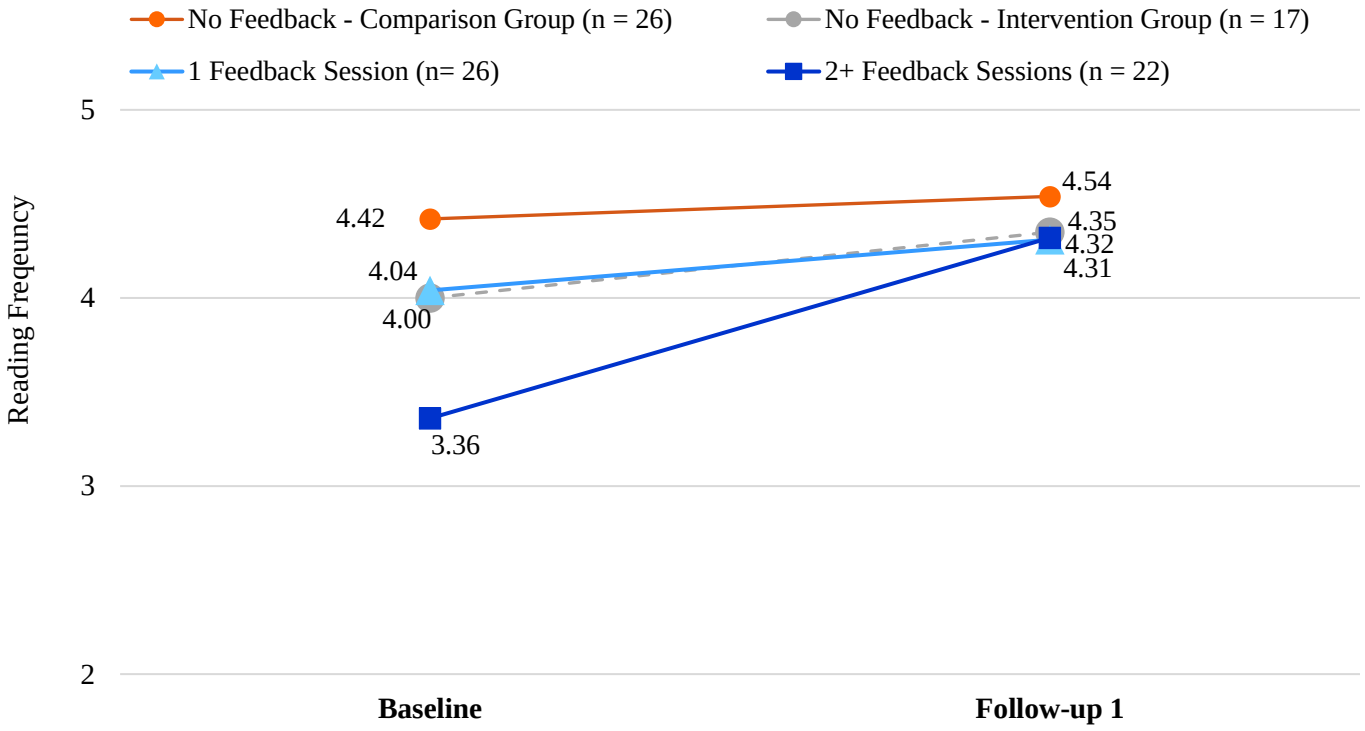
Reading frequency. We conducted a 4 (dosage level) x 2 (time) repeated measures ANOVA. We found a statistically significant time x group interaction, indicating that participants who received different levels of LENA dosage had different rates of progress with respect to reading frequency, $F(1, 87) = 3.63$, $p = .02$, $\eta_p^2 = .11$. Post-hoc comparisons showed a statistically significant difference in reading frequency scores between comparison group participants who did not get feedback and intervention group participants who received two or more feedback sessions ($p = .02$).

Take-away: Reading frequency by dosage level

Parents who received 2+ feedback sessions made significantly greater gains than did comparison group parents who did not receive feedback.

As shown in Figure 6, parents who eventually received two or more *LENA Feedback* sessions had the lowest scores at baseline, but they made the greatest gains over time. Therefore, it seems that *LENA Feedback* was able to "close the gap" in reading frequency across families. These results also indicate that intervention group parents who read with their child less often at baseline received more *LENA Feedback* than did parents who read with their child less frequently. One possible explanation for this is that these parents felt that they needed more help to support their child's language and literacy development and therefore took advantage of opportunities to complete a LENA and receive feedback more frequently. Overall, these results also suggest that multiple *LENA Feedback* sessions are optimal for increasing parent-child reading frequency and also that the intervention had an impact on families most in need of support.

Figure 6. Reading Frequency: Baseline and Follow-up 1 by Dosage (*n*=91)



Child-level Impact

Impact Research Question 2: *Do children in the DR+LENA group demonstrate greater increases in their oral language and communication skills than do children in the DR only group?*

Results indicate that, overall, children's language and literacy skills increased during their participation in Ready to Read, regardless of intervention condition:

- Children could understand and say significantly more words between baseline and follow-up.

Results also provide some anecdotal support for program impact:

- A slightly greater proportion of intervention children were above the ASQ cutoff at follow-up compared with the comparison group (91% versus 87%, respectively).

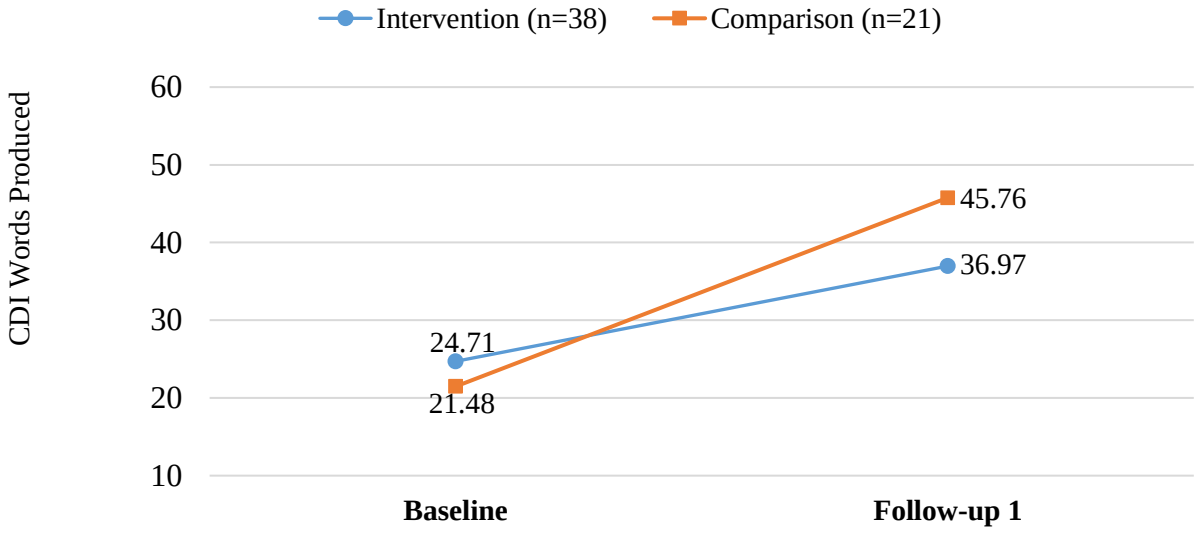
Analysis approach. To answer this research question related to children's language and literacy development, we conducted repeated measures ANOVAs to analyze group differences and changes over time for the CDI Words Produced outcome and conducted a chi-square test for ASQ-C categorical scores. Results are summarized below.

CDI Words Produced. CDI production scores were analyzed using a 2 (group) x 2 (time) repeated measures ANOVA. The time x group interaction was not statistically significant, which indicates similar rates of change over time between the intervention and comparison groups. There was, however, a statistically significant main effect of time, $F(1, 57) = 21.17$, $p < .001$, $\eta_p^2 = .34$, or a medium-sized increase over time. As shown in Figure 7, the comparison group made a slightly greater gains in words produced than did the comparison group, but both groups' scores increased over time.

Take-away:
Words produced increased significantly across all children between baseline and follow-up.

Notably, most intervention families speak Spanish or are bilingual, while comparison families are more likely to speak English only. Because some bilingual children temporarily have smaller vocabularies than monolingual children as they acquire language skills, the lack of group equivalence could account for this finding (Zero to Three, 2016). In addition, the CDI vocabulary checklist is available in English or Spanish, but does not specify how to answer if children say some words in English and others in Spanish, which could lead to inconsistencies in how bilingual families complete the assessment.

Figure 7. CDI Words Produced: Mean Scores by Condition and Time Point

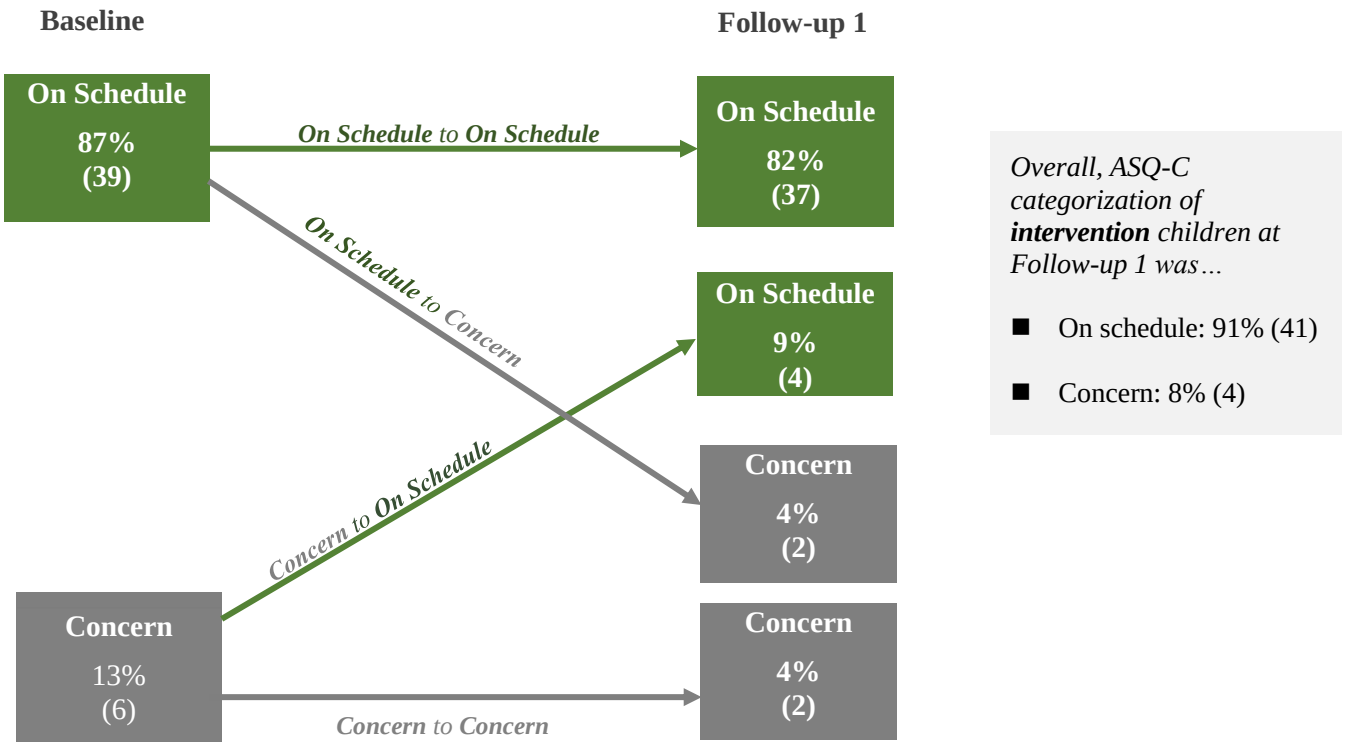


ASQ-C. The ASQ-C measure assigns a child to one of three categories related to development in verbal and nonverbal communication skills as "on schedule," "close to the cutoff," or "below the cutoff." To compare changes in children's categorizations over time and to determine whether there are differences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

Results of a Fisher's Exact chi-square test showed no statistically significant difference between groups regarding changes in ASQ-C scores, $p = .66$. This means there is no more variation in classification of children by study condition than we would expect due to chance. However, analyses conducted separately by group showed statistically significant changes over time. These are described below.

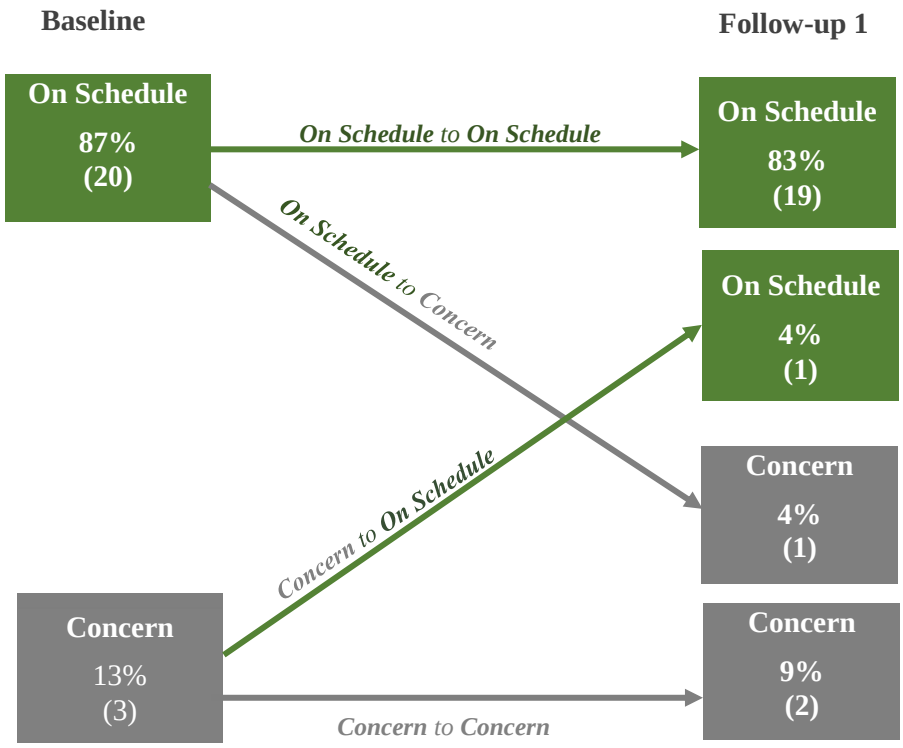
Intervention group. We conducted a chi-square test to assess change over time in ASQ-C classification among intervention group children specifically. Results showed "marginally" significant change over time ($p = .08$) -- the proportion of intervention children who were "on schedule" increased from 87% to 91% (see Figure 8).

Figure 8. Intervention children by ASQ-C categorization (concern/on schedule) and Time Point (*n*=45)



Comparison group. A chi-square test was also conducted to assess change over time in ASQ-C classification among comparison group children. Results showed statistically significant changes in classification (p = .03) – this involved two children "switching" categories between baseline and follow-up (Figure 9).

Figure 9. Comparison children by ASQ-C categorization (concern/on schedule) and Time Point (*n*=23)



Overall, ASQ-C categorization of ***comparison*** *children at Follow-up 1 was…*

- On schedule: 87% (20)
- Concern: 13% (3)

Results suggest that most children's ASQ-C scores remain *on schedule* during their participation in Ready to Read, with a few children per group switching classifications over time. At follow-up, a slightly larger proportion of intervention children versus comparison children were *on schedule* (91% versus 87%, respectively), though there is not a statistically significant difference in change *over time by group*. It is likely that children's ASQ-C classifications may be related to factors other than the *LENA Feedback* intervention.

Impact Research Question 4 (Exploratory): *Do children whose parents receive LENA Feedback demonstrate greater increases in their oral language and communication skills than children whose parents do not receive feedback?*

To address this exploratory question, we examined changes in child-level outcomes over time based on *LENA Feedback* dosage categorized into four levels: 0 feedback sessions (comparison group); 0 feedback sessions (intervention group), 1 feedback session, and 2 or more feedback sessions (as described in the parent-level results section). For child-level outcomes, this involved conducting the repeated measures ANOVAs and chi-square tests described previously using dosage as the independent variable. Results were non-significant and similar to those of the ITT analyses and are provided in Appendix B.

CONCLUSION: FINDINGS, LESSONS LEARNED, AND NEXT STEPS

Summary of Parent-Level Impact

Results from the parent impact analyses indicate that PLG parents increased their support of children's language and literacy development. Scores for all parent-level measures -- parent-child conversational turns, as well as parent-reported use of interactive reading behaviors, reading frequency, and storytelling frequency – went up significantly over time across the whole PLG sample.

Outcome	Evidence of Impact	Effect Size
Reading Frequency	• Statistically significant difference between intervention and comparison groups' growth over time • Greatest gains among those who received 2+ *LENA Feedback* sessions	Small
Storytelling Frequency	• Intervention: Statistically significant increase over time • Comparison: No significant increase over time	Medium
Interactive Reading Behaviors *(parent report)*	• Intervention: Statistically significant increase over time • Comparison: No significant increase over time	Medium

Group x time comparisons. The main analyses (2 x 2 ANOVAs) examined whether intervention group parents showed *greater* gains over time than did the comparison group. **Intervention group parents reported increasing how often they read with their children – more so than the comparison group.** Thus, it appears that the *LENA Feedback* intervention increased parent-child reading frequency. The size of this effect on reading frequency is "small," indicating that *LENA Feedback* related to the importance of talking with children may have a small influence on parent-child reading frequency.

Individual group analyses. For some parent-level outcomes, 2 x 2 ANOVAs showed no statistically significant differences between the groups' growth over time. However, descriptive statistics indicated that scores increased for the intervention group, while scores remained stable among children in the comparison group. Thus, we analyzed change over time for the intervention and comparison groups separately (using paired samples *t*-tests). **Intervention parents engaged in statistically significantly more frequent storytelling and interactive reading between baseline and follow-up; these results represented medium-sized effects, suggesting not only statistical, but practical, significance**. By contrast, storytelling and interactive reading did not change significantly over time for the comparison group. Thus, assignment to an intervention (*DR + LENA Feedback*) PLG may contribute to gains in parents' support of language and literacy that did not occur for comparison site (*DR* only) parents. It is important to note, however, that non-significant results for the comparison group could be a function of the small sample size. With a larger sample, we might see significant change over time for both the intervention and comparison groups, suggesting that factors other than the intervention (e.g., PLG "programming as usual," or child maturation) account for the growth over time.

Analyses by dosage. Parents who received two or more *LENA Feedback* sessions experience statistically significantly more growth than did comparison group parents who did not receive feedback. By contrast, there was no significant difference in the rate of change between parents who only received one *LENA Feedback* session and comparison group parents. Thus, it appears that multiple *LENA Feedback* sessions are needed to make an impact on parent-child reading frequency.

Overall, results indicate some preliminary support for the effectiveness of the *LENA Feedback* intervention due to its differentially positive impact on the support parents in the intervention group provide to their children in storytelling frequency, the use of interactive reading strategies, and reading frequency (particularly for those parents who received

more *LENA Feedback*). Although most analyses in which direct comparison of groups were made did not provide evidence that the LENA intervention has impacted parent-level outcomes more than PLG "programming as usual" (with the exception of reading frequency), lack of power to detect effects is a likely culprit. Most outcome measures had a relatively small number of participants who had data available at two or more timepoints, and this is particularly true of the comparison group. However, when analyzed separately, to examine change over time, parents in the intervention group showed significant increases from baseline to follow-up where comparison group participants' scores remained relatively flat. Based on the results of this study the relationship of between the *LENA Feedback* intervention and parents' support of their children's language and literacy development is somewhat tenuous but seems to be stronger for those who have the greatest room for growth.

Summary of Child-Level Impact

Overall, children's vocabularies, as measured by CDI Words Produced, increased during their participation in the study, and according to the ASQ-C, most children were "on track" in terms of developing communication skills at both baseline and follow-up assessments. However, we did not find statistically significant differences between the intervention and comparison groups in children's rates of growth with respect to these outcomes.

One possible explanation for these findings is that *LENA Feedback* provided to parents did not transfer into improved vocabularies and communication skills for their children. However, there are several unique aspects of these data sources that could also explain these results. First, the CDI Words Produced measure is perhaps the outcome most influenced by the lack of group equivalence on the basis of language, since there may be some short-term differences in how bilingual children develop language skills versus monolingual children. Also, it is important to consider that the ASQ-C is essentially a screening tool to identify possible developmental delays. Given that children with developmental concerns may require fairly intensive early intervention services, it may not be possible for *LENA Feedback* to provide the support necessary to change children's developmental status from the "concern" category to being "on schedule." Moreover, there is very little variation in Ready to Read children's ASQ-C scores – almost all children remain "on schedule" throughout their participation in the study, so there is little "room for improvement" on this measure.

Overall, results indicate no support for the effectiveness of the *LENA Feedback* intervention for increasing children's vocabulary and communication skills. It is possible that the small influence it may have on parents' support of children's language and literacy development is simply too distal or small of an impact to reach children or it may take longer to manifest itself in improved child outcomes than could be captured with available data.

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 high 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. Although this study was carefully designed in attempt to isolate the impact of *LENA Feedback*, program leaders believe that the overall emphasis on early language and literacy at PLGs may have influenced parent and child outcomes, which could explain why we observed gains across all study participants. Furthermore, because all participants received support through PLG "programming as usual," as well as *DR*, this made it difficult to detect an effect for the relatively small program component of *LENA Feedback.*

A factor which we believe facilitated the data collection process was recruiting and retaining a bi-lingual data collector who was able to build trust and rapport with both the PLG families and facilitators. In addition, the PLG facilitators proved to be an invaluable source of support for the study in a variety of ways, such as by telling families about the study, handling the distribution of LENAs in most sites, and maintaining program records. Consistent and high-quality coaching support provided to PLG facilitators by a project leader also appeared to be a key factor contributing to successful implementation of this project.

Next Steps

Based on the results of this study, *LENA Feedback* may be too expensive to continue to implement, considering its relatively small impact. Future opportunities to support families' early language and literacy skills could include identifying another type of individualized data (e.g., parent-child observations) that could be used during one-on-one coaching sessions with families in a simpler and less expensive manner. Based upon the forthcoming results of the Ready to Ready study in center-based settings, which tests the impact of *Cradling Literacy* training and coaching with teachers, the project team could also consider adapting *Cradling Literacy* for use with PLG families. A next step for the PLG program includes expanding PLGs to additional communities (e.g., rural areas).

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.

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

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

Gilkerson, J. & Richards, J. (2009). The Power Talk—2nd Edition—Impact of Adult Talk, Conversational Turns, and TV During the Critical 0-4 Years of Child Development; LENA Foundation Technical Report LTR-01-2.

Hart, B. and 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

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).

LENA Research Foundation. (n.d.). [LENA percentile data] Unpublished raw data.

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

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

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. https://doi.org/10.1093/biomet/70.1.41

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, Maryland: 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. doi: 10.1037/a0018719

U.S. 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

Xu, D., Yapanel, U., & Gray, S. (2009). Reliability of the LENA Language Environment Analysis System in Young Children's Natural Home Environment. Boulder, CO: LENA Foundation. Retrieved from http://www.lenafoundation.org/wp-content/uploads/2014/10/LTR-05-2_Reliability.pdf

Zero to Three. (2016). *Dual Language Development: Double the Benefit.* Retrieved from https://www.zerotothree.org/resources/303-dual-language-development-double-the-benefit

APPENDIX A. PARENT-LEVEL IMPACT ANALYSES BY DOSAGE

Measure	ANOVA Results	Mean score by time point and group
LENA: Conversational Turns	Time: $F(1, 109) = 8.94$, $p < .01$, $\eta_p^2 = .08$ Time x Group: $F(3, 109) = 0.60$, $p = .62$, $\eta_p^2 = .02$	Baseline No Feedback: Comparison (n = 33): 388.84 (170.32) Baseline No Feedback: Intervention (n = 19): 435.05 (163.28) Baseline 1 Session (n = 38): 376.50 (187.93) Baseline 2+ Sessions (n = 23): 378.83 (185.28) Post 1 No Feedback: Comparison: 440.93 (265.31) Post 1 No Feedback: Intervention: 453.58 (246.27) Post 1: 1 Session: 457.92 (238.24) Post 1: 2+ Sessions: 477.65 (239.65)
Home Literacy	Time: $F(1, 87) = 10.70$, $p < .01$, $\eta_p^2 = .11$ Time x Group: $F(3, 87) = 1.24$, $p = .30$, $\eta_p^2 = .04$	Baseline No Feedback: Comparison (n = 26): 5.42 (.99) Baseline No Feedback: Intervention (n = 17): 4.70 (1.69) Baseline 1 Session (n = 25): 4.64 (1.35) Baseline 2+ Sessions (n = 23): 5.35 (1.49) Post 1 No Feedback: Comparison: 5.58 (1.24) Post 1 No Feedback: Intervention: 5.35 (1.17) Post 1: 1 Session: 5.64 (1.35) Post 1: 2+ Sessions: 5.78 (1.00)
Storytelling Frequency	Time: $F(1, 88) = 9.67$, $p < .01$, $\eta_p^2 = .10$ Group: n.s. Time x Group: $F(3, 88) = 1.42$, $p = .24$, $\eta_p^2 = .04$	Baseline No Feedback: Comparison (n = 26): 3.54 (1.39) Baseline No Feedback: Intervention (n = 17): 3.41 (1.46) Baseline 1 Session (n = 26): 3.08 (1.41) Baseline 2+ Sessions (n = 23): 3.17 (1.30) Post 1 No Feedback: Comparison: 3.81 (.75) Post 1 No Feedback: Intervention:3.47 (1.33) Post 1: 1 Session: 3.81 (1.13) Post 1: 2+ Sessions: 4.00 (1.00)

APPENDIX B. CHILD-LEVEL IMPACT ANALYSES BY DOSAGE

Measure	ANOVA Results	Mean score (standard deviation) by time point and group
CDI	Time: $F(1, 50) = 23.70$, $p < .001$, $\eta_p^2 = .32$ Time x Group: $F(1, 3) = 0.84$, $p = .48$, $\eta_p^2 = .05$	Baseline No Feedback: Comparison (n = 18): 18.22 (18.66) Baseline No Feedback: Intervention (n = 11): 17.00 (15.38) Baseline 1 Session (n = 15): 25.33 (31.17) Baseline 2+ Sessions (n = 10): 19.30 (27.91) Post 1 No Feedback: Comparison: 41.56 (30.00) Post 1 No Feedback: Intervention: 35.55 (29.30) Post 1: 1 Session: 35.07 (28.67) Post 1: 2+ Sessions: 35.30 (23.06)

ASQ-C categorization (concern/on schedule) by Time Point

	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
0 feedback sessions (comparison group) (n = 23)*	9% (2)	4% (1)	4% (1)	83% (19)
0 feedback sessions (intervention group) (n = 14)	14% (2)	14% (2)	7% (1)	64% (9)
1 feedback session (n = 15)	--	7% (1)	7% (1)	86% (13)
2 + feedback sessions (n = 16)	--	6% (1)	--	94% (15)

* $p < .05$