

A Study of Classroom Literacy Interventions and Outcomes in Even Start

A Study of Classroom Literacy Interventions and Outcomes in Even Start

SEPTEMBER 2008

David Judkins, Westat

Robert St.Pierre, Principal Investigator, Abt Associates

Babette Gutmann, Project Director, Westat

Barbara Goodson, Abt Associates

Adrienne von Glatz, Westat

Jennifer Hamilton, Westat

Ann Webber, Westat

Patricia Troppe, Westat

Tracy Rimdzius, Institute of Education Sciences

NCEE 2008-4028
U.S. DEPARTMENT OF EDUCATION



U.S. Department of Education

Margaret Spellings

Secretary

Institute of Education Sciences

Grover J. Whitehurst

Director

National Center for Education Evaluation and Regional Assistance

Phoebe Cottingham

Commissioner

September 2008

This report was prepared for the Institute of Education Sciences under Contract No. ED-01-CO-0120. The project officer was Tracy Rimdzius in the National Center for Education Evaluation and Regional Assistance.

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

This publication is in the public domain. Authorization to reproduce it in whole or in part is granted. While permission to reprint this publication is not necessary, the citation should be: David Judkins, Robert St.Pierre, Babette Gutmann, Barbara Goodson, Adrienne von Glatz, Jennifer Hamilton, Ann Webber, Patricia Troppe, and Tracy Rimdzius. *A Study of Classroom Literacy Interventions and Outcomes in Even Start* (NCEE 2008-4028). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

To order copies of this report,

- Write to ED Pubs, Education Publications Center, U.S. Department of Education, P.O. Box 1398, Jessup, MD 20794-1398.
- Call in your request toll free to 1-877-4ED-Pubs. If 877 service is not yet available in your area, call 800-872-5327 (800-USA-LEARN). Those who use a telecommunications device for the deaf (TDD) or a teletypewriter (TTY) should call 800-437-0833.
- Fax your request to 301-470-1244.
- Order online at www.edpubs.org.

This report also is available on the IES website at <http://ies.ed.gov/ncee>.

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

EXECUTIVE SUMMARY

The Even Start Family Literacy Program was established in 1989 (P.L. 107-110, Sec. 1231) to help break the cycle of poverty and illiteracy for low-income families, by improving the literacy skills of parents and their young children (U.S. Department of Education 2003). Even Start projects offer family literacy services, defined as four integrated instructional components (P.L. 107-110, Sec. 9101 (20)):

- Early Childhood Education (ECE);
- Parenting Education (PE);
- Parent-Child Literacy Activities (PC); and
- Adult Education (AE).

Two previous studies of the Even Start Program showed that parents and children who participated in Even Start did not have better literacy outcomes than parents and children in a randomly assigned control group that did not receive Even Start services. The Even Start Classroom Literacy Interventions and Outcomes (CLIO) Study is the third randomized study of Even Start. As opposed to the earlier evaluations that investigated the effectiveness of Even Start relative to randomly assigned control groups in which parents and children were not enrolled in Even Start, the CLIO study was intended to intervene by offering the combination of research-based, literacy-focused early childhood education and parenting education curricula (the “CLIO combined curricula”). The CLIO study was intended to determine (1) whether the CLIO combined curricula were more effective than existing Even Start instructional services, and (2) whether research-based parenting education curricula that focus on child literacy (the “CLIO parenting curricula”) added value to research-based, literacy-focused early childhood education curricula (the “CLIO preschool curricula”).

This report presents 2-year impacts of the CLIO curricula on child language, literacy, and social competence; parenting skills; parent literacy; and instructional practices and participation in preschool and parenting classes.

Main Findings

The main findings from the CLIO impact analyses are that (1) the CLIO combined curricula had statistically significant, positive impacts on some of the hypothesized precursors to the development of children's early literacy skills, including instructional supports for literacy, child social competence, and parenting skills; but (2) the CLIO combined curricula did not have statistically significant impacts on any of the child language development and early literacy outcomes.

The CLIO combined curricula had statistically significant positive impacts on

- two of five measures of preschool instruction: support for print knowledge and literacy resources in the classroom;
- one of three measures of parenting instruction: the amount of parenting education time spent on child literacy;
- both measures of parenting outcomes: parent interactive reading skill and parent responsiveness to their child; and
- child social competence.

The CLIO combined curricula did not have statistically significant impacts on:

- three of five measures of preschool instruction: support for oral language, support for phonological awareness, and support for print motivation;
- two of three measures of parenting instruction: the amount of parenting education time spent on parenting skills not related to child literacy or the amount of parent-child time spent with parents and their children interacting on child literacy activities;
- monthly hours of preschool instruction received by children or monthly hours of parenting instruction received by parents;
- parent English reading skills (includes vocabulary); and

- child expressive language (in English or Spanish), receptive vocabulary, phonological awareness (Elision or Blending), print knowledge, or syntax and grammar.

The CLIO parenting curricula added value to the CLIO preschool curricula by increasing significantly the amount of parenting education time spent on child literacy, the amount of parenting education time spent on parenting skills not related to child literacy, and parent interactive reading skill. The CLIO parenting curricula did not significantly add value to the CLIO preschool curricula with respect to parent responsiveness, child literacy outcomes, or child social competence.

Background

The Even Start Family Literacy Program provides grants to local projects to provide family literacy services to low-income families. Family literacy services are defined as the integration of the four instructional services mentioned above with sufficient intensity in terms of hours and duration to make sustainable changes in a family. An important premise underlying the Even Start program is that the combination of early childhood education, parenting education, parent-child literacy activities, and adult education adds value to participant outcomes. That is, language and literacy outcomes for children in Even Start should be improved directly, through the effects of participation in preschool, and indirectly, through enhancements in both parenting skills and parent literacy. Parenting skills are expected to be enhanced through participation in parenting and parent-child activities, and parent literacy through participation in adult education literacy training.

Since the inception of Even Start in 1989, the U.S. Department of Education has sponsored three national evaluations of the program that focused on performance and effectiveness. Two of the three national evaluations included experimental studies that randomly assigned eligible and interested families to participate in Even Start or a control group of families who would delay participation in Even Start for at least 1 year (St.Pierre et al. 2003; St.Pierre et al. 1995). The results of these studies showed that Even Start projects were not effective at improving the literacy skills of participating preschool-age children and their parents. That is, literacy gains made by Even Start parents and children were no different from literacy gains made by control parents and

children. The control group for these randomized studies was composed of parents who wanted to enroll their children in Even Start but who were randomly assigned to participate in Even Start in the year following the evaluation. About two-thirds of these control parents were unable to arrange any other formal early childhood education (ECE) services during the period of the evaluation, so the control condition mostly corresponded to at-home care by parents or extended family members (St.Pierre et al. 2003, p. 162).

The absence of significant effects of Even Start on literacy skills, along with new requirements in the reauthorized Even Start legislation to base instruction on scientifically based reading research (Sec. 1231(2)(D)), prompted an examination of the Even Start model to determine how it could be improved. The lead investigators of the most recent national Even Start evaluation (St.Pierre, Ricciuti, and Rimdzius 2005) addressed several questions about Even Start's apparent ineffectiveness: (1) whether the Even Start model was fully implemented, (2) whether Even Start's instructional services were sufficiently intensive, (3) whether Even Start families participated sufficiently, and (4) whether the quality of Even Start's instruction and curriculum content was sufficient to lead to positive effects.

The CLIO study was, therefore, designed to test the extent to which research-based, literacy-focused curricula strengthen Even Start services and lead to significant impacts on parents and children.¹ Specifically, the CLIO study was designed to address two primary research questions:

- Is the combination of research-based, literacy-focused preschool, parenting, and parent-child curricula (the CLIO combined curricula) more effective than the existing combination of services in Even Start?
- Do research-based parenting and parent-child curricula (the CLIO parenting curricula) that focus on child literacy add value to the CLIO preschool curricula?

¹ This is consistent with Even Start's second legislative evaluation requirement (Sec. 1239 (2)), which is to identify effective programs that can be duplicated and used in providing technical assistance. CLIO is also consistent with the requirement for research (Sec. 1241) that examines successful family literacy services.

Thus, the study was an evaluation of the *incremental* effectiveness of providing the CLIO curricula to Even Start projects.

CLIO Study Design and Curricula

Through a competitive process, the CLIO study selected two combined preschool and parenting education curricula,² each of which were based on the most current research on the development of children’s early literacy skills. CLIO used these curricula in four combinations—two that implemented the combined research-based preschool and parenting curricula and two that implemented the research-based preschool curricula in combination with existing parenting education services. The CLIO study used an experimental design in which 120 Even Start projects were randomly assigned to implement one of the four CLIO curricula combinations or to be in a control group that provided their regular pre-CLIO instructional services (see table ES-1).

The CLIO combined curricula and CLIO preschool curricula were implemented in the sample of Even Start projects during program years 2004-2005 and 2005-2006. Implementation included summer training sessions for project directors and teachers in each year, as well as ongoing support for preschool and parenting education staff from the curriculum developers over the 2-year period.

The CIRCLE group at the University of Texas-Houston Health Sciences Center teamed with Abrams & Company Publishers to provide the Let’s Begin with the Letter People preschool curriculum to CLIO. Let’s Begin is a preschool curriculum that builds early literacy skills and uses 26 imaginary characters that represent the letters of the alphabet. The CIRCLE group provided the Play and Learning Strategies (PALS) parenting curriculum to CLIO. PALS focuses on responsive parenting and teaches parents techniques to build their children’s language and cognitive development.

² The study team decided not to include Even Start’s adult education component in the test of research-based curricula because (1) most projects provided a variety of adult education services at different levels (adult basic education (ABE), general equivalency diploma (GED), English as a second language (ESL)) to meet family needs, (2) a substantial portion of projects used community service providers to deliver adult education services, and (3) the research on effective adult education models is still in its infancy.

Table ES-1. Specification of the Five CLIO Study Groups

	Study group				
	Group 1	Group 2	Group 3	Group 4	Group 5
Even Start instructional component	LET'S BEGIN with the Letter People (ECE) CLIO preschool curriculum	LET'S BEGIN with the Letter People and Play and Learning Strategies (PALS) (ECE/PE) CLIO combined curriculum	Partners for Literacy (ECE) CLIO preschool curriculum	Partners for Literacy (ECE/PE) CLIO combined curriculum	Control
Early childhood education	LET'S BEGIN	LET'S BEGIN	Partners for Literacy	Partners for Literacy	As usual
Parenting education	As usual	PALS	As usual	Partners for Literacy	As usual
Parent-child joint literacy activities	As usual	PALS	As usual	Partners for Literacy	As usual
Adult education	As usual	As usual	As usual	As usual	As usual
NOTE: Shaded areas identify instructional components that were provided by the CLIO curriculum developers.					

The University of North Carolina at Chapel Hill provided the Partners for Literacy curriculum to CLIO. The preschool Partners curriculum is based on game-like activities conducted with pairs of children and instructional strategies designed to support children's cognitive and language development. The parenting Partners curriculum adapts the game-like activities and instructional strategies from the preschool curriculum and trains parents to use these with their children at home. The Partners curriculum also includes training in problem-solving skills for children and parents.

CLIO Contrasts

As discussed earlier, the CLIO study addressed two key research questions:

- 1) Is the combination of research-based, literacy-focused preschool, parenting, and parent-child curricula (the CLIO combined curricula) more effective than the existing combination of services in Even Start?

- 2) Do research-based parenting and parent-child curricula (the CLIO parenting curricula) that focus on child literacy add value to the CLIO preschool curricula?

The first research question was addressed analytically by combining projects that received the CLIO combined curricula (study groups 2 and 4 in table ES-1) and comparing their outcomes with those of control projects (study group 5). The study's second research question was addressed analytically by combining projects that received the CLIO combined curricula (study groups 2 and 4), and comparing their outcomes with those of projects that received the CLIO preschool curricula (study groups 1 and 3).

CLIO Data Collection and Outcome Constructs

The study team collected data over a 3-year period. The first year of data collection was 2003-2004, prior to implementation of the CLIO curricula. The second and third years of data collection (2004-2005 and 2005-2006) corresponded to the two CLIO curricula implementation years.³

The study team conducted the following types of data collection in all CLIO projects: direct assessments of child language and literacy; teacher ratings of child social competence; videotapes of parent-child interactions; interviews of parents; direct assessments of parent literacy; observations of classroom instruction in preschool, parenting education, and parent-child classes; surveys of teachers and project directors; and tallies of child and parent participation in instructional services. The study team also observed and rated the fidelity of implementation of the CLIO curricula. The outcome constructs used in the CLIO impact analyses are presented in table ES-2.

³ The CLIO study is also following children into kindergarten and first grade.

Table ES-2. CLIO Outcome Measures

	Outcome		Data collection instrument	Mode of data collection	Domain
CHILD	1	Expressive language: English	Individual Growth and Development Indicator (IGDI)	Child assessment	Emergent literacy
	2	Expressive language: Spanish			
	3	Receptive vocabulary	Peabody Picture Vocabulary Test (PPVT)		
	4	Phonological awareness: Elision	Comprehensive Test of Phonological and Print Processing (Preschool – CTOPPP)		
	5	Phonological awareness: Blending			
	6	Print knowledge			
	7	Syntax and grammar	Test of Language Development (TOLD-3)		
	8	Social competence	Teacher rating form	Teacher rating	Socio-emotional development
PARENT	9	Parent interactive reading skill	Read Aloud Together Profile & Parent Interview	Video observation, parent report	Parenting skills
	10	Parent responsiveness			
	11	Reading & vocabulary skill	Parent assessment battery	Parent assessment	Parent language & literacy
INSTRUCTIONAL	12	Support for oral language development	Observation Measures of Language and Literacy Instruction (OMLIT) and Parenting Education and Child/Parent Observation (PECAP)	Classroom observation	Preschool classroom instruction
	13	Support for print knowledge			
	14	Support for phonological awareness			
	15	Support for print motivation			
	16	Literacy resources in classroom			Parenting classroom instruction
	17	Parenting education time spent on child literacy			
	18	Parenting education time spent on parenting skills			Parent-child classroom instruction
	19	Parent-child time spent interacting on child literacy activities			
	20	Child: Monthly hours of preschool instruction received	Instructional Services Participation Form (ISPF)	Project report	Participation amount
	21	Parent: Monthly hours of parenting and parent-child instruction received			

Implementation of the CLIO Curricula

Fidelity to Planned CLIO Curricula. Fidelity of implementation to the CLIO curricula in the sample projects was rated both by independent observers and by the curriculum developers. Both sets of ratings indicated that, on average, implementation of the CLIO combined curricula and the CLIO preschool curricula only reached about 50 percent of full implementation. Fidelity ratings for the Let's Begin and PALS projects were generally higher than those for the Partners for Literacy projects, for both the preschool and parenting classrooms but particularly for preschool classrooms. Most of the average fidelity ratings by observers and developers were higher in 2006 than in 2005 with the exception of observer ratings for Partners for Literacy preschool classrooms.

Exposure to the CLIO Curricula. Participants (parents and children) in any intervention need a minimum level of exposure to the curriculum to obtain the hypothesized benefits. Even Start guidelines do not specify an expected level of exposure for children or parents, and the hours of instruction offered by local projects vary widely. In each implementation year, while projects reported that they offered preschoolers an average of 80 hours of preschool education per month, children in CLIO projects actually participated in preschool an average of 50 hours per month. Parents also received only partial exposure to the parenting curricula. Projects reported that they offered parents an average of 25 hours of parenting education and parent-child activities per month, but parents participated for an average of 13 hours of parenting education and parent-child activities per month. These levels of participation relative to hours of services offered are in line with what was documented in previous Even Start evaluations (St.Pierre et al. 2003, p. 129).

Control Projects. Project directors reported that about 75 percent of the CLIO control projects used a formal early childhood curriculum (most often High/Scope or Creative Curriculum), and about 60 percent used a formal parenting curriculum (most often locally developed). Observations of control classrooms showed that they spent about 45 percent of the day in activities that are often considered by developmental psychologists to have particularly high value for children because of the opportunities for children to construct knowledge and receive feedback on their interactions with materials, peers, and adults in the classroom (Bruner and Watson 1983). The remainder

of the control group day was spent in daily group activities including review of the calendar/weather/attendance, gross motor play and transition, and meals/snacks.

Impact Findings

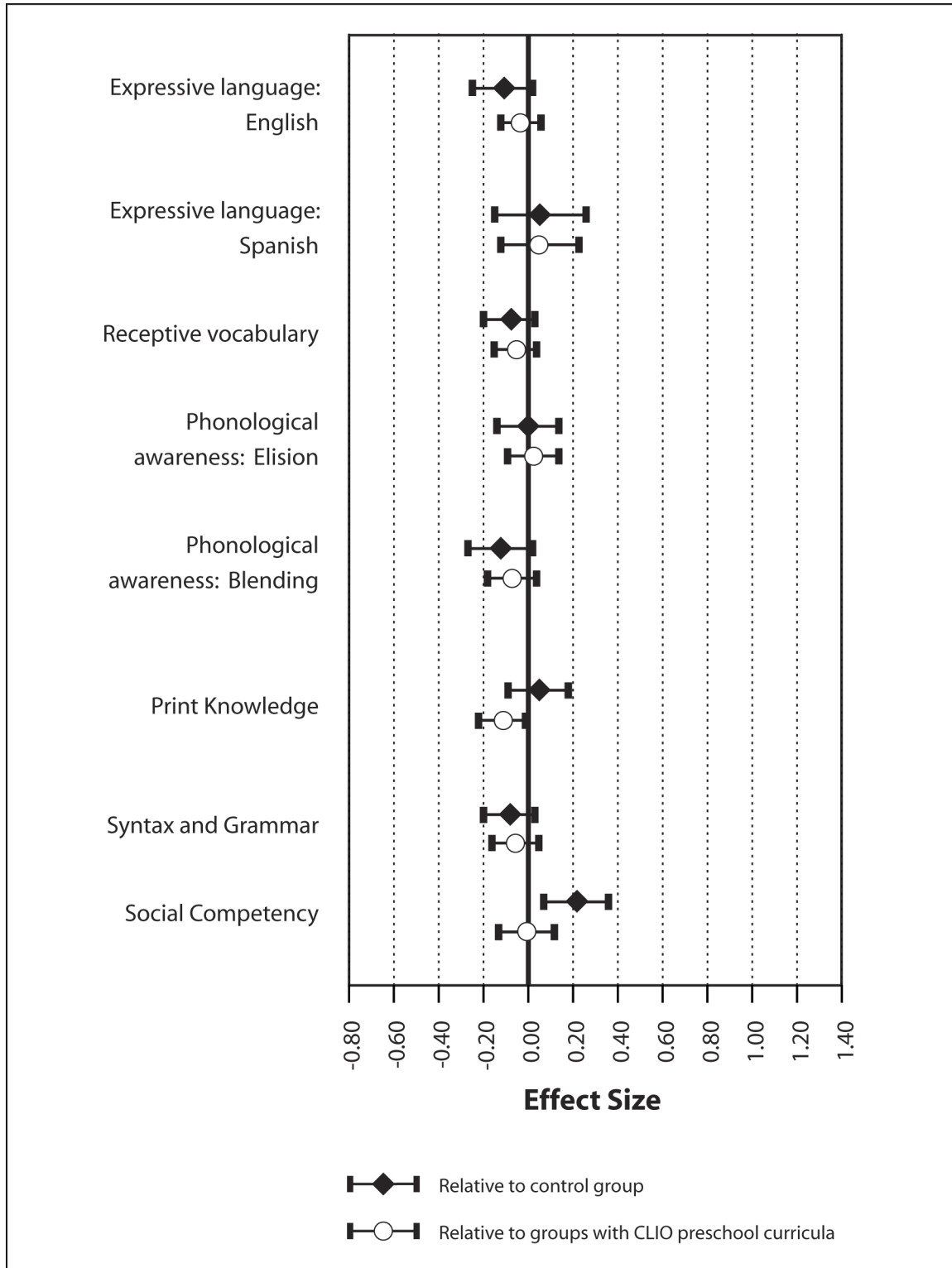
Impacts of the CLIO Combined Curricula. The study showed that Even Start projects assigned to the CLIO combined curricula did not exhibit better child language and literacy outcomes than Even Start projects assigned to the control group (figure ES-1). In the figures in this section, effect sizes for the combined curricula are indicated by filled diamonds (relative to the control group) and open circles (relative to the preschool curricula), and 95 percent confidence intervals⁴ are shown as horizontal bands on either side of the diamond or circle. Effect size indicates the difference in outcome between the average subject who received the treatment and the average subject who did not.⁵

There were no statistically significant impacts of the CLIO combined curricula on any of the seven measures of child language and literacy skills (six in English and one in Spanish), as can be seen by the fact that none of the confidence bands exclude zero, even before adjustment for multiple comparisons. Estimated effect sizes on emergent literacy outcomes were all smaller than 0.13 in absolute value, with confidence interval limits all bounded by 0.27 in absolute value. However, the CLIO combined curricula did have a statistically significant positive effect on child social competence (behavior in class) as rated by preschool teachers. The effect size of the impact of the CLIO combined curricula on child social competence was 0.22.

⁴ The confidence intervals may be interpreted as follows. If the experiment were to be independently repeated a very large number of times under the same general conditions, drawing on the same population of schools and students, and on every repetition both an effect estimate and a confidence interval on that estimate were calculated, then, over the long run, 95 percent of the confidence intervals would contain the long-run average of estimated effects.

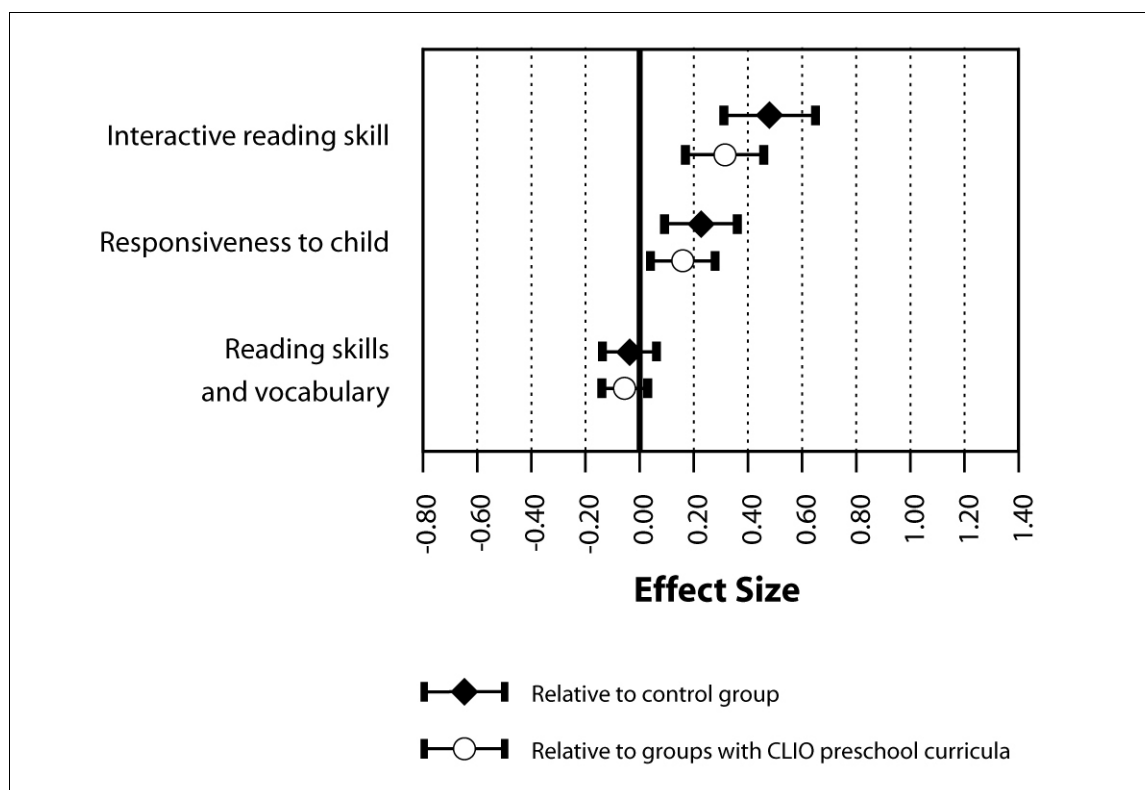
⁵ Effect size was calculated by taking the difference between the treatment and control group means and dividing that difference by the standard deviation of the control group's scores in 2005.

Figure ES-1. Effect Sizes for CLIO Combined Curricula on Child Outcomes Relative to Both the Control Group and the CLIO Preschool Curricula (average of spring 2005 and spring 2006)



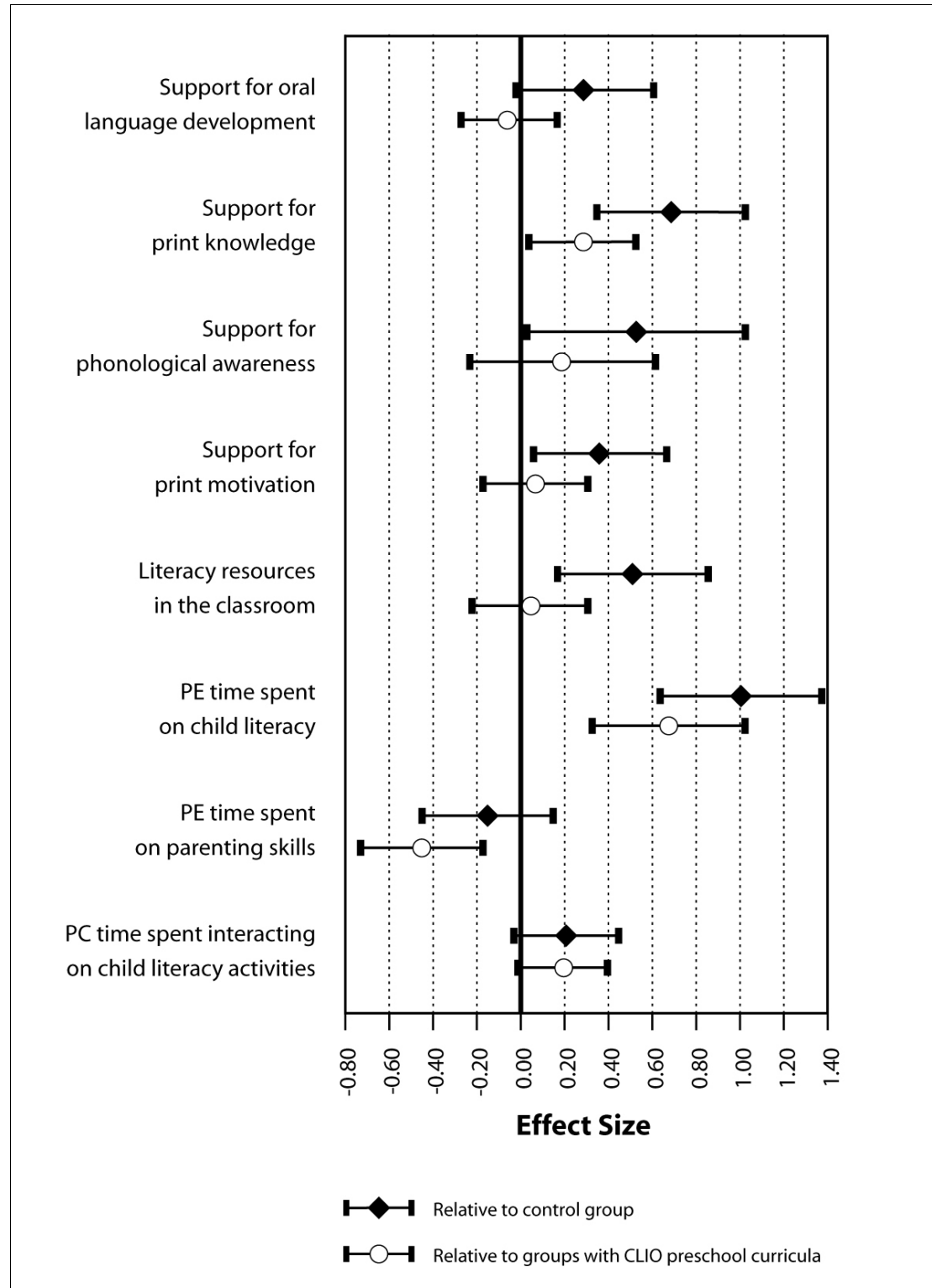
The CLIO combined curricula had a statistically significant positive impact on both of the parent outcomes examined (figure ES-2). The effect size of the impact on parent interactive reading skill was 0.48, and the effect size of the impact on parent responsiveness to their child was 0.22. Even though CLIO did not manipulate adult education curricula, the study assessed parent reading skills and vocabulary and showed that the CLIO combined curricula did not have a statistically significant impact on these skills (figure ES-2).

Figure ES-2. Effect Sizes for CLIO Combined Curricula on Parent Outcomes Relative to Both the Control Group and the CLIO Preschool Curricula (average of spring 2005 and spring 2006)



The CLIO combined curricula had a statistically significant positive impact on two of five measures of instructional support for literacy development in preschool classrooms (figure ES-3). The effect sizes of the statistically significant impacts on support for print knowledge and literacy resources in the classroom were 0.69 and 0.52, respectively. There was no statistically significant impact on the following three

Figure ES-3. Effect Sizes for CLIO Combined Curricula on Instructional Outcomes Relative to Both the Control Group and the CLIO Preschool Curricula (average of spring 2005 and spring 2006)



preschool instructional measures: support for oral language development, support for phonological awareness, or support for print motivation.⁶

The CLIO combined curricula had a positive impact on one of the three measures of parenting education and parent-child classroom instruction (figure ES-3). The effect size of the impact on the amount of parenting education time spent on child literacy was 1.01. There was no statistically significant impact on the amount of parenting education time spent on parenting skills not related to child literacy or the amount of parent-child time spent with parents and their children interacting on child literacy activities.

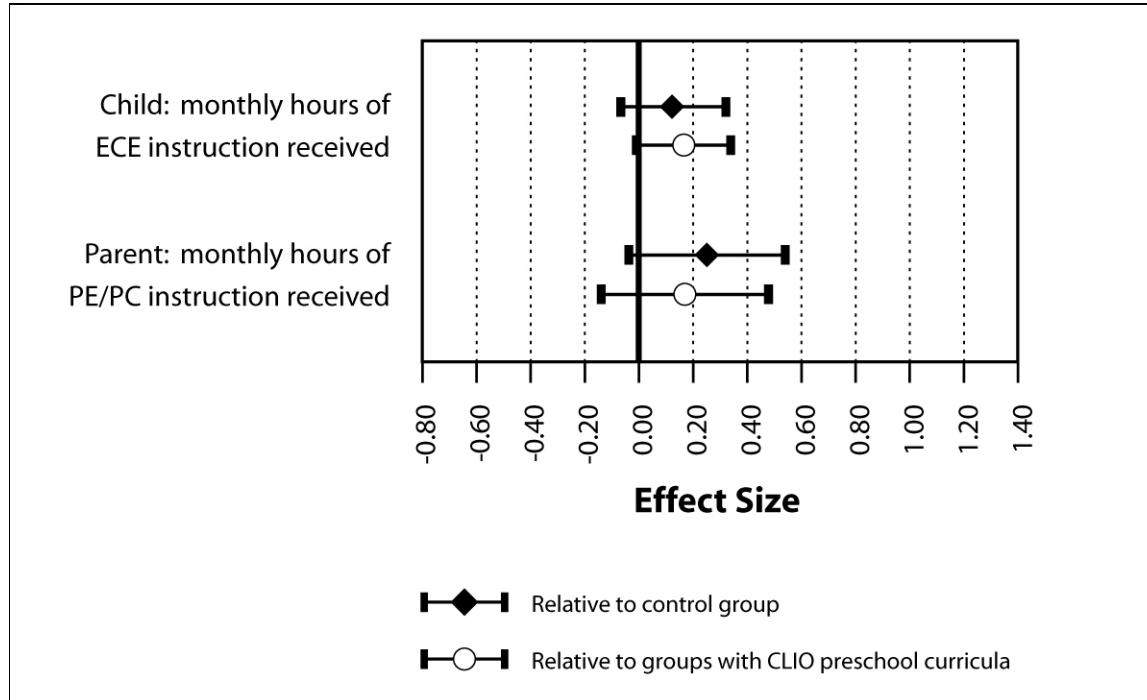
The study also examined whether the CLIO combined curricula had an impact on participation levels (figure ES-4). The results showed that there was no statistically significant impact of the CLIO combined curricula on either child levels of participation in preschool or parent levels of participation in parenting education or parent-child activities. Neither of the confidence bands exclude zero.

Added Value of the CLIO Parenting Curricula. CLIO parenting curricula did not add significantly to the effectiveness of the CLIO preschool curricula on any of the seven measures of child literacy skills or on child social competence (figure ES-1). That is, adding research-based parenting components focused on child literacy did not add significantly to children's outcomes beyond what was achieved with the CLIO preschool curricula. (In figures ES-1 through ES-4, the effect sizes for the added value of the CLIO parenting curricula are indicated by open circles.) The estimated effect sizes of the CLIO parenting curricula on emergent literacy outcomes were all smaller than 0.11 in absolute value, with confidence interval limits all bounded by 0.23 in absolute value.

However, the CLIO parenting curricula did have a statistically significant positive incremental effect on parent interactive reading skill (effect size of 0.30) (figure ES-2). The difference on parents' responsiveness to their child between the CLIO combined curricula and the CLIO preschool curricula, while similar in size to the statistically significant difference between the CLIO combined curricula and the control group, was not statistically significant.

⁶ Although the confidence bands for support for phonological awareness and support for print motivation exclude zero, the effect sizes are not significant once adjusted for multiple comparisons.

Figure ES-4. Effect Sizes for CLIO Combined Curricula on Participation Relative to Both the Control Group and the CLIO Preschool Curricula (average of spring 2005 and spring 2006)



There were statistically significant incremental effects of the CLIO parenting curricula on two of the instructional measures. The effect sizes of the incremental effects of the CLIO parenting curricula on the amount of parenting education time spent on child literacy and the amount of parenting education time spent on parenting skills not related to child literacy were 0.68 and -0.45, respectively (figure ES-3). There was no statistically significant incremental effect of the CLIO parenting curricula on how time was spent in parent-child classes or (as expected) in preschool classes.

Finally, the CLIO parenting curricula did not have a statistically significant incremental effect on child participation in preschool or on parent participation in parenting education (figure ES-4).

Secondary Analyses

Three secondary analyses were conducted to examine the variation in impacts of the CLIO curricula.

Year of Implementation. One hypothesis of the CLIO study was that impacts might be greater in the second year, when most projects could be assumed to have had 2 years to reach full implementation. With respect to child outcomes, there is evidence that the CLIO combined curricula had statistically significant negative effects on four of the seven children's language and literacy outcomes in the first year of implementation. By the second year, rough parity with the control group was achieved. There is little evidence of differential effects by year for child social competence, parent outcomes, instructional outcomes, and participation.

Analysis of Growth for Child and Parent Outcomes. While the primary impact analysis measures parent and child outcomes at the end of preschool, the study also examined impacts on the pattern of growth from fall to spring. The only significant finding was that the CLIO parenting curricula had a positive incremental effect on parent responsiveness to their child.

Interactions of Study Group with Ethnicity and Home Language. About half of all children in the CLIO sample spoke a home language other than English. An analysis of interactions found that impacts on children's emergent literacy did not vary significantly as a function of home language or ethnicity.

Summary

Prior studies have established that Even Start does not have statistically significant impacts on children's emergent literacy or on parent literacy. The CLIO study investigated whether the implementation of research-based, literacy-focused curricula would improve literacy outcomes for Even Start children and parents. Although there were positive impacts on some of the literacy supports in preschool classrooms, on time spent on child literacy in parenting education classes, on parenting skills, and on children's social competence, there were no statistically significant impacts on children's language and literacy. There was no evidence that the failure to

find impacts on these core outcomes was due to a lack of fidelity in the treatment classrooms or cross-over in the control classrooms.

ACKNOWLEDGMENTS

This report is the culmination of several years of design, data collection, and analysis. We gratefully acknowledge the contributions of a significant number of individuals in its preparation and production.

We benefited from the advice of a Technical Work Group. Members have included: Marilyn Box, (formerly) Mesa Public Schools Even Start; Gene Brody, University of Georgia; Thomas Cook, Northwestern University; Lilli Copp, (formerly) Florida State Department of Education; David Francis, University of Houston; Larry Hedges, Northwestern University; Sue Henry (formerly) New York State Department of Education; Chris Lonigan, Florida State University; Robin Morris, Georgia State University; Lynne Vernon-Feagans, University of North Carolina; Barbara Wasik, Johns Hopkins University; Lin Wrinkle, Taylor, Texas Even Start.

Numerous individuals at Westat and Abt Associates contributed in a multitude of ways to the successful conduct of this study. From Westat, design and analysis advice was provided by Alexander Ratnofsky and Camilla Heid. Recruitment was successfully led by Carin Celebuski. Data collection was capably managed by Juanita Lucas-McLean. Field managers over the rounds of data collection included Laura Collins, Katia Cruz, Julie Daft, Sabrina Daly, Karen Gray-Adams, Luis Romero, Sylvia Segovia, and Dawn Thomas-Banks. Participation data were collected and edited by Cathy Lease. The many data systems were designed and implemented under the direction of Patricia Nicchitta, with the assistance of Kanaka Durga and Ying Long. Analysis and statistical support was provided by Jiaquan Fan, Philip Fletcher, and Frank Jenkins. Expert editorial and production assistance was provided by Evarilla Cover and Saunders Freeland. From Abt Associates, Carolyn Layzer helped conceive and develop the OMLIT observation system and also helped train and manage its administration in the field. Anne Robertson played a similar role for the PECAP observation system. Cristofer Price provided analysis and technical support. Programming and database management for the classroom observation data were led by Nancy McGarry and Amanda Parsad.

We were also assisted in this undertaking by a number of individuals from two other organizations, XTRIA and the CDM Group. Both organizations provided several staff who capably assisted in the recruitment of Even Start projects and in the training of data collection staff.

Patricia McKee of the U.S. Department of Education's Office of Elementary and Secondary Education and her staff served as important liaisons with the Even Start State coordinators. The Even Start State coordinators provided support throughout recruitment and data collection. Most importantly, this study would not have been possible without the 3-year commitment of the Even Start projects that agreed to participate in the CLIO study. Their willingness to participate for 3 years is truly appreciated by all of us connected to this study.

Disclosure of Potential Conflicts of Interests⁷

The research team for this evaluation consists of a prime contractor, Westat, and a subcontractor, Abt Associates. Neither of these organizations or their key staff has financial interests that could be affected by findings from the Even Start Classroom Literacy Interventions and Outcomes Study.

No one on the Technical Work Group, convened to provide advice and guidance, has financial interests that could be affected by findings from the evaluation. One of the members of the study's Technical Work Group, Dr. Christopher Lonigan of Florida State University, is the lead author on the Test of Preschool Early Literacy. The CLIO study used components of a prepublication version of this assessment called the Preschool Comprehensive Test of Phonological and Print Processing (Pre-CTOPPP).

⁷ Contractors carrying out research and evaluation projects for IES frequently need to obtain expert advice and technical assistance from individuals and entities whose other professional work may not be entirely independent of or separable from the particular tasks they are carrying out for the IES contractor.

Contractors endeavor not to put such individuals or entities in positions in which they could bias the analysis and reporting of results, and their potential conflicts of interest are disclosed.

CONTENTS

	<u>Page</u>
Executive Summary	iii
Acknowledgments	xxi
Disclosure of Potential Conflicts of Interests	xxii
1. Background	1
The Even Start Program	1
Even Start Has Not Performed Up To Expectations	2
The CLIO Study: Seeking to Improve Even Start.....	3
Organization of the Report	6
2. Research Design	7
Selection of the CLIO Curriculum Developers.....	7
Establishment of the Five Study Groups for the CLIO Study	9
Implementation of the Curricula	11
Design of the Study to Address the Research Questions	11
Recruitment of Even Start Projects	12
Random Assignment of CLIO Projects	16
Study Projects at Baseline	17
Data Collection Schedule and Methods.....	22
Sample Sizes.....	26
Development of Outcome Measures.....	26
3. Curriculum Implementation	41
Curricula as Intended	41
CLIO Curricula Professional Development	51
Instructional Practices and Fidelity in CLIO Classrooms	52
Fidelity of Implementation	53
Exposure to the Curricula	59
Control Projects During CLIO Study	62

Contents (continued)

	<u>Page</u>
4. Analysis Methods	67
Use of Covariates	68
Multi-Level Modeling	70
Contrast Selection and Corrections for Multiple Comparison Testing.....	72
Reporting Findings	74
Statistical Power	76
Handling of Missing Data.....	78
5. Impacts of the CLIO Curricula on Children, Parents, Classroom Instruction, and Participation.....	81
Impacts on Children	83
Impacts on Parents.....	88
Impacts on Classroom Instruction.....	91
Impacts on Participation	96
Summary	97
6. Secondary Analyses of Impacts of the CLIO Curricula.....	101
Year of Implementation	101
Growth Instead of End-Point Status for Child and Parent Outcomes	102
Interactions of Study Group with Ethnicity and Home Language.....	105
7. Additional Analyses	107
Relationship Between Child Outcomes and Preschool Instruction.....	108
Relationship Between Child Outcomes and Parent Teaching Behaviors	112
Variation in Impacts by Fidelity of Implementation of the CLIO Curricula.....	113
Parent Participation and Parenting Outcomes	120
Participation in Preschool Education and Child Outcomes	125
Contrasts in Child Emergent English Literacy Across Study Groups Among Children with Substantial Participation.....	126
Bibliography	129

Contents (continued)

	<u>Page</u>
Appendix A. Descriptive Comparison of the Five Groups	A-1
Appendix B. Child and Parent Literacy Outcomes.....	B-1
Appendix C. Construction of the Child Social Competence Scale	C-1
Appendix D. Construction of Parenting Skills Scales.....	D-1
Appendix E. Preschool Instructional Outcome Variables.....	E-1
Appendix F. Parenting Instructional Outcome Variables.....	F-1
Appendix G. Fidelity Measurement.....	G-1
Appendix H. Final Multi-Level Outcome Models	H-1
Appendix I. Effect Size Calculation.....	I-1
Appendix J. Sample Sizes and Methods Used for Missing Data	J-1
Appendix K. ITT Tables for Single Developer Contrasts.....	K-1
Appendix L. Growth Analysis	L-1
Appendix M. Interactions of Curriculum with Home Language and Ethnicity	M-1
Appendix N. Simulation Study of Method for Exploring Fidelity- Adjusted Relationships	N-1
Appendix O. Participation and Participation-Adjusted Group Differences	O-1

List of Tables

	<u>Page</u>
Table ES-1 Specification of the Five CLIO Study Groups.....	viii
Table ES-2 CLIO Outcome Measures	x
Table 2-1 Specification of the Five CLIO Study Groups.....	10
Table 2-2 Stratification Variables Used in Random Assignment	16
Table 2-3 Percentage Distribution of CLIO Children and Even Start Parents by Race/Ethnicity	18
Table 2-4 Percentage Distribution of Educational Attainment for Mothers of CLIO Children and New Even Start Parents	19
Table 2-5 Average Monthly Hours of ECE Instruction Offered and Received and Average Monthly Hours of PE/PC Instruction Received, for CLIO and Even Start.....	20
Table 2-6 Educational Attainment of Even Start Staff in CLIO Projects: Spring 2004.....	20
Table 2-7 Overview of Data Collection	23
Table 2-8 Cross-sectional Child Sample Size by Child Age and Data Collection Cycle.....	27
Table 2-9 CLIO Outcome Measures	28
Table 3-1 Average Classroom-Level Fidelity Ratings by Study Group and Classroom Type for Spring 2005 and Spring 2006: Independent Observer Ratings.....	56
Table 3-2 Average Project-Level Fidelity Ratings by Study Group and Classroom Type for Spring 2005 and Spring 2006: Developer Ratings	57

List of Tables (continued)

	<u>Page</u>
Table 3-3 ECE Instruction in CLIO Classrooms (average of spring 2005 and spring 2006)	64
Table 4-1 Research Questions and Contrasts	73
Table 4-2 Effect Sizes and Percentile Standings	75
Table 4-3 Minimum Detectable Effect Sizes	77
Table 5-1 Unadjusted Outcome Means (average of Spring 2005 and Spring 2006 data) by Study Group, and Results of Overall Test for Any Differences Across the Five Groups	84
Table 5-2 Effects of CLIO Combined Curricula on Children (Research Question 1) (average of spring 2005 and spring 2006)	85
Table 5-3 Incremental Effects of CLIO Parenting Curricula on Children (Research Question 2) (average of spring 2005 and spring 2006)	87
Table 5-4 Effects of CLIO Combined Curricula on Parents (Research Question 1) (average of spring 2005 and spring 2006)	89
Table 5-5 Incremental Effects of CLIO Parenting Curricula on Parents (Research Question 2) (average of spring 2005 and spring 2006)	91
Table 5-6 Effects of CLIO Combined Curricula on Instruction (Research Question 1) (average of spring 2005 and spring 2006)	92
Table 5-7 Incremental Effects of CLIO Parenting Curricula on Instruction (Research Question 2) (average of spring 2005 and spring 2006)	94
Table 5-8 Effects of CLIO Combined Curricula on Participation (Research Question 1) (average of spring 2005 and spring 2006)	96
Table 5-9 Incremental Effects of CLIO Parenting Curricula on Participation (Research Question 2) (average of spring 2005 and spring 2006)	96

List of Tables (continued)

	<u>Page</u>
Table 5-10	Summary of Findings from Selected Meta-Analyses of the Effects on Cognitive Outcomes of Early Education for Disadvantaged Children.....98
Table 6-1	Effects of CLIO Combined Curricula on Child Outcomes by Year103
Table 6-2	Effects of CLIO Combined Curricula on Parent Outcomes by Year103
Table 6-3	Effects of CLIO Combined Curricula on Instructional Outcomes and Participation by Year104
Table 7-1	Significant Regression Coefficients (and standard errors) from Non-experimental Analyses of Relationships of ECE Classroom Instructional Variables and Parent Behaviors with Child Outcomes.....111
Table 7-2	Fidelity-Adjusted Estimates of Curriculum Effects Relative to Status Quo (Developer-rated Fidelity).....118
Table 7-3	Fidelity-Adjusted Estimates of Curriculum Effects Relative to Status Quo (Observer-rated Fidelity).....119
Table 7-4	Contrasts with the Control Group of CLIO Combined Curricula on Children with High Participation Levels (Research Question 1) (Combined Spring 2005 and Spring 2006).....128
Table A-1	Percentage Distribution of CLIO Children, by Study Group and Child and Family Demographics: Spring 2004, Spring 2005, and Spring 2006..... A-1
Table A-2	Percentage of CLIO Parents, by Study Group and Parent-Reported Literacy Levels: Spring 2004..... A-5
Table A-3	Average Language, Literacy, Social Competence, and Participation Outcomes for CLIO Children and Parents, by Study Group: Spring 2004 A-6

List of Tables (continued)

	<u>Page</u>
Table A-4	Average Size, Amount of Services, and Early Childhood Education Instructional Outcomes for CLIO Projects, by Study Group: Spring 2004 A-7
Table A-5	Percentage Distribution of Even Start Staff, by Study Group and Selected Demographics: Spring 2004 A-8
Table A-6	Average Project Score on the Early Childhood Environmental Rating Scale-Revised (ECERS-R) Measures by Subscale and Study Group: Spring 2004 A-9
Table A-7	Means and Balance Tests for CLIO Study Children, by Study Group and Child and Parent Covariate: Spring 2005 and Spring 2006..... A-10
Table A-8	Project-level Means and Balance Tests, by Study Group and Project-level Covariate: Spring 2005 and Spring 2006 A-11
Table B-1	CLIO Child Assessment Subtests: Reliability and ValidityB-4
Table B-2	CLIO Parent Assessment Subtests: Reliability and ValidityB-6
Table B-3	References for CLIO Child and Parent Assessment Subtests, Including Reliability/Validity DataB-7
Table B-4	References for Tests Cited in Concurrent Validity Columns of Tables B-1 and B-2.....B-8
Table D-1	Items Dominating Each Parenting Outcome, Fall 2004..... D-6
Table E-1	The Observation Measure of Language and Literacy Instruction (OMLIT).....E-6
Table E-2	Formal Inter-Rater Reliability Tests for Candidate Observer.....E-8

List of Tables (continued)

	<u>Page</u>
Table E-3	Inter-Rater Agreement on the Classroom Literacy Opportunities Checklist (OMLIT-CLOC).....E-11
Table E-4	Inter-Rater Agreement on the Quality of Instruction in Language and Literacy (OMLIT-QUILL).....E-11
Table E-5	Inter-Rater Agreement on the Read Aloud Profile (OMLIT-RAP).....E-12
Table E-6	Inter-Rater Agreement on the Classroom Literacy Instruction Profile (OMLIT-CLIP).....E-13
Table E-7	Inter-Rater Agreement on the Snapshot of Classroom Activities (OMLIT-SNAP)E-13
Table E-8	ECE Classroom Outcomes and Component OMLIT VariablesE-15
Table E-9	Internal Consistency and Inter-Rater Reliability of ECE Outcome Constructs, Spring 2004.....E-17
Table E-10	Correlations Among ECE Outcome Constructs, Spring 2004E-17
Table F-1	Definition and Reliability of PECAP Outcome Measures.....F-4
Table G-1	Fidelity Domains by CLIO Curriculum.....G-4
Table G-2	Correlation Between Project-level Observer and Developer Fidelity Ratings by Study Group and Classroom Type: Spring 2005 and Spring 2006.....G-6
Table H-1	Model for Child Expressive Language in English—Spring 2005 and Spring 2006 CombinedH-3
Table H-2	Model for Child Expressive Language in Spanish—Spring 2005 and Spring 2006 CombinedH-4

List of Tables (continued)

	<u>Page</u>
Table H-3	Model for Child Receptive Vocabulary—Spring 2005 and Spring 2006 CombinedH-5
Table H-4	Model for Child Elision Component of Phonological Awareness—Spring 2005 and Spring 2006 CombinedH-6
Table H-5	Model for Child Blending Component of Phonological Awareness—Spring 2005 and Spring 2006 CombinedH-7
Table H-6	Model for Child Print Awareness—Spring 2005 and Spring 2006 CombinedH-8
Table H-7	Model for Child Syntax and Grammar—Spring 2005 and Spring 2006 CombinedH-9
Table H-8	Model for Child Social Competence—Spring 2005 and Spring 2006 CombinedH-10
Table H-9	Model for Child Monthly Hours of ECE Instruction Received—Spring 2005 and Spring 2006 CombinedH-11
Table H-10	Model for Parent Interactive Reading Skill—Spring 2005 and Spring 2006 CombinedH-12
Table H-11	Model for Parent Responsiveness—Spring 2005 and Spring 2006 CombinedH-13
Table H-12	Model for Parent Reading and Vocabulary Skill—Spring 2005 and Spring 2006 CombinedH-14
Table H-13	Model for Parent Monthly Hour of PE and PC Instruction Received—Spring 2005 and Spring 2006 CombinedH-15
Table H-14	Model for OMLIT Oral Language Development—Spring 2005 and Spring 2006 CombinedH-16

List of Tables (continued)

	<u>Page</u>
Table H-15	Model for OMLIT Print knowledge—Spring 2005 and Spring 2006 Combined H-16
Table H-16	Model for OMLIT Phonological Awareness—Spring 2005 and Spring 2006 Combined H-17
Table H-17	Model for OMLIT Print Motivation—Spring 2005 and Spring 2006 Combined H-17
Table H-18	Model for OMLIT Literacy Resources in Classroom—Spring 2005 and Spring 2006 Combined H-18
Table H-19	Model for PE Time Spent on Child Literacy—Spring 2005 and Spring 2006 Combined H-18
Table H-20	Model for PE Time Spent on Parenting Skills—Spring 2005 and Spring 2006 Combined H-19
Table H-21	Model for PC Time Spent Interacting on Child Literacy Activities—Spring 2005 and Spring 2006 Combined H-19
Table J-1	Child and Parent Sample Sizes, Spring 2005 and Spring 2006 Combined J-6
Table J-2	Increases in Analyzable Child and Parent Sample Sizes Due to Imputation of Whole Parent Interviews J-8
Table J-3	Projects with No Observation Data by Observation Type and Year of Missing Data..... J-9
Table K-1	Effects of CLIO Combined Curricula on Children (Research Question 1) (average of spring 2005 and spring 2006) K-1
Table K-2	Incremental Effects of CLIO Parenting Curricula on Children (Research Question 2) (average of spring 2005 and spring 2006) K-1

List of Tables (continued)

	<u>Page</u>
Table K-3	Effects of CLIO Combined Curricula on Parents (Research Question 1) (average of spring 2005 and spring 2006) K-2
Table K-4	Incremental Effects of CLIO Parenting Curricula on Parents (Research Question 2) (average of spring 2005 and spring 2006) K-2
Table K-5	Effects of CLIO Combined Curricula on Instruction (Research Question 1) (average of spring 2005 and spring 2006) K-3
Table K-6	Incremental Effects of CLIO Parenting Curricula on Instruction (Research Question 2) (average of spring 2005 and spring 2006) K-3
Table K-7	Effects of CLIO Combined Curricula on Participation (Research Question 1) (average of spring 2005 and spring 2006) K-4
Table K-8	Incremental Effects of CLIO Parenting Curricula on Participation (Research Question 2) (average of spring 2005 and spring 2006) K-4
Table L-1	Results of Overall Test for Any Differences in Growth Across the Five Study Groups (Fall 2004 to Spring 2005).....L-3
Table L-2a	Effects of CLIO Combined Curricula on Children (Research Question 1) (Growth Fall 2004 to Spring 2005).....L-4
Table L-2b	Effects of CLIO Combined Curricula on Children (Research Question 1) (Growth Fall 2004 to Spring 2005).....L-4
Table L-3a	Incremental Effects of CLIO Parenting Curricula on Children (Research Question 2) (Growth Fall 2004 to Spring 2005).....L-5
Table L-3b	Incremental Effects of CLIO Parenting Curricula on Children (Research Question 2) (Growth Fall 2004 to Spring 2005).....L-5
Table L-4a	Effects of CLIO Combined Curricula on Parents (Research Question 1) (Growth Fall 2004 to Spring 2005).....L-6

List of Tables (continued)

	<u>Page</u>
Table L-4b. Effects of CLIO Combined Curricula on Parents (Research Question 1) (Growth Fall 2004 to Spring 2005).....	L-6
Table L-5a Incremental Effects of CLIO Parenting Curricula on Parents (Research Question 2) (Growth Fall 2004 to Spring 2005).....	L-7
Table L-5b Incremental Effects of CLIO Parenting Curricula on Parents (Research Question 2) (Growth Fall 2004 to Spring 2005).....	L-7
Table M-1 Results of Overall Test for Any Interaction of Curriculum with Ethnicity or Home Language (Combined Data from Spring 2005 and Spring 2006).....	M-2
Table N-1 Simulation Study Results for Fidelity-Adjustments	N-5
Table O-1 P-Values for Relationships Between Amount of Participation in Parenting Education and Parent Outcomes	O-2
Table O-2 P-Values for Relationships Between Participation in Preschool Education and Child Outcomes.....	O-4
Table O-3 Results of Overall Test for Differences Among Children with High Participation Levels Across the Five Groups (Combined Spring 2005 and Spring 2006) (High participation = 420 hours or more over 7 months).....	O-8
Table O-4a Contrasts with the Control Group of CLIO Combined Curricula on Children with High Participation Levels (Research Question 1) (Combined Spring 2005 and Spring 2006).....	O-9
Table O-4b Contrasts with the Control Group of CLIO Combined Curricula on Children with High Participation Levels (Research Question 1) (Combined Spring 2005 and Spring 2006).....	O-9

List of Tables (continued)

	<u>Page</u>
Table O-5a	Contrasts Among Experimental Groups with and without CLIO Parenting Curricula on Children with High Participation Levels (Research Question 2) (Combined Spring 2005 and Spring 2006) O-10
Table O-5b	Contrasts Among Experimental Groups with and without CLIO Parenting Curricula on Children with High Participation Levels (Research Question 2) (Combined Spring 2005 and Spring 2006) O-10

List of Figures

	<u>Page</u>
Figure ES-1 Effect Sizes for CLIO Combined Curricula on Child Outcomes Relative to Both the Control Group and the CLIO Preschool Curricula (average of spring 2005 and spring 2006)	xiii
Figure ES-2 Effect Sizes for CLIO Combined Curricula on Parent Outcomes Relative to Both the Control Group and the CLIO Preschool Curricula (average of spring 2005 and spring 2006)	xiv
Figure ES-3 Effect Sizes for CLIO Combined Curricula on Instructional Outcomes Relative to Both the Control Group and the CLIO Preschool Curricula (average of spring 2005 and spring 2006)	xv
Figure ES-4 Effect Sizes for CLIO Combined Curricula on Participation Relative to Both the Control Group and the CLIO Preschool Curricula (average of spring 2005 and spring 2006)	xvii
Figure 1-1 CLIO Model of Effects for Even Start Projects	6
Figure 2-1 Flow of Even Start Projects Through CLIO Recruitment and Randomization	15
Figure 2-2 CLIO Cross-Sectional Sample Sizes	27
Figure 3-1 Relationship of Measuring Fidelity of Implementation and Overall Instructional Practices	54
Figure 3-2 Percentage Distribution of Fidelity Ratings of CLIO Treatment Classrooms by Classroom Type for Spring 2005 and Spring 2006: Observer Ratings	58
Figure 3-3 Percentage Distribution of Fidelity Ratings of CLIO Treatment Projects by Classroom Type for Spring 2005 and Spring 2006: Developer Ratings	59

List of Figures (continued)

	<u>Page</u>
Figure 3-4	Distribution of CLIO Children by Average Monthly Hours of ECE Participation, Spring 2005 and Spring 200661
Figure 3-5	Distribution of CLIO Parents by Average Monthly Hours of PE/PC Participation, Spring 2005 and Spring 200662
Figure 5-1	Effect Sizes for CLIO Combined Curricula on Child Outcomes Relative to Both the Control Group and the CLIO Preschool Curricula (average of spring 2005 and spring 2006)86
Figure 5-2	Effect Sizes for CLIO Combined Curricula on Parent Outcomes Relative to Both the Control Group and the CLIO Preschool Curricula (average of spring 2005 and spring 2006)90
Figure 5-3	Effect Sizes for CLIO Combined Curricula on Instructional Outcomes Relative to Both the Control Group and the CLIO Preschool Curricula (average of spring 2005 and spring 2006)95
Figure 5-4	Effect Sizes for CLIO Combined Curricula on Participation Relative to Both the Control Group and the CLIO Preschool Curricula (average of spring 2005 and spring 2006)97
Figure 7-1	Relationship Between Developer-Rated ECE Fidelity and Emergent Child English Literacy by Study Group121
Figure 7-2	Relationship Between Observer-Rated ECE Fidelity and Emergent Child English Literacy by Study Group123
Figure J-1	Progression of Study-Eligible Children: Spring 2005 and Spring 2006 CombinedJ-3
Figure J-2	Progress of Study Eligible Parents: Spring 2005 and Spring 2006 CombinedJ-4
Figure O-1	Relationship of Parent Responsiveness (to Child) to Average Hours of Participation in Parenting Education O-3

List of Figures (continued)

	<u>Page</u>
Figure O-2 Relationship of Parent Interactive Reading Skill to Average Hours of Participation in Parenting Education	O-3
Figure O-3 Relationship of Emergent Literacy to Average Hours of Participation in Preschool Education (Spring 2005 and Spring 2006 Child Literacy, Average ECE Participation Across September 2004 through March 2005 and September 2005 through March 2006—All Study Groups)	O-5

1. BACKGROUND

In this chapter, we discuss the Even Start program, the purpose of the current study, and the organization of the remainder of the report.

The Even Start Program

The Even Start Family Literacy Program was established in 1989 with the goal of improving the academic achievement of low-income young children and their parents, especially in the area of reading (U.S. Department of Education 2003). Even Start projects offer four integrated instructional activities for low-income families:

- age-appropriate early childhood education to prepare children for success in school and life experiences (early childhood education, or ECE);
- training for parents regarding how to be the primary teacher for their children (parenting education, or PE);
- interactive literacy activities between parents and their children (parent-child literacy activities, or PC); and
- parent literacy training that leads to economic self-sufficiency (adult education, or AE).

The underlying premise of Even Start, and of the family literacy model more generally, is that these four instructional components are necessary for improved child literacy and are maximally effective when integrated into a unified program. That is, child language and literacy should be improved directly, through participation in ECE, and indirectly through improvements in both parenting skills and parent literacy. Parenting skills are expected to be improved through participation in PE and PC activities, and parent literacy through participation in AE. Also, Even Start services¹ are to be of “sufficient intensity in terms of hours, and of sufficient duration, to make sustainable changes in a family.”

¹ Family literacy services are defined in Sec. 203 of Title II of *The Workforce Investment Act of 1998*, Public Law 105-220, also known as the *Adult Literacy and Family Education Act of 1998*.

Even Start Has Not Performed Up To Expectations

Since 1989, the U.S. Department of Education (ED) has sponsored three national evaluations of the Even Start program that focused on performance and effectiveness. Two random assignment studies that were part of these evaluations (St.Pierre et al. 1995; St.Pierre et al. 2003) showed that Even Start projects were not effective at improving the literacy skills of participating preschool-age children and their parents. That is, literacy gains made by Even Start parents and children were no different from literacy gains made by control parents and children. The control group for these randomized studies was composed of parents who wanted to enroll their children in Even Start but who were randomly assigned to participate in Even Start in the year following the evaluation. About two-thirds of these control parents were unable to arrange any other formal ECE services during the period of the evaluation, so the control condition mostly corresponded to at-home care by parents or extended family members.²

The absence of significant effects on literacy skills prompted an examination of the Even Start model to determine how it could be improved. The lead investigators of the most recent national evaluation of Even Start (St.Pierre, Ricciuti, and Rimdzius 2005) addressed several questions that might explain Even Start's apparent ineffectiveness: (1) whether the Even Start model was fully implemented, (2) whether Even Start's instructional services were sufficiently intensive, (3) whether Even Start families participated sufficiently, and (4) whether the quality of Even Start's instruction and curriculum content was sufficient to lead to positive effects.

Data from the national evaluation showed that Even Start projects were, indeed, able to fully implement the program, and that Even Start projects offered instructional services at a level of intensity that is comparable to mainstream programs offering the individual parts of a family literacy program. However, the evaluation also documented that Even Start families participated at low levels and for a relatively short period of time. Further, evaluation data showed that Even Start's instructional services were not of uniformly high quality. In particular, the national evaluation found that (1) Even Start's early childhood education programming was not of higher quality than the instruction received by control children and was not of higher quality than the

² See page 162 of St.Pierre et al. (2003).

instruction received by Head Start children, (2) the adult education programs provided to Even Start parents varied widely in their quality and the extent to which they focused on literacy, and (3) the parenting education programs offered by Even Start projects were similar in content and delivery systems to mainstream parenting programs, which have little research evidence on the extent to which they are effective at enhancing either parenting skills or child literacy. On the basis of these findings, the evaluators suggested that one promising avenue for improving Even Start would be to increase the extent to which Even Start's instructional services focus on literacy.

The CLIO Study: Seeking to Improve Even Start

The Even Start Classroom Literacy Interventions and Outcomes (CLIO) Study is the third randomized study of Even Start. Prior studies have investigated the effectiveness of Even Start relative to control groups in which parents and children were not enrolled in Even Start. In contrast, the CLIO study examined the effectiveness of four different curricular packages against the "regular" Even Start program. These curricular packages featured research-based literacy instruction. Two of the packages focused solely on early childhood education instruction, while the other two packages combined instruction in early childhood education with instruction in Even Start's two parenting components. This approach is supported by the strengthened mandate of Even Start from the *Literacy Involves Families Together Act (LIFT 2001)* and the *No Child Left Behind Act (NCLB 2001)*, which call for Even Start projects to provide

- high-quality, intensive instructional programs,
- instructional programs based on scientifically based reading research, and
- reading readiness activities based on scientifically based reading research.

In addition, the CLIO study is consistent with Even Start's second legislative evaluation requirement, which is "to identify effective Even Start programs ... that can be duplicated and used in providing technical assistance to Federal, State, and local

programs.³ The CLIO design is also consistent with the research goals and methods with respect to the components of successful family literacy services:⁴

Preschool and parenting instruction were manipulated to construct four distinct experimental curricula:

- Two CLIO combined curricula that focused on child literacy both in preschool and parenting instruction. These were each a combination of a CLIO preschool curriculum and a CLIO parenting curriculum.
- Two CLIO preschool curricula that had an intense focus on child literacy but left parenting instruction alone. These were each a combination of a CLIO preschool curriculum and whatever approach to parenting instruction was already in use at the Even Start projects.

In keeping with Even Start's four-component family literacy approach, all of the Even Start projects participating in the CLIO study continued to provide AE to parents, but the AE instruction was not changed as part of the CLIO study.

The CLIO study addresses two primary research questions:

- 1) Is the combination of research-based, literacy-focused preschool, parenting, and parent-child curricula (the CLIO combined curricula) more effective than the existing combination of services in Even Start?
- 2) Do research-based parenting and parent-child curricula (the CLIO parenting curricula) that focus on child literacy add value to the CLIO preschool curricula?

Thus, CLIO is an evaluation of the incremental effectiveness of providing these research-based literacy-focused instructional services, over and above the existing instruction provided by Even Start projects.

³ Evaluation goal #2 under Sec. 1239 of the *No Child Left Behind Act of 2001*, Public Law 107-110.

⁴ Sec. 1241 of the *No Child Left Behind Act of 2001*, Public Law 107-110.

The conceptual model for the CLIO study (figure 1-1) builds on the premise that Even Start improves child outcomes both directly (through ECE participation) and indirectly (through parenting and adult education). The oblongs on the left hand side of the model illustrate the instructional services that CLIO sought to improve (ECE, PE, and PC). The next three boxes show that the CLIO curricula are hypothesized to improve the instructional practices of staff working with Even Start children and parents. Improved instructional practices are hypothesized to lead to short-term enhancements, by the end of preschool, in children's development and in parent behaviors and skills. Improvements in parenting skills also were hypothesized to enhance children's development.⁵ The AE instruction provided to Even Start parents was not changed as part of the CLIO study; however, the family literacy model assumes that AE produces improvements in parents' literacy and educational levels, which contribute to enhanced child development.⁶ This relationship is represented in the three boxes along the bottom of the model.

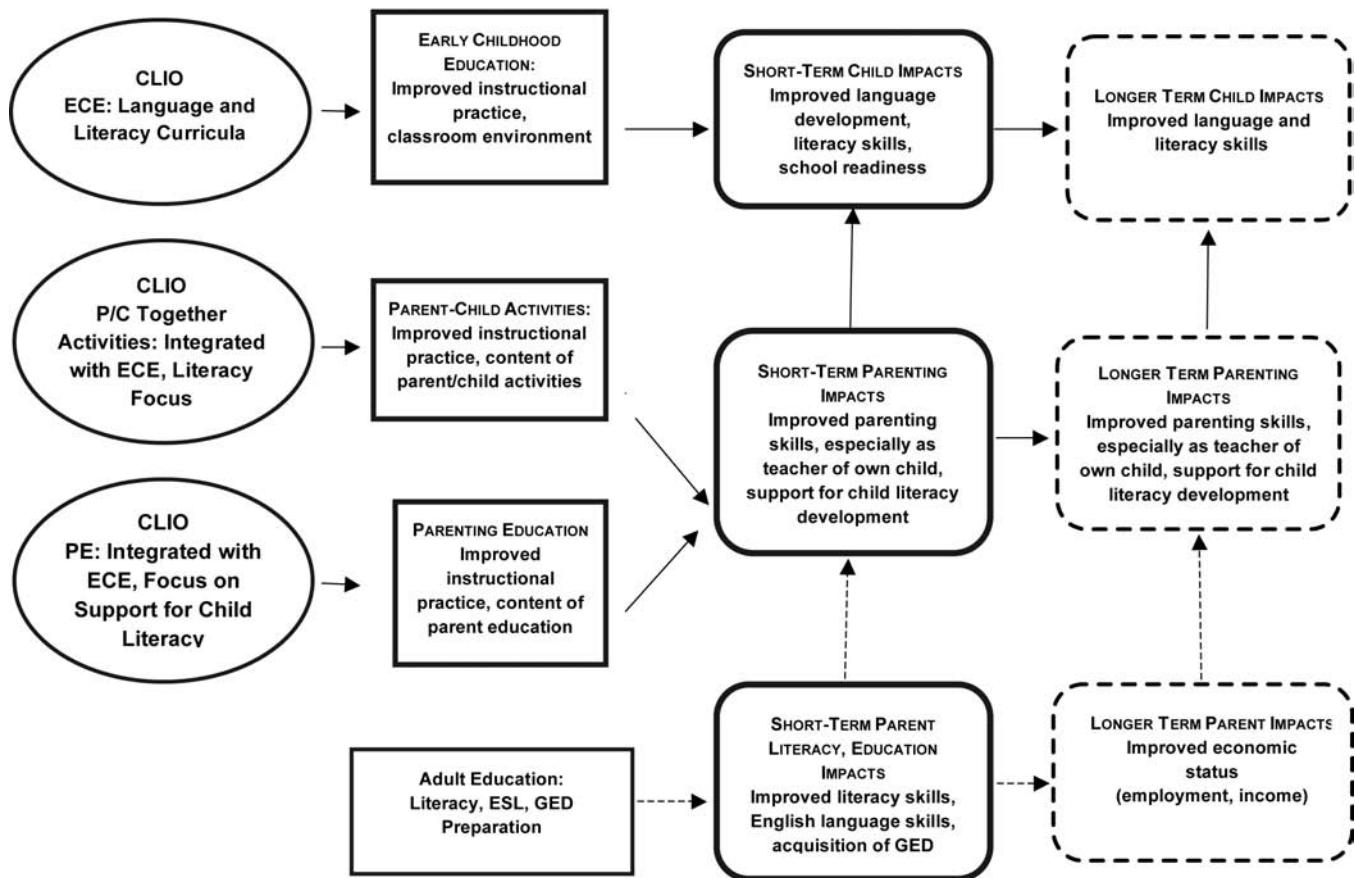
The model includes longer term impacts, since the hypothesis is that changes in children's development and skills by the end of preschool will result in improved reading and language skills in the early school grades. Longer term changes in parenting skills and in parents' improved literacy and education are hypothesized to support these improved child outcomes. IES is interested in exploring whether there is support for these hypotheses, and so is conducting a follow-up study to address the question: Do the CLIO curricula produce positive effects at the end of kindergarten or at the end of first grade?⁷

⁵ Although Even Start serves families with at least one child between birth and age 8, and most projects serve children throughout this entire age range, the study focused on preschool-age children and their families. At the time the study was designed, the Even Start program office focused on school readiness by attempting to improve the effectiveness of educational services for 3- and 4-year-olds. CLIO did not include infants and toddlers because we felt that (1) there was no conclusive evidence that formal instruction in language and literacy is helpful for that age group, (2) there was an absence of systematic curricula for children in this age group, and (3) there was only limited information about how services for infants and toddlers were administered in Even Start. CLIO did not include school-age children since Even Start's role for school-age children is predominantly one of coordination with public schools. It was deemed unlikely that Even Start projects would be able to bring about curriculum changes in the public schools.

⁶ The study team decided not to include Even Start's adult education component in the test of research-based curricula because (1) most projects provided a variety of adult education services at different levels (ABE, GED, ESL) to meet family needs, (2) a substantial portion of projects used community service providers to deliver adult education services, and (3) the research on effective adult education models is still in its infancy.

⁷ Followup data collection with children in kindergarten and first grade is being conducted, and findings from that data collection will appear in a later report.

Figure 1-1. CLIO Model of Effects for Even Start Projects



Organization of the Report

In the remainder of this report we fully describe the research design (chapter 2), provide a description of the CLIO curricula (chapter 3), discuss the methods for the analyses (chapter 4), and present the findings of our analyses (chapters 5 through 7).

2. RESEARCH DESIGN

In this chapter, we describe the selection and implementation of the CLIO curricula, recruitment and random assignment, the CLIO projects at baseline, the data collection schedule and methods, the sample design, and the development of the outcome measures.

Selection of the CLIO Curriculum Developers

The hypothesis underlying the CLIO study is that an increased focus on literacy in preschool and parenting instruction would improve parent and child outcomes for Even Start families. To select interventions that were literacy focused and based on research, a public process was used in which developers of preschool and parenting curricula were invited to submit proposals for review by an expert panel. The Request for Proposals was prepared, and proposals were solicited in spring 2003 from curriculum developers. Eight proposals were received, and the authors of the four highest rated proposals were invited to make oral presentations to the expert panel. The expert panel rated the proposals on several key criteria, including the quality of the proposed intervention, capability of the institution to meet the requirements of the study and bring the interventions to scale, and staff qualifications and experience.

The key criterion (worth 55 of the 100 possible points) related to the quality of the proposed intervention. Under this criterion, the proposals were judged on the extent to which the content of the interventions (both preschool and parenting components) focused on literacy (specifically the domains of oral language, phonological awareness, print recognition and conventions of print), was appropriate for the Even Start population, and integrated the preschool and parenting components. The proposals were also judged on the evidence that the interventions were effective (particularly with populations similar to Even Start).

The selection was based primarily on the ratings of the intervention content, specifically whether the content proposed was linked—either by previous research or expert judgment—to the study's targeted outcomes. Although the four highest rated proposals thoroughly documented their curricula's grounding in the research literature on emergent literacy, rigorous evidence of effectiveness was not extensive. After the

oral presentation round of the selection process, the expert panel judged three of the four remaining proposals as acceptable for inclusion in the study. Two proposals were chosen from the three based on the strength of the parenting component of the intervention.

The two curriculum developers that were selected each has a preschool curriculum coupled with an integrated parenting curriculum. The developers were responsible for implementing their curricula in two modes: (1) preschool only, and (2) preschool and parenting combined. The curricula tested in CLIO were the following:

- **Let's Begin with the Letter People® and Play and Learning Strategies (PALS)**—The CIRCLE group at the University of Texas-Houston Health Sciences Center teamed with Abrams & Company Publishers to develop and implement:
 - (1) Preschool only: Let's Begin with the Letter People, a preschool curriculum that is built around 26 imaginary characters that represent the letters of the alphabet. Let's Begin was augmented with teacher training from CIRCLE on effective practices in early literacy.
 - (2) Preschool and parenting combined: Let's Begin was linked with the Play and Learning Strategies (PALS) parenting curriculum. PALS was developed by CIRCLE for parents whose children are at risk for developmental delay and academic failure due to poverty, low family literacy, and other risk factors and teaches parents to understand where their child is on the developmental continuum and what techniques they can use to build their children's language skills, cognitive development, and school readiness.
- **Partners for Literacy**—The University of North Carolina at Chapel Hill, Frank Porter Graham Child Development Center, developed and implemented:
 - (1) Preschool only: ECE Partners for Literacy, a preschool curriculum based on game-like activities and interactive book reading conducted with pairs of children and designed to promote language development and emergent literacy. The curriculum is based on an earlier version that was used in the Abecedarian project.

- (2) Preschool and parenting combined: ECE/PE Partners for Literacy, which links the preschool curriculum with an integrated parenting education curriculum that uses many of the same activities, strategies, and materials as the preschool curriculum.

Effectiveness and Appropriateness of Selected Curricula

Let's Begin and PALS. At the time of selection, Let's Begin had been implemented and field tested in preschool classrooms with populations similar to Even Start across the country. The developers cited the positive results of an ongoing efficacy study conducted by Abrams & Company, the curriculum publisher. The curriculum has take-home materials available in Spanish, and teacher materials contain suggestions for adapting activities for English language learners. PALS was developed for disadvantaged families and had been implemented in both English and Spanish. The developer cited positive results from a recent randomized controlled experimental study (Landry, Smith, and Swank 2003). Most PALS materials are available in both English and Spanish.

Partners for Literacy. Partners for Literacy materials were developed for children from low-income families and at the time of selection had been used in preschool settings with populations similar to Even Start. Developers cited positive results from three randomized, controlled longitudinal research studies: the Abecedarian Project (Ramey et al. 1976), project CARE (Wasik, Ramey, Bryant, and Sparling 1990), and the Infant Health and Development Program (Ramey et al. 1992). Professionals and parents received materials for cultural responsiveness, and all materials for parents are available in English and Spanish. Teacher training also included time and materials devoted to teaching English language learners.

Establishment of the Five Study Groups for the CLIO Study

One hundred twenty Even Start projects were recruited and randomly assigned to one of five study groups: one of the four CLIO curricula or an "as is" control group that provided their regular pre-CLIO instructional services (table 2-1). Even Start's four instructional components define the rows of the table, while the four CLIO

Table 2-1. Specification of the Five CLIO Study Groups

	Study group				
	LET'S BEGIN with the Letter People (ECE)	LET'S BEGIN with the Letter People and Play and Learning Strategies (PALS)(ECE/PE)	Partners for Literacy (ECE)	Partners for Literacy (ECE/PE)	
Even Start instructional component	CLIO preschool curriculum	CLIO combined curriculum	CLIO preschool curriculum	CLIO combined curriculum	Control
Early childhood education	LET'S BEGIN	LET'S BEGIN	Partners for Literacy	Partners for Literacy	As usual
Parenting education	As usual	PALS	As usual	Partners for Literacy	As usual
Parent-child joint literacy activities	As usual	PALS	As usual	Partners for Literacy	As usual
Adult education	As usual	As usual	As usual	As usual	As usual
NOTE: Shaded areas identify instructional components that were provided by the CLIO curriculum developers.					

experimental groups and a control group define the columns. Shaded table cells identify instructional components that were provided by the CLIO curriculum developers. The CLIO study groups have the following characteristics:

- Study Group 1: Even Start projects were assigned the Let's Begin preschool curriculum. These projects provided their usual PE, PC, and AE instructional services.
- Study Group 2: Even Start projects were assigned both the Let's Begin preschool curriculum and the PALS parenting curriculum. These projects provided their usual AE instructional services.
- Study Group 3: Even Start projects were assigned the Partners for Literacy preschool curriculum. These projects provided their usual PE, PC, and AE instructional services.
- Study Group 4: Even Start projects were assigned both the Partners for Literacy preschool curriculum and the parenting curriculum. These projects provided their usual AE instructional services.

- Study Group 5: Even Start projects assigned to the control group provided each of the four instructional components as usual. This was an “as is” or “business as usual” control group.

Implementation of the Curricula

Once selected, the curriculum developers modified their existing curricula as appropriate, developed materials, prepared professional development and implementation plans, and piloted their curricula with a small number of Even Start projects in 2003-2004. Curricula were revised on the basis of the pilot test, and plans were made for large-scale implementation. In summer 2004, a 4-day centralized training session was held for each curriculum, attended by project directors and teachers from the assigned Even Start projects. In addition, the curriculum developers provided each participating Even Start project with on-going support and technical assistance over the life of the study.

Implementation of the CLIO curricula in Even Start classrooms began in the 2004-2005 school year, and during that year, ongoing support was provided to each project through telephone calls and on-site visits. In summer 2005, a second centralized training was held to retrain Even Start staff who had been trained in the previous summer and to provide an initial training to new staff members. Each of the four CLIO curricula was then implemented for a second year during 2005-2006, again with ongoing support of the curriculum developers. The implementation phase of CLIO concluded at the end of the 2005-2006 school year.

Design of the Study to Address the Research Questions

The CLIO study was designed to address two primary research questions:

- **Research Question 1:** Is the combination of research-based, literacy-focused preschool, parenting, and parent-child curricula (the CLIO combined curricula) more effective than the existing combination of services in Even Start?

- **Research Question 2:** Do research-based parenting and parent-child curricula (the CLIO parenting curricula) that focus on child literacy add value to the CLIO preschool curricula?

The first question was addressed analytically by combining projects in the second and fourth columns of table 2-1 (those assigned to the CLIO combined curricula), and comparing their outcomes with those of control projects (the fifth column). The second question was addressed analytically by combining projects in the second and fourth columns of table 2-1, and comparing their outcomes with those of projects in the first and third columns (those assigned to the CLIO preschool curricula).

In addition to these two primary research questions, the CLIO study examined several secondary questions:

- *Instructional Practices:* To what extent are particular preschool instructional practices associated with better child outcomes?
- *Parenting Practices:* To what extent are parenting practices associated with better child outcomes?
- *Fidelity of Implementation:* How much variation was there in the faithfulness with which CLIO projects implemented the assigned curricula? Were child and parenting outcomes better in projects with higher fidelity to their assigned curriculum?
- *Participation:* To what extent is participation associated with better outcomes?

Recruitment of Even Start Projects

Recruitment of Even Start projects for the CLIO study began with a careful screening of projects to determine which ones met the study's eligibility requirements. To be eligible for CLIO, an Even Start project had to

1. serve preschool children in a center-based instructional setting,
2. enroll a minimum of either five 3- and 4-year-olds in one center-based classroom, or eight 3- and 4-year-olds in two center-based classrooms,

3. provide at least 12 hours per week of center-based preschool instruction,
4. serve a majority of families who speak either English or Spanish,
5. be able to exert control over the curricula used in preschool classrooms, and
6. be willing to meet the study requirements, including being randomly assigned to one of the five study groups.

Exerting control over preschool curricula was an eligibility criterion because Even Start requires projects to build on existing services, where possible, to avoid duplication. Thus, many projects do not directly provide all of Even Start's instructional services, but rather, coordinate with other programs to provide some services. For example, during the latest national evaluation (St.Pierre, Ricciuti, and Rimdzius 2003), 22 percent of Even Start 3- and 4-year-olds who participated in center-based ECE received these services from Head Start programs. Projects that outsourced their instructional services in this way were not excluded from participating in CLIO, but few chose to do so, since the Even Start grantee often did not have control over the preschool curriculum. Hence, most CLIO projects were ones that provided their own preschool instruction. The study team did not, however, screen for eligibility based on (1) serving children and their parents in a center-based setting for the provision of parent-child activities or parenting education or (2) providing a specified number of hours per week of parent-child activities or parenting education instruction.

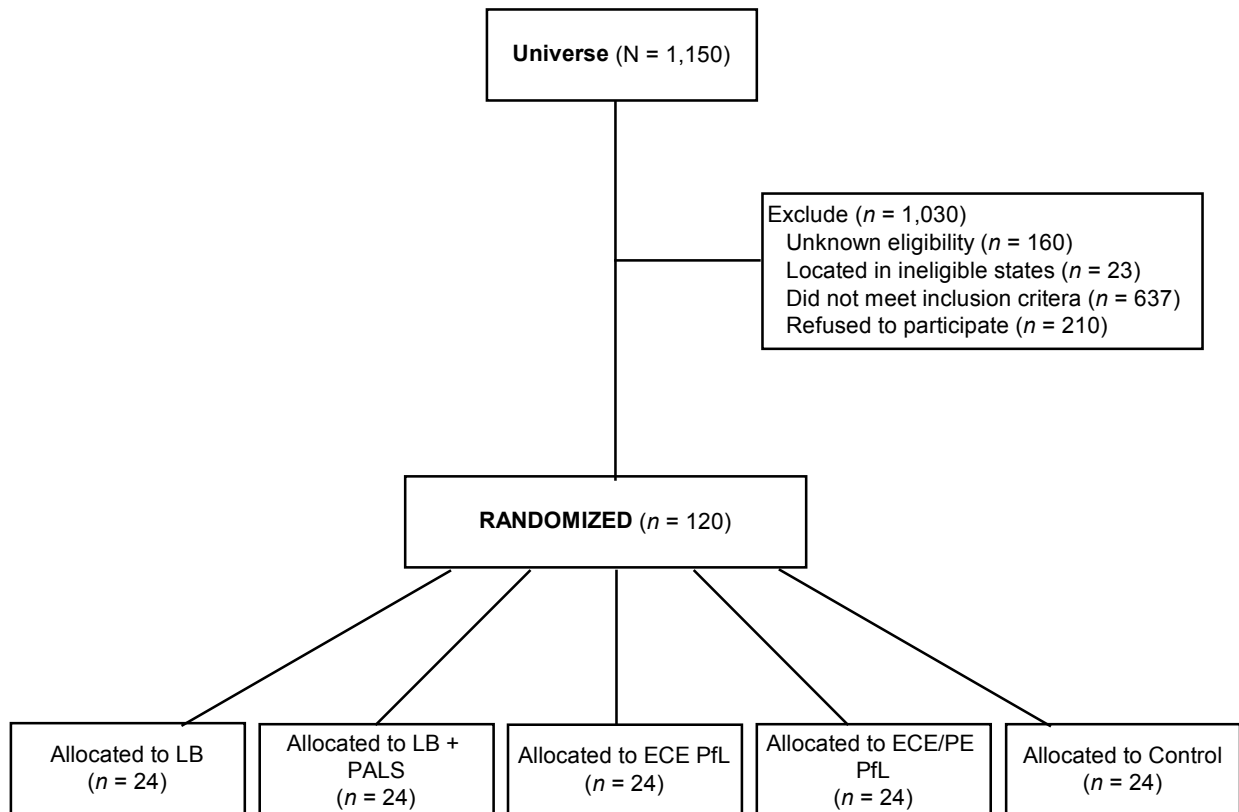
In spring 2003, 1,150 Even Start projects operated throughout the United States. Telephone calls were attempted with 1,127 of these projects, excluding 23 projects in Alaska, Hawaii, and Puerto Rico that were defined as ineligible due to the distances and prohibitive costs that would be associated with data collection. Telephone surveys were completed with 967 (86 percent) of the 1,127 Even Start projects in eligible states. Of these, 637 were ineligible to participate in the study for reasons such as not serving a sufficient number of preschool-age children, not offering preschool-age children at least 12 hours per week of center-based preschool instruction, serving primarily families that spoke languages other than English or Spanish, or not having a center-based ECE program.

During the screening process, we relaxed the requirements to broaden the eligibility pool by allowing (1) the enrollment of fewer numbers of children in center-based classrooms and (2) the provision of at least 10 hours (lowered from 12) of center-based preschool instruction. At the conclusion of the screening process, 330 Even Start projects were deemed eligible.

Of the eligible projects, 120 were willing to participate in the study and were randomly assigned to one of the five study groups. (See figure 2-1 for the flow of the projects through the recruitment process.)

CLIO was designed and implemented as a real-world study of literacy focused, research-based curricula in Even Start settings. Nationwide, Even Start settings vary widely on every aspect of the program. Even Start projects are mandated to offer instructional services that include early childhood education, adult literacy education, parenting education, and structured literacy interaction between parents and their children. They also are required to avoid duplication of services by building on existing community resources such as local adult education programs or Head Start. Even Start projects decide on the frequency and duration of instruction, whether instruction is primarily center-based or home-based, and whether to invent educational curricula from scratch, use published curricula, or use a hybrid of approaches. Based on the availability of local instructional services and the extent to which those services are perceived to be of high quality, project staff decide which activities will be supported by Even Start funds and which will be provided by collaborating agencies. Projects offer screening and referral services such as referrals for mental health counseling, services to battered family members, employment services, and screening or treatment for chemical dependency. Projects also offer support services such as transportation, flexible scheduling, childcare, nutrition assistance, health care, and meals to help families participate in the program. The CLIO curricula were implemented in a sample of 120 Even Start projects that were spread over 33 states and reflected this diversity. As a result, CLIO could not be a tightly controlled laboratory-type experiment.

Figure 2-1. Flow of Even Start Projects Through CLIO Recruitment and Randomization



NOTES:

All groups had 24 projects providing at least 1 year of data for the combined spring 2005 and spring 2006 analysis.

The study did not collect data in both implementation years (2004-05 and 2005-06) from 8 of the original 120 projects because (1) they dropped out of the study prior to the first implementation year (2004-05) and were replaced for 2005-06, (2) they lost funding prior to the first implementation year (2004-05) and were replaced for 2005-06, or (3) they lost funding for the second implementation year (2005-06) and had data for 2004-05 only.

The number of projects included in both years of data collection ranged from 21 to 23 projects out of 24.

Random Assignment of CLIO Projects

The CLIO random assignment plan focused on ways to minimize pre-existing differences among the five study groups. Before random assignment, 24 strata were formed, each containing exactly five projects. The variables used to form the strata were (1) size of project (number of 3- and 4-year-olds served), (2) proportion of children who were Spanish speakers, (3) year that the project was up for recompetition, and (4) region. (See table 2-2 for the cut points for each of these stratification variables.)⁸

Table 2-2. Stratification Variables Used in Random Assignment

Stratification variables	Categories
Size of project	3 categories: • large, defined as more than 28 3- and 4-year-olds • medium, defined as 10 through 28 3- and 4-year-olds • small, defined as 9 or fewer 3- and 4-year-olds
Proportion of Spanish-speaking children	3 categories: • very large, defined as more than 25 percent Spanish speakers • medium, defined as 8 percent through 25 percent Spanish speakers • small, defined as less than 8 percent Spanish speakers
Last year of current grant	3 categories: • 2003-2004 school year • 2004-2005 school year • other
Region	4 standard Census categories

The highest priority was placed on size of project. The large category contained only 13 projects. Among these, we generally found either a very large proportion or a very small proportion of Spanish speakers, so within it we created just two strata, substratified only on percentage Spanish while ignoring the other two variables. The other three large projects were then mixed in with the medium projects.

⁸ Simply crossing these four variables would have created 108 strata, more than could be used. Accordingly, extensive collapsing of preliminary strata was required. Because of the requirements of exactly five projects per stratum, we also sometimes had to make small changes in the thresholds.

Within the medium and small categories, we were able to use more categories of percentage Spanish and to pay some attention to the other variables. Once the 24 strata were finalized, the five projects in each were randomly assigned to the five study groups in early 2004.

A comparison of the resulting five groups of projects (mostly in terms of variables collected at the spring 2004 baseline) showed that the random assignment plan resulted in well-matched study groups. There were no statistically significant differences among the five groups on 55 of 58 spring 2004 variables that were examined (see tables A-1 through A-5 in appendix A).⁹ Additionally, there were no statistically significant differences among the five groups on 23 of the 25 variables examined in spring 2005 and spring 2006 (see table A-1)¹⁰

Study Projects at Baseline

The voluntary nature of participation and the eligibility criteria for the study meant that the CLIO sample was not nationally representative of Even Start projects. However, the 120 recruited projects were located in 33 states in all regions of the country and varied on characteristics such as population density, number of families served, percentage of families who are English language learners, and number of years as Even Start projects. In this section, we present descriptive statistics on the study sample, both to set a context for the evaluation and to provide a basis for assessing the study's external validity.

Race/Ethnicity. In spring 2004, prior to implementation of the CLIO curricula, 57 percent of CLIO children were identified as Hispanic. In 2000-2001 (the most recent period with national data), 46 percent of Even Start parents nationally were Hispanic (table 2-3). Although the CLIO sample is not nationally representative, the

⁹ Two of these three variables were included among the covariates chosen for the impact analysis. See chapter 4 for a full list of covariates.

¹⁰ Two types of tests were used for testing for baseline balance across the study groups. For baseline item response theory (IRT) scores, hierarchical linear modeling (HLM) was used with a two-level setup (project and child) and no covariates other than strata. For all other types of variables, a stratum-adjusted Kruskal-Wallis test was run on project-level averages. No weights were used at the project level. Multinomial variables like race were transformed into a series of binary recodes, each of which was tested separately.

movement from 46 percent to 57 percent continues a long trend of increases in the percentage of Hispanic families served by Even Start.

Table 2-3. Percentage Distribution of CLIO Children and Even Start Parents by Race/Ethnicity

	CLIO children spring 2004	Even Start parents 2000-01
Race/ethnicity		
White	22	30
Black	11	19
Hispanic	57	46
Other	9	5

NOTE: Detail may not sum to totals due to rounding.

SOURCES: U.S. Department of Education, Even Start Classroom Literacy Interventions and Outcomes Study, "Parent Interview," Spring 2004; U.S. Department of Education, Planning and Evaluation Service, Elementary and Secondary Education Division, *Third National Even Start Evaluation: Program Impacts and Implications for Improvement*, Washington, DC: 2003.

Maternal Education. As of spring 2004, 38 percent of CLIO children had mothers with a high school diploma, GED, or higher (table 2-4). In 2000-2001, 15 percent of new Even Start parents had this level of education. Additionally, in 2003-2004, 24 percent of new Even Start participants had a high school diploma, GED, or higher. The national statistics are based on parents from newly entering Even Start families, whereas the CLIO statistic is based on information about mothers from all Even Start families in each project, some of whom many have increased their education attainment as a result of participating in Even Start. Since prior research has shown that Even Start has a positive impact on GED attainment, it is not surprising that education attainment based on all families in the program, including those that have participated for many months, would be different from educational attainment based only on newly entering families. Of course, there may be other reasons as well for the high educational attainment of CLIO mothers relative to national Even Start figures.

Hours of Instruction. Amount of instruction offered and received is a statistic that was calculated in prior Even Start studies. The CLIO projects are similar to the 2000-2001 national sample in terms of participation in PE and PC activities. CLIO parents participated in PE and PC activities for an average of 10 hours a month in

Table 2-4. Percentage Distribution of Educational Attainment for Mothers of CLIO Children and New Even Start Parents

	Mothers of CLIO children spring 2004	New Even Start parents 2000-2001	New Even Start parents 2003-2004
Educational attainment			
HS, GED, or higher	38	15	24
Without HS/GED	62	84	76

NOTE: Detail may not sum to totals due to rounding.

SOURCES: U.S. Department of Education, Even Start Classroom Literacy Interventions and Outcomes Study, "Parent Interview," Spring 2004; U.S. Department of Education, Planning and Evaluation Service, Elementary and Secondary Education Division, *Third National Even Start Evaluation: Program Impacts and Implications for Improvement*, Washington, DC: 2003; U.S. Department of Education, Consolidated State Performance Report School Year 2004-05.

2003-2004, similar to the 2000-2001 national estimate of 11 hours a month (table 2-5). In 2003-2004, projects in the CLIO study offered preschool-age children an average of 84 hours of instruction each month, fairly similar to the national statistic of 76 hours a month in 2000-2001 (table 2-5). Children in the CLIO sample participated in preschool instruction an average of 42 hours a month during 2003-2004. Although this was only half of the amount offered to them, it nevertheless is greater than the national Even Start average of 33 hours a month of participation in preschool education in 2000-2001 (table 2-5).

Teacher Education. In spring 2004, 82 percent of the lead preschool teachers, 22 percent of the preschool aides, and 89 percent of the lead PE teachers in the CLIO sample had an associate's degree or higher (table 2-6).

Classroom Instruction. At baseline most CLIO projects (71 percent) reported that they used at least one formal curriculum in their preschool classrooms. These projects cited a wide range of instructional programming, including published comprehensive curricula, literacy-focused supplemental curricula, state curriculum frameworks, skills assessments linked to instructional strategies, informal non-published curricula, and local or other reading initiatives. Of the projects that used a formal curriculum, close to 70 percent reported using either High Scope or Creative Curriculum. About half (55 percent) reported that they used at least one formal curriculum for their parenting education sessions. Of those, about 40 percent used Parents as Teachers. Most projects (78 percent) reported that they did not use any formal curriculum in their parent-child interactive sessions.

Table 2-5. Average Monthly Hours of ECE Instruction Offered and Received and Average Monthly Hours of PE/PC Instruction Received, for CLIO and Even Start

	CLIO 2003-2004	Even Start 2000-2001
Projects ¹		
Hours of ECE instruction offered per month	84	76
Children ¹		
Hours of ECE instruction received per month	42	33
Parents		
Hours of parenting education and parent-child joint activities received per month	10	11

¹ Even Start hours for 2000-2001 represent instruction for 3- and 4-year-olds.

SOURCES: U.S. Department of Education, Even Start Classroom Literacy Interventions and Outcomes Study, "Instructional Services Participation Form, and Project Director Survey," Spring 2004; U.S. Department of Education, Planning and Evaluation Service, Elementary and Secondary Education Division, *Third National Even Start Evaluation: Program Impacts and Implications for Improvement*, Washington, DC: 2003

Table 2-6. Educational Attainment of Even Start Staff in CLIO Projects: Spring 2004

Staff position	Percent
Lead preschool teacher	
Educational attainment	
Less than associate's degree	18
Associate's degree or higher	82
Preschool aide	
Educational attainment	
Less than associate's degree	78
Associate's degree or higher	22
Lead PE teacher	
Educational attainment	
Less than associate's degree	11
Associate's degree or higher	89

SOURCE: U.S. Department of Education, Even Start Classroom Literacy Curricula and Outcomes Study, "Staff Survey," Spring 2004.

At baseline, all CLIO projects were observed with the Early Childhood Environment Scale-Revised Edition (ECERS-R; Harms, Clifford, and Cryer 1998) to assess the quality of the classroom environment, including use of space, materials and experiences to enhance children's development, schedule, and supervision. Specifically, the ECERS-R assesses 37 items that cover the six classroom areas: space and furnishings; personal care routines; oral language and reasoning skills; fine motor, gross motor, and creative activities; interactions among children and between children and staff; and program time and structure. Each item is ranked on a scale of 1 (inadequate conditions) to 7 (excellent conditions), with 3 representing minimal conditions and 5 representing good conditions.

Across the 37 items, CLIO projects scored an average of 4.88 on the ECERS-R measure, suggesting that, on average, CLIO projects have good classroom environments. No project received an "inadequate conditions" rating. Approximately 70 percent of projects received a rating of 4 or 5. These data suggest that CLIO classrooms are comparable to Head Start preschool classrooms. According to the Head Start Family and Child Experiences Surveys (FACES), the average overall ECERS-R score for Head Start classrooms was 4.91 in spring 2001 and 4.81 in fall 2003 (U.S. Department of Health and Human Services 2006).

Even Start Continues to Work With a Needy Population. The data presented in this section show that Even Start families in the CLIO sample continue to face many of the same difficulties that were identified in the Third National Even Start Evaluation (St.Pierre et al. 2003). Income and education levels are low compared to most of America: 58 percent of Even Start families in the CLIO projects have monthly income below \$1,500, and only 38 percent of Even Start mothers in the CLIO projects have a high school diploma. Furthermore, adult English fluency is poor. In spring 2004, 59 percent of CLIO parents report a native language other than English. Among the non-native speakers, just 7 percent claim to speak and understand English very well (compared to 88 percent of the native speakers). Literacy is also poor. Overall, just 40 percent of CLIO parents claim to read English very well. These self-reports of low fluency and literacy are borne out by the assessments. In particular, Even Start parents in the CLIO projects scored quite low on the Peabody Picture Vocabulary Test (PPVT)—70 on average using publisher norm scores with a national mean of 100 and a national standard deviation of 15.

Given the unexpectedly large number of CLIO mothers with a high school diploma, GED, or higher educational attainment (relative to the most recent available national Even Start figures), we also looked at the breakdown of published PPVT scores by mother's education. We found that even among mothers with higher educational attainment, receptive vocabulary scores were very low: mothers who attended some college had an average PPVT score of 84, those with a high school diploma averaged 77, and those without a high school diploma averaged 63.

Data Collection Schedule and Methods

In this section, we discuss the CLIO study's data collection schedule and provide an overview of the data collected by data collection cycle (see table 2-7).

The CLIO study's data collection schedule was as follows:

- **2003-2004: baseline year.** We collected baseline data on 3- and 4-year-olds and their parents in all Even Start projects participating in the CLIO study during the 2003-2004 project year. We also observed classroom instruction and collected information about the project in the spring of the baseline year.
- **2004-2005: first year of implementation.** We collected data on 3- and 4-year-olds and their parents in all Even Start projects participating in the CLIO study during the 2004-2005 project year. We also observed classroom instruction and collected information about the project in the spring of the first implementation year.
- **2005-2006: second year of implementation.** We collected data on 3- and 4-year-olds and their parents in all Even Start projects participating in the CLIO study during the 2005-2006 school year. We also observed classroom instruction and collected information about the project in the spring of the second implementation year.

Data were collected from (1) preschoolers (3- and 4-year-olds), (2) their parents, (3) classrooms, and (4) projects. Here we briefly describe the types of data collected and the methods for collecting these data. In the next section, we discuss the creation of outcome measures based on these data.

Table 2-7. Overview of Data Collection

	Data collection instrument	Baseline year		First implementation year		Second implementation year
		Fall 2003	Spring 2004	Fall 2004	Spring 2005	Spring 2006
Preschoolers						
Assessment	Child Assessment Battery	X	X	X	X	X
Social-emotional	Teacher–Child Rating Form	X	X	X	X	X
Videotape	Read Aloud Together Profile			X	X	X
Participation	Instructional Services Participation Form	X	X	X	X	X
Parents						
Assessment	Parent Assessment Battery	X	X	X	X	X
Interview	Parent Interview	X	X	X	X	X
Videotape	Read Aloud Together Profile			X	X	X
Participation	Instructional Services Participation Form	X	X	X	X	X
Classrooms						
Observation of instruction	Observation Protocols		X		X	X
Observation of fidelity	Observation Protocols				X	X
Survey	Teacher Survey		X		X	X
Projects						
Survey	Project Director Survey		X		X	X

Data Collected From Preschoolers

At each data collection cycle, trained field staff administered a battery of one-on-one child assessments. The battery covered the following domains: a test of expressive language, in both English and Spanish; a test of receptive vocabulary; two tests of phonological awareness, Elision and Blending; a test of print knowledge; and a test of syntax and grammar. These assessments are described in detail later in this chapter (see section entitled child outcomes). The study team conducted a week-long training prior to each data collection cycle to prepare field staff for data collection.

Assessments were conducted at each Even Start project, in a setting provided by the Even Start project staff.

Each preschooler's teacher was asked to complete a Teacher-Child Rating (TCR) form at each data collection cycle. The TCR captured information on each child's behavior and social skills. More information about the TCR is provided later in this chapter (see section entitled child outcomes). The trained field staff distributed and collected the TCRs while on-site.

Another aspect of the data collection was the videotaping of each preschooler and parent during the three data collection cycles in the two implementation years. Trained field staff videotaped the parent and child engaged in a book reading activity and in playing with a toy. Both the book and the toy were supplied by the field staff, who were trained in videotaping during a week-long training and given a script to follow. The videotaping took place at each Even Start project, in a setting provided by the Even Start project staff. More details on the videotaping can be found later in the chapter (see section entitled Parent Outcomes).

Finally, each preschooler's hours of participation in preschool education in the Even Start project was collected from Even Start project staff using a template developed for the CLIO study referred to as the Instructional Services Participation Form and described later in the chapter (see section entitled Instructional Outcomes). Projects submitted the participation information to the study team monthly.

Data Collected From Parents

At each data collection cycle, trained staff administered a battery of one-on-one parent assessments as well as a parent interview. The battery covered receptive vocabulary, basic reading skills, and comprehension. More detail on the assessments is provided later in the chapter (see section entitled Parent Outcomes). The interview collected parents' self-reported information about their reading and language activities with their child, the home literacy environment, their ratings of their child's behavior and social skills, and parent demographics. The study team conducted a week-long training prior to each data collection cycle to prepare field staff for data collection. (The same field staff conducted the child assessment, the parent assessments, the parent

interviews, and the videotaping.) Assessments and interviews were conducted at each Even Start project, in a setting provided by the Even Start project staff.

The videotaping is discussed above under data collected from preschoolers.

Finally, each parent's hours of participation in parenting education and parent-child activities in the Even Start project were collected from Even Start project staff using a template developed for the CLIO study referred to as the Instructional Services Participation Form and described later in the chapter (see section on Instructional Outcomes). Projects submitted the participation information to the study team monthly.

Data Collected From Classrooms

In the spring of each year, preschool education classes, parenting education classes, and parent-child activities classes were observed by trained staff. The study team provided extensive training in the use of the observation measures. There were two sets of staff for the observations, each trained separately: one set for the preschool classes, and one set for the parenting and parent-child classes. The observation protocols were designed to collect information on instructional practices in the classrooms and on the fidelity of implementation to the CLIO curricula. Observation of instructional practices is described in detail later in this chapter (see section on Instructional Outcomes), and observation for fidelity of implementation is discussed in chapter 4.

Each spring, teachers were asked to complete a short survey to provide information on their educational background, demographics, and professional development opportunities. The observers distributed and collected the teacher surveys while on-site to conduct the classroom observations.

Data Collected From Projects

Each Even Start project director was administered a survey to collect information regarding services in each of the four Even Start components, including numbers of families and children served, hours offered, and curricula used. Project directors were also asked to provide information on their educational background,

demographics, and professional development opportunities. The project director survey was conducted as a mail survey.

Sample Sizes

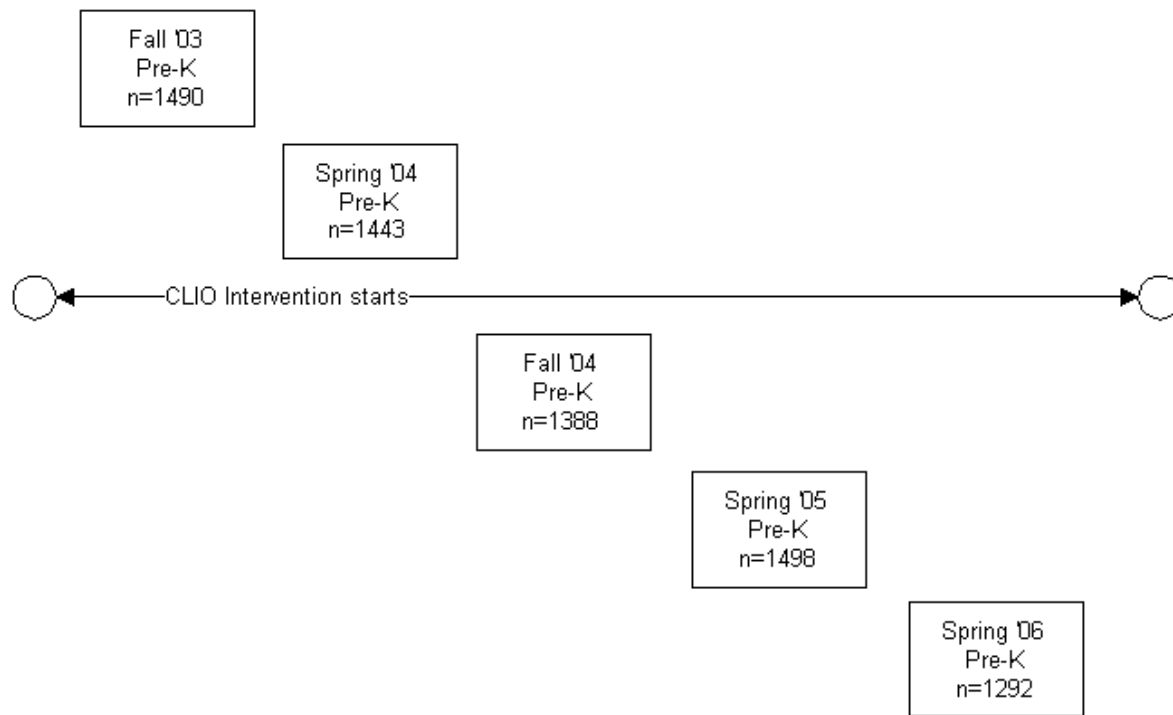
Children enrolled at CLIO projects were generally eligible for participation in the CLIO study if they were between 36 and 60 months of age at the time of assessment and were not yet attending kindergarten.¹¹ Children whose attendance at CLIO centers overlapped multiple data collection periods were assessed multiple times, but there was no effort to make the sample longitudinal. Sample sizes for analysis purposes are shown in figure 2-2 and table 2-8.

Development of Outcome Measures

The CLIO study collected a large amount of information in many different outcome domains to fully address the study's research questions. There are three broad measurement categories: (1) child outcomes, (2) parent outcomes, and (3) instructional outcomes. (Table 2-9 shows the outcome measures and the constructs they were selected to measure.) Child and parent outcomes align directly with the primary CLIO research goals of improving child language and literacy and parenting practices. While the instructional outcomes can be viewed as either mediating variables or as outcomes in their own right, the principal analysis for this study treated the instructional process variables as outcomes. However, secondary analyses were run with non-experimental techniques to explore the relationships between instructional process variables and child and parent outcomes.

¹¹ For spring data collection (2004, 2005, and 2006), the child must have turned 3 no later than March 1. For fall data collection (2003 and 2004), the child must have turned 3 no later than October 1. In spring 2006, children who had been assessed in spring 2005 and who were old enough to be in kindergarten were not included in the preschool sample, whether or not they were still in preschool. This change was made so that these children could participate in the follow-up data collection with the instruments designed for kindergarten students.

Figure 2-2. CLIO Cross-Sectional Sample Sizes



NOTE: Figure counts are limited to children who took at least one of the child assessments and were enrolled at a project for at least 28 days as of the assessment date.

Table 2-8. Cross-sectional Child Sample Size by Child Age and Data Collection Cycle

Child age at assessment	Baseline		CLIO curricula in implementation		
	Fall 2003	Spring 2004	Fall 2004	Spring 2005	Spring 2006
3	574	367	571	400	353
4	726	658	685	683	603
5 and 6	190	418	132	415	336
Total	1,490	1,443	1,388	1,498	1,292

NOTE: Counts are limited to children who took at least one of the child assessments and were enrolled at a project for at least 28 days as of the assessment date.

Table 2-9. CLIO Outcome Measures

	Outcome		Data collection instrument	Mode of data collection	Domain	
CHILD	1	Expressive language: English	Individual Growth and Development Indicator (IGDI)	Child assessment	Emergent literacy	
	2	Expressive language: Spanish				
	3	Receptive vocabulary	Peabody Picture Vocabulary Test (PPVT)			
	4	Phonological awareness: Elision	Comprehensive Test of Phonological and Print Processing (Preschool—CTOPPP)			
	5	Phonological awareness: Blending				
	6	Print knowledge				
	7	Syntax and grammar	Test of Language Development (TOLD-3)			
	8	Social competence	Teacher rating form	Teacher rating	Socio-emotional development	
PARENT	9	Parent interactive reading skill	Read Aloud Together Profile & Parent Interview	Video observation, parent report	Parenting skills	
	10	Parent responsiveness				
	11	Reading & vocabulary skill	Parent assessment battery	Parent assessment	Parent language & literacy	
INSTRUCTIONAL	12	Support for oral language development	Observation Measures of Language and Literacy Instruction (OMLIT) and Parenting Education and Child/Parent Observation (PECAP)	Classroom observation	ECE classroom instruction	
	13	Support for print knowledge				
	14	Support for phonological awareness				
	15	Support for print motivation				
	16	Literacy resources in classroom			PE classroom instruction	
	17	PE time spent on child literacy				
	18	PE time spent on parenting skills				PC classroom instruction
	19	PC time spent interacting on child literacy activities				
	20	Child: Monthly hours of ECE instruction received	Instructional Services Participation Form (ISPF)	Project report	Participation amount	
	21	Parent: Monthly hours of PE and PC instruction received				

A key consideration in creating outcome measures was the total number to construct. The experimental curricula were expected to possibly affect a broad range of child, parent, and instructional outcomes. This argued for a large number of outcome measures. However, false positive findings can be caused by running a large number of statistical tests. This argued for being parsimonious in selecting outcome measures. A target of about 20 outcomes was set so that the expected number of false positive

findings in the event of no true effects would not be more than one. Multiple comparison adjustments were used for the variety of contrasts of interest pertinent to each outcome, but these adjustments were not applied across outcomes, as to do so would have lowered statistical power too precipitously.

Multiple literacy subtests, targeting different aspects of literacy, were used for both parents and children. The literacy subtests were averaged together for parents but left separate for children. In both cases, there are substantial correlations among the tests, but given that the CLIO curricula do not systematically vary adult literacy education, parent subtests were averaged to reduce multiple comparison problems. The children's subtests were kept separate partly because of the different theory behind each subtest and partly to facilitate subsequent meta-analyses since other studies tend to report them separately.

Child Outcomes

Even Start projects provide ECE to children in low-income families to prepare them for success in school. Hence, the CLIO curricula were selected, in large part, for the strength of their preschool curricula, in particular the language and literacy dimensions. The CLIO child assessment battery was designed to measure early language and literacy development, including vocabulary, phonological awareness, and print knowledge. Research has shown these areas are important in the development of reading skills and predictive of school achievement (National Research Council 2001).

On each of the child outcome tests (except the IGDI and Spanish IGDI), multiple scoring procedures were applied, including a simple count of items correct as well as complex scoring similar to what is done in the National Assessment of Educational Progress (NAEP), the Early Childhood Longitudinal Study – Birth Cohort (ECLS-B), and several other large-scale testing programs. (Appendix B provides descriptions of the two scoring procedures.) Where the complex scoring was carried out, the scores were scaled to have a mean of 250 and a standard deviation of 50. The raw scores were left on the metric represented by the number of items asked. For all contrasts between study groups, treatment effects are expressed in terms of standard deviations within the control group.

Below we describe each of the outcome measures. See appendix B for psychometric data for each test.

Expressive Language. The IGDI Picture Naming subtest (Early Childhood Research Institute on Measuring Growth and Development 2003) measures expressive language by asking the child to recognize and name a series of common objects in English using picture cards. Following the test publisher's standard procedure, the assessor counted the number of cards the child named correctly in 1 minute, but did not track which cards the child failed to name correctly. The protocol called for the subtest to be administered to all children regardless of native language. A parallel subtest, the Spanish IGDI, was administered in Spanish to children from Spanish-speaking families.

Receptive Vocabulary. The Peabody Picture Vocabulary Test (PPVT-III) measures receptive vocabulary (Dunn and Dunn 1997) and has been widely used in other early childhood studies. In this subtest, the child demonstrates his or her understanding of the meaning of an English word by pointing to the correct picture. CLIO used a version of this test that was adapted by Westat for this study. This adapted version contained fewer items, and the words were divided into three sets. In the first set for every child, there were 14 words. If the child made fewer than three errors in the first set, then he/she was given an additional set (the ceiling set) of 10 more difficult words. At the other extreme, if the child made more than seven errors in the first set, then he/she was given an additional set (the basal set) of eight easier words.

Phonological Awareness: Elision. The Preschool CTOPPP (Lonigan, Wagner, Torgesen, and Rashotte 2002) was developed to measure phonological awareness in English in younger children. The Elision subtest measures the child's ability to recognize English word parts, such as components of compound words, syllables, and phonemes. The examiner reads a compound word and the child is asked to identify what is left when part of the word is taken away. The part taken away can be a whole word from a compound word (e.g., "toothbrush without brush") or phoneme (e.g., team without /m/). There were 18 items in this subtest. For the first nine items, the child identified what was left by pointing at a picture of it on a page with four pictures. For the last nine items, the child was asked to verbalize the remainder without visual aids. The first nine items were given to each child regardless of the child's error rate. After the first nine items, a run of three consecutive errors caused the subtest to be stopped.

Phonological Awareness: Blending. The blending subtest of the Preschool CTOPPP measures the child's ability to combine English word parts, such as components of compound words, syllables, and phonemes. The examiner says two parts of an English word and asks the child to put them together (e.g., "horse and shoe together is horseshoe"). There were 21 items in this subtest. In the first nine items, the child was asked to identify compound words by pointing at the answer from a page with four choices. The child went through all nine items regardless of the number of errors. For the final 12 items, the picture support was not available, and there was a skip-out rule tied to three consecutive errors.

Print Knowledge. The Print Awareness subtest of the Preschool CTOPPP assesses the child's ability to identify Roman alphabet symbols that represent letters and words used in English, to identify specific letters by name and by sound, and to produce letter sounds. On this subtest, there was no stopping rule. All children were asked all items.

CLIO used a research version of the Preschool CTOPPP available in the study's first data collection year (2003-2004). However, a slightly revised version of the test with normed scores has since been published by ProEd as the Test of Preschool Early Literacy (TOPEL). The TOPEL Phonological Awareness test combines the CTOPPP Elision and Blending subtests, contains fewer items, and uses different stopping rules. The TOPEL Print Knowledge test contains the same items as the CTOPPP Print Awareness subtest, but administers the items in a different order with different administration rules.

Syntax and Grammar. The TOLD-3 Grammatical Understanding subtest (Newcomer and Hammill 1997a; 1997b) measures the child's ability to comprehend the meaning of an English sentence, with an emphasis on syntax and morphology. Knowledge of syntax is important in constructing and understanding sentences. In this subtest, the assessor read a sentence aloud, and the child was asked to select one picture from three possible choices that correctly corresponded to the sentence. There were 24 items in this subtest. Six consecutive errors caused the subtest to be stopped.

Child Social Competence. Developing children's social competence is an important objective for early childhood programs, and the development of social skills and positive behaviors is associated with success in school. Positive behavior includes

cooperation with adults, friendly play, and sharing with other children. Problem behaviors include disruptive or overly aggressive behavior, hyperactivity, excessive shyness and social withdrawal. These negative behaviors are associated with problems in school and/or receipt of psychological help (Gresham and Elliott 1990).

A social competence scale was created from preschool teacher reports on the behavior and social skills of children in the study. The scale combines information from two different sets of items in the CLIO Teacher's Rating Form—cooperative behavior and problem behavior. (Appendix C contains details on the construction of the child social competence scale.)

Parent Outcomes

Two types of parent outcomes were of interest: improved parenting skills and improved parent literacy. The hypothesis was that both help parents to be their child's first teacher. As with instructional outcomes, parent outcomes were analyzed both as outcomes and as mediators for child outcomes.

Parenting Skills. The parenting curricula implemented by CLIO projects focused on showing parents how to be effective teachers of their child and emphasized teaching early reading skills. It has been shown (Whitehurst and Lonigan 1998) that the practice of specific behaviors during joint book reading can promote children's engagement in reading and help them better comprehend the story and understand the conventions of print. Mutual questioning and responding, making stories relevant to the child's life, giving praise and feedback, explaining, physically sharing the book, monitoring a child's understanding, and adjusting language are all behaviors that enhance children's literacy skills and comprehension. Given the second primary research question of determining the added value of a parenting curriculum with a focus on child literacy, we developed the instruments for measuring parenting behavior with a particular focus on those aspects of parenting that theory suggested should promote child literacy.

Parenting skills were measured by coding videotaped parent-child interactions and by parent self-report. Both of these measured parenting behaviors that were thought to be important in differentiating parents who were more or less effective teachers (See appendix D for more information on the coding of the interactions.) A

total of 90 variables describing parenting behaviors were measured in the spring of 2005 and 2006, so some distillation was required. Instead of sorting the variables *a priori* into groups based on the literature of the field, the data on these variables were empirically combined into two outcome scales:

- Parent interactive reading skill, and
- Parent general responsiveness to the child.

The procedures used in this process included variable clustering and factor analysis within clusters. These procedures result in an unequal weighting of the items assigned to each scale. The scale for parent interactive reading skill has 49 items, while the scale for parent general responsiveness has 41. The correlation between the two scales is 0.6.

Despite the lack of *a priori* grouping, these two scales align fairly well with two primary dimensions of teacher quality at the pre-k and elementary level recently identified by Hamre and Pianta (2005): instructional support and emotional support. (See appendix D for details on how these scales were created.) They also align well with the goals set for parents by the CLIO parenting curricula.

There is no prior information on the reliability or validity of these scales. Some of the relationships explored in chapter 7 support the validity of the second scale in terms of its relationship with the targeted child outcomes. In appendix D, we provide information on the training of the coders and the ways in which we established rater reliability.

Parent Language and Literacy. Improving parent literacy is one of Even Start's main goals. While AE was not varied as part of the CLIO curricula, parent language and literacy have a strong relationship with child outcomes. Further, parent literacy may be positively affected by participation in parenting education with a child literacy focus.

The CLIO parent assessment was designed to measure English language and literacy outcomes, including vocabulary, basic reading, phonics, and comprehension. A single outcome measure was created from four tests: the PPVT and three Woodcock-

Johnson subtests (Letter-Word Identification, Passage Comprehension, and Word Attack). Using spring 2004 data (first plausible value of each IRT score), Cronbach's alpha for the scale was 0.95 (see appendix B for information about scoring procedures and psychometric properties for each of these subtests).

The Peabody Picture Vocabulary Test (PPVT-III) (Dunn and Dunn 1997) measures receptive vocabulary. In this test, the parent demonstrates his or her understanding of the meaning of a word by pointing to the correct picture after the test administrator reads a test word aloud.

Three Woodcock-Johnson subtests were administered (Woodcock, McGrew, and Mather 2001). Letter-Word Identification measures basic reading skills and requires respondents to identify printed letters and words with an oral response. There are 76 items in this subtest. Word Attack measures the subject's skill in applying phonic and structural analysis skills to the pronunciation of unfamiliar printed words. The subject reads aloud letter combinations that are linguistically logical but that form nonsense words or low-frequency words in English. There are 32 items in this subtest. Passage Comprehension measures comprehension and vocabulary skills. In the first several items, respondents point to the picture represented by a phrase. The remaining items require reading a short passage and supplying an appropriate answer for a missing key word. There are 47 items in this subtest. The test publisher's skip-out rule of stopping after six errors in a row was used for all three subtests.

Instructional Outcomes

If the curricula work as designed, significant changes in instructional practices were expected to occur. These changes were hypothesized to be necessary (if not sufficient) for impacts on children and parents. Knowing whether such changes occurred would be useful in interpreting the level of impacts on children and parents. Moreover, an understanding of how instruction changed in treatment projects is important to designers of future curricula. Although there was not a strong hypothesis about whether introducing research-based, literacy focused curricula would increase participation levels, level of participation was measured, since it was also a possible mediator of impacts on children and parents. Accordingly, systems were developed to measure instructional practices and participation.

Instructional practices were evaluated based on direct observation by experienced education researchers using standardized schedules, as is described below in more detail. As with parenting, many variables were created from these observational measurement systems, and so some distillation was required. From these detailed measurements, we developed eight instructional outcomes. Of these, five focus on ECE, two on PE, one on PC. In addition, based on monthly information from projects, we developed two participation measures—one for children’s participation in ECE classes and one for parent participation in PE and PC classes. The alignment of particular measurements with the scales was straightforward because the intended scales guided the development of the measurement systems. However, not all the measurement elements worked as intended; some of these were dropped from the scales as discussed in appendices E and F to improve the reliability of the scales.

Preschool Instruction. The CLIO preschool curricula were intended to promote aspects of language development and emergent literacy skills that have been shown to be essential to proficient reading. As described by Whitehurst and Lonigan (1998), the elements of emergent literacy form two domains: (1) inside-out processes, which are rule-driven processes for rendering the written symbols of text into sound and vice versa (i.e., decoding and encoding), and (2) outside-in processes, which are sources of information outside the printed text, such as vocabulary, background knowledge, and contextual knowledge, that support understanding (and decoding) the text. Other terms used to describe emergent literacy are oral language and print motivation (outside-in), and phonological processing/sensitivity and print knowledge (inside-out). These terms are defined as follows:

- *oral language*: lexical/conceptual, semantic, and syntactic abilities,
- *print motivation*: interest in reading and writing activities,
- *phonological processing*: sensitivity to and ability to manipulate word sounds, and
- *print knowledge*: knowledge of units of print (letters, words), ability to translate print to sound and sound to print (letter-sound, and ultimately word-sound).

The primary mechanism hypothesized to improve developmental outcomes for Even Start children in CLIO is a preschool curriculum that focuses on teaching these

skills. The two preschool curricula were selected for CLIO, in part, because they include instructional activities in all four areas of emergent literacy, albeit differing in relative attention across areas.

To assess whether the curricula were successful in changing teaching activities in the critical areas of language and literacy, a classroom observation measure was developed specifically for the CLIO study: the Observation Measures of Language and Literacy Instruction, or OMLIT (Goodson, Layzer, Smith, and Rimdzius 2004, 2006). The OMLIT is a battery of six measures that focus on aspects of classroom practice that have been shown in research to support children's language development and acquisition of early literacy skills. Appendix E describes the development of and rationale for the OMLIT and provides psychometric information on the battery. There is no prior information on the reliability or validity of these measures.

Classroom observations using the OMLIT were conducted each spring in CLIO classrooms. Along with the ECE OMLIT, observers also completed the Arnett Rating of Caregiver Behavior for each lead teacher (Arnett 1989). The Arnett rated the teacher's engagement with, responsiveness to, and affect toward children in the classroom. In the spring 2004 baseline data collection only, observers also completed the ECERS-R, which rates overall classroom quality in six classroom areas. (See earlier discussion about the ECERS-R under classroom instruction, pages 22-23.)

Five outcome constructs were derived from the six ECE OMLIT measures to correspond to key elements of preschool instruction that are being manipulated by the curricula. (Appendix E provides details on how the preschool instructional outcome measures were constructed.) These included constructs for the four components of emergent literacy, and an additional construct—the adequacy of language and literacy resources in the classroom—which is commonly considered to be related to children's emergent literacy (although no strong research evidence exists to support this claim). The five instructional outcome variables are the extent to which the preschool classroom provides:

- support for oral language,
- support for phonological awareness,

- support for print knowledge,
- support for print motivation, and
- adequacy of literacy resources in the classroom.

PE Instruction. Even Start requires that each project provide PE instruction, designed to increase parents' knowledge about early childhood development and parenting behaviors and practices, toward the objective of helping parents contribute actively and constructively to the literacy development and school readiness of their children. Prior research has shown that Even Start projects often use PE to provide instruction in many areas that are not directly related to child literacy or how to effectively interact with their children. Examples of topics in other areas include adult life skills, household management, health, and nutrition (St.Pierre et al. 2003). Given that the core strategy for the CLIO parenting curricula was to intensify the focus on child literacy, the study team decided to measure how parenting education time was allocated across these three broad areas: how parents can directly promote child literacy, how parents can interact more effectively with their children (abbreviated as "parenting skills"), and other topics.

An observational measure called the Parenting Education and Child and Parent Observation (PECAP) was created specifically for the CLIO study to measure the amount and type of activities being undertaken both in PE and PC classes. Appendix F provides details on the development of and psychometric information for the PECAP. There is no prior information on the reliability or validity of these measures.

Based on observations made each spring with the PECAP, two outcomes were created for PE classes:¹²

- the percentage of PE class time spent on child literacy activities: reading/looking at books/letters; writing/emergent writing; oral language, songs, rhymes, sound games; and

¹² The PECAP also records time spent on adult-focused activities, such as parent health and on other activities such as play activities, but these variables were not used in the CLIO impact analysis.

- the percentage of PE class time spent on parenting skills: responding to and managing child behavior; home-school relations; ideas for home play; child development; child health, well-being, safety.

Most important, from the perspective of this evaluation, is the percentage of PE class time spent on child literacy, since that variable is closely aligned with the central thrust of the CLIO parenting curricula, as well as with Even Start's broader guidance (U.S. Department of Education 2003) that PE should be directed at skills that allow the parent to be the "primary teachers for their children." The percentage of PE class time spent on parenting skills also is important because these activities may, in the long run, contribute to a parent's ability to be a good teacher for his/her children. The PECAP was revised between spring 2004 and spring 2005, and there were considerable missing data in spring 2004, so no baseline data were available for this measure.

PC Literacy Activities. The legislation authorizing Even Start requires that projects provide interactive literacy activities for parents and their children (U.S. Department of Education 2003). These activities may take place in preschool classrooms, as part of PE classes, in separate PC classes, or during home visits. As with PE, prior Even Start evaluations have shown that many Even Start projects use PC time for a wide assortment of non-literacy activities (St.Pierre et al. 2003).

The PECAP observational measure (described above in the discussion of PE outcome variables) was used to describe the amount and type of PC literacy activities in CLIO projects. Unlike PE, where the focus is on child literacy activities in which parents are being shown how to be their child's teacher, in PC literacy activities the focus is on literacy activities where the parent gets the opportunity to practice what he/she learned in PE by interacting with his/her child. The PECAP records whether each activity involves (1) parents only, (2) children only, (3) both parents and children, with parent/child pairs interacting together, or (4) parents and children but without any interaction in parent/child pairs. The outcome variable created from the PECAP data was

- the percentage of PC time in which parents and children were interacting on activities that were directly related to child literacy.

As with PE outcomes, changes in the PECAP from spring 2004 to spring 2005 and missing data in spring 2004 mean that no baseline measurement was available for this measure.

Extent of Participation in Even Start. It was hypothesized that the amount of Even Start instructional services received by CLIO parents and children might vary by study group. Perhaps families in the experimental groups were more engaged by the CLIO curricula and therefore missed fewer days of Even Start than families in the control group. The Instructional Services Participation Form (ISPF) was used to collect participation information from CLIO projects monthly. CLIO project staff were asked to report the number of hours of instruction in which each child and parent participated in each of Even Start's four instructional components (ECE, PE, PC, AE). Projects submitted data via an on-line ISPF data collection system, email, and fax. Editing and follow-up were performed on an on-going basis. Two outcome variables were constructed from the ISPF data:

- the number of hours per month that a child participated in ECE, and
- the number of hours per month that a parent participated in PE and PC.

These two scales were created by counting the hours that a child or parent participated in Even Start across a 9-month period (September through May)¹³ and then dividing the total by nine—regardless of whether the participation was all within a single month or spread more evenly across months.¹⁴ Children enrolled for only brief periods have low participation scores, as do children enrolled for longer periods but with rare attendance. Children with high weekly attendance over a long enrollment period have high participation scores.

¹³ Although Even Start is intended to be a year-round program, projects may either not provide services in the summer, or the services may differ greatly from those provided during the school year. Reports from the summer months were particularly ambiguous. To reduce the impact of this ambiguity on the participation outcomes, participation was counted only from September through May of each year.

¹⁴ We also have a 7-month version of each participation measure. The 7-month versions average participation hours from September through March. The 7-month versions were used as putative causal agents in some analyses in chapter 7 and thus need to avoid containing data about the period following the literacy assessments. Since chapter 3 is about the process that could have affected results, the 7-month version is also used there. The 9-month versions were used as intervention outcomes in chapters 5 and 6.

3. CURRICULUM IMPLEMENTATION

This chapter discusses the implementation of the CLIO curricula. Our examination of CLIO curriculum implementation includes the following components:

- description of the curricula,
 - What was the intended curriculum—that is, the desired teaching practices in the preschool and parenting education classes?
 - What types of training and support were provided to project staff to help support high-fidelity implementation of each curriculum?
- fidelity of implementation of the curricula,
 - How was fidelity measured for each curriculum?
 - To what extent were the curricula implemented as planned?
- level of exposure of children and parents to the curricula, and
- types of instruction provided in the control group.

Answering questions about fidelity of implementation requires a definition of what constitutes the critical components of each curriculum (content and process) and a method for measuring the match between the component as intended and the component as implemented. These activities pose challenges for any study but especially for CLIO, where multiple curricula were implemented across a large and heterogeneous sample of classrooms.

Curricula as Intended

The following descriptions of the curricula expand on the descriptions in chapter 2.

Let's Begin with the Letter People plus Play and Learning Strategies

This combined curriculum consists of the Let's Begin with the Letter People preschool curriculum described below plus the Play and Learning Strategies parenting curriculum.

Let's Begin with the Letter People-Preschool Curriculum. As designed for the CLIO study, this curriculum included Let's Begin with the Letter People (Abrams and Company Publishers), a published early childhood education classroom curriculum, augmented by CIRCLE teacher training on developmentally appropriate techniques for promoting language and literacy skills in preschool (CIRCLE is the Center for Improving the Readiness of Children for Learning and Education, the state center for early childhood development at the University of Texas, Houston Health Sciences Center).

Let's Begin with the Letter People. This literacy program for pre-kindergarten children adheres to the research findings that the most effective instruction for preparing children to become lifelong readers and learners builds progressively on children's understanding and use of both spoken and written language, specifically focusing on the four critical domains of early literacy:

1. Oral language, oral comprehension, and vocabulary;
2. Phonological and phonemic awareness;
3. Letter recognition; and
4. Conventions of print.

The unique feature of Let's Begin with the Letter People is that it is built around 26 imaginary characters that represent the letters of the alphabet. The curriculum uses these characters to help children learn about letters, sounds, and concepts.¹⁵ The curriculum is organized into five themes: *All About Me; Getting Along With Others; Everyone Has Needs; Animals, Animals, Animals; and Nature All Around Us.* Each resource book contains five or six lessons, each dealing with a different facet of the

¹⁵ Letter People are inflatable characters, representing the letters of the alphabet. They wear a capital letter on their front and a lower case letter on their back to support letter identification. Each has a distinguishing characteristic that is readily associated with the sound of their letter (e.g., Mr. N has a noisy nose and likes to eat noodles)

theme, and is introduced with the help of a different Letter Person. For example, the All About Me theme contains the following six units:

1. What's My Name? (Mr. N).
2. What Will I Wear? (Ms. W).
3. My Body (Ms. P).
4. I'm Healthy (Mr. H).
5. All About My Senses (Mr. M).
6. We All Have Feelings (Ms. A).

Each unit, in turn, contains six lessons. Each lesson plan provides a daily menu of activities—beginning with a brief whole-class Meeting Circle activity, followed by a selection of individual and small-group Interest Center activities.

As an example, the first unit in the All About Me theme is called 'What's My Name?' Mr. N introduces this unit, which focuses on the letter N/n, and the /n/ sound, and introduces relevant vocabulary words. In the Meeting Circle Time, the storybook to be read to the class is *'From Ann to Zach'*, and the read-along book is Mr. N's *'What's My Name?'* An example of an activity in an Interest Center is making placemats individualized with each child's name in the art center.

Each Even Start classroom assigned to Let's Begin with the Letter People was provided with the following set of materials:

- Letter People and their accessories,
- Letter People finger puppets,
- Big and Little books and their story tapes,
- Musical materials,
- Family materials,
- Manipulative letters, numbers, and shapes,
- Me Bags,¹⁶ and
- Teacher resource materials.

¹⁶ Me Bags are bags that children take home and fill with meaningful objects, which they then explain to their classmates. Each Letter Person also has a Me Bag that is shared with the class to reinforce concepts being taught.

For additional information about Let's Begin with the Letter People, see http://www.abramsandcompany.com/lets_begin_with_letter_people.aspx.

CIRCLE Training. In conjunction with the Let's Begin with the Letter People curriculum, CIRCLE provided professional development to the early childhood education staff on providing developmentally appropriate literacy instruction. This training covered the following topics:

- rich language input,
- responsiveness to children's signals,
- maintaining and building on interests,
- choice-providing strategies,
- monitoring of children's behavior,
- language development,
- print and book awareness,
- motivation to read/read aloud,
- phonological awareness,
- letter knowledge and early word recognition, and
- written expression.

Play and Learning Strategies (PALS)-Parenting Curriculum. For the CLIO study, the Let's Begin with the Letter People curriculum was combined with the PALS parenting education program developed by the University of Texas-Houston Health Sciences Center. PALS also included a certification requirement for parents.

PALS is designed to teach parents responsive parenting strategies that can help their children build the cognitive and language skills they will need for success in school. PALS has previously been evaluated with parents whose children are at risk for developmental delay and academic failure due to poverty, low family literacy, and other risk factors. Parents who participate in the PALS curriculum learn two categories of behaviors:

1. behaviors that provide a supportive, dynamic interpersonal environment for learning, such as attending to children's communicative signals, responding promptly and warmly, and following children's interests to encourage further interactions with learning materials and people; and

2. behaviors that directly help children build their cognitive, language, and pre-literacy skills, such as teaching new words, using rich language to describe and connect objects and actions, building on children's utterances to model longer sentences, and reading books with children in ways that encourage their active participation.

To help parents learn these behaviors PALS instructors focus on the following four key concepts:

1. *The zone of proximal development*: this is the learning level at which adult support is critical for helping children accomplish goals beyond what they can do by themselves.
2. *Scaffolding*: when an adult builds upon what the child can do and helps him/her take the next step by expanding on what the child has said or done. As the child experiences this support consistently over time, greater cognitive and social skill development is expected to occur.
3. *Warm, responsive style*: children are not likely to sustain their participation in learning activities if parents ignore them or respond harshly to them. In contrast, when parents respond with warmth and encouragement to their children's efforts, children are more likely to stay engaged and attentive, thereby allowing them to learn much more from each interaction.
4. *Spiral learning*: children need repeated exposure to new concepts and behaviors over time, with multiple opportunities to practice and integrate their new knowledge into their existing skill set. Mastery happens gradually as children build a network of connections between new and old skills.

Parents learn these concepts so that they can be better teachers of their own children. In addition, parent educators are taught to rely on these concepts while working with Even Start parents. That is, during a parent education session, the PALS instructor will scaffold parent knowledge, use a warm and responsive instruction style, and will spiral back to previously learned topics.

Each PALS center-based session uses the following format: a review of the previous week's topic and discussion of the parents' use of that skill, introduction of a

new skill through discussion and an instructional videotape, practice of the new skill with their child while the parent educator videotapes and provides coaching, and discussion as the group and parent educator review the videotape together.

In addition to these center-based parenting education sessions, PALS also provides activities for home visits. Home visitors review and reinforce concepts that parents may be having difficulty implementing. Further, the home visitor provides an important link between the preschool classroom themes from Let's Begin with the Letter People and the PALS concepts, which are reinforced through the home activities. Materials provided to the parenting educator by the PALS program included the following.

- facilitator manuals in English and Spanish,
- video manuals in English and Spanish,
- home visitor manuals in English and Spanish,
- videotapes in English and Spanish,
- coaching toy bag,
- home visitor toy bag, and
- parent workbooks.

Certification. As part of the CLIO study, each Even Start parent educator who was implementing PALS was required to become certified as a PALS facilitator. The certification process included the requirement that the participant submit up to three videotapes of themselves leading PALS instructional sessions. The first of these tapes included the facilitator conducting a complete mock session (introduction, video, and summary, transition, filming, coaching, and reviewing). The second tape was also a complete session but this time with parents. A third tape was requested if the facilitator needed further follow-up suggestions. The tapes were reviewed and appropriate feedback provided to each facilitator-in-training before award certification was granted.

For further information about PALS, see
<http://www.childrenslearninginstitute.org/our-programs/program-overview/PALS/default.html>.

Partners for Literacy

Although some aspects of the Partners for Literacy curriculum existed prior to CLIO,¹⁷ the current version that includes both a preschool and a parenting curriculum was developed specifically for the CLIO study. The developers further refined some of the materials and strategies over the course of the 2 implementation years. Partners for Literacy was developed by Barbara Wasik and Joe Sparling from the Frank Porter Graham Child Development Center at the University of North Carolina, Chapel Hill.

Partners for Literacy—Preschool Curriculum. (Wasik and Sparling, University of North Carolina, Frank Porter Graham Child Development Center). This early childhood education curriculum focuses on language and literacy activities for children, coupled with specific instructional strategies for teachers. It encompasses four key early literacy domains: oral language and vocabulary, letter knowledge, phonological awareness, and concepts of print, as well as social and emotional development. Educators in a Partners preschool classroom are to create a literacy-rich classroom, promote language and literacy skills throughout the day, and provide both group and individualized instruction. Instruction is individualized by interacting with successive pairs of children throughout the day on both game-like activities and book reading. The activities and instructional strategies are described below.

LiteracyGames is one essential part of the curriculum. LiteracyGames is a series of 50 instructional activities (games) for 3-year-olds and 50 games for 4-year-olds, designed to be played by one adult and two children. The games address such early literacy domains as concepts of print, letter knowledge, oral language, phonological awareness, writing, and creativity. Each has a specific instructional goal and takes about 5 minutes. The games are engaging for children and involve considerable back-and-forth between adult and child. Children progress from easy to more difficult activities within each game.

Teachers focus on one or two games per week, depending on the length of the school day. Each game has three “cycles” allowing teachers to tailor the activities to

¹⁷ For example, some components of Partners for Literacy were used in the now-famous Abecedarian study out of the Frank Porter Graham Center.

the abilities of individual children. Teachers are expected to set aside a specific time of the day in which to play LiteracyGames. For example, during the LiteracyGames time the teacher would play a game with successive pairs of children while the rest of the class engaged in free play or center time supervised by the teacher assistant. Child progress on LiteracyGames is monitored using a Record of Mastery form.

Interactive Book Reading is another essential part of the Partners for Literacy preschool curriculum. Teachers are expected to read for 5 to 10 minutes with each child, every day. The teacher may choose to read with children individually or in pairs. The curriculum provides a set of children's books to complement those in the classroom and encourages teachers to create a lending library of books that children can take home. Two other types of books were designed to facilitate Interactive Book Reading: Conversation Books and Little Conversation Books.

Conversation Books are very simple books designed to stimulate conversation between an adult and child. They also serve to familiarize teachers with a specific interactive book reading strategy known as the 3S Strategy: See, Show, and Say (discussed in more detail below). Little Conversation Books are a collection of small books designed to build early literacy skills, specifically oral language, phonological awareness, print awareness, and alphabet knowledge. The sets include: ABC stories, Nursery Rhymes and Poems, Words in Words, Matching Stories, Sequencing Stories, and Problem Solving Stories.

Scaffolding new competencies is the foundation for all Partners for Literacy instructional strategies. That is, building on existing knowledge and skills and adding support to enable a child to move to a higher level of competence. Scaffolding can take many forms, including modeling, leading questions, prompting, and instructional conversations.

The 3N Strategy is a means of structuring interactions with children. It consists of three parts: *Notice* what a child is doing or is ready to do; *Nudge* or prompt, encourage, assist the child to take a step further; and *Narrate* or give praise, feedback, and acknowledgment. This strategy is used throughout the Partners for Literacy classroom day.

Interactive Book Reading Strategies. These strategies are all designed to facilitate an instructional conversation between the child and teacher.

- *3S Strategy: See, Show, Say.* See, Show, and Say refer to the different levels of response required of a child during any Interactive Book Reading session. A child is first asked to “see” or look at a specific feature of the book, such as a picture. If the child follows this direction, the teacher can then ask the child to “show” an object or word on the page. If the child can follow this direction, the teacher then asks the child to “say” a word or answer a question. This strategy helps develop active engagement on the part of the child.
- *Wh Questions: Who, What, When, Where, Why.* The teacher uses Wh questions when a child has even a minimum level of oral language. Wh questions are used in a progressively more complex manner by the teacher, always individualizing for the child. These questions facilitate an ongoing instructional conversation between teacher and child and help the child develop comprehension skills.
- *Expanded Book Reading* is used to increase motivation to read and promote comprehension. It includes a set of activities to facilitate literacy skills and comprehension such as discussing a book before or after it has been read or incorporating the book into other classroom activities such as art or music.

Enriched Caregiving. Enriched Caregiving, or teaching throughout the day, is intended to promote language, literacy, cognitive, and social development. Teachers can use Enriched Caregiving strategies to turn everyday routines, activities, and transitions into important learning opportunities. For example, a teacher could use lunchtime conversation to develop oral language skills or to reinforce a previous lesson that dealt with a particular letter sound.

Problem-Solving Strategies. Problem-solving strategies are used in many ways throughout the Partners for Literacy curriculum. These strategies are used to help children develop social skills and learn how to manage their own emotions. Teachers help children learn to recognize feelings, identify wants and needs, and develop simple problem-solving skills at age appropriate levels.

A typical day in a Partners for Literacy classroom would incorporate the types of activities common in many preschool classrooms such as circle time, free play, center time, and structured small group activities in addition to gross motor play, meals, etc. The curriculum calls for setting aside time for Partners for Literacy-specific activities such as LiteracyGames and interactive book reading—an hour per day for half-day programs and 2 hours per day for full-day programs.

Partners for Literacy—Parenting Curriculum. Partners for Literacy is a fully integrated curriculum, so that the parent component parallels the themes and activities in the preschool classroom. Parents are taught to use many of the same materials and instructional strategies with their children as are used in preschool classrooms.

LearningGames. LearningGames are short game-like activities designed for parents and children ages 3 to 5 that are similar to the LiteracyGames used in the preschool classrooms. LearningGames cover the same early literacy domains and are played in the same back and forth fashion as LiteracyGames. They are designed to complement the activities in preschool classrooms and give parents the opportunity to reinforce these activities at home.

Interactive Book Reading. The parenting Interactive Book Reading component of Partners for Literacy helps parents develop strategies to have a conversation with their children when reading a book together. These strategies promote early literacy skills. A variety of children’s storybooks are used by parents, as well as the Conversation Books and Little Conversation Books described earlier.

Instructional Strategies. Parents are taught the same instructional strategies as those used in preschool classrooms, including the *3N Strategy*, *3S Strategy*, *Wh Questions*, and *Expanded Book Reading*. Parents are also taught to use Enriched Caregiving strategies with their children. The curriculum also helps parents learn a problem-solving strategy that includes setting goals and defining possible solutions and their consequences. This strategy helps parents address everyday parenting concerns.

In a parenting session, parent educators use a specific protocol as the structure for each lesson: Read, Role Play, and Reflect. As each lesson or strategy is introduced, parents read and discuss the information together. They then role play and practice a specific strategy with a partner. Finally, parents work as a group to review

and reflect on their role play, as they prepare to work with their children during parent-child interactive sessions.

Parent-child interactive sessions provide parents with the opportunity to play LearningGames and read a book with their children under the guidance of instructional staff. Each interactive session is reviewed in the parenting session prior to the introduction of any new activity or concept.

Complete information about the Partners for Literacy preschool and parenting curricula can be found at www.fpg.unc.edu/~literacy.

CLIO Curricula Professional Development

As discussed in chapter 2, implementation of the curricula in the field was supported by the curriculum developers in three ways: annual group training, on-site mentoring, and ongoing support. For the CLIO study, this professional development was an integral part of the curricular interventions.

Training. Each developer conducted 4-day training sessions for project staff in the four study groups in the summer of 2004, before the first implementation year. The developers provided make-up trainings in the fall of 2004 for staff who were not able to attend the summer session. Developers again conducted training sessions during the summer of 2005. If project staff were not able to attend the summer or make-up sessions, developer staff often provided on-site training during mentoring visits. The main challenges were teacher turnover and the widely dispersed treatment projects. However, at least one staff member from all projects received training, and most teachers (more than 80 percent of early childhood education and parenting education staff) reported receiving training on the curricula.

The 4-day summer sessions included parallel tracks for preschool and parenting staff. The training consisted of extensive hands-on practice and small-group discussion in addition to large-group overviews and introductions of specific topics and teaching strategies. The Let's Begin training included sessions on the five themes and associated lesson plans, use of the huggables and other Letter People materials, and best practices related to developmentally appropriate literacy instruction. The Partners for

Literacy training included sessions on LiteracyGames, interactive book reading, enriched caregiving, and English language learners. In addition both developers had sessions devoted to classroom organization and management. Project staff received detailed manuals describing all aspects of each curriculum.

Training for parenting staff also included ample opportunity for hands-on practice and role-playing. PALS training focused on the concepts and behaviors to be taught to parents, reviewed the video lessons, and allowed staff to practice coaching parents. After training, teachers produced a videotape for developer staff to become certified. Partners for Literacy included sessions on teaching parents about LearningGames and book-reading strategies, enriching caregiving strategies at home, and coaching parents during parent-child activities. Parenting staff also received detailed curriculum manuals.

On-site mentoring and ongoing support. On-site mentoring consisted of visits to each treatment project from developer staff members. During these 2- or 3-day visits, developers observed project staff using the curriculum, provided feedback and coaching, and supported curriculum planning. On average, each treatment project received two mentoring visits per year.

Curriculum developers also provided CLIO projects with support throughout the implementation years. Let's Begin and PALS held regional monthly conference calls to allow project staff to discuss implementation issues with developer staff and among themselves. Developer staff were also available by phone and email on an as-needed basis. Partners for Literacy assigned a consultant to each of their projects. These consultants conducted the mentoring visits and stayed in regular contact with assigned projects by email and phone.

Instructional Practices and Fidelity in CLIO Classrooms

Two kinds of data on classrooms were collected to describe the types and quality of the instructional practices. First, observations were conducted of all classrooms in the sample, both treatment and control, to document the extent to which the children were exposed to instructional practices shown in previous research to be related to children's emergent literacy skills. Second, in only the treatment classrooms,

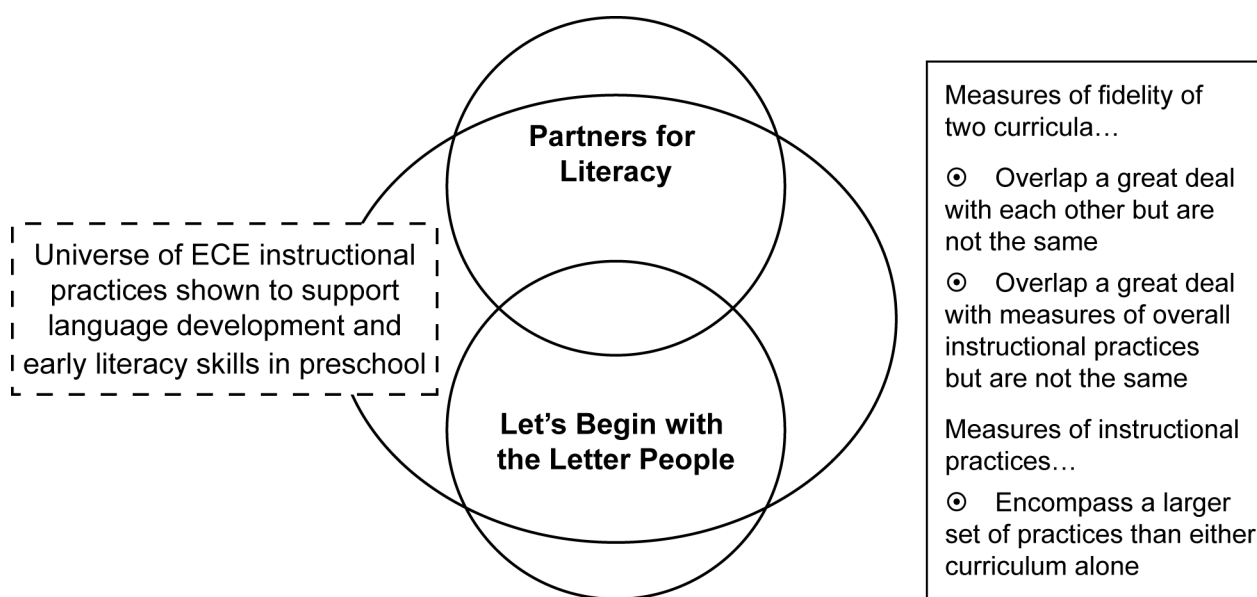
observations were conducted to evaluate the fidelity of implementation of the CLIO curricula. Since the CLIO curricula were chosen because they represented research-based instructional practices for supporting emergent literacy, it is highly likely that the CLIO curricula overlap substantially (1) with the universe of potentially effective early childhood education practices measured in the CLIO observations and (2) with each other.

Figure 3-1 represents this pictorially. In the figure, the smaller circles represent the instructional practices used by the two CLIO preschool curricula (as intended), while the larger circle represents all of the instructional practices that research has shown are linked to child literacy outcomes. The overlap between each of the smaller circles and the larger circle represents the extent to which each of the CLIO curricula embody the “best” practices as defined by the field. Further, the two smaller circles themselves overlap to a large degree, since the two CLIO preschool curricula use some common instructional practices across their multiple activities.

Fidelity of Implementation

Program fidelity is the degree of fit between the developer-defined components of the curriculum and its actual implementation in the classroom. In other words, how closely does the classroom instruction match the specifications of the curriculum that were set out in the curriculum materials and training? Research studies often measure fidelity because of the added value that an understanding of fidelity can bring to interpreting impacts of a curriculum. If a curriculum is poorly implemented, then any negative findings about that curriculum might indicate directions for curriculum modification or modification to teacher training procedures.

Figure 3-1. Relationship of Measuring Fidelity of Implementation and Overall Instructional Practices



The study measured the fidelity of implementation of the CLIO curricula in two ways: evaluation staff conducted independent observations of preschool and parenting classrooms, and the curriculum developers rated the level of implementation of both the preschool curriculum and parenting curriculum in each project. The rating system was the same for both sources of information. A 5-point Likert scale was used to rate each project's implementation of the curriculum as one of the following:

1. not appreciably implemented,
2. partially implemented,
3. half-way implemented,
4. almost fully implemented, or
5. fully implemented.

Ratings by Independent Observers. The first set of fidelity ratings was at the classroom level and was based on the classroom observations conducted by trained independent observers. The observations were conducted in the spring of each implementation year, which yielded two fidelity ratings by independent observers over

the life of the study.¹⁸ These corresponded to approximately 9 months and 20 months after training.

During the classroom observations, the observers completed fidelity checklists that were developed by the evaluation team with input from the developers themselves. The checklists were designed to closely match the specific recommendations made to teachers during summer trainings about how to put the curriculum into place in their classrooms. Separate checklists were designed for preschool and parenting classrooms, and observers rated them separately. (See appendix G for more detail on the development of the CLIO fidelity checklists and scoring rubrics.)

Ratings by Curriculum Developers. Each curriculum developer also rated the fidelity of implementation of the curriculum in each of their projects. Developers assigned their ratings at the project level. Fidelity was rated separately for the preschool curriculum and the parenting curriculum, and separately for each of the 2 implementation years of the CLIO study. Further, in each implementation year, the developers provided fidelity ratings twice, in the winter and the spring. Thus, the developers provided fidelity ratings of their projects four times over the life of the study, which described fidelity approximately 6 months, 11 months, 18 months, and 23 months after the initial training. According to developers, their fidelity ratings included information gained through records received from implementing projects, one-on-one emails and phone calls between developer and project staff, and coaching visits to projects by developer staff. (See appendix G for further information on fidelity measurement.)

Summary of Ratings. On average, curriculum implementation was rated as having reached about 50 percent of what represented full implementation of the models (or “3” on the 5-point rating scale). The average fidelity ratings were consistent across the two sets of raters (see table 3-1 for average ratings by the observers and table 3-2 for average ratings by the developers).

¹⁸ Four projects refused to implement their randomly assigned curriculum in one year or the other. Despite this refusal, all four projects did allow continued data collection, including classroom observations and observer fidelity measurement.

Table 3-1. Average Classroom-Level Fidelity Ratings^a by Study Group and Classroom Type for Spring 2005 and Spring 2006: Independent Observer Ratings

		Study group			
		Let's Begin with the Letter People (ECE)	Let's Begin and Play & Learning Strategies (ECE/PE)	Partners for Literacy (ECE)	Partners for Literacy (ECE/PE)
Early childhood education (ECE)					
2005	2.84	3.38	3.39	2.20	2.49
2006	2.85	3.61	3.55	2.00	2.20
Parenting education (PE)					
2005	3.30	†	3.28	†	3.32
2006	3.46	†	3.52	†	3.40

^a The fidelity ratings range from 1 (not appreciably implemented) to 5 (fully implemented).

† Not applicable.

¹ For the ECE classroom rating, "all study groups" is the average rating across all four study groups. For the PE classroom rating, "all study groups" is the average rating for the two study groups with PE components.

NOTES: The data are based on classrooms with at least one child who took a spring CLIO assessment in the relevant year. Observer fidelity ratings were conducted at the classroom level as part of the larger set of classroom observations of instructional practices.

SOURCE: U.S. Department of Education, Even Start Classroom Literacy Interventions and Outcomes Study, "Observer Fidelity Ratings," Spring 2005 and Spring 2006.

Table 3-2. Average Project-Level Fidelity Ratings^a by Study Group and Classroom Type for Spring 2005 and Spring 2006: Developer Ratings

	All study groups ¹	Study group			
		Let's Begin with the Letter People (ECE)	Let's Begin and Play & Learning Strategies (ECE/PE)	Partners for Literacy (ECE)	Partners for Literacy (ECE/PE)
Early childhood education (ECE)					
2005	2.99	3.30	3.72	2.46	2.45
2006	3.29	4.10	3.33	2.59	3.05
Parenting education (PE)					
2005	2.96	†	3.29	†	2.59
2006	3.26	†	3.41	†	3.10

^a The fidelity ratings range from 1 (not appreciably implemented) to 5 (fully implemented).

† Not applicable.

¹ For the ECE project rating, "all study groups" is the average rating across all four study groups. For the PE project rating, "all study groups" is the average rating for the two study groups with PE components.

NOTES: Developer fidelity ratings were recorded at the project level. In projects with more than one ECE or PE class, the fidelity scores for classrooms were averaged.

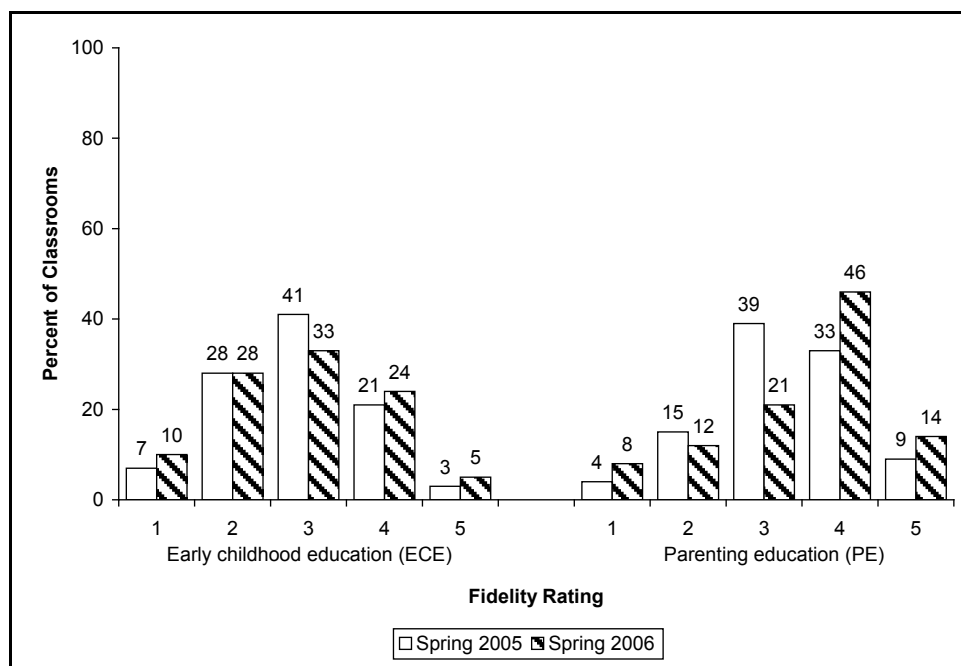
SOURCE: U.S. Department of Education, Even Start Classroom Literacy Interventions and Outcomes Study, "Developer Fidelity Ratings," Spring 2005 and Spring 2006.

A visual inspection of the fidelity data suggests that fidelity ratings varied by curriculum and by implementation year, as follows (tables 3-1 and 3-2).

- For both observer ratings and developer ratings, the fidelity of implementation for the projects implementing Let's Begin with the Letter People and PALS were generally higher than those for the Partners for Literacy projects.
- For both observer ratings and developer ratings, the average fidelity ratings were higher in the second year of implementation of the CLIO curricula. This was the expected pattern, since fidelity after nearly 2 years of experience implementing the curricula (2006) should be higher than fidelity after only a single year of implementation (2005). The one exception was the observer ratings of the fidelity of implementation in the Partners for Literacy preschool classrooms, which were lower in 2006 compared with 2005.

In 2005, across all study groups, observers rated 24 percent of preschool classrooms as achieving “4” or higher on the 5-point fidelity scale and 42 percent of parenting education classrooms as achieving “4” or higher on the fidelity scale (figure 3-2). In 2006, observers rated 29 percent of preschool classrooms and 60 percent of parenting education classrooms as achieving at least “4” on the 5-point fidelity scale.

Figure 3-2. Percentage Distribution of Fidelity Ratings of CLIO Treatment Classrooms by Classroom Type for Spring 2005 and Spring 2006: Observer Ratings

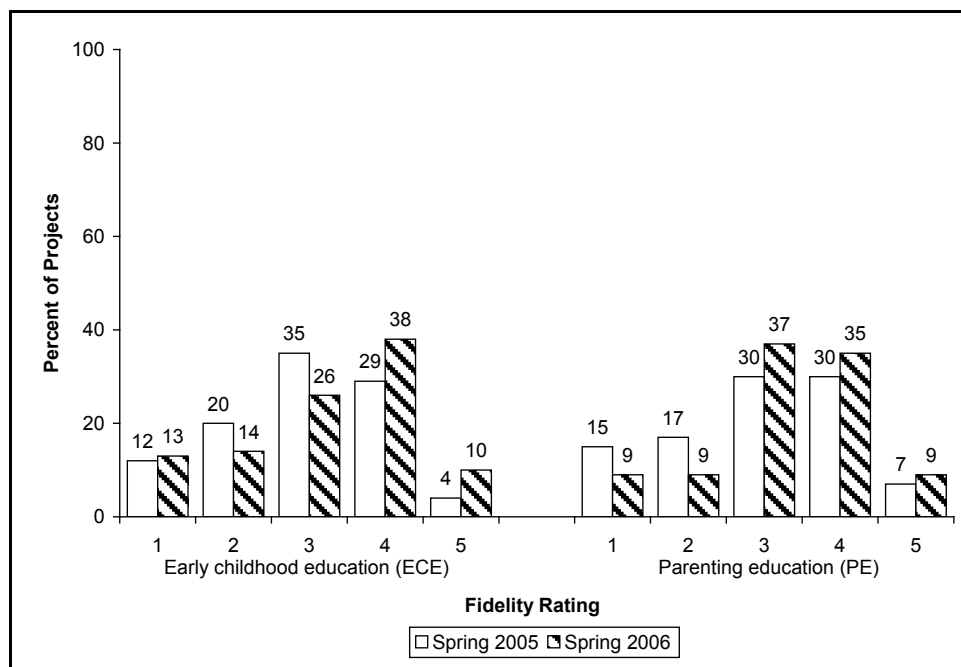


NOTES: The fidelity ratings range from 1 (not appreciably implemented) to 5 (fully implemented). The ECE rating includes classrooms in all four study groups, since all four had CLIO ECE curricula. The PE rating includes classrooms in the two study groups with CLIO PE curricula. The data are based on classrooms with at least one child who took a CLIO assessment in the relevant year. Detail may not sum to total due to rounding.

SOURCE: U.S. Department of Education, Even Start Classroom Literacy Interventions and Outcomes Study, “Observer Fidelity Ratings,” Spring 2005 and Spring 2006.

Across all study groups in 2005, developers rated 33 percent of projects as achieving at least “4” on the 5-point fidelity scale for the preschool curricula and rated 37 percent of projects as achieving at least “4” on the fidelity scale for parenting curricula (figure 3-3). In 2006, developers rated 38 percent of projects as “4” or above on the fidelity scale for preschool curricula and 44 percent of projects as “4” or above on the fidelity scale for parenting curricula.

Figure 3-3. Percentage Distribution of Fidelity Ratings of CLIO Treatment Projects by Classroom Type for Spring 2005 and Spring 2006: Developer Ratings



NOTES: The fidelity ratings range from 1 (not appreciably implemented) to 5 (fully implemented). Any rating that fell between two intervals was rounded down for the graph. The ECE rating includes projects in all four study groups, since all four had CLIO ECE curricula. The PE rating includes projects in the two study groups with CLIO PE curricula. Detail may not sum to total due to rounding.

SOURCE: U.S. Department of Education, Even Start Classroom Literacy Interventions and Outcomes Study, "Developer Fidelity Ratings," Spring 2005 and Spring 2006.

Exposure to the Curricula

Another important component of the implementation of any curriculum is exposure. In order to benefit from the curricula in significant ways, children and parents would need sufficient exposure to the curricula. Even Start guidelines do not specify an expected level of exposure for children or parents, and the hours of instruction offered by local projects vary widely. Curriculum developers were aware of this at the outset of the study, and the curricula were designed to accommodate this variation. To measure exposure to services in general and to the CLIO curricula, we collected monthly participation data from all CLIO projects (both treatment and control) on the number of hours of instruction for each preschooler and the number of hours of parenting instruction for each parent.

Preschool Participation. Projects reported that they *offered* an average of 80 hours of preschool education per month in both years of the study.¹⁹ The participation data show that in the first 7 months of the school year,²⁰ children received an average of 50 hours of instruction per month.²¹ Almost three-quarters of children attended preschool 60 hours or fewer per month, where 60 hours per month is the equivalent of a half-day program meeting 5 days a week (see figure 3-4). In approximately half of CLIO projects, no children participated in preschool education for 60 hours or more per month. This is consistent with participation data from the Third National Even Start Evaluation (St.Pierre et al. 2003).

Participation in Parenting Education. During the 2 implementation years, projects reported that they offered parents a monthly average of 14 hours of PE and 11 hours of PC.²² As was true for children and the preschool curricula, parents also received limited exposure to the parenting curricula. Over the 7-month period preceding spring data collection, parents received an average of 13 hours combined parenting instruction per month.²³ The majority of parents (over 80 percent) had 20 or fewer hours of combined parenting instruction per month (see figure 3-5). This is consistent with participation data from the Third National Even Start Evaluation (St.Pierre et al. 2003).

¹⁹ Preschool monthly hours offered ranged from 24 to 160 hours, with a standard deviation of 31.1.

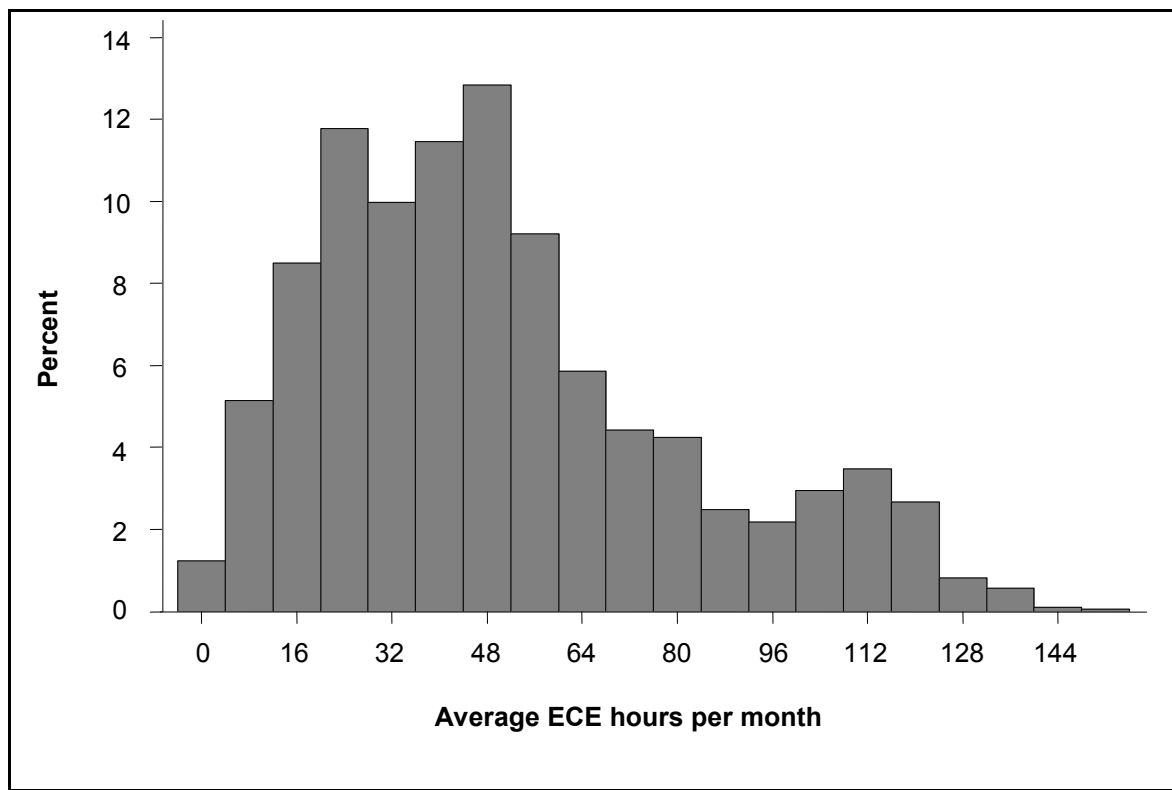
²⁰ We use 9-month participation as an outcome, but for purposes of understanding the intervention, the 7-month figures are more useful. Clearly participation in April and May cannot affect performance in March assessments.

²¹ Monthly preschool participation hours ranged from 0 to 152 hours, with a standard deviation of 31.5.

²² PE monthly hours offered ranged from 4 to 40 hours, with a standard deviation of 6.6. PC monthly hours offered ranged from 2 to 28 hours, with a standard deviation of 5.4.

²³ Monthly PE/PC participation hours ranged from 0 to 81 hours, with a standard deviation of 9.4. The high figure includes three parents for whom Even Start projects reported more than 68 hours per month, on average, of parent participation. Excluding these outliers, the monthly PE/PC participation hours ranged from 0 to 64, with a standard deviation of 9.2 hours.

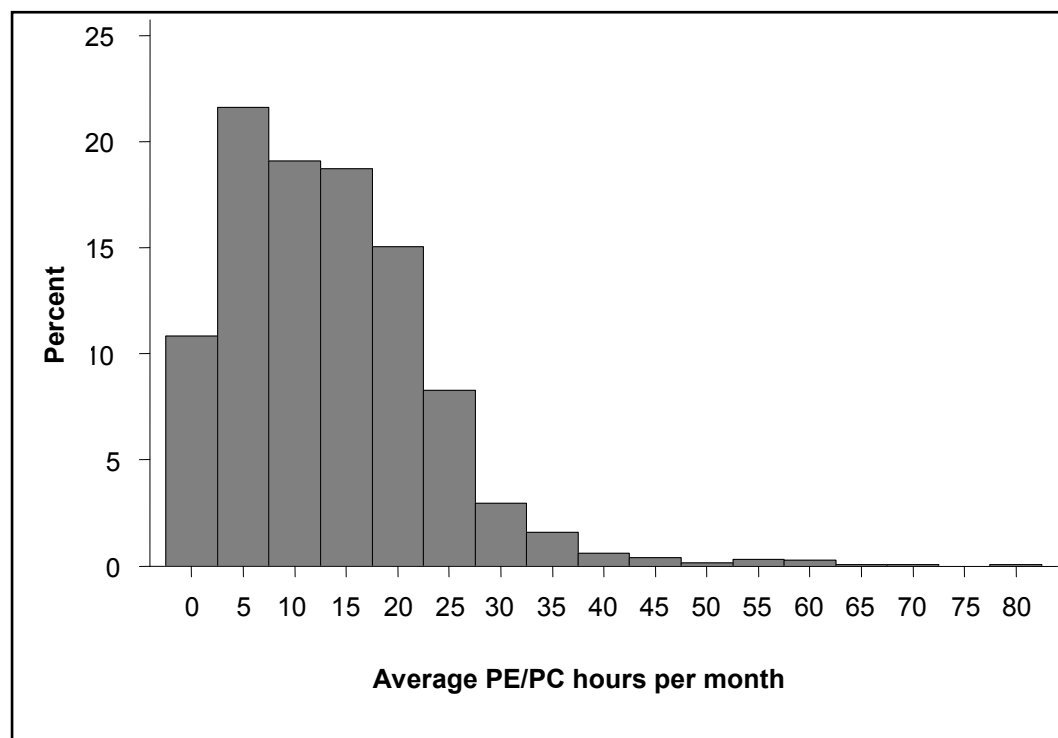
Figure 3-4. Distribution of CLIO Children by Average Monthly Hours of ECE Participation, Spring 2005 and Spring 2006



NOTE: Average monthly hours is calculated over the 7-month period of September through March. The data were run on children in the combined 2005/2006 analysis sample.

SOURCE: U.S. Department of Education, Even Start Classroom Literacy Interventions and Outcomes Study, "Instructional Services Participation Form (ISPF)," 2004-05 and 2005-06.

Figure 3-5. Distribution of CLIO Parents by Average Monthly Hours of PE/PC Participation, Spring 2005 and Spring 2006



NOTES: Average monthly hours is calculated over the 7-month period of September through March. The data are parent-level runs of children in the combined 2005/2006 analysis sample.

SOURCE: U.S. Department of Education, Even Start Classroom Literacy Interventions and Outcomes Study, "Instructional Services Participation Form (ISPF)," 2004-05 and 2005-06.

Control Projects During CLIO Study

While the CLIO curricula were being implemented in the four treatment groups, projects in the control group continued with their usual services in the four Even Start instructional areas. Understanding the control classrooms can help us interpret the impacts on parents and children.

The study has two sources of information about the instruction in the control classes. First, project directors in all of the CLIO projects, including control projects, were asked about any formal curricula used in their preschool or parent education programs. Second, the OMLIT observations provide detailed information on the instructional processes in both control and treatment projects. We consider the OMLIT

data to provide a richer and more valid estimate of the differences between the treatment and control projects than reports from project directors.

Based on project director reports of curricula used, we estimated that most control projects used a formal early childhood curriculum and about half used a formal parenting curriculum. Specifically, in both implementation years a majority of control projects reported using at least one formal preschool curriculum—71 percent in 2005 and 76 percent in 2006. Among the control projects that reported using a curriculum, the two most commonly used were High Scope and Creative Curriculum. About half of the control projects reported using one of these two curricula in each of the implementation years. Other curricula cited included a range of instructional programming, including published comprehensive curricula, literacy-focused supplemental curricula, informal non-published curricula, and local or other reading initiatives.

In both implementation years, over half of control projects reported using at least one formal PE curriculum—58 percent in 2005 and 64 percent in 2006. Parents as Teachers was the only formal curriculum named by more than a few projects. A number of other curricula were named by individual projects (e.g., MotherRead, the Bowdoin method), but most projects used materials or activities developed by the project either on its own or based on materials available to the field.

The OMLIT observations data from spring 2005 and spring 2006 were used to create a set of variables that provide a descriptive profile of both treatment and control classrooms (table 3-3). As a group, the control classrooms spent about 15 percent of the day in literacy-related activities, including reading, alphabet/print knowledge, phonological awareness, and emergent writing. For treatment classrooms, time spent in literacy-related activities ranged from 16 to 19 percent of the day. Including these literacy activities, control classrooms spent about 45 percent of the day in activities that are often considered by developmental psychologists to have particularly high value for children because of the opportunities for children to construct knowledge and receive feedback on their interactions with materials, peers, and adults in the classroom (Bruner and Watson 1983). Treatment classrooms spent from about 42 to 51 percent of the day in these high-value activities. The remainder of the day for all classrooms was spent in daily group activities, including review of the calendar, weather, and attendance; gross motor play and transition; and meals/snacks.

Table 3-3. ECE Instruction in CLIO Classrooms (average of spring 2005 and spring 2006)

Instructional variables	Study group					All study groups
	Let's Begin with the Letter People (ECE)	Let's Begin and Play & Learning Strategies (ECE/PE)	Partners for Literacy (ECE)	Partners for Literacy (ECE/PE)	Control	
Average minutes reading aloud per half-day ^a	13.5	15.2	19.4	13.8	11.7	14.6
% of "literacy rich" classrooms ^{b,c}	24.0	31.8	14.1	20.1	13.6	20.2
% time children are in high-level activities ^c	42.6	51.2	51.2	46.6	43.7	46.9
% time children are in literacy-related activities ^d	16.0	19.6	19.0	17.2	15.3	17.3
% time class is in whole group activity (excluding routines) ^d	62.0	61.1	59.4	59.1	60.4	60.4
% classrooms with high-quality support for new vocabulary ^a	4.0	4.6	6.4	4.2	2.3	4.2
% classrooms with high-quality support for oral language ^e	6.7	9.1	11.5	9.7	9.1	9.2
% classrooms with good or better quality of instructional support for children's letter-word knowledge ^f	32.0	48.5	28.2	29.2	13.6	29.3
% classrooms with good or better quality of instructional support for children's understanding of sound ^f	28.0	27.3	19.2	15.3	13.6	20.3
% classrooms with good or better quality of instructional support for children's writing ^f	21.3	34.8	10.3	12.5	15.9	18.5
% classrooms with good or better quality of instructional support for children's oral language ^f	10.7	16.7	16.7	12.5	13.6	14.0

^aFrom the Read Aloud Profile; *high-quality support for vocabulary* defined as at least two new vocabulary words discussed, with at least two comprehension supports (picture, gesture, semantic network)

^{b,c}From the Classroom Literacy Opportunities Checklist; *literacy-rich* defined as rating of "rich resources" on average score across seven areas of literacy resources available to children in classroom.

^dFrom the Classroom Snapshot; *high-level activities* include all learning and creative activities (excludes routines, gross motor play, unstructured socialization/horsing around).

^eFrom the Read Aloud Profile; *high-quality support for oral language* defined as at least two open-ended questions with time for children to respond (questions requiring prediction, theorizing, etc.).

^fFrom the Quality of Language/Literacy Instruction; *good or better quality* defined as rating of greater than 3 (on a 1-5 scale) where quality is defined as high-value activities, authentic and integrated activities, variety of activities, involvement of all children, opportunities for children to work with teacher in small groups, etc.

SOURCE: U.S. Department of Education, Even Start Classroom Literacy Interventions and Outcomes Study.

Data from OMLIT (*Observation Measures of Literacy Instruction*, Goodson et al. 2004).

The majority of the day, children in treatment and control classrooms were organized in one large group, all doing the same activity. On average, children were read-aloud to about 11 minutes a day in control classrooms and from about 13 to 19 minutes a day in treatment classrooms. In control classrooms, about 2 percent of the read aloud sessions included quality support for new vocabulary and about 10 percent included open-ended questions that support children in learning to use more abstract and complex thinking processes. In treatment classrooms, from 4 to 6 percent of read-aloud sessions included quality support for new vocabulary. Across the four overall ratings of the quality of the literacy and language instructional support shown in table 3-3, the percentage of control classrooms rated as good or better ranged from 10 to 16 percent. Instruction was generally conducted in English, and in the majority of preschool classrooms with English language learners the staff spoke only English. Thirty-five percent of the classrooms with English language learners had staff who spoke the same language as the ELL children.

4. ANALYSIS METHODS

Our primary analysis followed the classical intent-to-treat (ITT) analysis associated with randomized clinical trials (RCTs). Instead of using simple randomization-based tests, we used modeling procedures that condition on baseline covariates to improve power. Power is an important consideration in the CLIO design, since randomization was conducted at the project level, and each study group included only 24 projects.²⁴ Given that Even Start projects (rather than children or families) were the units of randomization, we reflected the cluster structure in all estimates of the precision of effect estimates. Also, given that some of the covariates are project-level covariates, and given the need to reflect the clustering in precision estimates, we used multi-level modeling procedures for the analysis.

To reduce the number of false positive findings without unduly compromising power, we identified a limited set of tests to be run for each outcome scale. We used the Bonferroni adjustment for multiple testing among the set of tests run for each of the 21 outcomes (discussed in chapter 2). (We did not, however, apply such an adjustment across outcomes.) In this report, test thresholds were set using a Bonferroni adjustment. Tests that were significant by this criterion are flagged with asterisks. Confidence intervals and *p*-values are also presented in all ITT-related tables. Neither of these are adjusted for the multiple comparisons in any way.

In this chapter, we discuss the use of covariates, describe the multi-level modeling, present the contrasts we selected and the corrections we made for multiple comparisons, and describe our procedures for handling missing data. In chapter 5, we present the findings from our planned ITT analyses. We then undertook secondary ITT analyses—which we describe in chapter 6—and additional more exploratory analyses—which we describe in chapter 7.

²⁴ Although these procedures are generally less robust than simple randomization-based tests, we conducted simulations that convinced us that the increase in power was worth the loss of robustness. Specifically, we conducted a simulation study of HLM and several alternatives in settings similar to what we expect in CLIO and found that the HLM package developed by Bryk and Raudenbush (2004) exploits the covariates to improve power while maintaining control over type I error rates, even when there are violations of some of the standard assumptions that underlie the statistical methods used by the package. We therefore chose this package for our analysis.

Use of Covariates

Valid results can be achieved in randomized designs without conditioning on any covariates. However, it is possible to increase power by conditioning on covariates, particularly when the number of randomized study units is small. In analyzing CLIO data, we conditioned on a limited set of covariates.²⁵

- **Project Level**

- Average pre-intervention score on the outcome of interest if available²⁶ (e.g., the project-level pre-intervention score on the IGDI was used as a covariate when analyzing IGDI as an outcome as shown in the first row of table H-1);
- Average pre-intervention score on other outcomes of interest²⁷ (e.g., the average of the project-level pre-intervention scores on the PPVT, the blending and elision components of the preschool CTOPPP, and social competence was used as a covariate when analyzing IGDI as an outcome as shown in the second row of table H-1);
- Pre-intervention child to teacher ratio;²⁸ and
- Year (flag for 2006 versus 2005).

- **Family Level**

- Maternal age in years;
- Mother is college graduate (could be associate's degree);
- Home language is not English;
- Household monthly income above \$1,500;
- Number of children in household under age 8;

²⁵ The family- and child-level covariates were not used in the analysis of the classroom instructional outcomes.

²⁶ Baseline values were not available for print awareness, syntax and grammar, the two parenting scales, and the three PE and PC instructional scales because the data for these scales were first collected in spring 2005. For print awareness, syntax and grammar, and parent responsiveness, baseline IGDI scores were used in place of the unavailable baseline scores. This decision was based on some exploratory analyses of spring 2005 and earlier data. The validation for the decision may be found in the tables of appendix H, such as table H-7 that shows a significant relationship between project-average baseline English IGDI scores and 2005/2006 child scores on syntax and grammar.

²⁷ The exact rules for forming this covariate varied by outcome. It was not used for instructional outcomes (including participation outcomes). When used, it was the average score from, excluding the outcome of interest, the four English language child emergent literacy assessments that had been administered at baseline, the child social competence score, and the parent score on reading and vocabulary skill.

²⁸ The choice for project-level ratios rather than classroom-specific ratios was dictated by the fact that we cannot reliably align children assessed in 2005 and 2006 with classrooms defined in 2003 and 2004.

- Number of people in household over 18 years old; and
- Respondent to parent interview is Hispanic.²⁹
- **Family Level Summaries of Child-Level Data**³⁰
 - Flag for whether any of videotaped children were classified by their parents as having special needs;
 - Flag for whether any of videotaped children are male;
 - Average age of videotaped children in months; and
 - Maximum number of times that any of sample children in the family moved in last year.
- **Child Level**
 - Child has special needs;
 - Child is male;
 - Child is Hispanic/Latino;
 - Child age in months; and
 - Number of times that child moved in last year.

The pre-intervention measurements used to define the project-level covariates were conducted in the fall of 2003 and the spring of 2004, prior to the introduction of the CLIO curricula. Data from the two rounds were averaged together to form the project-level covariate. The child- and family-level covariates were measured in the parent interview at about the same time as the child assessments.

We made most of our decisions about which covariates to use based on their performance in an analysis of spring 2004 baseline IGDI scores or for theoretical reasons.³¹ After the analysis of 2005 and 2006 had started, the decision was made to add two additional covariates. One of the additions was year (2005 versus 2006). The other was a second measure of pre-intervention project quality, which we did by averaging together the other baseline assessments.

²⁹ We only used parent ethnicity in the analysis of parent outcomes. For child-level outcomes, child-specific ethnic origin was used.

³⁰ Family-level summaries of child-level data were only used in analyzing parent-level outcomes. For child-level outcomes, child-specific covariates were used.

³¹ For example, child gender was not significant but was nonetheless retained as a covariate under the theory that it might be important for other outcomes such as the social competence scale.

Multi-Level Modeling

We analyzed the CLIO data using the HLM package developed by Bryk and Raudenbush. As discussed in chapter 2, the sample is clustered by project and classroom. Also, there was no limit on the number of eligible siblings from the same family. It was not uncommon to have two siblings from the same family in the sample, and some families had three siblings in the sample. Additionally, because we jointly analyzed spring 2005 and spring 2006 outcome data, some children appeared twice in analytic files. So there are five natural levels in the data: project, classroom, family, child, and child-year. An empirical study of spring 2004 baseline data showed that the between-classroom variance is much smaller than either between-project variance or within-classroom variance. Moreover, a comparison of a two-level model (random effects for project and child only) with a three-level model (random effects for project, classroom, and child) showed that estimated between-child variance was not strongly affected by the choice of the number of levels. It appears that omitting the classroom-level random effect causes the first component to be over-estimated, which is far less problematic than bias in the child-level component of variance.³² With respect to the family-level, we believe that intra-family correlation is much more important for parenting outcomes than for child outcomes. Accordingly, for child outcomes, we fit three-level models with random effects for project, child, and child-year and fixed effects for stratum, project-level covariates, family-level covariates, and child-level covariates. Similarly, for parent and parenting outcomes, we fit three-level models with random effects for project, parent, and parent-year and fixed effects for stratum, project-level covariates, and family-level covariates. All of the models were linear.

The general form of the model³³ for child and parent outcomes was

$$Y_{ijkt} = \alpha_i + \beta_j + X_{ijk}\gamma + Z_{ij}\delta + \lambda_t + a_{ij} + c_{ijk} + e_{ijkt},$$

³² This conclusion is also supported by Jenkins, Lee, Cheah, and Leytush (2006).

³³ Since there is just one project per combination of study group and stratum, there is no need for a separate index for project. In the language of HLM documentation, this model description is equivalent to saying that we used a three-level linear model with covariates at the person and project levels. We conducted a simulation (with a slightly simpler model) on which HLM performed very well. (See Fan and Judkins 2006.) In the simpler model, we left out the random effects for classrooms and children. We also had just a single covariate at each level. However, we weakened the assumptions about normally distributed errors and constant variances. We also introduced random nonresponse at the project level. In terms of striking a good balance between constraining type I errors at or below the nominal level and having high statistical power, HLM was as good as or better than any of the alternatives tested.

where:

- the indices stand respectively for study group (i), stratum (j), child or parent (k), and year (t);
- the terms in Greek letters are fixed effects (α for treatment effect, β for stratum effect, γ for effects of family and child covariates, δ for effects of measured project covariates, and λ for the effect of year);
- the terms in lower-case Latin letters are random effects (a for project-level random error, c for stable child-level random error, and e for year-specific child- or parent-level random error);
- the terms in upper-case Latin characters are measured variables (Y for outcome, X for child- and/or family-level covariate row vector, and Z for project-level covariate row vector);
- random effects at each level are assumed to be independently and identically normally distributed; and
- random effects at different levels are assumed to be independent of each other.

When analyzing instructional outcomes as outcomes, we used a simpler model. Since these outcomes do not exist at the child level, the model omitted the fixed effects of child-level covariates and the random effect at the child level. Accordingly, the models for instructional outcomes were of the form

$$Y_{ijt} = \alpha_i + \beta_j + Z_{ij}\delta + \lambda_t + a_{ij} + e_{ijt},$$

where:

- the indices stand respectively for study group (i), stratum (j), and year (t);
- the terms in Greek letters are fixed effects (α for treatment effect, β for stratum effect, δ for effects of measured project covariates, and λ for the effect of year);

- the terms in lower-case Latin letters are random effects (a for project-level random error, and e for year-specific classroom-level random error);
- the terms in upper-case Latin characters are measured variables (Y for outcome and Z for project-level covariate row vector);
- random effects at each level are assumed to be independently and identically normally distributed; and
- random effects at different levels are assumed to be independent of each other.

Contrast Selection and Corrections for Multiple Comparison Testing

To answer the two primary research questions, eight specific contrasts were tested in addition to an overall test for any differences among the five experimental groups (table 4-1).³⁴ The overall test tells us whether any of the curriculum variations are more effective than any of the others, but does not provide guidance on which curriculum might be worthy of wider support and usage.

Referring to the model equation, these contrasts may be more compactly designated as $\frac{\alpha_2 + \alpha_4}{2} - \alpha_5$, $\alpha_2 - \alpha_4$, $\alpha_2 - \alpha_5$, $\alpha_4 - \alpha_5$, $\frac{\alpha_2 + \alpha_4}{2} - \frac{\alpha_1 + \alpha_3}{2}$, $(\alpha_2 - \alpha_1) - (\alpha_4 - \alpha_3)$, $\alpha_2 - \alpha_1$, and $\alpha_4 - \alpha_3$, respectively, where the study groups are numbered as follows (same as in table 2-1):

1. Let's Begin (ECE),
2. Let's Begin & Play and Learning Strategies (ECE/PE),
3. Partners for Literacy (ECE),
4. Partners for Literacy (ECE/PE), and
5. Control.

³⁴ IES provided guidance in the choice of the most appropriate contrasts to answer the two main research questions. We restricted the analysis to just these eight contrasts so that we could control false-positive discovery rates using the Bonferroni adjustment without unduly sacrificing statistical power.

Table 4-1. Research Questions and Contrasts

Research questions	Contrasts
(A) Is the combination of research-based, literacy-focused preschool, parenting, and parent-child curricula (the CLIO combined curricula) more effective than the existing combination of services in Even Start?	Average of both CLIO combined curricula versus control group
	Let's Begin & Play and Learning Strategies (ECE/PE) versus Partners for Literacy (ECE/PE)
	Let's Begin & Play and Learning Strategies (ECE/PE) versus control group
	Partners for Literacy (ECE/PE) versus control group
(B) Do research-based parenting and parent-child curricula (the CLIO parenting curricula) that focus on child literacy add value to the CLIO preschool curricula?	Average of both CLIO combined curricula versus average of both preschool curricula
	The difference between Let's Begin & Play and Learning Strategies (ECE/PE) and Let's Begin (ECE) versus the difference between Partners for Literacy (ECE/PE) and Partners for Literacy (ECE)
	Let's Begin & Play and Learning Strategies (ECE/PE) versus Let's Begin (ECE)
	Partners for Literacy (ECE/PE) versus Partners for Literacy (ECE)

The key contrast for the first primary research question tests is $\frac{\alpha_2 + \alpha_4}{2} - \alpha_5$, which compares the average of the two CLIO combined curricula with the control group. The study averages the two curricula together because this provides better power than separate tests if the two CLIO curricula have similar effectiveness. The study also tests the two developers' combined curricula against each other in case one of them is more effective than the other. This test uses the contrast $\alpha_2 - \alpha_4$. Each developer's combined curriculum is also tested against the control group in case only one of them is effective. These tests use the contrasts $\alpha_2 - \alpha_5$ and $\alpha_4 - \alpha_5$.

The key contrast for the second primary research question tests is $\frac{\alpha_2 + \alpha_4}{2} - \frac{\alpha_1 + \alpha_3}{2}$, which compares the average of the two CLIO combined curricula with the average of the two CLIO preschool curricula. The study averages the two combined curricula together as well as the two CLIO preschool curricula together because this provides better power to detect the added value of a research-based, child-literacy focused parenting curricula than separate tests if the two CLIO curricula have

similar effectiveness. The study also tests the two developers' contrasts for added value of the parenting curricula against each other in case one of them has a higher add-on value than the other. This test uses the contrast $(\alpha_2 - \alpha_1) - (\alpha_4 - \alpha_3)$. The added value from each developer's parenting curriculum is also tested in case only one of them has substantial added value. These tests use the contrasts $\alpha_2 - \alpha_1$, and $\alpha_4 - \alpha_3$.

A Bonferroni adjustment³⁵ was made for the fact that eight contrasts and the overall test were run. The test count was nine. This procedure allows the contrasts to be discussed even if the overall test is not significant. The critical value for all contrast tests was thus based on a test size of $0.05/9=0.0056$. In other words, only contrast tests with a *p*-value smaller than 0.0056 are highlighted in this report as evidence for the effectiveness of the CLIO curricula although all tests are reported, whether significant or not. The standard *p*-value criterion of 0.05 was still used for the overall test.

Reporting Findings

Effect Size Calculation

Most estimates of CLIO effects are expressed in terms of "effect sizes." Effect sizes were calculated by scaling the contrasts discussed above as proportions of the population standard deviation within the control group in 2005. This was done to facilitate assessment of the practical importance of any statistically significant findings. A detailed description of the methodology used in this study may be found in appendix I.

Another way of interpreting effect sizes is to translate them into percentile standings as shown in table 4-2. An effect size of 1.0 means that 84 percent of the treated sample is scoring higher than the average subject in the control group. An effect size of 0.5 means that 69 percent of the treated sample is scoring higher than the average subject in the control group. An effect size of 0.2 means that 58 percent of the treated sample is scoring higher than the average subject in the control group.

³⁵ We considered using the Benjamini-Hochberg adjustment instead of a Bonferroni adjustment. However, the theory does not appear to be well developed for experiments with more than two experimental groups.

Table 4-2. Effect Sizes and Percentile Standings

Effect size	Percent of treated group scoring higher than comparison group
1.4	91.9
1.3	90
1.2	88
1.1	86
1.0	84
0.9	82
0.8	79
0.7	76
0.6	73
0.5	69
0.4	66
0.3	62
0.2	58
0.1	54
0	50

Confidence Intervals on Effect Sizes, *p*-Values, and Statistical Significance

Tables in chapters 5, 6, and 7, as well as some tables in various appendices, contain effect sizes, confidence intervals on effect sizes, and *p*-values for hypothesis tests. The confidence intervals and the *p*-values are not adjusted for multiple comparisons. As noted earlier, however, statistical significance was determined based on critical values that were adjusted for the fact that multiple tests were conducted for each outcome variable. In the tables, asterisks identify contrasts that are statistically significant after applying the Bonferroni adjustment for multiple comparisons. For example, the threshold for statistical significance for the analyses reported in table 5-2 is 0.0056 (rather than .05); asterisks denote any comparisons that are statistically significant using that threshold. Each table has a note indicating the *p*-value threshold for statistical significance applied to the analyses reported in that table.

Statistical Power

Table 4-3 reports on the minimum detectable effect sizes (MDES) given the variances that were actually observed in the analysis. (The minimum detectable effect sizes are generally smaller than had been projected during the design phase of the research.)

The first data column in the table shows the MDES for the first primary research question: the effect of CLIO combined curricula. Excluding the Spanish-language assessment of expressive language (for which the sample size was smaller) and participation, the average MDES for the CLIO combined curricula on the child outcomes was 0.25. The comparable figure for the parent outcomes was also 0.25. Power for classroom outcomes was much weaker as would be expected given that there are many fewer classrooms than children. The average for them was 0.63.

The second data column in the table shows the MDES size for the second primary research question: the incremental effect of CLIO parenting curricula. Excluding the Spanish-language assessment of expressive language (for which the sample size was smaller) and participation, the average MDES for the incremental effect of the CLIO parenting curricula on the child outcomes was 0.20. The comparable figure for the parent outcomes was 0.22. Power for classroom outcomes and participation was weaker.

The third data column in the table shows the MDES size for all contrasts in which any single study group is contrasted with another. Excluding the Spanish-language assessment of expressive language (for which the sample size was smaller) and participation, the average MDES for one-to-one group contrasts on the child outcomes was 0.29. The comparable figure for the parent outcomes was 0.26. Power for classroom outcomes and participation was weaker.

The fourth data column in the table shows the MDES for tests of the differential incremental effects of the two CLIO parenting curricula. Excluding the Spanish-language assessment of expressive language (for which the sample size was smaller) and participation, the average MDES for differential incremental effects on the

Table 4-3. Minimum Detectable Effect Sizes

Outcome measure	Contrasts			
	Average of both combined curricula versus control group	Average of both combined curricula versus average of both preschool curricula	Let's Begin and PALS (ECE/PE) versus Pfl (ECE/PE) Let's Begin and PALS (ECE/PE) versus control group Pfl (ECE/PE) versus control group Let's Begin and PALS (ECE/PE) versus Let's Begin (ECE) Pfl (ECE/PE) versus Pfl (ECE)	Difference between Let's Begin and PALS (ECE/PE) and Let's Begin (ECE) versus the difference between Pfl (ECE/PE) and Pfl (ECE)
Child outcomes				
Expressive language: English	0.24	0.19	0.29	0.38
Expressive language: Spanish	0.38	0.32	0.44	0.63
Receptive vocabulary	0.21	0.18	0.24	0.32
Phonological awareness: Elision	0.26	0.21	0.29	0.41
Phonological awareness: Blending	0.27	0.20	0.30	0.42
Print knowledge	0.25	0.20	0.30	0.42
Syntax and grammar	0.22	0.20	0.25	0.33
Child social competence	0.27	0.23	0.32	0.44
Parent outcomes				
Parent interactive reading skill	0.31	0.29	0.26	0.55
Parent responsiveness to child	0.25	0.22	0.30	0.43
Parent reading and vocabulary	0.19	0.15	0.22	0.31
Instructional outcomes				
Support for oral language development	0.58	0.43	0.63	0.84
Support for print knowledge	0.63	0.45	0.66	0.96
Support for phonological awareness	0.92	0.79	1.14	1.65
Support for print motivation	0.56	0.44	0.63	0.86
Literacy resources in the classroom	0.64	0.49	0.69	0.93
PE time spent on child literacy	0.68	0.65	0.89	1.28
PE time spent on parenting skills	0.55	0.52	0.68	1.03
PC time spent interacting on child literacy activities	0.44	0.38	0.55	0.77
Participation outcomes				
Child: monthly hours of ECE instruction received	0.36	0.33	0.45	0.69
Parent: monthly hours of PE/PC instruction received	0.53	0.57	0.69	1.16

NOTE: Contrasts shown here are the same as those found in table 4-1.

child outcomes was 0.39. This number is larger than for any of the other contrasts because of the double differencing. The comparable figure for the parent outcomes was 0.43. Power for classroom outcomes and participation was weaker.

All of these power calculations were approximated with the formula:

$$MDES = \frac{\Phi^{-1}(1 - .05/2/9) + \Phi^{-1}(.8) \left(\frac{U_{95} - L_{95}}{2} \right)}{\Phi^{-1}(1 - .05/2)} = 0.922(U_{95} - L_{95}),$$

where (L_{95}, U_{95}) is the 95 percent confidence interval on the effect size as actually estimated in chapter 5 and appendix K, Φ denotes the cumulative distribution function for the standard normal distribution, and the number 9 is the count of tests used in the Bonferroni adjustments for multiple testing.

Handling of Missing Data

There are a variety of types of missing data in the CLIO data system. We used a range of compensation strategies for these different varieties of missing data.

- We replaced two projects that either dropped out of the study or lost their Even Start funding during the summer of 2004. Data were collected at the replacement projects in spring 2006 only. (See figure 2-1.)
- We made no adjustments for six additional projects that lost their Even Start funding during the summer of 2005.

- We imputed missing items in the parent interviews (or whole parent interviews if need be) for children with substantial other data. This is important because the parent interviews are the source of most of the covariates.³⁶
- We dropped children with missing child emergent-literacy outcomes or social competence scores from the respective analyses. (However, if a child had a score for a particular emergent-literacy assessment, then that score was used without respect to the availability of scores on the other emergent-literacy and social competence scales.)
- We imputed parent assessment scores for those parents who took at least one of the four assessments but not all of them. For the impact analysis of parent language and literacy we analyzed the set of parents who took at least one assessment.
- For analysis of parenting, we analyzed those parents for whom we had both the book video and the toy video. We used imputed parent interview data when necessary to calculate the two parenting outcomes.
- For analysis of instructional outcomes, we ignored projects that had no activity for us to observe at the time of scheduled trips.
- For the analyses in chapter 7 where there can be missing data in the putative causal agent as well as in the outcome score, we discarded cases that are missing one or more of the variables involved.

Additional information on the methods used for each type of missing data can be found in appendix J. This appendix also provides information on child and parent sample sizes as well as project closures.

³⁶ Where we used imputation, we used complex methods that are designed to preserve covariance structures. These methods do not assume that participating parents are generally similar to nonparticipating parents. Rather, they assume that they are similar within small groups defined by the data that are available on the dyads such as the child assessment scores.

Given the low rates of imputation that were required and that, for the most part, we only imputed covariates rather than outcomes, we did not account for imputation variance in variance estimates. We considered using multiple imputation to improve the estimation of post-imputation variances as suggested by Rubin (1987), but that would have required multiple runs of HLM on parallel datasets—a complication that did not seem warranted given the level of imputation and the variables imputed.

5. IMPACTS OF THE CLIO CURRICULA ON CHILDREN, PARENTS, CLASSROOM INSTRUCTION, AND PARTICIPATION

This chapter presents detailed findings about the effects of the CLIO curricula on the 21 outcomes defined for this study: seven child emergent literacy outcomes, a teacher rating of child social competence, three measures of parenting skills and parent literacy, eight measures of preschool and parenting classroom instruction, and two measures of parent and child participation in Even Start's instructional services. Two sets of analyses are presented for each of these outcome areas, addressing the study's two main research questions.

The first main research question asks whether the integrated CLIO combined curricula produced better outcomes for children and parents compared with regular Even Start services. These combined curricula provided explicit, focused literacy instruction in the preschool classroom, linked with integrated child literacy-focused parenting curricula. This question was addressed through four analytic contrasts. The key contrast grouped the two CLIO combined curricula and compared them to the control group. The subsequent three contrasts provide detail on how the two CLIO combined curricula performed when compared to each other and to the control group. The four contrasts are listed below.

- both CLIO combined curricula vs. control,
- Let's Begin and PALS vs. ECE/PE Partners for Literacy,
- Let's Begin and PALS vs. control, and
- ECE/PE Partners for Literacy vs. control.

The second main research question for this study sought to estimate the added value of the CLIO parenting curricula over and above the CLIO preschool curricula. This was an interesting question because the Even Start program is based on the assumption that providing PE and PC will add value to the ECE program for children, despite the fact that prior researchers have not been able to provide strong evidence supporting this hypothesis (St.Pierre, Layzer, Goodson, and Bernstein 1997; St.Pierre et al. 2003; White, Taylor, and Moss 1992). This second main research question was addressed through four analytic contrasts. The key contrast compared both of the

CLIO combined curricula with both of the CLIO preschool curricula. The second contrast compares the added value of the PALS parenting curriculum against the added value of the Partners for Literacy parenting curriculum. The final two contrasts provide results when each of the CLIO combined curricula were tested against the corresponding CLIO preschool curricula. The four contrasts are listed below.

- added value of both of the CLIO parenting curricula to both of the preschool curricula,
- added value of PALS vs. added value of the Partners parenting curriculum,
- added value of PALS to the Let's Begin preschool curriculum, and
- added value of the Partners parenting curriculum to the Partners preschool curriculum.

Tests conducted for each developer separately are only discussed when there were statistically significant differences between the two developers' curricula. However, appendix K provides the results for each outcome of the developer-specific tests.

In this chapter, we distinguish between statistically significant and non-significant findings on the basis of having a *p*-value less than .0056 ($.05/9 = .0056$), since we conducted nine tests of relationships among the study groups for each outcome. The confidence intervals presented in this chapter are expressed in terms of effect size units and have no Bonferroni adjustment.

In addition to conducting hypothesis tests for the eight contrasts that are required to address the two primary research questions, we ran an overall test for each outcome to see if there were significant differences among the five study groups. This was done because of the desire to be sensitive to unanticipated patterns of differences among the study groups. However, the Bonferroni test for multiple comparisons was set up in a way that removes the requirement to have a significant result on the overall test before discussing contrasts, as is required with some procedures for multiple comparison testing. Specifically, we divided the nominal alpha level by nine rather than eight in assessing the significance of the eight contrasts. It is not required to adjust the significance level of the overall test, and it never happened that the overall test was

significant without at least one contrast also being significant or that an individual contrast was significant without the overall test also being significant. Accordingly, we only discuss significance tests on the contrasts and not the results of the overall tests (see table 5-1). (For the complete set of statistics on the full model for each of the 21 outcomes, see appendix H.)

Impacts on Children

Even Start projects provide ECE to children in low-income families with the goal of helping children learn to read and preparing them for success in school. In the Even Start model, impacts on children were hypothesized to occur through two pathways—directly, as a result of the CLIO preschool curricula, and indirectly, as a result of impacts on parenting skills. The CLIO curricula were selected for the strength of their preschool components, in particular the language and literacy dimensions.

Effects of CLIO Combined Curricula on Children (Research Question 1).

The combined curricula had no statistically significant effect on any of the six measures of child emergent literacy in English, or on the Spanish measure of expressive language. Table 5-2 provides the results for the first and most important contrast that compared the outcomes of projects that were assigned to implement the CLIO combined curricula (Let's Begin and PALS, and ECE/PE Partners for Literacy) against the outcomes of the control group of Even Start projects that were assigned to implement their usual instructional services. Estimated effect sizes on emergent literacy outcomes were all smaller than 0.13 in absolute value, with confidence interval limits all bounded by 0.27 in absolute value. The impacts of the two different versions of the CLIO combined curricula were not statistically different from each other (see table K-1 in appendix K).

The CLIO combined curricula had a statistically significant positive effect on children's social competence, as rated by classroom teachers (table 5-2). The effect size for the two CLIO curricula combined was 0.22, with no statistically significant difference between the two developers (see table K-1 in appendix K).

Table 5-1. Unadjusted Outcome Means (average of Spring 2005 and Spring 2006 data) by Study Group, and Results of Overall Test for Any Differences Across the Five Groups

Outcome measure	Study group					Overall test across 5 groups
	Let's Begin (ECE)	Let's Begin and PALS (ECE/PE)	Partners for Literacy (ECE)	Partners for Literacy (ECE/PE)	Control	
Child Outcomes						
Expressive language: English	15.8	15.3	16.4	14.0	16.3	0.312
Expressive language: Spanish	16.0	16.6	14.9	16.5	16.0	>.500
Receptive vocabulary	254.0	248.7	254.2	241.6	253.4	>.500
Phonological awareness: Elision	251.5	253.4	249.9	242.1	249.2	>.500
Phonological awareness: Blending	251.4	249.9	257.2	244.4	254.7	0.181
Print knowledge	256.7	259.7	260.2	240.1	248.2	0.009*
Syntax and grammar	251.8	251.1	251.5	240.5	251.4	>.500
Child social competence	247.3	253.6	254.5	252.2	241.8	0.008*
Parent Outcomes						
Parent interactive reading skill	0.18	0.28	-0.07	0.36	-0.12	0.000*
Parent responsiveness to child	-0.20	-0.20	-0.48	-0.29	-0.51	0.000*
Parent reading and vocabulary	259.8	250.3	248.9	243.0	247.8	0.296
Instructional Outcomes						
Support for oral language development	45.3	47.2	47.3	44.8	43.9	0.085
Support for print knowledge	51.0	57.8	53.5	52.3	49.5	0.000*
Support for phonological awareness	55.3	59.8	53.6	53.5	51.2	0.228
Support for print motivation	48.3	53.7	56.8	54.3	47.4	0.004*
Literacy resources in the classroom	53.2	53.9	50.5	51.1	48.3	0.005*
PE time spent on child literacy	26.5	42.2	23.4	40.7	16.3	0.000*
PE time spent on parenting skills	41.0	35.1	42.4	19.6	32.0	0.000*
PC time spent interacting on child literacy activities	19.2	18.7	17.9	28.0	18.4	0.088
Participation Outcomes						
Child: monthly hours of ECE instruction received	46.1	47.3	40.2	44.6	40.2	>.500
Parent: monthly hours of PE/PC instruction received	12.2	12.6	9.8	12.0	9.8	0.324

*Statistically significant at $p < 0.05$.

NOTE: Study group means are not covariate adjusted, but the p -values are from an overall chi-square test that is covariate adjusted.

**Table 5-2. Effects of CLIO Combined Curricula on Children (Research Question 1)
(average of spring 2005 and spring 2006)**

Child outcomes	Average of the two CLIO combined curricula vs. control		
	Effect size	95% CI	<i>p</i> -Value
Expressive language: English	-0.11	-0.25,0.02	0.081
Expressive language: Spanish	0.05	-0.15,0.26	>.500
Receptive vocabulary	-0.09	-0.20,0.03	0.128
Phonological awareness: Elision	0.00	-0.14,0.14	>.500
Phonological awareness: Blending	-0.13	-0.27,0.02	0.083
Print knowledge	0.05	-0.09,0.18	>.500
Syntax and grammar	-0.08	-0.20,0.03	0.159
Social competence	0.22*	0.07,0.36	0.003

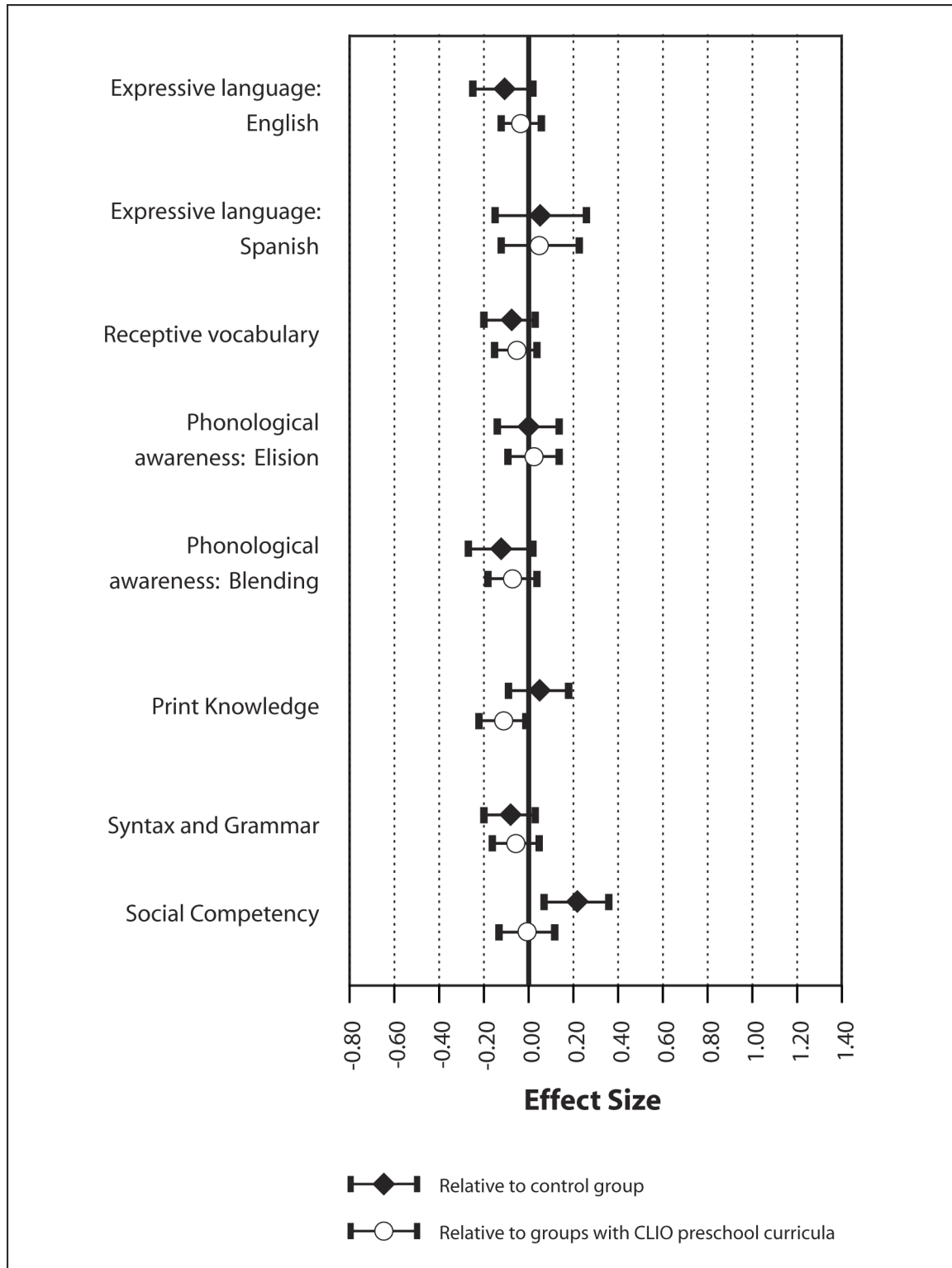
*Statistically significant at $p < 0.0056$, a value chosen according to Bonferroni's rule to control false positive discovery rates across the nine contrasts run for each outcome.

These results are also shown graphically in figure 5-1. The horizontal error bands with center solid diamonds repeat the information in table 5-2. The error bands represent 95 percent confidence intervals on the effect sizes. Up and down the chart, it is shown that there were no positive effects of the CLIO combined curricula on emergent literacy, as can be noted by the fact that none of the confidence bands exclude zero, even before adjustment for multiple comparisons. The only effect that is statistically significant is for social competence. To put the effect size on social competence into some context, the natural uncontrolled difference in social competence between the boys and girls in this study is equivalent to an effect size of 0.39.

Note that no attempt was made to determine whether the CLIO curricula closed (or widened) the gap in social competence between girls and boys. The natural uncontrolled difference merely expresses the well-known gender maturity gap on the same scale as the CLIO effect.³⁷ Similar information may be found in table H-8, in which it is shown that the coefficient for gender in the model of social competence is one of the highest in the model.

³⁷ See for example Walker (2004).

Figure 5-1. Effect Sizes for CLIO Combined Curricula on Child Outcomes Relative to Both the Control Group and the CLIO Preschool Curricula (average of spring 2005 and spring 2006)



Incremental Effects of CLIO Parenting Curricula on Children (Research Question 2). The CLIO parenting curricula had no statistically significant incremental effect on the six measures of child emergent literacy in English, or on the Spanish measure of expressive language. That is, adding research-based parenting components focused on child literacy did not add significantly to children's outcomes beyond what was achieved with the CLIO preschool curricula. Table 5-3 provides the results of the first and most important contrast, which compares the average outcomes for the two CLIO combined curricula with the average outcomes for the two CLIO preschool curricula. The estimated effect sizes on emergent literacy outcomes were all smaller than 0.11 in absolute value with confidence interval limits all bounded by 0.23 in absolute value. There were no statistically significant differences between the two developers' curricula (see table K-2 in appendix K).

Table 5-3. Incremental Effects of CLIO Parenting Curricula on Children (Research Question 2) (average of spring 2005 and spring 2006)

Child outcomes	Average of the two CLIO combined curricula vs. average of the two CLIO preschool curricula		
	Effect size	95% CI	<i>p</i> -Value
Expressive language: English	-0.04	-0.14,0.06	>.500
Expressive language: Spanish	0.05	-0.12,0.23	>.500
Receptive vocabulary	-0.05	-0.15,0.04	0.290
Phonological awareness: Elision	0.02	-0.09,0.14	>.500
Phonological awareness: Blending	-0.07	-0.18,0.04	0.194
Print knowledge	-0.11	-0.22,-0.01	0.035
Syntax and grammar	-0.06	-0.16,0.05	0.287
Social competence	-0.01	-0.13,0.12	>.500

*Statistically significant at $p < 0.0056$, a value chosen according to Bonferroni's rule to control false positive discovery rates across the nine contrasts run for each outcome.

CLIO parenting curricula had no statistically significant incremental effect on child social competence, nor was there any statistically significant difference in effectiveness between the two developers' curricula on this outcome measure (table 5-3 and table K-2 in appendix K).

Figure 5-1 also displays the results for research question 2. The horizontal error bands with center open circles repeat the information in table 5-3. The error bands represent 95 percent confidence intervals on the effect sizes. Up and down the chart, it is

shown that there were no statistically significant incremental effects of the CLIO parenting curricula on any of the child outcomes, including social competence.

Impacts on Parents

The parenting curricula implemented in CLIO focused on teaching parents to be effective teachers of their children and to support their child's development of emergent literacy skills. A number of aspects of parenting have been shown to be related to children's emergent literacy development, primarily in descriptive and correlational research. First, children from homes where parents engage in elaborated conversations with them, model the uses of literacy, and engage them in activities that promote basic understandings about literacy (e.g., shared book reading) have more well-developed language and literacy-related skills than children from homes where these activities are less frequent (Hart and Risley 1995; Snow, Barnes, Chandler, Hemphill and Goodman 1991). Second, shared-reading interventions in which parents read aloud to children had a moderate, significant effect on children's oral language outcomes (Arnold, Epstein, Lonigan and Whitehurst 1994; Huebner 2000; Lonigan, Anthony, Bloomfield, Dyer, and Samwel 1999; Whitehurst et al. 1988; Whitehurst et al. 1994). (Third, the affective relationship between a parent and child has been shown to make a difference in children's learning—children learn best when the adults in their lives are responsive to their cues and needs (Pianta 2003).

Impacts on parents were investigated using measures of parenting skills that were based on observing the parent and child reading together and playing with a toy, as well as some self-reported home behaviors of the parents. Two outcome measures were developed from these observations—a scale that measures parents' interactive reading skills and a scale that measures parents' responsiveness to their child. In addition, standardized tests were administered to parents to measure their reading skills and vocabulary based on the possibility that the CLIO parenting curricula might have the unintended consequence of affecting parent literacy.

Effects of CLIO Combined Curricula on Parents (Research Question 1).

The CLIO combined curricula had statistically significant positive effects on both measures of parenting skills, including an effect of 0.48 on parent interactive reading skill and an effect of 0.22 on parent responsiveness to child (table 5-4). The CLIO

combined curricula had no statistically significant effect on parent reading skills and vocabulary. There were no statistically significant differences on any of the three parent outcomes between the two developers' curricula (see table K-3 in appendix K).

**Table 5-4. Effects of CLIO Combined Curricula on Parents (Research Question 1)
(average of spring 2005 and spring 2006)**

Parent outcomes	Average of the two CLIO combined curricula vs. control		
	Effect size	95% CI	p-Value
Interactive reading skill	0.48*	0.31,0.65	0.000
Responsiveness to child	0.22*	0.09,0.36	0.002
Reading skills and vocabulary	-0.04	-0.14,0.06	>.500

*Statistically significant at $p < 0.0056$, a value chosen according to Bonferroni's rule to control false positive discovery rates across the nine contrasts run for each outcome.

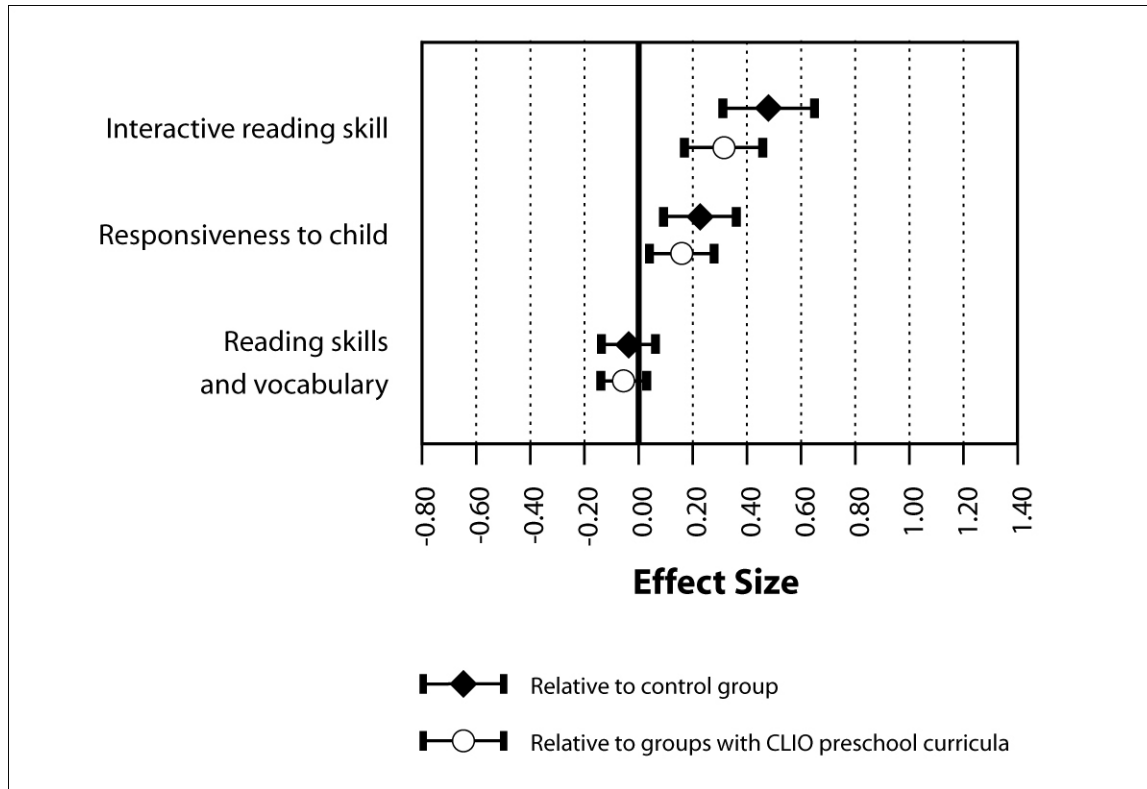
The main results are also shown graphically in figure 5-2. The horizontal error bands with center solid diamonds repeat the information in table 5-4. The error bands represent 95 percent confidence intervals on the effect sizes and show statistically significant effects on interactive reading skill and responsiveness to child, but no statistically significant effect on parent literacy. To put the effect sizes for parenting into some context, the natural uncontrolled difference in parent interactive reading skill between mothers with and without a postsecondary degree³⁸ in this study was equivalent to an effect size of 0.46.³⁹ Similarly, the natural uncontrolled difference in parent responsiveness between mothers with and without postsecondary education in this study was equivalent to an effect size of 0.38.⁴⁰

³⁸ This is the same binary variable as used as a covariate in all the analyses. It is coded one if the mother has an associate's, bachelor's, or graduate degree, and zero otherwise. About 9 percent of CLIO mothers have college degrees. This percentage is well balanced across the five study groups.

³⁹ The uncontrolled difference in parental interactive reading skill across parental education levels is highly significant statistically in a multi-level model that controls on study group and stratum. There is no evidence of any interaction for parental interactive reading skill between study group and parental education.

⁴⁰ This effect size is also highly significant in a multi-level model. Moreover, there is no significant evidence that the relationship between parental responsiveness and parental education varies by study group.

Figure 5-2. Effect Sizes for CLIO Combined Curricula on Parent Outcomes Relative to Both the Control Group and the CLIO Preschool Curricula (average of spring 2005 and spring 2006)



Note that no attempt was made to determine whether the CLIO parenting curricula closed (or widened) the gap in parenting behaviors between mothers of varying education levels. The natural uncontrolled differences merely indicate how the outcomes vary between groups with sharp differences in personal capital, both in terms of what young adults have before the decision to enter college and what they have after college attendance. Similar information may be found in tables H-10 and H-11, in which it is shown that the coefficients for mother's education in models for the two parenting outcomes are higher than those for any other covariate.

Incremental Effects of CLIO Parenting Curricula on Parents (Research Question 2). Conceptually, we expected any effects on parenting to be caused by the CLIO parenting curricula rather than by the CLIO preschool curricula. As expected, the CLIO parenting curricula had a statistically significant positive incremental effect (e.s.=0.30) on parent interactive reading skill when preschool instruction was held

constant (table 5-5). Although the 95 percent confidence interval for the effect on parent responsiveness for adding the CLIO parenting curricula to the preschool curricula does not contain zero, the hypothesis test slips out of the critical region after the Bonferroni correction.

Table 5-5. Incremental Effects of CLIO Parenting Curricula on Parents (Research Question 2) (average of spring 2005 and spring 2006)

Parent outcomes	Average of the two CLIO combined curricula vs. average of the two CLIO preschool curricula		
	Effect size	95% CI	<i>p</i> -Value
Interactive reading skill	0.30*	0.15,0.46	0.000
Responsiveness to child	0.16	0.04,0.28	0.010
Reading skills and vocabulary	-0.06	-0.14,0.03	0.180

*Statistically significant at $p < 0.0056$, a value chosen according to Bonferroni's rule to control false positive discovery rates across the nine contrasts run for each outcome.

There was no statistically significant incremental effect of CLIO parenting curricula on parent reading skills and vocabulary (table 5-5). The point estimate is negative, and the upper limit of the 95 percent confidence interval is 0.03.

When holding preschool instruction constant, there were no statistically significant differences between the two developers in the incremental effects of their parenting curricula on any of the three parent outcomes (see table K-4 in appendix K).

Figure 5-2 also displays the results for research question 2. The horizontal error bands with center open circles repeat the information in table 5-5. The error bands represent 95 percent confidence intervals on the effect sizes.

Impacts on Classroom Instruction

It was hypothesized that preschool curricula that included a systematic focus on supporting the development of language and early literacy skills would promote better emergent literacy outcomes for Even Start children. One of the main criteria for choosing the two CLIO preschool curricula (Let's Begin and ECE Partners for Literacy) was that they included instructional activities in the major areas of emergent literacy. It was hypothesized that if the curricula were implemented as designed, changes in

instructional practices would be apparent and meaningfully large and would ultimately lead to improved child outcomes. Further, it was hypothesized that increasing the extent to which parenting curricula focus on child literacy skills would lead to improvements in parents' teaching skills, and consequently to improved child outcomes.

Effects of CLIO Combined Curricula on Classroom Instruction (Research Question 1).

Preschool Instruction. The CLIO combined curricula had statistically significant and positive effects on two of the five instructional outcomes: support for print knowledge (e.s.=0.69) and literacy resources in the classroom (e.s.=0.52) (table 5-6). The positive effect on support for print knowledge was greater for Let's Begin and PALS than for ECE/PE Partners for Literacy (see table K-5 in appendix K). There were no other statistically significant differences in effects on preschool instruction between the two developers.

Table 5-6. Effects of CLIO Combined Curricula on Instruction (Research Question 1) (average of spring 2005 and spring 2006)

Instructional outcomes	Average of the two CLIO combined curricula vs. control		
	Effect size	95% CI	p-Value
Support for oral language development	0.29	-0.02,0.61	0.061
Support for print knowledge	0.69*	0.35,1.03	0.000
Support for phonological awareness	0.53	0.03,1.03	0.034
Support for print motivation	0.36	0.06,0.67	0.017
Literacy resources in the classroom	0.52*	0.17,0.86	0.003
PE time spent on child literacy	1.01*	0.64,1.38	0.000
PE time spent on parenting skills	-0.15	-0.45,0.15	0.314
PC time spent interacting on child literacy activities	0.21	-0.03,0.45	0.079

*Statistically significant at $p < 0.0056$, a value chosen according to Bonferroni's rule to control false positive discovery rates across the nine contrasts run for each outcome.

Parenting Education. The CLIO combined curricula had a statistically significant positive effect (e.s.=1.01) on the percentage of PE class time spent on child literacy (table 5-6). This effect was hypothesized, since the CLIO curricula provided PE activities that focused specifically on child literacy, while prior research showed that typical Even Start PE classes spent considerable time on activities not directly related to

child literacy (e.g., nutrition, parent health issues, social conversation). The CLIO combined curricula did not have a statistically significant effect on the amount of PE time spent on parenting skills (table 5-6), nor was there a statistically significant difference between the two curricula on how time was used in PE classes (see table K-5 in appendix K).

Parent-Child Literacy Activities. The CLIO combined curricula did not have a statistically significant effect on the percentage of PC class time spent with parents and children interacting on child literacy activities, nor was there a statistically significant difference on this outcome between the two developers' curricula (table 5-6 and table K-5 in appendix K).

The main results are also shown graphically in figure 5-3. The horizontal error bands with center solid diamonds repeat the information in table 5-6. The error bands represent 95 percent confidence intervals on the effect sizes. There are a wide variety of estimated effects, with the strongest being for PE class time spent on child literacy activities.⁴¹

Incremental Effects of CLIO Parenting Curricula on Classroom Instruction (Research Question 2).

Preschool Instruction. The CLIO parenting curricula were not hypothesized to affect the instruction in preschool classrooms, but corresponding tests were nonetheless run for reasons of symmetry in the testing procedure. The addition of the CLIO parenting curricula had no statistically significant incremental effect on any of the five measures of preschool classroom instruction when compared with projects that were assigned to the CLIO preschool curricula, nor was there a statistically significant difference between the two curricula (table 5-7 and table K-6 in appendix K).

Parenting Education. Compared with projects assigned to the CLIO preschool curricula, projects that were assigned the CLIO combined curricula had a statistically significant positive effect on the percentage of PE class time spent on child literacy (e.s.=0.68) (table 5-7). As discussed in chapter 2, we split parenting education

⁴¹ Although the confidence bands for support for phonological awareness and support for print motivation exclude zero, the effect sizes are not significant once adjusted for multiple comparisons.

class time into three categories. Increased time on child literacy must come out of one or both the remaining categories of parenting skills and “other,” a residual category which consists mostly of personal and adult-focused activities such as life skills, leisure activities, social conversation, and meals. There was a statistically significant negative effect on the percentage of PE class time spent on parenting skills (e.s.= -0.45). None of the differences between the developers on these outcomes was statistically significant (see table K-6 in appendix K).

Table 5-7. Incremental Effects of CLIO Parenting Curricula on Instruction (Research Question 2) (average of spring 2005 and spring 2006)

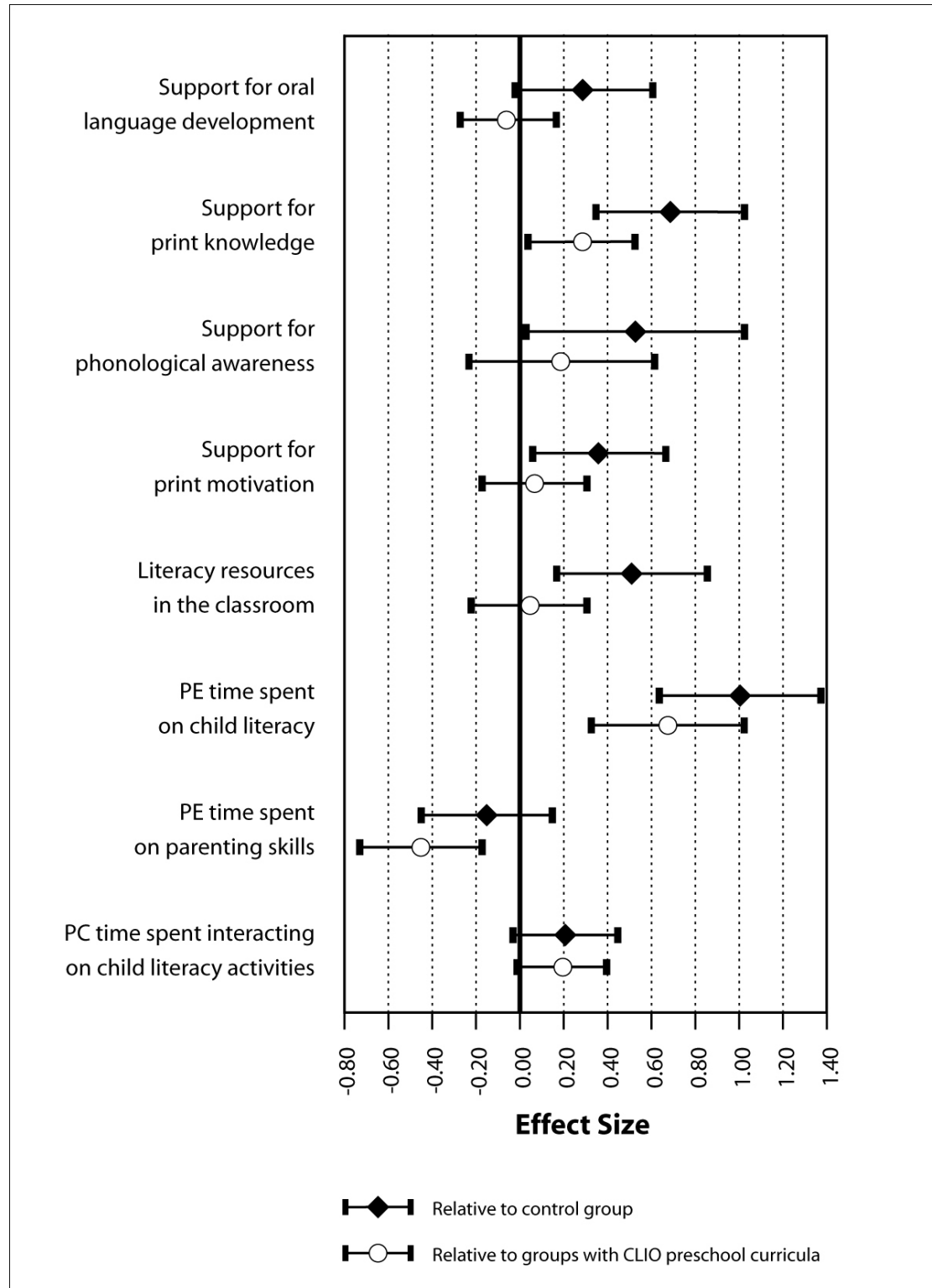
Instructional outcomes	Average of the two CLIO combined curricula vs. average of the two CLIO preschool curricula		
	Effect size	95% CI	p-Value
Support for oral language development	-0.06	-0.29,0.17	>.500
Support for print knowledge	0.29	0.04,0.53	0.020
Support for phonological awareness	0.19	-0.23,0.62	>.500
Support for print motivation	0.07	-0.17,0.31	>.500
Literacy resources in the classroom	0.05	-0.22,0.31	>.500
PE time spent on child literacy	0.68*	0.33,1.03	0.000
PE time spent on parenting skills	-0.45*	-0.73,-0.17	0.002
PC time spent interacting on child literacy activities	0.20	-0.01,0.40	0.056

*Statistically significant at $p < 0.0056$, a value chosen according to Bonferroni's rule to control false positive discovery rates across the nine contrasts run for each outcome.

Parent-Child Literacy Activities. When compared with projects that were assigned the CLIO preschool curricula, projects that were assigned the CLIO combined curricula had no statistically significant incremental effect on the percentage of PC class time in which parents were interacting with children on activities directly related to child literacy (table 5-7). There was no statistically significant difference between the two developers (see table K-6 in appendix K).

Figure 5-3 also displays the results for research question 2. The horizontal error bands with center open circles repeat the information in table 5-7. The error bands represent 95 percent confidence intervals on the effect sizes.

Figure 5-3. Effect Sizes for CLIO Combined Curricula on Instructional Outcomes Relative to Both the Control Group and the CLIO Preschool Curricula (average of spring 2005 and spring 2006)



Impacts on Participation

The study tested the hypothesis that families in projects that were assigned to implement the CLIO combined curricula would be more engaged by the new curricula and therefore would participate more consistently in instructional services.

Effects of CLIO Combined Curricula on Participation (Research Question 1). The CLIO combined curricula did not have a statistically significant effect on child participation in preschool or on parent participation in parenting education (table 5-8). This finding was consistent across both developers (see table K-7 in appendix K).

Table 5-8. Effects of CLIO Combined Curricula on Participation (Research Question 1) (average of spring 2005 and spring 2006)

Participation outcomes	Average of the two CLIO combined curricula vs. control		
	Effect size	95% CI	p-Value
Child: monthly hours of ECE instruction received	0.12	-0.07,0.32	0.196
Parent: monthly hours of PE/PC instruction received	0.25	-0.04,0.54	0.080
*Statistically significant at $p < 0.0056$, a value chosen according to Bonferroni's rule to control false positive discovery rates across the nine contrasts run for each outcome.			

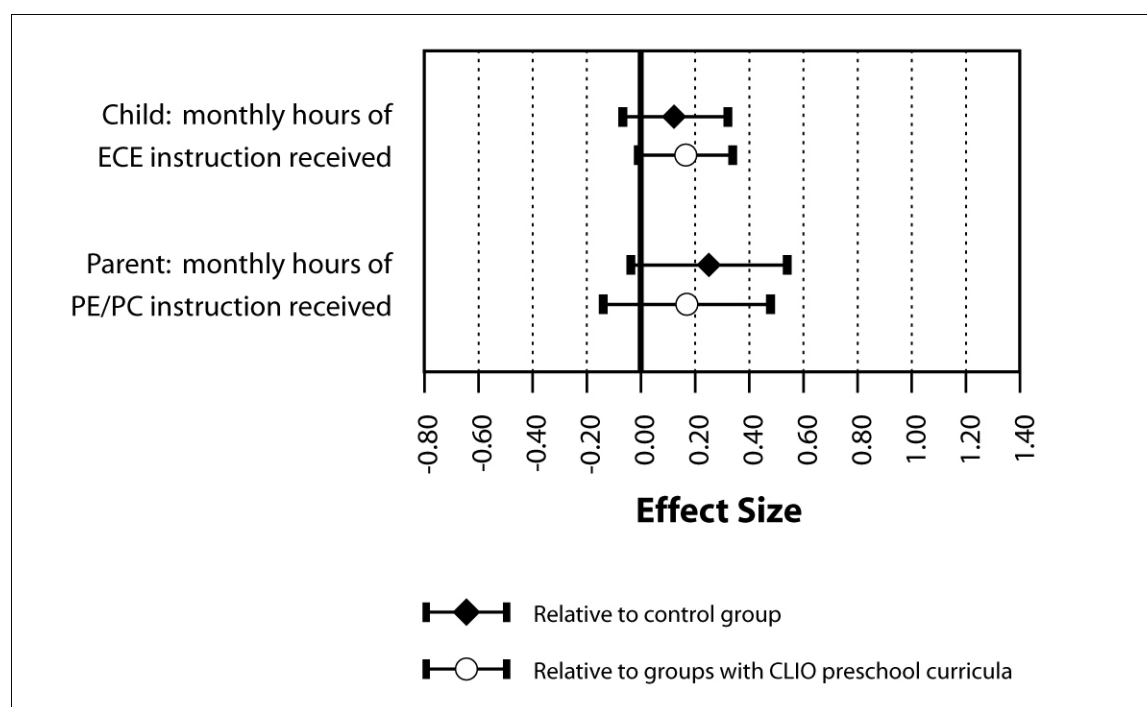
Incremental Effects of CLIO Parenting Curricula on Participation (Research Question 2). The CLIO parenting curricula did not have a statistically significant incremental effect on child participation in preschool or on parent participation in parenting education (table 5-9). This finding was consistent for both developers (see table K-8 in appendix K).

Table 5-9. Incremental Effects of CLIO Parenting Curricula on Participation (Research Question 2) (average of spring 2005 and spring 2006)

Participation outcomes	Average of the two CLIO combined curricula vs. average of the two CLIO preschool curricula		
	Effect size	95% CI	p-Value
Child: monthly hours of ECE instruction received	0.16	-0.01,0.34	0.064
Parent: monthly hours of PE/PC instruction received	0.17	-0.14,0.48	0.279
*Statistically significant at $p < 0.0056$, a value chosen according to Bonferroni's rule to control false positive discovery rates across the nine contrasts run for each outcome.			

Figure 5-4 also displays the results for both research questions 1 and 2. The horizontal error bands with center filled diamonds repeat the information in table 5-8, while those with open circles repeat the information in table 5-9. The error bands represent 95 percent confidence intervals on the effect sizes. Estimated effects of the CLIO combined curricula on participation are not statistically significant with respect to either comparison.

Figure 5-4. Effect Sizes for CLIO Combined Curricula on Participation Relative to Both the Control Group and the CLIO Preschool Curricula (average of spring 2005 and spring 2006)



Summary

Effectiveness of CLIO Combined Curricula. Even Start projects that were assigned the CLIO combined curricula were not more effective at improving child language and literacy than Even Start projects that provided regular or “as is” preschool and parenting instructional services. The only child outcome that improved significantly in the projects that implemented the CLIO combined curricula was social competence (behavior in class) as rated by preschool teachers.

In addition to not reaching the threshold for statistical significance, the estimated effect sizes are small compared to other interventions focusing on improving child language and literacy outcomes. Findings from a review of meta-analyses in the field of early childhood intervention studies are summarized in table 5-10. The point of the table is not to pinpoint the average effect of preschool programs, but rather to show that in many studies, preschool programs for disadvantaged children have had positive effects on cognitive outcomes, in the range of 0.3 to 0.6 standard deviations.⁴²

Table 5-10. Summary of Findings from Selected Meta-Analyses of the Effects on Cognitive Outcomes of Early Education for Disadvantaged Children

Study authors	Effect size after 1 or 2 years of preschool	Outcome area	Number of studies
Jacob, Creps and Boulay (2004)	.33 sd	Cognitive achievement (reading, writing, spelling, math, verbal development, school readiness)	47
Gorey (2001)	.65 sd	Academic achievement (e.g., Woodcock-Johnson Revised, Iowa Test of Basic Skills, California Achievement Tests, grades)	35
Nelson and Westhues (2003)	.52 sd	Cognitive (achievement tests, grades, IQ, teacher ratings)	34
Collins (1984)	.33 sd	Cognitive outcomes	49
U.S. Department of Health and Human Services (1983)	.34 sd	Cognitive outcomes	71
Casto and White (1984)	.43 sd	IQ and other cognitive variables	26

Although there were no statistically significant impacts on child literacy outcomes, the CLIO curricula showed statistically significant positive effects on three of Even Start's hypothesized pathways to better child outcomes: (1) improved parenting skills (improved interactive reading skills and responsiveness to their child); (2) greater instructional support for literacy development in preschool classrooms (greater support for print knowledge and richer literacy resources, although no statistically significant effects on three other aspects of preschool instruction); and (3) greater instructional focus on child literacy in PE classes (greater percentage of time spent on child literacy,

⁴² The principal reason for expressing the estimated effects in the preceding tables and figures of this chapter in effect sizes was to facilitate the comparison with well-conducted meta-analyses such as those in this table.

but no statistically significant effect on the measure of PC instruction). Thus, positive effects on precursor variables such as improved parenting skills and better classroom instruction did not translate into impacts on child language or literacy skills. In addition, the CLIO curricula did not have a statistically significant effect on participation by parents and children. Low participation levels have been suggested as another factor explaining Even Start's lack of impacts.

Incremental Effectiveness of CLIO Parenting Curricula. The results showed that the CLIO parenting curricula did not add significantly to the effectiveness of the CLIO preschool curricula on children. There were no statistically significant incremental effects on the seven measures of child language and literacy skills. Nor was such an effect found on child's social competence. CLIO parenting curricula had a positive incremental effect on parent interactive reading skill (but no statistically significant incremental effect on parent responsiveness to their child).

The CLIO parenting curricula had incremental effects on two of the instructional measures. In projects with CLIO combined curricula, a greater percentage of PE class time was spent on child literacy, and a smaller percentage of PE class time on general parenting skills. There was no statistically significant effect on how time was spent in PC sessions or in preschool classes. Although the changes in PE instruction resulting in more focus on child literacy and the concomitant improvement in parent interactive reading skills were hypothesized as possible pathways to better child outcomes, this was not the case. The CLIO parenting curricula did not lead to improved child language or literacy skills, child social competence, or participation in Even Start compared to CLIO preschool curricula combined with locally selected parenting curricula.

6. SECONDARY ANALYSES OF IMPACTS OF THE CLIO CURRICULA

In addition to the primary ITT analysis presented in chapter 5, we carried out three secondary analyses of the impacts of the CLIO curricula. They were designed to answer the following questions:

- Did treatment impacts vary by year of implementation?
- Do results differ if we study children's growth in emergent literacy over the project year instead of emergent literacy at the end of the year?
- Did treatment impacts vary by home language or ethnicity of students?

These analyses are within the ITT framework and use nearly the same models as the primary impact analyses. The main reason for separating them from the analysis in chapter 5 involves the research mode. Chapter 5 is more deductive, starting out from fixed hypotheses and using analysis procedures that provide fairly strong protection against false discoveries from multiple testing. Chapter 6 is more inductive, in that we developed the hypotheses after reviewing the results of the primary analysis. These analyses are therefore more exploratory in nature, with less tight control over false discovery rates. In addition, the growth analysis could only be performed on students with long-term participation. Thus, there is potential selection bias if the treatments affected the length of family utilization of Even Start services.

Year of Implementation

Although early power calculations indicated that it was best to analyze the 2 years of data together, there might be a considerable ramp-up time for implementing the CLIO curricula. Thus, impacts might be greater in the second year, when more projects could be assumed to have had 2 years to reach full implementation. To test this hypothesis, we replicated part of the ITT analysis separately for the each of the 2 years of data. In the analysis by year, we compared projects assigned to the CLIO combined curricula with those projects assigned to the control group and looked for evidence of change from 2005 to 2006. We averaged the two experimental groups with the CLIO combined curricula in order to improve power to detect time interactions. The groups assigned to the CLIO preschool curricula were omitted from this analysis.

The results are shown in tables 6-1 through 6-3. In each, the effect estimates with asterisks are significant after Bonferroni correction. We used a test count of three for the correction given that we looked at each year separately and contrasted the 2 years.

With respect to child outcomes (table 6-1), the most striking pattern is that there is considerable evidence that the CLIO combined curricula had negative effects on four of the seven child emergent literacy outcomes in the first year of implementation. By the second year, rough parity with the control group was re-achieved. The negative effects in the first year involved English vocabulary, phonological awareness, and grammar. For social competence, there appears to be little evidence of a change in the effectiveness of the CLIO combined curricula. Although the estimated effect is only significant in 2006, the estimated change in effectiveness is not statistically significant.

With respect to parent outcomes (table 6-2), instructional outcomes, (table 6-3) and participation (table 6-3), there is little evidence of differential effects by year. The combination of changes in child impacts with stability in parenting and instructional impacts is difficult to reconcile.

Growth Instead of End-Point Status for Child and Parent Outcomes

In addition to looking at the impacts of the CLIO curricula on the status of children's emergent literacy skills and social competence, parents' parenting skills, and parents' literacy at the end of preschool, we examined impacts on the pattern of growth in these areas from fall to spring. That is, we asked if the CLIO curricula affected the rate at which children and parents obtained new skills as opposed to looking only at their spring level of skill. This analysis was conducted for the 2004-2005 project year only.⁴³ (We could not repeat this analysis in the 2005-2006 project year, because child assessments were conducted only in the spring of that year.)

⁴³ The growth analyses are based on a substantially smaller sample of children than the primary ITT analysis. See appendix L for more information.

Table 6-1. Effects of CLIO Combined Curricula on Child Outcomes by Year

Outcome	2005			2006			Change		
	Effect size	95% CI	<i>p</i> -Value	Effect size	95% CI	<i>p</i> -Value	Effect size	95% CI	<i>p</i> -Value
Expressive language: English	-0.23*	(-0.39,-0.06)	0.006	0.01	(-0.14,0.16)	>.500	0.24*	(0.06,0.42)	0.008
Expressive language: Spanish	0.04	(-0.26,0.34)	>.500	0.07	(-0.26,0.40)	>.500	0.03	(-0.45,0.50)	>.500
Receptive vocabulary	-0.17*	(-0.30,-0.03)	0.012	0.00	(-0.15,0.15)	>.500	0.17	(0.00,0.34)	0.045
Phonological awareness: Elision	-0.16	(-0.33,0.00)	0.051	0.18	(-0.02,0.38)	0.077	0.34*	(0.10,0.58)	0.006
Phonological awareness: Blending	-0.25*	(-0.42,-0.08)	0.003	0.01	(-0.17,0.19)	>.500	0.26*	(0.07,0.45)	0.007
Letter and sound recognition	0.04	(-0.11,0.19)	>.500	0.05	(-0.11,0.20)	>.500	0.01	(-0.14,0.15)	>.500
Syntax and grammar	-0.21*	(-0.38,-0.04)	0.012	0.05	(-0.09,0.19)	>.500	0.26*	(0.05,0.47)	0.011
Social competence	0.15	(-0.03,0.32)	0.099	0.30*	(0.13,0.48)	0.001	0.16	(-0.04,0.36)	0.107

* Statistically significant at the 0.0167 level (0.05/3).

Table 6-2. Effects of CLIO Combined Curricula on Parent Outcomes by Year

Outcome	2005			2006			Change		
	Effect size	95% CI	<i>p</i> -Value	Effect size	95% CI	<i>p</i> -Value	Effect size	95% CI	<i>p</i> -Value
Parent interactive reading skill	0.53*	(0.32,0.75)	0.000	0.42*	(0.20,0.65)	0.000	-0.11	(-0.39,0.17)	>.500
Parent responsiveness	0.27*	(0.11,0.42)	0.001	0.17	(-0.02,0.36)	0.064	-0.10	(-0.31,0.11)	>.500
Reading and vocabulary score	-0.07	(-0.18,0.04)	0.197	-0.02	(-0.13,0.10)	>.500	0.06	(-0.06,0.17)	>.500

* Statistically significant at the 0.0167 level (0.05/3).

Table 6-3. Effects of CLIO Combined Curricula on Instructional Outcomes and Participation by Year

Outcome	2005			2006			Change		
	Effect size	95% CI	p-Value	Effect size	95% CI	p-Value	Effect size	95% CI	p-Value
Instruction									
Support for oral language development	0.37	(-0.08,0.82)	0.099	0.21	(-0.38,0.80)	>.500	-0.16	(-1.00,0.67)	>.500
Support for print knowledge	0.74*	(0.28,1.19)	0.002	0.63*	(0.18,1.09)	0.006	-0.11	(-0.71,0.50)	>.500
Support for phonological awareness	0.34	(-0.17,0.85)	0.187	0.74	(-0.05,1.53)	0.061	0.40	(-0.46,1.26)	>.500
Support for print motivation	0.01	(-0.47,0.49)	>.500	0.75*	(0.29,1.20)	0.002	0.74	(0.02,1.45)	0.039
Literacy resources in classroom	0.41	(0.00,0.81)	0.042	0.64*	(0.17,1.10)	0.006	0.23	(-0.29,0.74)	>.500
PE time spent on child literacy	0.68*	(0.15,1.21)	0.011	1.39*	(0.92,1.86)	0.000	0.71	(0.01,1.40)	0.041
PE time spent on parenting skills	0.07	(-0.43,0.57)	>.500	-0.41	(-0.85,0.04)	0.069	-0.48	(-1.22,0.27)	0.198
PC time on interactive literacy	0.04	(-0.37,0.45)	>.500	0.38	(0.03,0.72)	0.030	0.34	(-0.26,0.93)	0.263
Participation									
Child participation	0.12	(-0.10,0.34)	0.289	0.14	(-0.09,0.37)	0.233	0.02	(-0.21,0.25)	>.500
Parent participation	0.17	(-0.16,0.51)	0.296	0.34	(-0.03,0.72)	0.067	0.17	(-0.23,0.57)	>.500

* Statistically significant at the 0.0167 level (0.05/3).

When we ran the growth analysis on the 6 English emergent literacy scales, child social competence, the 2 parenting scales, and parent reading and vocabulary skill, only 1 of the 10 overall tests⁴⁴ was statistically significant. Moreover, only 1 of 80 contrasts (10 outcomes * 8 contrasts per outcome) was significant after Bonferroni correction.⁴⁵ Both the significant overall test and the significant contrast were for parent responsiveness. Recall that effects were already found for this outcome scale in the primary ITT analysis—the average of the CLIO combined curricula was found to be better than the control curricula in promoting parent responsiveness. In the growth analysis, the comparison of the CLIO combined curricula with the control group lost statistical significance, but it was replaced by a significant contrast between the CLIO combined curricula with the average of the CLIO preschool curricula. (See appendix L for tables displaying the results of this analysis.)

Interactions of Study Group with Ethnicity and Home Language

We explored interactions of study group with both ethnicity and home language. The CLIO preschool curricula focused on English-language emergent literacy. Yet, about 50 percent of children in the sample spoke a language other than English at home (see table A-1). Additionally, the growth analysis above suggests that vocabulary growth rates differed for Hispanic children (see appendix L).

All of the ITT models included binary covariates for Hispanic background and self-reported home language other than English or a mix of English and Spanish. This secondary ITT analysis tested for interactions of the five study groups with these two covariates. The models were conducted for the six English-language child emergent literacy outcomes only. For each outcome, we ran a chi-square test with eight degrees of freedom, testing whether the impacts for any of the study groups varied as a function of ethnicity or home language.

⁴⁴ By overall test, we mean a test of whether any of the five study groups are different from each other.

⁴⁵ The same eight contrasts were run for each outcome scale as discussed in chapter 5. A Bonferroni adjustment of 9 was applied, meaning that the *p*-value for a contrast had to be smaller than $0.05/9=0.0055$ in order to be considered significant. This is the same Bonferroni adjustment as used in chapter 5.

The results showed that the interactions of study group by home language and ethnicity were not statistically significant for any of the six child outcomes. That is, the impacts did not vary significantly as a function of home language or ethnicity. (See appendix M for tables displaying the results of these analyses.)

7. ADDITIONAL ANALYSES

In this chapter, we report results from analyses that examine potential sources of variation in child and parent outcomes, using a broader set of tools than just the ITT methods of chapters 5 and 6. The emphasis is on finding sources of variation in child emergent literacy, but we include some additional results for child social competence and for parenting outcomes as well. We conducted six additional analyses:

1. One analysis examines the relationship to children's outcomes of the provision of "high-quality" ECE to children, where "high-quality" is defined as teacher instruction that aligns with scientifically based practices to support early literacy.
2. A second analysis examines the relationship to children's outcomes of the parent behaviors that are (a) shown in the primary ITT analyses to be affected by the treatment curricula and (b) that are assumed to support children's learning.
3. A third analysis uses data on classroom practices and child outcomes to try to determine whether any of the CLIO curricula would have shown stronger effects if they had been implemented with higher fidelity to the developers' ideals.
4. A fourth analysis examines the relationship to parent outcomes of participation in parenting education.
5. A fifth analysis examines the relationship to children's outcomes of participation in preschool education.
6. A sixth analysis explores how participation in preschool education might interact with the effects of curriculum on child outcomes.

The methods for all six analyses have weaknesses. The first, second, fourth, and fifth analyses above ignore study group assignments entirely and test other putative causal agents for the child impacts, specifically, measures of preschool classroom practices, parenting practices, participation in parenting education, and participation in preschool education. These four analyses are thus purely observational studies. The third and sixth analyses do involve conditioning on study group assignment, but they also involve restricting and/or reweighting the sample in ways

that permit selection biases to enter the analysis. We view all the analyses in this chapter as exploratory.

Relationship Between Child Outcomes and Preschool Instruction

In exploring the relationships between aspects of classroom instruction and child outcomes, we used a broad set of instructional variables derived from the OMLIT measures that were used in preschool classroom observations. These included the five composite instructional variables that were used as outcomes in the ITT analyses as well as dozens of additional instructional variables from the OMLIT related to the following areas:⁴⁶

- ratings of the quality and frequency of instructional support for early literacy skills (emergent writing, print knowledge, phonological awareness, oral language);
- richness of the literacy resources in the classroom (environmental print, books, writing materials, manipulatives that teach children about print or writing);
- support for ELL children (whether classrooms have at least one adult who speaks the language of ELL children in that class, integration of ELL children and their home language with English-language activities);
- distribution of classroom day across different types of activities (group or circle time, block play, creative play, dramatic play, fine motor play, block play, sensory play, or gross motor play);
- proportion of classroom time spent on teaching different early literacy skills (print knowledge, sounds, emergent writing, print motivation);

⁴⁶ The classroom instruction variables were based on the OMLIT measures: the Read Aloud Profile (use of dialogic reading techniques), the Classroom Snapshot (frequencies of activities and child groupings), the Classroom Literacy Opportunities Profile (adequacy of literacy resources in the classroom), the Classroom Literacy Instruction Profile (frequency of different types of literacy activities and knowledge afforded), and the Quality of Instruction in Language and Literacy (overall ratings of the quality of support for oral language development, writing, print knowledge, function/features of print, print motivation, phonological awareness, and practices with ELL children).

- proportion of classroom activities conducted with small group of children (five or fewer);
- number of books read aloud in classroom and overall amount of reading aloud;
- extent to which read alouds (a) include dialogic reading techniques, (b) are done with a small group of children, and (c) are conducted in the home language of ELL children;
- proportion of literacy activities selected by child versus teacher; and
- Physical environment of the classroom: adequacy of space, organization of the classroom in learning centers, level of organization of the materials in the classroom, materials arranged to allow child choice.

The statistical method we used for these explorations was to fit multi-level models to the child outcomes in terms of the instructional variables, as well as the same covariates used in the ITT analysis. The form of these models was

$$Y_{ijkt} = \theta Q_{ijkt} + \beta_j + X_{ijk}\gamma + Z_{ij}\delta + \lambda_t + a_{ij} + c_{ijk} + e_{ijkt},$$

where Q_{ijkt} is the instructional practice (i.e., the putative causal agent) experienced in year t by child k within the project within stratum j assigned to curriculum i ; θ is the coefficient for Q_{ijkt} ; and all the rest of the symbols are defined the same as in chapter 4 for the ITT analysis. The covariates used are also the same as discussed in chapter 4 for child outcomes. The exact covariates used for each outcome may be read from the corresponding table in appendix H (tables H-1 through H-8).

Each of the eight outcome variables was related to dozens of instructional variables. In order to dampen the problem of false discoveries with such a large number of tests, a Bonferroni adjustment was made within each cluster of eight tests related to the same instructional variable. In table 7-1, estimates of θ are shown that are significantly different from zero using a p -value of $0.05/8=0.00625$.

Note that these coefficients are difficult to interpret. Neither the dependent nor the independent variables in these analyses had their distributions standardized. To try to make reading of them a little more meaningful, we have added the population

standard deviations in the row and column headings. We discuss the general patterns below.

Twelve of the instructional variables (out of about five dozen) were statistically significant after adjustment for multiple comparisons to one or more of the child outcomes. Key findings from these analyses include the following:

- One of the five instructional composite variables that were used as outcomes in the ITT analysis—support for print knowledge—was significantly and positively related to children’s phonological awareness (blending) (table 7-1, first major row).
- The classroom instruction variables that were most consistently related to improved emergent literacy outcomes for children were overall ratings of the quality and frequency of activities that support print knowledge and functions/features of print (table 7-1, second and third major rows). These predictors were related more often to measures of children’s English blending skills.
- In terms of groups of variables that were tested but are not shown in the table, no statistically significant associations with emergent literacy outcomes for children were found for classroom time allocation, literacy resources in the classroom, teacher engagement and positive responsiveness to children, among many other aspects.
- No statistically significant associations were found between any of the classroom instructional variables and ratings of children’s social competence (table 7-1, Social competence column).
- Statistically significant negative associations were found between some aspects of classroom instruction (i.e., instruction supporting the development of child’s oral language, instruction supporting the development of child’s print knowledge, frequency of activities to support print knowledge, and language/literacy activities to promote print knowledge and oral language) and child results on the Spanish version of the IGDI (table 7-1).

Table 7-1. Significant Regression Coefficients (and standard errors) from Non-experimental Analyses of Relationships of ECE Classroom Instructional Variables and Parent Behaviors with Child Outcomes

Predictor (Population standard deviation in control group in 2005)	Child outcome variable							
	IGDI		PPVT (7.6)	CTOPPP			TOLD grammar 5.5)	Social compe- tence rating (9.1)
	English (8.7)	Spanish (6.9)		Elision (3.9)	Blending (5.6)	Print awareness (10.0)		
(1) Composites representing language/literacy instruction that supports development of child's:								
▪ Oral language (10.5)		-.054 (.019)						
▪ Print knowledge (8.7)		-.109 (.024)			.047 (.016)			
(2) Quality of activities intended to support:								
▪ Print knowledge (.52)				.33 (.09)	.53 (.19)			
▪ Sounds in words (.63)				.28 (.08)				
▪ Functions/features of print (.81)					.67 (.18)		.35 (.12)	
▪ Average across five types of activities (.54))					.80 (.23)			
(3) Frequency of activities intended to support:								
▪ Print knowledge (.65)		-.95 (.30)			.66 (.20)			
▪ Functions/features of print (.87)			.45 (.14)		.61 (.19)		.38 (.12)	
(4) Language/literacy activities: Knowledge afforded								
▪ Print knowledge (.052)	5.6 (1.6)	-6.5 (2.3)			4.0 (1.4)			
▪ Oral language (.070)		-8.8 (3.0)						
▪ Child-selected literacy activity (na)						-14.2 (4.9)		
(5) Dialogic reading								
▪ Strategies supporting comprehension (.30)	-1.42 (.48)							
▪ Book read in language not English (.10)	-3.41 (.95)							
(6) Parenting behavior								
▪ Reading: parent responsiveness (.97)	1.07 (.16)	.73 (.17)	.95 (.14)	.44 (.07)	.77 (.11)	1.31 (.18)	.58 (.11)	1.12 (.22)
▪ Family rules (TV programming) (.32)							1.04 (.37)	

Relationship Between Child Outcomes and Parent Teaching Behaviors

We tested the relationships between emergent literacy outcomes and parent teaching behaviors, both in the home and in the structured parent/child interaction session. The statistical procedures closely paralleled those of the prior section. The statistical method we used for these explorations was to fit multi-level models to the child outcomes in terms of the parent teaching behaviors, as well as the same covariates used in the ITT analysis. The form of these models was

$$Y_{ijkt} = \theta Q_{ijkt} + \beta_j + X_{ijk}\gamma + Z_{ij}\delta + \lambda_t + a_{ij} + c_{ijk} + e_{ijkt},$$

where Q_{ijkt} is the parent teaching behavior (i.e., the putative causal agent) experienced in year t by child k within the project within stratum j assigned to curriculum i ; θ is the coefficient for Q_{ijkt} ; and all the rest of the symbols are defined the same as in chapter 4 for the ITT analysis. The covariates used are also the same as discussed in chapter 4 for child outcomes. The exact covariates used for each outcome may be read from the corresponding table in appendix H (tables H-1 through H-8). Because there were eight child outcomes tested against each parent teaching behavior, a Bonferroni adjustment was used. Relationships were classified as statistically significant if there was an associated p -value less than $0.05/8=0.00625$.

Analyses were run with seven parent teaching behaviors as putative causal agents. Only two of the seven are represented in table 7-1. The other five analyses uncovered no significant associations. Key findings from the analysis include the following:

- Parents' responsiveness to their child was statistically significantly related to all eight child outcomes, including measures of both English and Spanish language and literacy development and social competence (table 7-1, sixth major row). This is in contrast to the finding for teachers, where a rating of the teacher's responsiveness to children was not significantly related to any of the child outcomes.
- Parent interactive reading skill was not statistically significantly associated with any measures of emergent child literacy.

- Four types of household rules were examined in relationship to child outcomes. The only significant finding⁴⁷ was that children in families that had rules about which TV programs could be watched had higher scores on a test of English grammar (table 7-1, sixth major row).

Variation in Impacts by Fidelity of Implementation of the CLIO Curricula

As noted in chapter 2, a secondary research question concerned whether child and parenting outcomes are better in projects with higher fidelity to their assigned curriculum. As discussed in chapter 3, on average, fidelity reached about 50 percent of what represented full implementation. The fact that there were no statistically significant effects of CLIO curricula on emergent child literacy in the ITT analysis of chapter 5 heightened interest in this question of potential effectiveness in a fuller implementation.

Despite this interest in it, it is a very difficult question to answer. This difficulty is due to five factors. First, the CLIO curricula are high-dimensional interventions, meaning that teachers are supposed to change many aspects of their instruction. This high dimensionality makes it difficult to express fidelity as a unidimensional measure. Yet, the concept of “higher fidelity” requires a unidimensional measure. This is particularly problematic for the experimental groups with combined curricula. Second, fidelity means something different in each of the four experimental groups, yet they must be expressed in a common scale in order to address questions about improved fidelity across experimental groups. Third, the sample sizes are small. Fidelity is a project-level concept, yet there are only 24 projects in each study group. Fourth, the projects in the control group engage in some of the same practices that are part of the experimental curricula, so the control group projects do not necessarily have zero fidelity. Fifth, fidelity is self-selected, so a strategy must be developed to guard against selection biases. Teachers who do a better job of implementing the CLIO curricula might be better teachers regardless of the curriculum they use. Without strong measurements of pre-existing teacher skill to use as covariates, there is a danger that fidelity-adjusted estimates of curriculum effects will instead be estimates of the differences between effective and ineffective teachers.

⁴⁷ The other three rules involved hours of TV, schedule for eating, and schedule for bedtime.

After considering a variety of techniques in the planning and exploratory phases of this project, we settled on one primary approach that is explained below in detail. We also describe and report on a less sophisticated graphical approach that seems to indicate that fidelity as measured in this study has little relationship to child outcomes.

Prior to discussing the approaches and findings, we briefly remind the reader of the study's measures of implementation fidelity for each curriculum. The observer-rated fidelity measures included items from the OMLIT (appendix E) in addition to items that measured curriculum-specific measures of teacher practice. This overlap of quality criteria and the fact that the OMLIT was used in classrooms in the control group as well as for classrooms in the four experimental groups allows the development of pseudo-fidelity scores for the control group, at the classroom level. For developer-rated fidelity, measured at the project level, the fidelity of the control group is zero; no contact was ever allowed between the developers and the control group.

Primary Approach

The primary approach involved changes to the ITT analysis in several ways: (1) we replaced the main effects for treatment group with a set of interactions of treatment group with fidelity, (2) we added covariates to try to reduce selection biases, (3) we narrowed the focus to the two study groups with CLIO combined curricula to increase power (since the combined curricula encompassed three Even Start components), and (4) along with the change in focus, we changed the adjustments for multiple comparisons.

More specifically, for child outcomes we fit the model

$$Y_{ijkt} = F_{ijt}\theta_i + \beta_j + X_{ijk}\gamma + Z_{ij}\delta + \lambda_t + a_{ij} + c_{ijk} + e_{ijkt},$$

where Y_{ijkt} is the raw-score⁴⁸ outcome for child k in year t in the project within stratum j assigned to curriculum i ; F_{ijkt} is either the developer-rated or the observer-rated fidelity

⁴⁸ As discussed in appendix B, we calculated both simple scores and complex IRT scores for all the child outcomes other than the IGDI assessments and the Pre-CTOPPP Print Awareness subtest. The IRT scores had been strongly optimized for the ITT analysis in ways that make them inappropriate for most other usages.

or pseudo fidelity of either the classroom or project, scaled to lie between 1 for classrooms/projects with highest ranked fidelity and 0 for classrooms/projects with lowest ranked fidelity; θ_i is the fidelity-adjusted effect of curriculum i ; and the other terms are as defined in chapter 4. The same covariates that had been used in the ITT analysis were also included in this analysis. However, two additional covariates were added: years of experience for the Even Start project director and education of lead preschool teacher, both measured concurrently with other data collection.

Appendix N includes results from a simulation study that shows that fitting this model produces the desired results when there is no selection bias and all the other model assumptions are satisfied. Under these conditions, estimates of the fidelity-adjusted effects of the curricula are unbiased, and corresponding test statistics are more highly significant than the test statistics for the ITT estimates of chapter 5.

A very similar model was fit for parent outcomes:

$$Y_{ijkt} = F_{ijt}\theta_i + \beta_j + X_{ijk}\gamma + Z_{ij}\delta + a_{ij} + c_{ijk} + e_{ijkt},$$

where Y_{ijkt} is the outcome for parent k in year t in the project within stratum j assigned to curriculum i ; F_{ijt} is the fidelity or pseudo fidelity to the parenting curriculum of the parenting classroom in which that parent sat on the observation day in that year, scaled to lie between 1 for projects with highest ranked fidelity and 0 for projects with lowest ranked fidelity; θ_i is the fidelity-adjusted effect of parenting curriculum i relative to the projects with lowest ranked fidelity or pseudo-fidelity scores; and the other terms are as defined in chapter 4. Since two of the experimental groups did not have CLIO parenting curricula, the scores used for them are pseudo-fidelity scores, formed in the same way as on the control group based on PECAP observation scores. The same covariates that had been used in the ITT analysis were also included in this analysis. However, one additional covariate was added: years of experience for the Even Start project director.

Both the child models and the parent models were fit with both developer-rated fidelity and observer-rated fidelity. Developer ratings were scored at the project level, and observer ratings were scored at the classroom level, but all were analyzed at the classroom level. For this analysis, the project-level fidelity ratings provided by the developers were assigned to each classroom in the project.

We decided to focus this fidelity-adjusted analysis on the contrasts associated with the first research question as laid out in chapter 2. We did this because we thought that fidelity might matter most in the study groups with more comprehensively reformed curricula. This narrow focus also allowed us to use a more liberal Bonferroni adjustment. We attached an asterisk denoting statistical significance to any tests with a p -value less than $0.025 = .05/2$, given that only two tests were run for each outcome and each rater. In presenting the fidelity-adjusted estimates, we also found it useful to juxtapose them with ITT estimates. These ITT estimates are slightly different from those of chapter 5 even though they were run with the same covariates as in chapter 5 because they were run on the same raw scores as the fidelity-adjusted estimates. The criterion for statistical significance on the ITT estimates is based on the same Bonferroni test count of 2 as used to classify the statistical significance of the fidelity-adjusted estimates.

Note that because of the addition of covariates and interaction terms (between the models of this section and those in chapter 5), the p -values associated with the ITT estimates can be different from those associated with fidelity-adjusted estimates. While some competing procedures for fidelity adjustment change point estimates without changing p -values, the procedure used in this report changes both.

Although the focus was on the two study groups with the CLIO combined curricula, data from all five study groups were used to fit these models. This procedure allowed us to estimate the coefficients on the covariates with greater precision.

The estimates of θ_i obtained from these models were scaled by the same population standard deviations used to scale the ITT estimates in chapter 5, as explained in chapter 4. The resulting fidelity-adjusted effect sizes are intended to be interpreted as the difference in average outcome between projects with highest ranking fidelity on the one hand and control projects and experimental projects with lowest ranking fidelity on the other hand. If projects with highest ranking fidelity to a curriculum have higher average outcomes than the norm for the experimental group assigned this curriculum, and if projects with lowest ranking fidelity have lower average outcomes than that same norm, then the fidelity-adjusted effect sizes should be larger than the corresponding ITT effect sizes.

For the most part, the fidelity-adjusted impact estimates are similar to the ITT estimates. Table 7-2 displays results based on developer ratings, and table 7-3, results based on observer ratings. The confidence intervals and *p*-values are not adjusted for multiple comparisons; the rules for declaring statistical significance, on the other hand, do reflect adjustment for multiple comparisons as discussed above. Examining the adjusted estimates based on developer fidelity ratings in table 7-2 first, we see that the set of estimates significantly different from zero was unchanged by the fidelity adjustment. Using observer-rated fidelity leads to two changes in statistical significance. Greater fidelity to the Partners for Literacy parenting curriculum is associated with higher parent responsiveness. On the other hand, greater fidelity to the Partners for Literacy preschool curriculum is associated with lower Spanish vocabulary scores among children from Spanish-speaking homes.

The hypothesis for the fidelity-adjusted analyses was that fidelity would be positively related to impacts (fidelity-adjusted impacts would be more positive or less negative). We were unable to develop a formal statistical test for this hypothesis.⁴⁹ Considering all 40 comparisons between ITT and fidelity-adjusted effect estimates visual inspection indicates that 25 have larger fidelity-adjusted estimates; 11 have larger ITT estimates; and 4 are tied. For child outcomes, 19 of 32 estimate pairs have a larger fidelity-adjusted effect than ITT effect. Among parent outcomes, six of eight estimated pairs have a larger fidelity-adjusted effect than ITT effect, and in only one instance is the fidelity-adjusted effect the smaller of the two.

⁴⁹ There are serious obstacles to constructing a test. Running a Z-test or *t*-test on each pair of alternate estimates is unsatisfactory for two reasons. First, it ignores the very strong correlation between the alternate estimates of each effect, making it far too conservative. Second, it ignores the issue of multiple comparisons, making it too liberal. We are unaware of any way of saying whether the procedure is too conservative or too liberal when both faults are considered at the same time, but we would have no confidence in the results. A more subtle procedure would be to run a significance test across the set of 40 pairs (or just the 32 child outcome pairs). This test would not be affected by the correlation within pairs. However, the significance test requires that the pairs themselves be independent, an assumption that seems violated in these data. For example, it seems reasonable that chance differences in estimated Elision effects would be correlated with chance differences in estimated Blending effects.

	Let's Begin and PALS (ECE/PE)				ECE/PE Partners for Literacy			
	ITT effect	Fidelity adjusted			ITT effect	Fidelity adjusted		
	size with	Effect			size with	Effect		
Outcome	raw	size	95% CI	p-Value	raw	size	95% CI	p-Value
	scores				scores			
Child outcomes								
Expressive language: English	-0.14	-0.09	(-0.30,0.12)	0.397	-0.09	0.05	(-0.10,0.21)	0.490
Expressive language: Spanish	0.06	0.28	(0.00,0.56)	0.051	0.04	0.15	(-0.15,0.44)	0.340
Receptive vocabulary	-0.07	-0.01	(-0.16,0.13)	0.870	-0.05	-0.02	(-0.19,0.14)	0.771
Phonological awareness:								
Elision	0.07	0.08	(-0.07,0.23)	0.283	-0.03	0.03	(-0.12,0.17)	0.694
Phonological awareness:								
Blending	-0.11	-0.17	(-0.37,0.03)	0.092	-0.10	-0.17	(-0.41,0.07)	0.172
Print knowledge	0.11	0.09	(-0.11,0.29)	0.388	-0.11	-0.27	(-0.53,0.00)	0.049
Syntax and grammar	-0.01	-0.01	(-0.15,0.12)	0.844	0.00	0.06	(-0.12,0.24)	0.510
Social competence	0.25*	0.25*	(0.10,0.41)	0.002	0.14	0.14	(-0.06,0.35)	0.178
Parent outcomes								
Parent interactive reading skill	0.45*	0.52*	(0.24,0.80)	0.001	0.50*	0.64*	(0.35,0.94)	0.000
Parent responsiveness	0.26*	0.40*	(0.20,0.60)	0.000	0.18	0.24	(0.00,0.47)	0.053
NOTE: The ITT estimates are based on raw scores rather than IRT scores to make them more comparable.								
* Statistically significant at the 0.025 level (0.05/2).								

Table 7-3. Fidelity-Adjusted Estimates of Curriculum Effects Relative to Status Quo (Observer-rated Fidelity)

Outcome	Let's Begin and PALS (ECE/PE)				ECE/PE Partners for Literacy			
	ITT effect size with raw scores	Fidelity adjusted			ITT effect size with raw scores	Fidelity adjusted		
		Effect size	95% CI	<i>p</i> -Value		Effect size	95% CI	<i>p</i> -Value
Child outcomes								
Expressive language: English	-0.14	-0.06	(-0.24,0.11)	0.478	-0.09	-0.25	(-0.55,0.04)	0.093
Expressive language: Spanish	0.06	0.03	(-0.29,0.35)	0.852	0.04	-0.44*	(-0.77,-0.11)	0.009
Receptive vocabulary	-0.07	-0.03	(-0.17,0.11)	0.689	-0.05	-0.13	(-0.39,0.13)	0.327
Phonological awareness: Elision	0.07	0.12	(-0.03,0.27)	0.118	-0.03	-0.01	(-0.27,0.25)	0.936
Phonological awareness: Blending	-0.11	-0.05	(-0.23,0.13)	0.583	-0.10	-0.01	(-0.33,0.31)	0.949
Print knowledge	0.11	0.17	(0.00,0.35)	0.055	-0.11	-0.33	(-0.61,-0.04)	0.025
Syntax and grammar	-0.01	0.02	(-0.13,0.17)	0.786	0.00	0.13	(-0.08,0.35)	0.234
Social competence	0.25*	0.30*	(0.10,0.50)	0.003	0.14	0.00	(-0.29,0.28)	0.976
Parent outcomes								
Parent interactive reading skill	0.45*	0.39*	(0.12,0.65)	0.005	0.50*	0.50*	(0.22,0.79)	0.001
Parent responsiveness	0.26*	0.32*	(0.11,0.54)	0.004	0.18	0.30*	(0.08,0.53)	0.009
NOTE: The ITT estimates are based on raw scores rather than IRT scores to make them more comparable.								
* Statistically significant at the 0.025 level (0.05/2).								

Secondary Approach

Given the results for child outcomes from the first approach that were contrary or only weakly consistent with expectations, we explored the meaning of the preschool fidelity measurements with another approach. The secondary approach is graphical rather than model-based. Because it is graphical, it does not easily lend itself to application on a large set of outcomes. Accordingly, we summarized the most important child outcomes of emergent English literacy into a single scale. The raw scores for the six assessments of different aspects of this concept were standardized with the 2005 means and standard deviations within the control group and then averaged together by year for each child with scores on all of them. This averaging was done to reduce noise and the number of graphs to be viewed. These were then averaged up to the project level by year to create a measure of project-level springtime emergent literacy. These were then graphed against annual project-level preschool fidelity for each of the five study groups. Each graph had on the order of 47 points.

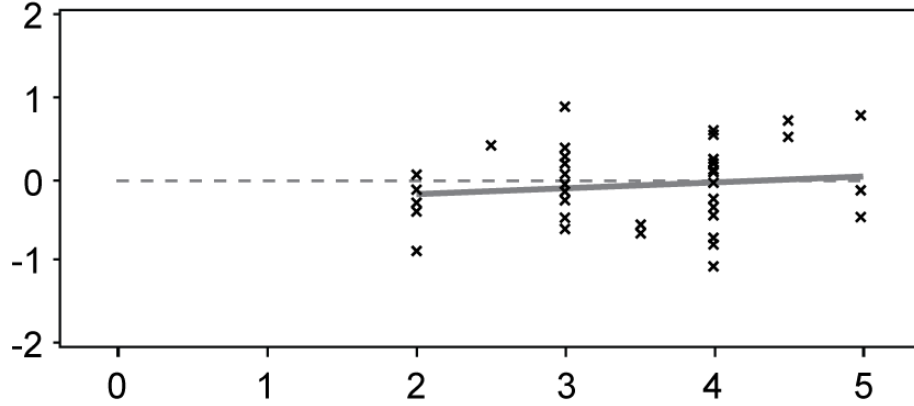
Figure 7-1 shows the results using developer-rated preschool fidelity, and figure 7-2 is the same except for using observer-rated preschool fidelity. There is a separate graph within each figure for each study group. The line shows a linear regression of the double average (project and literacy subdomain) against project-level preschool fidelity. As is apparent from the graphs, the study's measures of preschool fidelity at the project level do not appear to be related to average emergent literacy at the project level. This is true for both developer-rated fidelity and for observer-rated fidelity. Given this finding, it is then not surprising that the first approach produced inconsistent results.

Parent Participation and Parenting Outcomes

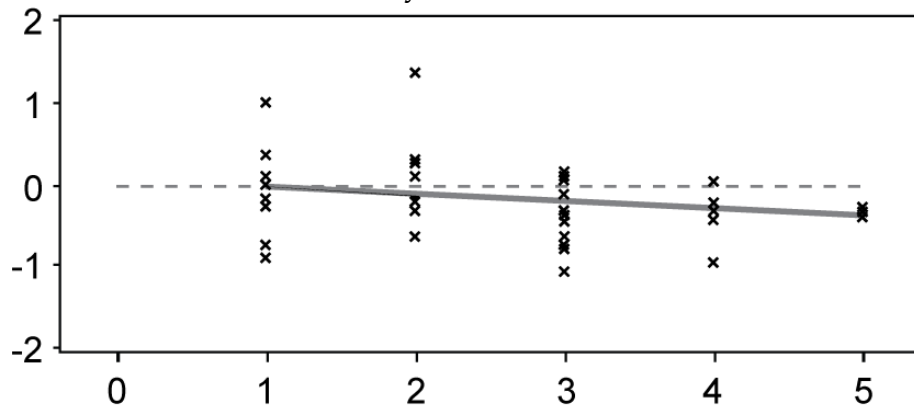
To test the relationships between amount of participation in parenting education and parenting outcomes, we fit models similar to those in the ITT analysis, replacing the term for study group with a measure of participation. Amount of participation was measured as the average number of hours per month that the parent participated in the relevant instructional services over the preceding 7 months.

Figure 7-1. Relationship Between Developer-Rated ECE Fidelity and Emergent Child English Literacy by Study Group

Let's Begin and PALS (ECE/PE)



ECE/PE Partners for Literacy



Let's Begin (ECE)

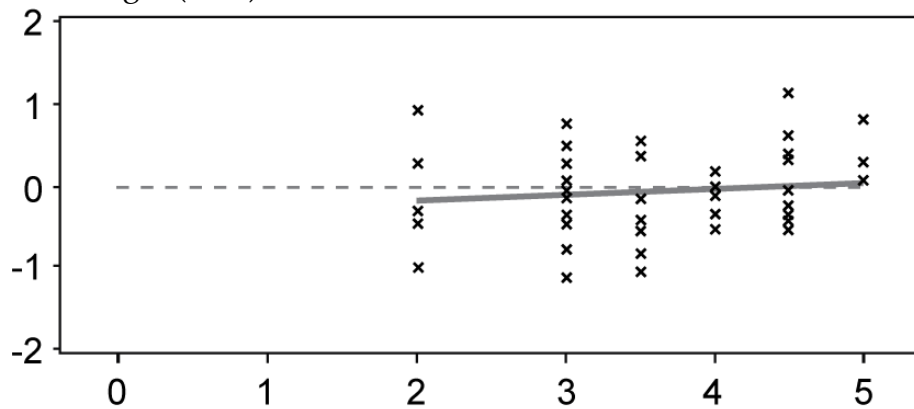
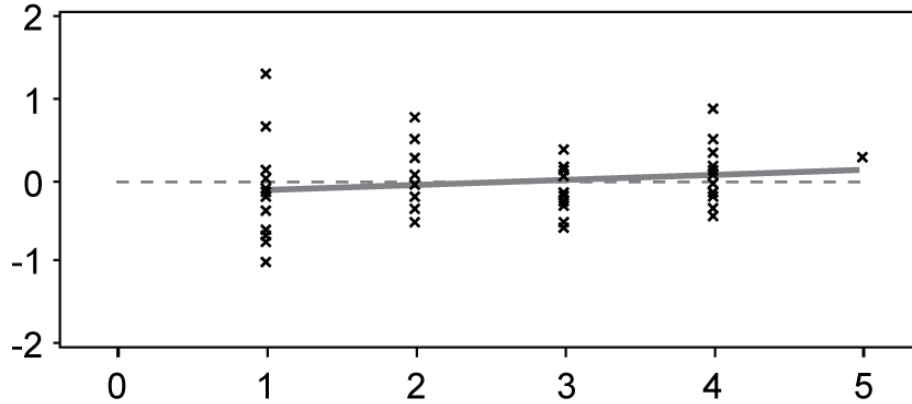
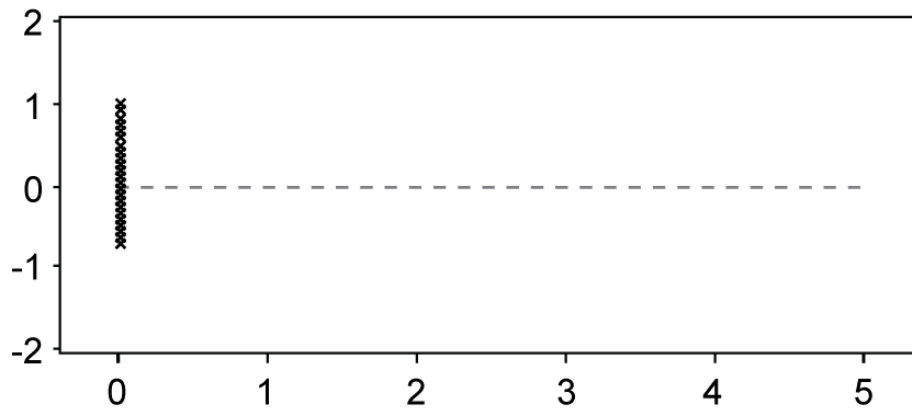


Figure 7-1. Relationship Between Developer-Rated ECE Fidelity and Emergent Child English Literacy by Study Group (continued)

Partners for Literacy (ECE)



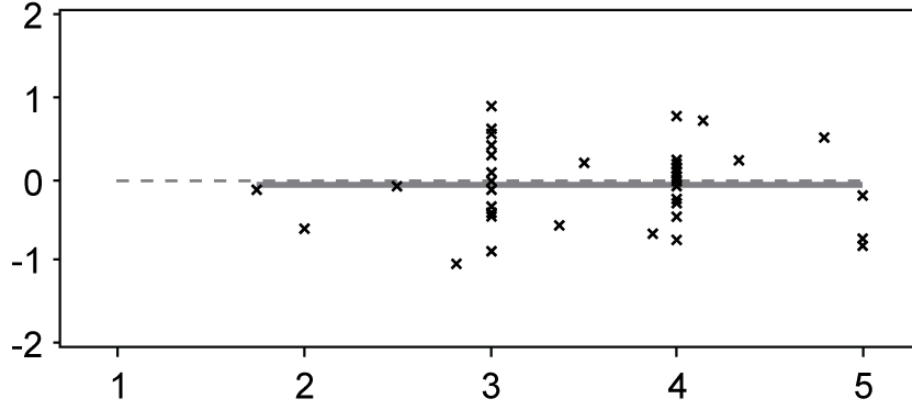
Control



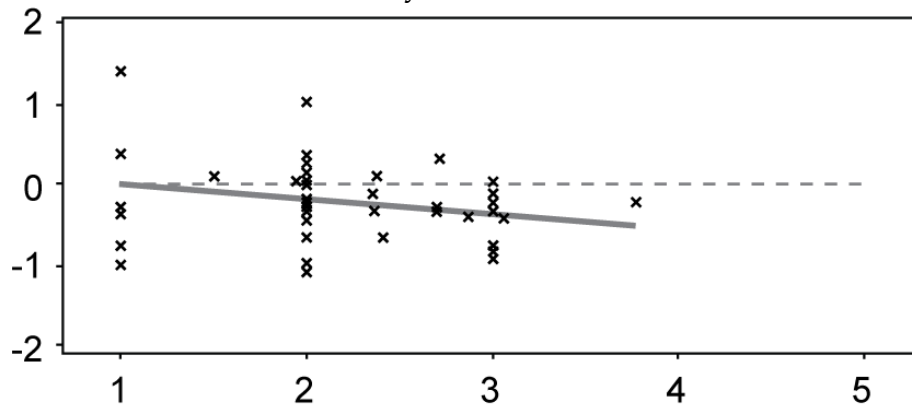
Notes: Horizontal axes in all graphs reflect developer-rated fidelity to the preschool curriculum. Vertical scales refer to standardized English emergent literacy in the children of the project. Each x represents a project.

Figure 7-2. Relationship Between Observer-Rated ECE Fidelity and Emergent Child English Literacy by Study Group

Let's Begin and PALS (ECE/PE)



ECE/PE Partners for Literacy



Let's Begin (ECE)

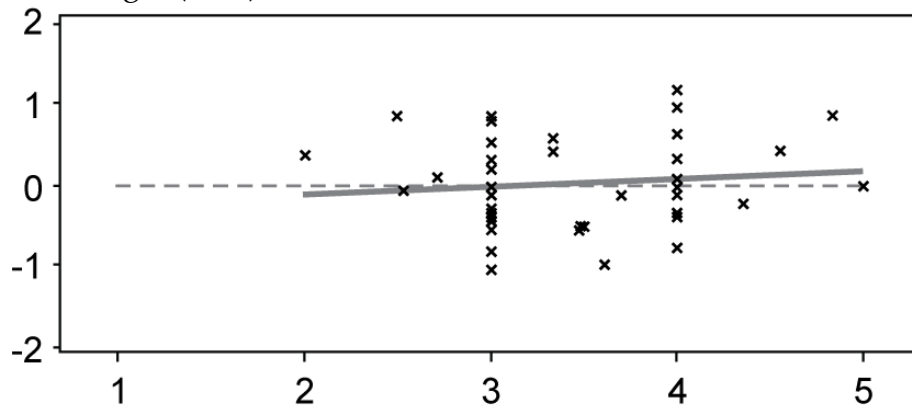
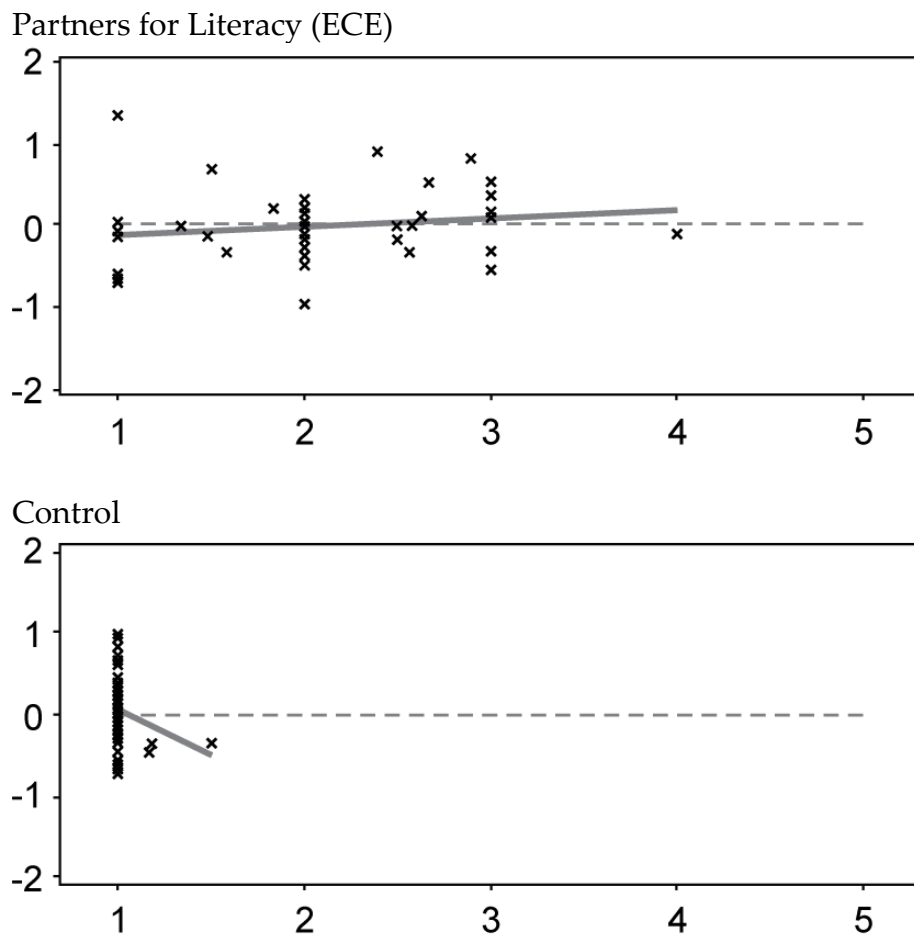


Figure 7-2. Relationship between Observer-Rated ECE Fidelity and Emergent Child English Literacy by Study Group (continued)



Notes: Horizontal axes in all graphs reflect observer-rated fidelity to the preschool curriculum. Vertical scales refer to standardized English emergent literacy in the children of the project. Each x represents a project.

The statistical procedures closely paralleled those of the sections relating teaching behaviors to child outcomes. The statistical method we used for these explorations was to fit multi-level models to the parenting outcomes in terms of the parent participation in parenting and parent-child education at the Even Start project, as well as the same covariates used in the ITT analysis. The form of these models was

$$Y_{ijkt} = \theta Q_{ijkt} + \beta_j + X_{ijk}\gamma + Z_{ij}\delta + \lambda_t + a_{ij} + c_{ijk} + e_{ijkt},$$

where Y_{ijkt} is the parent outcome; Q_{ijkt} is the parent participation (i.e., the putative causal agent) experienced in year t by parent k within the project within stratum j assigned to curriculum i ; θ is the coefficient for Q_{ijkt} ; and all the rest of the symbols are defined the same as in chapter 4 for the ITT analysis. The covariates used are also the same as discussed in chapter 4 for parent outcomes. The exact covariates used for each outcome may be read from tables H-10 and H-11. The term θ was judged to be significant if the corresponding p -value was less than 0.025, incorporating a Bonferroni adjustment with a test count of two since there are two parent outcomes.

We found a statistically significant relationship between hours of parental participation in parenting education and the quality of their responsiveness to their children. We did not obtain statistically significant evidence for a relationship between parental participation in parenting education and parent interactive reading skill. (For more information on this analysis, see appendix O.)

Participation in Preschool Education and Child Outcomes

To test the relationships between amount of participation in preschool education and child outcomes, we fit models similar to those in the ITT analysis, replacing the term for study group with a measure of participation. Amount of participation was measured as the average number of hours per month that the child participated in preschool over the preceding 7 months. (For information on missing participation data, see appendix J, p. J-10.)

The statistical procedures closely paralleled those of the sections relating teaching behaviors to child outcomes. The statistical method we used for these explorations was to fit multi-level models to the child outcomes in terms of the child participation in preschool education at the Even Start project, as well as the same covariates used in the ITT analysis. The form of these models was

$$Y_{ijkt} = \theta Q_{ijkt} + \beta_j + X_{ijk}\gamma + Z_{ij}\delta + \lambda_t + a_{ij} + c_{ijk} + e_{ijkt},$$

where Y_{ijkt} is the child outcome; Q_{ijkt} is the child participation (i.e., the putative causal agent) experienced in year t by child k within the project within stratum j assigned to

curriculum i ; θ is the coefficient for Q_{ijkt} ; and all the rest of the symbols are defined the same as in chapter 4 for the ITT analysis. The covariates used are also the same as discussed in chapter 4 for child outcomes. The exact covariates used for each outcome may be read from tables H-1 through H-8. The term θ was judged to be significant if the corresponding p -value was less than 0.00625, incorporating a Bonferroni adjustment with a test count of eight since there are eight child outcomes.

We found a statistically significant and positive relationship between participation in preschool education and child scores on five of six English emergent-literacy outcomes. A more detailed analysis suggested that the relationship appears to be confined to levels of participation above about 85 ECE hours per month, a level experienced by 26 percent of study children. For children who attended preschool for the equivalent of a school-day program (6 hours a day, 5 days a week, or around 120 hours per month over 7 months), the differential in emergent literacy is around a third of a standard deviation. This level of participation was obtained by only 3 percent of study children. (For more details on this analysis, see appendix O.)

Contrasts in Child Emergent English Literacy Across Study Groups Among Children with Substantial Participation

Under the theory that low participation may have masked curriculum effects, we investigated the association between the CLIO curricula and emergent English literacy among children with substantial participation in preschool education. In this analysis, we used a threshold of 420 hours over 7 months, which corresponds to half-day programming 5 days per week. This cutoff seemed reasonable in terms of the natural relationship between child emergent literacy and child preschool participation as measured in the baseline data. Children above this threshold on participation tended to score slightly higher on emergent literacy than children below the threshold. However, introducing the threshold led to substantial sample losses that resulted in difficulties in fitting the intended multi-level models. Only about 30 percent of enrolled children attend Even Start levels this often. The final models fit were of the form

$$Y_{ijkt} = \begin{cases} \alpha_i + X_{ijk}\gamma + Z_{ij}\delta + \lambda_t + a_{ij} + c_{ijk} + e_{ijkt} & \text{given } Q_{ijkt} > 420; \\ \text{unknown} & \text{otherwise.} \end{cases},$$

Note that is nearly same as the ITT model in chapter 4, except for the fact that the β_j term for the randomization strata has been dropped and that the model is only fit among children with high participation levels. Dropping this term resulted in better model diagnostics, but the analysis would have been more compelling had this step not proven necessary. Although there is no absolute requirement to condition on strata in a stratified design if there are many other covariates available for conditioning, there is a strong tradition of doing so.⁵⁰ The same covariates were used in the X_{ijk} and Z_{ij} terms as in the ITT analysis. The exact covariates used for each outcome may be read from tables H-1 and H-3 through H-7. Testing procedures for comparing the adjusted study group means α_i were the same as in the ITT analysis described in chapter 4.

The contrasts among the α_i have not been labeled as effects because of the possibility that initial randomization of projects may not have induced a randomization of children with high participation levels. Although the analyses of child participation as a function of study group failed to find a statistically significant effect, it is not possible to rule out the possibility of an effect. The confidence interval for child participation in table 5-8 is fairly broad, and so it is possible that high participants in study groups with the combined curricula are not comparable to the high participants in the control group.

No statistically significant differences in emergent literacy were observed among the five study groups among those children with higher participation. Table 7-4 shows the contrasts between the study groups with combined curricula and the control group. The confidence intervals are bounded by potential effect sizes of -0.40 and +0.28. Thus, no evidence was found that CLIO curricula would be better for child emergent English literacy than are the Even Start curricula even if participation were higher. (See appendix O for additional detail on methodology and the full results.)

⁵⁰ For example, in a classic two-way experimental design with treatments and blocks, it would be rare to see someone run a one-way ANOVA on the data. There is a strong tradition that calls for a two-way ANOVA to be fit on such a design. The problem with running a one-way ANOVA on a two-way design is that variances on treatment effects are over-estimated if the stratification was useful, and therefore tests can become too conservative. However, if enough other covariates are added in an ANCOVA analysis, this conservativeness may be overcome. In particular, if the stratification variable only weakly predicts the outcome but other covariates are available that strongly predict the outcome, then a reasonable argument may be made for breaking with tradition, particularly if the dataset is small, and the design matrix is ill-conditioned, as ours appears to have been.

Table 7-4. Contrasts with the Control Group of CLIO Combined Curricula on Children with High Participation Levels (Research Question 1) (Combined Spring 2005 and Spring 2006)

Child outcomes	Average of the two CLIO combined curricula vs. control		
	Effect size	95% CI	<i>p</i> -Value
Expressive language: English	0.05	(-0.13,0.22)	>.500
Receptive vocabulary	-0.09	(-0.29,0.11)	>.500
Phonological awareness: Elision	0.04	(-0.20,0.28)	>.500
Phonological awareness: Blending	-0.09	(-0.33,0.15)	>.500
Print knowledge	0.01	(-0.22,0.25)	>.500
Syntax and grammar	-0.18	(-0.40,0.05)	0.107
*Statistically significant at $p<0.0056$, a value chosen according to Bonferroni's rule to control false positive discovery rates across the nine contrasts run for each outcome.			

Bibliography

- Achenbach, T.M., Edelbrock, C., and Howell, C.T. (1987). Empirically Based Assessment of the Behavioral/Emotional Problems of 2- and 3-Year-Old Children. *Journal of Abnormal Child Psychology*, 15: 629-650.
- Alexander, K.L., and Entwisle, D.R. (1988). Achievement in the First Two Years of School: Patterns and Processes. *Monographs of the Society for Research in Child Development*, 53(2), Serial No. 218.
- Angrist, J.D., Imbens, G.W., and Rubin, D.B. (1996). Identification of Causal Effects Using Instrumental Variables (With Discussion). *Journal of the American Statistical Association*, 91: 444-472.
- Arnett, J. (1989). Caregivers in Day-Care Centers. *Journal of Applied Developmental Psychology*, 10: 541-552.
- Arnold, D.H., Epstein, J.N., Lonigan, C.J., and Whitehurst, G.J. (1994). Accelerating Language Development Through Picture Book Reading: Replication and Extension to a Videotape Training Format. *Journal of Educational Psychology*, 86(2): 235-243.
- Bankson, N.W. (1990). *Bankson Language Test—Second Edition*. Austin, TX: PRO-ED.
- Barnard, J., Frangakis, C.E., Hill, J.L., and Rubin, D.B. (2003), Principal Stratification Approach to Broken Randomized Experiments: A Case Study of School Choice Vouchers in New York (With Discussion). *Journal of the American Statistical Association*, 98: 299-323.
- Barnett, W.S. (1995). Long-Term Effects of Early Childhood Programs on Cognitive and School Outcomes. *The Future of Children*, 5(3): 25-30.
- Bonate, P.L. (2000). *Analysis of Pretest-Posttest Designs*. Boca Raton: Chapman & Hall/CRC.
- Bruner, J., and Watson, R. (1983). *Child's Talk: Learning to Use Language*. New York: W.W. Horton.

- Carrow-Woolfolk, E. (1995). *Oral and Written Language Scales: Listening Comprehension and Oral Expression*. Circle Pines, MN: American Guidance Service, Inc.
- Casto, G., and White, K. (1984). The Efficacy of Early Intervention Programs With Environmentally At-Risk Infants. *Journal of Children in Contemporary Society*, 17: 37-50.
- Collins, R.C. (1984). *Head Start: A Review of Research With Implications for Practice in Early Childhood Education*. Presentation at the Annual Meeting of the American Educational Research Association. ERIC Document Reproduction Service No. ED 245 833.
- Committee for Proprietary Medicinal Produces. (2004). Committee for Proprietary Medicinal Produces (CPMP): Points to Consider on Adjustment for Baseline Covariates. *Statistics in Medicine*, 23: 701-709.
- Cook, T.D., and Campbell, D.T. (1979). *Quasi-Experimentation: Design and Analysis Issues for Field Settings*. Boston: Houghton Mifflin.
- DeVellis, R.F. (1991). *Scale Development: Theory and Applications*. Newbury Park: Sage.
- Domitrovich, C.E., and Greenberg, M.T. (2000). The Study of Implementation: Current Findings From Effective Programs That Prevent Mental Disorder in School-Aged Children. *Journal of Educational and Psychological Consultation*, 11(2): 193-221.
- Dunn, L.M., and Dunn, L.M. (1997). *PPVT-III. Peabody Picture Vocabulary Test, Third Edition*. Circle Pines, MN: American Guidance Service.
- Dusenbury, L., Brannigan, R., Falco, M., and Hansen, W.B. (2003). A Review of Research on Fidelity of Implementation: Implications for Drug Abuse Prevention in School Settings. *Health Education Research*, 18: 237-256.
- Early Childhood Research Institute on Measuring Growth and Development. (2003). *Picture Naming Individual Growth and Development Indicator*. Minneapolis, MN: Center for Early Education and Development, University of Minnesota.
- Elliot, S.N., Gresham, F.M., Freeman, R., and McCloskey, G. (1988). Teacher and Observer Ratings of Children's Social Skills: Validation of the Social Skills Rating Scales. *Journal of Psychoeducational Assessment*, 6: 152-161.

- Fan, J., and Judkins, D. (2006). Robust Covariate Control in Cluster-Randomized Trials. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 2988-2994.
- Glass, G.V., McGaw, B., and Smith, M.L. (1981). *Meta-analysis in Social Research*. Beverly Hills: Sage.
- Gomby, D.S, Culross, P.L., and Behrman, R.E. (1999). Home Visiting: Recent Program Evaluations-Analysis and Recommendations. *The Future of Children*, 9(1): 4-26.
- Goodson, B.D., Layzer, C., Smith, W.C., and Rimdzius, T. (2004). *Observation Measures of Language and Literacy Instruction (OMLIT)*. Cambridge, MA: Abt Associates Inc.
- Goodson, B.D., Layzer, C., Smith, W.C., and Rimdzius, T. (2006). *Observation Measures of Language and Literacy Instruction (OMLIT)*. Cambridge, MA: Abt Associates Inc.
- Goodson, B.D., Layzer, J.I., St.Pierre, R.G., Bernstein, L.S., and Lopez, M. (2000). Effectiveness of a Comprehensive Five-Year Family Support Program on Low-Income Children and Their Families: Findings From the Comprehensive Child Development Program. *Early Childhood Research Quarterly*, 15(1): 5-39.
- Gorey, K. (2001). Early Childhood Education: A Meta-Analytic Affirmation of the Short- and Long-Term Benefits of Educational Opportunity. *School Psychology Quarterly*, 16(1): 9-30.
- Gorsuch, R.L. (1980). Factor Score Reliabilities and Domain Validities. *Educational and Psychological Measurement*, 40(4): 895-897.
- Gresham, F.M., and Elliott, S.N. (1990). *Social Skills Rating System*. Circle Pines, MN: American Guidance Service.
- Hamre, B.K., and Pianta, R.C. (2005). Can Instructional and Emotional Support in the First-Grade Classroom Make a Difference for Children At Risk of School Failure?" *Child Development*, 76: 949-967.
- Harms, T., Clifford, R., and Cryer, D. (1998). *Early Childhood Environment Scale-Revised Edition*. New York and London: Teachers College Press.

- Hart, B., and Risley, T. (1995). *Meaningful Differences in the Everyday Experiences of Young American Children*. Baltimore, MD: Brookes Publishing.
- Heckman, J., and Robb, R. (1986). Alternative Methods for Solving the Problem of Selection Bias in Evaluating the Impact of Treatments on Outcomes. In H. Rainer (Ed.), *Drawing Inferences from Self-Selected Samples*. Berlin: Springer-Verlag.
- Heckman, J.J. (1996). Comment. *Journal of the American Statistical Association*, 91: 459-462.
- Hoenig, J.M., and Heisey, D.M. (2001). The Abuse of Power: The Pervasive Fallacy of Power Calculations for Data Analysis. *The American Statistician*, 55: 19-24.
- Huebner, C.E. (2000). Promoting Toddlers' Language Development Through Community-Based Intervention. *Journal of Applied Developmental Psychology*, 21(5): 513-535.
- Hunter, J.E., and Schmidt, F.L. (2004). *Methods of Meta-Analysis* (2nd ed.). Thousand Oaks: Sage.
- Innocenti, M.S., Hollinger, P.D., Escobar, C.M., and White, K.R. (1993). The Cost-Effectiveness of Adding One Type of Parent Involvement to an Early Intervention Program. *Early Education and Development*, 4: 306-326.
- Jacob, R., Creps, C., and Boulay, B. (2004). *Meta-Analysis of Research and Evaluation Studies in Early Childhood Education*. Presentation at the Annual Meeting of the Association for Public Policy Analysis and Management. Abt Associates Inc., Cambridge, MA.
- Jenkins, F., Lee, H., Cheah, B., and Leytush, O. (2006). Robustness of Hierarchical Linear Modeling of Complex Survey Data When Higher Levels of Aggregation Are Left Out of the Model. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 3169-3176.
- Judkins, D., Krenzke, T., and Fan, Z. (2005). *Vector Imputation at Westat*. Presented at the Statistical Society of Canada Meetings in Saskatoon, Saskatchewan.
- Kaufman, A.S., and Kaufman, N.L. (1990). *Kaufman Brief Intelligence Test*. Circle Pines, MN: American Guidance Service, Inc.

- Kaufman, A.S., and Kaufman, N.L. (1985). *Kaufman Test of Educational Achievement*. Circle Pines, MN: American Guidance Service, Inc.
- Landry, S.H., Smith, K.E., and Swank, P.R. (2003). The Importance of Parenting During Early Childhood for School Age Development. *Developmental Neuropsychology*, 24(2-3): 559-591.
- Layzer, J.I., and Goodson, B.D. (2001). *National Evaluation of Family Support Programs: Final Report. Volume A. Meta-Analysis*. Cambridge, MA: Abt Associates Inc.
- Layzer, J.I., Layzer, C., Goodson, B.D., and Price, C. (2006). *Impacts of Preschool Language and Literacy Curricula in Subsidized Child Care: Evaluation of Project Upgrade*. Cambridge, MA: Abt Associates, Inc. Available at http://www.acf.hhs.gov/programs/opre/cc/upgrade_miami_dade/index.html
- Let's Begin with the Letter People early childhood curriculum, Abrams & Company web site. website. Available at http://www.abramsandcompany.com/lets_begin_with_letter_people.aspx.
- Lonigan, C.J., Anthony, J.L., Bloomfield, B.G., Dyer, S.M., and Samwel, C.S. (1999). Effects of Two Shared-Reading Interventions on Emergent Literacy Skills of At-Risk Preschoolers. *Journal of Early Intervention*, 22(4): 306-322.
- Lonigan, C.J., Wagner, R.K., Torgesen, J.K., and Rashotte, C.A. (2002). *The Preschool Comprehensive Test of Phonological and Print Processing*. Tallahassee, FL: Florida State University.
- Lonigan, C. (2006). Development, Assessment and Promotion of Preliteracy Skills. *Early Education and Development*, 17(1): 91-114.
- Lonigan, C.L., Schatschneider, C., and Westberg, L. (in preparation). *Results of the National Early Literacy Panel Research Synthesis: Identification of Children's Skills and Abilities Linked to Later Outcomes in Reading, Writing, and Spelling*. Washington, DC: National Institute for Literacy.
- McGrew, K.S., and Woodcock, R.W. (2001). *Technical Manual*. Woodcock-Johnson III. Itasca, IL: Riverside Publishing.
- Mislevy, R.J. (1984). Estimating Latent Distributions. *Psychometrika*, 49: 359-381.

- Mislevy, R.J. (1991). Randomization-Based Inference About Latent Variables From Complex Samples. *Psychometrika*, 56: 177-196.
- Missall, K.N., and McConnell, S.R. (2004, April). *Technical Report: Psychometric Characteristics of Individual Growth and Development Indicators: Picture Naming, Rhyming, and Alliteration*. University of Minnesota: Center for Early Education and Development.
- National Research Council (2001). *Eager to Learn: Educating Our Preschoolers*. Committee on Early Childhood Pedagogy. B.T. Bowman, M.S. Donovan and M.S. Burns (Eds.). Commission on Behavioral and Social Sciences and Education. Washington, DC: National Academy Press.
- Nelson, G., and Westhues, A. (2003). A Meta-Analysis of Longitudinal Research on Preschool Prevention Programs for Children. *Prevention & Treatment*, 6: 31.
- Newcomer, P.L., and Hammill, D.D. (1997a). *Test of Language Development—Primary: Third Edition (TOLD-P:3)*. Austin, TX: PRO-ED, Inc.
- Newcomer, P.L., and Hammill, D.D. (1997b). *Examiner's Manual. Test of Language Development, Primary (Third Edition)*. Austin, TX: PRO-ED, Inc.
- Partners for Literacy early childhood and parenting curriculum, Frank Porter Graham Child Development Center website. Available at <http://www.fpg.unc.edu/~literacy>.
- Piesse, A., Judkins, D., and Fan, Z. (2005). Item Imputation Made Easy. *Proceedings of the Section on Survey Research Methods of the American Statistical Association*, 3476-3479.
- Pianta, R.C. (2003). *Standardized Classroom Observations From Pre-K to Third Grade: A Mechanism for Improving Quality Classroom Experiences During the P-3 Years*. University of Virginia (pp. 3-4).
- Play and Learning Strategies (PALS) parenting curriculum, University of Texas Health Sciences Center, Children's Learning Institute website. Available at <http://www.childrenslearninginstitute.org/our-programs/program-overview/PALS/default.html>.

- Ramey, C.T., and Campbell, F.A. (1988). Preventive Education for High-Risk Children: Cognitive Consequences of the Carolina Abecedarian Project. *American Journal of Mental Deficiency, 88*(5): 515-523.
- Ramey, S.L., and Ramey, C.T. (1992). Early Educational Intervention With Disadvantaged Children—To What Effect? *Applied and Preventive Psychology, 1*: 130-140.
- Ramey, C.T., Bryant, D.M., Wasik, B.H., Sparling, J.J., Fendt, K.H., and LaVange, L.M. (1992). Infant Health and Development Program for Low Birth Weight, Premature Infants: Program Elements, Family Participation, and Child Intelligence. *Pediatrics, 3*: 454-465.
- Ramey, C.T., Collier, A.M., Sparling, J.J., Loda, F.A., Campbell, M.D., Ingram, M.D., and Finkelstein, N.W. (1976). The Carolina Abecedarian Project: A Longitudinal and Multidisciplinary Approach to the Prevention of Developmental Retardation. In T.D. Tjossem (Ed.), *Intervention Strategies for High Risk Infants and Young Children* (pp. 629-665). Baltimore, MD: University Park Press.
- Raudenbush, W.W., Bryk, A.S., Cheong, Y.K., and Congdon, R.T., Jr. (2004). *HLM6: Hierarchical Linear and Nonlinear Modeling*. Lincolnwood, IL: Scientific Software International.
- Rosenthal, R. (1994). Parametric Measures of Effect Size. In H. Cooper and L.V. Hedges (Eds.), *The Handbook of Research Synthesis*. New York: Russell Sage Foundation.
- Rubin, D.B. (1987). *Multiple Imputation for Nonresponse in Surveys*. New York: John Wiley.
- SAS Institute Inc. (2007). *SAS OnlineDoc® 9.1.3*. Carey, NC: SAS Institute, Inc.
- Snow, C.E., Barnes, W.S., Chandler, J., Hemphill L., and Goodman, I.F. (1991). *Unfulfilled Expectations: Home and School Influences on Literacy*. Cambridge, MA: Harvard University Press.
- Snow, C.E., Burns, M.S., and Griffin, P. (Eds.). (1998). *Preventing Reading Difficulties in Young Children*. Washington, DC: National Academy Press.

- Sorongon, A.G. (2007, March). *Predictive Validity of Measures of Preschool Children's Cognitive and Social Skills to Kindergarten School Performance Indicators*. Presented at the Society for Research in Child Development Biennial Conference in Boston, MA.
- St.Pierre, R.G., Ricciuti A.E., and Rimdzius T. (2005). Effects of a Family Literacy Program on Low-Literate Children and Their Parents: Findings From an Evaluation of the Even Start Family Literacy Program. *Developmental Psychology*, 41(6): 953-970.
- St.Pierre, R.G., Ricciuti, A.E., Tao, F., Creps, C., Swartz, J., Lee, W., Parsad, A., and Rimdzius, T. (2003). *Third National Even Start Evaluation: Program Impacts and Implications for Improvement*. Cambridge, MA: Abt Associates Inc. for the U.S. Department of Education, Planning and Evaluation Service, contract EA97049001.
- St.Pierre, R.G., Layzer, J.I., Goodson, B.D., and Bernstein, L.S. (June 1997). *National Impact Evaluation of the Comprehensive Child Development Program: Final Report*. Cambridge, MA: Abt Associates Inc. for the U.S. Department of Health and Human Services, Administration on Children, Youth and Families, contract 105-90-1900.
- St.Pierre, R.G., Swartz, J.P., Gamse, B., Murray, S., Deck D., and Nickel P. (January, 1995). *National Evaluation of the Even Start Family Literacy Program: Final Report*. Cambridge, MA: Abt Associates Inc. for the U.S. Department of Education, Planning and Evaluation Service, contract LC90062001.
- Tao, F., Gamse, B., and Tarr, H. (1998). *Second National Evaluation of the Even Start Family Literacy Program: Final Report*. Arlington, VA: Fu Associates Ltd.
- Tobler, N.S., and Stratton, H.H. (1997). Effectiveness of School-Based Drug Prevention Programs: A Meta-Analysis of the Research. *Journal of Primary Prevention*, 18(1): 71-128.
- U.S. Department of Health and Human Services, Administration on Children, Youth and Families (1983). *The Effects of the Head Start Program on Children's Cognitive Development: Preliminary Report of the Head Start Evaluation, Synthesis and Utilization Project*. Washington, DC: Author. Available at http://eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/2e/9f/60.pdf

- U.S. Department of Education (2003). *Guidance for the William F. Goodling Even Start Family Literacy Programs*. Washington, DC: Office of Elementary and Secondary Education.
- U.S. Department of Health and Human Services, Administration for Children and Families (2003). *Head Start Faces 2000. A Whole-Child Perspective on Program Performance. Fourth Progress Report*. Washington, DC: Author.
- U.S. Department of Health and Human Services, Administration for Children and Families. (2006). *FACES 2003 Research Brief*. Washington, DC: Author.
- Walker, S. (2004). Teacher Reports of Social Behavior and Peer Acceptance in Early Childhood: Sex and Social Status Differences. *Child Study Journal* 34(1): 13-28.
- Wasik, B.H., Ramey, C.T., Bryant, D.M., and Sparling, J.J. (1990). A Longitudinal Study of Two Early Intervention Strategies: Project CARE. *Child Development*, 61(6): 1682-1696.
- Wechsler, D. (1992). *Wechsler Individual Achievement Test*. San Antonio, TX: The Psychological Corporation.
- Wechsler, D. (1991). *Wechsler Intelligence Scale for Children—Third Edition*. San Antonio, TX: The Psychological Corporation.
- Weikart, D.P., Bond, J.T., and McNeil, J.T (1978). *The Ypsilanti Perry Preschool Project: Preschool years and Longitudinal Results Through Fourth Grade*. Ypsilanti, MI: High/Scope.
- White, K.R., Taylor, M.J., and Moss, V.D. (1992). Does Research Support Claims About the Benefits of Involving Parents in Early Intervention Programs? *Review of Educational Research*, 62(1): 91-125.
- Whitehurst, G.J., Arnold, D.S., Epstein, J.N., Angell, A.L., Smith, M., and Fischel, J.E. (1994). A Picture Book Reading Intervention in Day Care and Home for Children From Low-Income Families. *Developmental Psychology*, 30(5): 679-689.
- Whitehurst, G.J., Falco, F.L., Lonigan, C.J., Fischel, J.E., DeBaryshe, B.D., Valdez-Menchaca, M.C., and Caulfield, M. (1988). Accelerating Language Development Through Picture Book Reading. *Developmental Psychology*, 24(4): 552-559.

- Whitehurst, G.J., and Lonigan, C.J. 1998). Child Development and Emergent Literacy. *Child Development*, 69(3): 848-872.
- Whitehurst, G.J., and Lonigan, C.J. (2001). Emergent Literacy: Development From Prereaders to Readers. In S.B. Neuman and D.K. Dickinson (Eds.), *Handbook of Early Literacy Research*, (pp. 11-29). New York: Guilford Press.
- Williams, K.T., and Wang, J.J. (1997). *Technical References to the Peabody Picture Vocabulary Test—Third Edition (PPVT-III)*. Circle Pines, MN: American Guidance Service, Inc.
- Woodcock, R.W., McGrew, K.S., and Mather, N. (2001). *Woodcock-Johnson III Tests of Achievement*. Itasca, IL: Riverside Publishing.
- Yoshikawa, H. (1995). Long-Term Effects of Early Childhood Programs on Social Outcomes and Delinquency. *The Future of Children*, 5(3): 51-75.
- Zill, N. (1990). *Behavior Problems Index based on parent report (memorandum)*. Washington, DC: Child Trends.
- Zimmerman, I.L., Steiner, V.G., and Pond, R.E. (1992). *Preschool Language Scales—Third Edition*. San Antonio, TX: The Psychological Corporation.

APPENDIX A DESCRIPTIVE COMPARISON OF THE FIVE GROUPS

Table A-1. Percentage Distribution of CLIO Children, by Study Group and Child and Family Demographics: Spring 2004, Spring 2005, and Spring 2006

Demographics	Total	Study group					Balance test <i>p</i> -value
		Let's Begin with the Letter People (ECE)	Play & Learning Strategies (ECE/PE)	Partners for Literacy (ECE)	Partners for Literacy (ECE/PE)	Control	
Child's age							
Spring 2004							
3–3.5 years	7.34	6.91	4.88	10.34	7.55	6.62	0.1598
3.5–4 years	19.08	17.45	16.03	21.00	21.94	18.75	0.3982
4–4.5 years	20.96	20.36	24.04	23.51	17.63	18.75	0.8073
4.5–5 years	28.02	29.82	27.18	25.08	27.34	31.25	0.4968
5 years or older	24.60	25.45	27.87	20.06	25.54	24.63	0.2177
Spring 2005							
3–3.5 years	8.18	7.69	3.88	8.61	11.24	9.18	0.0270
3.5–4 years	19.40	20.74	23.62	17.80	16.57	18.71	0.0317
4–4.5 years	22.45	23.08	19.74	22.85	21.60	25.17	0.3020
4.5–5 years	26.57	27.09	25.89	26.41	27.22	26.19	0.8890
5 years or older	23.40	21.40	26.89	24.33	23.37	20.75	0.5446
Spring 2006							
3–3.5 years	8.72	9.22	6.85	7.14	11.46	8.96	0.5842
3.5–4 years	18.92	16.04	23.79	20.00	18.58	16.85	0.3014
4–4.5 years	22.39	23.21	19.76	23.93	22.53	22.22	0.6890
4.5–5 years	27.49	31.40	30.24	24.64	23.72	29.39	0.3593
5 years or older	22.03	20.14	19.35	24.29	23.72	22.58	0.4615
Average age in months							
Spring 2004	54	54	55	52	54	54	0.1591
Spring 2005	53	53	54	54	53	53	0.6722
Spring 2006	53	53	53	53	53	53	0.6318
Child's sex							
Spring 2004							
Male	50.87	54.95	47.39	50.78	50.00	51.47	0.5035
Female	49.13	45.05	52.61	49.22	50.00	48.53	
Spring 2005							
Male	51.36	56.19	53.40	53.12	46.75	47.62	0.1263
Female	48.64	43.81	46.60	46.88	53.25	52.38	

Table A-1. Percentage Distribution of CLIO Children, by Study Group and Child and Family Demographics: Spring 2004, Spring 2005, and Spring 2006 (continued)

		Study group					Balance test <i>p</i> - value
		Let's Begin with the Letter People (ECE)	Play & Learning Strategies (ECE/PE)	Partners for Literacy (ECE)	Partners for Literacy (ECE/PE)	Control	
Demographics	Total						
Child's sex (cont'd)							
Spring 2006							
Male	49.74	47.10	53.63	45.00	53.36	50.54	0.2395
Female	50.26	52.90	46.37	55.00	46.64	49.46	
Child's race/ethnicity ¹							
Spring 2004							
White	22.49	32.60	26.95	19.29	17.71	15.83	0.5379
Black	11.17	9.52	4.26	21.22	8.49	11.20	0.6615
Hispanic	57.23	53.48	59.57	48.87	69.00	56.37	0.8306
Asian	2.72	1.47	1.06	4.50	2.21	4.25	0.4461
Other	6.38	2.93	8.16	6.11	2.58	12.36	0.2736
Spring 2005							
White	20.36	34.78	23.95	14.24	15.98	13.95	0.1479
Black	10.97	7.36	4.85	21.07	9.47	11.22	0.9539
Hispanic	61.19	55.52	64.08	56.38	68.34	61.22	0.2338
Asian	2.41	1.00	0.97	3.56	2.07	4.42	0.2379
Other	5.07	1.34	6.15	4.75	4.14	9.18	0.3422
Spring 2006							
White	22.10	31.06	25.00	13.93	22.92	17.56	0.1630
Black	10.42	9.90	5.24	15.71	11.86	8.96	0.4033
Hispanic	59.50	56.66	63.31	57.50	58.10	62.37	0.0893
Asian	2.51	0.68	1.21	7.50	2.37	0.72	0.2906
Other	5.47	1.71	5.24	5.36	4.74	10.39	0.2666
Child has a disability or special need ²							
Spring 2004							
Yes	14.38	14.91	11.85	13.97	12.59	18.82	0.9442
No	85.62	85.09	88.15	86.03	87.41	81.18	
Spring 2005							
Yes	12.68	14.05	10.36	9.79	15.38	13.95	0.7608
No	87.32	85.95	89.64	90.21	84.62	86.05	
Spring 2006							
Yes	11.09	12.97	10.08	8.21	13.04	11.11	0.3585
No	88.91	87.03	89.92	91.79	86.96	88.89	

Table A-1. Percentage Distribution of CLIO Children, by Study Group and Child and Family Demographics: Spring 2004, Spring 2005, and Spring 2006 (continued)

		Study group					Balance test <i>p</i> - value
		Let's Begin with the Letter People (ECE)	Play & Learning Strategies (ECE/PE)	Partners for Literacy (ECE)	Partners for Literacy (ECE/PE)	Control	
Demographics	Total						
Mother's educational attainment							
Spring 2004							
9 th grade or less	36.32	33.82	34.39	37.14	43.68	32.34	0.6815
Grade 10-12	22.92	19.85	20.70	25.71	23.47	24.54	0.4659
Special education diploma ³	2.54	1.47	2.11	2.54	3.61	2.97	0.6624
High school diploma/GED/or higher	38.22	44.85	42.81	34.60	29.24	40.15	0.6268
Spring 2005							
9 th grade or less	37.86	39.80	35.92	36.20	40.24	37.07	0.9443
Grade 10-12	19.21	19.73	16.18	17.51	19.82	23.13	0.3334
Special education diploma ³	2.47	0.67	3.88	2.67	2.96	2.04	0.3115
High school diploma/GED/or higher	40.46	39.80	44.01	43.62	36.98	37.76	0.9986
Spring 2006							
9 th grade or less	36.29	36.86	38.31	33.21	37.15	36.20	0.5609
Grade 10-12	19.96	14.33	20.16	23.21	21.34	21.15	0.8099
Special education diploma ³	4.07	4.78	5.65	2.50	5.93	1.79	0.9419
High school diploma/GED/or higher	39.69	44.03	35.89	41.07	35.57	40.86	0.7747
Language spoken at home							
Spring 2004							
English	39.57	46.55	41.11	41.64	27.80	40.44	0.1553
Spanish	46.64	38.55	49.83	41.32	57.40	46.69	0.2532
English and Spanish	8.19	10.55	6.62	5.36	10.83	8.09	0.9109
Other	5.60	4.36	2.44	11.67	3.97	4.78	0.4570
Spring 2005							
English	38.05	43.48	38.51	39.76	33.14	35.71	0.2383
Spanish	47.37	42.14	49.51	40.65	55.62	48.64	0.8939
English and Spanish	9.13	9.70	10.03	9.79	7.10	9.18	0.7233
Other	5.45	4.68	1.94	9.79	4.14	6.46	0.2606
Spring 2006							
English	38.06	43.00	38.31	35.00	37.94	35.84	0.1840
Spanish	43.75	42.32	43.15	43.57	45.45	44.44	0.3211
English and Spanish	12.71	10.92	15.32	9.29	10.28	17.92	0.5005
Other	5.47	3.75	3.23	12.14	6.32	1.79	0.4510

Table A-1. Percentage Distribution of CLIO Children, by Study Group and Child and Family Demographics: Spring 2004, Spring 2005, and Spring 2006 (continued)

Demographics	Total	Study group					Balance test <i>p</i> -value
		Let's Begin with the Letter People (ECE)	Play & Learning Strategies (ECE/PE)	Partners for Literacy (ECE)	Partners for Literacy (ECE/PE)	Control	
Monthly household income ⁴							
Spring 2004							
\$500 or less	12.76	10.67	7.31	18.40	14.34	12.45	0.2658
\$501 to \$1,000	22.04	16.21	26.15	21.18	27.09	19.50	0.1254
\$1,001 to \$1,500	24.98	22.53	29.62	21.53	27.49	24.07	0.6949
More than \$1,500	40.22	50.59	36.92	38.89	31.08	43.98	0.2137
Spring 2005							
\$500 or less	12.11	8.70	15.21	12.17	13.91	10.20	0.4182
\$501 to \$1,000	19.53	19.40	19.42	19.29	21.30	18.03	0.6544
\$1,001 to \$1,500	26.00	24.41	24.27	26.41	27.81	26.87	0.7315
More than \$1,500	42.36	47.49	41.10	42.14	36.98	44.90	0.2938
Spring 2006							
\$500 or less	12.49	9.22	15.32	11.43	16.21	11.11	0.2018
\$501 to \$1,000	18.18	21.16	19.76	16.43	18.58	15.05	0.3507
\$1,001 to \$1,500	26.24	29.35	23.39	22.50	24.11	31.18	0.5171
More than \$1,500	43.09	40.27	41.53	49.64	41.11	42.65	0.6093

¹Black includes African American; Hispanic includes Latino; Other includes American Indian/Alaska Native, Pacific Islander/ Native Hawaiian, and more than one race. Race categories exclude Hispanic origin unless specified.

²For example, physical, emotional, language, hearing, or learning difficulties identified by a doctor or other health or education professional.

³The definition and requirements for a special education diploma vary by state. For example, in some states, students receive such a diploma if they meet all graduation requirements but do not pass a high school exit test. In others, it is essentially given to all students when they reach a certain age and have not graduated.

⁴Includes income before taxes and other deductions in the past month for all members of the household.

NOTES: ECE = Early childhood education curriculum. PE = Parenting education curriculum. Detail may not sum to totals because of rounding. The table is limited to children with a parent interview in a given year. A stratum-adjusted Kruskal-Wallis test run on project-level averages was used to test baseline balance across study groups. No weights were used at the project level.

Multinomial variables like race were transformed into a series of binary recodes, each of which was tested separately.

SOURCE: U. S. Department of Education, Even Start Classroom Literacy Interventions and Outcomes Study, "Parent Interview," Spring 2004, Spring 2005, Spring 2006.

Table A-2. Percentage of CLIO Parents, by Study Group and Parent-Reported Literacy Levels: Spring 2004

		Study group					Balance test <i>p</i> -value
		Let's Begin with the Letter People (ECE)	Play & Learning Strategies (ECE/PE)	Partners for Literacy (ECE)	Partners for Literacy (ECE/PE)	Control	
Literacy level	Total						
Literacy in native language ¹							
Speaks native language well or very well	96.42	97.74	96.88	93.75	96.09	98.00	0.9069
Reads native language well or very well	91.94	92.48	90.63	88.75	93.85	94.00	0.7817
Writes native language well or very well	89.64	90.23	88.75	85.63	91.62	92.00	0.7430
English literacy							
Understands English well or very well	57.03	58.27	60.45	64.71	44.05	56.35	0.1400
Speaks English well or very well	50.95	56.30	50.75	58.82	39.29	48.41	0.0390
Reads English well or very well	56.20	61.02	55.22	62.63	46.83	54.37	0.4932

¹Limited to parents with English as a second language. English is a second language for 59.41 percent of CLIO parents.

NOTES: ECE = Early childhood education curriculum. PE = Parenting education curriculum. A stratum-adjusted Kruskal-Wallis test run on project-level averages was used to test baseline balance across study groups. No weights were used at the project level. Multinomial variables like race were transformed into a series of binary recodes, each of which was tested separately.

SOURCE: U.S. Department of Education, Even Start Classroom Literacy Interventions and Outcomes Study, "Parent Interview," Spring 2004.

Table A-3. Average Language, Literacy, Social Competence, and Participation Outcomes for CLIO Children and Parents, by Study Group: Spring 2004

Outcome ¹	Total	Study group					Balance test <i>p</i> -value
		Let's Begin with the Letter People (ECE)	Play & Learning Strategies (ECE/PE)	Partners for Literacy (ECE)	Partners for Literacy (ECE/PE)	Control	
Children							
Language and literacy ¹							
Expressive language: English	15	16	16	15	12	16	0.6787
Expressive language: Spanish	17	16	17	15	17	17	0.7069
Receptive vocabulary	253	263	255	254	238	256	0.1458
Phonological awareness:							
Elision ²	262	271	267	259	251	261	0.0226
Phonological awareness:							
Blending ³	272	272	272	271	280	265	>0.500
Social competence	252	250	252	252	259	246	>0.500
Monthly hours of Even Start participation ⁴	42	52	44	40	40	37	0.3248
Parents							
Reading and vocabulary score ⁵	250	261	254	251	237	244	0.3923
Monthly hours of Even Start participation ³	10	10	12	10	11	8	0.4839

¹The child receptive vocabulary, Elision, and Blending assessments and the parents reading and vocabulary score were linearly transformed to have a mean of 250 and a standard deviation of 50.

²The Elision subtest measures the child's ability to recognize word parts, such as components of compound words, syllables, and phonemes.

³The Blending subtest measures the child's ability to combine word parts, such as components of compound words, syllables, and phonemes.

⁴Monthly hours of participation is the sum of the hours a child or parent participated in Even Start across a 7-month period divided by seven, regardless of whether the participation was all within a single month or spread more evenly across months. Although Even Start is intended to be a year-round program, projects may either not provide services in the summer or the services may differ greatly from those provided during the school year. As a result, only participation counts from September 2003 through March 2004 are included in the calculations.

⁵The reading and vocabulary score is based on four subtests that measure vocabulary, basic reading, phonics, and comprehension.

NOTES: ECE = Early childhood education curriculum. PE = Parenting education curriculum. To test for baseline balance for IRT scores, HLM was used with a two-level setup (project and child) and no covariates other than strata. For all other types of variables, a stratum-adjusted Kruskal-Wallis test run on project-level averages was used to test baseline balance across study groups. No weights were used at the project level.

SOURCE: U.S. Department of Education, Even Start Classroom Literacy Interventions and Outcomes Study, "Child Assessment Battery, Teacher Rating Forms, Parent Assessment Battery, and Instructional Services Participation Form (ISPF)," Spring 2004.

Table A-4. Average Size, Amount of Services, and Early Childhood Education Instructional Outcomes for CLIO Projects, by Study Group: Spring 2004

Size, services, and outcomes	Total	Study group					Balance test <i>p</i> -value
		Let's Begin with the Letter People (ECE)	Play & Learning Strategies (ECE/PE)	Partners for Literacy (ECE)	Partners for Literacy (ECE/PE)	Control	
Project size ¹							
Number of families	32	27	26	36	33	36	0.0086
Number of preschool children	21	19	21	23	23	20	0.5370
Hours of early childhood education preschool services offered per week ²	21	25	17	20	21	20	0.2673
Early childhood education instructional outcomes ³							
Support for oral language development	46.29	45.11	48.17	43.98	44.39	49.70	0.2832
Support for print knowledge	47.38	43.58	47.99	49.02	46.54	49.56	0.2033
Support for phonological awareness	47.47	46.13	50.97	47.16	43.79	49.10	0.1049
Support for print motivation	47.75	48.75	48.57	45.07	46.64	49.73	0.7792
Literacy resources in classroom	45.91	46.99	45.49	43.40	43.83	49.79	0.3056

¹Includes families and preschool children in home-based and center-based services.

²Limited to center-based instructional services.

³Constructs were derived from the OMLIT measures to correspond to the four key components of emergent literacy. The study group means from 2004 and 2005 are interpreted relative to the 2004 control group mean. Constructs were scaled to have a mean of 50 and standard deviation of 10. If, for example, a study group mean of 52 can be interpreted as 2/10 standard deviation units higher than the 2004 control group mean.

NOTES: Preschool children are 3- to 5-year-old children. ECE = Early childhood education curriculum. PE = Parenting education curriculum. A stratum-adjusted Kruskal-Wallis test run on project-level averages was used to test baseline balance across study groups. No weights were used at the project level.

SOURCE: U.S. Department of Education, Even Start Classroom Literacy Interventions and Outcomes Study, "Project Director Survey and Observation Measures of Language and Literacy Instruction (OMLIT)," Spring 2004.

Table A-5. Percentage Distribution of Even Start Staff, by Study Group and Selected Demographics: Spring 2004

		Study group					Balance test <i>p</i> -value
		Let's Begin with the Letter People (ECE)	Play & Learning Strategies (ECE/PE)	Partners for Literacy (ECE)	Partners for Literacy (ECE/PE)	Control	
Demographics	Total						
Lead early childhood education teacher							
Educational attainment							
Less than a bachelor's degree	39.05	34.48	37.50	35.29	50.00	37.50	0.4487
Bachelor's degree	45.56	41.38	46.88	52.94	38.24	47.50	0.8928
Graduate degree	15.38	24.14	15.63	11.76	11.76	15.00	0.2476
License attainment							
Has license ¹	86.90	79.31	100.00	79.41	91.18	84.62	0.0912
Does not have license	13.10	20.69	0.00	20.59	8.82	15.38	
Bilingual status							
Is bilingual	28.99	24.14	25.00	35.29	35.29	25.00	0.2662
Is not bilingual	71.01	75.86	75.00	64.71	64.71	75.00	
Early childhood education aide							
Educational attainment							
High school diploma or GED	78.05	83.33	76.19	77.78	84.62	68.42	0.9265
Associate's or bachelor's degree	21.95	16.67	23.81	22.22	15.38	31.58	
License attainment							
Has license ¹	33.12	32.26	30.00	38.24	28.95	35.29	0.9245
Does not have license	66.88	67.74	70.00	61.76	71.05	64.71	
Bilingual status							
Is bilingual	38.92	25.81	42.86	37.84	45.00	42.11	0.7205
Is not bilingual	61.08	74.19	57.14	62.16	55.00	57.89	
Lead parenting education teacher							
Educational attainment							
Less than graduate degree	67.62	65.00	75.00	50.00	72.73	69.57	0.6831
Graduate degree	32.38	35.00	25.00	50.00	27.27	30.43	
License attainment							
Has license ¹	84.00	95.00	82.61	85.71	68.18	90.48	0.4691
Does not have license	16.00	5.00	17.39	14.29	31.82	9.52	
Bilingual status							
Is bilingual	28.57	20.00	20.83	37.50	31.82	34.78	0.3499
Is not bilingual	71.43	80.00	79.17	62.50	68.18	65.22	

¹Includes possession of child development certificate, a teacher certificate, or other certificate.

NOTES: ECE = Early childhood education curriculum. PE = Parenting education curriculum.

A stratum-adjusted Kruskal-Wallis test run on project-level averages was used to test baseline balance across study groups. No weights were used at the project level. Multinomial variables like race were transformed into a series of binary recodes, each of which was tested separately.

SOURCE: U.S. Department of Education, Even Start Classroom Literacy Interventions and Outcomes Study, "Staff Survey," Spring 2004.

Table A-6. Average Project Score on the Early Childhood Environmental Rating Scale-Revised (ECERS-R) Measures by Subscale and Study Group: Spring 2004

		Study group					Balance test <i>p</i> -value
		Let's Begin with the Letter People (ECE)	Play & Learning Strategies (ECE/PE)	Partners for Literacy (ECE)	Partners for Literacy (ECE/PE)	Control	
Demographics	Total						
Overall score	4.88	4.97	5.04	4.74	4.75	4.88	0.7786
Subscale							
Language-reasoning	5.15	5.28	5.20	5.12	4.77	5.38	0.4549
Space & furnishings	4.67	4.80	4.97	4.33	4.55	4.64	0.3701
Personal care routines	4.33	4.76	4.38	4.03	4.40	4.04	0.5475
Activities	4.34	4.44	4.42	4.17	4.11	4.53	0.5921
Interactions	5.77	5.54	5.90	5.76	5.70	5.97	0.8487
Program structure	5.02	4.99	5.36	5.01	4.99	4.73	0.8385

NOTES: ECE = Early childhood education curriculum. PE = Parenting education curriculum. Detail may not sum to totals because of rounding. The ECERS-R consists of 37 items that fall within the scope of the 6 subscales listed above. Each item is ranked on a scale from 1 (inadequate conditions) to 7 (excellent conditions). A score of 5 represents good conditions. A stratum-adjusted Kruskal-Wallis test run on project-level averages was used to test baseline balance across study groups. No weights were used at the project level.

SOURCE: U.S. Department of Education, Even Start Classroom Literacy Interventions and Outcomes Study, "ECERS-R," Spring 2004.

Table A-7. Means and Balance Tests for CLIO Study Children, by Study Group and Child and Parent Covariate: Spring 2005 and Spring 2006

Covariate	Total	Study group					Balance test <i>p</i> -value
		Let's Begin with the Letter People (ECE)	Play & Learning Strategies (ECE/PE)	Partners for Literacy (ECE)	Partners for Literacy (ECE/PE)	Control	
Number of children in household under age 8	2.03	1.95	2.07	2.01	2.06	2.07	0.4015
Number of adults in household	2.13	2.13	2.12	2.15	2.14	2.13	0.8446
Responding parent's age in years	30.47	30.47	30.44	30.68	29.70	31.03	0.6990
Mother is college graduate (could be associate's degree)	0.08	0.08	0.09	0.10	0.06	0.09	0.3747
Home language is not English	0.51	0.46	0.50	0.52	0.56	0.52	0.1760
Monthly household income above \$1,500 ¹	0.42	0.43	0.42	0.45	0.37	0.44	0.9735
Number of times child moved in last year	0.42	0.47	0.38	0.43	0.40	0.42	0.5279
Number of times family moved in last year	0.43	0.47	0.39	0.44	0.41	0.43	0.6675
Child has special needs ²	0.12	0.14	0.10	0.09	0.14	0.13	0.4673
Child is male	0.51	0.52	0.55	0.49	0.49	0.49	0.2784
Child is Hispanic/Latino	0.60	0.55	0.64	0.56	0.65	0.61	0.2815
Child age in months	52.79	52.55	52.89	53.23	52.62	52.61	0.5690
Responding parent is Hispanic/Latino	0.59	0.53	0.61	0.54	0.64	0.61	0.4589
Any children in videotaped parent-child interaction have special needs	0.13	0.15	0.12	0.10	0.15	0.14	0.5046
Any children in videotaped parent-child interaction are male	0.54	0.54	0.59	0.53	0.53	0.53	0.3084
Average age in months of children in videotaped parent-child interaction	52.86	52.62	52.86	53.48	52.88	52.41	0.6566

¹Includes income before taxes and other deductions in the past month for all members of the household.

²For example, physical, emotional, language, hearing, or learning difficulties identified by a doctor or other health or education professional.

NOTES: ECE = Early childhood education curriculum. PE = Parenting education curriculum. Data are based on children who took at least one English language and literacy assessment during spring 2005 and/or spring 2006.

Means were run at the child level. A stratum-adjusted Kruskal-Wallis test run on project-level averages was used to test baseline across study groups. No weights were used at the project level.

SOURCE: U.S. Department of Education, Even Start Classroom Literacy Interventions and Outcomes Study.

Table A-8. Project-level Means and Balance Tests, by Study Group and Project-level Covariate: Spring 2005 and Spring 2006

Project-level covariates	Total	Study group					Balance test <i>p</i> -value
		Let's Begin with the Letter People (ECE)	Play & Learning Strategies (ECE/PE)	Partners for Literacy (ECE)	Partners for Literacy (ECE/PE)	Control	
Child to teacher ratio	7.00	6.45	6.61	7.88	6.80	7.26	0.2759
Average baseline score ¹							
Expressive language: English (IGDI)	13.86	14.56	15.02	13.17	12.40	14.18	0.1894
Expressive language: Spanish (IGDI)	13.46	13.67	13.30	13.25	13.37	13.68	0.8153
Receptive vocabulary (PPVT)	247.53	255.48	248.97	243.53	240.16	249.51	0.1690
Phonological awareness: Elision ²	253.79	260.82	255.32	253.66	247.71	251.46	0.1088
Phonological awareness: Blending ³	264.63	266.09	265.00	268.42	265.84	257.80	0.1606
Social competence	237.38	237.76	236.92	241.78	233.82	236.61	0.7480
ECE hours	38.09	46.85	35.55	39.70	34.50	33.85	0.3426
PE/PC hours	73.70	78.36	74.64	69.47	82.22	63.79	0.6929
Parent reading & vocabulary skills ⁴	252.46	259.44	255.83	250.23	248.19	248.60	0.6566
OMLIT Oral Language development	46.32	45.03	48.17	43.98	44.75	49.70	0.2907
OMLIT Print knowledge	47.38	43.46	47.99	49.02	46.87	49.56	0.2674
OMLIT Phonological awareness	47.46	46.06	50.97	47.16	44.01	49.10	0.1283
OMLIT Print motivation	47.75	48.41	48.57	45.07	46.97	49.73	0.8726
OMLIT Literacy resources in classroom	45.86	46.53	45.49	43.40	44.11	49.79	0.3048
Average of child English literacy, child social competence, and parent reading & vocabulary skill baseline scores							
	251.26	255.60	252.93	251.18	247.11	249.46	0.2839
Average of child English literacy, child social competence, and parent reading & vocabulary skill baseline scores, excluding....							
Expressive language: English (IGDI)	251.16	255.92	252.41	251.53	247.14	248.79	0.1413
Receptive vocabulary (PPVT)	252.00	255.62	253.72	252.71	248.50	249.45	0.2146
Phonological awareness: Elision	250.75	254.55	252.45	250.69	246.99	249.06	0.5918
Phonological awareness: Blending	248.58	253.50	250.51	247.73	243.37	247.79	0.1778
Social competence	254.03	259.17	256.13	253.06	249.77	252.03	0.4290
Parent reading & vocabulary skills	251.01	254.83	252.35	251.37	246.90	249.63	0.3628

¹The child receptive vocabulary, Elision, and Blending assessments and the parents reading and vocabulary score were linearly transformed to have a mean of 250 and a standard deviation of 50.

²The Elision subtest measures the child's ability to recognize word parts, such as components of compound words, syllables, and phonemes.

³The Blending subtest measures the child's ability to combine word parts, such as components of compound words, syllables, and phonemes.

⁴The reading and vocabulary score is based on four subtests that measure vocabulary, basic reading, phonics, and comprehension.

NOTES: ECE = Early childhood education curriculum. PE = Parenting education curriculum. Means are based on projects that participated in CLIO during spring 2005 and/or spring 2006. A stratum-adjusted Kruskal-Wallis test run on project-level averages was used to test baseline across study groups. No weights were used at the project level.

SOURCE: U.S. Department of Education, Even Start Classroom Literacy Interventions and Outcomes Study.

APPENDIX B

CHILD AND PARENT LITERACY OUTCOMES

Child Literacy Outcomes

For information on incomplete child assessments, see appendix J, p. J-5.

Scoring Procedures

Simple Scoring Procedure. For the IGDI, Spanish IGDI, and Pre-CTOPPP Print Awareness subtests, the simple score is the number of correct responses. For the Pre-CTOPPP Elision, Pre-CTOPPP Blending, and TOLD-3 Grammar subtests, the simple score is the number correct even if the subtest was stopped due to a run of errors (this assumes the child would have gotten none of the remaining items correct). Since test items were arranged roughly in order of item difficulty, this should be a fair assumption, but it will not do full justice to some children.

For the PPVT, the simple score depends on whether the basal set was administered. If the child basal set was not administered, the child was given full credit for the eight items in it. This procedure is based on the fact that the basal set was not administered only if the child made fewer than eight errors (out of 14) on the first set and the fact that the words in the basal set are easier than those in the first set.

Complex Scoring Procedure. Although the simple scoring procedure yields scores that can be used to estimate unbiased differences in literacy across groups, it is an inefficient procedure that fails to use information about the discriminatory power of the items within each subtest. In contrast, three-parameter item-response theory (IRT) modeling with empirical Bayes (EB) shrinkage to group-specific means exploits that information to get more accurate estimates of the literacy ability of each child. However, the procedure can give badly biased estimates for group means other than for the means of targeted random groups, so it will only be used for the ITT analysis involving comparison of randomly assigned groups of projects.

Three-parameter IRT modeling with EB shrinkage to group-specific means is based on theory due to Mislevy (1984). The fact that it can be badly biased for group contrasts other than those between targeted random groups was also established by

Mislevy (1991). This procedure is more efficient than the simple procedure because it weights items differentially by their discriminatory power and vulnerability to guessing, and it uses auxiliary data about how similar subjects performed on the test.

The procedure gives higher weight to items with greater discriminatory power. Note that discriminatory power is not the same thing as item difficulty. It has more to do with the relevance of the item. Although test designers strive to make all items in a test highly relevant, they cannot succeed in making them of equal relevance. To illustrate the difference between difficulty and relevance, consider adding a math question to a literacy test. The math question can be easy or hard, but it will have low relevance to reading literacy. The complex scoring procedure automatically downweights the less relevant questions, while the simple scoring procedure counts them equally.

The complex procedure gives higher weight to items that are less vulnerable to guessing. Consider that in the Blending and Elision tests, the items vary with respect to the probability that a child can guess the correct answer. That is because half the items are multiple choice and half are open-ended. The complex scoring procedure automatically downweights the items that are easier to guess, while the simple scoring procedure counts them equally.

The third way that the complex method delivers greater efficiency is through use of auxiliary data. This is accomplished by “shrinking” outlying scores back toward group means, where the groups are defined in terms of auxiliary data. For example, the language spoken in the child’s home is a form of auxiliary data. The procedure discounts extraordinarily good and poor performances within groups, acknowledging that with respect to future performance, there is likely to be regression to the group mean. The extent of the shrinkage depends on the number of items in the test and the group sample size. If the number of items is very large and the group sample size is small, there will be very little shrinkage because the reliability of each person’s simple score will be high. On the other hand, if the number of items is small and the group sample size is large, then there will be substantial shrinkage.

We defined the groups by group, data collection cycle,¹ child age in months, and whether the child’s home language is neither English nor a mix of English with some other language. Among these “conditioning variables,” study group and data collection cycle are most important. By using these as conditioning variables for the shrinkage step in the scoring, we ensure that the effects of the curricula are not shrunk. The other two conditioning variables are used to improve precision. From other studies and the early phases of this study, we know that children develop substantially month by month. Also, children in homes where only a foreign language is spoken score much lower on the English literacy assessments. So the imputation and shrinkage is based on the experience of other children of the same age with similar language backgrounds within the same experimental group within the same assessment cycle.

Because of the imputation and shrinkage, the analysis of the complex scores is also more complicated. Complex scores are not independent across children, so simple analysis procedures would underestimate the standard errors on estimated effects. In order to estimate the standard errors on estimated effects well, it is necessary to generate multiple “plausible values” for each child. We generated five plausible values for each child. The variance across these five values was added to naïve variance estimates in order to get good total variance estimates. The variance across the plausible values essentially represents estimated measurement error. If the child’s ability was so well measured by the test that there was no need for imputation or shrinkage, then the plausible values for the child would all be identical.

The complex scoring procedure was used for the ITT analysis. This provided asymptotically unbiased estimates of curriculum effects—estimates that have lower variances than could be achieved with the simple scoring, thereby increasing the statistical power to detect effects. The variance improvement was especially pronounced for the PPVT, given the administration rules. However, the complex procedure was not used for nonexperimental analyses because the complex procedure would tend to wash out effects due to other factors such as parenting practices.

¹ There were five cycles of data collection as discussed in chapter 2. All five cycles were scored in a single run in order to have consistent estimates of item parameters in the IRT model.

Psychometric Properties

Psychometric data are reported for each subtest in table B-1. References for child assessment subtests, including reliability and validity, are included in table B-3; table B-4 includes references for tests cited in the Concurrent Validity column.

Table B-1. CLIO Child Assessment Subtests: Reliability and Validity

Subtest	Reliability	Concurrent validity
IGDI Picture Naming-English	Test-retest reliability (rxx): .67 Alternate forms reliability, for preschool children (rxx): .44 to .78	Correlation with PPVT-III, ages 3 to 5 (rxx): .56 to .75 Correlation with Preschool Language Scale-3, ages 3 to 5 (rxx): .63 to .79
IGDI Picture Naming-Spanish	Information not available from publisher	Information not available from publisher
PPVT-III, publisher version	Internal consistency, publisher (alpha) age 3: .93 age 4: .95 age 5: .94 Test-retest reliability (rxx), ages 2 to 5: .92 Alternate forms reliability (rxx): age 3: .90 age 4: .94 age 5: .93	Correlation with measures of cognitive ability (Wechsler Intelligence Scale for Children-III), ages 7 to 14 (rxx): .91 verbal IQ .82 performance IQ .90 full scale IQ Correlation with measures of oral language (Oral and Written Language Scales), ages 3 to 5 (rxx): .83 oral expression .66 listening comprehension .82 oral composite
PPVT-III, adaptive version for CLIO	Internal consistency, CLIO administration (alpha): Spring 2005— age 3: .86 age 4: .87 age 5: .87 Spring 2006— age 3: .82 age 4: .87 age 5: .85	Correlation with Academic Knowledge task from Woodcock-Johnson III, for kindergarten-age children (rxx): .58
Preschool CTOPPP Print Awareness	Internal consistency, publisher (alpha): age 3: .89 age 4: .94 age 5: .95	Information not available from publisher

Table B-1. CLIO Child Assessment Subtests: Reliability and Validity (continued)

Subtest	Reliability	Concurrent validity
Preschool CTOPPP Elision	Internal consistency, publisher (alpha): age 3: .78 age 4: .87 age 5: .85	Preschool CTOPPP Elision
Preschool CTOPPP Blending	Internal consistency, publisher (alpha): age 3: .88 age 4: .89 age 5: .86	Preschool CTOPPP Blending
Grammatical Understanding (TOLD:P3)	Internal consistency, publisher (alpha): age 4: .86 age 5: .82 Test-retest reliability, for children in kindergarten to grade 2 (rxx): .81	Correlation with Bankson Language Test-Second Edition, for children in grades 1-3 (rxx): .79 semantic knowledge .64 morphological/syntactic rules .67 overall language quotient

NOTE: Unless otherwise noted, reliability and validity information are supplied by test publishers.

Parent Literacy Outcomes

The parent assessment scale consisted of four subtests combined into a single outcome measure: the PPVT and three Woodcock-Johnson subtests (Letter-Word Identification, Passage Comprehension, and Word Attack). For information on missing parent assessments, see appendix J, p. J-6 and J-7.

Scoring

The parent assessments were scored three different ways. First, a simple raw score was calculated for each parent on each test. Second, a complex score was based on 3-parameter IRT scoring with empirical Bayes shrinkage to group-specific means, as discussed for child outcomes. These were then linearly transformed to a metric with mean and standard deviation on the entire set of parents of 250 and 50, respectively. The IRT scores were used in the ITT analysis, and the simple raw scores were used in the nonexperimental analyses. The third scoring was based on publisher procedures; these were used to provide context in the section in chapter 2 entitled Even Start Continues to Work With a Needy Population.

Psychometric Properties

Table B-2 presents psychometric data for each subtest. References for parent assessment subtests, including reliability and validity, are included in table B-3; table B-4 includes references for tests cited in the Concurrent Validity column.

Table B-2. CLIO Parent Assessment Subtests: Reliability and Validity

Subtest	Reliability	Concurrent validity
Woodcock-Johnson Letter-Word Identification	Reliability coefficient, ages 20-29 (rxx): .91 Test-retest reliability, <1 year, ages 19-44 (rxx): .90	WJ-III Achievement Clusters are compared to Kaufman Test of Educational Achievement (KTEA) and Wechsler Individual Achievement Test (WIAT), for children in grades 1-8 (rxx). - Basic Reading (LWI & WA subtests) correlated with KTEA Reading Composite (.66) and WIAT Basic Reading (.82) - Reading Comprehension (PC & another test not part of CLIO) correlated with KTEA Reading Comprehension (.62) and WIAT Reading Comprehension (.79)
Woodcock-Johnson Passage Comprehension	Reliability coefficient, ages 20-29 (rxx): .75 Test-retest reliability, <1 year, ages 19-44 (rxx): .84	
Woodcock-Johnson Word Attack	Reliability coefficient, ages 20-29 (rxx): .83 Test-retest reliability, <1 year, ages 19-44 (rxx): NA	
PPVT-III	Internal consistency, publisher (alpha) ages 19-24: .94 ages 25-30: .97 ages 31-40: .97 Test-retest reliability, ages 26-57 (rxx): .93 Alternate forms reliability (rxx): ages 19-24: .92 ages 25-30: .94 ages 31-40: .94	Correlation with Kaufmann Brief Intelligence Test (K-BIT) Vocabulary Test, for adults with mean age of 25 (rxx): .82

Table B-3. References for CLIO Child and Parent Assessment Subtests, Including Reliability/Validity Data

Subtest	References
IGDI test	Early Childhood Research Institute on Measuring Growth and Development (2003). <i>Picture Naming Individual Growth and Development Indicator</i> . Minneapolis, MN: Center for Early Education and Development, University of Minnesota.
IGDI-reliability/validity data	Missall, K.N., and McConnell, S.R. (2004 April). <i>Technical Report: Psychometric Characteristics of Individual Growth and Development Indicators: Picture Naming, Rhyming, and Alliteration</i> . Center for Early Education and Development, University of Minnesota.
PPVT-III (publisher) test	Dunn, L.M., and Dunn, L.M. (1997). <i>PPVT-III. Peabody Picture Vocabulary Test, Third Edition</i> . Circle Pines, MN: American Guidance Service.
PPVT-III (pub)-reliability/validity data	Williams, K.T., and Wang, J.J. (1997) <i>Technical References to the Peabody Picture Vocabulary Test – Third Edition (PPVT-III)</i> . Circle Pines, MN: American Guidance Service, Inc.
PPVT (adaptive)-validity	Sorongon, A.G. (2007 March). <i>Predictive Validity of Measures of Preschool Children’s Cognitive and Social Skills to Kindergarten School Performance Indicators</i> . Presented at the Society for Research in Child Development Biennial Conference in Boston, MA.
Pre-CTOPPP test	Lonigan, C.J., Wagner, R.K., Torgesen, J.K., and Rashotte, C.A. (2002). <i>The Preschool Comprehensive Test of Phonological and Print Processing</i> . Tallahassee, FL: Florida State University.
TOLD test	Newcomer, P.L., and Hammill, D.D. (1997a). <i>Test of Language Development – Primary: Third Edition (TOLD-P:3)</i> . Austin, TX: PRO-ED, Inc.
TOLD-reliability/validity data	Newcomer, P.L., and Hammill, D.D. (1997b). <i>Examiner’s Manual (1997b). Test of Language Development, Primary (Third Edition)</i> . Austin, TX: PRO-ED, Inc.
Woodcock-Johnson test	Woodcock, R.W., McGrew, K.S., and Mather, N. (2001). <i>Woodcock-Johnson III Tests of Achievement</i> . Itasca, IL: Riverside Publishing.
WJ-reliability/validity data	McGrew, K.S., and Woodcock, R.W. (2001). <i>Technical Manual. Woodcock-Johnson III</i> . Itasca, IL: Riverside Publishing.

Table B-4. References for Tests Cited in Concurrent Validity Columns of Tables B-1 and B-2

Subtest	References
Child Assessment (table B-1)	
IGDI	Dunn, L.M., and Dunn, L.M. (1997). <i>PPVT-III. Peabody Picture Vocabulary Test, Third Edition</i> . Circle Pines, MN: American Guidance Service
	Zimmerman, I.L., Steiner, V.G., and Pond, R.E. (1992). <i>Preschool Language Scales—Third Edition</i> . San Antonio, TX: The Psychological Corporation.
PPVT-III	Wechsler, D. (1991). <i>Wechsler Intelligence Scale for Children—Third Edition</i> . San Antonio, TX: The Psychological Corporation.
	Carrow-Woolfolk, E. (1995). <i>Oral and Written Language Scales: Listening Comprehension and Oral Expression</i> . Circle Pines, MN: American Guidance Service, Inc.
TOLD	Bankson, N.W. (1990). <i>Bankson Language Test—Second Edition</i> . Austin, TX: PRO-ED
Parent Assessment (table B-2)	
Woodcock-Johnson	Kaufman, A.S., and Kaufman, N.L. (1985) <i>Kaufman Test of Educational Achievement</i> . Circle Pines, MN: American Guidance Service, Inc.
	Wechsler, D. (1992). <i>Wechsler Individual Achievement Test</i> . San Antonio, TX: The Psychological Corporation.
PPVT-III	Kaufman, A.S., and Kaufman, N.L. (1990) <i>Kaufman Brief Intelligence Test</i> . Circle Pines, MN: American Guidance Service, Inc.

APPENDIX C

CONSTRUCTION OF THE CHILD SOCIAL COMPETENCE SCALE

The social competence scale for CLIO was created using classroom teacher reports on individual children and their behavior and social skills. The scale combines information from two different sets of items in the CLIO Teacher's Rating Form—cooperative behavior and problem behavior. This form combined items from two scales developed for the Head Start Family and Child Experiences Survey (FACES) (U.S. Department of Health and Human Services 2003). These scales were called the Cooperative Classroom Behavior Scale and the Behavior Problems Scale on FACES. (For information on missing teacher rating forms, see appendix J, p. J-6.

The FACES Cooperative Classroom Behavior Scale consists of 12 items adapted from the Personal Maturity Scale (Alexander and Entwisle 1988) and the Social Skills Rating System (Elliott, Gresham, Freeman, and McCloskey 1988). Each item is rated on a 3-point scale: never, sometimes, or very often. On the FACES study, a Cronbach alpha of 0.88 was reported. Using spring 2004 CLIO data, we obtained a value of 0.89 for the same statistic.

The FACES Behavior Problems Scale consists of 14 items concerning aggression, hyperactivity, and withdrawal adapted from the Personal Maturity Scale (Alexander and Entwisle 1988), the Child Behavior Checklist for Preschool-Aged Children, Teacher Report (Achenbach, Edelbrock, and Howell 1987) and the Behavior Problems Index (Zill 1990). These items use a 3-point scale: not true, somewhat or sometimes true, very true or often true." On the FACES study, a Cronbach alpha of 0.86 for the 14 items was reported. Using spring 2004 CLIO data, we obtained a value of 0.84 for the same statistic.

Although the 26 items in the CLIO Teacher's Rating Form had been used to form two scales in FACES (and in fact, the 14 items in the FACES Behavior Problems Scale had been further decomposed into three subscales for aggressive, hyperactive, and withdrawn behavior), the strong focus of the CLIO curricula on literacy activities rather than on behavior modification led us to want a single scale to measure social competence. The fact that we have 20 outcomes for child literacy, parent literacy, parenting skills, instructional practices, and participation means that we are already at some risk of false positive findings. We hoped that the two FACES scales could be

combined into a single social competence scale for CLIO. We lose some sensitivity with this decision because the two scales probably do tap slightly different latent behavioral factors, but the two scales are substantially correlated, so the possible loss in sensitivity is more than compensated by the improvement in specificity.

To test the feasibility of this combination, we ran four-parameter logistic IRT models¹ on spring 2004 data for all 26 items together as well as separately for the 12 items in the FACES Cooperative Classroom Behavior Scale and the 14 items in the FACES Behavior Problems Scale. The Cronbach's alpha for the single scale with 26 items was 0.92; moreover, the correlation between the two FACES scales was 0.57. However, two items did not fit into the combined scale well. Their correlations with the combined scale were 0.16 and 0.24, respectively. We dropped these two items and fit a new IRT model for the remaining 24 items. Although the Cronbach's alpha for the single scale with 24 items rounded to the same 0.92 as the original scale with 26 items, the scale is more unidimensional, cohesive, and interpretable without these two items.

This scale was named "social competence." The technical details of the scoring parallel those for the child assessments, as discussed in appendix B. IRT models were fit on the combined 2005/2006 data and linearly transformed to have a mean of 250 and a standard deviation of 50 on the combined dataset (when averaged across the five plausible values). Scores for the baseline year were obtained by applying the same scoring algorithm without any new IRT modeling.

As with the child assessments, we also computed a simple raw score. This simple raw score excludes the same two items excluded by the IRT scoring. In parallel with the procedures for the literacy scores, the IRT scores were used in the ITT analysis, and the simple raw scores were used in the nonexperimental analysis.

¹ We used four-parameter models rather than the three-parameter models used for the child literacy assessments because the socio-emotional items are on 3-point Likert scales rather than binary indicators.

APPENDIX D

CONSTRUCTION OF PARENTING SKILLS SCALES

Data on parenting skills were collected from staged parent-child interactions and from parental self-report. There were two staged interactions for each parent-child dyad.¹ One involved having the parent and child read a book together, where the parent had the choice of an English- or Spanish-language edition of the book. The other involved having them play with a toy together. The toy was chosen to elicit play-acting from the parent and child. Both interactions were videotaped, and the videotapes were then coded in their entirety. Self-reports were gathered through a parent interview administered at the Even Start project. (Information on missing data can be found in appendix J, p. J-7 for the parent interview and p. J-9 for the videotapes.)

Joint Book Reading. Three separate coding systems were applied to the joint book reading task. One focused on the mechanics of reading, another on behaviors with emotional overtones, and a third on summarization. There is considerable overlap in the behavioral dimensions measured by the three. These three systems were called Reading Aloud Profile – Together (RAPT), the Contingency Scoring Sheet, and Quality Indicators.

The RAPT is based on the OMLIT-RAP, the instrument developed for measuring instructional behavior during book reading (a description of the OMLIT-RAP is provided in appendix E, p. E-4). Fifty-five specific behaviors are measured in the RAPT,² with some items focusing on parent behaviors and others focusing on child behaviors. The behaviors are grouped by when they took place: “during pre-reading,” “during reading,” or “during post-reading.” They include such items as whether the parent tracks the print with a finger during reading and whether the child points to

¹ For parents with multiple sample children, the staged interaction usually involved the parent and all her/his children present rather than having separate sessions for the parent to interact with each child individually. However, the coding was done separately by child. In 2005, 8 percent of the interactions involved more than one child; in 2006, 7 percent involved more than one child. We have no direct information about the impact on parenting outcomes of the number of children in the interaction. However, we have some indirect information. Parents with multiple children in the study had statistically significantly lower responsiveness; this lower responsiveness could either be a true effect or an artifact of the measurement process caused by having multiple children present in most of the staged interactions. No significant association was found between number of children in the study and parent interactive reading skill.

² There were 57 items in the fall 2004 version of the RAPT, but two of these were dropped for spring 2005 and spring 2006.

pictures or words. If a particular behavior was observed at least once during the book reading task, the corresponding item on the RAPT was checked.

The “Contingency Scoring Sheet” instrument included eight 7-point Likert scales, five of which characterize the parent’s behavior and three of which characterize the child’s behavior. The scales were rated globally, based on the entire sum of observed behavior during the task. The scales are:

- Parental Supportiveness: Emotional availability and physical/affective presence;
- Parental Stimulation of Cognitive Development: Effortful teaching to enhance perceptual, cognitive, and linguistic development;
- Parental Intrusiveness: Parental control of child rather than recognizing and respecting the validity of the child’s perspective;
- Parental Negative Regard: Expression of discontent with, anger toward, disapproval of, and/or rejection of the child;
- Parental Detachment: Lack of awareness of, attention to, and engagement with the child;
- Child Engagement of Parent: Child (a) shows, initiates, and/or maintains interaction and (b) communicates positive regard and/or affect to the parent;
- Child’s Sustained Interest: Child displays focus, excitement, interest, question asking, relating to personal experience; and
- Child Negativity Toward Parent: Shows frustration, anger, hostility, or dislike toward parent.

The Quality Indicators consist of three 5-point Likert scales. These quality items focus on three aspects of the reading interaction: (1) the degree to which the parent introduced and contextualized new vocabulary to support the child’s learning; (2) the extent to which the parent used open-ended questions that invite the child to engage in prediction, imagination, and/or rich description; and (3) the depth of child’s engagement with the reading activity. Whereas the RAPT items indicate simply whether certain behaviors were observed, the Quality Indicators provide information

on the frequencies of higher and lower quality behavior based on the full interaction over the course of the reading task.

Toy Interaction Activity. For the toy activity, the same Contingency Scoring as described above was used to code the interactive play.

Training of Coders. Training took place over several days. The first day focused on the RAPT coding sheet. This included section-by-section overview and practice using actual videotaped interactions. The next 2 days focused on the Contingency Scales. Starting with the Parental Behaviors scale, training covered one scale at a time. When viewing the videos after being trained on a scale, the coder would then code all the previous scales. For example, first the "Supportiveness" scale was covered; coders watched a video and then coded only supportiveness. Then the "Cognitive Stimulation" scale was covered; coders watched another video and coded cognitive stimulation and supportiveness. Successive scales were added in this fashion until coders were able to code all scales at once.

After being trained on the scales, coders participated in several days of group practice coding, which included the RAPT and the contingency scales. Following this exercise, the coders worked on individual practice tapes (generally around 10), which were then reviewed with the supervisor. Once the coder was coding successfully on the practice tapes, the coder moved on to coding reliability tapes. Starting with 10 reliability tapes, the target was at least 85 percent reliability on these tapes. If the coder did not reach this level of reliability, then any problem areas were worked on and the coder was given 10 more reliability tapes. A coder was not allowed to begin coding for the study until he/she reached a minimum of 85 percent reliability.

Throughout the coding for the study, the supervisor randomly checked the coding to ensure continued reliability. Also the supervisor reviewed all entered paperwork. If any coding seemed to be out of line with what was recorded, then the coder was consulted and, if necessary, modifications were made and that coder was subject to further reliability coding.

Parent Interview. A total of 16 items from the parent interview were used in this analysis. The items from the parent interview relate to the frequency and type of reading activities engaged in with the child at home and were included because they

reflect the parent's engagement in teaching activities at home and were, therefore, believed to be related to parenting skills. The parent interview items were:

- In the past week, have you or someone in your family:
 - Told your child a story?
 - Helped your child learn the names of letters, words, or numbers?
 - Practiced writing the letters of the alphabet with your child?
 - Practiced the sounds that letters make with your child?
 - Discussed new words with your child?
 - Helped your child learn songs or music?
 - Practiced writing or spelling your child's name?
 - Talked about rhyming words (e.g., mat, sat) with your child?
- How many times have you or someone in your family read to your child in the past week?
- When you read to your child do you:
 - Stop reading and ask them to tell you what is in the picture?
 - Stop reading and point out letters?
 - Stop reading and ask what will happen next?
 - Read the entire story as they listen without interrupting?
- When you read to your child do you:
 - Read the same story, over and over?
 - Ask them to read with you?
- How often did your child ask you to read books to him/her in the past week?

Extracting Parent Outcomes. Across all the above parenting measures, there are a total of 90 variables. As discussed earlier, we needed to compress this information into a small number of parenting scales in order to avoid problems with multiple comparison testing. We ended up choosing a complex method for forming two such scales. Before discussing it, we review why a simpler solution was rejected.

The rejected method was to simply sum or average items from the same instrument. An investigation of the fall 2004 data quickly showed that this would not be

the best solution. For example, the scale for the Contingency Scoring Sheet for parental stimulation of cognitive development during the book reading was more closely related to the RAPT items about the mechanics of reading than to the other 15 scales³ in the Contingency Scoring Sheet. As another example, a RAPT item about whether the child loses interest or walks away before the book reading has been completed is more closely related to scales in the Contingency Scoring Sheet such as child negativity toward parent and child sustained interest than with other items in the RAPT.

The approach we chose is a combination of variable clustering and factor analysis. Using fall 2004 data, we experimented with a variety of factorizations with different numbers of factors and different rotations of the factors. One criterion was that the scales should be positively associated with each of the positive items in the item pool and negatively correlated with each of the negative items in the item pool.⁴ None of the experiments we ran yielded a set of factors that met this criterion. We then combined variable clustering with factor analysis. Using variable clustering software, we divided the item pool into two clusters. We then used factor analysis to extract the first factor within each cluster. This technique resulted in two scales that had the proper direction of correlation with all items. Table D-1 shows the names that we assigned to the two scales and the set of items that dominate each scale. Both scales are standardized to have a mean of zero and a standard deviation of one in fall 2004. The fall 2004 linear transformations were applied to spring 2005 and spring 2006 data without any new variable clustering analysis or factor analysis.

There are 49 items in the first scale and 41 in the second. Item weights vary substantially within each set. In the first scale, the relative variance of the absolute weights is 0.33. When standardized to sum to 49, the weights in the first scale vary from 0.22 to 2.66. Of the 49 items, 22 have a scaled weight greater than 1. In the second scale, the relative variance of the absolute weights is 0.48. When standardized to sum to 41, the weights in the second scale vary from 0.22 to 2.56. Of the 41 items, 15 have a scaled weight greater than 1.

³ The 15 scales consist of the remaining 7 scales from the reading session, and the 8 scales from the toy session.

⁴ We used our judgment to classify specific behaviors as positive or negative with respect to utility in parenting.

Table D-1. Items Dominating Each Parenting Outcome, Fall 2004

Outcome 1: Parent interactive reading skill	Outcome 2: Parent responsiveness
Reading task – Contingency scoring: • Parent cognitive stimulation	Reading task – Contingency scoring: • Parent supportiveness • Child engagement of parent • Child negativity toward parent
Reading task – RAPT – Prior to reading • Child verbally responds to questions from parent about book • Parent captures child’s attention – expresses interest in book	
Reading task – Quality indicator • Quality of open-ended questions and techniques for eliciting responses to them	
Reading task – RAPT – During reading • Child labels, names pictures • Child makes comments related to text, pictures or parent’s comments • Parent discusses/expands on meaning of illustration or text; offers new information • Parent expands on child’s comments/questions about the story	Reading task – RAPT – During reading • Child verbally responds to questions from the parent about book • Parent directs child’s attention to illustration
	Toy task – Contingency scoring: • Parent supportiveness • Parent cognitive stimulation • Child engagement of parent • Child sustained interest

The items listed for each outcome are those with loadings of greater than |0.50|.

The parent interview items did not contribute heavily to either scale. Of the 16 items, 15 are in the parent responsiveness scale. Their average scaled weight is 0.51, and none have a scaled weight greater than 1.0.

The correlation between the two scales is +0.6. Note that this correlation is slightly larger than the correlation between the two subscales of social competence, yet we advocate keeping these two parenting scales separate while collapsing the social competence scales. The difference is because of the targeting of the curricula. Because the parenting curricula target parenting practices more strongly than the preschool curricula target child social competence, it seems appropriate to have two scales for the first and just one for the second. The biggest difference between the scales is that the

parent responsiveness scale captures those elements of the parent-child interaction that are common across the read aloud and toy play sessions, which mostly involve reciprocal warmth and affection, while parent interactive reading skills captures those elements of parent-child interaction that are unique to the read aloud, which include the use of open-ended questions and enthusiasm for reading.

Siblings. For parents with multiple children in the study, the two parenting scores were averaged across the sample dyads with nonmissing parenting scores. Parenting behaviors do vary by child within a family, but there are compelling reasons to average out some of this variability. First, the analysis software cannot properly handle the covariances that are sure to exist in the parenting behaviors of a parent to her/his individual children. Second, many of the components of the parenting index are not child specific.

Alignment with parenting curricula. The formal goals of both of the parenting curricula used in the CLIO study are similar. For example, both Partners for Literacy and Play and Learning Strategies (PALS) focus on teaching parents how to encourage both the emotional and cognitive development of their children. To reach these goals, the two curricula take slightly different approaches; Partners for Literacy focuses on enriched care giving, while PALS stresses instruction in child development and responsiveness to behavioral cues. However, despite these different instructional paradigms, parents in both curricula are taught to encourage emotional development through positive emotional support. Similarly, parents are taught how to encourage cognitive development through use of rich vocabulary words and early reading instruction, including joint book reading.

The two main goals of both parenting curricula are succinctly captured in the two parenting outcomes. The emotional development goal is captured in the parent responsiveness outcome that measures the extent to which parents relate positively to their child and are sensitive to their child's emotions. The cognitive development goal is captured in the parent interactive reading skill outcome that measures the degree of discussion, conversation, and verbal engagement between the parent and the child.

Reliability and Validity. There are no pre-existing data on the reliability or validity of the two scales. Nor is there any way that we know of to estimate the reliability of the scale from the data collected. Although we have some data on the

consistency of the coding across coders, we have no data on the variability of the behavior across days, settings, book/toy choices, moderators, and so on. Since factor analysis was used to derive the scales rather than forming a simple average of a set of related items, the common and simple expedient of giving a Cronbach's alpha can give severe underestimates of reliability because the unequal weighting of the items improves reliability, and the alpha is an appropriate indicator of reliability only when items are equally weighted. An alternative has been developed to measure reliability of scales based on unequal weighting, but Gorsuch (1980) noted that even this solution is flawed when some of the weights are very small as is the case in these scales. A solution suggested by Gorsuch is to focus on items with large weights. The overall alpha for the 49 items in parent interactive reading skill is 0.79. Among the 22 items in it with larger than average weights, the alpha coefficient is 0.84. The overall alpha for the 41 items in parent responsiveness is 0.55. Among the 15 items in it with larger than average weights, the alpha coefficient is 0.80.

In terms of validity, chapter 7 includes analyses in which the two parenting scales were used as putative causal variables rather than as outcome variables. These analyses found a statistically significant and positive relationship between the parent responsiveness scale and all of the targeted child outcomes. On the other hand, we did not find any statistically significant relationships between parent interactive reading skill and child emergent literacy.

Coder: _____ CHILD ID# _____

Date: _____ Child's Name: _____

READ ALOUD PROFILE – TOGETHER (OMLIT-RAPT)

A. PRE-Reading Activities				B. Behavior DURING Reading						C. POST-Reading Activities			
A1. Caregiver (circle all that apply)		A2. Child (circle all that apply)		B1. Caregiver (circle all that apply)		B2. Child (circle all that apply)				C1. Caregiver (circle all that apply)		C2. Child (circle all that apply)	
1	Ensures child is comfortable, can see book	1	Expresses interest, excitement	1a	Tracks print with finger, labels punctuation	1b	1a	Attends to picture/story	1b	1	Asks questions about child's interest in book	1	Asks to read book again
2	Captures child's attention – expresses interest in book	2	Verbally responds to questions from parent about book	2a	Uses gestures, dramatic voices, props, tone of voice to interest child	2b	2a	Verbally responds to questions from parent about book	2b	2	Allows child to look at book	2	Responds to questions, expands on parent's comments about book
3	Labels, reads, directs attention to features of book such as title, author, illustrations or illustrator	3	Tells parent things about book, point out features of book	3a	Directs child's attention to illustrations	3b	3a	Points to pictures, words	3b	3	Answers child's questions about story or related topics	3	Comments on story/illustrations
4	Points to features of book such as title, author, illustrations or illustrator, tracks print	4	Asks questions about the book	4a	Asks story-related <u>close-ended</u> questions, not recall	4b	4a	Labels, names pictures	4b	4	Expands on child's comments about story/illustrations	4	Asks questions about story or related topics
5	Tells child sounds/letters to listen for, look for	5	Expands on parent's comments about book	5a	Discusses/expands on meaning of illustrations or text; offers new info	5b	5a	Repeats words/parts of story	5b	5	Reviews/reinforces vocabulary in book	5	Tries to "read" book on own – turning pages, exploring pictures
6	Reminds child of similar books s/he has read/ if s/he has read same book before	6	Tells parent things about the story line	6a	Expands on child's comments/questions about the story	6b	6a	Acts out/makes sounds related to story	6b	6	Asks for recall of information about story	6	No post-reading activities (without codes 1-6)
7	Responds to questions, expands on child's comments about book	7	No pre-reading activities (without codes 1-6)	7a	Comments on sound, letters, sound-letter links	7b	7a	Connects story to own life	7b	7	Asks questions about story that relate to child's own experiences		
8	Expands on book through close-ended questions, discussion, vocabulary, and/or background knowledge			8a	Highlights new vocabulary	8b	8a	Makes comments <u>related</u> to text, pictures or parent's comments	8b	8	Asks story-related <u>open-ended</u> questions		
9	Relates text to child's experiences/asks story related questions about child's experiences			9a	Asks recall questions about earlier parts of the story	9b	9a	Asks questions <u>related</u> to text, pictures or parent's comments	9b	9	Summarizes/retells story <u>without</u> child involvement		
10	Asks story-related <u>open-ended</u> questions			10a	Relates text to child's experiences/asks story related questions about child's experience	10b	10a	Tries to "read" book on own – turning pages, exploring pictures	10b	10	Summarizes/retells story <u>with</u> child involvement		
11	No pre-reading activities before reading begins			11a	Asks story-related <u>open-ended</u> questions	11b	11a	Tries to "read" book on own – telling story	11b	11	No post-reading activities (without codes 1-10)		
				12a	Has child join in reading/ completing text on own	12b	12a	Loses interest or walks away before book is completely read	12b	Length of Interaction: _____			
		13a	No Reading activities (without codes 1-12)	13b	13a	No Reading activities (without codes 1-12)	13b						

Quality Indicators for RAPT					
Story-related Vocabulary	☐ 1 (Minimal)	☐ 2	☐ 3 (Moderate)	☐ 4	☐ 5 (Extensive)
	Some story-related vocabulary words are introduced/discussed but the definition of one or more of the words is misleading or wrong. OR No new vocabulary introduced or discussed.		Two or three story-related vocabulary words are introduced or discussed and the definition is accurate. <i>Both of the following supports are given for each word:</i> i. A picture, gesture, or other concrete visual aid is used; or ii. The word is linked to a rich network of related words or concepts.		Six or more story-related vocabulary words are introduced or discussed and the definition of each vocabulary word is accurate. <i>Both of the following supports are given for each word:</i> i. A picture, gesture, or other concrete visual aid is used; and ii. Each word is linked to a rich network of related words or concepts.
Use of Open-Ended Questions ^a	☐ 1 (Minimal)	☐ 2	☐ 3 (Moderate)	☐ 4	☐ 5 (Extensive)
	Parent poses only one open-ended question. Parent rarely/never provides opportunity for child to respond (not allowing much time, not restating question or not acknowledging child's response). OR Parent poses no open-ended questions.		Parent poses two or three open-ended questions. Parent consistently shows interest in/actively encouraging child's response (e.g., pausing for child, restating question, scaffolding, or acknowledging child's response).		Parent poses at least four open-ended questions. Parent consistently shows interest in/actively encouraged child's responses (e.g., pausing for child, restating question, scaffolding, or acknowledging child's response).
Depth of Parent-Child Discussion	☐ 1 (Minimal)	☐ 2	☐ 3 (Moderate)	☐ 4	☐ 5 (Extensive)
	Parent engages child in no or low-level discussion only; no extended discussion before, during or after reading. Parent/child discussion consists mainly of short comments, management statements.		Parent engages child in one extensive discussion before, during or after reading. Parent/child discussion involves at least 3 turns (1 turn is one back-and-forth) Parent/child discussion lasts at least 2 minutes.		Parent engages child in extensive discussion at least twice before, during or after reading Parent/child discussion involves at least 3 turns (1 turn is one back-and-forth) Parent/child discussion lasts at least 2 minutes.
☐ Read Aloud ends before book is completed. Explain Circumstances: _____ _____ _____					
Reading Aloud Profile - Together (RAPT). WESTAT Rockville, MD, (c) 2004. Reprint only with permission of authors. Even Start Classroom Literacy Interventions and Outcomes Study (CLIO)					

Contingency Scoring Sheet

Book Reading

(CLIO Spring 2005)

Coder: _____		CHILD ID#: _____	
Date: _____		Child's Name: _____	
I. PARENT'S BEHAVIOR			
Supportiveness		Stimulation of Cognitive Development	
1 2 3 4 5 6 7 NC			
Intrusiveness			
1 2 3 4 5 6 7 NC		1 2 3 4 5 6 7 NC	
Negative Regard		Detachment	
1 2 3 4 5 6 7 NC		1 2 3 4 5 6 7 NC	
II. CHILD'S BEHAVIOR			
Engagement of Parent		Negativity toward Parent	
1 2 3 4 5 6 7 NC		1 2 3 4 5 6 7 NC	
Sustained Interest in Book		Read this book before? ☐ Yes ☐ No	
		If yes; How many times? _____	
1 2 3 4 5 6 7 NC		_____	

Contingency Scoring Sheet

Toy Reading

(CLIO Fall 2004)

Coder: _____		CHILD ID#: _____	
Date: _____		Child's Name: _____	
I. PARENT'S BEHAVIOR			
Supportiveness		Stimulation of Cognitive Development	
CHILD ID#: _____ Date: _____			
1 2 3 4 5 6 7 NC			
Intrusiveness			
1 2 3 4 5 6 7 NC		1 2 3 4 5 6 7 NC	
Negative Regard		Detachment	
1 2 3 4 5 6 7 NC		1 2 3 4 5 6 7 NC	
II. CHILD'S BEHAVIOR			
Engagement of Parent		Negativity toward Parent	
1 2 3 4 5 6 7 NC		1 2 3 4 5 6 7 NC	
Sustained Interest in Toys		Were others present? ☐ Yes ☐ No Is this a twin/sibling case? ☐ Yes ☐ No If yes, indicate Twin ID#: _____	
1 2 3 4 5 6 7 NC			

APPENDIX E

PRESCHOOL INSTRUCTIONAL OUTCOME VARIABLES

Introduction

The preschool instructional outcome measures were derived from a new observation instrument developed for the CLIO study—Observation Measures of Language and Literacy Instruction, or OMLIT (Goodson, Layzer, Smith, and Rimdzius 2004). The OMLIT focuses directly on aspects of early childhood education instructional practice that, based on professional opinion and research, support children’s acquisition of early literacy skills. (The OMLIT also provides general descriptive information about the classroom organization and activities.) This appendix first describes the development of the OMLIT measures, how the measures were administered in the CLIO study, and how the observers were trained to use the measures. The appendix then discusses the OMLIT constructs that were developed to describe instructional practices, for the impact and supplemental analyses. A copy of the OMLIT is provided at the end of this appendix. (For information on missing data, see appendix J, p. J-9.)

Foundations and Development of the OMLIT

Both of the preschool curricula implemented in CLIO were intended to promote early literacy skills which, in theoretical writings, expert opinions, and best practice documents, have been considered to be precursors to later conventional literacy skills. The field has been influenced by two documents that considered the available research on which early literacy skills appear to be foundational for later literacy. The report of the National Research Council’s panel on preventing reading difficulties in young children (Snow, Burns, and Griffin 1998), identified weaknesses in oral language, phonological awareness, and alphabet knowledge as predictors of later significant reading problems. Whitehurst and Lonigan (1998) identified skills in four domains—oral language, print and letter knowledge, and phonological processing—as encompassing two aspects (outside-in and inside-out skills) of emergent literacy that are related to later conventional forms of reading and writing. These domains are defined as follows:

- *oral language*: lexical/conceptual, semantic, and syntactic abilities;

- *phonological processing*: sensitivity to and ability to manipulate the sounds in words;
- *print knowledge*: knowledge of the units of print (letters, words) and ability to translate print to sound and sound to print (letter-sound and ultimately word-sound); and
- *print motivation*: interest in reading and writing activities.

More recently, the National Early Literacy Panel completed a systematic review of research that involved the measurement of one or more early literacy skills assessed when children were between birth and 5 years of age or in kindergarten, and the measurement of one or more conventional literacy skills assessed when children were in kindergarten or older (Lonigan, Schatschneider, Westberg, in review).¹ Based on review of 300 primary research articles, the Panel identified a set of 11 early literacy skills that were found to have significant and meaningful relationships with later measures of conventional literacy (decoding, print comprehension, and spelling). The early literacy skills most strongly and consistently related to later literacy included alphabet knowledge, phonological awareness, writing own name, phonological memory, and rapid automatized naming. These skills were predictive of later literacy in multiple studies, even when demographic and other child characteristics were accounted for. Other skills that were consistently although less strongly related to later literacy outcomes included concepts of print, oral language, print knowledge, and visual processing. (The number of relevant studies varied by measure; for example, the most studies were available on alphabet knowledge ($N = 52$ studies) and phonological awareness ($N = 69$ studies)). The OMLIT was developed to measure aspects of early childhood education classrooms and instructional practice that support the development of these important early literacy skills. At the time the OMLIT was developed, there was some consensus among professionals about specific classroom practices that are linked to the development of early literacy skills (Lonigan 2006; Whitehurst and Lonigan 2001). However, there was not a strong research base that systematically tested the impacts of different classroom practices on the acquisition of early literacy skills in preschool. Thus, the OMLIT was developed based on a combination of theory, professional opinion, and research. The rationale and

¹ The NELP looked at published scientific studies that could provide correlational evidence showing the relationship between early skill attainment and later literacy growth in decoding, reading comprehension, or spelling.

supporting research for each of the OMLIT measures is provided in the OMLIT Training Manual, which can be obtained from Abt Associates, Inc. (Goodson and Layzer (2005).

The final version of the OMLIT was developed over a 12-month period with multiple rounds of pilot-testing and revision. The OMLIT includes five observation measures and a Classroom Description of the classroom on the day of the observation. The measures are briefly described below and summarized in table E-1.

- **The Classroom Description** provides contextual information for the observation. Before the observation begins, the observer records the staff assigned to the classroom and present, the number of children enrolled and their home language and special education status, and any current classroom theme. At the end of the observation period, the observer records the language(s) used in instruction by each staff member across the observation period.
- **The Classroom Literacy Opportunities Checklist (OMLIT-CLOC)** is an inventory of classroom literacy resources that observers complete at the end of a half day of observation. It provides a rating of the extent to which a classroom is a literacy-rich environment and is divided into 10 sections: (1) physical layout of classrooms, (2) print environment, (3) books and reading area, (4) writing resources, (5) listening area, (6) literacy toys, (7) cultural diversity, (8) literacy in other centers, (9) numerals, and (10) curriculum theme.
- **The Snapshot of Classroom Activities (OMLIT-SNAP)** is a time-sampled description of classroom activities and groupings, integration of literacy in other activities, and language in the classroom. It has two sections. The Environment section describes the number of children and adults present, as well as the type of adult (staff, parents). The Activities section describes activities that are taking place. Then, for each activity, the observer records the number of children and adults in that activity, whether any adult or child is talking, whether he/she is speaking English or another language, whether any literacy materials are used (text, writing, letters, and if the teacher is singing with the children (distinguished on the measure because of its potential as a phonological awareness/oral language support).

- **The Read Aloud Profile (OMLIT-RAP)** is a description of staff behavior when reading aloud to children. (In CLIO, observers conducted a RAP observation each time staff read aloud to at least two children (up to five RAPs could be completed per classroom, if staff read five books to two or more children during the half-day observation.) The RAP records adult behavior during the read-aloud session in four categories: (1) pre-reading (set-up) behavior, (2) behavior while reading the book, (3) post-reading behavior, and (4) the language the adult uses when talking to children during the read aloud. The RAP also records: (1) role of the adult involved in the read-aloud (e.g., teacher, aide, etc.), (2) characteristics of the book being read, and (3) number of children involved in the read-aloud. The RAP also includes three quality indicators that summarize particular aspects of the read-aloud: (1) the degree to which the adult introduces and contextualizes new vocabulary to support children's learning; (2) the extent to which the adult uses open-ended questions that invite children to engage in prediction, imagination, and/or rich description; and (3) the quality and length of any post-reading book-related activities that the adult organizes (beyond oral discussion).
- **The Classroom Literacy Instruction Profile (OMLIT-CLIP)** is a time-sampled description of instructional methods in literacy activities. At specified intervals, the observer determines if the lead teacher is involved in a literacy activity. If so, then the observer follows the teacher and records data on any literacy activities that occur in a 10-minute period. For each activity, the observer codes seven characteristics of the activity: (1) type of literacy activity, (2) number of children involved, (3) language spoken by teacher, (4) language spoken by children, (5) instructional style, (6) text support, and (7) literacy knowledge afforded. If the teacher is not involved in a literacy activity, the observer then observes the aide/assistant, if there is one, to record any other literacy activities occurring. If there are no literacy activities, the observer records the type of nonliteracy activity in which the teacher was involved. If the literacy activity involves discussion between the adult and children, three quality indicators are completed that rate the cognitive challenge and depth of the discussion and the extent of the children's participation.

- **The Quality of Instruction in Language and Literacy (OMLIT-QUILL)** is a rating of classroom staff on the use of high-quality instructional practices in six areas of language and literacy instruction: (1) opportunities to engage in literacy/language activities, (2) opportunities to engage in writing activities, (3) activities to promote letter/word knowledge, (4) activities to promote oral language, (5) activities to promote functions/features of print, and (6) activities to promote understanding of sounds. Additional items are scored if there are English language learners in the classroom. These items rate the extent to which the instructional practices integrate the ELL children and their language in the classroom literacy activities.

Administration of the OMLIT

All of the early childhood education classrooms in all CLIO projects—treatment and control—were observed three times over the course of the study (fall 2004, prior to the intervention, spring 2005 and spring 2006). At each observation point, classrooms were observed for half a day (typically, 3.5 hours). In half-day programs, the observations typically took place in the morning, starting when the children arrived and ending when children were dismissed. In school-day or extended-day programs, the observations also took place in the morning, starting when the children arrived and ending at lunchtime.

Observations were scheduled for days when there were no field trips or other special events planned and when the lead teacher was expected to be present. If the lead teacher was unexpectedly absent on the observation day, the observation was rescheduled.

Training of Classroom Observers

The OMLIT observations were conducted by trained observers. The observers were college graduates who were either working in the field of educational research or had some experience in preschool settings. Prior to each wave of classroom observations, candidate observers were hired and attended a centralized multi-day training on the OMLIT. The training curriculum included extended practice coding

Table E-1. The Observation Measure of Language and Literacy Instruction (OMLIT)

Measure	Focus	Description	Schedule	Detail
Classroom Description	Overall classroom environment	Classroom context: # and ages of children enrolled/present, staff present; staff languages; home language of children; classroom theme	One per classroom. Language(s) of instruction completed at end of observation; other data completed at start of observation	Context, children and adults present, children's home language, classroom theme, languages of instruction
Classroom Literacy Opportunities Checklist (OMLIT-CLOC)	Overall classroom environment	Inventory of classroom literacy resources available to children. 56 items, rated on 3-point scale	One completed by end of observation	Completed on ongoing basis during observation
Snapshot of Classroom Activities (OMLIT-SNAP)	Overall classroom environment	Children & adults in each class activity, focus on literacy, integration of literacy materials in other activities, languages spoken in classroom, # adults/children	Time-sampled: a record of the classroom completed every 15 minutes over the observation session. 4 SNAPS are completed in each hour of observation.	Each Snapshot is a picture of what is happening in the classroom at each 15-minute mark. Snapshot represents a moment in time, rather than the range of activities and interactions across the 15-minute period. The first snapshot is conducted when the first child arrives, and subsequent snapshots are conducted every 15 minutes after the first began.
Read Aloud Profile (OMLIT-RAP)	Behavior of adults in classroom who read aloud to children	Instructional practices when staff read aloud to children, focus on dialogic reading practices	Event-sampled: whenever staff is reading aloud to at least 2 children, up to 5 RAPs coded per observation.	When shared book reading begins, RAP is coded continuously until end of read aloud.

Table E-1. The Observation Measure of Language and Literacy Instruction (OMLIT) (continued)

Measure	Focus	Description	Schedule	Detail
Classroom Literacy Instruction Profile (OMLIT-CLIP)	Behavior of teacher, assistants directing or involved in literacy activities w/children	Literacy activities in classroom and instructional methods used by staff	Time-sampled: a 10-minute observation conducted every 15 minutes over the observation session, following each Snapshot. 4 CLIPs are completed in each hour of observation.	After each SNAP, a CLIP is coded continuously over the subsequent 10 minutes. During the 10 minutes, if a focal adult becomes involved in a literacy-related activity with children, information about the literacy activity is recorded. The record of literacy-related activities continues for 10 minutes and could involve multiple literacy events.
Quality of Instruction in Language and Literacy (OMLIT-QUILL)	Instruction practices of adults for literacy activities w/children	Rating of frequency and quality of literacy instruction and support for children's language/literacy development (print knowledge, phonological sensitivity, understanding the features/functions of print, print motivation)	Once based on entire observation	During the observation, notes are kept about relevant literacy activities in the classroom. At the end of the observation session, the notes, along with information coded on the Snapshot, the RAP and the CLIP, are used to complete ratings of QUILL items.

SOURCE: Goodson, B.D., Layzer, C.J., Smith, W.C., and Rimdzius, T. (2004 and 2006). *Measures Developed as Part of the Even Start Classroom Literacy Interventions and Outcomes (CLIO) Study*, under contract number ED-01-CO-0120, as administered by the Institute of Education Sciences, U.S. Department of Education.

paper-and-pencil vignettes and videotaped clips of real-time and unscripted read-alouds. As part of the training, candidate observers conducted a practice observation and a reliability observation of a real-time early childhood classroom.

To be allowed in the field, candidate observers were required to (1) pass formal paper-and-pencil and/or video reliability tests for five of the six OMLIT measures and (2) achieve a criterion level of inter-rater agreement with a master coder,

when both the candidate observer and the master coder conducted a real-time observation of the same early childhood education classroom. Table E-2 summarizes the reliability tests that each candidate observer had to pass before being hired to conduct the actual observations of CLIO classrooms. The purpose of estimating this type of inter-rater reliability was to determine the accuracy of individual candidate observers for purposes of training and hiring.

Table E-2. Formal Inter-Rater Reliability Tests for Candidate Observer

Measure	Paper-and pencil reliability test	Live observation reliability test
Classroom Description	At least 75% exact agreement with master coding of written description of example classroom context	At least 75% exact agreement with Classroom Description completed by master coder ^a in the field reliability observation
Classroom Literacy Opportunities Checklist (OMLIT-CLOC)		At least 75% exact agreement with CLOC coding completed by master observer, for literacy-related materials and resources in the classroom in the field reliability observation
Snapshot of Classroom Activities (OMLIT-SNAP)	At least 75% exact agreement with master coding of written descriptions of classroom activities (type of activities, #'s of staff and children)	At least 75% exact agreement with SNAP coding completed by master observer, for activities and groupings of staff and children observed in the field reliability observation
Read Aloud Profile (OMLIT-RAP)	At least 75% exact agreement with master coding of videotapes of real-time live read-alouds	At least 75% exact agreement with RAP coding completed by master observer, for any staff-child reading observed in the field reliability observation
Classroom Literacy Instruction Profile (OMLIT-CLIP)	At least 75% exact agreement with master coding of written descriptions of language and literacy activities with children	At least 75% exact agreement with CLIP coding completed by master observer, for any staff-directed literacy activity observed in the field reliability observation
Quality of Instruction in Language and Literacy (OMLIT-QUILL)	At least 75% exact agreement with master coding of written descriptions of classroom activities that support development of early literacy skills of children oral language development	At least 75% exact agreement with QUILL coding completed by master observer, for frequency and quality of literacy instruction and support for language/literacy development during the field reliability observation

^aMaster coders were OMLIT developers and trainers.

Paper-and-Pencil Tests. Reliability was assessed via paper-and-pencil tests for three of the OMLIT measures—the SNAP, the CLIP, and the QUILL. The OMLIT developers prepared and coded in advance written scenarios describing classroom events (“criterion” coding). Although this type of paper-and-pencil test does not

simulate the “live” action in a classroom, it does provide a measure of how well candidate observers understood the coding definitions for the various activities and specialized literacy data.

The accuracy of coding of written scenarios was determined by comparing the coding done by the candidate observers to the criterion coding of the same scenarios. Agreement was defined by the exact match between a candidate observer’s coding and the master coding. To calculate observers’ percentage agreement with the master coding, a procedure was used to reduce inflation in inter-rater reliability due to chance agreement. Observers were credited for correctly coding instances of behaviors and were penalized for incorrectly coding behaviors that did not occur. Observers were not credited for abstaining from marking behaviors that did not occur. For each candidate observer, percentage agreement was calculated for each code individually, and an aggregate overall percentage agreement across multiple codes was also calculated. Overall percentage agreement was used to judge whether each observer had met the criterion for employment. On the SNAP, the mean overall percentage agreement across the codes was 98 percent. On the CLIP, the mean overall percentage agreement across the codes was 79 percent. On the QUILL, the mean overall percentage agreement across the codes was 76 percent.

Coding Videotaped Clips. For the OMLIT RAP, candidate observers practiced using seven videotapes of adults reading to preschool children and then coded three videotape clips for reliability purposes. All of the video clips were pre-coded in advance by the OMLIT developers (the “criterion” coding). Agreement was assessed for instructional behavior during the read aloud and for the overall quality ratings. Agreement was calculated as the exact match between the coding of the video clip by the candidate observer and the criterion coding, averaged across the three video clips. The mean percentage agreement was 96 percent on the adult behaviors during the read-aloud and 94 percent on the quality ratings.

Reliability Observations. All candidate observers who achieved at least 75 percent agreement on all of the paper-and-pencil tests and the coding of the videotape clips then conducted a full observation in an early childhood classroom, alongside a master coder. By comparing the coding of the candidate observer and the master coder, another measure of inter-rater agreement could be calculated, based on a real-life administration of the OMLIT. Using the same methods to calculate inter-rater

reliability, candidate coders had to achieve at least 75 percent agreement to be allowed into the field.

Field Reliability of the OMLIT

In each wave of data collection, a small number of experienced OMLIT observers each was paired with the trained classroom observers to measure the field-based reliability of the OMLIT, using actual CLIO classrooms.² The purpose of these dual observations was to assess the reliability of the OMLIT itself, as opposed to the accuracy of individual observers. Tables E-3 to E-7 show the inter-rater agreement that was calculated for each OMLIT measure, based on the dual observations in the field. For each of the OMLIT measures, reliability was calculated as the percentage exact agreement between the coding of the two observers in the same CLIO classroom.

Classroom Literacy Opportunities Checklist (CLOC). Scores on the CLOC include an average score across all items and average scores on each of eight types of literacy resources. The average percentage agreement for the CLOC ratings was 82 percent across all items and 75 percent or higher on each of the eight subscores (table E-3).

Quality of Instruction in Language and Literacy (QUILL). Inter-rater agreement was computed separately for frequency and quality of the different types of language/literacy activities. As shown in table E-4, average percentage agreement on the coding of the frequency of literacy activities ranged from 67 percent to 88 percent, with overall average agreement of 76 percent. On the quality ratings, the average agreement ranged from 68 percent to 94 percent across the separate items and the overall rating. The overall average agreement was 72 percent for the separate items (excluding the overall score, which had the highest average agreement—94 percent). This reliability was based on exact matches between the ratings given by the two coders.

² Each of the master observers was assigned to conduct paired observations with three to four of the newly trained observers. The pairing was done randomly, i.e., master observers were paired with the newly trained observers in a “blind” match. Once the pairings were made, the master observer accompanied the newly trained observer to one of the CLIO classrooms to which the observer had been assigned.

Table E-3. Inter-Rater Agreement on the Classroom Literacy Opportunities Checklist (OMLIT-CLOC)

Codes^a (# items)	Average % agreement^b
Total across all items (56)	82
Physical layout of classrooms (5)	90
Print environment (8)	77
Books/reading area/listening area (16)	77
Writing resources (5)	81
Literacy toys and materials (7)	83
Cultural diversity (3)	76
Literacy in other centers (3)	75
Curriculum theme (9)	76

^aEach item rated on a scale of 1-3.

^bBased on exact agreement between coding from 90 paired observations conducted over the three waves of data collection—baseline (fall 2004) and spring 2006 and 2006.

Table E-4. Inter-Rater Agreement on the Quality of Instruction in Language and Literacy (OMLIT-QUILL)

Codes	Average % agreement^b
Frequency ratings for literacy activities^a	
All literacy/language activities	82
Writing activities	88
Activities to promote letter/word knowledge	82
Activities to promote oral language	67
Activities to promote functions/features of print	67
Activities to promote understanding of sounds	71
Quality ratings for instruction in literacy^a	
All language and literacy activities	94
Writing activities	85
Activities to promote letter/word knowledge	85
Activities to promote oral language	87
Activities to promote functions/features of print	68
Activities to promote understanding of sounds	69

^aFrequency rated on a scale of 1-4; quality rated on a scale of 1-5.

^bBased on exact agreement between coding from 90 paired observations conducted over the three waves of data collection—baseline (fall 2004) and spring 2006 and 2006.

Read Aloud Profile (RAP). Agreement on instructional behavior across the entire book reading (before, during, and after reading a book) was 90 percent (table E-5). For the quality indicators, average inter-rater reliability was 84 percent agreement for exact matches between the ratings given by the two coders.

Table E-5. Inter-Rater Agreement on the Read Aloud Profile (OMLIT-RAP)

Codes	Average % agreement^a
Instructional strategies	
Pre-reading strategies used by teacher	89
Reading strategies used by teacher	85
Post-reading strategies used by teacher	97
Pre-reading, reading, post-reading codes combined	90
Quality indicators	
Vocabulary links	83
Adult use of open-ended questions	83
Depth of post-reading activity	85

^aBased on exact agreement between coding from 90 paired observations conducted over the three waves of data collection—baseline (fall 2004) and spring 2006 and 2006.

Classroom Literacy Instruction Profile (CLIP). The CLIP involves a two-stage coding protocol. First, the observer determines if any of the classroom staff are involved in a literacy activity. Then, if there is a literacy activity, the observer codes seven characteristics of the literacy activity, based on observing the full literacy activity or reaching the end of the 10-minute CLIP observation period. If no staff member is involved in a literacy activity during the observation segment, the observer records only the type of nonliteracy activity that the classroom is involved in. Inter-rater reliability for the CLIP was computed for whether the two coders agreed on *when* a staff member was involved in a literacy activity, and, for observation segments where the two raters agreed that the teacher was involved in a literacy activity, the average percentage agreement on the seven characteristics of the literacy activity. On average, the inter-rater agreement on the occurrence of a literacy event was 85 percent (table E-6). In the instances when both observers identified a literacy activity, observers agreed 95 percent of the time on the type of literacy activity and the literacy knowledge afforded. The inter-rater agreement on the quality ratings averaged 92 percent across the two ratings for Cognitive Challenge and Depth of Discussion (based on exact matches between the ratings given by the two coders).

Snapshot of Classroom Activities (SNAP). The Environment section on the SNAP includes a count of the numbers of children and adults present in the classroom. Agreement was above 80 percent for exact matches on each of the codes on the Environment (table E-7). On the Activities section of the SNAP, children and adults are allocated across 15 activities. The average inter-rater agreement on type of activities

Table E-6. Inter-Rater Agreement on the Classroom Literacy Instruction Profile (OMLIT-CLIP)

Codes	Average % agreement^a
Occurrence of literacy event	
Staff involved in literacy event or not	85
Rate of literacy activities (total # literacy events/# CLIPs)	94
Characteristics of literacy events	
Type of literacy activity	98
Number of children involved	96
Language spoken by teacher	97
Language spoken by children	97
Instructional style	97
Text support	98
Literacy knowledge afforded	96
Quality ratings	
Cognitive challenge	92
Depth of discussion	93

^aBased on exact agreement between coding from 90 paired observations conducted over the three waves of data collection—baseline (fall 2004) and spring 2006 and 2006.

Table E-7. Inter-Rater Agreement on the Snapshot of Classroom Activities (OMLIT-SNAP)

Codes	Average % agreement^a
Environment	
Total # children present	88
Type of adults present: teachers/aides	81
Type of adults present: other	87
All codes on Environment	85
Activities	
Type of activity	82
Number of children in activity	57
Number of teachers in activity	80
Number of aides in activity	81
Number of other adults in activity	91
Print integrated in other activities	89
Any language by child/adult in each activity	71
Talk	
Any adult talk	100
Any child talk	100
Any adult/child talk	100

^aBased on exact agreement between coding from 90 paired observations conducted over the three waves of data collection—baseline (fall 2004) and spring 2006 and 2006.

occurring in the SNAP was 82 percent. Coders agreed on the presence of print materials in activities 89 percent of the time. Coders agreed perfectly on whether the provider was talking to the children during the SNAP and in what language.

OMLIT Outcome Constructs

For the purposes of analysis, constructs were derived from the multiple OMLIT measures to correspond to key elements of the classroom that are being manipulated by the CLIO curricula. These included constructs for the four key components of emergent literacy as well as adequacy of language and literacy resources in the classroom. These five constructs were developed by identifying *on a conceptual basis* the set of individual teaching practices from across the OMLIT battery of measures that, on the basis of the research, are linked to children’s development in that domain. These constructs are shown in table E-8, along with the specific teaching behaviors that make up each.

The teaching behaviors within each domain are on different scales—some are proportions of time, some are counts, etc. Therefore, to build scales, we converted all of the individual teaching behaviors into standard scores with a mean of 0 and a standard deviation of 1. We then examined the internal consistency of the resulting scales using the Cronbach’s alpha statistic. The results of these analyses are shown in table E-9. The constructs with the fewest behaviors had the lowest internal consistency, as would be expected. We also computed Cronbach’s alphas for the final CLIO constructs (derived from the reliability analyses) in a second OMLIT data set from 162 child care center classrooms in Miami (Layzer, Layzer, Goodson, and Price, 2006). Cronbach’s alphas in the second sample of classrooms were similar to those for the CLIO data.

Correlations among the five constructs are shown in table E-10. The adequacy of literacy resources in the classrooms is related to only one of the constructs about instructional practices—print knowledge. This is not unexpected, since the ratings on both “support for print knowledge” and literacy resources depend to some extent on the availability of print/text materials in the classroom. The four constructs for level of instructional support for different components of emergent literacy were correlated with each other at statistically significant levels that ranged from 0.15 to 0.39.

Table E-8. ECE Classroom Outcomes and Component OMLIT Variables

Outcome construct	Instructional behaviors (OMLIT variables) in each construct	Field reliability of behavior ^a (%)
Support for oral language	Read Aloud Profile (OMLIT-RAP):	
	• Time in reading across observation period	74
	• # books read in observation period	99
	• % read alouds with different supports for comprehension of text	87
	• % read alouds with open-ended questions	83
	• Quality of open-ended questions, vocabulary supports, post-reading	92
	Literacy activities (OMLIT-CLIP):	
	• Time on oral language activities	88
	• % oral language activities with small groups	79
	• Quality of teacher/child discussion	93
	Rating of frequency/quality of support for oral language development (OMLIT-QUILL):	
	• Frequency of oral language activities	67
	• Quality of oral language activities	87
Support for phonological awareness	Read Aloud Profile (OMLIT-RAP):	
	• % read alouds with discussion of sounds	91
	Literacy activities (OMLIT-CLIP):	
	• Time on sounds	90
	• % activities on sounds with small groups	89
	Classroom activities (OMLIT-SNAP):	
	• Proportion classroom time on sounds, singing	77
	Rating of frequency and quality of support for phonological awareness (OMLIT-QUILL):	
	• Frequency of activities to support phonological awareness	71
	• Quality of activities to support phonological awareness	69
Support for print knowledge	Read Aloud Profile (OMLIT-RAP):	
	• % read-alouds with discussion of print concepts	83
	Classroom activities (OMLIT-SNAP):	
	• Proportion classroom time in activities with text, letters	80
	• Proportion classroom time in activities with writing (copying, emergent)	78
	• % text, writing activities in small groups	76
	• % activities with print involved	89
	Literacy activities (OMLIT-CLIP):	
	• Time on print knowledge activities	92
	• % print knowledge activities with small groups	87
	• Time on emergent writing activities	80
	• Time on copying/tracing activities	82
	• % print knowledge activities with small groups	88

Table E-8. ECE Classroom Outcomes and Component OMLIT Variables (continued)

Outcome construct	Instructional behaviors (OMLIT variables) in each construct	Field reliability of behavior^a (%)
Support for print knowledge (cont'd)	Rating of frequency and quality of support for print knowledge (OMLIT-QUILL):	
	• Frequency of activities to support writing	88
	• Quality of activities to support writing	85
	• Frequency of print knowledge activities to support print knowledge	82
	• Quality of activities to support print knowledge	85
	• Frequency of activities to support understanding functions/features of print	67
	• Quality of activities to support understanding functions/features of print	68
Support for print motivation	Read Aloud Profile (OMLIT-RAP):	
	• % read alouds with support for print motivation	95
	• Number of RAPs	99
	• Number of minutes of reading aloud	75
	Literacy activities (OMLIT-CLIP):	
	• Time on activities involving print motivation	94
Adequacy of literacy resources	• % activities on print motivation with small groups	93
	Literacy resources in the classroom (OMLIT-CLOC):	
	• environmental print	77
	• text materials	78
	• writing resources	81
	• rich, integrated theme	76
	• literacy manipulatives	82
	• integration of print in other centers	71

^aBased on exact agreement between paired observers in 90 paired observations conducted over the 3 waves of data collection—baseline (spring 2004), spring 2005, and spring 2006.

Table E-9. Internal Consistency and Inter-Rater Reliability of ECE Outcome Constructs, Spring 2004

Construct	# Items in construct	Cronbach's alpha		CLIO inter-rater reliability ^c (N=33 paired observations)
		CLIO ^a (N=199 classrooms)	Miami ^b (N=162 classrooms)	
Support for oral language	14	.84	.80	.87
Support for print knowledge	16	.84	.82	.89
Support for phonological awareness	4	.58	.61	.83
Support for print motivation	5	.73	.72	.89
Adequacy of literacy resources in class	7	.75	.73	.80

^aData from observations conducted in spring 2004 in CLIO classrooms.

^bIn Miami child care center data, Cronbach's alpha derived from same set of OMLIT variables that are included in the final version of constructs derived from the CLIO data.

^cBased on exact agreement between coding from paired observations.

Table E-10. Correlations Among ECE Outcome Constructs, Spring 2004^a

	Support for print knowledge	Support for phonological awareness	Support for print motivation	Adequacy of literacy resources
Support for oral language	0.15*	0.12	0.37***	0.10
Support for print knowledge		0.39***	0.32***	0.25***
Support for phonological awareness			0.19**	0.13
Support for print motivation				0.10

^a N = 199 Even Start projects

*= $p < .05$, **= $p < .01$, ***= $p < .001$

Finally, each of the instructional constructs was aggregated to the project level and was re-scaled to a more convenient metric. The aggregation enabled the analyses to be conducted at the project level. The re-scaling enhanced the interpretability of results by ensuring that the value from any subgroup could be easily interpreted relative to the 2004 control group mean. After aggregation and re-scaling, the 2004 control group mean and standard deviation for each construct was 50, and 10, respectively. If, for example, a study group mean from 2005 was 52, then that score

could be interpreted as being 2/10 standard deviation units higher than the 2004 control group mean. Aggregation and re-scaling of the constructs entailed the following seven steps:

1. A project-level value for each item³ was calculated as the mean of the item across all classes nested within a project.
2. The 2004 control group mean and standard deviation was calculated for each item.
3. Each item was standardized by subtracting the 2004 control group mean and dividing by the 2004 control group standard deviation of the item.
4. Each of the five constructs was created as the sum of relevant standardized items.
5. The 2004 control group mean and standard deviation was calculated for each OMLIT construct.
6. Each construct was standardized by subtracting the 2004 control group mean and dividing by the 2004 control group standard deviation of the construct. After completion of this step, the 2004 control group mean and standard deviation were zero and one, respectively.
7. Each construct was rescaled by multiplying by 10, and adding 50. After completion of this step, the 2004 control group mean and standard deviation were 50 and 10, respectively. The resulting scores are such that the 2005 control group mean and the study group means from 2004 and 2005 are interpreted relative to the 2004 control group mean.

³ The items (specific teaching behaviors) corresponding to each OMLIT construct are described in table E-8.



Observation Booklet Spring 2006

Observation Measures of Language and Literacy Instruction (OMLIT)

Authors:

Barbara D. Goodson, Ph.D.

Carolyn J. Layzer, Ph.D.

W. Carter Smith, Ph.D.

Abt Associates Inc.

Tracy Rimdzius

U.S. Dept. of Education

February – March 2006

Measures developed as part of the Even Start Classroom Literacy Interventions and Outcomes (CLIO) Study,
under contract ED-01-CO-0120, as administered by the Institute of Education Sciences, U.S. Department of Education
Prime contractor: Westat

*For information about the OMLIT measures,
Contact Tracy Rindzius, U.S. Dept of Education
tracy.rindzius@ed.gov
202-208-7085*

Early Childhood Education Classroom Description



Part 1: Identifying Information

Name ID#	Date of Observation ____ / ____ / ____ mm dd yyyy Time Observation Began ____ : ____ am pm Time Observation Ended ____ : ____ am pm
Observer: _____ Project / Classroom: _____	

Part 2: Staff List (teachers /assistants /regular staff)

Staff Name	Staff Role	Staff ID#
(1) _____	(1) _____	(1) _____
(2) _____	(2) _____	(2) _____
(3) _____	(3) _____	(3) _____
(4) _____	(4) _____	(4) _____

Part 3: Classroom Context

Number of Children Enrolled (by age group)

_____	Infants & toddlers (under 3 yrs)
_____	Preschool (3 – 5 yrs)
_____	School age (6+ yrs)
_____	Total

Primary Home Language of the Children (% should add to 100)

_____ %	English only
_____ %	Spanish only /bilingual Spanish-English
_____ %	Other language (1) only /bilingual other language–English
	<i>Specify language (1):</i> _____
_____ %	Other language (2) only /bilingual other language–English
	<i>Specify language (2):</i> _____

Any Children with Diagnosed Special Needs?

☐ Yes
 ☐ No
 ☐ Don't know

Classroom Theme: Any theme, topic, unit (for day, for week, for month) that class is focusing on?

☐ No theme

Describe theme _____

Early Childhood Education Classroom Description



Part 4: Post-Observation Summary

Language of Instruction of Staff: Select one response for each staff member present during observation

Staff (1) ID# _____	Staff (2) ID# _____	Staff (3) ID# _____	Staff (4) ID# _____	
☐	☐	☐	☐	English only
☐	☐	☐	☐	Spanish only
☐	☐	☐	☐	Primarily English, some Spanish
☐	☐	☐	☐	Primarily Spanish, some English
☐	☐	☐	☐	English and Spanish equally
☐	☐	☐	☐	Primarily English, another language (<i>specify</i> : _____)
☐	☐	☐	☐	Primarily another language, some English (<i>specify</i> : _____)
☐	☐	☐	☐	English and another language equally (<i>specify</i> : _____)
☐	☐	☐	☐	English and multiple other languages (<i>specify</i> : _____)
☐	☐	☐	☐	Other combination of languages (<i>specify</i> : _____)
☐	☐	☐	☐	NA—not in classroom on day of observation

Language of Other Adults in Classroom: Select one response for the other adults that were present

N/A:

- ☐ No other adults in classroom
- ☐ No ELL children in the class
- ☐ Other adults in classroom speak only English

Other adults in classroom speak additional language(s)

- ☐ Adults speak language(s) of **all** ELL groups in classroom
- ☐ Adults speak language(s) of **some** but not all ELL groups in classroom
- ☐ Adults speak languages of **none** of ELL groups in classroom

Other adults include adults who are not regular staff but who work with the children, such as parent volunteers and other center staff such as director. Do NOT include special visiting musicians, health care professionals, etc.

Indication that Observation Day Was Not Typical:

Describe any special events or unusual circumstances that indicate that the day was not typical: _____

No RAPs:

- ☐ No RAPs coded because no read-alouds occurred

Snapshot of Classroom Activities



Number	Children Present in Classroom
	Infants & toddlers (under 3 yrs old)
	Preschool (3 – 5 yrs old)
	School Age (5+ yrs old)
	Total Children (all ages)

☐ **Whole Group Activity (1-14)**

Number	Staff Present in Classroom
	Teacher
	Assistant teacher/ Aide
Number	Other Adults Present in Classroom
	Parent dropping off/picking up/visiting
	Volunteer (parent or other)
	Other adult (visitor, other center staff)
	Total Adults (staff and other adults)

No Snapshot Coded:

☐ RAP

☐ Gross motor group activity

☐ Other reason

Snapshot #1

Time: ____ : ____

am pm

Talk	E	S	O	None	E	S	O	None
Ad/C	☐	☐	☐	☐	☐	☐	☐	☐
C/C	☐	☐	☐	☐				

Activity		# of Children and Adults in Activity				Literacy Resources/Activities	
		Children	Teachers	Aides	Other Adults		
						L	Describe
1	Reading/text/alphabet/vocabulary (w/ print)					L	
1a	Alphabet/numerals					L	
1b	Sounds/singing						
1c	Oral language/vocabulary (no print)						
2	Emergent writing/copying/tracing					L	
3	Science/nature						
4	Math concepts/attributes/colors						
5	Dramatic play						
6	Creative play						
7	Block play						
8	Fine motor play						

Activity		# of Children and Adults in Activity				Literacy Resources/Activities	
		Children	Teachers	Aides	Other Adults		
						L	Describe
9	Sensory play						
10	Meeting time						
11	Games with rules						
12	TV/video/ computer						
13	Gross motor play						
14	Other activity						
15	Meals/routines/ management						
Not in class						Notes:	

1. Reading/ text/ vocabulary (with print)

Looking at books or pictures, adult reading aloud, children reading together without adult, emergent reading (pretending to read), shared reading activities. Listening to stories on audiotape or CD. Teaching children new vocabulary words **with print support** for vocabulary (e.g., printed word).

1a. Alphabet/ numerals

Recognizing letter/numeral forms, letter/sound correspondence. **Always involves print.**

1b. Sounds/ singing

Sounds of words with no print. All singing (may or may not have print, e.g., words of song displayed).

1c. Oral language/ vocabulary

Discussions, new concepts and vocabulary with either **no print** or **no print emphasis**.

2. Emergent writing/ copying/ tracing

Child(ren) writing, includes pretend writing, scribbling, invented spelling. Child dictation to teacher. Tracing letter or number templates. Practice in correctly writing numerals/distinguishing numerals. **Always involves print**

3. Science/ nature

Formal and informal communication of science or nature. Science examples: astronomy, working with pets, collecting leaves, feeding pets, magnets, health & safety.

4. Math concepts/ attributes/ colors

Formal and informal communication of math concepts, attributes, or colors. Shapes, counting, measuring, patterns, amount. Identifying and matching non-geometric shapes (animals, familiar objects). Identifying and matching colors and color names.

5. Dramatic play

Pretend or make-believe play; dress-up, playing with dolls; assigning roles; zooming cars and trucks.

Note: Includes acting out stories/playing with puppets, figures of people/animals, and stuffed animals in pretend environments.

6. Creative play

Arts and crafts – creating visual art (painting, drawing, sculpting clay & play dough, cutting and pasting). **Note:** Always code PlayDoh® as “*Creative play*.”

Music – instruments, formal and informal movement/dance activities.

7. Block play

All building with blocks and other large building materials.

Note: Once construction is done, and blocks are part of a completed pretend environment with cars, trucks, figures of people, code as “*Dramatic play*.”

8. Fine motor play

Manipulation of materials, such as puzzles, stringing beads, sewing cards, woodworking, LEGOs®, Lincoln Logs, interconnecting building pieces.

9. Sensory play

Manipulating sand, water, and textured materials such as beans, rice, shaving cream, where objective is learning about qualities of materials and not constructing a particular object.

10. Meeting time

Routines or daily rituals as part of group or circle time.

Includes activities such as calendar, day of the week, weather, the day's activities, etc. Also includes discussions, such as sharing by children with questions from teacher, peers.

11. Games with rules

Playing board games, card games, and video games (e.g., Nintendo, Game Boy, Play Station) that are **not** explicitly educational.

12. TV/ video/ computer

Watching commercial television programs, video tapes/DVDs or computer programs which may or may not be educational.

Note: If activity involves computer, circle “Computer” in description box.

13. Gross motor play

Large muscle play – active outdoor play and indoor physical activity (tunnels, gymnastics). Include outdoor walks here (e.g., walking to and from a destination such as the library).

Note: Code organized dance/ movement activity as “*Creative play*.”

14. Other activity

Special activities that are not part of the regular activities on list, such as special events/destinations, field trips, student assessments (e.g., school assembly, library, fire station, ice cream store).

Note: The activity should be specified in the description box.

15. Meals/ Routines/ Transitions/ Conversation/ Management/ No activity

Meals/ snacks: Engaged in the act of eating a meal/snack, and/or meal/snack preparation and clean up.

Routines/ transitions: Arriving/departing, napping/sleeping, physical care/ hygiene (including first aid, toileting), setting-up or cleaning-up of activities/materials, lining-up.

Conversation/ management: Any talking or interaction between adult and child, between children, or between adults outside of a listed activity. Conversation may be positive or negative. Examples: adult managing a child's behavior, comforting a child, or chatting. Children may be interacting in nonproductive ways.

Uninvolved/ administration: Not involved in any activity listed above and not interacting with another person. Child roaming aimlessly around classroom, having a tantrum, otherwise unengaged. Teacher/other adult doing administrative work, monitoring overall classroom activity from a distance.

Not in class

List children and staff who have left classroom and, if known, where they have gone. Do **NOT** include these children or staff in counts at top of form.

Classroom Literacy Instruction Profile MLIT-CLIP

No CLIP coded: ☐ Gross motor group activity ☐ Other reason	Context: ☐ Child-selected activity ☐ RAP ☐ Teacher-selected activity ☐ Meal/snack/routine	Staff ID# _____ ☐ All Non-Literacy Activity (10 min.)	CLIP #1 Expected Start Time ____ : ____ am pm
---	---	--	---

Other Staff in Literacy Events: ☐ Staff ID _____ Column A: ____ ☐ Staff ID _____ Column A: ____ ☐ Staff ID _____ Column A: ____	Describe Literacy Event ("Title"): _____	Literacy Event #1 Start ____ : ____ am pm End ____ : ____ am pm
---	---	---

Literacy Event #1

A. Literacy Activity (circle one)	B. Literacy Knowledge Afforded (circle one)	C. Teacher's Instructional Style (circle one)	D. Text Support/Context for Literacy Instruction (circle all that apply; <i>specify</i>)	E. Number of Children in Activity w/Teacher (circle one E1-E4; E5 if applicable)	F. Child(ren)'s Talk (circle all that apply)	G. Teacher's Involvement With Child(ren) (circle all that apply)
1 Teacher presents information or explains about print/text/ language, or Reads text to children (when not RAP)	1 Sounds —NO PRINT used— (phonological awareness)	1 Performing/presenting (child(ren) listen/watch)	1 Whole/connected text (e.g., book, story) _____	1 One Child	1a <u>Talk with teacher:</u> English 1b Spanish, or other	1a <u>Teacher language:</u> English 1b Spanish, or other language
2 Teacher writing	2 Letters Shape & name; numerals; some sight words (e.g., child's own name) (orthographic awareness)	2 Directing child(ren)'s response (usually closed-ended questions/commands)	2 Isolated text—sentence, word, letter(s), numeral(s), or word part _____	2 Two children	2a <u>Talk with peers:</u> English 2b Spanish, or other	2a <u>Focus of language:</u> To group 2b To one child 2c To children in turn
3a <u>Focused oral language:</u> Language games, rhymes, songs, storytelling [NOT songs for transitions or management]	3 Sounds & letters together —PRINT used— (phonological-orthographic awareness)	3 Making suggestions/ offering materials	3 Environmental print/ functional text _____	3 Small group (3-5 children)	3a <u>Talk with group:</u> English 3b Spanish, or other	3 No teacher language
3b Discussion, dialogue	4 Comprehension of text/ story	4 Observation/listening (teacher within 3 feet of child(ren); 1 or more minutes observing and/or listening)	4 TV/video _____	4 Large group (6+ children)	4 No child language	
3c Circle time routines: weather, calendar, job chart, daily songs		5 Group discussion (> 1 child or 1 child & group; various question types) * Quality Rating	5 Computer/interactive instructional technology _____	5 Whole group		
4 Child reading/emergent reading; shared reading (teacher with one child, taking turns reading)	5 Vocabulary and background knowledge	6 Individual discussion (single child only; various question types) * Quality Rating	6 Picture(s), representative object(s), illustration(s), or icon(s) _____			
5a <u>Child writing:</u> Emergent writing	6 Conventions of text, format, emergent writing NOT letter shapes (Print awareness)		6			
5b Copying, tracing	6		6			
6 Child tagging/matching (e.g., word to object, word to icon, word to picture, letter to picture, letter to letter, word to word, etc.)	7 Print motivation		7 Connected to classroom theme			
	8 Oral communication/ listening skills		8 Other _____			
7 Literacy assessment	☐ For B1-B8: Check if content is incorrect					9 None

Quality Indicators for Discussion/Dialogue on OMLIT-CLIP

ONLY IF C5-C6 CODED						#1: Literacy Event #2
1. Cognitive Challenge ^a	☐ 1 (Minimal)	☐ 2	☐ 3 (Moderate)	☐ 4	☐ 5 (Extensive)	
(Focus on teacher's language)	Content/topics of discussion/conversation are mostly (more than 75% of time) about management or routines.		Content/topics are mix of management/routines and other topics (about 50% of each).		Content/topics of discussion/conversation are primarily rich or abstract ^b —about the physical world or about human motivation.	
	Adult asks <i>no</i> open-ended questions and only 1 closed-ended question that requires little thought or dialogue.	0 open-ended and 2+ closed-ended	Adult asks <i>one</i> open-ended question and may ask multiple closed-ended questions that require little thought or dialogue.	2 open-ended	Adult asks <i>three or more</i> open-ended questions that require children to use imagination, make predictions, generate hypotheses, etc.	
	Adult does not extend the conversation beyond the here-and-now.		Adult extends conversation beyond the here-and-now for a <i>single topic</i> —talks about past/future, ideas, language, or books.		Adult extends conversation beyond the here-and-now for <i>more than one topic</i> —talks about past/future, ideas, language, or books.	
Code feature below <i>only if</i> discussion relates to children's experience; otherwise code NA						
	Briefly mentions an experience children have had (or a related book/class activity), but does not elaborate or invite children's responses OR does not relate topic to children's experiences at all.		Relates activity to experiences children have had (in or out of class); uses this to encourage/invite discussion among children. Checks to see that most children recall the experience that is referred to but does not go beyond yes/no answers from children.		Relates activity to experiences children have had (in or out of class); uses this to encourage/invite discussion among children. Goes beyond checking to be sure that most/all children recall the related event—asks for comments, details that describe that event.	
2. Depth of Discussion	☐ 1 (Minimal)	☐ 2	☐ 3 (Moderate)	☐ 4	☐ 5 (Extensive)	
(Focus on teacher's attempts to draw out child's response)	Adult poses a couple of questions or responds to a child's comment minimally, without engaging children in multiple turn-taking opportunities; OR the discussion may start but lasts only briefly before adult commences another activity or engages in management-related discourse.		Engages children in discussion marked by turn-taking, but children's turns are short (yes/no, single word, etc.).		Engages children in discussion marked by multiple turn-taking (including longer child turns, more than 3 child turns).	
	Adult <i>never</i> elaborates on children's remarks (by providing details, description, or explanation) AND <i>never</i> asks children to elaborate (by asking for details, description, explanation).	Elaborates or asks for elaboration <i>once</i>	Adult elaborates on children's remarks (by providing details, description, or explanation) OR asks children to elaborate (by asking for details, description, explanation) <i>twice</i> .	Elaborates or asks for elaboration <i>3 times</i>	Adult elaborates on children's remarks (by providing details, description, or explanation) OR asks children to elaborate (by asking for details, description, explanation) <i>four or more times</i> .	

a Cognitive Challenge: Defined by extent of cognitive abstractions and cognitive extensions in discussion/dialogue with children

b Cognitive abstraction = talking about generalizations about the physical world and about human motivations)

Cognitive extension = talking about past and future or about things that are "non-present"—e.g., weekend plans, talk about books, or about ideas

Classroom Literacy Instruction Profile MLIT-CLIP

No CLIP coded: ☐ Gross motor group activity ☐ Other reason	Context: ☐ Child-selected activity ☐ RAP ☐ Teacher-selected activity ☐ Meal/snack/routine	Staff ID# _____ ☐ All Non-Literacy Activity (10 min.)	CLIP #1 Expected Start Time ____ : ____ am pm
---	---	--	---

Other Staff in Literacy Events: ☐ Staff ID _____ Column A: ____ ☐ Staff ID _____ Column A: ____ ☐ Staff ID _____ Column A: ____	Describe Literacy Event ("Title"): _____	Literacy Event #2 Start ____ : ____ am pm End ____ : ____ am pm
---	---	---

Literacy Event #2

A. Literacy Activity (circle one)	B. Literacy Knowledge Afforded (circle one)	C. Teacher's Instructional Style (circle one)	D. Text Support/Context for Literacy Instruction (circle all that apply; <i>specify</i>)	E. Number of Children in Activity w/Teacher (circle one E1-E4; E5 if applicable)	F. Child(ren)'s Talk (circle all that apply)	G. Teacher's Involvement With Child(ren) (circle all that apply)
1 Teacher presents information or explains about print/text/ language, or Reads text to children (when not RAP)	1 Sounds —NO PRINT used— (phonological awareness)	1 Performing/presenting (child(ren) listen/watch)	1 Whole/connected text (e.g., book, story) _____	1 One Child	1a <u>Talk with teacher:</u> English 1b Spanish, or other	1a <u>Teacher language:</u> English 1b Spanish, or other language
2 Teacher writing	2 Letters Shape & name; numerals; some sight words (e.g., child's own name) (orthographic awareness)	2 Directing child(ren)'s response (usually closed-ended questions/commands)	2 Isolated text—sentence, word, letter(s), numeral(s), or word part _____	2 Two children	2a <u>Talk with peers:</u> English 2b Spanish, or other	2a <u>Focus of language:</u> To group 2b To one child 2c To children in turn
3a <u>Focused oral language:</u> Language games, rhymes, songs, storytelling [NOT songs for transitions or management]	3 Sounds & letters together —PRINT used— (phonological-orthographic awareness)	3 Making suggestions/ offering materials	3 Environmental print/ functional text _____	3 Small group (3-5 children)	3a <u>Talk with group:</u> English 3b Spanish, or other	3 No teacher language
3b Discussion, dialogue	4 Comprehension of text/ story	4 Observation/listening (teacher within 3 feet of child(ren); 1 or more minutes observing and/or listening)	4 TV/video _____	4 Large group (6+ children)	4 No child language	
3c Circle time routines: weather, calendar, job chart, daily songs		5 Group discussion (> 1 child or 1 child & group; various question types) * Quality Rating	5 Computer/interactive instructional technology _____	5 Whole group		
4 Child reading/emergent reading; shared reading (teacher with one child, taking turns reading)	5 Vocabulary and background knowledge	5	5			
5a <u>Child writing:</u> Emergent writing	6 Conventions of text, format, emergent writing NOT letter shapes (Print awareness)	6 Individual discussion (single child only; various question types) * Quality Rating	6 Picture(s), representative object(s), illustration(s), or icon(s) _____			
5b Copying, tracing	6	6	6			
6 Child tagging/matching (e.g., word to object, word to icon, word to picture, letter to picture, letter to letter, word to word, etc.)	7 Print motivation		7 Connected to classroom theme			
	8 Oral communication/ listening skills		8 Other _____			
7 Literacy assessment	☐ For B1-B8: Check if content is incorrect			9 None		

Quality Indicators for Discussion/Dialogue on OMLIT-CLIP

ONLY IF C5-C6 CODED						#1: Literacy Event #3
1. Cognitive Challenge ^a	☐ 1 (Minimal)	☐ 2	☐ 3 (Moderate)	☐ 4	☐ 5 (Extensive)	
(Focus on teacher's language)	Content/topics of discussion/conversation are mostly (more than 75% of time) about management or routines.		Content/topics are mix of management/routines and other topics (about 50% of each).		Content/topics of discussion/conversation are primarily rich or abstract ^b —about the physical world or about human motivation.	
	Adult asks <i>no</i> open-ended questions and only 1 closed-ended question that requires little thought or dialogue.	0 open-ended and 2+ closed-ended	Adult asks <i>one</i> open-ended question and may ask multiple closed-ended questions that require little thought or dialogue.	2 open-ended	Adult asks <i>three or more</i> open-ended questions that require children to use imagination, make predictions, generate hypotheses, etc.	
	Adult does not extend the conversation beyond the here-and-now.		Adult extends conversation beyond the here-and-now for a <i>single topic</i> —talks about past/future, ideas, language, or books.		Adult extends conversation beyond the here-and-now for <i>more than one topic</i> —talks about past/future, ideas, language, or books.	
Code feature below <i>only if</i> discussion relates to children's experience; otherwise code NA						
	Briefly mentions an experience children have had (or a related book/class activity), but does not elaborate or invite children's responses OR does not relate topic to children's experiences at all.		Relates activity to experiences children have had (in or out of class); uses this to encourage/invite discussion among children. Checks to see that most children recall the experience that is referred to but does not go beyond yes/no answers from children.		Relates activity to experiences children have had (in or out of class); uses this to encourage/invite discussion among children. Goes beyond checking to be sure that most/all children recall the related event—asks for comments, details that describe that event.	
2. Depth of Discussion	☐ 1 (Minimal)	☐ 2	☐ 3 (Moderate)	☐ 4	☐ 5 (Extensive)	
(Focus on teacher's attempts to draw out child's response)	Adult poses a couple of questions or responds to a child's comment minimally, without engaging children in multiple turn-taking opportunities; OR the discussion may start but lasts only briefly before adult commences another activity or engages in management-related discourse.		Engages children in discussion marked by turn-taking, but children's turns are short (yes/no, single word, etc.).		Engages children in discussion marked by multiple turn-taking (including longer child turns, more than 3 child turns).	
	Adult <i>never</i> elaborates on children's remarks (by providing details, description, or explanation) AND <i>never</i> asks children to elaborate (by asking for details, description, explanation).	Elaborates or asks for elaboration <i>once</i>	Adult elaborates on children's remarks (by providing details, description, or explanation) OR asks children to elaborate (by asking for details, description, explanation) <i>twice</i> .	Elaborates or asks for elaboration <i>3 times</i>	Adult elaborates on children's remarks (by providing details, description, or explanation) OR asks children to elaborate (by asking for details, description, explanation) <i>four or more times</i> .	

a Cognitive Challenge: Defined by extent of cognitive abstractions and cognitive extensions in discussion/dialogue with children

b Cognitive abstraction = talking about generalizations about the physical world and about human motivations)

Cognitive extension = talking about past and future or about things that are "non-present"—e.g., weekend plans, talk about books, or about ideas

Classroom Literacy Instruction Profile MLIT-CLIP

No CLIP coded: ☐ Gross motor group activity ☐ Other reason	Context: ☐ Child-selected activity ☐ RAP ☐ Teacher-selected activity ☐ Meal/snack/routine	Staff ID# _____ ☐ All Non-Literacy Activity (10 min.)	CLIP #1 Expected Start Time ____ : ____ am pm
---	---	--	---

Other Staff in Literacy Events: ☐ Staff ID _____ Column A: ____ ☐ Staff ID _____ Column A: ____ ☐ Staff ID _____ Column A: ____	Describe Literacy Event ("Title"): _____	Literacy Event #3 Start ____ : ____ am pm End ____ : ____ am pm
---	---	---

Literacy Event #3

A. Literacy Activity (circle one)	B. Literacy Knowledge Afforded (circle one)	C. Teacher's Instructional Style (circle one)	D. Text Support/Context for Literacy Instruction (circle all that apply; <i>specify</i>)	E. Number of Children in Activity w/Teacher (circle one E1-E4; E5 if applicable)	F. Child(ren)'s Talk (circle all that apply)	G. Teacher's Involvement With Child(ren) (circle all that apply)
1 Teacher presents information or explains about print/text/ language, or Reads text to children (when not RAP)	1 Sounds —NO PRINT used— (phonological awareness)	1 Performing/presenting (child(ren) listen/watch)	1 Whole/connected text (e.g., book, story) _____	1 One Child	1a <u>Talk with teacher:</u> English 1b Spanish, or other	1a <u>Teacher language:</u> English 1b Spanish, or other language
2 Teacher writing	2 Letters Shape & name; numerals; some sight words (e.g., child's own name) (orthographic awareness)	2 Directing child(ren)'s response (usually closed-ended questions/commands)	2 Isolated text—sentence, word, letter(s), numeral(s), or word part _____	2 Two children	2a <u>Talk with peers:</u> English 2b Spanish, or other	2a <u>Focus of language:</u> To group 2b To one child 2c To children in turn
3a <u>Focused oral language:</u> Language games, rhymes, songs, storytelling [NOT songs for transitions or management]	3 Sounds & letters together —PRINT used— (phonological-orthographic awareness)	3 Making suggestions/ offering materials	3 Environmental print/ functional text _____	3 Small group (3-5 children)	3a <u>Talk with group:</u> English 3b Spanish, or other	3 No teacher language
3b Discussion, dialogue	4 Comprehension of text/ story	4 Observation/listening (teacher within 3 feet of child(ren); 1 or more minutes observing and/or listening)	4 TV/video _____	4 Large group (6+ children)	4 No child language	
3c Circle time routines: weather, calendar, job chart, daily songs		5 Group discussion (> 1 child or 1 child & group; various question types) * Quality Rating	5 Computer/interactive instructional technology _____	5 Whole group		
4 Child reading/emergent reading; shared reading (teacher with one child, taking turns reading)	5 Vocabulary and background knowledge	6 Individual discussion (single child only; various question types) * Quality Rating	6 Picture(s), representative object(s), illustration(s), or icon(s) _____			
5a <u>Child writing:</u> Emergent writing	6 Conventions of text, format, emergent writing NOT letter shapes (Print awareness)		7 Connected to classroom theme			
5b Copying, tracing			8 Other _____			
6 Child tagging/matching (e.g., word to object, word to icon, word to picture, letter to picture, letter to letter, word to word, etc.)	7 Print motivation		9 None			
	8 Oral communication/ listening skills					
7 Literacy assessment	☐ For B1-B8: Check if content is incorrect					

A. Literacy Activity

- 1. Teacher presents information (or explains about) print / text / language, or reads text to children):**
 - presenting information to children
 - explaining about **print / text**: The teacher is explaining something about a written text to child(ren).
 - explaining about the **sounds** of language: The teacher explains something about the sounds a letter or word makes, for example.
 - explaining about **alphabet**: The teacher points out or explains about individual letters or the whole alphabet. This can include talking about numerals, which are symbols that represent numbers.
 - reading**: The teacher is reading a text—it could be a poem, a story that the class has composed together, a book (>5 RAPs), or some other non-book text.
- 2. Teacher writing:** The teacher is writing letter(s), word(s), or sentence(s) while child(ren) observe. This includes teacher writing as child dictates (e.g., caption).
- 3. Focused oral language:**
 - 3a)** The teacher is leading language play involving rhyming, alliteration, attending to rhythm of speech (clapping out syllables); other games involving oral language; nursery rhymes; storytelling.
 - 3b)** The teacher is conducting some kind of focused oral language activity (not social conversation or comforting) such as an in-depth discussion, interviewing, or dialogue.
 - 3c)** The teacher is leading circle time routine activities such as calendar, job chart, weather, daily songs, etc.
- 4. Child reading/emergent reading; shared reading:**

Child(ren) responding to text, e.g., reading word on flashcard, looking at a book, pretending to read, reading. Shared reading is when the teacher is with one child, and they are reading (a book) together.
- 5. Child writing:**
 - 5a) emergent writing:** Child(ren) writing, includes pretend writing, scribbling, invented spelling (e.g., teacher taking dictation from child). Also includes child(ren) creating written messages, lists, plans for the day (emergent or conventional writing)
 - 5b) copying, tracing:** Child(ren) tracing or copying print (e.g., teacher may be monitoring or assisting).
- 6. Child tagging / matching:** Child(ren) placing printed tags onto corresponding (referent) objects, icons, or locations. Can be matching word to object/icon/picture; letter to picture, letter to letter, word to word, etc.
- 7. Literacy assessment:** Teacher is conducting a literacy assessment.

B. Literacy Knowledge Afforded

- 1. Sounds (phonological awareness):** Instruction targets sounds (including, but not limited to, sounds of words or letters). Instruction does not involve use of print.
- 2. Letters & words (orthographic awareness):** Instruction focuses on learning to recognize the form (shape) of letters and sight words. It can include naming the letters to identify them and learning to form the letters (i.e., to write them, trace them, copy them, etc.). It can include learning to recognize numerals (circle if numerals) and one's own name.
- 3. Sounds and letters (phonological-orthographic awareness):** Instruction links letters with sounds. Helps children begin to decode written words.
- 4. Comprehension of text/story:** Instruction emphasizes meaning of text, skills for understanding text. Answering questions about text content, summarizing content, predicting, formulating questions about text.
- 5. Vocabulary and background knowledge:** Instruction builds knowledge of new concepts and new words, builds on what children already know to give new vocabulary meaning. Teacher introduces new concepts/words in play/ other activity (w/ or w/out explicit definition); compares/ contrasts new concepts or words to known ones.
- 6. Conventions of text, format, emergent writing:** Instruction points out **format &/or purpose of print, signs, symbols** in environment; teacher explains print conventions (e.g., title, author on book cover; capital letters at beginning of sentences, period at end; spaces between words). Includes beginning-middle-end concept and emergent writing but **NOT** letter shapes/names [=B2].
- 7. Print motivation:** Activity serves to make reading or other interaction with print enjoyable so that child would want to engage in it again. Code **emergent reading** opportunities here. However, if another code in this dimension fits, the other code should be used.
- 8. Oral Communication / Listening Skills:** Purpose of activity is to encourage children to use language to communicate, to use new vocabulary, and/or to learn to listen carefully/ attentively.

C. Instructional Styles

- 1. Performing/Presenting:** Teacher reads, speaks, without expecting response from children at that time.
- 2. Directing child(ren)'s response:** Teacher tells students what to do, explains or demonstrates, with little conversation from children. Includes recitation and/or call-and-response (class responds to closed-ended questions in unison, in a back and forth manner), and singing.
- 3. Making suggestions, offering materials:** The teacher suggests ways of extending activities (e.g., to include or extend literacy activity) or offers materials (e.g., literacy materials).

- 4. Observation/Listening:** Teacher is observing or listening to child(ren) engaged in literacy activity for 1 minute or more; teacher is within 3 feet of child(ren) in activity.
- 5. Relates to child's experience:** Teacher attempts to relate the topic to something in the children's experience (as in the read aloud question type, "distancing"). "Distancing" questions attempt to help children link the topic to something they already know about. For example, "Remember on our trip to the firehouse when we saw the fire truck? What did the firemen use to reach the top windows of the building?"
- 6. Group Discussion:** Teacher talks with 2 or more students, encourages input from small group of children or whole class; includes responsive, dialogic conversation; may involve lots of interaction and children responding to each other. Questions may be a mixture of open-ended, recall, etc. There are multiple conversational turns.
- 7. Individual Discussion:** Teacher talks with one child gathering knowledge about the child; extending the child's knowledge—includes higher-level questioning, a variety of question types.

D. Text Support for Literacy Activity/Context

- 1. Whole/Connected text:** The text consists of 3 or more connected sentences (story, expository text, song). This can be a book.
- 2. Isolated text:** The target of instruction is a: sentence (Text is no longer than 2 sentences), word (a single word or set of unconnected words is the target of instruction.), a letter or letters, numeral(s), or parts of a single word (e.g., syllable(s), onset-rime, phoneme(s)).
- 3. Environmental print/ Functional text:** Items of environmental/functional print such as poster, sign, menu, catalog, classroom calendar, chore chart, sign-up sheet used as focus of literacy activity.
- 4. TV/Video:** The focus of the literacy activity is a television show or segment, teacher participating in TV/video viewing and/or commenting, directing attention to TV/video content.
- 5. Computer / Interactive instructional technology:** A computer or other interactive instructional technology provides textual, visual, and/or auditory support for the literacy activity.
- 6. Picture(s), representative object(s), illustration(s), or icon(s):** Pictures or object(s) that represent a concept or thing, illustrations, or icons in room is/are the focus of literacy activity; teacher engaged with object(s), pictures, illustrations, or icons; teacher directing attention to object(s), pictures, illustrations, or icons as part of the literacy activity.
- 7. Connected to classroom theme:** The activity is connected to a current classroom theme (e.g., sea life, construction, the universe).
- 8. Other**
- 9. None:** No text, TV, or objects being used. Note: this is possible if the focus is on sound, for example.

OMLIT-CLIP Rules

A. PRE-Reading (set-up)		B. Reading		C. POST-Reading (extension)		D. Adult Reading Book		E. Adult Language with Children		F. Number of Children Reading (circle one of F1-4; also F5 if applies)		G. Book Characteristics (circle one for each)	
(circle all that apply)		(circle all that apply)		(circle all that apply)		(circle all that apply)		(circle all that apply)		(circle one of F1-4; also F5 if applies)		(circle one for each)	
1	Guides book choice; discusses children's book choice(s)	1	Tracks print/discusses English print conventions	1	Answers children's questions about the book or related topics	1	Teacher	1	English	1	One child	1a	<u>Type of book</u>
2	Points to features of the book such as the title, illustrations, author	2	Uses props/dramatic voices/gestures			2	Assistant/ Aide	2	Spanish	2	Two children	1b	Picture book
2a	Discusses/defines concepts of print such as the title, illustrator, author	3	Directs children's attention to illustrations/text/story (e.g., asks questions about; discusses/expands on meaning; offers new information)	2	Expands on children's comments about the book or illustrations	3	Other Adult	3	Other language	3	Small group (3-5 children)	1c	Alphabet book
3	Reminds children of similar books they have read or that they have read same book before	4	Comments on sounds, letters, and/or sound-letter links in the book	3	Comments on sounds, letters, and/or sound-letter links in the book	Vocabulary & Supports:						1d	Counting book
4	Comments on sounds, letters, sound/letter links, or tells children to listen and look for them in the book	5	Highlights new book-related vocabulary	4	Reviews/reinforces book-related vocabulary with or without print reference							1e	Chapter book
5	Introduces book-related vocabulary	6	Relates the book to other activities in class, class theme									2a	Reference book
6	Relates the book to other activities in class, class theme	7	Expands on children's comments about the book	5	Summarizes/retells the story without child involvement							2b	Yes No
7	Talks about events and/or features to listen, look for in the book	8	Answers children's questions about the book or related topics	6	Summarizes/retells the story with child involvement							Open-ended Questions:	
8	Introduces background information related to the book (book (with or without child input)	9	Has children join in reading/completing text on their own or as a group (choral reading)	7	Asks for recall of information about the book	4	Large group (6+ children)	3a	<u>Language</u>				
9	Narrates/tells the story in advance of reading	10	Asks recall questions about earlier parts of the book	8	Asks book-related questions about children's experiences outside of classroom activities	5	Whole class	3b	English only				
10	Relates the book to children's experiences outside of classroom activities	11	Relates the book to children's experiences/Asks book-related questions about children's experiences outside of classroom activities	9	Asks book-related <u>open-ended</u> questions (requires prediction, expanded response, thinking, and/or analysis)			3c	Spanish only				
11	No pre-reading experiences or activities (without any codes 1-10)	13	Picture walk	10	Organizes post-reading book- related activity (beyond oral discussion)			3d	Eng & Spanish				
		14	Reads text straight through (without any codes 1 – 13)	11	No post-reading activities or extension occurs (without any codes 1-10)				3e	Other language			
												☐	Eng & other language
													Book is read in different language
													<u>Words/page</u>
												4a	0 words
												4b	1 word
												4c	2-10 words
												4d	>10 words
												5a	<u>Book on tape</u>
												5b	Yes No
													<u>Related to class theme</u>
												6a	Yes
												6b	No
												6c	Don't know

Quality Indicators for OMLIT-RAP

RAP #1					
1. Story-related vocabulary <i>Code as "1" if no A5, B5, or C4 is circled. Code item as "1" if no new vocabulary introduced.</i>	☐ 1 = Minimal	☐ 2	☐ 3 = Moderate	☐ 4	☐ 5 = High
	Some story-related vocabulary words are introduced/discussed, but the definition of one or more of the words is misleading or wrong.	1 story-related vocabulary word and no comprehension supports.	One story-related vocabulary word is introduced or discussed and the definition is accurate AND <u>At least one</u> of the following comprehension supports is given for the word: - A picture, gesture, or other concrete visual aid is used; - The word is linked to a rich network of related words or concepts.	2 story-related vocabulary words with partial comprehension supports.	At least 2 story-related vocabulary words are introduced or discussed and the definition of each vocabulary word is accurate AND Both of the following comprehension supports are given for each word: - A picture, gesture, or other concrete visual aid is used; and - Each word is linked to a rich network of related words or concepts.
2. Adult use of open-ended questions <i>Code item as "1" if no open-ended questions (no A10, B11, B12, C8 or C9 circled).</i>	☐ 1 = Minimal	☐ 2	☐ 3 = Moderate	☐ 4	☐ 5 = High
	Adult poses only one open-ended question and does not provide opportunity for children to respond to question (child not given time to respond, or adult moves on after child has responded).	1 open-ended question and time for children to respond.	Adult poses two open-ended questions and provides opportunity for children to respond to one but not both of the questions.	2-3 open-ended questions and time for response to at least 2.	Adult poses at least four open-ended questions and consistently shows interest in/actively encourages children's responses (e.g., pausing for children, restating question, calling on particular children, acknowledging children's response).
3. Depth of Post-Reading <i>Code item as "1" if no C1-C10 is circled.</i>	☐ 1 = Minimal	☐ 2	☐ 3 = Moderate	☐ 4	☐ 5 = High
	No post-reading extension or activities. (Post-reading coded as C11)	Relates to the book but lasts LESS than 5 minutes	Discussion and/or activity that - Relates to the story/book but does not extend its meaning or comprehension AND - Lasts at least 5 minutes.	Extends comprehension and lasts LESS than 10 minutes	Discussion and/or activity that - Extends the meaning of the text and reinforces comprehension of the story/book AND - Lasts at least 10 minutes.

A. PRE-Reading

- Guides book choice; discusses children's book choice(s):** Adult encourages children to choose the book; talks about their choice with them. Helps them make appropriate choice.
- Points to features of the book such as the title, illustrations, author:** Points to title, author, illustrator, or illustration on front of book (or points to chapter title in a chapter book)
- Discusses/defines concepts of print such as the title, illustrator, author:** Defines, describes meaning of concepts of print such as title, author, illustrator, or illustration.
- Reminds children of similar books they have read or that they have read same book before:** Calls attention to books by the same author or illustrator, on same topic, etc. OR reminds children they've already read same book before. Ex: "What was another book that we read about ducks?" or "...by Eric Carle?"
- Comments on sounds, letters, sound-letter links, or tells children to listen and look for them in the book:** Talks about sounds they will hear in the story, especially sounds they may have been learning about in class. Or talks about letters they will see in the book, especially letters they have been learning about. Ex: "During the story, when you hear the 'buh' sound, raise your hand." Or "This story has a lot of words that begin with the letter 'g.' Let me know when you see one."
- Introduces book-related vocabulary:** Highlights or explains new vocabulary. Ex: "This book is about a fish called a 'sunfish.' Sunfish have fins. Fins are what they use to move around in the water. When we read the book, you will see pictures of sunfish and we can pick out their fins." Does NOT include concepts of print such as illustrator, author, etc.
- Relates the book to other activities in class, class theme:** Calls attention to the book's relation to class activities or theme. Ex: "This duck likes to eat fish. What does our pet turtle like to eat?" Or, "Remember last week when we went to the fire station? This story is about firemen like the ones we met."
- Talks about events and/or features to listen, look for in the book:** Helps children anticipate things that will happen in the book. Ex: "At the end we'll talk about all the different things that the caterpillar likes to eat. What do we think his favorite food is?"
- Introduces background information related to the book (with or without child input):** Describes what the book is about. Ex: "This book is about a birthday party that Little Bear has with all his forest friends." May or may not invite child discussion.
- Narrates/tells the story in advance of reading:** Recites all or major parts of the book (e.g., nursery rhyme in book based on rhyme) before actually reading the book.
- Relates the book to children's own experiences outside of classroom activities:** Links book to children's experiences outside of class. For example: "Have you ever fed the ducks in the park before? What sound did they make?" or "What kinds of things do you like to do on a rainy day?"
- No pre-reading experiences or activities (without any codes 1-10):** The adult may alert children to the reading activity but does not provide any of the above-listed experiences/activities.

B. Reading

- Tracks print/discusses English print conventions:** Adult moves finger along the page below the line of print or points sequentially to words in text while reading. Discusses punctuation, directionality of English print.
- Uses props / dramatic voices / gestures:** Uses props (e.g., hand puppets, stuffed animals, items in the story), gestures, or different voices to tell story.
- Directs children's attention to illustrations/text/story (e.g., asks questions about; discusses/expands on meaning; offers new information):** Points to, or in some way calls children's attention to the book's illustrations, details of the illustrations. Engages children in (brief) discussion about the meaning of text or illustrations and/or offers new information about the story that may not be written in the text or depicted in illustration (e.g., explaining about something unfamiliar in the text or pictures).
- Comments on sounds, letters, and/or sound-letter links in the story/book:** Calls children's attention to sounds, letters, or sounds and their corresponding letters in the book text.
- Highlights new book-related vocabulary:** Calls attention to, defines, and/or gives examples to help children understand unfamiliar words in the book.
- Relates the book to other activities in class, class theme:** Places the book in context by mentioning the class theme and/or how the book fits into the class activities.
- Expands on children's comments about the book:** When child makes a comment, adult extends by asking child for elaboration or restating child's comment.
- Answers children's questions about the book or related topics:** Allows children to ask questions about the book and then responds to those questions.
- Has children join in reading/completing text on their own or as a group (choral reading):** Pauses and/or indicates to children in some way that they should recite words/numbers, phrases, or longer chunks of the text aloud with the teacher.
- Asks recall questions about earlier parts of the book:** Asks children to recall events, characters, attributes from earlier in the story.
- Relates the book to children's experiences/Asks book-related questions about children's experiences outside of classroom activities:** Extends children's understanding by tapping into their own experiences to help them comprehend the story.
- Asks book-related open-ended questions (requires prediction, expanded response, thinking, and/or analysis):** Probes children's comprehension by asking questions about the story that require children to predict (e.g., "What do you think will happen next? What if...?"); elaborate responses; engage in more thought or analysis of the story.
- Picture walk:** "Walks through" the book without reading text; turns pages and describes aspects of the illustrations, and/or asks children about the illustrations. May or may not "tell" the story.

14. Reads text straight through (without any codes 1-13):

The adult does not engage children in any of the activities or behaviors listed above while reading the story.

C. POST-Reading

- Answers children's questions about the book or related topics:** Teacher reads, speaks, without expecting response from children at that time.
- Expands on children's comments about the book or illustrations:** Allows children to ask questions about the book and then responds to those questions.
- Comments on sounds, letters, and/or sound-letter links in the book:** Calls children's attention to sounds, letters, or sounds and their corresponding letters in the book text.
- Reviews/reinforces book-related vocabulary with or without print reference:** The teacher suggests ways of extending activities (e.g., to include or extend literacy activity) or offers materials (e.g., literacy materials).
- Summarizes/retells the story without child involvement:** Re-tells plot of story to remind children, help children who didn't understand what the story meant.
- Summarizes/retells the story with child involvement:** Involves children in retelling plot of story.
- Asks for recall of information about the book:** Asks children to recall events, characters, attributes from the story just read aloud.
- Asks book-related questions about children's experiences outside of classroom activities:** Extends children's understanding by tapping into their own experiences to help them comprehend the story.
- Asks book-related open-ended questions (requires speculation, expanded response, thinking, and/or analysis):** Probes children's comprehension by asking questions about the book that require speculation, longer or more elaborated responses, more thought, or analysis of the story. Ex: "What do you think would have happened if...?"
- Organizes post-reading book-related activity (beyond oral discussion):** The teacher suggests ways of extending activities (e.g., to include or extend literacy activity) or offers materials (e.g., literacy materials).
- No post-reading activity or extension occurs (without any codes 1-10):** The adult does none of the above-listed extensions or activities after reading the book.

Note: Discussion of concepts of print during post-reading should be coded as A-2a.

Quality of Language and Literacy Instruction OMLIT-QUILL

Item	Frequency Rating	Overall Quality Rating					Item Score
		☐ 1 = Minimal	☐ 2	☐ 3 = Moderate	☐ 4	☐ 5 = High	
1. Opportunities to engage in language and literacy activities ^a Domain = writing, letter/word knowledge, oral language, functions/features of print, sound in words. ^b Small groups = 3-5 children; large groups = 6+ children ^c Rich language = rare vocabulary, extended sentences, new words. Abstract concepts = non-present topics (prediction, analysis)	☐ No opportunities ☐ Minimal (one) opportunity ☐ Moderate number of (a few) opportunities ☐ Extensive number of (many) opportunities	Language and literacy activities <i>rarely/never</i> higher-quality; typically lower-quality, such as worksheets, tracing/copying, recitation, lecture <i>Little/no variety</i> in language and literacy activities provided (only 1 domain of activities) ^a Language and literacy (not solely oral language) <i>rarely/never</i> integrated into activities with goals other than literacy Language and literacy activities <i>rarely/never</i> conducted with children in <u>small groups/individual children</u> ^b Staff work with <i>only a few/a small percentage</i> of the children in language and literacy activities over the day Staff <i>rarely/never</i> use rich language with children, talk about abstract concepts, or talk about language itself ^c Staff <i>rarely/never</i> positive, enthusiastic engaged in language and literacy activities		Language and literacy activities <i>sometimes</i> higher-quality and <i>sometimes</i> lower-quality (about 50% of each) <i>Some variety</i> in language and literacy activities provided (3 domains of activities) ^a Language and literacy (not solely oral language) <i>sometimes</i> integrated into activities with goals other than literacy Language and literacy activities <i>sometimes</i> conducted with <u>small groups/ individual children</u> and sometimes with <u>large groups</u> ^b Staff work with <i>up to half</i> of the children in language and literacy activities over the day Staff <i>sometimes</i> use rich language with children, <i>sometimes</i> talk about abstract concepts, and <i>sometimes</i> talk about language itself ^c Staff <i>sometimes</i> positive, enthusiastic, engaged in language and literacy activities and <i>sometimes</i> not		Language and literacy activities <i>often/consistently</i> higher-quality, such as songs, rhymes, reading aloud, games, extended 1-1 discussions/ dialogue, journals <i>Wide variety</i> in language and literacy activities provided (5 domains of activities) ^a Language and literacy (not solely oral language) <i>often</i> integrated into activities with goals other than literacy Language and literacy activities are <i>often/consistently</i> conducted with <u>small groups/individual children</u> ^b Staff work with <i>most/all</i> of the children in language and literacy activities over the day Staff <i>often/consistently</i> use rich language with children, talk about abstract concepts, and talk about language itself ^c Staff <i>often/consistently</i> positive, enthusiastic, engaged in language and literacy activities	Scores on these features are the average from items 2 – 6.
Snapshots		CLIPS/RAPs		Other Language and Literacy Activities			
ALL LANGUAGE AND LITERACY							

Quality of Language and Literacy Instruction OMLIT-QUILL

Item	Frequency Rating	Overall Quality Rating					Item Score
		☐ 1 = Minimal	☐ 2	☐ 3 = Moderate	☐ 4	☐ 5 = High	
2. Opportunities to engage in writing	☐ No opportunities ☐ Minimal (one) opportunity ☐ Moderate number of (a few) opportunities ☐ Extensive number of (many) opportunities	Writing activities <i>rarely/never</i> higher-quality; usually lower quality, such as worksheets, tracing/copying		Writing activities <i>sometimes</i> higher-quality and <i>sometimes</i> lower-quality (about 50% of each)		Writing activities <i>often/consistently</i> higher-quality, such as emergent writing, captioning, dictation with teacher, writing own name on work, book-making, journals	
^a Children writing on their own = using invented or phonetic spelling and irregular letter forms Note: Writing can include writing numerals, not just letters.		<i>Little/no variety</i> in writing activities provided (either only 1 activity or only 1 type of activity)		<i>Some</i> variety in writing activities provided (3 different types of activities)		<i>Wide</i> variety in writing activities provided (5+ different types of activities)	
		Writing <i>rarely/not</i> integrated into activities with goals other than literacy		Writing <i>sometimes</i> integrated into activities with goals other than literacy		Writing <i>often</i> integrated into activities with goals other than literacy	
		Writing activities <i>rarely/never</i> conducted with children in <u>small groups/individual children</u>		Writing activities <i>sometimes</i> conducted with <u>small groups/ individual children</u> and sometimes with children in large groups		Writing activities are <i>often/consistently</i> conducted with <u>small groups/individual children</u>	
		Staff work with <i>only a few/a small percentage</i> of the children in writing activities over the day		Staff work with <i>some/up to half</i> of the children in writing activities over the day		Staff work with <i>most/all</i> of the children in writing activities over the day	
		In <i>few/no</i> writing activities, writing is done by children themselves rather than by adults		In <i>some</i> writing activities, writing is done by children themselves rather than by adults		In <i>most/all</i> writing activities, writing is done by children themselves rather than by adults	
		Staff <i>rarely/never</i> allow or encourage children to write on their own; ^a <i>usually</i> insist on conventional letter formation/spelling		Staff <i>sometimes</i> allow or encourage children to write on their own ^a and <i>sometimes</i> insist on conventional letter formation/spelling		Staff <i>often/consistently</i> allow or encourage children to write on their own ^a rather than insisting on conventional letter formation/spelling	
Snapshots		CLIPS			Other Writing Activities		
WRITING							

Quality of Language and Literacy Instruction

OMLIT-QUILL

Item	Frequency Rating	Overall Quality Rating					
		☐ 1 = Minimal	☐ 2	☐ 3 = Moderate	☐ 4	☐ 5 = High	Item Score
3. Attention to/promotion of letter/word knowledge ^a	☐ No opportunities ☐ Minimal (one) opportunity ☐ Moderate number of (a few) opportunities ☐ Extensive number of (many) opportunities	Activities promoting letter/word knowledge <i>rarely/never</i> higher-quality; usually lower quality, such as drills, flashcards, worksheets		Activities promoting letter/word knowledge <i>sometimes</i> higher-quality and <i>sometimes</i> lower-quality (about 50% of each)		Activities promoting letter/word knowledge <i>often/consistently</i> higher-quality, such as reading alphabet books, having children write own name, helping child locate classroom job by calling attention to key letters, games such as letter bingo, letter wall	
^a Letter/word knowledge: attention to same/different in letters, names, words; associating letter names and letter shapes; letter-sound matches		<i>Little/no variety</i> in activities to promote letter/word knowledge (either only 1 activity or only 1 type of activity)		<i>Some variety</i> in activities to promote letter/word knowledge (3 different types of activities)		<i>Wide variety</i> in activities to promote letter/word knowledge (5+ different types of activities)	
		Letter/word knowledge <i>rarely/not</i> integrated into activities with goals other than literacy		Letter/word knowledge <i>sometimes</i> integrated into activities with goals other than literacy		Letter/word knowledge <i>often</i> integrated into activities with goals other than literacy	
		Activities to promote letter/word knowledge <i>rarely/never</i> conducted with children in <u>small groups/individual children</u>		Activities to promote letter/word knowledge <i>sometimes</i> conducted with <u>small groups/ individual children</u> and sometimes with children in large groups		Activities to promote letter/word knowledge are <i>often/consistently</i> conducted with <u>small groups/individual children</u>	
		Staff promote letter/word knowledge with <i>only a few/a small percentage</i> over the day		Staff promote letter/word knowledge with <i>some/up to half</i> of the children over the day		Staff promote letter/word knowledge with <i>most/all</i> of the children over the day	
Snapshots		CLIPs/RAPs			Other Examples of Attention to Letter/Word Knowledge		

LETTER/WORD KNOWLEDGE

Quality of Language and Literacy Instruction OMLIT-QUILL

Item	Frequency Rating	Overall Quality Rating						Item Score
		☐ 1 = Minimal	☐ 2	☐ 3 = Moderate	☐ 4	☐ 5 = High		
4. Opportunities/encouragement of oral language to communicate ideas and thoughts <								

Quality of Language and Literacy Instruction

Item	Frequency Rating	Overall Quality Rating						Item Score
		☐ 1 = Minimal	☐ 2	☐ 3 = Moderate	☐ 4	☐ 5 = High		
5. Attention to the functions and features of print^a ^a Functions of print: labeling, naming, categorizing, describing. Features of print: directionality (i.e., print goes from left to right, top to bottom). Note: functional print on display in the classroom is <u>not</u> sufficient; staff must engage in <i>active behaviors</i> to draw children's attention to the functions/features of print	☐ No activities ☐ Minimal (one) activity ☐ Moderate number of (a few) activities ☐ Extensive number of (many) activities	Activities that draw attention to the functions/features of print <i>rarely/never</i> higher-quality; usually lower quality, such as direct instruction in absence of authentic, meaningful text		Activities that draw attention to the functions/features of print <i>sometimes</i> higher-quality and <i>sometimes</i> lower-quality (about 50% of each)		Activities that draw attention to the functions/features of print are <i>often/consistently</i> higher-quality, such as being part of reading aloud, working with authentic print materials.		
		<i>Little/no variety</i> in activities that draw attention to the functions/features of print (either only 1 activity or only 1 type of activity)		<i>Some variety</i> in activities to draw attention to the functions/features of print (3 different types of activities)		<i>Wide variety</i> in activities to draw attention to the functions/features of print (5+ different types of activities)		
		Attention to the functions/features of print is <i>rarely/not</i> integrated into activities with goals other than literacy		Attention to functions/features of print is <i>sometimes</i> integrated into activities with goals other than literacy		Attention to functions/features of print is <i>often integrated</i> into activities with goals other than literacy		
		Activities that draw attention to functions/features of print <i>rarely/never</i> conducted with children in <u>small groups/individual children</u>		Activities that draw attention to functions/features of print <i>sometimes</i> conducted with <u>small groups/ individual children</u> and sometimes with children in large groups		Activities that draw attention to functions/features of print <i>often/ consistently</i> conducted with <u>small groups/individual children</u>		
		Staff work with <i>only a few/a small percentage</i> of the children in activities that draw attention to the functions/ features of print		Staff work with <i>some/up to half</i> of the children in activities that draw attention to the functions/features of print		Staff work with <i>most/all</i> of the children in activities that draw attention to the functions/features of print		
Snapshots		CLIPs/RAPs			Other Examples of Attention to Features/Functions of Print			
FUNCTIONS/FEATURES OF PRINT								

Quality of Language and Literacy Instruction OMLIT-QUILL

Item	Frequency Rating	Overall Quality Rating						Item Score
		☐ 1 = Minimal	☐ 2	☐ 3 = Moderate	☐ 4	☐ 5 = High		
6. Attention to sounds in words ^a throughout the day 								

**Quality of Language and Literacy Instruction
OMLIT-QUILL**

Language and Literacy Strategies with English-Language Learners (ELLs)						☐ No ELL children in classroom	
Item	Frequency	Overall Quality Rating					
		☐ 1 = Minimal	☐ 2	☐ 3 = Moderate	☐ 4	☐ 5 = High	Item Score
7. ELL children intentionally included in activities, conversations	☐ Check and skip item if <u>all</u> children in class are ELLs.	ELL children <i>rarely/never</i> integrated with English-speaking children in activities		ELL children <i>sometimes</i> integrated with English-speaking children in activities and <i>sometimes</i> segregated		ELL children <i>often/regularly</i> integrated with English-speaking children in activities	
		ELL children <i>rarely/never</i> encouraged/ supported to join conversations with English-speaking children		ELL children <i>sometimes</i> encouraged/ supported to join conversations with English-speaking children and <i>sometimes</i> not		ELL children <i>often/regularly</i> encouraged/ supported to join conversations with English-speaking children	
8. Development of both home language(s) and English supported for ELL children		Staff <i>rarely/never</i> positive about having ELL children in the classroom		Staff <i>sometimes</i> positive about having ELL children in the classroom (or only some staff appear positive)		All staff <i>consistently</i> positive about having ELL children in the classroom	
		ELL children <i>never</i> encouraged OR <u>sometimes</u> forced to try using English		ELL children <i>sometimes</i> encouraged but <u>never</u> forced to try using English		ELL children <i>regularly</i> encouraged but <u>never</u> forced to try using English	
		→ Code only if > 1 ELL child in class or ≥ 2 ELLs with same home language ELL children <i>rarely/never</i> encouraged to use their home language with each other; are <i>actively</i> discouraged		→ Code only if > 1 ELL child in class or ≥ 2 ELLs with same home language ELL children <i>rarely</i> encouraged to use their home language with each other but <i>not discouraged</i>		→ Code only if > 1 ELL child in class or ≥ 2 ELLs with same home language ELL children <i>often/consistently</i> encouraged to use their home language with each other	
Snapshots		CLIPs/RAPs			Other Examples		

ELL INSTRUCTION: INCLUSION AND SUPPORT FOR HOME LANGUAGE

**Quality of Language and Literacy Instruction
OMLIT-QUILL**

Language and Literacy Strategies with English-Language Learners (ELLs)					☐ No ELL children in classroom		
Item	Language	Overall Quality Rating					
		☐ 1 = Minimal	☐ 2	☐ 3 = Moderate	☐ 4	☐ 5 = High	Item Score
9. Home language(s) of ELL children integrated into language and literacy activities	☐ No English used (only ELL children's home language used). [Skip item.]	<i>No</i> staff members speak ELL children's home language(s) AND <i>no</i> other adults used as translators		<i>No</i> staff members speak ELL children's home language(s) AND other adults only <i>sometimes</i> used as translators		<i>At least one</i> staff member speaks ELL children's home language(s) OR other adults <i>often/regularly</i> used as translators	
		ELL children's home language(s) <i>rarely/never</i> integrated with English in print-based language and literacy activities		ELL children's home language(s) <i>sometimes</i> integrated with English in print-based language and literacy activities		English/ELL children's home language(s) <i>often/regularly</i> integrated in print-based language and literacy activities	
		English and ELL children's home language(s) <i>rarely/never</i> integrated in oral language activities (songs, rhymes, language games)		English and ELL children's home language(s) <i>sometimes</i> integrated in oral language activities (songs, rhymes, language games)		English/ELL children's home language(s) <i>often/regularly</i> integrated in oral language activities (songs, rhymes, language games)	
10. Language and literacy materials/methods appropriate for ELL children	☐ No English used (only ELL children's home language used). [Code features]	<i>Few/no</i> <u>text</u> materials in language and literacy activities in English and in ELL children's home language(s)		<i>Some</i> <u>text</u> materials in language and literacy activities in English and in ELL children's home language(s)		<i>Most/all</i> <u>text</u> materials in language and literacy activities in English and in ELL children's home language(s)	
		<i>Few/no</i> other print materials in classroom (labels, posters, charts) include both English and home language(s) of ELL children		<i>Some</i> other print materials in classroom (labels, posters, charts) include both English and home language(s) of ELL children		<i>Many/most</i> print materials in classroom (labels, posters, charts) include English and home language(s) of ELL children	
		Methods used to teach English to ELL children <i>rarely/never</i> explicit and contextualized ^a		Methods used to teach English to ELL children <i>sometimes</i> explicit and contextualized ^a and <i>sometimes</i> not		Methods used to teach English to ELL children <i>usually/consistently</i> explicit and contextualized ^a	
		<i>Few/no</i> books available to children/read aloud are appropriate for English language learners ^b		<i>Some</i> books available to children/read aloud are appropriate for English language learners ^b		<i>Many/most</i> books available to children/read aloud are appropriate for English language learners ^b	
Snapshots		CLIPs/RAPs			Other Examples		
ELL INSTRUCTION: INTEGRATION INTO LITERACY ACTIVITIES/MATERIALS							

^a Explicit = emphasis on key words, oral description of actions, events. Contextualized = use of gestures, images, objects
^b Predictable; clearly illustrated; clear, repetitive themes, plots

Classroom Literacy Opportunities Checklist MLIT-CLOC

Physical Layout of Classroom		1	2	3	Notes
1	The room is arranged in distinct centers for different activities (e.g., dramatic play, blocks, books, science, math, art or music).	No distinct centers 1	1-2 distinct centers 2	3+ distinct centers 3	
<i>Note: Two or more combined areas count as one. Circle time area is <u>not</u> a center.</i>					
2	Materials and objects in the room appear well organized (i.e., clearly marked, sorted in a systematic way, and stored in designated areas).	Room cluttered with materials, objects that are not marked, sorted, stored 1	Some materials marked, sorted, stored; others not 2	Most materials, objects are clearly marked, sorted, and stored 3	
3	The classroom layout (i.e., space, furniture placement) is designed so that whole-group, small-group, and individual instruction can easily occur (e.g., a large rug for whole group, tables for small group, or individual desks for individual instruction).	Only 1 group size possible 1	Two group sizes possible 2	All group sizes possible 3	
4	The classroom layout allows children to choose materials and participate in activities independently (e.g., low shelves and easy-to-open cabinets contribute to lots of choice).	Restricts choice 1	Allows some choice 2	Allows lots of choice 3	
5	The space is sufficient in size for the number of children (i.e., children are able to move around freely or set up separate activity areas), with adequate lighting and no noxious or unpleasant odors.	Insufficient space, inadequate light and/or odors 1	Sufficient space and either adequate light or no odors 2	Sufficient space, adequate light, and no odors 3	

Classroom Literacy Opportunities Checklist MLIT-CLOC

Print Environment		1	2	3	Notes
6	Examples of children's writing are on display (i.e., actual letters or words formed by children, or lines, marks or squiggles that appear to imitate print, or formal letter-formation exercises). <i>Note: Writing must be <u>child's own attempt</u>.</i>	No writing on display 1	Only child's own name on work or art on display 2	Writing on display other than just child's own name 3	
7	Child names are matched with photographs of child/child and family or representative object (e.g., animal, color, etc.) posted in classroom. <i>Note: Children must be identified individually.</i>	No names with photographs 1	Names matched to other representative icon 2	Yes, names matched to photographs 3	
8	Examples of functional print that include words/letters are visible (i.e., print used for a purpose) (e.g., calendar, weather chart, job chart, written rules, words of songs, daily menu, daily schedule, etc.). <i>Note: Do not include labels on types of materials here (e.g., "Blocks", "Writing Center").</i>	None 1	1-3 examples 2	4+ examples 3	
9	Examples of functional print that include numerals are visible (e.g., posters, charts, other visual displays that include numbers, such as calendar, daily schedule, birthday dates). <i>Note: Do not count a regular classroom clock. Do not count puzzles, toys.</i>	None 1	1-3 examples 2	4+ examples 3	
10	Labels for groups of toys, materials or areas/centers in the classroom. <i>Note: Labels for the same area or things in 2 or more languages count as 1 label.</i>	None 1	1-6 labels 2	7+ labels 3	
11	Environmental print--posters with print, children's writing, other print materials--are at or below <u>children's eye level</u> . <i>Note: Eye level is defined as the height of an adult's waist.</i>	<u>None</u> at eye level 1	<u>Some</u> at eye level 2	<u>Most or all</u> at eye level 3	
12	There is at least one <u>alphabet</u> chart at or below children's eye level. <i>Note: Chart must include the entire alphabet; an alphabet rug and alphabet table count and should be coded as a "2".</i>	No chart(s) 1	Chart(s) but not at eye level 2	Chart(s) and at eye level 3	
13	There is at least one <u>numeral</u> chart at or below children's eye level. <i>Note: If numerals chart on rug or table, code as "2".</i>	No numeral chart 1	Chart(s) but not at eye level 2	Chart(s) at eye level 3	

Accessible = Can be reached by children on their own.

Classroom Literacy Opportunities Checklist MLIT-CLOC

Literacy Toys and Materials		1	2	3	Notes
14	There are toys and/or materials accessible to children that include <u>words/letters</u> (e.g., toys - puzzles, blocks, board games, card games; materials - magnetic letters, letter stamps, letter cards, word cards). <i>Note: Do not include alphabet chart or alphabet rug.</i>	None 1	1-3 toys, materials 2	4+ toys, materials 3	
15	There are toys and/or materials accessible to children that include <u>numerals</u> (e.g., toys - puzzles, blocks, board games, card games; materials - magnetic numerals, number stamps, numeral molds for sand play, numeral flash cards). <i>Note: Materials must have numerals on them. Do not include numeral chart.</i>	None 1	1-3 toys, materials 2	4+ toys, materials 3	

Classroom Literacy Opportunities Checklist MLIT-CLOC

Books and Reading Area		1	2	3	Notes
16	Total number of books in the classroom <i>If NO BOOKS in the classroom, Skip to item 28</i>	None 1	1-20 books 2	21+ books 3	
17	There is a separate and distinct <u>reading area with books</u> for children to choose from. <i>Note: Area should consist of more than just circle area.</i> <i>If NO area, Skip to item 21</i>	No 1	Books but not separate area 2	Yes, books and separate, distinct area 3	
18	Number of children that the reading area accommodates.	1 child only 1	2-3 children 2	4+ children 3	
19	The reading area is a place where children can sit comfortably to read (e.g., soft furnishings including rugs, pillows, cushions, or couches).	No 1	Sitting area with only a rug 2	Yes 3	
20	Number of books accessible to children <u>in the reading area</u> (i.e., on low shelves, in baskets).	1-7 books 1	8-20 books 2	21+ books 3	
21	There are books on display in <u>open-faced shelving</u> (in the reading area <i>and/or</i> the rest of the classroom). <i>Note: Book covers must face out.</i>	None 1	1-3 books on display 2	4+ books on display 3	
22	Books accessible to children in the classroom are in <u>good condition</u> (i.e., pages are not torn, covers are not missing, print is not faded).	<u>None</u> in good condition 1	<u>Some</u> in good condition 2	<u>Most or all</u> in good condition 3	
23	Books accessible to children in the classroom represent a <u>variety of types</u> (e.g., stories, poetry, non-fiction, wordless, alphabet/counting books, children's magazines (not for cutting up), etc.).	1 type 1	2 types 2	3+ types 3	
24	Books accessible to children in the classroom that present primarily <u>factual</u> information or <u>non-fiction</u> subject matter (e.g., reference books, dictionaries, science, history, biographies, etc.)	None 1	1-3 books 2	4+ books 3	
25	Books accessible to children in the classroom that are in <u>languages</u> other than English, or in both English and other language.	None 1	1-3 books 2	4+ books 3	
26	Books accessible in the classroom cover a range of <u>reading abilities/difficulty levels</u> from very easy to challenging, appropriate for less and more advanced readers (e.g., wordless picture books, 1 word picture books, picture books with 2-10 words, picture books with >10 words, or chapter books).	All 1 level 1	2 levels 2	3 or more levels 3	
27	The classroom has <u>Big Books</u> accessible to children.	None 1	Big Books <u>not</u> accessible to children 2	Big Books accessible to children 3	

Accessible = Can be reached by children on their own.

Classroom Literacy Opportunities Checklist



Listening Area	1	2	3	Notes
28 There is an area for listening to recorded books and listening materials accessible for children's use. <i>Note: Audio equipment must be in working order to count. Radio does not count as listening equipment.</i> <i>If NO area, Skip to item 31</i>	Neither listening area nor listening materials 1	Materials or area, <u>not both</u> 2	Materials and area 3	
29 Amount of listening equipment available. <i>Note: Determine by number of headphones accessible.</i>	Player but no headphones available 1	Sufficient equipment for only 1 child 2	Sufficient equipment for 2+ children 3	
30 There are books to look at along with the recordings of books.	No 1	Books but don't match recordings 2	Yes 3	

Classroom Literacy Opportunities Checklist MLIT-CLOC

Writing Supports		1	2	3	Notes
31	There is a separate writing area (table and chairs) with writing materials accessible to children.	Neither materials nor separate area 1	Materials or area, <u>not both</u> 2	Materials and area 3	
32	There are tools in the classroom accessible to children to help them practice writing <u>words/letters</u> (e.g., stencils, templates, tracing sheets, worksheets/workbooks). <i>Note: tools can be in the writing area or somewhere else in the classroom</i>	None 1	1-3 tools 2	4+ tools 3	
33	There are tools in the classroom accessible to children to help them practice writing <u>numerals</u> (e.g., stencils, numeral templates, tracing sheets, worksheets/workbooks). <i>Note: tools can be in the writing area or somewhere else in the classroom</i>	None 1	1-3 tools 2	4+ tools 3	
34	There are a variety of types of <u>paper</u> in the classroom accessible to children (e.g., large newsprint, colored paper, index cards, dry erase board, etc.). <i>Note: paper can be in the writing area or somewhere else in the classroom</i>	None 1	1-3 types 2	4+ types 3	
35	There are a variety of types of <u>writing utensils</u> in the classroom accessible to children in the classroom (e.g., pencils, crayons, chalk, markers, etc.). <i>Note: writing utensils can be in the writing area or somewhere else in the classroom</i>	None 1	1-3 types 2	4+ types 3	
36	The classroom has a mailbox, message board or other ways for children to learn about writing for a purpose. <i>Note: The mailbox or message board is used as a way for children to exchange messages; it is not a place to store work done by each student (i.e., not regular cubbies).</i>	None 1	1 way of using writing 2	2+ ways of using writing 3	

Accessible = Can be reached by children on their own.

Classroom Literacy Opportunities Checklist



Literacy Materials Outside of the Reading and Writing Areas (If NO centers/areas, SKIP to item 40)		1	2	3	Notes
37	Literacy materials in dramatic play area: There are <u>books and/or other literacy materials</u> in the dramatic play area (e.g., notepads, phone book, shopping lists, pencils, magazines, how-to books, cookbooks). Note: If classroom does not have dramatic play area, code as "1".	No books or literacy materials in dramatic play 1	Either books or literacy materials in dramatic play, <u>not both</u> 2	Books and literacy materials in dramatic play 3	
38	Literacy materials in other areas: There are <u>books and/or other literacy materials</u> in centers/areas other than the book, listening, writing, or dramatic play areas/centers (e.g., art, science, blocks, woodworking, outdoor playground, etc.). *Art: books about art, art journal. *Science area: Notepads, notebooks, places to record observations, paper, pencils. *Block area: Small traffic signs, maps, labeled photos of buildings/construction sites. *Woodworking area: Tool catalogues, home repair magazines. *Outdoors: Paper/markers to make signs, colored chalk, maps on fences, bird & tree guides.	No other areas with books and/or literacy materials 1	1 other area with books and/or literacy materials 2	2+ other areas with books and/or literacy materials 3	
39	There are materials to encourage <u>storytelling</u> in an area <u>other than</u> the dramatic play area (e.g., felt boards with story characters, puppets, story cards).	No other areas have storytelling materials 1	1 other area has storytelling materials 2	2+ other areas have storytelling materials 3	

Diversity in Literacy Materials		1	2	3	Notes
40	<u>Books/text materials</u> accessible in the classroom show a variety of diverse groups of people (e.g., ethnicity, age, disability). Note: Animals and cartoon characters do not count; Must be real people or depictions of humans.	None 1	1-3 examples of diversity in books 2	4+ examples of diversity in books 3	
41	Posters or other visual displays are in a <u>language</u> other than English (e.g., posters, signs, labels, etc.). Note: Count each type of visual display in another language as one instance.	None 1	1-3 posters 2	4+ posters 3	
42	<u>Toys and materials</u> in the classroom represent other cultures, ethnic groups, types of people (e.g., clothing, food, decorative objects, dolls and other toys). Note: Do <u>NOT</u> count books.	None 1	1-3 toys, materials 2	4+ toys, materials 3	

Accessible = Can be reached by children on their own.

Classroom Literacy Opportunities Checklist MLIT-CLOC

Instructional Technology	1	2	3	Notes
43 There are computers for the children to use.	None 1	1-2 computers 2	3+ computers 3	
44 Other interactive technology for children to use (e.g., LeapFrog SchoolHouse™ materials, See & Say™ materials, etc.).	None 1	Sufficient for 1-2 children to use at same time 2	Sufficient for 3+ children to use at same time 3	

Curriculum Theme (Code at end of observation)	1	2	3	Notes
45 Classroom has a curriculum theme. <i>Not truly a content theme:</i> a color, a letter, a number, or something similar. <i>Common concepts:</i> holidays, seasons, dinosaurs. <i>Unusual or rich concept:</i> construction, maps, professions. <i>If there is NO theme, SKIP to item 52</i>	No evidence of a content theme 1	Theme based on common concepts 2	Unusual or rich concept 3	
Describe topic of theme.				
46 Theme is evident in <u>reading/text</u> materials (i.e., books on display in classroom, books read aloud).	No 1		Yes 3	
47 Theme is evident in <u>art projects</u> related to theme.	No 1		Yes 3	
48 Theme is evident in <u>children's work</u> on display in classroom (e.g., art projects, children's writing, photographs, etc.).	No 1		Yes 3	
49 Theme is evident in <u>dramatic play</u> materials (e.g., props, costumes, related books placed in dramatic play area).	No 1		Yes 3	
50 Theme is evident in <u>commercial or teacher-made posters/displays</u> (e.g., charts, writing, graphs, photographs with captions).	No 1		Yes 3	
51 Theme is evident in <u>classroom activities</u> related to current theme (e.g., songs, learning centers, field trips).	No 1		Yes 3	

Classroom Literacy Opportunities Checklist MLIT-CLOC

Literacy Resources Outside of Classroom		1	2	3	Notes
52	Book area/library outside of the classroom.	No book area/library outside class	Book area used but NOT during observation	Book area/library outside of class used during observation	
	<i>If NO book area/library outside of classroom, SKIP to Item 54</i>	1	2	3	
53	Number of books accessible in book area/library outside of the classroom	1-7 books	8-20 books	21+ books	
		1	2	3	
54	Computer lab/computer room or area outside of classroom.	No computer area	Computer area used but NOT during observation	Computer area used during observation	
	<i>If NO computer area outside of classroom, SKIP to End</i>	1	2	3	
55	Number of computers accessible to children outside of the classroom.	1-2 computers	3-4 computers	5+ computers	
		1	2	3	

Accessible = Can be reached by children on their own.

APPENDIX F

PARENTING INSTRUCTIONAL OUTCOME VARIABLES

This appendix provides details on how the outcome variables for parenting education were created. The variables involved are

- PE time spent on child literacy,
- PE time spent on parenting skills, and
- PC time spent interacting on child literacy activities.

All three were developed from observational data. The form used to collect these data was called the PECAP (Parenting Education and Child and Parent Observation). The general rule was that one or two trips by independent observers from the evaluation contractors were made each year to observe parenting classes at each participating Even Start center. Two trips rather than one trip would have been authorized for the most part only when PE and PC classes were not on the same day. Trips usually involved a single day of observation. (For information on missing data, see appendix J, p. J-9.)

The PECAP

The PECAP is an observation measure that describes a PE or PC class in terms of topics discussed, types of activities, size of parent and parent/child groups, and type of instructional approaches used by staff. The PECAP includes the following sections: (1) identifying information about the observer and the class; (2) the classroom context (number of parents and staff, location of class); (3) a class observation taken every 5 minutes during which the observer takes 3 to 5 seconds to scan the class and then uses a structured checklist to record the activities being conducted, parent grouping, instructional leader, and class format at that moment; (4) an observation summary to describe instructional materials, language of instruction, and level of parent engagement; and (5) questions asked of the instructor to determine the number of parents enrolled, their ages and languages, and instructional goals for parents.

Training of Observers of Instructional Services for Parents

About 25 observers were trained in spring 2004 and spring 2005 and 17 in spring 2006 to administer the PECAP in Even Start PE and PC classes. The observers were experienced field staff; most had worked previously for Westat, the Census Bureau, or another research contractor. Many of the observers had taken college classes or had degrees, and some were retired educators. Several of the observers were bilingual, that is, English/Spanish speakers.

Each of the spring training sessions followed the same format. The trainers provided an overview of the instructional activity to be observed, identified key issues related to observing in each setting, and then presented a review of the relevant measure. Following an explanation of each category of activities on the PECAP form and the appropriate codes for each category, the observers practiced coding using several sets of written scenarios, each describing a moment in a typical class or session. Videotapes were used in each training session to provide examples of real classes or sessions. Directions for coding the observations were discussed in tandem with selections from videotapes. A debriefing was held after every practice session during which questions were addressed and coding conventions were clarified.

The PECAP's reliability was assessed for use in PE classes and PC sessions by having observers code written classroom scenarios similar to the practice scenarios. The coding done by each observer for each scenario was compared to a criterion coding established by the trainers. Written scenarios were used to establish reliability in the training because it was not possible to find a sufficient number of parenting classrooms near the training site to do live reliability coding. As assessed in this manner, the overall reliability (percentage of items on which there was exact agreement between observers and criterion coding) of the PECAP in 2006 was .93 for PE class scenarios and .96 for PC session scenarios.

PECAP Revisions

The PECAP measure was changed between 2004 and 2005 and again between 2005 and 2006. Major changes included the following:

- The PECAP form was substantially reformatted to improve ease of recording.
- Detail was added to the PECAP form to help better describe the types of language and literacy activities that might be occurring in a class. At the same time, the number of topics and activities was reduced from 20 to 15 by combining many of the non-literacy topics.
- A section was added to the PECAP Form to allow collection of information on whether the session was for parents only, for children only, for parents and children interacting together, or for parents and children not interacting. The number of parents/children participating was recorded for each of these categories.
- For the revised PECAP, interaction was defined as (1) parent talking with child or (2) parent working very closely with a child on an activity. So, parents and children sharing an activity without talking, such as watching an entertaining videotape, would not be coded as interacting if they were all watching quietly, even if some children were being held by their mothers.

Outcome Measures and Inter-Rater Reliability in the Field

Following the spring 2006 training, the inter-rater reliability of the PECAP was assessed by having two observers (one of the observers trained in 2006 and one of the trainers from the research team) simultaneously code classrooms in the field, providing a measure of reliability based on co-observation of the same classroom. The reliability (percentage of items on which there was exact agreement in the field between the observer and the trainer) assessed in this way for the entire PECAP is .87.

Data from the PECAP were used to create three of the outcome measures used in the CLIO evaluation. Because the PECAP was changed from 2004 to 2005 and again from 2005 to 2006, only the 2006 PECAP data were used to construct outcome measures for the study; none of the 2004 or 2005 PECAP data were used in any way in this evaluation. The definition and reliability of each PECAP outcome measure is shown in table F-1.

Table F-1. Definition and Reliability of PECAP Outcome Measures

Outcome Measure	Reliability (percent of items on which there was exact agreement in the field between observer and trainer) (%)
The percentage of PE class time spent on child literacy activities: reading/looking at books/letters; writing/emergent writing; oral language, songs, rhymes, sound games	
• All child literacy topics	67
• Reading	75
• Writing	67
• Oral language	52
The percentage of PE class time spent on parenting skills: responding to and managing child behavior; home-school relations; ideas for home play; child development; child health, well-being, safety	82
The percentage of PC time in which parents and children interacted on activities directly related to child literacy	89



Spring 2006

**Even Start Classroom Literacy
Interventions and Outcomes Study
(CLIO)**

**Parenting Education and Child and Parent Observation
(OMLIT-PECAP)**

Project ID: _____

Teacher: _____

Date: _____

Observer: _____

Observation complete: ☐

Westat

Time Observation Began ____: ____ AM PM

Date of Observation ____/____/____
mm dd yyyy

Time Observation Ended ____: ____ AM PM

PARENTING EDUCATION and CHILD and PARENT OBSERVATION CLASSROOM DESCRIPTION

Part 1: Identifying Information

Observer ID#: _____ Observer Name: _____

Project/Class ID#: _____

Part 2: Staff List (teachers, assistants, regular staff)

Staff First Name	Staff Class Role	Other Staff Role	Staff ID#
(1) _____	(1) _____	(1) _____	(1) _____
(2) _____	(2) _____	(2) _____	(2) _____
(3) _____	(3) _____	(3) _____	(3) _____
(4) _____	(4) _____	(4) _____	(4) _____

Part 3: Classroom Context

Type of class/session: ☐ Parenting Education ☐ Parent-Child Activity

Number of parents attending _____	Primary Home Language of the Parents (% should add to 100)
Number of parents enrolled _____	_____ % English only
Number of children attending _____	_____ % Spanish primarily
Number of children enrolled _____	_____ % Bilingual Spanish/English
	_____ % Other language primarily
	Specify language #1 _____
	Specify language #2 _____
	_____ % Bilingual other language/English
	Specify language #1 _____
	Specify language #2 _____

Part 4: Teacher question

Ask the lead teacher the following question and record the response verbatim:

What are the two most important things you want parents in this class to learn?

B. Teaching Method Definitions

1. **Direct Instruction:** Staff member **tells or shows** parents and/or children what to do and/or how to do it, includes lecture, question and answer, and demonstration/modeling.
2. **Coaching:** Staff member **closely guides and supports one or two parents** and/or children in trying out new behaviors.
3. **Facilitating/Monitoring:** Staff member **assists** parents and/or children, and **actively observes** in order to help as needed.
4. **Non-instructional:** Staff member is not acting as an instructor, may be away from the group or occupied with non-instructional activities.

C. Learning Context Definitions

1. **Staff-directed instruction:** Participants are listening to, watching, responding to staff.
2. **Discussion:** Two or more participants are engaged in a conversation related to topic.
3. **Role play/Present:** Parents are preparing for or engaging in a role-play activity; parent is presenting to the group.
4. **Videotape:** Parents and/or children are watching a videotape.
5. **Hands-on/Play:** Parents are engaged in making or doing something (e.g., could be arts and crafts or completing a form); parents and/or children are playing or reading together.
6. **Non-instructional:** The context is not instructional, e.g., transition, social conversation, meal and snack time.

D. Parent Child Participation Definitions

1. **Parents only:** Only parents are involved in this topic/activity.
2. **Children only:** Only children are involved in this topic/activity.
3. **Parents w/children interacting:** Parents and children are in the same activity, and they are interacting. Code interaction when parents and children are talking to each other or when parents and children are working closely together on the same activity, such as when a parent guides a child's hand as she/he writes an alphabet letter.
4. **Parents w/children not interacting:** Parents and children are in the same activity, but they are not interacting. For example, parents and children are **not** interacting when sitting together, not talking, listening to a teacher read a book. Parents and children singing together as part of a group are also coded as **not** interacting.

E. Literacy Link Definitions

1. **Parent-child talk:** Check this box if any parent is **talking to any child** during the time of the snapshot.
2. **Child Literacy Link:** Check this box if the topic/activity is related in any way to child language and/or literacy development. Topics 1, 2, and 3 will always have a check in Child Literacy Link, because they are about activities related to reading, writing, and talking.
3. **Parent Literacy Link:** Check this box if the topic/activity involves parent literacy activities in any way. Parent literacy includes parents' reading, writing, and learning new vocabulary words either as English language learners or native English speakers. Always check this if a parent is reading to a child.

A. Definitions for Topics and Activities

[Usually topics are coded in PE and activities in PC, but this is not always true.]

1. **Reading/looking at books/text with children.** Importance of reading books to young children and helping children become aware of letters and numbers. Parents and children looking at text of any kind, including looking at books together; adults reading aloud to children, and children pretending to read (emergent reading), making books. (*Letters and numbers are both considered text.*)
2. **Writing/emergent writing with children.** Importance of providing children with paper and writing/drawing tools. Parents helping their children write, including pretend writing, scribbling, invented spelling; child dictating to an adult
3. **Oral language, songs, rhymes, sound games with children.** Importance of talking to children and providing a rich oral language environment. Activities not focused on written text in which children are encouraged to talk or listen; all songs, rhymes, and sound games
4. **Parenting/responding to and managing child behavior.** How to read children's signals; maintain their attention; respond to children's misbehavior; how to set limits
5. **Child health and safety.** Children's health issues, e.g., preventive medical care, good nutrition; ensuring child safety, e.g., using seatbelts, providing safe toys; promoting children's emotional well-being
6. **School-home relations.** How to interact effectively with children's teachers and child care providers; also includes discussion of school rules, regulations, and procedures
7. **Parent health, well-being/social support.** Adult health issues and ways to maintain good physical and/or emotional health; group social/emotional support
8. **Life skills/cultural content.** Skills and knowledge needed in everyday life (e.g., literacy and ESL, budget, home management, etc.); discussion of cultural customs, such as holiday celebrations
9. **Program planning/leisure activities for parents.** Parents plan program activities, such as field trips; engage in administrative activities related to Even Start program; participate in leisure activities for own enjoyment (e.g., beauty makeovers, knitting, cooking)
10. **Transitions, routines/social conversation/ meals and snacks.** Non-instructional activities such as arriving/ departing, setting-up/cleaning-up, social conversation among parents, eating meals and snacks
11. **A,B. Other PE topic or activity (specify).** Specify and code any other parenting education instructional topic or activity that is not described by codes 1-9
12. **Other PC play activities (staff selected).** Play activities (other than specific child literacy activities) children and parents might engage in that have been selected by staff in parent-child session. Play activities could include arts and crafts activities, dramatic play, playing with blocks, games with rules, and active physical play
13. **Other PC play activities (free play).** Play activities (other than specific child literacy activities) children and parents might engage in, selected by children or parents from a variety of options. Play activities may include arts and crafts activities, dramatic play, playing with blocks, games with rules, and active physical play
- 14.,
15. **Double coding.** Use these rows for double coding topics/activities 1-9 when more than one topic/activity is present, but both fit only one of the categories of topics/activities 1-9. For example, if, at the same time, one group of parents is talking about dental health and the another about child vaccination schedules, code one of these in the row for #5 (Child health and safety) and the second in row #14. Note which category of topic/activity is being double coded (e.g., #5) and provide a brief description of the topic or activity.

Parenting Education and Child and Parent Observation Snapshot MCT-PECAP

Number	Adults and Children Present:
	Teacher(s), Assistant(s)
	Other adults (e.g., visiting speaker)
	Parents
	Total Adults
	Total Children

Staff Codes:

T = Teacher: Lead teacher, other teacher
A = Assistant: Assistant, aide
O = Other: Visiting speaker, program director

Snapshot # ____

Time ____ : ____ am pm

A. Topics/Activities		B. Teaching Method (enter staff codes)					C. Learning Context (circle numbers)						D. Parent Child Participation (enter numbers)					E. Literacy Links (check all that apply)		
		Direct instruction	Coaching	Facilitating/ Monitoring	Non- instructional		1 Staff-directed instruction	2 Discussion	3 Role play/Present	4 Videotape	5 Hands-on/Play	6 Non- instructional	Parents only	Children only	Parents w/children interacting	Parents w/children not interacting		Parent-child talk	Child literacy link	Parent literacy link
1	Reading/looking at books/text with children						1	2	3	4	5	6								
2	Writing/emergent writing with children						1	2	3	4	5	6								
3	Oral language, songs, rhymes, sound games with children						1	2	3	4	5	6								
4	Parenting/responding to and managing child behavior						1	2	3	4	5	6								
5	Child health and safety						1	2	3	4	5	6								
6	School-home relations						1	2	3	4	5	6								
7	Parent health, well-being/social support						1	2	3	4	5	6								
8	Life skills/cultural content						1	2	3	4	5	6								
9	Program planning/leisure activities for parents						1	2	3	4	5	6								
10	Transitions, routines/social conversation/meals, snacks						1	2	3	4	5	6								
11A	Other PE topic or activity (specify) _____						1	2	3	4	5	6								
11B	Other PE topic or activity (specify) _____						1	2	3	4	5	6								
12	Other PC play activities (staff selected)						1	2	3	4	5	6								
13	Other PC play activities (free play)						1	2	3	4	5	6								
14	Code # _____						1	2	3	4	5	6								
15	Code # _____						1	2	3	4	5	6								

(Blank rows #14 and #15 are available for double coding topics/activities 1-9.)

Description of Topics/Activities _____

Comments _____

PECAP OBSERVATION SUMMARY (To be completed at the end of classroom observation)

A. LITERACY-RELATED MATERIALS & EQUIPMENT USED IN CLASS (check all that apply)

☐ 1. Textbooks, workbooks, other commercially produced curriculum materials	☐ 9. Blackboard, whiteboard, flipchart
☐ 2. Photocopy hand-outs from printed materials, e.g., pages from parenting books, directions for making toys	☐ 10. Paper, crayons, pencils, paints, and other materials used for drawing, painting, or writing
☐ 3. Audio visual equipment (e.g., video, TV, audio tape player)	☐ 11. Children's books
☐ 4. Newspapers/magazine articles	☐ 12. Early literacy/numeracy items (e.g., toys such as alphabet puzzles and blocks, magnet letters, or games with letters and/or numbers)
☐ 5. Life skills material (maps, bus schedules, tax forms, etc.)	☐ 13. Computer equipment
☐ 6. Graphics, photographs, paintings, posters, etc.	☐ 14. Other (<i>Specify</i>): _____
☐ 7. Student journals, student reports	☐ 15. Other (<i>Specify</i>): _____
☐ 8. Displays and materials developed by instructor/program staff	☐ 16. No materials were used

B. LENDING LIBRARY FOR USE OUTSIDE CLASS (check all that apply)

☐ Children's books	☐ Videos	☐ Parenting materials
☐ Other (<i>Specify</i>): _____	☐ No materials available for loan	

C. LANGUAGE used by staff in instruction (circle one)	LANGUAGE used by most of parents (circle one)
1 = English primarily	1 = English primarily
2 = Spanish primarily	2 = Spanish primarily
3 = Substantial amounts of both English & Spanish	3 = Substantial amounts of both English & Spanish
4 = Language(s) other than English or Spanish primarily	4 = Language(s) other than English or Spanish primarily
5 = Substantial amounts of both English and other language(s)	5 = Substantial amounts of both English and other language(s)

D. INTERPRETER

Did anyone translate?	☐ Yes	☐ No	IF YES,
Was the person who translated	☐ an interpreter?	☐ a parent?	☐ Translated from Spanish ☐ Translated from other language (<i>Specify</i>): _____

E. LINKS BETWEEN CLASSES

Was any link mentioned or observed between the parenting education class and the parent-child activity session?	☐ Yes Describe link: ☐ No
---	--

F. INDICATION THAT OBSERVATION DAY WAS NOT TYPICAL:

Describe any special events or unusual circumstances that indicate that the day was not typical:
--

G. QUALITY OF INSTRUCTION AND INTERACTION (circle one number for each item)		Not at all 1	Somewhat 2	Quite a bit 3	Very much 4
1.	Instructor listens attentively when parents speak to him/her	1	2	3	4
2.	Instructor provides parents with specific feedback about what they are doing well and what they are doing incorrectly	1	2	3	4
3.	Instructor encourages parents to take leadership roles in the group	1	2	3	4
4.	Instructor provides verbal praise to parents	1	2	3	4
5.	Instructor asks parents about which topics they want to discuss	1	2	3	4
6.	Instructor talks to parents in ways they can easily understand	1	2	3	4
7.	Instructor involves parents in the learning process (e.g., calls on individual students, asks questions)	1	2	3	4
8.	Instructor provides clear examples to illustrate the topic of instruction	1	2	3	4
9.	Instructor speaks warmly to the parents or parents and children	1	2	3	4
10.	Instructor gives clear and thorough answers to parents' questions	1	2	3	4
11.	Instructor shows enthusiasm about parents' efforts to try new skills	1	2	3	4
12.	Instructor uses multiple formats (e.g., discussion, role play, reading) to teach and illustrate important concepts	1	2	3	4
13.	Instructor relates topics to parents' daily lives	1	2	3	4
14.	Instructor encourages parents to try new behaviors	1	2	3	4
15.	Instructor asks parents to tell how they use ideas from the class in their daily lives	1	2	3	4
16.	When introducing a new topic or activity, instructor clearly explains what parents are expected to do	1	2	3	4
17.	Instructor varies the content and/or format of instruction during the class period	1	2	3	4
18.	Instructor breaks down new behaviors into smaller steps to help parents understand what they are being asked to do	1	2	3	4
19.	Instructor welcomes parents' questions and comments	1	2	3	4
20.	Instructor asks review questions to ensure that parents understand the lesson before moving on to the next topic or activity	1	2	3	4

APPENDIX G

FIDELITY MEASUREMENT

This appendix describes the methods used to construct measures of the fidelity of implementation of the CLIO curricula, the findings on fidelity of implementation in the CLIO projects over time, and information on the psychometric properties of the implementation rating scales. The concept of fidelity of implementation targets the degree to which a curriculum as actually implemented in an educational setting meets the specification of the crucial components of the curriculum. This specification is typically provided by the curriculum developer. Chapter 3 presents descriptive statistics about fidelity, and fidelity-adjusted estimates of curriculum effects are discussed in chapter 7.

Sample for the Fidelity Ratings

In the CLIO study, fidelity of implementation was assessed only for the specific curricula provided to projects in the four study groups (Let's Begin/C.I.R.C.L.E and PALS and the two versions of Partners for Literacy). Thus we measured fidelity only in Even Start projects that were assigned to one of the four CLIO curricula. A discussion of ratings used in the fidelity-adjusted analysis for treatment and control projects appears later in this appendix.

Constructing Fidelity Ratings

The study measured the fidelity of implementation of each CLIO curriculum in two ways: the curriculum developers rated the level of implementation of both the preschool and parenting education curriculum at the project level, and evaluation staff conducted independent observations of the preschool and parenting education classrooms. The rating system was the same for both sources of information. A 5-point Likert scale was used to rate each project's implementation of the curriculum as follows:

1. has not appreciably implemented any components of the curriculum,
2. partially implemented (some curriculum components implemented as specified by the model),

3. moderate implementation (about half of the curriculum components implemented as specified by the model),
4. mostly implemented (majority of curriculum components implemented as specified by the model),
5. fully implemented (all curriculum components implemented as specified by the model).

Each method of assessing fidelity had advantages and disadvantages. Evaluation staff conducted independent observations using a standardized instrument. Observers also participated in standardized training. Observer ratings, however, were based on one observation per year. Developer ratings were based on multiple data points and incorporated a variety of information, including multi-day site visits as well as regular one-on-one contact with projects. Developer ratings were not based on standardized data collection, however.

Developer fidelity ratings. All curriculum developers provided a project-level rating of the fidelity of implementation of their curriculum in CLIO classrooms. Developers rated the fidelity of implementation of these curricula in preschool and parenting classes separately, twice during 2005 (the first year of implementation) and twice during 2006 (the second year of implementation). The four ratings provide measures of fidelity of implementation after approximately 6 months, 11 months, 18 months, and 23 months of implementation.

Developer ratings were based on information gained through written records received from the projects, one-on-one emails and phone calls between the developer and project staff, and coaching visits to projects by developer staff.

CLIO observer fidelity ratings. The second set of fidelity ratings was based on 1-day classroom observations conducted by trained CLIO observers in spring 2005 and spring 2006. These ratings reflect fidelity of implementation after approximately 9 months and 20 months of implementation. Fidelity checklists were developed for each preschool and parenting curriculum by the evaluation team, with input from the developers themselves. The checklists were designed to closely match the training that the developers provided to teachers. The ratings by independent observers represent fidelity on the day of the observation.

These fidelity checklists grouped items into a series of key curriculum domains. For example, Let's Begin included eight domains: curriculum theme, print and letter knowledge, classroom materials, circle time activities, high-quality classroom practices, classroom climate, teacher assessments, and home literacy connection. Partners for Literacy (ECE) included six domains: use of literacy games, use of interactive book reading, emphasis on early literacy skills, classroom materials/setting, enriched care giving/teacher-child relationships, and assessments.

Domains included checklist items specific to the curriculum, and for some domains, items from the study's overall classroom observations (i.e., ECE OMLIT and PECAP) that overlapped with developer-specified features of their curriculum. For example, the literacy games domain for Partners for Literacy (ECE) included checklist items about whether literacy games were played and how many pairs of children worked with a teacher, as well as OMLIT items measuring the quality of cognitive challenge and depth of discussion. As another example, the print and letter knowledge domain for Let's Begin included information on curriculum-specific conditions such as the placement and use of the Letter Wall as well as the OMLIT item about whether toys or materials accessible to children include words or letters.

Each classroom was rated on the percentage of items satisfied within each domain. Each domain then was given a weight based on the judgment of the evaluation team about the importance of the domain to the curriculum model. Table G-1 shows the domains, domain components by source, and domain weights for each curriculum. The domain-specific percentages were then weighted by the domain weights to obtain a weighted score. The evaluation team established scoring bands to recode the weighted score into the same 5-point Likert scale used by the developers.

To provide an indication of the validity of the observer parenting fidelity ratings, observers rated fidelity in parenting classrooms in the two CLIO preschool curricula study groups. As expected, the parenting fidelity ratings for projects in these study groups were much lower than parenting fidelity ratings in the study groups that were also provided with CLIO parenting curricula. The fact that the observer fidelity ratings for parenting instruction were sometimes greater than 1 in the projects without CLIO parenting curricula might be interpreted as a sign that there was some

Table G-1. Fidelity Domains by CLIO Curriculum

Domain	Number of domain items			Domain weight
	Checklist items	OMLIT/PECAP items	Total	
Let’s Begin with the Letter People: ECE				
Curriculum theme	2	0	2	2
Print & letter knowledge	7	3	10	2
Classroom materials	6	2	8	1
Circle time activities	7	1	8	1
High-quality classroom practices (CIRCLE training)	0	9	9	.5
Classroom climate	2	2	4	.5
Teacher assessments	2	0	2	.5
Home literacy connection	4	0	4	.5
Partners for Literacy (PfL): ECE				
Use of literacy games	2	2	4	2
Use of interactive book reading	2	11	13	2
Classroom materials	5	1	6	1
Use of enriched caregiving	1	1	2	1
Emphasis on early literacy skills	0	10	10	1
Use of PfL assessments	2	0	2	.5
Play & Learning Strategies (PALS): PE				
Certification/session	2	0	2	2
Components of instruction: practice skills	8	0	8	2
Components of instruction: review, introduce, & plan	6	0	6	1
Classroom materials	7	0	7	1
Classroom climate	3	11	14	.5
Partners for Literacy (PfL): PE				
Use of PfL materials	5	0	5	1
Use of PfL strategies	6	0	6	1
Instructor characteristics	3	5	8	.5

cross-contamination of parenting instruction when preschool instruction is deliberately changed. However, it may also reflect the fact that the observer fidelity rating includes some measures from the PECAP.

Correlations Between Observer and Developer Fidelity Ratings

Since the study obtained fidelity ratings from two different sources (curriculum developers and evaluation observers), we explored the consistency of the ratings by examining the correlation between developer ratings and observer ratings for the same time period (spring of each year of the study).¹ As table G-2 indicates, the correlations between the two rating scales for all preschool classrooms regardless of study group were statistically significant overall (at $p < .05$). However, when each study group was examined separately, the correlation between the two types of ratings was statistically significant only for the preschool ratings for Let's Begin and PALS curriculum (2005 only) and preschool Partners for Literacy curricula (2005 and 2006). There were no significant correlations for the parenting education ratings.

The correlations between the two fidelity ratings could indicate that the two systems are measuring different aspects of the curricula or one system includes a variety of data and the other system is a one-time assessment.

Stability of Fidelity Ratings Over Time

We also looked at the stability of the fidelity ratings at the project level. We examined the time-to-time correlation of the developer ratings from spring 2005 to spring 2006.² All time-to-time correlations from both developers were positive and statistically significant ($p < .05$). Time-to-time correlations of the preschool fidelity ratings for Partners for Literacy were .54, and correlations for Let's Begin were .41. The time-to-time correlations of the fidelity ratings for parenting were .63 for Partners for Literacy and .44 for PALS.

¹ Kendall's Tau-b is the correlational measure used. It is used for ordinal data and takes ties into account. For this analysis, the observer fidelity ratings were averaged up to the project level then correlated with the project-level developer ratings. If a spring developer rating was missing, we used the January rating if available.

² We ran Spearman rank order correlations and Kendall's Tau-b correlations for this analysis. The correlation coefficients for the Spearman runs were slightly higher than those from the Kendall's Tau-b run. We report the Kendall's Tau-b correlation coefficients in the text.

Table G-2. Correlation Between Project-level Observer and Developer Fidelity Ratings^a by Study Group and Classroom Type: Spring 2005 and Spring 2006

	All study groups ¹	Study group			
		Let's Begin with the Letter People (ECE)	Let's Begin and Play & Learning Strategies (ECE/PE)	Partners for Literacy (ECE)	Partners for Literacy (ECE/PE)
Early childhood education (ECE)					
2005	0.48*	0.04	0.41*	0.62*	0.19
2006	0.39*	0.21	0.32	0.55*	0.19
Parenting education (PE)					
2005	0.10	†	-0.09	†	0.28
2006	-0.01	†	0.11	†	-0.09

^a The fidelity ratings range from 1 (not appreciably implemented) to 5 (fully implemented).

†Not applicable.

* $p < 0.05$.

NOTES: Kendall's Tau-b is the correlational measure used. For this analysis, the observer fidelity ratings were averaged up to the project level then correlated with the project-level developer ratings. If a spring developer rating was missing, we used the January rating if available.

SOURCE: U.S. Department of Education, Even Start Classroom Literacy Interventions and Outcomes Study, "Observer and Developer Fidelity Ratings," Spring 2005 and Spring 2006.

The correlation of the preschool ratings across time points for Partners for Literacy, was 0.33 and significant ($p < .05$). The correlation of the parenting ratings across time was 0.35, which was not significant. For Let's Begin and PALS, the correlations for the preschool ratings and parenting ratings were 0.20 and 0.25, respectively, and neither was statistically significant.

Fidelity Ratings Used in the Fidelity-Adjusted Analysis

For the fidelity-adjusted analysis described in chapter 7, the observer and developer fidelity data were analyzed at the classroom level.³ For classrooms in projects that had CLIO combined curricula, parenting fidelity ratings were not included for

³ For this analysis, the project-level fidelity ratings provided by the developers were assigned to each classroom in the project. This allowed us to conduct the fidelity-adjusted analysis of outcomes at the classroom level, where most of the child data also reside.

analysis of child outcomes, and preschool fidelity ratings were not included for analysis of parenting outcomes. As mentioned earlier, CLIO fidelity ratings reflected the degree to which the CLIO curricula were being implemented as intended. Classrooms in control projects were assigned a developer-rated fidelity score of null, but could be assigned positive observer-rated fidelity scores. This is because the observer-rated fidelity scores were based partially on developer-specific checklists and partially on the OMLIT/PECAP observation measures used in all projects. Accordingly, one fidelity score for ECE classrooms and one for PE classrooms were developed for each control project based only on OMLIT/PECAP items, respectively. These scores are the same that would be obtained for a classroom in a project assigned to an experimental group in which the observer did not check any of the developer-specific items. We call these OMLIT/PECAP-only scores “pseudo fidelity” scores, one for the ECE curriculum and one for the PE curriculum. Finally, the same cutpoints for translating these continuous scales to the ordinal 1-through-5 score were applied.

APPENDIX H

FINAL MULTI-LEVEL OUTCOME MODELS

Although the results from the multi-level models that are most pertinent to the research questions are given in chapter 5, we present more information about these models in this appendix. In addition to serving as documentation of the modeling, readers may be specifically interested in the relationships of some of the covariates to the outcomes.

There is a table in the appendix for each of the 21 formal outcomes. Each table shows the descriptions of the variables in the models, the point estimates for their coefficients in the model, the standard error on that point estimate, the corresponding t -statistic to test whether that coefficient is significantly different from zero, the degrees of freedom available for the variance estimate in the denominator of the t -statistic, and the p -value for the two sided-hypothesis. None of the p -values in this appendix are adjusted for multiple comparisons. All variables in the models are shown with the exception of the stratum indicators. (In the interest of space, these were suppressed.) The reader should be aware that the coefficients for the balance of the variables are sensitive to the fact that the dummy variables for the strata were part of these models.

The first block of variables in all the tables correspond to project-level covariates. The estimates, therefore, show $\hat{\delta}$, in the notation of chapter 4. The next two blocks of variables in tables H-1 through H-13 correspond to child- and parent-level covariates. The estimates, therefore, show $\hat{\gamma}$, in the notation of chapter 4. The fourth block of variables in tables H-1 through H-13 and the second block of variables in tables H-14 through H-21 correspond to study groups. The estimates, therefore, show $\hat{\alpha}$, in the notation of chapter 4. As mentioned above, $\hat{\beta}$ is not shown.

The last three rows of tables H-1 through H-9 give the components of variance at the various levels of random effects for child outcomes. In the notation of chapter 4, the level-1 (child-year) component of variance is equal to $\text{Var}(e_{ijkt})$; the level-2 (child) component of variance is equal to $\text{Var}(c_{ijk})$; and the level-3 (project) component of variance is equal to $\text{Var}(a_{ij})$. Similarly, the last three rows of tables H-10 through H-13 give the components of variance at the various levels of random effects for parent

outcomes. In the notation of chapter 4, the level-1 (parent-year) component of variance is equal to $\text{Var}(e_{ijkt})$; the level-2 (parent) component of variance is equal to $\text{Var}(c_{ijk})$; and the level-3 (project) component of variance is equal to $\text{Var}(a_{ij})$. Also, the last two rows of tables H-14 through H-21 give the components of variance at the two levels of random effects for instructional outcomes. In the notation of chapter 4, the level-1 (project-year) component of variance is equal to $\text{Var}(e_{ijt})$, and the level-2 (project) component of variance is equal to $\text{Var}(a_{ij})$.

As an example of how these tables may be of interest, note that the child to teacher ratio is significant in only one of the nine child outcome models. As another example, note that while boys score below girls on a number of areas of emergent literacy, they score far below girls on social competence. Finally, note that mother's education is not only positively related to all measures of emergent English literacy; it is also significantly and positively related to child social competence.

One note of caution about comparing coefficients across the tables of this appendix is that scaling for the dependent variables is not constant across the variables. See table A-3 and its endnotes for information on the scaling. Also note that the covariates within a table may have different scales, although the coding of each independent variable is constant across tables. The binary variables are coded as 1 for yes and 0 for no.

Table H-1. Model for Child Expressive Language in English—Spring 2005 and Spring 2006 Combined

Model term	estimate	s.e.	t-ratio	d.f.	p-value
Average baseline score for IGDI	0.29	0.10	3.05	89	0.003
Average of child English literacy, child social competence, and parent reading & vocabulary skill baseline scores excluding English IGDI	0.01	0.02	0.54	89	0.588
Child to teacher ratio	0.00	0.08	0.06	89	0.954
Number of children in household under age 8	-0.22	0.12	-1.85	2742	0.064
Number of adults in household	-0.36	0.14	-2.58	2742	0.010
Parent age in years (responding parent)	0.06	0.03	2.26	2742	0.024
Mother is college graduate (could be associate's degree)	1.19	0.41	2.91	2742	0.004
Home language is not English	-3.87	0.36	-10.86	2742	0.000
Household income above \$1500	1.49	0.24	6.21	2742	0.000
Number of times that child moved in last year	-0.33	0.19	-1.77	2742	0.076
Child has special needs	-1.14	0.48	-2.36	2742	0.019
Child is male	0.26	0.25	1.05	2742	0.296
Child is Hispanic/Latino	-2.99	0.55	-5.44	2742	0.000
Child age in months	0.43	0.02	24.64	2742	0.000
Flag for 2006 assessment	0.24	0.27	0.89	2742	0.375
Study group A2 (Let's Begin & PALS)	-1.19	0.74	-1.62	89	0.109
Study group B2 (Partners for Literacy ECE/PE)	-0.83	0.67	-1.23	89	0.222
Study group A1 (Let's Begin)	-1.10	0.66	-1.67	89	0.098
Study group B1 (Partners for Literacy ECE)	-0.20	0.66	-0.31	89	0.758
Variance components	estimate	s.e.	Chi-square	d.f.	p-value
level-1 (child-year)	18.85	1.17	NA	NA	NA
level-2 (child)	17.57	1.66	4686.50	2363	0.000
level-3 (project)	2.73	0.62	312.28	89	0.000

Table H-2. Model for Child Expressive Language in Spanish—Spring 2005 and Spring 2006 Combined

Model term	estimate	s.e.	t-ratio	d.f.	p-value
Baseline score for Spanish IGDI	0.19	0.11	1.69	54	0.097
Average of child English literacy, child social competence, and parent reading & vocabulary skill baseline scores	-0.08	0.04	-2.13	54	0.037
Child to teacher ratio	0.01	0.10	0.06	54	0.955
Number of children in household under age 8	-0.23	0.17	-1.34	1498	0.181
Number of adults in household	0.19	0.16	1.15	1498	0.249
Parent age in years (responding parent)	0.06	0.03	2.35	1498	0.019
Mother is college graduate (could be associate's degree)	0.80	0.54	1.49	1498	0.137
Home language is not English	-0.48	0.39	-1.22	1498	0.223
Household income above \$1500	0.89	0.30	2.93	1498	0.004
Number of times that child moved in last year	0.03	0.21	0.12	1498	0.903
Child has special needs	-1.90	0.48	-3.92	1498	0.000
Child is male	-0.17	0.29	-0.59	1498	0.555
Child is Hispanic/Latino	3.15	1.88	1.68	1498	0.092
Child age in months	0.31	0.02	16.25	1498	0.000
Flag for 2006 assessment	-0.48	0.28	-1.71	1498	0.087
Study group A2 (Let's Begin & PALS)	0.43	0.80	0.54	54	0.591
Study group B2 (Partners for Literacy ECE/PE)	0.28	0.84	0.34	54	0.739
Study group A1 (Let's Begin)	0.49	0.84	0.59	54	0.560
Study group B1 (Partners for Literacy ECE)	-0.50	0.79	-0.63	54	0.530
Variance components	estimate	s.e.	Chi-square	d.f.	p-value
level-1 (child-year)	24.01	2.27	NA	NA	NA
level-2 (child)	5.56	2.22	1681.67	1285	0.000
level-3 (project)	2.61	0.74	208.47	54	0.000

Table H-3. Model for Child Receptive Vocabulary—Spring 2005 and Spring 2006 Combined

Model term	estimate	s.e.	t-ratio	d.f.	p-value
Average baseline score for PPVT	0.29	0.09	3.33	89	0.002
Average of child English literacy, child social competence, and parent reading & vocabulary skill baseline scores excluding PPVT	0.19	0.15	1.21	89	0.229
Child to teacher ratio	0.04	0.42	0.09	89	0.927
Number of children in household under age 8	-1.52	0.85	-1.78	395	0.075
Number of adults in household	-1.58	0.82	-1.93	208	0.054
Parent age in years (responding parent)	0.42	0.15	2.84	115	0.006
Mother is college graduate (could be associate's degree)	4.41	3.06	1.44	191	0.151
Home language is not English	-22.25	2.30	-9.67	259	0.000
Household income above \$1500	7.79	1.48	5.25	938	0.000
Number of times that child moved in last year	-2.01	1.08	-1.86	41	0.070
Child has special needs	-4.99	2.19	-2.28	152	0.024
Child is male	-0.95	1.70	-0.56	189	0.576
Child is Hispanic/Latino	-9.73	2.87	-3.39	137	0.001
Child age in months	2.74	0.10	26.59	251	0.000
Flag for 2006 assessment	1.52	1.39	1.09	611	0.277
Study group A2 (Let's Begin & PALS)	-4.94	3.17	-1.56	89	0.122
Study group B2 (Partners for Literacy ECE/PE)	-3.26	3.40	-0.96	89	0.341
Study group A1 (Let's Begin)	-2.81	2.70	-1.04	89	0.302
Study group B1 (Partners for Literacy ECE)	-0.69	3.06	-0.22	89	0.823
Variance Components	estimate	s.e.	Chi-square	d.f.	p-value
level-1 (child-year)	746.07	56.58	NA	NA	NA
level-2 (child)	457.19	75.12	4116.22	2361	0.000
level-3 (project)	27.30	12.17	175.18	89	0.000

Table H-4. Model for Child Elision Component of Phonological Awareness—Spring 2005 and Spring 2006 Combined

Model term	estimate	s.e.	t-ratio	d.f.	p-value
Average baseline score for elision	0.07	0.13	0.53	76	0.598
Average of child English literacy, child social competence, and parent reading & vocabulary skill baseline scores excluding elision	0.33	0.19	1.75	73	0.084
Child to teacher ratio	0.54	0.51	1.06	89	0.291
Number of children in household under age 8	-0.52	0.91	-0.57	310	0.569
Number of adults in household	-1.69	0.98	-1.73	220	0.085
Parent age in years (responding parent)	0.14	0.16	0.91	878	0.366
Mother is college graduate (could be associate's degree)	5.65	3.53	1.60	65	0.114
Home language is not English	-17.79	2.95	-6.03	84	0.000
Household income above \$1500	6.72	1.88	3.58	225	0.001
Number of times that child moved in last year	-2.04	0.95	-2.15	2736	0.031
Child has special needs	-7.46	2.74	-2.72	138	0.008
Child is male	-4.67	1.69	-2.77	86	0.007
Child is Hispanic/Latino	-5.65	2.83	-2.00	2602	0.046
Child age in months	2.48	0.12	21.29	190	0.000
Flag for 2006 assessment	-1.93	2.05	-0.95	1310	0.345
Study group A2 (Let's Begin & PALS)	3.40	3.69	0.92	89	0.361
Study group B2 (Partners for Literacy ECE/PE)	-3.34	3.67	-0.91	89	0.365
Study group A1 (Let's Begin)	-0.29	3.72	-0.08	89	0.939
Study group B1 (Partners for Literacy ECE)	-1.83	4.07	-0.45	78	0.654
Variance components	estimate	s.e.	Chi-square	d.f.	p-value
level-1 (child-year)	1609.68	105.12	NA	NA	NA
level-2 (child)	47.22	103.53	2564.26	2358	0.002
level-3 (project)	49.53	21.87	206.29	89	0.000

Table H-5. Model for Child Blending Component of Phonological Awareness—Spring 2005 and Spring 2006 Combined

Model term	estimate	s.e.	t-ratio	d.f.	p-value
Average baseline score for blending	0.13	0.12	1.12	64	0.267
Average of child English literacy, child social competence, and parent reading & vocabulary skill baseline scores excluding blending	0.40	0.15	2.61	89	0.011
Child to teacher ratio	1.05	0.48	2.20	89	0.030
Number of children in household under age 8	-1.16	0.97	-1.20	61	0.237
Number of adults in household	-1.13	1.13	-1.00	1449	0.320
Parent age in years (responding parent)	0.41	0.17	2.40	66	0.019
Mother is college graduate (could be associate's degree)	2.18	3.41	0.64	191	0.524
Home language is not English	-12.98	2.38	-5.45	142	0.000
Household income above \$1500	6.00	2.21	2.72	98	0.008
Number of times that child moved in last year	-0.67	0.98	-0.69	272	0.494
Child has special needs	-10.76	3.19	-3.37	43	0.002
Child is male	-0.91	1.59	-0.57	102	0.569
Child is Hispanic/Latino	-4.89	3.06	-1.60	501	0.110
Child age in months	2.67	0.12	21.98	162	0.000
Flag for 2006 assessment	3.33	2.23	1.49	2729	0.136
Study group A2 (Let's Begin & PALS)	-6.25	4.03	-1.55	89	0.124
Study group B2 (Partners for Literacy ECE/PE)	-6.14	4.28	-1.43	89	0.155
Study group A1 (Let's Begin)	-5.76	3.81	-1.51	89	0.134
Study group B1 (Partners for Literacy ECE)	0.27	4.41	0.06	89	0.951
Variance components	estimate	s.e.	Chi-square	d.f.	p-value
level-1 (child-year)	1330.99	94.33	NA	NA	NA
level-2 (child)	295.78	102.04	3040.35	2351	0.000
level-3 (project)	67.35	20.51	230.70	89	0.000

Table H-6. Model for Child Print Awareness—Spring 2005 and Spring 2006 Combined

Model term	estimate	s.e.	t-ratio	d.f.	p-value
Average baseline score for English IGDI	1.20	0.58	2.08	89	0.040
Average of child English literacy, child social competence, and parent reading & vocabulary skill baseline scores excluding English IGDI	-0.01	0.18	-0.06	89	0.952
Child to teacher ratio	0.35	0.56	0.62	89	0.534
Number of children in household under age 8	-1.64	0.90	-1.81	241	0.071
Number of adults in household	-1.60	1.09	-1.47	54	0.147
Parent age in years (responding parent)	0.53	0.16	3.34	1186	0.001
Mother is college graduate (could be associate's degree)	4.45	3.37	1.32	176	0.188
Home language is not English	-4.72	2.55	-1.85	2745	0.064
Household income above \$1500	7.17	1.54	4.66	2165	0.000
Number of times that child moved in last year	-3.00	1.03	-2.90	74	0.005
Child has special needs	-6.10	2.44	-2.50	68	0.015
Child is male	-5.14	1.67	-3.08	72	0.003
Child is Hispanic/Latino	-8.52	2.97	-2.87	127	0.005
Child age in months	3.15	0.12	25.88	481	0.000
Flag for 2006 assessment	7.83	1.87	4.19	271	0.000
Study group A2 (Let's Begin & PALS)	8.66	3.76	2.31	89	0.023
Study group B2 (Partners for Literacy ECE/PE)	-3.41	4.88	-0.70	89	0.486
Study group A1 (Let's Begin)	8.10	3.92	2.07	89	0.041
Study group B1 (Partners for Literacy ECE)	9.57	4.45	2.15	89	0.034
Variance components	estimate	s.e.	Chi-square	d.f.	p-value
level-1 (child-year)	803.40	62.15	NA	NA	NA
level-2 (child)	652.91	71.09	4455.81	2366	0.000
level-3 (project)	101.41	24.11	305.28	89	0.000

Table H-7. Model for Child Syntax and Grammar—Spring 2005 and Spring 2006 Combined

Model term	estimate	s.e.	t-ratio	d.f.	p-value
Average baseline score for English IGDI	0.86	0.40	2.15	89	0.034
Average of child English literacy, child social competence, and parent reading & vocabulary skill baseline scores excluding English IGDI	0.24	0.12	1.93	76	0.057
Child to teacher ratio	-0.29	0.39	-0.75	89	0.456
Number of children in household under age 8	-0.51	1.15	-0.44	30	0.663
Number of adults in household	-0.96	1.05	-0.91	41	0.366
Parent age in years (responding parent)	0.51	0.18	2.82	35	0.008
Mother is college graduate (could be associate's degree)	2.68	3.60	0.74	63	0.460
Home language is not English	-11.34	2.91	-3.90	39	0.001
Household income above \$1500	5.95	2.30	2.59	15	0.021
Number of times that child moved in last year	-1.48	1.36	-1.09	12	0.298
Child has special needs	-4.22	3.04	-1.39	43	0.172
Child is male	-3.67	1.62	-2.26	139	0.025
Child is Hispanic/Latino	1.84	2.85	0.65	141	0.520
Child age in months	3.30	0.11	31.21	913	0.000
Flag for 2006 assessment	-0.29	1.73	-0.17	1316	0.865
Study group A2 (Let's Begin & PALS)	-2.01	3.17	-0.64	36	0.529
Study group B2 (Partners for Literacy ECE/PE)	-5.72	3.16	-1.81	89	0.073
Study group A1 (Let's Begin)	-1.47	2.72	-0.54	89	0.591
Study group B1 (Partners for Literacy ECE)	-0.90	2.80	-0.32	89	0.749
Variance components	estimate	s.e.	Chi-square	d.f.	p-value
level-1 (child-year)	1251.07	88.09	NA	NA	NA
level-2 (child)	280.47	94.94	3092.19	2355	0.000
level-3 (project)	10.01	10.02	146.88	89	0.000

Table H-8. Model for Child Social Competence—Spring 2005 and Spring 2006 Combined

Model term	estimate	<i>s.e.</i>	<i>t</i> -ratio	d.f.	<i>p</i> -value
Average baseline score for social competence	0.37	0.11	3.29	88	0.002
Average of child English literacy, child social competence, and parent reading & vocabulary skill baseline scores excluding social competence	-0.21	0.10	-2.15	88	0.034
Child to teacher ratio	-0.25	0.47	-0.52	88	0.602
Number of children in household under age 8	-0.77	1.12	-0.69	112	0.492
Number of adults in household	1.53	1.13	1.36	714	0.174
Parent age in years (responding parent)	0.16	0.16	1.00	2821	0.320
Mother is college graduate (could be associate's degree)	9.18	3.27	2.80	121	0.006
Home language is not English	4.88	3.17	1.54	254	0.124
Household income above \$1500	2.86	2.14	1.34	132	0.184
Number of times that child moved in last year	-2.66	1.13	-2.35	108	0.020
Child has special needs	-17.80	3.09	-5.76	2729	0.000
Child is male	-17.63	2.25	-7.85	61	0.000
Child is Hispanic/Latino	6.73	3.24	2.08	2821	0.038
Child age in months	1.80	0.15	11.96	686	0.000
Flag for 2006 assessment	0.75	2.66	0.28	2797	0.777
Study group A2 (Let's Begin & PALS)	13.70	4.22	3.25	88	0.002
Study group B2 (Partners for Literacy ECE/PE)	7.81	4.75	1.64	88	0.103
Study group A1 (Let's Begin)	9.77	3.94	2.48	88	0.015
Study group B1 (Partners for Literacy ECE)	12.40	4.18	2.97	88	0.004
Variance components	estimate	<i>s.e.</i>	Chi-square	d.f.	<i>p</i> -value
level-1 (child-year)	1500.05	107.06	NA	NA	NA
level-2 (child)	432.38	121.94	3338.11	2438	0.000
level-3 (project)	85.78	26.58	233.91	88	0.000

Table H-9. Model for Child Monthly Hours of ECE Instruction Received—Spring 2005 and Spring 2006 Combined

Model term	estimate	s.e.	t-ratio	d.f.	p-value
Average baseline score for ECE hours	0.69	0.06	11.52	90	0.000
Child to teacher ratio	0.15	0.53	0.27	90	0.785
Number of children in household under age 8	0.49	0.41	1.22	2877	0.225
Number of adults in household	0.01	0.44	0.02	2877	0.982
Parent age in years (responding parent)	0.20	0.06	3.28	2877	0.001
Mother is college graduate (could be associate's degree)	-0.52	1.38	-0.37	2877	0.708
Home language is not English	0.81	1.08	0.75	2877	0.456
Household income above \$1500	0.93	0.80	1.16	2877	0.247
Number of times that child moved in last year	-2.12	0.51	-4.13	2877	0.000
Child has special needs	0.81	1.09	0.74	2877	0.459
Child is male	-0.24	0.73	-0.32	2877	0.747
Child is Hispanic/Latino	0.82	1.79	0.46	2877	0.646
Child age in months	0.42	0.10	4.34	2877	0.000
Flag for 2006 assessment	3.75	1.15	3.25	2877	0.002
Study group A2 (Let's Begin & PALS)	4.16	3.44	1.21	90	0.230
Study group B2 (Partners for Literacy ECE/PE)	3.19	3.36	0.95	90	0.346
Study group A1 (Let's Begin)	-2.12	3.80	-0.56	90	0.578
Study group B1 (Partners for Literacy ECE)	-0.24	3.22	-0.07	90	0.941
Variance components	estimate	s.e.	Chi-square	d.f.	p-value
level-1 (child-year)	285.46	18.10	NA	NA	NA
level-2 (child)	5.66	16.62	2564.12	2475	0.104
level-3 (project)	139.65	20.48	1412.90	90	0.000

Table H-10. Model for Parent Interactive Reading Skill—Spring 2005 and Spring 2006 Combined

Model term	estimate	<i>s.e.</i>	<i>t</i> -ratio	d.f.	<i>p</i> -value
Average of child English literacy, child social competence, and parent reading & vocabulary skill baseline scores	0.00	0.00	-0.34	90	0.734
Child to teacher ratio	0.01	0.01	1.46	90	0.148
Number of children in household under age 8	-0.02	0.03	-0.66	2388	0.51
Number of adults in household	0.03	0.03	1.01	2388	0.312
Parent age in years (responding parent)	0.01	0.00	1.45	2388	0.147
Mother is college graduate (could be associate's degree)	0.34	0.08	4.41	2388	0
Home language is not English	-0.14	0.07	-2.08	2388	0.037
Household income above \$1500	0.16	0.04	4.46	2388	0
Number of times that family moved in last year	-0.02	0.02	-0.64	2388	0.52
Any of assessed children in family have special needs	0.01	0.05	0.11	2388	0.917
Any of assessed children in family are male	0.05	0.04	1.13	2388	0.258
Parent is Hispanic/Latino	0.22	0.08	2.85	2388	0.005
Average age of assessed children in family in months	0.00	0.00	-1.10	2388	0.272
Flag for 2006 assessment	0.09	0.04	2.47	2388	0.014
Study group A2 (Let's Begin & PALS)	0.36	0.08	4.56	90	0
Study group B2 (Partners for Literacy ECE/PE)	0.41	0.09	4.76	90	0
Study group A1 (Let's Begin)	0.26	0.08	3.15	90	0.003
Study group B1 (Partners for Literacy ECE)	0.02	0.07	0.26	90	0.794
Variance components	estimate	<i>s.e.</i>	Chi-square	d.f.	<i>p</i> -value
level-1 (parent-year)	0.45	0.04	NA	NA	NA
level-2 (parent)	0.37	0.04	4039.23	2064	0.000
level-3 (project)	0.03	0.01	202.42	90	0.000

**Table H-11. Model for Parent Responsiveness—Spring 2005 and Spring 2006
Combined**

Model term	estimate	s.e.	t-ratio	d.f.	p-value
Average baseline score for IGDI	-0.01	0.01	-0.64	89	0.521
Average of child English literacy, child social competence, and parent reading & vocabulary skill baseline scores excluding English IGDI	0.01	0.00	2.26	89	0.026
Child to teacher ratio	0.01	0.01	0.79	89	0.43
Number of children in household under age 8	-0.05	0.03	-1.75	2387	0.079
Number of adults in household	0.00	0.02	0.03	2387	0.978
Parent age in years (responding parent)	0.00	0.00	-1.10	2387	0.27
Mother is college graduate (could be associate's degree)	0.34	0.06	5.84	2387	0
Home language is not English	-0.23	0.06	-3.99	2387	0
Household income above \$1500	0.20	0.04	5.42	2387	0
Number of times that family moved in last year	0.01	0.02	0.32	2387	0.75
Any of assessed children in family have special needs	-0.15	0.05	-2.78	2387	0.006
Any of assessed children in family are male	-0.13	0.04	-3.22	2387	0.002
Parent is Hispanic/Latino	0.02	0.09	0.21	2387	0.836
Average age of assessed children in family in months	0.01	0.00	4.15	2387	0
Flag for 2006 assessment	-0.09	0.04	-2.46	2387	0.014
Study group A2 (Let's Begin & PALS)	0.26	0.07	3.53	89	0.001
Study group B2 (Partners for Literacy ECE/PE)	0.17	0.08	2.12	89	0.037
Study group A1 (Let's Begin)	0.17	0.08	2.16	89	0.033
Study group B1 (Partners for Literacy ECE)	-0.05	0.08	-0.64	89	0.522
Variance components	estimate	s.e.	Chi-square	d.f.	p-value
level-1 (parent-year)	0.38	0.03	NA	NA	NA
level-2 (parent)	0.34	0.04	4138.25	2064	0.000
level-3 (project)	0.03	0.01	210.21	89	0.000

Table H-12. Model for Parent Reading and Vocabulary Skill—Spring 2005 and Spring 2006 Combined

Model term	estimate	s.e.	t-ratio	d.f.	p-value
Average baseline score for parent reading and vocabulary skill	0.37	0.05	7.11	89	0
Average of child English literacy and child social competence baseline scores	0.02	0.11	0.17	89	0.869
Child to teacher ratio	0.12	0.29	0.42	89	0.679
Number of children in household under age 8	-0.39	0.72	-0.54	2488	0.59
Number of adults in household	-3.24	0.80	-4.06	2488	0
Parent age in years (responding parent)	-0.16	0.11	-1.42	1008	0.155
Mother is college graduate (could be associate's degree)	10.67	1.96	5.44	1081	0
Home language is not English	-26.45	2.83	-9.33	2488	0
Household income above \$1500	8.08	1.28	6.30	2488	0
Number of times that family moved in last year	-0.42	0.64	-0.66	162	0.512
Any of assessed children in family have special needs	1.55	1.82	0.85	1185	0.395
Any of assessed children in family are male	1.48	1.30	1.13	860	0.258
Parent is Hispanic/Latino	-29.56	3.74	-7.91	2488	0
Average age of assessed children in family in months	0.19	0.06	3.21	2488	0.002
Flag for 2006 assessment	0.36	1.10	0.33	600	0.745
Study group A2 (Let's Begin & PALS)	-1.61	2.51	-0.64	89	0.523
Study group B2 (Partners for Literacy ECE/PE)	-2.30	2.09	-1.10	89	0.276
Study group A1 (Let's Begin)	2.80	2.68	1.05	89	0.299
Study group B1 (Partners for Literacy ECE)	-1.86	2.29	-0.81	89	0.419
Variance components	estimate	s.e.	Chi-square	d.f.	p-value
level-1 (parent-year)	163.34	14.28	NA	NA	NA
level-2 (parent)	550.72	28.67	9311.46	2152	0.000
level-3 (project)	30.99	9.34	222.54	89	0.000

Table H-13. Model for Parent Monthly Hour of PE and PC Instruction Received—Spring 2005 and Spring 2006 Combined

Model term	estimate	s.e.	t-ratio	d.f.	p-value
Average baseline score for receipt of PE/PC hours	0.06	0.02	3.85	90	0
Child to teacher ratio	-0.22	0.14	-1.59	90	0.115
Number of children in household under age 8	-0.18	0.13	-1.39	2686	0.164
Number of adults in household	-0.02	0.19	-0.11	2686	0.914
Parent age in years (responding parent)	0.05	0.02	2.07	2686	0.039
Mother is college graduate (could be associate's degree)	-0.66	0.46	-1.44	2686	0.15
Home language is not English	0.03	0.36	0.10	2686	0.924
Household income above \$1500	0.23	0.29	0.80	2686	0.425
Number of times that parent moved in last year	-0.40	0.20	-2.06	2686	0.039
Any of assessed children in family have special needs	0.50	0.32	1.56	2686	0.119
Any of assessed children in family are male	-0.37	0.21	-1.81	2686	0.07
Parent is Hispanic/Latino	0.36	0.64	0.57	2686	0.567
Average age of assessed children in family in months	0.08	0.01	8.32	2686	0
Flag for 2006 assessment	0.93	0.67	1.39	2686	0.164
Study group A2 (Let's Begin & PALS)	1.32	1.14	1.16	90	0.251
Study group B2 (Partners for Literacy ECE/PE)	2.16	1.18	1.83	90	0.071
Study group A1 (Let's Begin)	1.31	1.67	0.78	90	0.437
Study group B1 (Partners for Literacy ECE)	-0.14	1.34	-0.11	90	0.917
Variance components	estimate	s.e.	Chi-square	d.f.	p-value
level-1 (parent-year)	34.41	2.43	NA	NA	NA
level-2 (parent)	3.58	2.30	2620.33	2326	0.000
level-3 (project)	22.57	3.26	1761.62	90	0.000

Table H-14. Model for OMLIT Oral Language Development—Spring 2005 and Spring 2006 Combined

Model term	estimate	<i>s.e.</i>	<i>t</i> -ratio	d.f.	<i>p</i> -value
Baseline score for the outcome	0.25	0.05	4.59	90	0.000
Child to teacher ratio	0.00	0.22	0.01	90	0.989
Flag for 2006 assessment	-1.50	1.52	-0.99	196	0.324
Study group A2 (Let's Begin & PALS)	3.87	1.99	1.95	90	0.054
Study group B2 (Partners for Literacy ECE/PE)	2.34	1.82	1.28	90	0.203
Study group A1 (Let's Begin)	2.55	1.83	1.39	90	0.167
Study group B1 (Partners for Literacy ECE)	4.98	1.85	2.69	90	0.009
Variance components	estimate	<i>s.e.</i>	Chi-square	d.f.	<i>p</i> -value
level-1 (project-year)	116.97	NA	NA	NA	NA
level-2 (project)	0.08	NA	74.91	90	>.500

Table H-15. Model for OMLIT Print knowledge—Spring 2005 and Spring 2006 Combined

Model term	estimate	<i>s.e.</i>	<i>t</i> -ratio	d.f.	<i>p</i> -value
Baseline score for the outcome	0.26	0.07	3.67	90	0.001
Child to teacher ratio	0.03	0.23	0.14	90	0.892
Flag for 2006 assessment	0.04	1.09	0.04	196	0.971
Study group A2 (Let's Begin & PALS)	9.05	1.73	5.23	90	0.000
Study group B2 (Partners for Literacy ECE/PE)	3.88	1.83	2.12	90	0.036
Study group A1 (Let's Begin)	3.17	1.98	1.60	90	0.112
Study group B1 (Partners for Literacy ECE)	4.39	1.92	2.29	90	0.024
Variance components	estimate	<i>s.e.</i>	Chi-square	d.f.	<i>p</i> -value
level-1 (project-year)	69.98	NA	NA	NA	NA
level-2 (project)	10.27	NA	114.85	90	0.040

Table H-16. Model for OMLIT Phonological Awareness—Spring 2005 and Spring 2006 Combined

Model term	estimate	<i>s.e.</i>	<i>t</i> -ratio	d.f.	<i>p</i> -value
Baseline score for the outcome	0.20	0.08	2.34	90	0.022
Child to teacher ratio	-0.41	0.33	-1.24	90	0.218
Flag for 2006 assessment	6.11	2.21	2.77	196	0.007
Study group A2 (Let's Begin & PALS)	7.88	3.38	2.33	90	0.022
Study group B2 (Partners for Literacy ECE/PE)	3.34	3.05	1.10	90	0.276
Study group A1 (Let's Begin)	4.00	3.06	1.31	90	0.195
Study group B1 (Partners for Literacy ECE)	3.12	2.91	1.07	90	0.287
Variance components	estimate	<i>s.e.</i>	Chi-square	d.f.	<i>p</i> -value
level-1 (project-year)	269.11	NA	NA	NA	NA
level-2 (project)	19.19	NA	99.01	90	0.242

Table H-17. Model for OMLIT Print Motivation—Spring 2005 and Spring 2006 Combined

Model term	estimate	<i>s.e.</i>	<i>t</i> -ratio	d.f.	<i>p</i> -value
Baseline score for the outcome	0.05	0.09	0.52	90	0.604
Child to teacher ratio	0.15	0.39	0.38	90	0.705
Flag for 2006 assessment	6.90	2.36	2.93	196	0.004
Study group A2 (Let's Begin & PALS)	6.18	3.11	1.99	90	0.049
Study group B2 (Partners for Literacy ECE/PE)	7.01	3.30	2.13	90	0.036
Study group A1 (Let's Begin)	0.79	2.87	0.28	90	0.783
Study group B1 (Partners for Literacy ECE)	9.87	3.30	2.99	90	0.004
Variance components	estimate	<i>s.e.</i>	Chi-square	d.f.	<i>p</i> -value
level-1 (project-year)	301.09	NA	NA	NA	NA
level-2 (project)	0.39	NA	81.64	90	>.500

Table H-18. Model for OMLIT Literacy Resources in Classroom—Spring 2005 and Spring 2006 Combined

Model term	estimate	<i>s.e.</i>	<i>t</i> -ratio	d.f.	<i>p</i> -value
Baseline score for the outcome	0.14	0.08	1.89	90	0.061
Child to teacher ratio	-0.01	0.18	-0.03	90	0.978
Flag for 2006 assessment	0.77	0.83	0.93	196	0.357
Study group A2 (Let's Begin & PALS)	6.00	1.67	3.60	90	0.001
Study group B2 (Partners for Literacy ECE/PE)	3.32	1.88	1.76	90	0.081
Study group A1 (Let's Begin)	5.28	1.71	3.08	90	0.003
Study group B1 (Partners for Literacy ECE)	3.19	1.67	1.92	90	0.058
Variance components	estimate	<i>s.e.</i>	Chi-square	d.f.	<i>p</i> -value
level-1 (project-year)	38.52	NA	NA	NA	NA
level-2 (project)	23.22	NA	192.08	90	0.000

Table H-19. Model for PE Time Spent on Child Literacy—Spring 2005 and Spring 2006 Combined

Model term	estimate	<i>s.e.</i>	<i>t</i> -ratio	d.f.	<i>p</i> -value
Child to teacher ratio	0.34	0.72	0.48	90	0.636
Flag for 2006 assessment	7.81	3.71	2.10	193	0.037
Study group A2 (Let's Begin & PALS)	26.43	6.26	4.22	90	0.000
Study group B2 (Partners for Literacy ECE/PE)	24.15	5.13	4.71	90	0.000
Study group A1 (Let's Begin)	10.54	4.93	2.14	90	0.035
Study group B1 (Partners for Literacy ECE)	6.00	5.33	1.13	90	0.264
Variance components	estimate	<i>s.e.</i>	Chi-square	d.f.	<i>p</i> -value
level-1 (project-year)	755.30	NA	NA	NA	NA
level-2 (project)	149.86	NA	123.30	90	0.011

Table H-20. Model for PE Time Spent on Parenting Skills—Spring 2005 and Spring 2006 Combined

Model term	estimate	s.e.	t-ratio	d.f.	p-value
Child to teacher ratio	0.20	0.69	0.29	90	0.776
Flag for 2006 assessment	-14.63	4.15	-3.53	193	0.001
Study group A2 (Let's Begin & PALS)	2.91	6.18	0.47	90	0.638
Study group B2 (Partners for Literacy ECE/PE)	-12.61	4.99	-2.53	90	0.014
Study group A1 (Let's Begin)	8.70	6.00	1.45	90	0.150
Study group B1 (Partners for Literacy ECE)	10.49	6.18	1.70	90	0.092
Chi-square					
Variance components	estimate	s.e.	square	d.f.	p-value
level-1 (project-year)	988.86	NA	NA	NA	NA
level-2 (project)	64.76	NA	90.00	100.1367	0.218

Table H-21. Model for PC Time Spent Interacting on Child Literacy Activities—Spring 2005 and Spring 2006 Combined

Model term	estimate	s.e.	t-ratio	d.f.	p-value
Child to teacher ratio	-0.11	0.44	-0.25	91	0.807
Flag for 2006 assessment	-4.33	2.49	-1.74	192	0.083
Study group A2 (Let's Begin & PALS)	0.51	3.38	0.15	91	0.880
Study group B2 (Partners for Literacy ECE/PE)	9.98	3.91	2.55	91	0.013
Study group A1 (Let's Begin)	0.88	2.95	0.30	91	0.767
Study group B1 (Partners for Literacy ECE)	-0.30	3.38	-0.09	91	0.931
Chi-square					
Variance components	estimate	s.e.	square	d.f.	p-value
level-1 (project-year)	360.77	NA	NA	NA	NA
level-2 (project)	4.99	NA	93.36	91	0.412

APPENDIX I

EFFECT SIZE CALCULATION

Most estimates of CLIO effects are expressed in terms of “effect sizes.” In these calculations, a variant of Glass’s Δ method was used (Glass, McGaw, and Smith 1981). As discussed by Rosenthal in his review of parametric measures of effect size (Rosenthal 1994), this method is very similar to the better known Cohen’s d . The only difference between them is in the choice of estimation method for the population standard deviation of the outcome being investigated. With Glass’s Δ method, the population standard deviation is calculated within the control group with $N-1$ as the variance divisor. For this study, the population standard deviation was calculated in the control group in spring 2005 with $N-1$ as the variance divisor. With Cohen’s d , the population standard deviation would have been calculated by pooling across all five study groups and using N as the variance divisor.

The choice of $N-1$ was made because it provides unbiased estimates of the population variance. The choice was made to use only the control group because of the possibility that the interventions could have changed population standard deviations (such as causing parents to have more similar reading-aloud habits). Data from 2006 were not used in estimating the population standard deviation because the population represented by the 2006 sample was slightly narrower than that represented by the 2005 sample. As discussed in chapter 2, the 2006 sample excluded children old enough for kindergarten.

The formula used to calculate effect sizes for child and parent outcomes with simple scoring was

$$es = \frac{f(\hat{\alpha}_1, \dots, \hat{\alpha}_5)}{\hat{S}},$$

where $f(\hat{\alpha}_1, \dots, \hat{\alpha}_5)$ is one of the tested contrasts of the adjusted treatment group means estimated by HLM using both 2005 and 2006 data, such as $\alpha_2 - \alpha_4$ (in the notation

defined in chapter 4); $\hat{S} = \sqrt{\frac{\sum_{jk} (Y_{5jk,05} - \bar{Y}_{5,05})^2}{n_{5,05} - 1}}$ is the population standard deviation of the outcome variable in the spring of 2005 in the control group; $n_{5,05}$ is the control group

sample size (children or parents) in the spring of 2005; j indicates summation across strata (which in this case are synonymous with Even Start projects since there is only one control project per stratum); k indicates summation on either parents or children without respect to classrooms or family structures; and $\bar{Y}_{5,05}$ is the average response on the outcome within the control group in the spring of 2005.

A similar formula was used for instructional outcomes, the only difference lying in the method used to calculate the population standard deviation. The formula used to calculate the population standard deviation for instructional outcomes was

$$\hat{S} = \sqrt{\frac{\sum_j (Y_{5j,05} - \bar{Y}_{5,05})^2}{n_{5,05} - 1}},$$

where $n_{5,05}$ is the control group sample size (projects) in the spring of 2005; j indicates summation across strata (which in this case are synonymous with Even Start projects since there is only one control project per stratum); and $\bar{Y}_{5,05}$ is the average level of the instructional outcome within the control group in the spring of 2005.

APPENDIX J

SAMPLE SIZES AND METHODS USED FOR MISSING DATA

This appendix provides detailed information on the methods used to handle each type of missing data.

Dropout Projects

One project dropped out of the study and one project lost funding prior to the spring 2005 data collection. Because of the deep stratification used in this study, these projects left two cells in the two-way layout of stratum and study group with sample sizes of zero. Reassuringly, we established in the simulation study discussed in chapter 3 that, if the projects dropped out completely at random, then HLM still produces valid results. Nonetheless, we thought that the analysis might work better with no empty cells. So we replaced the two projects with new projects for the second implementation year, including the spring 2006 data collection. This is another advantage to using the 2 years of data together in that there were no empty cells on the joint 2005/2006 analysis.

The only complication is that we did not have baseline project-level covariates on the two replacement projects. We used the teacher-student ratio as reported by the project director more recently, but we needed to impute the project-level average scores for the spring 2004 child assessments. We did this by assigning the stratum-level average 2004 scores to the replacement projects.¹

By the time of spring 2006 data collection, another six projects lost their Even Start grants (either by failure to reapply or through denial of funding) and thus dropped out of the study. No replacements were made for them. Another 13 projects had difficulty recruiting 3- and 4-year-olds in one year or another. This led to significant gaps in their data for the relevant year. Although we cannot rule out the possibility that treatment assignment was somehow linked to these dropouts and recruitment failures,

¹ We had to do some form of imputation in order to use data from the two replacement projects and still control on covariates in the analysis. Using the average baseline covariates from the four original projects in each of the strata with a dropout project seemed better to us than using the global average of all 118 original projects. We also considered using baseline data from the dropout projects themselves. However, we decided that average stratum-level prediction was probably a sounder approach than single-project prediction.

we also have no evidence to suggest that there was a linkage. In the end, we assumed that the missing data elements from these 21 projects were missing at random.

Four projects chose not to implement their assigned curriculum in one or both implementation years. Three of these projects were in the Partners for Literacy CLIO combined curricula study group, and one was in the Partners for Literacy CLIO preschool curricula study group. Two of the four projects did not implement their curriculum in spring 2005 and spring 2006. The other two did not implement their curriculum in spring 2006. Despite this lack of implementation, the four projects participated in data collection and were included in all analyses in their assigned study group.

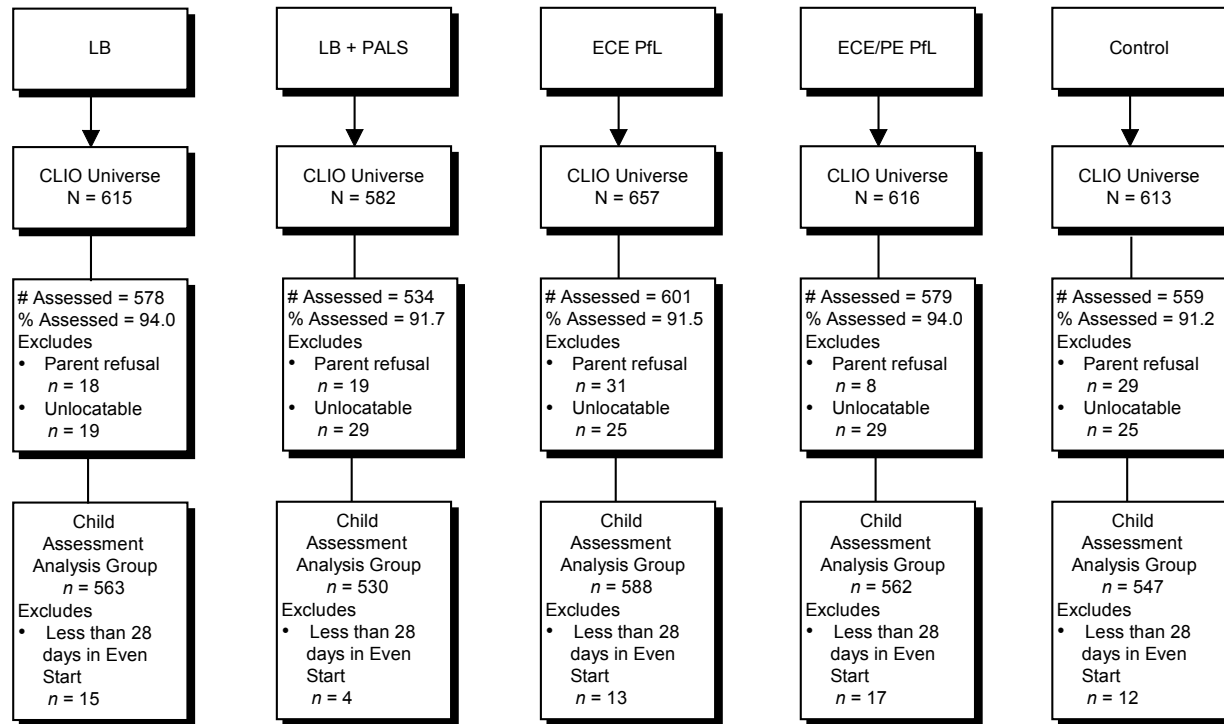
Sample Sizes

According to CLIO projects, there were approximately 3,080 preschool-aged children enrolled in CLIO study projects during the spring 2005 and spring 2006 assessment periods.² As figure J-1 illustrates, ECE Partners for Literacy projects had the largest number of children in spring 2005 and spring 2006 ($N = 657$) and Let's Begin and PALS projects had the fewest ($N = 582$). Overall, 2,851 children (or 92.5 percent) took at least one CLIO English language and literacy test during the spring 2005 and spring 2006 assessment periods. About 100 children were not assessed because of parent refusal. The remaining 127 children were enrolled in Even Start but were not at the project site at the time the CLIO assessments were administered.

The percentage of children assessed did not vary considerably by study group. The study team assessed between 91 and 92 percent of the CLIO universe in the control, ECE Partners for Literacy, and Let's Begin and PALS projects. In Let's Begin and ECE/PE Partners for Literacy projects, the study team assessed 94 percent of the CLIO universe.

² Preschool-age children are children who were at least 36 months old as of March 1 of the relevant assessment year and not yet in kindergarten.

Figure J-1. Progression of Study-Eligible Children: Spring 2005 and Spring 2006 Combined

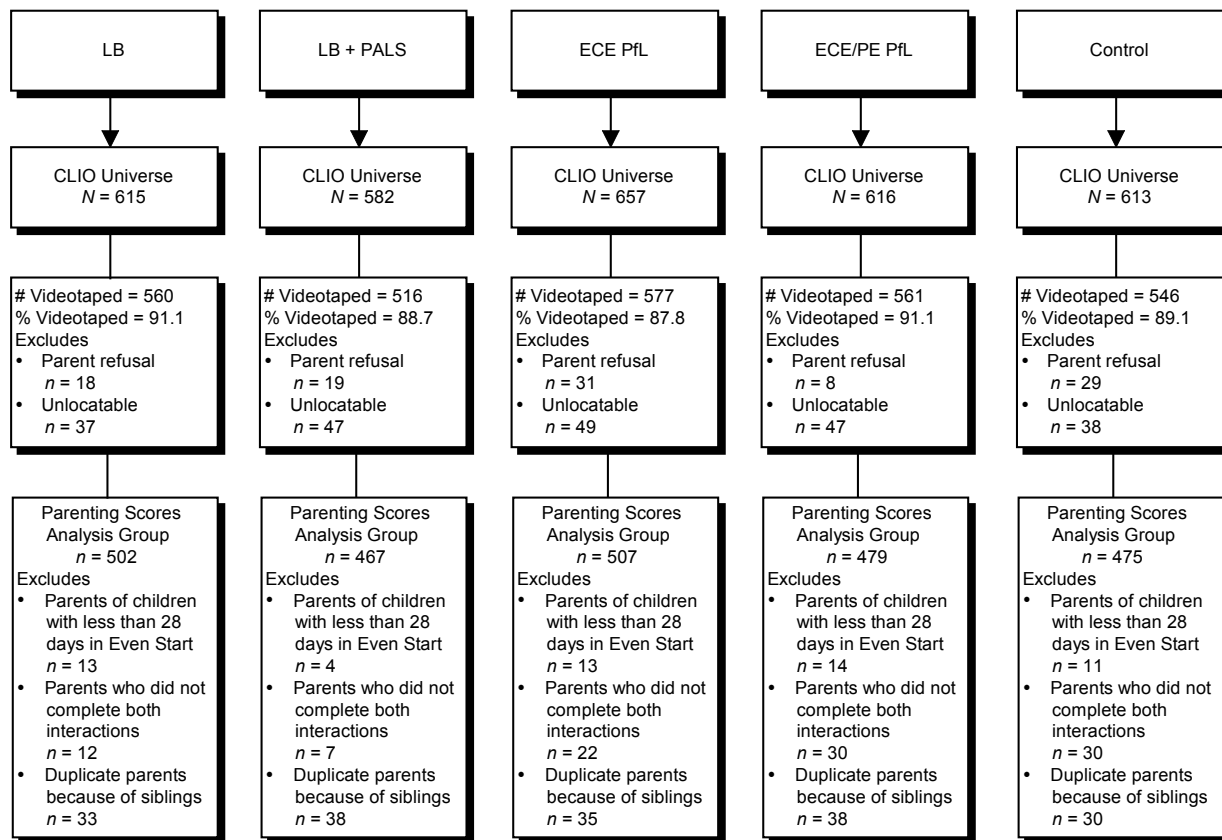


Notes: The CLIO universe was children enrolled during the spring assessment period and at least 36 months old as of March 1 of the relevant year. Number assessed consisted of children who took at least one CLIO child language and literacy test.

We limited our analyses to children who were enrolled in Even Start for 28 days or more. This reduced the sample size to 2,790 children who took at least one English language and literacy test during spring 2005 and spring 2006. This amounts to 530 children in Let's Begin and PALS projects, nearly 550 children in the control projects, about 560 children in ECE/PE Partners for Literacy and Let's Begin projects, and approximately 590 children in ECE Partners for Literacy projects.

For the parent sample, we assumed one parent per child. Thus, the CLIO parent universe equals the child universe of 3,080 individuals (see figure J-2). In looking at the progression of CLIO parents through the study, we focus on parents who participated in the videotaped parent-child interactions as this was a key source of information for the parenting responsiveness and interactive reading skills scores.

Figure J-2. Progress of Study Eligible Parents: Spring 2005 and Spring 2006 Combined



Notes: The CLIO universe was parents of children enrolled during the spring assessment period and at least 36 months old as of March 1 of the relevant year. For the universe, we assumed one parent per child, but removed duplicate parent records due to siblings for the parenting scores analysis. The parenting scores analysis also removes parents of children with little exposure to Even Start and parents who did not complete both the book and toy videotape interactions.

Across all study groups, nearly 90 percent of parents ($n=2,760$) participated in the videotaped interactions.³ This percentage did not vary noticeably by study group. For the parenting scores analysis, we excluded parents of children with less than 28 days in Even Start. We also removed parents who did not complete both the read aloud and toy portions of the interaction since the responsiveness measure drew on elements of the interaction common across both portions such as reciprocal warmth and affection. Finally, for parents with multiple children, we averaged the parenting scores, thereby dropping duplicate parent records due to siblings.

³ For parents with multiple sample children, the staged interaction usually involved the parent and all her/his children. However, the coding was done separately by child.

As a result, 2,430 parents were available for the parenting scores analysis. The sample sizes ranged from 467 parents from Let's Begin and PALS projects to 507 parents from the ECE Partners for Literacy projects.

Table J-1 shows additional 2-year sample sizes for the various outcome scales and for selected combinations of them. The sample sizes reflect children with parental consent who were enrolled for at least 28 calendar days as of their child assessment date. The number of children whose data were used in any analysis was 2,941. Somewhat smaller numbers were available for each specific analysis. The smallest sample size was 1,539 for the Spanish expressive language assessment, which was only administered to children whom project staff identified as being from Spanish-speaking homes. The largest sample was 2,930 children with data on level of parent participation. We obtained "substantial" data (something other than participation data or a parent interview) on 2,931 children.

Incomplete Child Assessment Battery

We have partial data for the English-language emergent literacy assessments on 31 children. For the ITT analyses, we used the data on all children with data on any particular assessment, so the sample size for each subtest was slightly different. Although we have substantial information of one type or another (mostly social competence scores) on another 141 children who were not assessed, we did not make use of that information for any adjustments. Instead, we assumed that the children who completed a subtest are a random sample of all those eligible to take the subtest at the study projects. There was no imputation of entire subtests. In addition, there are children whose subtests were ended prematurely for some reason such as fatigue or misbehavior. The IRT scoring procedure uses the partial information obtained in these subtests to develop subtest scores. When we instead use raw scores, the counts of children with completed subtests drops slightly because prematurely ended subtests do not produce valid raw scores. The difference can be seen by comparing the second and fourth rows of table J-1. Using the IRT scoring procedures allowed us to use partial assessments on 127 (=2,759-2,632) children for whom it was impossible to calculate raw scores.

Table J-1. Child and Parent Sample Sizes, Spring 2005 and Spring 2006 Combined

Scale	Child sample size	Parent sample size
1) At least one of the 6 English emergent-literacy IRT scores	2,790	
2) All of the 6 English emergent-literacy IRT scores	2,759	
3) At least one of the 6 English emergent-literacy raw scores	2,790	
4) All of the 6 English emergent-literacy raw scores	2,632	
5) Spanish expressive language score	1,539	
6) Social competence IRT score	2,864	
7) Social competence raw score	2,729	
8) Child participation	2,919	
9) Parenting scores	2,604	2,430
10) Parent reading and vocabulary score	2,720	2,531
11) Parent participation	2,930	2,728
12) "Substantial" (any child emergent literacy, social competence, parenting or parent literacy)	2,931	
13) Any	2,941	

NOTE: Numbers include families with actual or imputed parent interviews.

Missing Teacher Rating Forms

A total of 2,864 social competence IRT scores were derived from teacher rating forms for children with parental consent. (The IRT modeling procedure yielded a social competence score provided that at least one item is rated.) Using a raw scoring procedure, the number is 125 smaller.

Missing Parent Assessment

We had 2,531 parents who took at least one of the four English language literacy assessments that are used to build the parent reading and vocabulary skill outcome. Of these, 2,530 completed all four. To impute the single missing parent

assessment, we used a hotdeck with cells defined by study group and quintiles of the average of however many of the English-language parent-literacy assessments were taken. Within a cell, the single parent with a missing parent-literacy assessments was randomly matched to a parent who completed all four. The one missing assessment score was then cloned, and the cloned record was associated with the parent who took some but not all of the assessments.

Missing Parent Interviews

A total of 2,559 parent interviews were collected in spring 2005 and spring 2006 among children with consent, a minimum of 28 calendar days enrollment as of emergent-literacy assessment date, and substantial other information. Since there were 2,739 parents with children with substantial data, there was a need to account for 180 missing parent interviews. The corresponding number of children was 181 on whom substantial data had been collected but for whom no parent interview was obtained. As mentioned above, parent interviews are needed to compute covariates that were used in the ITT analyses and to compute the two parenting outcomes. Rather than discard the substantial data, whole interviews were imputed for these children. Specifically, we imputed entire parent interviews for all consented eligible children where the parent interview was missing but either 1) at least one of the English child emergent-literacy assessments was completed, 2) at least one of the English parent literacy assessments was completed, 3) both the book and toy videos were completed, or 4) the teacher rating form was completed.

Since some children attend Even Start centers for more than a year, we first searched to see if a parent interview was obtained at one of the prior rounds of data collection. If so, we simply cloned that parent interview into the round where it was missing. For the remaining missing parent interviews, we cloned randomly selected parent interviews from similar families in the same experimental group. Similarity within an experimental group was defined in terms of child emergent literacy, if available. Otherwise, it was defined in terms of child social competence. Using these procedures, we were able to boost the analyzable sample sizes as shown in table J-2.

Table J-2. Increases in Analyzable Child and Parent Sample Sizes Due to Imputation of Whole Parent Interviews

Scale	Child sample size	Parent sample size
Expressive language: English	72	
Expressive language: Spanish	21	
Receptive vocabulary	72	
Phonological awareness: Elision	72	
Phonological awareness: Blending	72	
Letter and sound recognition (Print Awareness)	72	
Syntax and grammar	72	
Social competence	178	
Parent interactive reading skill		5
Parent responsiveness		5
Parent reading & vocabulary skill		8
Child: Monthly hours of ECE instruction received	178	
Parent: Monthly hours of PE and PC instruction received	178	

Incomplete Parent Interviews

In examining item nonresponse rates on the parent interview, family income is one of the most frequently missing items, with 9 percent of respondents failing to provide the information. Since modeling procedures require nonmissing data on all covariates, power was maximized by imputing all the covariates that we used in the analysis. This was done with new software that imputes all missing items within an entire questionnaire in one operation. This software was designed to simultaneously preserve complex covariance structures across items, reduce nonresponse bias on marginal means for individual items, and preserve the natural discreteness, smoothness, or lumpiness of each individual item. Further information on this software may be found in Piesse, Judkins, and Fan (2005). Although this software can be used to make multiple imputations, only a single imputation for each missing datum was saved and used in the analyses.

Missing Videotapes of Parent-Child Read-Together Session or Toy-Play Session

A total of 2,604 parent-child dyads were videotaped in both session types in spring 2005 or spring 2006. There were 327 dyads that were missing the videotape data on at least one session but for whom we had at least some information about emergent literacy, social competence, or parent reading and vocabulary skill. We did not use these data. Instead, we assumed that the 2,604 are a random sample of all eligible dyads attending the study projects at the time. The impact analysis for parenting was based on the double sessions for the 2,430 unique parents in these dyads.

Missing Observation Measures of Instruction

We were not always able to schedule an observation trip to coincide with classes, particularly for the parenting education and parent-child classes, as these were infrequent at some projects. Missing classroom observations occurred more often in projects with CLIO curricula than projects in the control group. Across spring 2005 and spring 2006, seven projects were missing preschool observations (see table J-3). Three of the seven projects were in the Partners for Literacy CLIO combined curricula group; two were in the Let's Begin CLIO preschool curricula group; one was in the Partners for Literacy CLIO preschool curricula group; and one was in the control group.

Table J-3. Projects with No Observation Data by Observation Type and Year of Missing Data

Year(s) of missing data	Observation type		
	Preschool	Parenting education	Parent-child
Spring 2005	3	3	7
Spring 2006	4	6	5
Spring 2005 and Spring 2006	0	1	0

NOTE: In spring 2005, 118 projects participated in the study, and in spring 2006, 114 projects participated. Across the 2 years, 120 projects participated in the study providing data for one or both years.

We had no parenting education observation data in spring 2005 or spring 2006 for nine projects. Additionally, for one project we did not have an observation in either year. Four of the projects missing these data were in the Let's Begin CLIO

preschool curricula group; three were in the Partners for Literacy CLIO combined curricula group; two were in the Partners for Literacy CLIO preschool curricula group; and one was in the control group.

Over the 2 implementation years, we did not observe parent-child classrooms in 12 projects. Four of these projects were in the Let's Begin CLIO preschool curricula group; three were in the Let's Begin and PALS CLIO combined curricula group; two were in the Partners for Literacy CLIO combined curricula group; one was in the Partners for Literacy CLIO preschool curricula group; and two were in the control group.

We made no adjustment for the missing data from the one project missing parenting education in both years. For projects with only 1 year of observation data, we used the non-missing data for the combined analysis.

Missing Participation Data

Despite our quality assurance, there was some uncertainty on whether the absence of a child from a particular monthly ISPF indicates missing data or zero participation. There were enough of these blanks that some kind of imputation was required in order to have an acceptable sample size for analysis. For months outside the official enrollment period, we assumed zero participation. For months within the official enrollment period for a child, we considered whether we had participation hours of any type (ECE, PE, PC, or AE) for the month. If we had one type of participation hours for a child's family in a month, then we assumed that any missing reports for other forms of participation by that child's family were actually reports of zero participation. If, on the other hand, we had no participation data of any sort for a child's family in a particular month, but we did have such data in other months, then we imputed the child's average (by type of participation) for other months to the month with a blank report. This also covered the cases where a project failed to provide any data at all, since there were no children who were only enrolled for a month when the project failed to provide participation data. In the baseline year (2003-04) 4.7 percent of children required imputation for missing participation data for at least 1 month; in the first implementation year (2004-05) 1.9 percent of children required imputation; and in the second implementation year (2005-06) 4.9 percent required imputation. In addition to imputing missing participation data, we deleted outlying values, where outlying was

defined in terms of average participation during months enrolled rather than average enrollment across all nine reference months.

Missing Baseline Project Average Scores

We discussed previously the procedure for imputing baseline project-average scores for the replacement projects. However, there were other instances where we needed to impute baseline project-average scores. Since some of the projects had small enrollments, it sometimes happened that no one took one of the assessments in the spring of 2004 or the fall of 2003. This was particularly common with the Spanish version of the expressive language assessment because there several projects where the first Spanish-speaking families were enrolled after baseline assessments. In order to be able to use the baseline project average of an assessment as a covariate in the analysis of 2005 and 2006 assessments, it was necessary to impute how well the children at the project would have tested on each missing assessment had any taken it. We did this by building multivariate models for each of the baseline assessments in terms of the assessments that were administered and other variables. These models were fit at the project level. Demographic variables were averaged up to the project level so that we can use variables like the ethnicity mixture of a project without having to fit multi-level models. Predictions from the models were used as imputed values without adding on any random variation.

APPENDIX K

ITT TABLES FOR SINGLE DEVELOPER CONTRASTS

Table K-1. Effects of CLIO Combined Curricula on Children (Research Question 1) (average of spring 2005 and spring 2006)

Child outcomes	Let's Begin and PALS vs. Partners for Literacy (ECE/PE)			Let's Begin and PALS (ECE/PE) vs. control			Partners for Literacy (ECE/PE) vs. control		
	Effect size	95% CI	p-Value	Effect size	95% CI	p-Value	Effect size	95% CI	p-Value
Expressive language: English	-0.04	-0.22,0.14	>.500	-0.14	-0.30,0.03	0.102	-0.09	-0.25,0.06	0.216
Expressive language: Spanish	0.02	-0.22,0.26	>.500	0.06	-0.17,0.29	>.500	0.04	-0.20,0.28	>.500
Receptive vocabulary	-0.04	-0.17,0.09	>.500	-0.11	-0.23,0.02	0.090	-0.07	-0.21,0.07	>.500
Phonological awareness: Elision	0.15	-0.01,0.30	0.064	0.07	-0.09,0.24	>.500	-0.07	-0.23,0.09	>.500
Phonological awareness: Blending	0.00	-0.17,0.17	>.500	-0.13	-0.29,0.03	0.110	-0.13	-0.30,0.05	0.162
Print knowledge	0.22	0.04,0.40	0.013	0.16	0.01,0.31	0.036	-0.06	-0.24,0.11	>.500
Syntax and grammar	0.08	-0.06,0.22	0.261	-0.04	-0.18,0.09	>.500	-0.12	-0.26,0.01	0.070
Social competence	0.12	-0.05,0.29	0.166	0.28*	0.11,0.45	0.001	0.16	-0.01,0.33	0.064

*Statistically significant at $p < 0.0056$, a value chosen according to Bonferroni's rule to control false positive discovery rates across the nine contrasts run for each outcome.

Table K-2. Incremental Effects of CLIO Parenting Curricula on Children (Research Question 2) (average of spring 2005 and spring 2006)

Child outcomes	[Let's Begin and PALS (ECE/PE) Minus Let's Begin (ECE)] vs. [Partners for Literacy (ECE/PE) Minus Partners for Literacy (ECE)]			Let's Begin and PALS (ECE/PE) vs. Let's Begin (ECE)			Partners for Literacy (ECE/PE) vs. Partners for Literacy (ECE)		
	Effect size	95% CI	p-Value	Effect size	95% CI	p-Value	Effect size	95% CI	p-Value
Expressive language: English	0.06	-0.15,0.27	>.500	-0.01	-0.16,0.14	>.500	-0.07	-0.21,0.07	0.308
Expressive language: Spanish	-0.12	-0.46,0.22	>.500	-0.01	-0.25,0.23	>.500	0.11	-0.13,0.36	>.500
Receptive vocabulary	0.01	-0.16,0.18	>.500	-0.05	-0.17,0.08	>.500	-0.06	-0.18,0.07	>.500
Phonological awareness: Elision	0.11	-0.11,0.34	>.500	0.08	-0.07,0.23	0.287	-0.03	-0.20,0.13	>.500
Phonological awareness: Blending	0.12	-0.11,0.35	0.292	-0.01	-0.16,0.14	>.500	-0.13	-0.29,0.03	0.102
Print knowledge	0.25	0.03,0.48	0.025	0.01	-0.15,0.17	>.500	-0.24*	-0.40,-0.08	0.003
Syntax and grammar	0.09	-0.09,0.27	0.314	-0.01	-0.16,0.14	>.500	-0.10	-0.23,0.03	0.115
Social competence	0.17	-0.06,0.41	0.141	0.08	-0.09,0.25	>.500	-0.09	-0.27,0.08	0.291

*Statistically significant at $p < 0.0056$, a value chosen according to Bonferroni's rule to control false positive discovery rates across the nine contrasts run for each outcome.

Table K-3. Effects of CLIO Combined Curricula on Parents (Research Question 1) (average of spring 2005 and spring 2006)

Parent outcomes	Let's Begin and PALS vs. Partners for Literacy (ECE/PE)			Let's Begin and PALS (ECE/PE) vs. control			Partners for Literacy (ECE/PE) vs. control		
	Effect size	95% CI	<i>p</i> -Value	Effect size	95% CI	<i>p</i> -Value	Effect size	95% CI	<i>p</i> -Value
Interactive reading skill	-0.05	-0.11,0.01	0.072	0.45*	0.25,0.65	0.000	0.50*	0.29,0.71	0.000
Responsiveness to child	0.09	-0.08,0.25	0.303	0.26*	0.12,0.41	0.001	0.18	0.01,0.35	0.032
Reading skills and vocabulary	0.02	-0.11,0.14	>.500	-0.04	-0.15,0.08	>.500	-0.05	-0.17,0.07	>.500

*Statistically significant at $p < 0.0056$, a value chosen according to Bonferroni's rule to control false positive discovery rates across the nine contrasts run for each outcome.

Table K-4. Incremental Effects of CLIO Parenting Curricula on Parents (Research Question 2) (average of spring 2005 and spring 2006)

Parent outcomes	[Let's Begin and PALS (ECE/PE) Minus Let's Begin (ECE)] vs. [Partners for Literacy (ECE/PE) Minus Partners for Literacy (ECE)]			Let's Begin and PALS (ECE/PE) vs. Let's Begin (ECE)			Partners for Literacy (ECE/PE) vs. Partners for Literacy (ECE)		
	Effect size	95% CI	<i>p</i> -Value	Effect size	95% CI	<i>p</i> -Value	Effect size	95% CI	<i>p</i> -Value
Interactive reading skill	-0.35	-0.64,-0.05	0.018	0.12*	0.08,0.17	0.000	0.48*	0.29,0.67	0.000
Responsiveness to child	-0.15	-0.38,0.09	0.212	0.08	-0.09,0.25	>.500	0.23	0.06,0.40	0.006
Reading skills and vocabulary	-0.09	-0.26,0.08	0.285	-0.10	-0.22,0.02	0.092	-0.02	-0.14,0.11	>.500

*Statistically significant at $p < 0.0056$, a value chosen according to Bonferroni's rule to control false positive discovery rates across the nine contrasts run for each outcome.

Table K-5. Effects of CLIO Combined Curricula on Instruction (Research Question 1) (average of spring 2005 and spring 2006)

Instructional outcomes	Let's Begin and PALS vs. Partners for Literacy (ECE/PE)			Let's Begin and PALS (ECE/PE) vs. control			Partners for Literacy (ECE/PE) vs. control		
	Effect size	95% CI	p-Value	Effect size	95% CI	p-Value	Effect size	95% CI	p-Value
Support for oral language development	0.14	-0.19,0.48	>.500	0.37	-0.01,0.74	0.048	0.22	-0.12,0.56	0.197
Support for print knowledge	0.55*	0.23,0.87	0.001	0.96*	0.60,1.33	0.000	0.41	0.03,0.80	0.032
Support for phonological awareness	0.43	-0.24,1.09	0.200	0.74	0.11,1.37	0.019	0.31	-0.26,0.88	0.272
Support for print motivation	-0.05	-0.39,0.30	>.500	0.34	0.00,0.68	0.044	0.38	0.03,0.74	0.031
Literacy resources in the classroom	0.30	-0.07,0.67	0.104	0.67*	0.30,1.03	0.001	0.37	-0.05,0.78	0.074
PE time spent on child literacy	0.09	-0.44,0.62	>.500	1.06*	0.56,1.55	0.000	0.96*	0.56,1.37	0.000
PE time spent on parenting skills	0.49	0.13,0.85	0.008	0.09	-0.29,0.48	>.500	-0.40	-0.71,-0.08	0.011
PC time spent interacting on child literacy activities	-0.38	-0.70,-0.05	0.020	0.02	-0.25,0.29	>.500	0.40	0.09,0.70	0.010

*Statistically significant at $p < 0.0056$, a value chosen according to Bonferroni's rule to control false positive discovery rates across the nine contrasts run for each outcome.

Table K-6. Incremental Effects of CLIO Parenting Curricula on Instruction (Research Question 2) (average of spring 2005 and spring 2006)

Instructional outcomes	[Let's Begin and PALS (ECE/PE) Minus Let's Begin (ECE)] vs. [Partners for Literacy (ECE/PE) Minus Partners for Literacy (ECE)]			Let's Begin and PALS (ECE/PE) vs. Let's Begin (ECE)			Partners for Literacy (ECE/PE) vs. Partners for Literacy (ECE)		
	Effect size	95% CI	p-Value	Effect size	95% CI	p-Value	Effect size	95% CI	p-Value
Support for oral language development	0.37	-0.08,0.83	0.098	0.13	-0.21,0.46	>.500	-0.25	-0.56,0.06	0.110
Support for print knowledge	0.68	0.16,1.20	0.009	0.63*	0.27,0.98	0.001	-0.05	-0.41,0.30	>.500
Support for phonological awareness	0.34	-0.55,1.24	>.500	0.37	-0.30,1.03	0.272	0.02	-0.55,0.59	>.500
Support for print motivation	0.45	-0.02,0.92	0.051	0.30	0.00,0.59	0.042	-0.16	-0.52,0.21	>.500
Literacy resources in the classroom	0.07	-0.44,0.57	>.500	0.08	-0.26,0.42	>.500	0.01	-0.37,0.40	>.500
PE time spent on child literacy	-0.09	-0.78,0.60	>.500	0.63	0.12,1.15	0.014	0.73*	0.26,1.19	0.002
PE time spent on parenting skills	0.54	-0.01,1.10	0.049	-0.18	-0.60,0.23	>.500	-0.73*	-1.10,-0.35	0.000
PC time spent interacting on child literacy activities	-0.42	-0.84,0.00	0.042	-0.01	-0.27,0.24	>.500	0.41	0.08,0.73	0.013

*Statistically significant at $p < 0.0056$, a value chosen according to Bonferroni's rule to control false positive discovery rates across the nine contrasts run for each outcome.

Table K-7. Effects of CLIO Combined Curricula on Participation (Research Question 1) (average of spring 2005 and spring 2006)

	Let's Begin and PALS vs. Partners for Literacy (ECE/PE)			Let's Begin and PALS (ECE/PE) vs. control			Partners for Literacy (ECE/PE) vs. control		
Participation outcomes	Effect size	95% CI	<i>p</i> -Value	Effect size	95% CI	<i>p</i> -Value	Effect size	95% CI	<i>p</i> -Value
Child: monthly hours of ECE instruction received	0.03	-0.21,0.28	>.500	0.14	-0.09,0.37	0.224	0.11	-0.12,0.34	>.500
Parent: monthly hours of PE/PC instruction received	-0.12	-0.46,0.21	>.500	0.19	-0.14,0.52	0.246	0.31	-0.03,0.65	0.064

*Statistically significant at $p < 0.0056$, a value chosen according to Bonferroni's rule to control false positive discovery rates across the nine contrasts run for each outcome.

Table K-8. Incremental Effects of CLIO Parenting Curricula on Participation (Research Question 2) (average of spring 2005 and spring 2006)

	[Let's Begin and PALS (ECE/PE) Minus Let's Begin (ECE)] vs. [Partners for Literacy (ECE/PE) Minus Partners for Literacy (ECE)]			Let's Begin and PALS (ECE/PE) vs. Let's Begin (ECE)			Partners for Literacy (ECE/PE) vs. Partners for Literacy (ECE)		
Participation outcomes	Effect size	95% CI	<i>p</i> -Value	Effect size	95% CI	<i>p</i> -Value	Effect size	95% CI	<i>p</i> -Value
Child: monthly hours of ECE instruction received	0.10	-0.28,0.47	>.500	0.21	-0.07,0.49	0.125	0.12	-0.12,0.35	>.500
Parent: monthly hours of PE/PC instruction received	-0.33	-0.96,0.30	>.295	0.00	-0.48,0.48	>.500	0.33	-0.06,0.73	0.088

*Statistically significant at $p < 0.0056$, a value chosen according to Bonferroni's rule to control false positive discovery rates across the nine contrasts run for each outcome.

APPENDIX L

GROWTH ANALYSIS

In addition to looking at the impacts of the CLIO curricula on the status of children's emergent literacy skills at the end of preschool, we examined impacts on the pattern of growth in children's language skills from fall to spring. That is, we asked if the CLIO curricula affected the *rate* at which children obtained new skills.

Methods. We ran the growth analysis on the six English emergent literacy scales, child social competence, the two parenting scales, and parent reading and vocabulary skill. We used HLM as in the primary analysis but fit two-level models rather than three-level models, since we could not include year. The covariates were the same as were used in the regression models for the primary analysis, as were the contrasts that were tested and the adjustments for multiple comparisons.

The general form of the growth model¹ for child and parent outcomes was

$$G_{ijk} = \alpha_i + \beta_j + X_{ijk}\gamma + Z_{ij}\delta + a_{ij} + e_{ijk} ,$$

where:

- the indices stand respectively for study group (i), stratum (j), and child or parent (k);
- the terms in Greek letters are fixed effects (α for treatment effect, β for stratum effect, γ for effects of family and child covariates, and δ for effects of measured project covariates);
- the terms in lower-case Latin letters are random effects (a for project-level random error and e for year-specific child- or parent-level random error);
- the terms in upper-case Latin characters are measured variables (G for outcome growth from fall to spring, X for child- and/or family-level covariate row vector, and Z for project-level covariate row vector);

¹ Since there is just one project per combination of study group and stratum, there is no need for a separate index for project. In the language of HLM documentation, this model description is equivalent to saying that we used a three-level linear model with covariates at the person and project levels.

- random effects at each level are assumed to be independently and identically normally distributed; and
- random effects at different levels are assumed to be independent of each other.

Caveats. We know that there is a bias toward understating CLIO effects with this analysis because the fall measurements were made far enough into the program year so that children might have already received some benefit from the CLIO curricula. Another source of bias is the fact that only long-term participants could have both fall and spring assessments. Although there is no evidence that the CLIO curricula affect participation in any manner, the possibility of some small biases cannot be ruled out.

The growth analyses are based on a substantially smaller sample of children than the primary ITT analyses. The exact sample sizes vary by outcome, but taking the IGD1 test of expressive vocabulary for example, the primary ITT analysis has a sample size of 2,785 child assessments over 2 years on a unique set of 2,483 children across 120 projects. In contrast, the growth analysis was based on double assessments of 1,098 unique children across 113 projects. The smaller sample size is attributable to a combination of fall-to-spring turnover in the families served and low attendance, which reduced the chances of obtaining two data points on children who were enrolled at both time points. The reduced sample size decreases the power of the analyses. On the other hand, power concerns might be mitigated by the fact that test reliability was in the range of 80 to 92 percent, so if the overlap were high enough, and most of the benefits came later rather than early, it was possible that a change-score analysis could be more powerful.

Results. Only 1 of the 10 overall tests² was significant when we ran the growth analysis on the six English emergent literacy scales, child social competence, the two parenting scales, and parent reading and vocabulary skill (table L-1). Moreover, only 1 of 80 contrasts (10 outcomes * 8 contrasts per outcome) was significant after Bonferroni correction (tables L-2 through L-5).³ Both the significant overall test and the

² By overall test, we mean a test of whether any of the five study groups are different from each other.

³ The same eight contrasts were run for each outcome scale as discussed in chapter 5. A Bonferroni adjustment of 9 was applied, meaning that the *p*-value for a contrast had to be smaller than 0.05/9=0.0055 in order to be considered significant. This is the same Bonferroni adjustment as used in chapter 5.

significant contrast were for parent responsiveness. Recall that effects were already found for this outcome scale in the primary ITT analysis—the average of the CLIO combined curricula was found to be better than the control curriculum in promoting parent responsiveness. In the growth analysis, the comparison of the CLIO combined curricula with the control group lost significance, but it was replaced by a significant contrast between the CLIO combined curricula with the average of the CLIO preschool curricula.

Most of the covariates that were important in the ITT analysis of end-of-year status were not significant in the growth analysis. Only a few family and child characteristics are associated with differential growth in language skills from fall to spring of a single school year. Among those we examined, only ethnicity and number of adults in the household have predictive value for growth over the year. Hispanic children improve on their receptive and expressive English vocabularies faster than other children. Also children in households with multiple adults show a higher growth rate in expressive vocabulary.

Table L-1. Results of Overall Test for Any Differences in Growth Across the Five Study Groups (Fall 2004 to Spring 2005)

Outcome measure	<i>p</i> -value for overall test across 5 groups
Child outcomes	
Expressive language: English	>.500
Receptive vocabulary	>.500
Phonological awareness: Elision	0.282
Phonological awareness: Blending	>.500
Print knowledge	0.374
Syntax and grammar	>.500
Child social competence	>.500
Parent outcomes	
Parent interactive reading skill	0.091
Parent responsiveness to child	0.009*
Parent reading and vocabulary	>.500

**Table L-2a. Effects of CLIO Combined Curricula on Children (Research Question 1)
(Growth Fall 2004 to Spring 2005)**

Child outcomes	Average of the two CLIO combined curricula vs. control		
	Effect size	95% CI	<i>p</i> -Value
Expressive language: English	0.04	(-0.09,0.17)	>.500
Receptive vocabulary	-0.09	(-0.26,0.09)	0.311
Phonological awareness: Elision	-0.14	(-0.48,0.20)	>.500
Phonological awareness: Blending	-0.18	(-0.43,0.07)	0.152
Print knowledge	0.09	(-0.14,0.32)	>.500
Syntax and grammar	-0.05	(-0.37,0.27)	>.500
Social competence	0.10	(-0.22,0.41)	>.500

*Statistically significant at $p < 0.0056$, a value chosen according to Bonferroni's rule to control false positive discovery rates across the nine contrasts run for each outcome.

Table L-2b. Effects of CLIO Combined Curricula on Children (Research Question 1) (Growth Fall 2004 to Spring 2005)

Child outcomes	Let's Begin and PALS vs. Partners for Literacy (ECE/PE)			Let's Begin and PALS (ECE/PE) vs. control			Partners for Literacy (ECE/PE) vs. control		
	Effect size	95% CI	<i>p</i> -Value	Effect size	95% CI	<i>p</i> -Value	Effect size	95% CI	<i>p</i> -Value
Expressive language: English	-0.07	(-0.21,0.06)	0.284	0.01	(-0.13,0.15)	>.500	0.08	(-0.08,0.24)	0.312
Receptive vocabulary	0.02	(-0.21,0.25)	>.500	-0.08	(-0.28,0.12)	>.500	-0.10	(-0.32,0.12)	>.500
Phonological awareness: Elision	0.18	(-0.12,0.48)	0.233	-0.05	(-0.38,0.28)	>.500	-0.23	(-0.63,0.18)	0.257
Phonological awareness: Blending	-0.07	(-0.36,0.23)	>.500	-0.21	(-0.49,0.07)	0.130	-0.14	(-0.44,0.15)	>.500
Print knowledge	0.16	(-0.07,0.39)	0.166	0.17	(-0.10,0.44)	0.198	0.01	(-0.23,0.25)	>.500
Syntax and grammar	0.14	(-0.20,0.49)	>.500	0.02	(-0.32,0.37)	>.500	-0.12	(-0.51,0.26)	>.500
Social competence	-0.18	(-0.52,0.16)	0.293	0.01	(-0.34,0.36)	>.500	0.19	(-0.18,0.55)	0.312

*Statistically significant at $p < 0.0056$, a value chosen according to Bonferroni's rule to control false positive discovery rates across the nine contrasts run for each outcome.

Table L-3a. Incremental Effects of CLIO Parenting Curricula on Children (Research Question 2) (Growth Fall 2004 to Spring 2005)

Child outcomes	Average of the two CLIO combined curricula vs. Average of the two CLIO preschool curricula		
	Effect size	95% CI	p-Value
Expressive language: English	-0.01	(-0.10,0.08)	>.500
Receptive vocabulary	-0.09	(-0.26,0.07)	0.254
Phonological awareness: Elision	0.10	(-0.16,0.37)	>.500
Phonological awareness: Blending	0.02	(-0.19,0.23)	>.500
Print knowledge	0.01	(-0.18,0.20)	>.500
Syntax and grammar	-0.09	(-0.36,0.18)	>.500
Social competence	0.09	(-0.16,0.34)	>.500

*Statistically significant at $p < 0.0056$, a value chosen according to Bonferroni's rule to control false positive discovery rates across the nine contrasts run for each outcome.

Table L-3b. Incremental Effects of CLIO Parenting Curricula on Children (Research Question 2) (Growth Fall 2004 to Spring 2005)

Child outcomes	[Let's Begin and PALS (ECE/PE) Minus Let's Begin (ECE)] vs. [Partners for Literacy (ECE/PE) Minus Partners for Literacy (ECE)]			Let's Begin and PALS (ECE/PE) vs. Let's Begin (ECE)			Partners for Literacy (ECE/PE) vs. Partners for Literacy (ECE)		
	Effect size	95% CI	p-Value	Effect size	95% CI	p-Value	Effect size	95% CI	p-Value
Expressive language: English	0.00	(-0.19,0.19)	>.500	-0.01	(-0.15,0.12)	>.500	-0.01	(-0.14,0.12)	>.500
Receptive vocabulary	-0.09	(-0.40,0.21)	>.500	-0.14	(-0.39,0.11)	0.263	-0.05	(-0.24,0.14)	>.500
Phonological awareness: Elision	0.22	(-0.19,0.63)	0.275	0.21	(-0.16,0.58)	0.245	-0.01	(-0.30,0.28)	>.500
Phonological awareness: Blending	-0.10	(-0.52,0.32)	>.500	-0.03	(-0.28,0.22)	>.500	0.08	(-0.26,0.41)	>.500
Print knowledge	0.00	(-0.33,0.33)	>.500	0.01	(-0.22,0.24)	>.500	0.01	(-0.26,0.29)	>.500
Syntax and grammar	0.17	(-0.27,0.62)	>.500	-0.01	(-0.39,0.38)	>.500	-0.18	(-0.49,0.13)	0.242
Social competence	-0.06	(-0.52,0.39)	>.500	0.06	(-0.27,0.38)	>.500	0.12	(-0.23,0.47)	>.500

*Statistically significant at $p < 0.0056$, a value chosen according to Bonferroni's rule to control false positive discovery rates across the nine contrasts run for each outcome.

Table L-4a. Effects of CLIO Combined Curricula on Parents (Research Question 1) (Growth Fall 2004 to Spring 2005)

Parent outcomes	Average of the two CLIO combined curricula vs. control		
	Effect size	95% CI	<i>p</i> -Value
Interactive reading skill	0.10	(-0.14,0.35)	>.500
Responsiveness to child	-0.07	(-0.26,0.11)	>.500
Reading skills and vocabulary	-0.09	(-0.18,0.00)	0.046

*Statistically significant at $p<0.0056$, a value chosen according to Bonferroni's rule to control false positive discovery rates across the nine contrasts run for each outcome.

Table L-4b. Effects of CLIO Combined Curricula on Parents (Research Question 1) (Growth Fall 2004 to Spring 2005)

Parent outcomes	Let's Begin and PALS vs. Partners for Literacy (ECE/PE)			Let's Begin and PALS (ECE/PE) vs. control			Partners for Literacy (ECE/PE) vs. control		
	Effect size	95% CI	<i>p</i> -Value	Effect size	95% CI	<i>p</i> -Value	Effect size	95% CI	<i>p</i> -Value
Interactive reading skill	-0.20	(-0.52,0.11)	0.200	0.00	(-0.24,0.25)	>.500	0.20	(-0.13,0.54)	0.222
Responsiveness to child	0.12	(-0.03,0.28)	0.105	-0.01	(-0.21,0.19)	>.500	-0.14	(-0.33,0.06)	0.169
Reading skills and vocabulary	0.02	(-0.08,0.11)	>.500	-0.08	(-0.18,0.02)	0.111	-0.10	(-0.20,0.01)	0.058

*Statistically significant at $p<0.0056$, a value chosen according to Bonferroni's rule to control false positive discovery rates across the nine contrasts run for each outcome.

Table L-5a. Incremental Effects of CLIO Parenting Curricula on Parents (Research Question 2) (Growth Fall 2004 to Spring 2005)

Parent outcomes	Average of the two CLIO combined curricula vs. average of the two CLIO preschool curricula		
	Effect size	95% CI	<i>p</i> -Value
Interactive reading skill	0.22	(0.02,0.41)	0.025
Responsiveness to child	0.16*	(0.05,0.26)	0.003
Reading skills and vocabulary	-0.04	(-0.12,0.04)	>.500

*Statistically significant at $p < 0.0056$, a value chosen according to Bonferroni's rule to control false positive discovery rates across the nine contrasts run for each outcome.

Table L-5b. Incremental Effects of CLIO Parenting Curricula on Parents (Research Question 2) (Growth Fall 2004 to Spring 2005)

Parent outcomes	[Let's Begin and PALS (ECE/PE) Minus Let's Begin (ECE)] vs. [Partners for Literacy (ECE/PE) Minus Partners for Literacy (ECE)]			Let's Begin and PALS (ECE/PE) vs. Let's Begin (ECE)			Partners for Literacy (ECE/PE) vs. Partners for Literacy (ECE)		
	Effect size	95% CI	<i>p</i> -Value	Effect size	95% CI	<i>p</i> -Value	Effect size	95% CI	<i>p</i> -Value
Interactive reading skill	-0.46	(-0.88,-0.03)	0.030	-0.01	(-0.25,0.22)	>.500	0.44	(0.11,0.77)	0.007
Responsiveness to child	0.11	(-0.12,0.34)	>.500	0.21	(0.04,0.38)	0.011	0.10	(-0.04,0.24)	0.153
Reading skills and vocabulary	0.01	(-0.13,0.14)	>.500	-0.04	(-0.15,0.08)	>.500	-0.04	(-0.14,0.05)	>.500

*Statistically significant at $p < 0.0056$, a value chosen according to Bonferroni's rule to control false positive discovery rates across the nine contrasts run for each outcome.

APPENDIX M

INTERACTIONS OF CURRICULUM WITH HOME LANGUAGE AND ETHNICITY

We explored interactions of curriculum with both ethnicity and home language based on three factors: (1) 46 percent of children in the CLIO sample came from Spanish-speaking homes, and in another 5 percent of homes the primary language was another non-English language; (2) CLIO curricula focused on English-language emergent literacy; and (c) the finding in appendix L that vocabulary growth rates were different for Hispanic children from those of other children.

All of the ITT models included binary covariates for Hispanic background and self-reported home language other than English or a mix of English and Spanish. The additional analyses of this appendix tested for interactions of curriculum with these two covariates. These supplemental analyses were only conducted for the six English-language emergent literacy outcomes. For each outcome, we ran a chi-square test with eight degrees of freedom, testing whether the impacts for any of the curricula varied as a function of either of the language covariates. The results showed that the interactions of curriculum with home language and ethnicity were not significant for any of the six child outcomes. That is, the impacts did not vary significantly as a function of home language or ethnicity.

Methods. We fit models for the six English emergent literacy scales of the form:

$$Y_{ijkt} = \alpha_i + \alpha_{Si} + \alpha_{OLi} + \beta_j + X_{ijk}\gamma + Z_{ij}\delta + \lambda_t + a_{ij} + c_{ijk} + e_{ijkt},$$

where:

- the indices stand respectively for study group (i), stratum (j), child or parent (k), and year (t);
- the terms in Greek letters are fixed effects (α for treatment effect, α_{Si} for differential treatment effect among Hispanic children, α_{OLi} for differential treatment effect among children from families where neither English nor and mixture of English and Spanish is the language usually

spoken at home, β for stratum effect, γ for effects of family and child covariates, δ for effects of measured project covariates, and λ for the effect of year);

- the terms in lower-case Latin letters are random effects (a for project-level random error, c for stable child-level random error, and e for year-specific child- or parent-level random error);
- the terms in upper-case Latin characters are measured variables (Y for outcome, X for child- and/or family-level covariate row vector, and Z for project-level covariate row vector);
- random effects at each level are assumed to be independently and identically normally distributed; and
- random effects at different levels are assumed to be independent of each other.

We tested the hypothesis that among the eight interaction terms, $\{\alpha_{S1}, \alpha_{S2}, \alpha_{S3}, \alpha_{S4}, \alpha_{OL1}, \alpha_{OL2}, \alpha_{OL3}, \alpha_{OL4}\}$, at least one was different from zero against the null hypothesis that all eight interaction terms are zero. We used the standard asymptotic chi-square test supplied by HLM for this purpose with eight degrees of freedom. Table M-1 shows the p -values for the six tests with no adjustment for multiple testing. None were significant.

Table M-1. Results of Overall Test for Any Interaction of Curriculum with Ethnicity or Home Language (Combined Data from Spring 2005 and Spring 2006)

Outcome measure	p -value for overall test across 5 groups
Expressive language: English	0.393
Receptive vocabulary	0.349
Phonological awareness: Elision	>.500
Phonological awareness: Blending	>.500
Print knowledge	>.500
Syntax and grammar	>.500

APPENDIX N

SIMULATION STUDY OF METHOD FOR EXPLORING FIDELITY- ADJUSTED RELATIONSHIPS

As stated in chapter 7, the methodology used for exploring fidelity-adjusted relationships was to fit models of the form

$$Y_{ijkt} = F_{ijt}\theta_i + \beta_j + X_{ijk}\gamma + Z_{ij}\delta + \lambda_t + a_{ij} + c_{ijk} + e_{ijkt},$$

where Y_{ijkt} is the raw-score¹ outcome for child or parent k in year t in the project within stratum j assigned to curriculum i ; F_{ijkt} is the fidelity or pseudo fidelity of the curriculum implementation to which the child was exposed, scaled to lie between 1 for highest ranking fidelity and 0 for lowest ranking fidelity; θ_i is the fidelity-adjusted effect of curriculum i . The terms in Greek letters are other fixed effects (β for stratum effect, γ for effects of family and/or child covariates, δ for effects of measured project covariates, and λ for the effect of year); the terms in lower-case Latin letters are random effects (a for project-level random error, c for stable child- or parent-level random error, and e for year-specific child- or parent-level random error); the terms in upper-case Latin characters are measured variables (Y for outcome, X for child- and/or family-level covariate row vector, and Z for project-level covariate row vector); and it is assumed that all random errors are independent of each other and normally distributed.

The interpretation of the $\hat{\theta}_i$ from this model depends strongly, of course, on how fidelity is defined and measured. As noted in chapter 7 and appendix G, there are particularly difficult issues surrounding adherence to OMLIT/PECAP items that are included in each developer's fidelity rating. The observer-rated values of F_{ijkt} for projects in the four experimental groups do reflect credit for adherence to developer-selected OMLIT/PECAP items in addition to fidelity to the proprietary aspects of each curriculum. As a result, a case can be made for giving credit to control-group projects for their OMLIT/PECAP scores as well, rather than setting $F_{ijkt} = 0$ for control-group projects. If OMLIT/PECAP credit is given within the control group so that nonzero

¹ As discussed in appendix B, we calculated both simple scores and complex IRT scores for all the child outcomes other than the IGDI assessments and the Pre-CTOPPP Print Awareness subtest. The IRT scores had been strongly optimized for the ITT analysis in ways that make them inappropriate for most other usages.

pseudo fidelity scores are defined for some children in the control group, then $\hat{\theta}_i$ should be interpreted as sort of the maximum benefit that could be expected from the curriculum. A slightly more realistic upper bound on the potential benefit of the curriculum is obtained by setting $F_{ijkt} = 0$ for the control group project regardless of OMLIT/PECAP scores. In this case, $\hat{\theta}_i$ should be interpreted as the benefit of the curriculum if it were put in place with *perfect* fidelity in projects that would otherwise have *average* OMLIT/PECAP scores.

Of course, the ITT estimates are the most realistic estimates of the benefit of the curricula. They may be obtained from this model by setting $F_{ijkt} = 1$ for all projects in the experimental groups regardless of either OMLIT/PECAP scores or fidelity to proprietary aspects of the curricula and $F_{ijkt} = 0$ for the control group project regardless of OMLIT/PECAP scores. In this case, $\hat{\theta}_i$ should be interpreted as the benefit of the curriculum if it were put in place with *average* fidelity in projects that would otherwise have *average* OMLIT/PECAP scores. These estimates essentially treat whatever fidelity was observed as the most likely level of fidelity in future implementations.

To verify that the analysis procedure works as intended, we conducted a small simulation study. The simulation study does not reflect all the features of the CLIO study, but we believe that the important aspects are reflected. The simplified model we used to simulate the population was

$$Y_{ijk} = F_{ij}\theta_i + \beta_j + a_{ij} + e_{ijk}.$$

We set the number of strata to be 24, the number of study groups to be 5, the number of projects to be 120, and the number of children per project to be 20. We set $\text{var } a_{ij} = 0.1$, $\text{var } e_{ijk} = 1.0$, $(\beta_1, \dots, \beta_{24}) = \left(\frac{1-12.5}{11.5}, \dots, \frac{24-12.5}{11.5} \right)$, $(\theta_1, \theta_2, \theta_3, \theta_4, \theta_5) = (0.5, 0.5, 0.25, 0.25, 0.25)$. We also drew random values for the fidelity levels rather than fixing them at particular values. We used the multinomial distribution for these draws. We simulated three scenarios. In the first scenario, fidelity was perfect in the experimental groups and null in the control group (group 5). In the second scenario, fidelity in the experimental groups followed a multinomial distribution with mean²

² Fidelity takes the integer values 1 through 5 with the indicated probabilities, which were then transformed to values of 0, 0.25, 0.5, 0.75, and 1.

(0.1,0.2,0.4,0.2,0.1) while fidelity in the control group was null (equal to 1 on the 1 to 5 scale). In the third scenario, fidelity in the experimental groups followed a multinomial distribution with mean (0.1,0.2,0.4,0.2,0.1) while fidelity in the control group followed a multinomial distribution with mean (0.3,0.6,0.1,0,0).

For each scenario, we applied three estimation methods using the SAS procedure MIXED (SAS Institute, 2007). In the ITT method, we fit the model:

$$Y_{ijk} = \begin{cases} \theta_i + \beta_j + a_{ij} + e_{ijk} & \text{for } i \in \{1, 2, 3, 4\} \\ \beta_j + a_{ij} + e_{ijk} & \text{for } i = 5 \end{cases}.$$

For the fidelity-adjusted but not pseudo-fidelity-adjusted method, we fit the model:

$$Y_{ijk} = \begin{cases} F_{ij}\theta_i + \beta_j + a_{ij} + e_{ijk} & \text{for } i \in \{1, 2, 3, 4\} \\ \beta_j + a_{ij} + e_{ijk} & \text{for } i = 5 \end{cases}$$

For the fidelity- and pseudo-fidelity-adjusted method, we fit the model:

$$Y_{ijk} = F_{ij}\theta_i + \beta_j + a_{ij} + e_{ijk}$$

Note that this is the method actually used for the estimates presented in chapter 7.

We drew 100 populations for each of the three fidelity structures and then analyzed each three times, using the three analysis procedures. So there are a total of nine combinations of scenario and analysis method to consider. The results are summarized in table N-1.

In the first scenario of perfect fidelity in the experimental groups and null fidelity in the control group, the three analytic methods produce identical results, all of which are good. Estimated effects are close to perfect (0.5,0.5,0.25,0.25) and t -values are high.

In the second scenario of variable fidelity in the experimental groups and null fidelity in the control group, the three analytic methods produce divergent results. Estimated effects with the ITT method are diluted, and t -values are lower, indicating a reduction in statistical power to find effects. Using either of the corrected methods

reverses the dilution and increases power, although power is still lower than in the first scenario. The method that does not adjust for pseudo fidelity is slightly more powerful than the method that does adjust for it.

In the third scenario of variable fidelity in the experimental groups and variable pseudo fidelity in the control group, the three analytic methods diverge further. Estimated effects with the ITT method are more strongly diluted, and power is further reduced. Only the third method fully reverses the dilution of effects. It is also the most powerful. This is the method used in chapter 7.

Table N-1. Simulation Study Results for Fidelity-Adjustments

Scenario	Study group	ITT analysis		Fidelity-adjusted without pseudo-fidelity adjustment		Fidelity-adjusted with pseudo-fidelity-adjustment	
		Adjusted group mean	<i>t</i> -Value	Adjusted group mean	<i>t</i> -Value	Adjusted group mean	<i>t</i> -Value
Perfect fidelity in experimental groups with null fidelity in control group	1	0.504	5.24	0.504	5.24	0.504	5.24
	2	0.505	5.38	0.505	5.38	0.505	5.38
	3	0.262	2.73	0.262	2.73	0.262	2.73
	4	0.264	2.76	0.264	2.76	0.264	2.76
Variable fidelity in experimental groups with null fidelity in control group	1	0.234	2.33	0.475	3.26	0.466	2.91
	2	0.248	2.46	0.500	3.32	0.489	2.98
	3	0.131	1.36	0.260	1.78	0.250	1.55
	4	0.129	1.31	0.248	1.69	0.239	1.48
Variable fidelity in experimental groups with variable pseudo-fidelity in control group	1	0.184	1.87	0.430	2.99	0.479	3.03
	2	0.197	1.99	0.466	3.21	0.516	3.23
	3	0.084	0.84	0.225	1.50	0.275	1.67
	4	0.082	0.84	0.222	1.52	0.271	1.69

APPENDIX O

PARTICIPATION AND PARTICIPATION-ADJUSTED GROUP DIFFERENCES

In this appendix, we discuss our research into the role of child and parent participation in Even Start. From this broad area, we chose to focus on several specific questions:

- Is parent participation in parenting classes associated with parent responsiveness to their child and parent interactive reading skill?
- Is child participation in preschool education associated with emergent literacy and social competence?
- What is the relationship between curriculum and child outcomes among children with substantial participation in preschool education?

For the analyses of participation, we ignored study group and simply studied the natural relationships between participation and study outcomes with only weak confounder control. For the relationship between curriculum and child outcome among children with high participation in preschool education, we used ITT models of the same sort used in chapters 5 and 6, but the sample sizes were dramatically censored by a targeted study outcome (high participation). With the censored sample sizes, there is less certainty that the initial randomization of study projects results in the randomization of the child samples. Given this reduced certainty, the group differences are not labeled as “effects.”

Parent Participation and Parenting Outcomes

To test the relationships between amount of participation in parenting education and parenting outcomes, we fit models similar to those in the ITT analysis, replacing the term for study group with a measure of participation. Amount of participation was measured as the average number of hours per month that the parent participated in the relevant instructional services over the preceding 7 months.

The results (table O-1) show that a greater amount of participation in parenting education is associated with higher levels of parent responsiveness to

children during joint reading. (Since there are two tests, a p -value of 0.025 is the Bonferroni standard of evidence.) Amount of participation in parenting education was not associated with parent interactive reading skill.

Table O-1. P-Values for Relationships Between Amount of Participation in Parenting Education and Parent Outcomes

Parenting outcome	Parent participation in parenting education
Interactive reading skill	0.267
Responsiveness	0.005*
* Significant at $p < 0.025$.	

To get a feeling for the strength of the relationship, we fit reduced models for the parenting outcomes in which participation was omitted and then graphed the residuals from that model against monthly participation (figures O-1 and O-2). In addition to the points, a nonparametrically smoothed line is shown as well as a horizontal line as a reference for independence. From figure O-1, a positive relationship between participation and parent responsiveness is visible only among parents who attend on the order of 24 hours or more per month of PE/PC instruction, a level attained by only about 10 percent of parents. Among this very small group, there appears to be a fairly steady differential of about 0.3 standard deviations. The comparable figure for parent interactive reading skill (figure O-2) indicates no relationship with participation in parenting education.

Participation in Preschool Education and Child Outcomes

To test the relationships between amount of participation in preschool education and child outcomes, we fit models similar to those in the ITT analysis, replacing the term for study group with a measure of participation. Amount of participation was measured as the average number of hours per month that the child participated in preschool over the preceding 7 months.

Figure O-1. Relationship of Parent Responsiveness (to Child) to Average Hours of Participation in Parenting Education

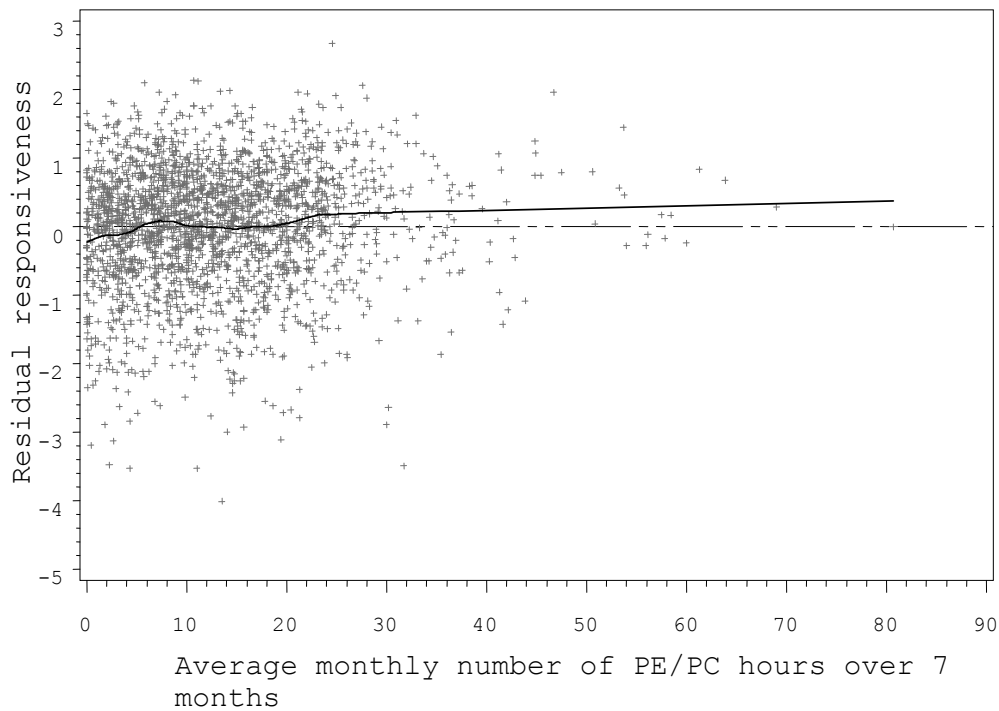
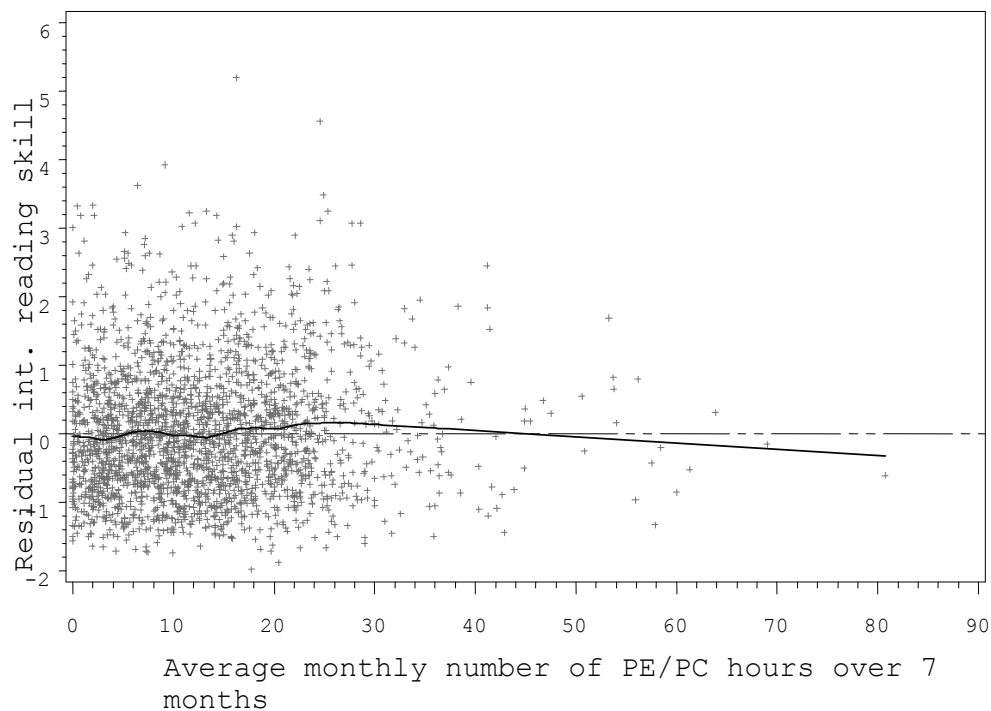


Figure O-2. Relationship of Parent Interactive Reading Skill to Average Hours of Participation in Parenting Education



There was a statistically significant association between hours of participation in preschool and five of six English emergent-literacy outcomes (table O-2). (Since there are eight tests, a p -value of 0.00625 is the Bonferroni standard of evidence.) To help understand the relationship between amount of participation in preschool education and child outcomes, amount of participation in preschool education was plotted against children's scores on a composite emergent literacy outcome measure. The composite child outcome was computed by averaging rescaled scores on the six English-language assessments.¹ The relationship between the composite emergent literacy outcome score and amount of participation in preschool education, for 2005 and 2006 combined, is based on a model run with covariates (the same set of covariates that were used in the ITT analyses).²

Table O-2. P-Values for Relationships Between Participation in Preschool Education and Child Outcomes

Outcomes	Child participation in preschool education
Child outcomes	
Expressive language: IGDI (English)	0.002*
Expressive language: IGDI (Spanish)	0.871
Receptive vocabulary: PPVT	0.002*
Phonological awareness: Elision (CTOPP)	0.000*
Phonological awareness: Blending (CTOPP)	0.256
Letter and sound recognition: Print awareness (CTOPP)	0.000*
Syntax and grammar: TOLD	0.000*
Social competence: Teacher rating	0.029

Notes: Raw assessment scores rather than IRT assessment scores were used for this table.

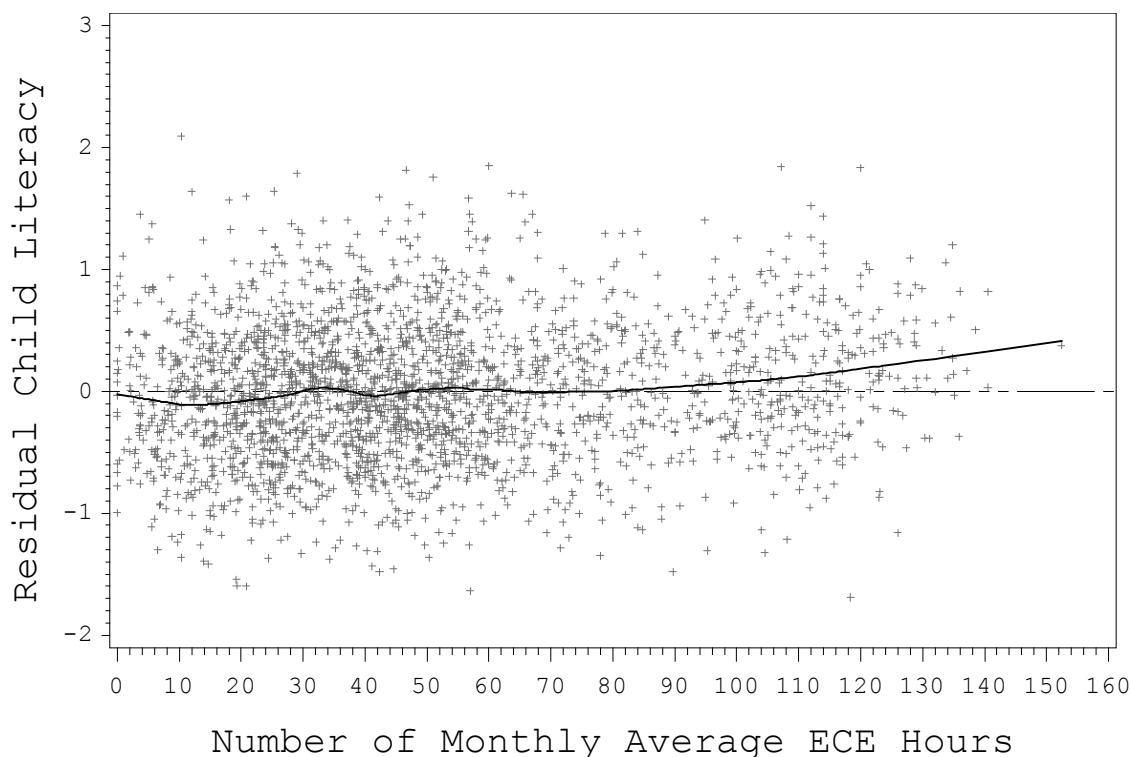
*Statistically significant at $p < 0.00625$.

¹ The composite combined scores have a mean of zero and a standard deviation of 1.

² The relationship of emergent literacy and participation is substantially stronger before covariate adjustments are made. Both participation in preschool education and emergent literacy are related significantly and positively to child age. Participation is also related to year of implementation (participation tends to be higher in the second year) and to household stability (participation tends to be higher for children who have not moved in the 6 months prior to testing).

The relationship between participation in preschool education and the adjusted scores on the emergent literacy composite is not linear (see figure O-3). For children who attended preschool classes for fewer than 85 hours per month over 7 months, there is essentially no relationship between participation and emergent literacy scores. Above 85 hours per month, a level attained by 26 percent of children, the smoothed line gradually pulls away from the horizontal line. For children who attended preschool for the equivalent of a school-day program (6 hours a day, 5 days a week, or around 120 hours per month over 7 months), the differential in emergent literacy is around a third of a standard deviation. This level of participation was obtained by just 3 percent of study children. Thus, although the estimates of association between preschool participation and child outcomes are statistically significant, figure O-3 shows that the relationship appears to be confined to levels of participation reached by only a small proportion of children in the sample.

Figure O-3. Relationship of Emergent Literacy to Average Hours of Participation in Preschool Education (Spring 2005 and Spring 2006 Child Literacy, Average ECE Participation Across September 2004 through March 2005 and September 2005 through March 2006—All Study Groups)



Contrasts in Child Emergent English Literacy Across Study Groups Among Children with Substantial Participation

The finding from the previous section that participation is positively related to emergent literacy replicates results from an earlier study of participation among Even Start participants. In the third national Even Start evaluation (St.Pierre et al. 2003), children who participated more intensively in preschool education scored higher on standardized literacy measures. Other previous research (e.g., Barnett 1995; Ramey and Ramey 1992; Ramey, Bryant, Wasik, Sparling, Fendt, and LaVange 1992) also has shown that children who participated more intensively in early childhood education scored higher on standardized literacy measures.

There was considerable interest throughout the analysis planning process in the interaction of curriculum with child participation. We investigated the association between the CLIO curricula and emergent English literacy among children with substantial participation in preschool education even though we had failed to find any effects of CLIO curricula on emergent English literacy on the total sample. We did this by essentially repeating the ITT analyses of chapter 5 after dropping out children with low levels of participation from all five study groups, including the control group.

There are three caveats to mention with regard to this analysis. The first is the difficulty of determining the threshold for substantial participation. The second is that the CLIO curricula may have had effects on participation that were too small to be detected but that were still large enough to bias the restricted analysis. The third is that the sample size for the restricted analysis was considerably smaller. We expand briefly on each of these caveats and our methodology before presenting the results.

Setting the participation threshold. We used a threshold of 420 hours over 7 months (60 hours per month), a level that corresponds fairly closely to a standard full-week half-day schedule.³ Figure O-3 shows the relationship between participation in preschool education and emergent literacy in 2005 and 2006 after adjusting for a

³ There is no strong research indicating what constitutes a “threshold” level of exposure required for an impact on children of ECE. At the same time, in some of the most well-known research demonstrating positive effects of ECE, such as on the Ypsilanti-Perry Preschool, classrooms provided children with half-day programming. Therefore, for the current study, to test whether impacts are higher if only children with substantial exposure are included in the analysis, a cut-off equivalent to half-day ECE was used.

standard set of covariates. As previously discussed, there is only a visible positive relationship between participation and outcomes among those with at least 85 hours per month. We used the smaller threshold corresponding to 60 hours per month for a couple of reasons. First, creating a graph (not shown) parallel to figure O-3 on spring 2004 data supports a relationship at about 60 hours per month. Second, there was also the question of sample size. As shown in figure O-3, the sample size is very sparse above 85 hours per month.

Selection biases. If there were an effect of the CLIO curricula on participation in preschool education, then that would lead to selection biases when analyzing only those children with substantial participation in preschool education. The implications of this on the restricted analysis are unknown. If some children were pushed over the threshold by the CLIO curricula, then the size and direction of the bias would depend on the backgrounds of those children.

Sample size loss. Subsetting the child samples from spring 2005 and spring 2006 down to those with at least 420 hours of participation resulted in strong sample size reductions. For the print awareness assessment, for example, the number of assessments dropped from 2,788 to 844,⁴ the number of unique children from 2,486 to 780, and the number of projects with at least one assessed child from 120 to 66. These sample size losses obviously reduce power to find significant differences among study groups. Moreover, even once the confidence intervals are adjusted to reflect the smaller sample sizes, the reduction in the number of projects makes the methods (restricted maximum likelihood) used to fit the multi-level models of less certain quality.

Methods. Our initial approach was to run the same multi-level models for the six English-language emergent literacy assessments as described in chapter 4 for the primary ITT analysis of chapter 5 on the restricted dataset of children in 2005 and 2006 with at least 420 hours of participation in a year. (If a child had at least 420 hours from September 2004 through March 2005, then his/her spring 2005 assessment data were used. Similar rules were used for the spring 2006 assessment data. Data were used from both years if the child met the participation level in both years.) However, we had to

⁴ This number was fairly evenly spread across the five study groups: 164, 173, 190, 169, and 148, as one would expect given the lack of a significant CLIO effect on participation, as shown in table 5-1.

drop stratum⁵ from the model because many of the cells in the two-way layout of study group and stratum were now empty.⁶

Results

Tables O-3 through O-5 parallel tables 5-1 through 5-3, reflecting tests on only those children with high participation. The only other difference between the tables of this appendix and of chapter 5 is that results are given only for the six assessments of emergent literacy in English.

There are no statistically significant results after Bonferroni adjustment. That is, there is no association between CLIO curricula and emergent literacy in English among children with high participation levels.⁷

Table O-3. Results of Overall Test for Differences Among Children with High Participation Levels Across the Five Groups (Combined Spring 2005 and Spring 2006) (High participation = 420 hours or more over 7 months)

Outcome measure	<i>p</i> -value for overall test across 5 groups
Expressive language: English	0.244
Receptive vocabulary	0.279
Phonological awareness: Elision	0.301
Phonological awareness: Blending	0.161
Print knowledge	0.107
Syntax and grammar	0.076

⁵ Recall from chapter 4 that the ITT models have 23 dummy variables to remove the variance due to the 24 randomization strata.

⁶ We considered fitting interaction models suggested by Efron and Feldman (1991). This approach would retain all observations and estimate the dose-response relationship between participation and emergent literacy as in figure O-3 separately for each study group and then comparing the curves. We rejected this approach because it would not be easy to apply this approach to CLIO given the deep stratification of the CLIO projects, the complexly clustered structure of the data, and the small number of children with high levels of participation. Moreover, these methods are just as vulnerable to biases caused by intervention effects on participation as the method we used.

⁷ Statistical power to find effects was good as evidenced by the confidence intervals, but selection bias may still be an issue, and a different threshold might have yielded different results.

Table O-4a. Contrasts with the Control Group of CLIO Combined Curricula on Children with High Participation Levels (Research Question 1) (Combined Spring 2005 and Spring 2006)

Child outcomes	Average of the two CLIO combined curricula vs. control		
	Effect size	95% CI	p-Value
Expressive language: English	0.05	(-0.13,0.22)	>.500
Receptive vocabulary	-0.09	(-0.29,0.11)	>.500
Phonological awareness: Elision	0.04	(-0.20,0.28)	>.500
Phonological awareness: Blending	-0.09	(-0.33,0.15)	>.500
Print knowledge	0.01	(-0.22,0.25)	>.500
Syntax and grammar	-0.18	(-0.40,0.05)	0.107

*Statistically significant at $p < 0.0056$, a value chosen according to Bonferroni's rule to control false positive discovery rates across the nine contrasts run for each outcome.

Table O-4b. Contrasts with the Control Group of CLIO Combined Curricula on Children with High Participation Levels (Research Question 1) (Combined Spring 2005 and Spring 2006)

Child Outcomes	Let's Begin and PALS vs. Partners for Literacy (ECE/PE)			Let's Begin and PALS (ECE/PE) vs. control			Partners for Literacy (ECE/PE) vs. control		
	Effect size	95% CI	p-Value	Effect size	95% CI	p-Value	Effect size	95% CI	p-Value
Expressive language: English	-0.10	(-0.30,0.09)	0.288	-0.01	(-0.20,0.19)	>.500	0.10	(-0.10,0.30)	>.500
Receptive vocabulary	-0.09	(-0.32,0.13)	>.500	-0.14	(-0.36,0.08)	0.200	-0.04	(-0.28,0.19)	>.500
Phonological awareness: Elision	0.22	(-0.05,0.49)	0.096	0.15	(-0.12,0.41)	0.272	-0.07	(-0.35,0.21)	>.500
Phonological awareness: Blending	-0.05	(-0.31,0.21)	>.500	-0.12	(-0.39,0.16)	>.500	-0.07	(-0.34,0.21)	>.500
Print knowledge	0.28	(0.00,0.55)	0.037	0.15	(-0.12,0.42)	0.253	-0.13	(-0.40,0.15)	>.500
Syntax and grammar	0.18	(-0.16,0.52)	0.277	-0.09	(-0.37,0.19)	>.500	-0.27	(-0.55,0.02)	0.054

*Statistically significant at $p < 0.0056$, a value chosen according to Bonferroni's rule to control false positive discovery rates across the nine contrasts run for each outcome.

Table O-5a. Contrasts Among Experimental Groups with and without CLIO Parenting Curricula on Children with High Participation Levels (Research Question 2) (Combined Spring 2005 and Spring 2006)

Child outcomes	Average of the two CLIO combined curricula vs. average of the two CLIO preschool curricula		
	Effect size	95% CI	p-Value
Expressive language: English	-0.10	(-0.25,0.04)	0.154
Receptive vocabulary	-0.16	(-0.31,0.00)	0.036
Phonological awareness: Elision	-0.11	(-0.32,0.09)	0.265
Phonological awareness: Blending	-0.21	(-0.39,-0.02)	0.023
Print knowledge	-0.15	(-0.35,0.05)	0.124
Syntax and grammar	-0.15	(-0.35,0.05)	0.119

*Statistically significant at $p < 0.0056$, a value chosen according to Bonferroni's rule to control false positive discovery rates across the nine contrasts run for each outcome.

Table O-5b. Contrasts Among Experimental Groups with and without CLIO Parenting Curricula on Children with High Participation Levels (Research Question 2) (Combined Spring 2005 and Spring 2006)

Child outcomes	[Let's Begin and PALS (ECE/PE) Minus Let's Begin (ECE)] vs. [Partners for Literacy (ECE/PE) Minus Partners for Literacy (ECE)]			Let's Begin and PALS (ECE/PE) vs. Let's Begin (ECE)			Partners for Literacy (ECE/PE) vs. Partners for Literacy (ECE)		
	Effect size	95% CI	p-Value	Effect size	95% CI	p-Value	Effect size	95% CI	p-Value
Expressive language: English	0.01	(-0.26,0.29)	>.500	-0.10	(-0.29,0.10)	>.500	-0.11	(-0.31,0.10)	0.288
Receptive vocabulary	-0.07	(-0.37,0.23)	>.500	-0.19	(-0.40,0.02)	0.060	-0.12	(-0.35,0.10)	0.254
Phonological awareness: Elision	0.32	(-0.06,0.71)	0.082	0.05	(-0.21,0.31)	>.500	-0.28	(-0.58,0.03)	0.062
Phonological awareness: Blending	0.09	(-0.26,0.44)	>.500	-0.16	(-0.40,0.08)	0.178	-0.25	(-0.52,0.02)	0.054
Print knowledge	0.36	(-0.01,0.74)	0.045	0.03	(-0.23,0.29)	>.500	-0.33	(-0.61,-0.05)	0.017
Syntax and grammar	0.43	(-0.06,0.92)	0.071	0.06	(-0.28,0.41)	>.500	-0.37	(-0.65,-0.08)	0.008

*Statistically significant at $p < 0.0056$, a value chosen according to Bonferroni's rule to control false positive discovery rates across the nine contrasts run for each outcome.

