MATHEMATICA
Policy Research, Inc.

Los Angeles Universal Preschool Programs, Children Served, and Children's Progress in the Preschool Year: Final Report of the First 5 LA Universal Preschool Child Outcomes Study

Final Report

June 24, 2009

John M. Love
Sally Atkins-Burnett
Cheri Vogel
Nikki Aikens
Yange Xue
Maricar Mabutas
Barbara Lepidus Carlson
Emily Sama Martin
Nora Paxton
Margaret Caspe
Susan Sprachman
Kathy Sonnenfeld

ACKNOWLEDGMENTS

Carol Shibley—and our word processing experts Donna Dorsey, Linda Heath, and Jane Nelson.

C ONTENTS

Chapter   **Page**

Chapter   **Page**

Chapter   **Page**

T A B L E S

FIGURES

_____

CHAPTER I

INTRODUCTION

A. BACKGROUND AND CONTEXT OF THE STUDY

In 1998, California voters passed Proposition 10, the citizen initiative that levied a tax on tobacco products and earmarked the revenues to advance child and family development, health, education, and safety from pregnancy until children enter kindergarten. Funds flow from the state treasury to the California Children and Families Commission, known as First 5 California, and to the 58 county commissions. The largest of these, First 5 LA, has as its mission "to make significant and measurable progress towards our vision by increasing the number of children from the prenatal stage through age 5 who are physically and emotionally healthy, safe, and ready to learn" (First 5 LA 2008).

The largest of First 5 LA's programs is the Los Angeles Universal Preschool (LAUP) Program. In February 2004, First 5 LA adopted a 10-year Universal Preschool Master Plan to increase the number of preschool slots in Los Angeles County and created LAUP to implement the plan. To meet its goals, First 5 LA committed $580 million over five years to expand and improve existing preschool programs and build new facilities. The goal was to serve 4-year-olds throughout the county through center-based preschools and home-based family child care providers. LAUP operates several types of programs in different target areas. "Early Launch" programs include the approximately 120 licensed preschool centers and 73 family child care home providers funded in the first round of LAUP funding in March 2005.

LAUP then funded "Second Round" programs beginning in January 2006. These programs were funded approximately one year after the Early Launch programs and were targeted specifically to certain geographic areas. They were located in 34 of the most underserved zip codes in the county, which LAUP had identified using analyses conducted for its Universal Preschool Master Plan as Areas of Greatest Service Need (AGSN). The geographic areas were so designated because of their acute need for additional preschool capacity: in 16 of these, the number of 4-year-olds exceeded licensed preschool spaces by at least 1,000. In the other 18 zip codes, capacity was at least 500 spaces short of the number of 4-year-olds and, in addition, the majority of the elementary schools in the area were designated by the state as "low performing." Furthermore, almost one-third of all 4-year-olds in the county (about 50,000) live in these 34 zip codes.

First 5 LA also received funding from the California First 5 Commission to be one of nine counties to operate a special demonstration program, the Power of Preschool (PoP) Demonstration Project. PoP programs are designed to "demonstrate the impact of voluntary, quality preschool on children's school readiness in counties throughout California" (LAUP 2008a). The PoP grant did not specify any particular preschool curriculum. Beginning in fall 2006, LAUP implemented these programs in 13 cities within the AGSN zip codes.[1] As of December 2008, 240 LAUP programs in 61 zip codes received some PoP funding.

As of 2006, LAUP had expended or committed funds to support 10,217 high-quality preschool spaces for 4-year-olds across Los Angeles County, with the goal to "have 70 percent of all 4-year-olds and their families in LA County actively participating in a LAUP program" by 2014 (LAUP 2006). As of June 2008, LAUP was funding 201 preschool centers serving 6,210 4-year-olds (with another 1,397 children served by 123 family child care home providers) (LAUP 2008a). The center providers included 129 Early Launch providers and 72 Second Round centers. LAUP began Round 2.5 in July 2007; its programs were not included in this study.

LAUP supports a diverse delivery system, funding preschools in public (both traditional and charter) and private schools, Head Start centers, and both small and large family child care homes. LAUP preschool centers employ a variety of approaches and curricula, including Montessori, Reggio Emilia, High/Scope, Creative Curriculum, and many others. The common characteristic of all LAUP preschools is that they must meet LAUP's quality criteria. Before LAUP grants any preschool a contract, it examines teachers' credentials, class sizes and child-adult ratios, the learning and development environment of the classroom, curriculum, staff stability, working conditions, regulatory compliance, and licensing status. LAUP hires external reviewers to rate preschools using a 5-Star Quality Rating Scale; only those receiving a Star rating of 3 or higher qualify for full funding from LAUP.[2] If a program is not yet ready to be a 3-Star program, LAUP provides individual coaching to each site through its Starting Points Coaching program. Once programs are operational, LAUP provides support through coaches who work with teachers to review and enhance their classrooms' quality. LAUP's coaching model partners provider staff with trained coaches who "model appropriate interactions, connect providers with community and early

---

[1] These PoP demonstration programs initially were referred to as Preschool for All (PFA) programs; some counties still use the PFA designation. The overarching goal of PoP programs is to increase the number of high-quality spaces for preschool-age children in California while gathering information that may guide the development of a statewide preschool system. PoP programs vary by county and are not required to follow a particular curriculum or format as a condition of their funding. LAUP programs receiving PoP funds were not intended to differ in any systematic way (such as curricula used or auspice) from other LAUP programs. PoP programs follow quality criteria that closely relate to LAUP's Star Rating System. See First 5 California's website for further details (www.ccfc.ca.gov).

[2] To read more about the rating scale, go to http://www.laup.net/index.php?page =for_families/quality_standards. LAUP uses the Early Childhood Environment Rating Scale-Revised to assess centers' learning and development environments.

childhood resources, and customize their coaching style to focus on the individual needs and goals of the providers" (LAUP 2008a).

Los Angeles County residents are diverse in language, race, ethnicity, and country of origin. Among the 10 million residents of the county, an estimated 224 languages are spoken. Approximately 156,000 4-year-old children live in Los Angeles County. According to LAUP (2006), the 4-year-olds are almost two-thirds Hispanic or Latino (61 percent), 18 percent white, 9 percent African American, 8 percent Asian, and 4 percent other races/ethnicities. About 44 percent of public school kindergarteners were considered English Language Learners (ELLs) in 2004-2005. Spanish is the most predominant non-English language, but substantial numbers of ELLs speak Armenian, Korean, and Cantonese. Approximately 60 percent of all public school students (K-12) receive free or reduced-price meals. Nearly half of Los Angeles County's children live in low-income families, and Latino children account for nearly three-quarters of the children in poverty (Los Angeles County Children's Planning Council 2004). Although low-income families tend to be clustered in the south and southeast portions of the county, they constitute important proportions of many neighborhoods where LAUP preschool programs are located. As noted previously, LAUP factored these unique demographics into its plan for targeting delivery of preschool services.

B. DESCRIPTION OF THE UNIVERSAL PRESCHOOL CHILD OUTCOMES STUDY

In February 2007, First 5 LA contracted with Mathematica Policy Research, Inc. (MPR) and its subcontractors, Juárez and Associates, American Institutes for Research (AIR), and Berkeley Evaluation and Assessment Research (BEAR) Center, to conduct the First 5 LA/LAUP Universal Preschool Child Outcomes Study (UPCOS).[3] We conducted the study in two phases. The spring 2007 pilot phase (Phase 1) examined the feasibility, reliability, and validity of selected child development measures in the large, culturally and linguistically diverse population of children served by LAUP programs.

In early spring 2007, we selected a purposive sample of 418 4-year-olds enrolled in 14 LAUP programs. The children were predominantly Latino (60 percent), and 28 percent spoke primarily Spanish at home. Data collection included direct child assessments and a teacher interview encompassing ratings of children's behavior and development. In addition, through focus groups and cognitive interviews with parents, we developed a better understanding of how parents interpreted child behavior rating items. Our psychometric analyses of the direct child assessments and teacher rating scales, along with results of the parent focus groups, led to firm conclusions about the psychometric qualities of measures to be used in Phase 2 (Vogel et al. 2008). We then recommended a set of measures for Phase 2 that significantly benefited from the pilot phase experience (see Atkins-Burnett et al. 2007 for specific recommendations for the measures used in Phase 2).

The second phase of the study (fall 2007 and spring 2008) addressed the quality and overall implementation of LAUP programs, children's growth from fall to spring, and the

---

[3] BEAR's participation in the study concluded after Phase 1 was finished.

relationships between family characteristics and children's development over time. In addition to the information it will provide to First 5 LA and LAUP, the study was designed to contribute to the larger early childhood program evaluation field through increased understanding of issues in assessing young children from diverse linguistic and cultural backgrounds.

C. FOCUS AND RESEARCH QUESTIONS OF THIS REPORT

The focus of this report is the study's second phase (fall 2007 and spring 2008). Phase 2 examined the quality, intensity, and overall implementation of LAUP programs (including classroom quality and teaching activities); documented the characteristics of the representative sample of teachers and the children and families enrolled in the programs; and measured children's behavior and development across the full range of domains related to school readiness. We analyzed children's fall-spring changes and examined the relationships between child and family characteristics and children's school readiness outcomes. Because the sample was selected to be representative of all LAUP center-based programs, we can generalize the results to all LAUP center-based programs, classrooms, and children. We include a separate report on the PoP programs in Appendix E.

After describing the characteristics of the representative sample of children and families, we report our findings related to the three broad questions this study addresses, which are described in more detail in Chapter II:

1. What is the overall level and range of quality in the implementation of LAUP/PoP center-based programs?

2. How do children enrolled in LAUP/PoP center-based programs develop from fall to spring?

3. How are characteristics of children and families related to school readiness outcomes?[4]

D. ROADMAP TO THE REPORT

In Chapter II, we describe the study methods used, including how we sampled programs and families; what child, parent, teacher, and classroom measures we used; what data collection procedures we followed; and what research questions the study was designed to address. In Chapter III, we present the results, beginning with descriptive data on programs, teachers, and classrooms, followed by descriptions of children and families. We then report what we learned about parents' involvement and satisfaction with the program. We conclude the results chapter with findings about the relationships between child and family characteristics and the children's school readiness outcomes. Chapter IV summarizes

---

[4] See discussion in Chapter II for information about research questions that we considered but did not address in this study.

the findings and presents our assessment of their implications for programs and recommendations for future study. In the appendices, we provide details on the measures used, the routing procedure used for determining the language of assessment, what we learned about a new experimental observational measure of teacher-child interactions, and our report on the PoP programs.

Chapter II

UPCOS Study Methods

A. PHASE 2 SAMPLING DESIGN AND SAMPLE, RESPONSE RATES, AND ATTRITION

1. Sampling Programs and Children/Families

First 5 LA's goal was to conduct Phase 2 of UPCOS with a sample of center-based programs and children representative of all those served by LAUP. After considering the sample of children and programs that would be necessary to provide the appropriate degree of statistical power for the analyses we planned, we set a target of obtaining a sample of 96 programs with about 1,920 children (or 20 children per program). To allow for possible refusals, we selected a sample of 120 programs that were listed as fully operational in May 2007, with 2,400 children.[1] We obtained data from LAUP on all the funded programs from which to draw the sample. We then employed a stratified two-stage sampling design in which, in the first stage, we selected a stratified random sample of programs and, in the second stage, selected a random sample of children within each sampled program.

To ensure a representative sample, we accounted for such factors as types of programs and geographic areas. The first factor was whether a program was an Early Launch or Second Round program. As described in Chapter I, Early Launch programs included the approximately 120 licensed preschool programs funded in the first round of LAUP funding in March 2005. The Second Round programs were those located in 34 of the most underserved zip codes in Los Angeles County, which LAUP had identified as AGSNs. In addition, LAUP selected a subset of the cities that include the AGSNs to participate in a special demonstration program funded by the state First 5 Commission, the PoP Demonstration Project. In sampling Stage 1, we stratified programs into five explicit strata based on a cross-classification of three variables: (1) whether a program was in an AGSN, (2) whether an AGSN program was Early Launch or Second Round, and (3) whether an AGSN program was a PoP program or not.[2]

---

[1] In this document, we use the term "program" interchangeably with "center" or "school."

[2] Second Round programs were all in AGSN zip codes. All PoP programs were located in AGSNs.

We then selected 120 programs across the five strata (Table II.1). To increase the power of analyses by domains of interest, we oversampled PoP programs; that is, we selected all PoP programs but only some of the non–PoP programs. We selected programs in the non–PoP stratum by using a without-replacement sequential sampling with probability proportional to size (PPS) methodology developed by Chromy (1979), in which the measure of size is the number of children in a program. In other words, larger programs had a higher probability of selection. Accordingly, any very large programs were selected with certainty. We then removed these large programs from the sampling strata before selecting the non-certainty programs.

Table II.1.   UPCOS Program-Level Sample

Stratum Number	Stratum Description	Number of Programs in Frame	Number of Programs Sampled	Number of Certainty Selections	Number of Non-Certainty Selections
1	Non–AGSN, Early Launch	80	57	24	33
2	AGSN, Early Launch, PoP	10	10	10	0
3	AGSN, Early Launch, Non–PoP	29	20	7	13
4	AGSN, Second Round, PoP	16	16	16	0
5	AGSN, Second Round, Non–PoP	24	17	7	10
Total	**All Strata**	**159**	**120**	**64**	**56**

Before beginning the selection process, we sorted programs in the three non–PoP strata according to the percentage of enrolled children who spoke a language other than English or Spanish so that the sampled programs would be more representative in terms of this language variable. Among these 120 selected programs, we determined that 1 was ineligible (because it was no longer operating). Among the 119 eligible programs, 98 agreed to participate—thus yielding 2 more than our target of 96.

To select the children within each of the programs, we followed four principles to guide our sampling design: (1) be inclusive and not exclude any children on the basis of language or disability; (2) include children born between December 3, 2002, and December 2, 2003, to maximize the probability that our sample would include only children who would enter kindergarten the following year (in fall 2008); (3) explicitly stratify based on the primary language spoken at home (Spanish or not Spanish) but without oversampling; and (4) sort on class sessions (if a classroom has more than one) and on gender within class sessions. At the second stage, we selected children in sampled programs by using a without-replacement sequential sampling with equal probability of selection methodology, developed by Chromy:

For the non-certainty programs, we selected 24 children within each program. (If the total number of children within the program was 24 or fewer, we selected all of them.)

For the programs selected with certainty, we selected children with a sample size proportional to the total number of children within the program, with a maximum sample size of 30 per program.

Out of the 3,376 children in the 98 eligible and participating programs, we selected 1,954. Among these, we deemed 68 children to be ineligible because either their birthdays fell outside the eligible range or they left the classroom after sample selection but before the scheduled data collection. Among the 1,886 eligible children, we obtained parental consent for including 1,724 eligible children in the fall UPCOS sample. Children were deemed eligible for the spring 2008 data collection if they remained in a program that was sampled for the UPCOS study even if they changed classrooms.

With enrollment incomplete in many programs at the beginning of program services in the fall and in order to accommodate families that enrolled after that date, we did not select samples of children until the scheduled period of data collection. Two weeks before data collection was to begin at a program, we obtained the program's child rosters and randomly selected the sample of children within the program.

2. Program and Family Recruitment Process

To maximize programs' cooperation, MPR worked closely with LAUP senior managers and coaches to introduce the study to programs and encourage participation. Our first contact with the programs was a letter from LAUP that explained the study and invited participation. MPR followed up with a letter and a Frequently Asked Questions (FAQ) summary that explained the purpose of the study, introduced our procedures, and described the confidentiality protections we would follow.

Within a week of mailing the MPR letter, our study coordinators contacted program directors by telephone. The 13 coordinators who remained in contact with the programs throughout the fall data collection period worked either directly for MPR or for our subcontractors, AIR and Juárez and Associates. In addition to introducing the study to the program director and arranging for the assignment of a center liaison to work with our staff, the coordinators completed a data collection plan for each program that included:

1. Obtaining information on the number of classrooms in the program, the number of children in each class, and the languages of the families

2. Developing plans for receiving a roster of eligible children and selecting the sample

3. Making arrangements for the distribution of consent forms

4. Making arrangements for conducting child assessments and scheduling parent interviews

As discussed above, of the 120 sampled programs, 98 (82 percent) were eligible and agreed to participate. As we prepared for data collection, we found that 1 program was no longer operational and that 21 did not agree to participate. Each participating program received $250 to help defray the cost of its center liaison's efforts.

Many programs distributed parental consent forms to parents two to three weeks before the data collection week. Some programs elected to distribute the forms earlier (for example, during the summer in coordination with the distribution of other enrollment forms). If the consent rate was less than 75 percent within a week of the scheduled start of data collection, we offered to help the programs contact parents and collect forms. If consent rates were still low three days before the start of data collection, we increased our efforts to contact parents and encourage participation. Approximately one-third of all programs required such assistance. In particular, key field staff conducted information meetings with parents and/or program staff or personally distributed and collected consent forms.

3. Response Rates and Attrition

We obtained informed consent from 91.4 percent of the parents of eligible children. The final sample of eligible children whose parents consented to participate totaled 1,724. Of these, 171 children were not available for the spring 2008 data collection. Most such children had left their programs, although some were in one program whose director decided not to participate in the spring. Thus, the spring sample consisted of 97 centers and 1,555 children. In Table II.2, we summarize the sample size and retention rates for each wave of data collection.

Table II.2.   UPCOS Sample and Retention Rates

Original Fall 2007 Sample in 98 Programs	Eligible Children	Parental Consent Obtained	Fall 2007 Assessment Data Obtained	Spring 2008 Assessment Data Obtained
1,954	1,886	1,724	1,657	1,555
	100%	91.4%	96.0%	90.1%

B. MEASURES USED WITH CHILDREN, PARENTS, TEACHERS, AND CLASSROOMS

Based on a careful review of available measures (reported in the project's Technical Memorandum No. 1, Atkins-Burnett et al. 2007), and in part on the psychometric and other analyses we conducted with the Phase 1 pilot data in spring 2007 (as reported in the project's Technical Memorandum No. 2, Vogel et al. 2008), and in consultation with First 5 LA, LAUP, and First 5 LA's Research Advisory Committee (RAC),[3] we identified the measures to be used in Phase 2 for collecting data from children, parents, teachers, and center classrooms.

---

[3] The First 5 LA Universal Preschool Research Advisory Committee (RAC) is made up of research experts representing varied depth and breadth of local and national early childhood research and related expertise who provide input and advice to First 5 LA on matters relating to research and evaluation.

1.   Child Assessment Measures

Several sources informed our recommendations for child measures, including the work of the National Education Goals Panel (Kagan et al. 1995) and subsequent frameworks (Love et al. 1994; Love 2003), recommendations of the First 5 LA RAC, and discussions with the staff at First 5 LA and LAUP. Our review and recommendations also took into consideration measures currently in use in the 2006 wave of the Head Start Family and Child Experiences Survey (West et al. 2007) and those developed for the preschool wave of the Early Childhood Longitudinal Study-Birth Cohort (ECLS-B), the Early Childhood Longitudinal Study-Kindergarten Cohort (ECLS-K), the Preschool Curriculum Evaluation Research (PCER) study, and other National Center for Education Statistics school surveys.

An important consideration for the selection of child development measures was ensuring coverage of all the important domains of development central to school readiness. Thus, we selected measures that fit within the five domains identified by the National Education Goals Panel framework for children's early development and learning (Kagan et al. 1995):

1. Cognition and general knowledge (for example, literacy, mathematics, problem solving)

2. Language development

3. Approaches to learning

4. Social and emotional development

5. Physical well-being and motor development

In addition, as described in the fall 2007 memorandum from MPR to First 5 LA (Atkins-Burnett et al. 2007), we selected child assessment measures with the following features:

1. Collectively covered all domains of children's early development and learning that comprise school readiness

2. Demonstrated evidence of use with diverse populations (including linguistic, racial/ethnic, and ability diversity)

3. Demonstrated evidence of reliability and validity

4. Balanced the "tried and true" with the new

5. Demonstrated sensitivity to the effects of interventions

6. Has been shown to be age-appropriate for 4-year-olds

7. Is currently available (as much as possible) in both Spanish and English

II: UPCOS Study Methods

With these considerations, and grounded in the analysis of the Phase 1 pilot data, we selected the measures in Table II.3 for use in Phase 2. Our measurement plan followed a

Table II.3. Child Assessment Measures

Measure[a]	Purpose and Domain Assessed
Direct Child Assessments	
Spanish Pre-LAS: Tío Simón Dice	Language screener: Spanish receptive proficiency
Spanish Pre-LAS: Exposición de Arte	Language screener: Spanish expressive proficiency
English Pre-LAS: Simon Says	Language screener: English receptive proficiency
English Pre-LAS: Art Show	Language screener: English expressive proficiency
Receptive One-Word Picture Vocabulary Test-English Edition (ROWPVT) and Spanish Bilingual Edition (ROWPVT-SBE)[b,c]	Language development: Receptive vocabulary
Expressive One-Word Picture Vocabulary Test-English Edition (EOWPVT) and Spanish Bilingual Edition (EOWPVT-SBE) [b]	Language development: Expressive vocabulary
ECLS-B Math [b]	Cognitive development: Mathematics concepts
Height and Weight	Physical development; health status
Executive Functioning 1. Walk-a-Line Slowly Task 2. Pencil Tapping Task	Executive functioning and self-control
Rapid Letter Naming [b]	Literacy development: Letter identification
Woodcock-Johnson III-Test 7 Spelling (WJ-III) and Woodcock-Muñoz III-Test 7 Ortografía (WM-III)	Literacy development and fine motor skills: Early writing and spelling/ability to take dictation/fine motor control
Environmental Print Awareness-spring only	Literacy development: Ability to identify words in the environment
Leiter-R Examiner Rating Scales both fall and spring	Assessor rating of children's attention, activity level, and sociability
Parent Ratings	
Health Ratings	LA County Health Survey items
Preschool Kindergarten Behavior Scales-2 (PKBS-2)-adapted	Rating positive social skills
Teacher Ratings	
Preschool Kindergarten Behavior Scales-2 (PKBS-2)	Rating positive social skills
Social Skills Rating System (SSRS)	Rating problem behaviors
Desired Results Developmental Profile-Revised (DRDP-R)	Comprehensive ratings of child development
ECLS-K Approaches to Learning	Rating positive approaches to learning

[a]Measures are listed in order of administration. [b]Denotes conceptual scoring. [c]The ROWPVT-SBE was administered only in the fall.

multimethod approach. Teachers, parents, the assessors, and the children themselves provided data about each child's development. The table spells out the general purpose of each measure, while Appendix A describes each measure in greater detail.[4]

2. Classroom Observation Measures

In addition, we identified instruments for observing LAUP classrooms. First 5 LA articulated three goals for the classroom observations:

1. Obtain information on quality in the programs

2. Obtain information on teachers' instructional practices

3. Learn about the experiences of English Language Learners (ELLs), including how teachers interact with them and what languages they use, for what activities, and how often

After reviewing the literature on classroom observation measures that would be suitable for accomplishing these goals, we selected one measure and developed another. We selected the Classroom Assessment Scoring System (CLASS) based on several considerations.[5] The CLASS, which examines classroom interactions, is a measure of classroom climate and teacher instructional practice. It does not depend on the presence of any particular materials or curriculum. The "focus is on what teachers do with the materials they have and on the interactions they have with students" (Pianta, La Paro, and Hamre 2006). The preschool CLASS assesses three domains: (1) Emotional Support for Children (including Class Climate, teacher sensitivity, and regard for student perspectives); (2) Classroom Organization (including Learning Formats, Time Use and Productivity, and Behavior Management); and (3) Instructional Support for Learning (including the level of concept development, the quality of feedback to children, and the modeling of language). Development of the CLASS was based on extensive literature reviews and on data from the classroom observations instudies conducted by the National Institute of Child Health and Human Development (NICHD) (Pianta et al. 2002) and the National Center for Early Development and Learning's (NCEDL) Multi-State Pre-K Study (Early et al. 2005). The CLASS is used widely and has shown relationships to important child outcomes such as performance on tests of literacy and increased child engagement in learning (Howes et al. 2005; Pianta 2003; Pianta,

---

[4] One additional measure was pilot tested and used on an experimental basis in Phase 2. This was the Respect for Differences Scale. Because of the experimental nature of this parent rating scale, we do not report findings from it in this report. However, in Appendix F we present a brief report on the development and psychometric characteristics of the instrument and illustrate what it is capable of revealing about children's respect for other cultures.

[5] We also considered using the Early Childhood Environment Rating Scale-Revised (ECERS-R; Harms, Clifford, and Cryer 1998). However, because LAUP uses the ECERS-R as part of the determination of programs' Star rating, we decided that an additional measure would provide more information about LAUP classes participating in UPCOS. As described in Chapter III, we also conducted analyses using the ECERS-R scores and Star ratings provided by LAUP.

La Paro, and Hamre 2006). Concurrent validity has been demonstrated with the CLASS and the Early Childhood Environment Rating Scale-Revised (ECERS-R; Harms, Clifford, and Cryer 1998), the Snapshot (Ritchie et al. 2002), and the percentage of time children are not engaged (Pianta, La Paro, and Hamre 2006).

In consideration of First 5 LA's goal to examine the language support for ELL children, particularly in the absence of an existing scale that would capture the extent to which English or other languages are used in conversations between adults and children in the classroom, we revised a section of the Child-Caregiver Observation System (C-COS) (Boller, Sprachman, and the Early Head Start Research Consortium 1998) developed by MPR for use in the Early Head Start Research and Evaluation Project. MPR and Bank Street College had adapted the C-COS for use in studies of family, friends, and neighbor care. We made additional adaptations for UPCOS and named the current adaptation the Language Interaction Snapshot (LISn). We specifically designed it to capture information about the language(s) teachers used with children in the classroom, the types of language used in different settings, and the adults who interacted with ELL children. The LISn is described in more detail in Appendix D.

C. DATA COLLECTION APPROACH

1. Data Collection Components: Fall 2007, Winter 2008, and Spring 2008

We collected UPCOS data in three waves: (1) fall 2007 baseline data collected from parents, children, and teachers; (2) late winter/early spring 2008 classroom observations and teacher interviews; and (3) follow-up data from these groups in late spring 2008. Data collection in the fall and spring for parents and children took place between September 10 and November 9, 2007, and between April 14 and June 13, 2008. The mean interval separating the fall and spring assessments was 30.8 weeks (SD = 3.26) or 7.1 months (SD = .75). We allowed teachers an additional month to complete their observations and ratings.

Interviews with Parents. The parent interviews took 30 to 40 minutes and were completed as a computer-assisted personal interview (CAPI). In fall 2007, the center liaisons and teachers scheduled most of the parent interviews as in-person interviews, with approximately 15 percent conducted by telephone. In spring 2008, we changed our procedure, offering parents a choice of mode; 28 percent of parent interviews were completed by telephone. Both English and Spanish versions of the questionnaire were available. In fall 2007, 58.9 percent of parent interviews were conducted in English, 41.0 percent in Spanish, and fewer than 1 percent (15 interviews) in Korean. In spring 2008, 57.2 percent of parent interviews were conducted in English, 42.3 percent in Spanish, and again fewer than 1 percent (7 interviews) in Korean. Parents received a $25 gift card for completing the interview.

Direct Assessments of Children. Project-trained assessors conducted one-on-one assessments with the sample children during the day at times agreed upon with the program administration; in addition to allowing assessments during class time, some centers permitted assessment during rest times, outdoor play times, or extended-day sessions. The text of the items and instructions for assessors were programmed into laptop computers, and assessors

entered scores directly into the computers, permitting the automatic calculation of any measurement basal and ceiling scores. Children viewed the test items on paper easels. Assessors started the testing with procedures designed to minimize stress by giving the child some stickers and ensuring that the child was comfortable before starting the assessments. In addition, the assessment battery was designed with particular attention to making Spanish-speaking children comfortable. For example, for Spanish-speaking children, the assessor conducted the two subscales of the Pre-LAS screening in Spanish before giving the English version so that children could gain confidence in their abilities and the assessor could build rapport with them.[6] Although the length of the assessment varied with the skill and attentiveness of the child, the median assessment time was 45 minutes in the fall and 40 minutes in the spring. We distributed books to all children in the class as a thank you, whether or not they were in the study sample.

MPR, in consultation with First 5 LA staff and consultant Michael López, developed a four-stage procedure to ensure that children were routed into an assessment-administration path that would present them with instructions in their dominant language. (In Appendix B, we present details on the routing process.) Child assessments were completed only in English or Spanish; with two exceptions, conceptual scoring enabled assessors to ask questions and accept and score responses in either language throughout the assessment battery.[7] The Woodcock-Johnson Test 7 (Spelling/writing) was administered only in English and the Woodcock-Muñoz III Test 7 (Ortografía) only in Spanish. In addition, because California public elementary schools permit instruction only in English, we decided to focus the spring assessment of receptive vocabulary on English proficiency; thus, we administered the spring ROWPVT only in English.

Teacher-Completed Measures. We asked teachers to complete a self-administered teacher child report (TCR) for each child in their class. All teachers received an advance letter explaining the purpose of the TCRs. The letter also provided them with a unique password and ID number so they could access the web version of the instrument if they preferred to complete it online. We also offered teachers the option of using a paper version. The majority of teachers chose the web option, with 69 percent of TCRs completed on the web in the fall and 68 percent in the spring. The TCR on paper was available in an English/Spanish bilingual version, but the web version was available only in English. Each

---

[6] Items on the English version were re-ordered so that they did not match the order of the Spanish version, thereby reducing an order effect that would inflate scores on the English Pre-LAS.

[7] Debate surrounds the use of conceptual scoring. Some find it unacceptable because it allows one of two responses (Spanish or English word) to be scored as correct and prevents comparisons of scores of the assessments with those of the publisher's national norms. However, Spanish-speaking children are typically compared with standard scores for a representative English-speaking sample. Not allowing the conceptual scoring and administering in only one language would downwardly bias the estimates of overall ability (except in the case of measuring language proficiency separately for each language) because the child may know the information or skill in one language and not know the word for it in the other language. We chose to use conceptual scoring in UPCOS because of the belief that the resulting estimates of children's skills and abilities will be more valid than would be the case if children's language proficiency in one language or another were allowed to add error to the scoring.

TCR took about 10 minutes to complete. Where permitted by program policy, MPR paid teachers $10 for each TCR completed.

The TCRs included standard rating scales (Table II.3) that requested frequency ratings focused primarily on aspects of children's social-emotional development and approaches to learning. In addition, we conducted special training for teachers in the PoP classrooms on the Desired Results Developmental Profile-Revised (DRDP-R). Teachers observed children's behavior during specified interactions with materials, other children, or the teacher over several weeks. Teachers then used the DRDP-R scoring form to provide their assessments of children's language, literacy, cognitive, and mathematics development, as well as of their social-emotional development, approaches to learning, and motor development.

We also asked teachers to complete a brief demographic survey about themselves. Called the Brief Teacher Questionnaire (BTQ), the survey was available in both English and Spanish. The teacher who completed the TCR and/or the lead teacher completed the BTQ only once.

Classroom Observations and Teacher Interviews. The classroom observations and teacher interviews took place primarily between March 17 and April 11, 2008, with a small number of classroom observations (N = 3) and teacher interviews (N = 36) completed between mid-April and the first week of May. Except for afternoon-only classes, the classroom observations took place over the course of a full morning session. Observers completed the 30-minute teacher interview with each lead teacher in the class either immediately after the observation or in a follow-up telephone call.

2. Field Staff Training

All field staff attended rigorous training sessions to learn the instruments they would administer and the procedures for conducting their components of the data collection. At the end of each training, field staff completed rigorous certifications. In addition, to ensure continuing adherence to the assessment protocols, we conducted quality control observations during the field periods.

Child Assessments and Parent Interviews. In late August, the field staff assigned to conduct the fall child assessments and parent interviews participated in five and one-half days of training in Los Angeles. The training consisted of four distinct components: (1) working with programs, parents, and children (with and without special needs); (2) using computer systems for sample management and data collection; (3) conducting parent interviews; and (4) carrying out child assessments. To help staff develop the understanding and competencies needed to collect high-quality data, each training session included in-depth discussions with the trainers. We used explanatory videos and provided practice sessions with each data collection instrument (including role play and mock interviewing). Data collection staff also underwent training in the administrative aspects of the project, record keeping, and the collection of materials from teachers.

On the final days of training, we conducted special sessions to certify field staff in administering the child assessments and parent interviews. We invited parents (whom we

ensured were not part of our sample) from nearby communities to bring their 4-year-olds to the training site to enable trainees to administer the child assessments to children similar to those in the UPCOS sample. Quality control monitors observed the child assessments and followed a structured protocol to record notes when data collectors (1) strayed from the script (miswording or paraphrasing), (2) made gesturing errors, (3) provided non-neutral encouragement, or (4) coached the child. The median score for all trainees during the initial certification session was 6 errors, or an average error rate of less than 2 percent. One assessor-trainee made more than 40 errors across the entire assessment; that trainee was not certified and was excused from the study. We trained 53 assessors/interviewers in the fall 2007 training and certified 52 on the child assessments and parent interviews. In early April, the experienced field staff attended a one-day refresher training session that highlighted changes to the parent interview and child assessments and included recertification. All staff were recertified for the spring data collection. Staff from MPR and Juárez and Associates conducted quality control visits during each field period to ensure adherence throughout the data collection period to the standards in place at the end of training. During the fall 2007 data collection, lead trainers conducted one quality control visit during the first week of data collection and a second visit in the middle of the field period. In spring 2008, field supervisors conducted the first quality control visit, and a lead trainer conducted the second.

Classroom Observations. In March 2008, we invited 14 of the most qualified field staff from the team conducting the parent and child assessments and interviews to conduct the classroom observations. The two instruments used to examine classroom teaching and instruction (the CLASS and LISn) require extensive training. The CLASS observation training took place over three and one-half days; the LISn training required one and one-half days. Both training sessions included discussion of how to think about the structure and organization of a classroom and how to use the measurement tools. To help observers understand, trainers made extensive use of videos combined with observer classroom practice.

All observers received certification at the culmination of both the CLASS and LISn trainings, following procedures recommended by the CLASS developer and thereby ensuring that only reliable observers would conduct the study observations. At the end of the second day of CLASS training, observers took a reliability test in which they watched and coded five 20-minute classroom segments. To assess inter-rater reliability of the CLASS items, we followed the developers' protocol and considered ratings falling within 1 point of the master coding scheme to constitute "agreement." Across observers, rater agreement (as calculated for the best four videotapes) ranged from 78 to 93 percent. To further ensure consistency across CLASS ratings, we also established reliability in the field. On the final day of training, CLASS observers visited classrooms (in the study sample) in groups of three accompanied by a "gold-standard" and reliable MPR trainer.[8] As with the videos, we based reliability on a calculation of the percentage of items on which the trainee's rating fell within 1 point of the gold-standard coder's rating. Agreement ranged from 38 to 95 percent. For those observers

---

[8] Data collected by the gold-standard observer during field reliability were used as study data to avoid additional burden on the LAUP centers.

who fell below 80 percent reliability or those whom the trainers wanted to confirm as reliable, a second field session (Field Reliability 2) took place. In Table II.4, we present the inter-rater percentage agreement scores for the CLASS for all observers. Information about the rater training and reliability on the LISn is presented in Appendix D.

Table II.4. **Inter-Rater Agreement (Within 1 Rating Point) on the CLASS for Training Reliability Tapes and Field Reliability Observations**

Observer	Training Reliability: Percentage Agreement Within 1 Point in Best 4 Tapes	Field Reliability 1: Percentage Agreement Within 1 Point	Field Reliability 2: Percentage Agreement Within 1 Point[9]
1	90	88	68
2	93	85	NA
3	78	95	90
4	90	85	NA
5	85	78	83
6	83	70	78
7	83	80	80
8	80	38	94
9	78	90	NA
10	83	80	NA
11	78	80	86

3. Data Collection Procedures

Fall 2007 Baseline Data Collection. We developed the schedule for the fall data collection to align as closely as possible with the start of the school year, with year-round classes and classes that started in late August or early September scheduled for the beginning of the field period. Nonetheless, some schools were unable to accommodate the team's visit during the scheduled week, and we honored their requests to change the timing. We completed all fall data collection between September 10 and November 9, 2007.

MPR assigned 52 certified interviewers/assessors, 29 of whom were English-Spanish bilingual, to the fall data collection. Teams, typically consisting of one team leader and two assessors/interviewers, conducted the child assessments and parent interviews. The composition of the teams fluctuated, with assignments based on program locations and the percentage of Latino children in the programs. Although space was tight in many programs, MPR field staff worked with program staff to identify the best conditions possible for conducting the child assessments and parent interviews; staff used quiet corners of hallways, staff lounges, or partially blocked off parts of the classroom as necessary. Parent interviews

---

[9] For this reliability visit, we observed a bilingual English/Spanish classroom. The gold-standard trainer, however, was not a bilingual English/Spanish speaker, whereas trainees 1 and 6 were. We considered their lower reliability scores acceptable for two reasons: (1) the two trainees were reliable with one another, and (2) the trainees believed that their lower agreement with the gold-standard occurred because they witnessed events that the gold-standard missed because of language.

usually took place at the start and end of the school day, in the evening, and during the children's lunch period. Any parent interviews not completed during the week of the data collection visit were attempted by telephone. Any child assessments not completed during the week were attempted at a future date but as close to the week of the original visit as possible.

Spring 2008 Data Collection. In general, the data collection procedures for the spring mirrored the procedures for the fall, with the following two changes: (1) although we attempted to schedule programs to maintain the sequence followed in the fall data collection, some programs requested earlier or later visits; and (2) our field staff contacted parents during the weekend before the scheduled school visit either to conduct a telephone interview or schedule the in-person visit. We used 46 field staff in the spring, 25 of whom were bilingual. Data collection took place between April 14 and June 13, 2008.

Winter/Spring 2008 Observations. Trained and reliable CLASS observers visited 88 of the 97 LAUP classrooms sampled for the classroom observations in March and April 2008.[10] The majority of observations occurred between March 11 and April 7 (3 classrooms were observed between April 28 and June 4). For cost reasons, we randomly selected 1 classroom from each program for the CLASS observations. From the 97 programs, we selected a subsample of 20 for conducting LISn observations as a pilot study (described in Appendix D). We trained 11 observers to conduct CLASS observations only and trained 3 other observers in both the CLASS and LISn measures. On the day of the classroom visit, observers completed all observational ratings, recorded their scores on the observation form, and completed notes in a booklet developed by MPR. On average, the observation period lasted between two and three hours for up to five CLASS rating cycles or, in the subsample participating in the LISn pilot study, consisted of a mix of CLASS cycles and LISn snapshots. In classrooms where LISn observations were carried out in addition to CLASS, only three or four CLASS rating cycles were completed. In Table II.5, we show the schedule for data collection in CLASS-LISn combined classrooms.

In accordance with CLASS guidelines, observers watched the "who, what, and how" of everything that happened at the classroom level for 20 minutes, with particular attention to the teachers' instructional behaviors and interactions with children. Thereafter, observers spent 10 minutes considering their scoring guidelines and recorded and derived scores for each dimension based on their informed judgments. They also completed the follow-up classroom context sheet to record global classroom content and structure during the 20-minute observation period. These activities make up one CLASS observation cycle. During the class sessions, observers were trained to complete a total of five cycles. However, in classrooms where outdoor recess took place or where the LISn was also conducted,

---

[10] We were unsuccessful in scheduling observations at 9 of the centers, for a number of reasons. Of the 88 completed CLASS observations, we completed 85 with the lead teacher and 3 with the assistant teacher when the lead teacher was absent and the observation could not be rescheduled.

Table II.5.   Data Collection Schedule for CLASS and LISn Observations

	Observation Instrument	Time in Minutes
Cycle 1	CLASS observe	20
	CLASS record	10
	LISn A	7 (5 to observe; 2 for global coding)
	LISn B	7
	LISn C	7
Cycle 2	CLASS observe	20
	CLASS record	10
	LISn A	7
	LISn B	7
	LISn C	7
Cycle 3	CLASS observe	20
	CLASS record	10
	LISn A	7
	LISn B	7
	LISn C	7
Cycle 4	CLASS observe	20
	CLASS record	10
	Additional LISn, if time	7
Cycle 5 (if a CLASS-only classroom)	CLASS observe	20
	CLASS record	10

observers may have had the opportunity to observe only three or four CLASS cycles (CLASS observations do not occur during outdoor recess). Overall, 47.0 percent of classrooms were observed for five CLASS cycles, 40.5 percent were observed for four CLASS cycles, and 12.5 percent for three CLASS cycles.

D. STUDY DESIGN, RESEARCH QUESTIONS, AND ANALYTIC APPROACH

We developed a descriptive design for the UPCOS study in order to provide maximum information about children, families, teachers, and classroom practices that would be useful for a formative evaluation of this new program. The study yields rich information about who is served, how services are implemented, and how children develop over time. Given that the study did not include a comparison group that did not receive the program intervention, we cannot interpret changes over time as caused by participation in LAUP programs; other, unmeasured variables could be the cause of observed change. Furthermore, families could freely enroll in programs of their choice and did so in non-random ways. Thus, according to our analysis, children who were clustered within a given program tended to be more like each other than like those in another program. The extent to which children or families within programs are more similar than those in different programs introduces a "design effect" that reduces the power of our analyses to detect differences and, more important, to

disentangle reasons for any differences that may be present.[11] Because we did not randomly assign children to programs, any differences at the end of the year cannot be attributed to features of the programs since they may be due to whatever factors are related to the reasons for children to be enrolled in particular programs (including geographic or family choice). Before we describe the research questions and our approach for answering each one, we provide a brief summary of the weighting, scoring procedures, and basic statistical techniques used in the study.

1. Weighting Procedures

Because we conducted UPCOS with a representative sample, we wanted to construct sample weights so that our analyses of the data we collected would truly represent the population of children and programs from which we sampled. The spring 2008 sample weights represent the original (fall 2007) study population that remained eligible through the spring; that is, the weights represent the sample still enrolled in one of the sampled eligible programs. We constructed the weights in two phases: first at the program level and then at the child level. The program weights account for the probability of selection of programs, their eligibility, and their willingness to participate. By spring 2008, 97 of the 120 originally selected programs were eligible and still participating. The base child weights account for the probability of selection of children within programs, their eligibility, and whether we obtained parental consent. By the spring, among the 1,954 children originally selected, we counted 1,555 eligible children with consent in the 97 remaining programs. Finally, we adjusted the base child weight for whether a child assessment was completed in both the fall and spring among the eligible children with parental consent. The result was 1,438 such children. The product of the program- and child-level weights is the final child-level weight.

2. Scoring Procedures and General Analytic Approaches

Given the diverse sample and the high percentage of Spanish-speaking households, we conceptually scored the majority of our child assessment measures by using the publishers' scoring rules for published assessments (EOWPVT-SBE and ROWPVT-SBE). Recognizing that the standardization samples differ for the English-only version and the Spanish-Bilingual version and that a few items are not administered in the Spanish-Bilingual edition because of potential bias in the items, we used item response theory (IRT) methods to scale the measures for all the children so that we had a common scale.[12] IRT methods do not

---

[11] We calculated intraclass correlation coefficients (ICCs) for each of the school readiness outcome measures within classrooms and found most fell above the 10 percent level we had assumed in our power calculations (the range in the fall was from a low of 1.6 percent for parent-rated internalizing problem behaviors to a high of .31 for teacher-rated social interaction). In general, scores were even more clustered for the spring measures. A high ICC indicates that children of similar ability levels were grouped in classrooms in a nonrandom way, thereby reducing the effective sample size and our ability to detect small differences between groups.

[12] The IRT scaling procedure means that only approximations of the scores can be used to compare the entire sample to national norms; however, it permits stronger estimates of change over time and allows comparisons to be made among children within the sample.

require complete data but instead estimate scores based on (1) the child's responses to the questions that are administered, and (2) the item difficulties as estimated from information gained from the entire sample. In the case of the ECLS-B Mathematics assessment, we used the item difficulties from the national ECLS-B preschool sample when scoring the assessments. For the child assessment outcome measures that differed by language (WJ-III and WM-III Spelling tests), we scored each according to standard scoring procedures as specified in the manuals.

We collected demographic information from parent interviews. In addition, parents provided information about a variety of child and family characteristics, such as family structure, children's health, languages spoken at home, household income, parent mental health, time in the United States, involvement in the program, challenges to involvement in the program, and satisfaction with the program.

For the hierarchical linear models (HLMs) we describe below, in addition to variables at the child/family level, we attempted to correct for selection bias in the sample by creating measures that best capture program-level composition possibly related to children's outcomes. The measures included the averages and variability in fall baseline ability, average socioeconomic status of families in the program, percentage of recent immigrant families in the program, and percentages of Spanish speakers in the classroom.

In examining classroom quality, we obtained LAUP data on the Star quality ratings and the ECERS-R scores for the classrooms in the program. LAUP uses ECERS-R scores to determine the Star rating, along with other factors such as teacher education and child-teacher ratios. We aggregated classroom-level ECERS-R scores to the program level.[13] From the teacher interviews, we obtained teacher education and experience working with young children. We collected data on group size and child-adult ratio during our classroom observations. In our model examining classroom quality, we also included variables that, we hypothesized, might be related to differences in quality, such as the mean socioeconomic status of families in the program (aggregated from parent-reported income), AGSN status, and program auspices.

We used two basic analytic approaches: descriptive and multivariate analysis. Descriptive analysis included an examination of means and variation in child and family characteristics and children's school readiness scores overall and across key subgroups. Multivariate approaches used multiple regression as well as HLM techniques. HLM is appropriate for analyzing data measured at several levels (in this case, at the individual child and program levels). It allows us to represent the structure of the data and examine the residual variability at each level. With children nested within programs, the observations are not independent of one another, but HLM accounts for the lack of independence and adjusts the standard errors for the estimates accordingly. We used a fixed effects model to examine child-level associations.

---

[13] We were unable to match classroom level ECERS scores from LAUP to our sample and therefore aggregated scores to the program level for these analyses.

3.   Research Questions

Here, we elaborate on our approaches for answering the study's three overarching research questions and identify specific subquestions for exploration in Chapter III. The first question we address asks:

1. What is the overall level and range of quality in the implementation of LAUP/PoP programs?

We answer the question by describing the characteristics of study classrooms and programs (such as curricula used, average class size, child-adult ratio, teacher reports of different instructional approaches, and so on). Our quality indicator is the CLASS, as described earlier, and we examine overall program quality by using means and variability of the CLASS subscales. We further explore associations with these observed quality ratings by using univariate and multivariate regression models that examine the relationship of teacher education and experience, group size and child-adult ratio, program characteristics, and the Star quality ratings that LAUP assigned to programs.

In addition to the analysis of classroom quality, we examine other aspects of implementation as reported by parents and teachers, including parent involvement with the LAUP program, parent reports of challenges to involvement, and parent reports of satisfaction with the program. We also compare parent satisfaction across family subgroups (ethnicity).

The second question we address in this study asks:

2.   How do children enrolled in LAUP/PoP programs develop from fall to spring?

Using descriptive statistics, we examine children's development in the spring followup on each of the school readiness measures and report means and standard deviations of scores in our sample relative to the fall baseline. To account for the clustering of children within programs, weighted estimates and significance tests were computed using SUDAAN®.[14] In addition to overall analyses of children's progress between the fall and spring measures, we examine how those scoring at the highest and lowest quartiles performed. This approach allows us to understand whether progress is occurring uniformly across the enrolled children or is concentrated in segments of the population of LAUP children. Given the strong influence of language during the early years and the fact that a large portion of the sample speaks a language other than English, we also present children's school readiness in the fall and spring separately by language subgroup. We contrast school readiness of children in four language groups as derived from fall 2007 parent reports: (1)

---

[14] SUDAAN®, a statistical software package that specializes in providing efficient and accurate analysis of data from complex studies, is ideal for the proper analysis of data from surveys and experimental studies, since its procedures properly account for complex design features, such as correlated observations, clustering, weighting, and stratification (Research Triangle Institute 2001). Variance estimation was done using the Taylor Series method.

English only or primarily, (2) Spanish only, (3) Spanish primarily, and (4) other language only or primarily.[15] For simplicity of presentation, in this report we refer to these groups as English, Spanish only, Spanish primarily, and Other.

Our third question asks:

3. How are characteristics of children and families related to school readiness outcomes?

Using a multivariate HLM analysis to answer this question, we were able to examine development in the context of all of the child characteristics, then child and family characteristics simultaneously, and, finally, the self-selection of families and children into programs after controlling for the individual characteristics of children and families.

Level 1 of the HLM models included child and family characteristics, as well as time between assessments and performance in the fall, to explain children's development in the spring. At the child level, the key domains included in the models were the child's fall score on the respective outcome measure; number of weeks between assessments; language category based on step 1 of our routing criteria (Appendix B); demographic information, including maternal education and employment, number of parents born outside the United States, and recency of immigration; household poverty measures; and parent mental health. Because children and families were not randomly assigned to programs but instead self-selected into them, we included controls at the program level (including means and standard deviations of child measures in the fall, mean socioeconomic status, percentage of recent immigrants, and concentration of Spanish speakers based on teacher report). In Table II.6, we summarize the variables included in each of the models predicting children's spring school readiness scores. In Table II.7, we list the spring school readiness outcome measures that we employed. (Appendix A provides information on each measure used in the study.)

First 5 LA had hoped that this study would address a fourth research question, one that would explore the relationships between program characteristics (such as classroom quality and teacher characteristics) and children's school readiness outcomes. In nonexperimental research designs such as UPCOS, researchers approach this type of question with considerable caution. Because families self-select into the programs, it is possible, even likely, that family characteristics (such as family income or parent education) will be associated with the type of preschool center the child enrolls in. We examined the possibility of this "selection bias" in our sample by calculating the ICCs for children's initial abilities in the fall and the correlations between average initial abilities of children in the classroom and certain

---

[15] The language routing procedure we developed for determining the language of assessment for each child is described in Appendix B. This leads to more language "groups" than these four. However, for analyzing the possibly differential performance of children in various language groups, it was important to have a group designation that would remain constant from fall to spring. These four groups derived from the parent report at the time of the fall baseline parent interview provide that consistency.

Table II.6   Explanatory Variables Included in Models Predicting Children's Spring School Readiness Outcomes

Construct	Coding and Reference Group
Child initial ability	We included the child's fall performance on the outcome of interest as a covariate in the model. For example, models predicting children's performance on the EOWPVT in the spring include the child's fall EOWPVT score in the analysis.
Assessment time interval	Number of weeks between the child's fall and spring assessments.
Child Characteristics	
Child gender	Male Female (referent group)
Child age in months	Child age in months as of September 1, 2007, which we designated as the start of the program year.
Child race/ethnicity	White African American, non-Hispanic Hispanic/Latino (referent group) Asian/Pacific Islander Other/multiple race Given the demographics of the sample, the referent group is Hispanic/Latino for all HLM analyses.
Child language group	English only and primarily (referent group) Spanish only Spanish primarily Other language only/primarily Given that instruction in California kindergartens is in English only, the referent group is English-only and -primarily for all HLM analyses.
Family Characteristics	
Mother's education level	High school diploma or higher Less than high school diploma (referent group)
Mother's employment status	Working full- or part-time Not in workforce (referent group)
Parent's immigration status	Mother and father born in the United States (referent group) One parent born outside the United States Both parents born outside the United States
Parent's recency of immigration	Less than or equal to 5 years More than 5 years (referent group)
Household dependency ratio	Household dependency is a continuous variable constructed from the number of children divided by number of adults in the household.
Family poverty ratio	The family poverty ratio is treated as a continuous variable: Less than 50% of the poverty threshold 50 to 99% of the poverty threshold 100 to 129% of the poverty threshold 130 to 184% of the poverty threshold 185 to 239% of the poverty threshold At or above 240% of the poverty threshold
Parent mental health	Severity of parent depressive symptoms is a continuous variable. Scores can range from 0 to 36.

Table II.6 *(continued)*

Construct	Coding and Reference Group
High risk	Children in families with three or more of the following risk factors: maternal education less than high school, mother not employed, family income below the poverty threshold, either parent not born in the United States, either parent living in the United States for 5 or fewer years, single parent, household size of 5 or more, mother at risk for depression, teenage mother, and high exposure to violence
	Children in families with 2 or fewer risk factors (referent group)
Program Characteristics	
Percentage Spanish speakers	25% or fewer of the children speak Spanish 26–74% of the children speak Spanish (referent group) 75% or more of the children in the program speak Spanish
Program-level socioeconomic status	Mean of the family poverty ratio within the program
Percentage of recent immigrants	Percentage of children in the program with a parent who has been in the United States for 5 or fewer years
Mean program-level child ability	Z-score of the fall mean for children in the program on a given outcome measure
Variation in child ability	Z-score for the variance in ability in the program on a given outcome measure

Table II.7 Child School Readiness Outcome Measures, Spring 2008

Construct	Measure
Language and Literacy	Expressive vocabulary—Expressive One Word Picture Vocabulary Test Spanish Bilingual Edition (EOWPVT-SBE) Receptive vocabulary—Receptive One Word Picture Vocabulary Test (ROWPVT) Letter knowledge—Rapid Letter Naming
Mathematics	ECLS-B Mathematics
Social-Emotional	Teacher Report—Preschool Kindergarten Behavior Scale (PKBS) PKBS total positive social skills PKBS social cooperation PKBS social interaction PKBS social independence Teacher Report—Social Skills Rating System (SSRS) SSRS problem behaviors Parent Report—Adapted Preschool Kindergarten Behavior Scale (PKBS) Social cooperation Social interaction/independence Externalizing problem behaviors Internalizing problem behaviors
Approaches to Learning	Teacher Reports of Children's Approaches to Learning (ECLS-K) Executive Functioning Pencil Tapping Walk-a-Line Slowly

classroom characteristics. The results indicate that the ICCs for children's initial abilities were around .2, suggesting that a substantial proportion (20 percent) of the variance in children's initial abilities was associated with program characteristics even before the program could have influenced the children's performance. Furthermore, the average initial abilities of children in the classroom were strongly correlated with classroom characteristics such as mean income of families in the classroom and concentration of ELL children in the classroom, with the magnitude of the correlations ranging from .5 to .7. These findings demonstrate that families enrolled their children in programs in nonrandom ways and either made choices based on program characteristics or enrolled children in particular centers for reasons we don't understand, but that resulted in children with similar characteristics being clustered together in centers. Therefore, we decided it would be misleading, and inappropriate, to examine the associations between program characteristics and the gains that children made from fall to spring.

CHAPTER III

FINDINGS: LAUP PROGRAMS, CHILDREN, AND FAMILIES

A. OVERVIEW OF PROGRAMS, CLASSROOMS, AND TEACHERS

In this chapter, we report the findings from our analysis of the UPCOS data collected in 2007-2008. We first provide an overview of the LAUP programs, their teachers, and classroom activities. We then describe the children and families enrolled in the programs. Next, we report on the children's development from fall to spring in all the domains of school readiness that we measured, with a particular focus on differentiating the progress of the lower- and higher-performing children and children from the various language groups included in the study. We then report on parent involvement and conclude the chapter with an analysis of the relationships of the characteristics of children and families (as well as program composition) to children's school readiness outcomes. In Chapter IV, we summarize the findings and discuss their implications. Because the large number of tables would disrupt the flow of the presentation, we put Tables III.5 through III.7 and III.9 through III.49 in Appendix C; some of the key results, however, are displayed in figures within this chapter.

1. Program Characteristics

Our sample of 98 programs included 23 of the 26 PoP programs (24 percent of the sample; Table III.1). A total of 57 programs were located in AGSNs within Los Angeles County. About two-thirds of programs in the study sample (67) were from the first round of funding, or Early Launch, while about one-third (31) were funded in the Second Round. Although programs frequently blended funding, almost 40 percent of programs were school-based.

In the 98 participating programs, the study sample included 187 classrooms. Most programs were small: half (51 percent) had two classrooms; one-third (34 percent) had one classroom.[1] The maximum number of classrooms in any single program was six, with just one program reporting that figure. Among the 82 programs reporting enrollment data,

---

[1] We use the term "classroom" to refer to individual class sessions, not a physical space in the program.

program size ranged from 15 to 116 children, with an average enrollment of 37 children. The average classroom had 6.3 children to each adult, and the average class size was 17.7, based on our classroom observations (Table III.13 in Appendix C). Class size ranged from 9 to 28 children. Teachers also reported enrollment in each class during the teacher interview. Karoly et al. (2008) reported that in their representative state sample, 77 percent of the classrooms had child:staff ratios no larger than 10:1; 91 percent of LAUP classrooms had child:staff ratios of 10:1 or better, which meets the accreditation standard set by the National Association for the Education of Young children and more than meets California's state child care licensing requirements. However, 83 percent were at 8:1 or better, the requirement for an LAUP five-Star rating.

Table III.1   Characteristics of Programs (Weighted)

Characteristic	Number	Percentage
Area of Greatest Service Need		
AGSN	57	58.2
Non–AGSN	41	41.8
Type of Program		
PoP	23	23.5
Non–PoP	75	76.5
Funding Cycle of Program		
Early Launch	67	68.4
Second Round	31	31.6
Program Auspices		
Operated by school district	39	39.8
Not operated by school district	59	60.2
Sample Size	**98**	

Source: LAUP administrative data.

2.   Teachers

The study included 138 lead teachers of 1,434 children.[2] Slightly more than half of children (52 percent) received instruction from Latino teachers, and 57 percent of children received instruction from teachers who reported speaking a language other than English at home (mainly Spanish, although 5 percent of children had teachers who reported speaking English and another language at home; Table III.2). Sixty-two percent of children had teachers who reported reading to them in both English and Spanish, and another 34 percent

―――――――――――――――――――

[2] Teacher responses were weighted at the child level. The results in this section represent the actual number and the weighted percentage of children served by teachers with a given characteristic. We provide teacher characteristics weighted at the program level in Appendix C (Tables III.5 to III.7); because the proportions relative to the child weight are similar, we do not describe them in the text.

of children had teachers who read in English only.[3] Similarly, half the children (54 percent) had teachers who made classroom instructional presentations in both languages, and 41 percent of children had teachers who made classroom presentations in English only. Approximately 5 percent of children had teachers who read or gave instructional presentations in a language other than English or Spanish.

Table III.2.   Distribution of Children Served by LAUP Lead Teachers: Demographic Characteristics and Language Use (Weighted)

Teacher Characteristic	Number of Children	Percentage of Children[b]
Gender		
Female	1,387	98.3
Male	24	1.7
Race/Ethnicity[a]		
Latino	675	51.8
African American, non–Latino	234	15.4
White, non–Latino	258	17.2
Asian, non–Latino	91	7.0
Other	105	8.6
Language Spoken at Home[a]		
English	590	43.3
Spanish	14	1.4
English and Spanish	570	45.4
Other language	62	4.9
English and other language	67	5.1
Language Used to Read to Children		
English	439	34.4
English and Spanish	713	61.7
English and another language	36	3.9
Language Used for Presentations[a]		
English	594	41.0
English and Spanish	720	53.9
English and another language	58	5.1
Sample Size	**1,434**	

Source: Teacher interviews, fall 2007.

[a]Categories are mutually exclusive.
[b]Categories with missing data do not sum to 100 percent.

---

   [3] We note that a higher proportion reported reading in both English and Spanish (62 percent) than reported speaking Spanish (or English and Spanish) at home (47 percent). It is not clear whether this reflects fluency in both languages or includes some English-speaking teachers who could read children's Spanish-language books.

On average, children's lead teachers reported 15 years of experience in working with children and a median experience level of 12 years, 10 of which were teaching preschool (Table III.3). Eighty-eight percent of children's teachers held an associate's degree or higher, and 61 percent held a bachelor's degree or higher (Table III.4), in contrast to statewide data indicating that 67 percent of lead teachers held at least an AA degree and 42 percent held a BA or higher degree (Karoly et al. 2008). Fewer than 1 percent of children had a teacher whose highest educational attainment was high school or high school equivalency.

Table III.3.   Experience of Children's Lead Teachers (Weighted)

	Mean (Standard Deviation)	Median
Years working with children	14.7 (1.2)	12
Years teaching preschool	11.5 (0.8)	10
Sample Size	**1,434**	

Source: Teacher interview, fall 2007.

Table III.4.   Education and Credentials of Children's Lead Teachers (Weighted)

Teacher's Education and Credentials	Percentage of Children[b]
Highest Grade of School Completed	
High school diploma/equivalent	0.6
Some college but no degree	11.9
Associate's degree	27.1
Bachelor's degree	44.6
Graduate or professional school, no degree	7.7
Graduate or professional degree	8.2
Field in Which Obtained Highest Degree[a]	
Child development/developmental psychology	37.0
Early childhood education	27.1
Other	35.9
College Courses Included 6 or More Classes in Early Childhood or Child Development	99.0
Has a Child Development Associate Credential	48.3
Has a State-Awarded Preschool Certificate	87.2
Has a Teaching Certificate or License	73.6
Sample Size	**1,434**

Source: Teacher interview, fall 2007.

[a]Includes only those teachers with an associate's degree or higher (N = 1,211).

[b]Categories with missing data do not sum to 100 percent.

III: Findings: LAUP Programs, Children, and Families

Among classrooms with teachers holding a college degree (AA or higher), slightly more than one-third (37 percent) of children had a teacher whose highest degree was in child development or developmental psychology. An additional 27 percent of children's teachers with a college degree obtained their highest degree in early childhood education. Nearly all of children's teachers (99 percent) had taken six or more courses in early childhood or child development, and 87 percent had obtained a state-issued preschool certificate. Seven in 10 children had a teacher with a teaching certificate or license.

3. Classroom Curricula and Activities

We asked teachers to name the primary curriculum they used with the children in their classroom.[4,5] The most commonly used published curricula were High/Scope, used in teaching 18 percent of the children, and Creative Curriculum, used in teaching 13 percent of the children (Table III.8). Fewer than 10 percent of the children were taught by teachers using any other published curriculum. Teachers of about one-third of the children reported using a nonspecific curriculum, including those described as "thematic," "Armenian," and "my own." In all, more than half (58 percent) of the curricula reported by teachers addressed

Table III.8.   Children Taught with Each Primary Curriculum (Weighted)

Curriculum Name	Children	
	Number	Percentage
High/Scope	232	17.9
Creative Curriculum	156	13.3
Houghton Mifflin	102	8.7
DLM	63	7.7
Open Court PreK	57	5.0
Scholastic Curriculum	73	4.5
Reggio Emilia	41	3.8
Montessori	45	2.2
Step by Step	17	1.6
Let's Begin with the Letter People	19	1.3
Doors to Discovery	12	0.8
Locally Designed	5	0.3
Other, nonspecific curriculum	478	33.0
Sample Size	**1,300**	

Source: Teacher interview winter 2008.

---

[4] Estimates represent only the primary curriculum teachers reported using in their classroom. In some cases, teachers in the same program reported using different primary curricula.

[5] LAUP has standards for the curricula that programs use. The Operating Guidelines specify that curricula must be (1) research based, (2) comprehensive, (3) integrated, (4) inquiry based, (5) responsive and sensitive, (6) incorporating family involvement, (7) consistent with California's pre-K Learning and Development Guidelines, and (8) based on ongoing screening and assessment (LAUP 2008b).

more than one domain of development. Almost all teachers (94 percent) reported receiving training and support in curriculum use, and they reported receiving a median of 21 hours of training and support in the past 12 months. The training and support came from the curriculum developers (32 percent), other staff at the center (24 percent), LAUP coaches (14 percent), or other sources (22 percent). Few staff noted receiving training or support from staff at other centers (5 percent) or faculty from a school of education (3 percent)(not shown).

Most children had teachers who reported that they undertook reading and language activities daily or almost daily (Table III.9 in Appendix C). The most common reading and language activities included listening to the teacher read stories where children see the print (100 percent), working on letter naming (92 percent), writing their own name (90 percent), discussing new words (84 percent), learning about conventions of print (81 percent), and practicing writing letters (78 percent). At least 75 percent of children had teachers who reported engaging in the above activities daily or almost daily. Children's teachers were very unlikely to read stories in which children could not see the print (37 percent reported never doing so).

All children's teachers reported counting out loud with children daily or almost daily, and most reported that they engaged in various mathematics activities daily or almost daily (Table III.10, Appendix C). At least 75 percent of children had teachers who reported engaging in all but three mathematics activities daily or almost daily. Somewhat fewer teachers reported engaging in the following three activities daily or almost daily: working with rulers or other measuring instruments (71 percent), using music to understand mathematics concepts (68 percent), and engaging in activities related to telling time (68 percent).

4. Screening and Assessment

Approximately one-third of the children (32 percent) were in classrooms where teachers reported that none of the children were screened for health or developmental problems, whereas 41 percent were in classrooms where teachers reported that all of the children were screened (not shown). In the majority of these cases (81 percent), the screening was completed by the teacher, whereas 28 percent were screened by a teaching assistant, health professional, supervisory or support staff, or often, by parents (screening could be completed by more than one type of staff). More than one-third of the children (34 percent) were in classrooms where the primary screening tool was Ages and Stages Questionnaires: A Parent-Completed Child Screening System (ASQ; Bricker and Squires 1999) (Table III.11, Appendix C). Other screening tools that were less frequently reported included other parent-report measures such as the Child Development Inventory (CDI; Ireton 1992) (7 percent) or the Parents' Evaluation of Developmental Status (PEDS; Glascoe 2006) (4 percent), as well as directly administered screening tools, including the Developmental Indicators for Assessment of Learning 3rd Edition (DIAL; Mardell-Czudnowski and Goldenberg 1998) (3 percent) and the Brigance Preschool Screen for Three- and Four-Year-Old Children (Brigance 2005) (1 percent). Just under 10 percent of the children were in classrooms that reported a Level 2 Health Screening as their main screening tool. Forty-two percent of the children were in classrooms that reported using "other" screening tools. These included a

combination of a teacher-constructed tool and an ongoing assessment tool such as the DRDP-R.

Slightly more than half of the children (53 percent) were in classrooms where teachers reported that all children were assessed. Almost 95 percent of these children were assessed by their teacher. Children were also assessed by an outside testing group (5 percent), a specialist in the program (2 percent), or others (11 percent). A large majority of the children (82 percent) were assessed using the DRDP-R as the primary assessment tool. The Creative Curriculum Developmental Continuum Assessment Toolkit for Ages 3-5 (Dodge, Colker, and Heroman 2002) was the only other assessment that was used with more than 5 percent of the children as the primary assessment. The most frequently reported methods used for assessments included classroom observation and work samples (98 percent), portfolios (93 percent), anecdotal records (85 percent), and family input (75 percent). Teachers of half the children also reported using standardized tests (50 percent). Almost all the children's teachers (97 percent) reported using the information gained from screenings and assessment. When asked how they use this information, teachers reported using it to identify the child's developmental level (97 percent), determine a child's strengths and weaknesses (88 percent), individualize activities (86 percent), make referrals for special services (74 percent), report to parents (73 percent), identify home activities for parents (73 percent), and report information to funding sources (33 percent) (not shown).

5. Classroom Quality

The classroom observation measure used in this study, the Classroom Assessment Scoring System or CLASS used in UPCOS (see Chapter II and Appendix A), measures classroom quality across three domains of interaction: (1) Emotional Support, (2) Classroom Organization, and (3) Instructional Support (Pianta, La Paro, and Hamre 2006). The Emotional Support domain assesses the emotional supportiveness of the classroom environment across four dimensions (Positive Climate, Negative Climate, Teacher Sensitivity, and Regard for Student Perspectives), while the three dimensions of Classroom Organization (Behavior Management, Productivity, and Instructional Learning Formats) capture the teacher's ability to organize the classroom and manage behavior to make efficient use of class time for instruction (See Box III.1). Instructional Support measures the quality of instructional practices used in the classroom (Concept Development, Quality of Feedback, and Language Modeling). Observers rated each dimension on a 7-point scale, with anchor behavioral descriptions for low (1 to 2), middle (3 to 5), and high (6 to 7) scores.

Although the CLASS does not rely on normative data, the CLASS Technical Appendix (Pianta et al. 2006) reports mean scores from several large-scale studies that used CLASS or its precursor, the Classroom Observation System

---

Box III.1 CLASS Domains and Dimensions

Domain: Emotional Support
    Positive Climate
    Negative Climate
    Teacher Sensitivity
    Regard for Student Perspectives

Domain: Classroom Organization
    Behavior Management
    Productivity
    Instructional Learning Formats

Domain: Instructional Support
    Concept Development
    Quality of Feedback
    Language Modeling

III: Findings: LAUP Programs, Children, and Families

(COS), which provide one basis of comparison. COS instruments rated more than 1,500 preschool classrooms in studies such as the National Center for Early Development and Learning's Multi-State Study of Prekindergarten and State-Wide Early Education Programs and the "My TeachingPartner Study." We also compare UPCOS data with a recent RAND study of child care quality in California (Karoly et al. 2008). In Table III.12 (Appendix C), we present the mean CLASS dimension scores of LAUP classrooms alongside scores obtained in other studies of preschool programs.

Mean Emotional Support scores on the Positive Climate, Regard for Student Perspectives, and Teacher Sensitivity dimensions were all 0.7 to 1 point higher than the averages reported in the CLASS Technical Appendix, and the Negative Climate dimension was about 0.35 point lower (Table III.12 in Appendix C), indicating that LAUP classrooms compared favorably with those studies.[6] Because the dimension scores were all more positive than the CLASS averages reported in the CLASS Technical Appendix, LAUP classrooms also scored higher on the Emotional Support domain than those CLASS Technical Appendix averages. LAUP classrooms scored higher on each of the Emotional Support dimensions (and on the overall domain) than the statewide means in the RAND study as well, although the differential was smaller, ranging from 0.2 to 0.4 (with a Negative Climate score that was 0.2 points lower).

The trend continued in the Classroom Organization domain, with LAUP classrooms receiving quality ratings ranging from 0.2 to 1.4 points higher than the averages listed in the CLASS Technical Appendix for all three dimensions: Behavior Management, Productivity, and Instructional Learning Formats. Similarly, LAUP classrooms scored higher than those in the RAND study by 0.2 to 0.6 points, and 0.4 on the dimension overall. Finally, in the Instructional Support domain, LAUP classrooms received average ratings on Concept Development, Quality of Feedback, and Language Modeling dimensions that were either equivalent to those in the CLASS Technical Appendix or higher by up to half a point. In this domain, LAUP classrooms scored slightly lower than those in the RAND study, by 0.1 point, due to lower means on Concept Development (0.3 points lower), and Quality of Feedback (also 0.3 points lower), offset somewhat by a 0.4 higher score on Language Modeling. In sum, classroom quality scores in LAUP classrooms observed during the study, for the most part, were on a par with or higher than the averages reported elsewhere for the measure (Tables III.12 and III.13 in Appendix C).

We also looked at each dimension and the three domains to assess the proportion of classrooms receiving low (1 to 2), middle (3 to 5), and high (6 to 7) ratings. In the Emotional Support domain and all of its dimensions, half of the classrooms received a high rating

---

[6] With few exceptions, the mean and median scores were similar on each dimension and for overall domain scale scores. Therefore, we report the means here for ease of comparison with previously published results from the CLASS Technical Appendix and the RAND study. Exceptions were the Teacher Sensitivity dimension (mean of 5.4, median of 6) the Behavior Management dimension (mean of 5.5, median of 6), and the Quality of Feedback dimension (mean of 2.5, median of 2). In all cases, both the mean and the median of LAUP classroom scores on the dimensions were higher than the means reported in the CLASS Technical Appendix.

(Figure III.1; for the reverse-coded Negative Climate, nearly all classrooms had a low score). More than half of the classrooms had high ratings on Behavior Management and Productivity, and one-third had high ratings for Instructional Learning Formats. For the total Classroom Organization domain, one-quarter of classrooms were highly ranked, and nearly three-quarters were mid-ranked (Figure III.2). In fact, half of classrooms received a Classroom Organization rating of 5.

Ratings were lower on all dimensions in the Instructional Support domain than in the other two domains (Figure III.3), consistent with findings reported in the CLASS Technical Appendix. Observers rated slightly more than one-third of classrooms (37 percent) at a middle level of Instructional Support and the remainder at low levels. This result is explained primarily by most classrooms scoring low on Concept Development (74 percent) and Quality of Feedback (60 percent), although half of classrooms earned a middle rating on Language Modeling. The recent RAND study indicated that 24 percent of the programs scored at or above 3.2 on the Instructional Support domain. LAUP classrooms were similar, with 26 percent scoring that high.

Figure III. 1 CLASS Emotional Support Ratings
54, 0, 45.9
Low Middle High

Figure III. 2 CLASS Classroom Organization Ratings
72.7, 26.6, 0.7
Low Middle High

Figure III. 3 CLASS Instructional Support Ratings
63.2, 36.7, 0
Low Middle High
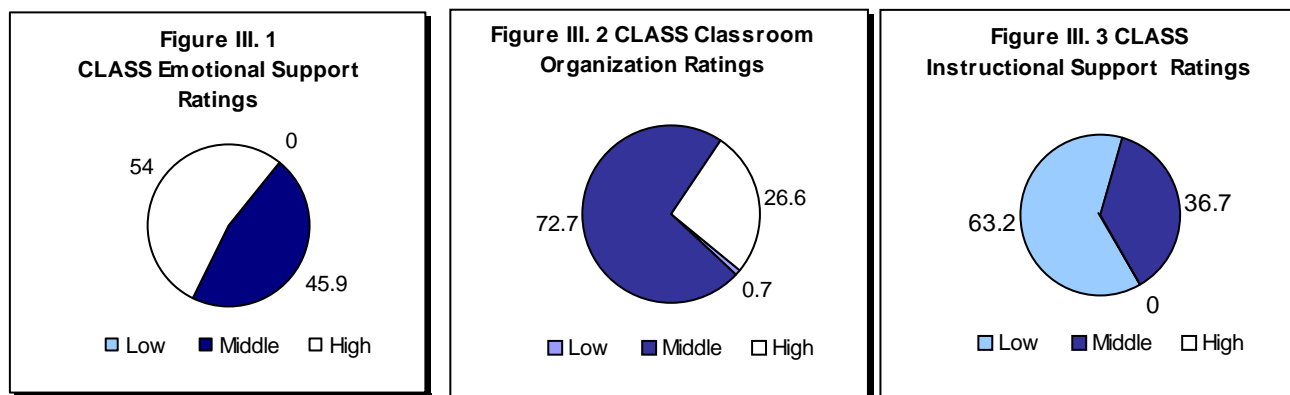
6. Relationships Between CLASS Domain Scores and Teacher and Program Characteristics

To investigate the relationship between Star ratings (which draw mainly on staff qualifications and ECERS-R scores) and observed classroom quality as measured by the CLASS scale, we examined the Star ratings and their components independently. In Table III.14 (Appendix C), we present three models predicting each domain of the CLASS (all weighted with classroom-level weights). Model 1 is a simple ordinary least squares (OLS) regression regressing standardized (z-score) of CLASS domain scores on Star ratings. Model 2 is a multivariate OLS specification that predicts CLASS domain scores by using indicators for the lead teacher education (whether a teacher holds at least a bachelor's degree, whether a teacher's education is less than an associate's degree), and specifying the teacher's number of years of experience in working with young children, the class size and child:staff ratio, and

the average ECERS-R score received by the program.[7] Finally, Model 3 adds variables to Model 2 that include program location (in an AGSN or not), the funding round, the mean socioeconomic status of families in the program, and an indicator for whether a school district administers the program.

We found that the Star ratings alone had a significant relationship only with the Emotional Support domain and explained only 3 percent of the variation in those ratings. The Star ratings were not significant in the models for Instructional Support or Classroom Organization.

The variance in the Classroom Organization domain could not be explained with any statistical certainty by teacher degree, but the Emotional Support and the Instructional Support domains were associated with teacher education. Teachers holding at least a bachelor's degree scored approximately one-third of a standard deviation higher than those with an AA on the Emotional Support domain. Teachers holding at least an associate's degree scored almost a standard deviation higher than those without an AA on Instructional Support. Holding a degree beyond the AA was not related to any increased score on the Instructional Support domain.

After controlling for teacher education, experience, the ECERS-R scores, and child:adult ratio, as well as after controlling for auspices, AGSN status, and the average socioeconomic status of families enrolled in the program, each domain was associated with class size. Larger classrooms were associated with lower ratings on each domain such that, for each additional child, the rating was 0.04 to 0.09 standard deviations lower. This suggests that decreasing class size by five children would have an effect on the Emotional Support domain similar to the effect of having a teacher with a BA (rather than an AA) in the classroom.

After controlling for class size, the ECERS-R, and teacher characteristics, the child:staff ratio had a weak relationship with the Classroom Organization domain. The relationship between the ratio and observed quality was not significant after controlling for program characteristics in Model 3. In other words, when aspects of the program context were added to the model, they explained some of the variance in classroom quality that had been associated with the child:staff ratio in models that did not control for these programmatic characteristics.

The mean ECERS-R score was significant in explaining each CLASS domain score. After controlling for program characteristics (AGSN, average socioeconomic status, funding round, and auspice), the ECERS-R showed a stronger relationship to Emotional Support and Classroom Organization and a weaker relationship to Instructional Support. Overall, a 1-point increase in the ECERS-R was associated with a 0.29 to 0.63 standard deviation increase in a CLASS domain score.

---

[7] The LAUP external assessors collected the ECERS-R data as part of determining the Star ratings.

After controlling for other program and teacher characteristics, Early Launch AGSN classrooms scored higher on Instructional Support by more than half a standard deviation when compared with the Second Round AGSN programs. After controlling for other characteristics, classrooms in a school-based environment (school auspice) scored more than two-fifths of a standard deviation higher on the Instructional Support domain.

Two key themes emerge from this analysis. First, it appears that the Star rating, although based mainly on teachers' education and ECERS-R score, did not predict scores on CLASS domains as well as some of its components (such as class size). Although the Star ratings include group size and ratio, they are the same for all three levels of the Star ratings. The component measures of the star rating explained more than 14 percent of the variance in each of the CLASS domain scores, whereas the explanatory value of the Star ratings on their own was much more limited (close to 0 and only significant for the Emotional Support domain). Group size is the only variable that was consistently significantly related with all CLASS quality domains. Second, we observed relationships between the program funding round and the Instructional Support domain, and between auspice and Instructional Support. Specifically, after controlling for teacher and classroom characteristics, we noted that classrooms in the Second Round programs had significantly lower scores on Instructional Support than the Early Launch programs, and school-based programs had higher Instructional Support scores than other programs.

B. CHILDREN AND FAMILIES OF LAUP AND THEIR PROGRAM EXPERIENCE

In the first part of this section of the report, we present what we learned about the characteristics of the LAUP children and families as they entered the program, overall and then by language group. In Section B.2, we describe children's development in the domains of school readiness discussed in Chapter II for the overall sample and the subsamples of children scoring in the highest and lowest quartiles of each measure's distribution in fall 2007. Section B.3 highlights the development of children by language group. We conclude the section with a summary of our findings.

1. Characteristics of the Children and Families Served by LAUP

Children and families represented a variety of racial/ethnic backgrounds (Tables III.15 and III.16 in Appendix C). Most children were of Latino descent (75 percent), followed by African American (8 percent), Asian (7 percent), white (7 percent), and other groups (4 percent). English and Spanish languages predominated, with 53 percent of children speaking English-only or primarily, 29 percent Spanish-primarily, and 13 percent Spanish-only.[8] Five percent spoke languages other than English and Spanish, including, for example, Armenian

---

[8] The groups are constructed based on parent report of languages spoken at home on the consent form, and are more specific than the questionnaire items reported in Table III.16. This maintained the same groups for analysis independent of the primary language of test administration. We examined the English-primarily language group (23 percent of the overall group) separately but found it did not differ from the English-only group on outcomes; thus, for analysis purposes the two groups are combined in this report. Henceforth we refer to the groups as English, Spanish-only, Spanish-primarily, and other-language.

and Mandarin. Nearly two-thirds of parents were immigrants, mostly from Mexico, and three-quarters of children lived in households in which at least one parent was born outside the United States (Table III.15), although nearly all children were born in the United States (Table III.16).

The children were split about evenly between boys and girls. At the fall assessments, children's ages averaged about 52 months, or almost 4-and-a-half years. By the spring assessments, children averaged about 59 months, or 5 years of age.

Families faced a range of challenges but also exhibited strengths in the face of the challenges at the start of children's LAUP experience. More than one-third of mothers and 40 percent of fathers lacked a high school diploma or GED, although nearly all families had at least one parent employed full-time, usually the father (Table III.15). Despite the high rate of employment among LAUP families, nearly 40 percent lived below the federal poverty line; 20 percent of families had moved at least once during the past year, and 4 percent had moved at least twice during that time (Table III.16). In the face of poverty and some instability in housing, we also found that 73 percent of children lived with both parents and 58 percent of parents were married. Households were small, on average, and children lived in households with 2.5 children and 2.5 adults. Children's home environments provided support for their development (Table III.17). For example, LAUP parents reported eating dinner together as a family an average of more than five days a week and noted that children maintained regular bedtimes more than four days a week. One challenge parents faced was limited understanding of English, with 40 percent of mothers reporting they did not understand English well or at all. However, about 40 percent of parents reported reading daily to their child in the fall, and just under half of children had more than 25 books in the home.

We asked parents about the time their children spent in LAUP programs each week. Parents reported a range from 3 to greater than 40 hours, with a median of 15.5 hours per week, and an average of 24 hours a week. LAUP funds 17.5 hours per child per week (3.5 hours per day); however, parents may obtain additional hours of care through subsidies or by paying for the hours. About 22 percent of children were cared for by someone other than their parents before or after their LAUP classes. Relative care was the most common type of additional non–LAUP child care (14 percent); 5 percent received before- or after-program care in another center-based program, and 3 percent received care from a nonrelative in a private home (not shown). Not counting the time spent in an LAUP program (regardless of whether the hours were funded by LAUP or by other means), children spent an average of 15 hours per week in the care of someone other than their parents.

We compared the zip codes of program locations and families' homes to obtain a rough measure of the distances families traveled to attend the program. We found that 52 percent of families lived in the same zip code as the program their child attended, and fewer than 5 percent lived in a zip code more than 10 miles from the program's zip code (not shown).

Most families had resources to support their own and their children's physical and mental health (Table III.18). Almost all children (93 percent) had health insurance, with the majority covered by a public/government plan. Virtually all children (97 percent) had

undergone a medical checkup in the past year, and three-quarters saw a dentist during that time. The majority of parents (73 percent) reported few depressive symptoms, suggesting positive mental health. By way of comparison, 48 percent of parents in FACES were classified as not depressed (ACF 2006). About half (53 percent) of parents described their physical health as excellent or very good. Parents also reported on their sources of social support when faced with six emotional, financial, and parenting problems. Seventy-six percent of parents reported sources of support for all six types of problems. About 10 percent of parents reported receiving community services, such as English as a Second Language classes, mental health services, and help with medical care (not shown in tables). A minority of parents and children were exposed to crime or violence in the home or neighborhood within the past year. About 13 percent of parents had witnessed a nonviolent crime in the past year, and a similar percentage had witnessed a violent crime. Five percent of children had witnessed domestic violence in the past year, and 2 percent of children were victims of violence in the home during that period.

When we examined family characteristics by language group, we found notable differences apart from the expected differences in race/ethnicity and immigration status. Parent education in the Spanish-only and Spanish-primarily groups was markedly lower compared to the English and other-language groups, as were rates of mothers' full-time employment (Tables III.19 and III.20, Appendix C). Across language groups, most children were born in the United States, and there were few differences in the number of children and adults in the household (Table III.21, Appendix C). However, children from Spanish-only and Spanish-primarily groups were more likely to live with both parents than the English group, although rates of marriage were similar (except for the other-language group, which was higher than the other three). Rates of poverty were higher in the two Spanish-language groups relative to the English and other-language groups, with 72 percent of the Spanish-only group living below the poverty line and 57 percent of the Spanish-primarily group doing so, compared to 22 percent of the English group and 27 percent of the other-language group. Notable differences in home environment by language group were in frequency of reading to the child and number of books in the home (Parents in the Spanish-only and Spanish-primarily groups read at lower rates than English and other-language groups) (Table III.22, Appendix C).

Children across language groups differed in the amount of time spent in LAUP per week. Children in the Spanish-only and Spanish-primarily groups attended fewer hours per week than children in the English or other-language groups. According to parent report, 22 percent of children in the Spanish-only group and 16 percent in the Spanish-primarily group attended preschool for 12 or fewer hours per week, while this was reported for only 7 percent of the children in the English group and 8 percent of the children in the other-language group (not shown). Although LAUP funded only half-day programs, some of the centers offered full-day programs; the children involved in these programs differed by language group. More parents of children in the English-language group (43 percent) and the other-language group (37 percent) reported that their children attended preschool for more than 24 hours a week; this compared with 18 percent for the Spanish-only group and 28 percent for the Spanish-primarily group. To investigate whether poverty might be related to time in LAUP programs, we constructed a set of contingency tables and calculated chi-

square statistics. We found a nonsignificant chi-square between a binary measure of poverty level and a categorical measure of hours per week in LAUP. Examining the relationship of poverty to language group was significant; however, we did not find a systematic difference by poverty status in attendance in LAUP programs.

Most children had health insurance, regardless of language group, but the English and the other-language groups were most likely to have private insurance, while the Spanish-only and Spanish-primarily groups were most likely to have public or government plans (Table III.23, Appendix C).

2. Children's Development from Fall 2007 to Spring 2008

In Table III.24, we present scores on each of the language and literacy and mathematics measures for the population of LAUP children. In Table III.25, we array similar information for the total scores for social-emotional development and approaches to learning and, in Table III.26, for physical and motor development. Table III.27 presents information on language and literacy development fall to spring for the children scoring in the lowest and highest quartiles on these measures in the fall; Table III.28 presents social-emotional development for the lowest and highest quartile groups. In this section, we present our analyses of children's progress from fall to spring on all the school readiness measures. In analyses of the full sample of LAUP children, all the fall-spring changes were statistically significant at $p < .001$. In analyses of change within the four language groups, a few changes were not significant, and we note these in the text.

Language and Literacy Development. Children in LAUP programs made progress over the course of the preschool year in language and literacy development, including performance in letter naming, receptive vocabulary, expressive vocabulary, and early writing skills. We used two subscales of the Preschool Language Assessment Screener (Pre-LAS) in English and Spanish as part of an extensive language routing procedure (for more details, see Appendix B). Children who passed the Pre-LAS in English went on to receive instructions in the next assessment in English. We found that children increased their English proficiency from fall to spring; that is, a greater proportion passed the English Pre-LAS in the spring than in the fall (78.6 versus 61.9 percent).

Children increased the number of letters of the English alphabet that they correctly named from fall to spring. The assessment of letter naming used two different forms with 30 letters that included all 26 letters of the English alphabet in either upper or lower case, and 4 additional letters in the opposite case to aid in creating an overlap between forms. Letters that look the same in upper and lower case (for example, c, s, o) were also used to support the overlap in the scale construction. After analyzing the forms together, the scores were scaled to range from 0 to 52 (26 upper- and 26 lower-case letters). Children scored a mean of 17 in the fall (Figure III.4).[9] By spring, the number increased to 30 of the total of 52. The

---

[9] The scores are based on the IRT probabilities scaled to a range of 0-52 to represent the number of "letters." We refer to "letters" in quotation marks, however, as the scores are based on the IRT analyses and are not literally the number of letters named.

raw score for the fall was 8.5 out of 30 letters, but the median was 4, indicating that most children knew very few letters. In the spring, the raw score mean was 18.1 out of 30 letters. In the spring, a little more than 5 percent of children could identify all letters of the alphabet.

In addition to examining the overall means (representing the average of the program), we looked to see if children who entered with stronger and weaker skills made similar rates of progress over the course of the year.[10] In letter naming, the bottom quartile of children scored 2.5 and 19.8 "letters" correct, on average, in the fall and the spring, respectively, while the top quartile of children scored 35.3 and 42.0. Children in the lower quartile increased 18 scale score points ("letters") in the spring, but they did not reach the level that the top-quartile children had achieved in the fall.

Figure III.4.   Number of Letters Named, Fall and Spring
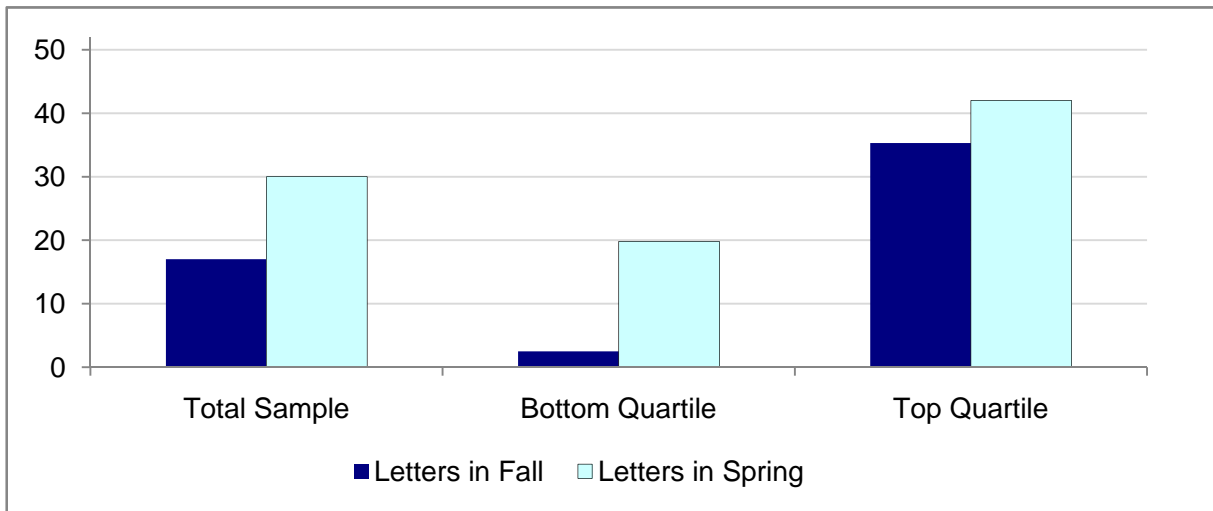
The average standard score on the children's skills in early writing (Woodcock-Johnson III Spelling subtest) in English increased from fall to spring. Standard scores compare each child's performance to that of a national sample of children of the same age and of diverse income levels. The scores are normed to have a mean of 100 and a standard deviation of 15. An increase in the standardized score for children from fall to spring indicates that children increased their early writing skills by more than the average for children of the same age nationally. In fact, children tested in English at both times performed at the national mean in the fall (101.1); in the spring, the average score was about one-half a standard deviation above the national mean, at 108.1 (Table III.24).[11] Children who were administered the early

---

[10] Children in the lowest quartile differed in a number of ways in addition to scoring low on the particular measure. Children who scored in the bottom quartile on two or more of the measures different from the rest of the children in that they had mothers who were less likely to have completed high school, were more likely to be Spanish speaking, were less likely to read to the child at least three days a week, and were less likely to be married. Low-scoring children were also more likely to live in an AGSN zip code.

[11] For comparability of the sample over time, we include in the tables the fall and spring scores of children who took the same version of the Spelling subtest over time (either the English Woodcock-Johnson

writing assessment in Spanish (Woodcock-Muñoz III "Spelling") did not, however, perform as strongly as their peers. Some children who took the Spanish version of the subtest in the fall were able to take the English version in the spring; therefore, the spring estimates on the Spanish version presented in Tables III.24, III.27 and III.29 reflect the performance of a smaller sample of children who did not make progress in acquiring English. Children who took the Spanish version of the "Spelling" subtest continued to score below the mean in both the fall and spring, with many children scoring near one standard deviation below the mean. Scoring one standard deviation below the mean is an important threshold for indicating either delayed performance or a child possibly at educational risk.

Although the average standard score on the Woodcock-Johnson III Spelling subtest in English increased for children in both the bottom and top quartiles, the fall-spring increase was much greater for children in the bottom quartile (Figure III.4). In other words, children in the bottom quartile showed greater improvement in their scores (of course, they had more room to grow). Children in the top quartile performed more than one standard deviation above the national mean in the fall (120.5); in the spring, their average score decreased by 2.7 points, to 117.8. In contrast, children in the bottom quartile performed more than one standard deviation below the national mean in the fall (80.5); in the spring, their average score increased by 18.7 points, to 99.2. Thus, the bottom quartile of children made progress relative to same-age peers in early writing, nearly reaching national norms.

We assessed children's receptive vocabulary on the conceptually scored Receptive One Word Picture Vocabulary Tests (ROWPVT; ROWPVT-SBE).[12] Children who were administered the English version of the ROWPVT in the fall obtained an IRT-derived mean score of 49.8 in the fall, and 58.3 in the spring (the spring assessment was administered only in English).[13] Children who took the Spanish bilingual version of the ROWPVT in the fall scored an IRT-derived mean of 47.0. In the spring, those children obtained a mean IRT-derived score of 49.6. Using the sample mean ages in the fall and spring, the IRT-derived ROWPVT scores for the children taking the English version in the fall are equivalent to a standard score of 88 in the fall and 95 in the spring (both relative to an English-speaking national sample), and those who took the Spanish bilingual edition in the fall were equivalent

---

(*continued*)

III, or the Spanish Woodcock-Muñoz III). This approach excludes the 206 children who took the Spanish version in the fall but who were able to take the English version in the spring. The average spring score for this group of children was 105.2.

[12] We calculated IRT scores for the ROWPVT and the EOWPVT that allowed us to compare fall to spring performance for the full analytic sample. Some items in the Spanish-Bilingual Edition are not administered. IRT is able to estimate scores with some missing item-level data. The comparisons of IRT scores to national norms (either English or Spanish, depending on the version administered) are based on the mean equivalent number correct for the mean IRT score. The mean child age was then used to look up the comparable standard score.

[13] The ROWPVT was conceptually scored in Spanish and English in the fall but was administered only in English in the spring. Because the scores are based on different assessments and are not directly comparable across time, we did not include a figure.

to a standard score of 83 and 81.[14] Children's performance on assessments of expressive vocabulary from the fall to the spring reflects an increase in their skills and proficiency in expressive language. However, relative to the gains observed in the nationally representative sample of English-speaking children, the gains obtained by the UPCOS sample were less than expected. The average child's proficiency in expressive vocabulary was 43.3 in the fall and 49.2 in the spring (IRT-derived and conceptually scored at both time points). Using the sample mean ages in the fall and spring, the IRT-derived EOWPVT scores are equivalent to a standard score of 78 in the fall and 72 in the spring (both relative to a nationally representative English-only sample). Children who entered the program with strong skills in expressive vocabulary demonstrated greater proficiency than those who entered the program with weaker skills—IRT-derived scores of 56.2 versus 29.7 in the fall and 59.5 versus 39.5 in the spring—and both groups of children showed progress, with the highest quartile increasing 3.3 points and the lowest quartile increasing nearly 10 points. However, the large increase for the lowest quartile did not bring them to the fall baseline levels of the highest quartile group (Figure III.5).

Figure III.5.   Expressive Vocabulary Fall and Spring
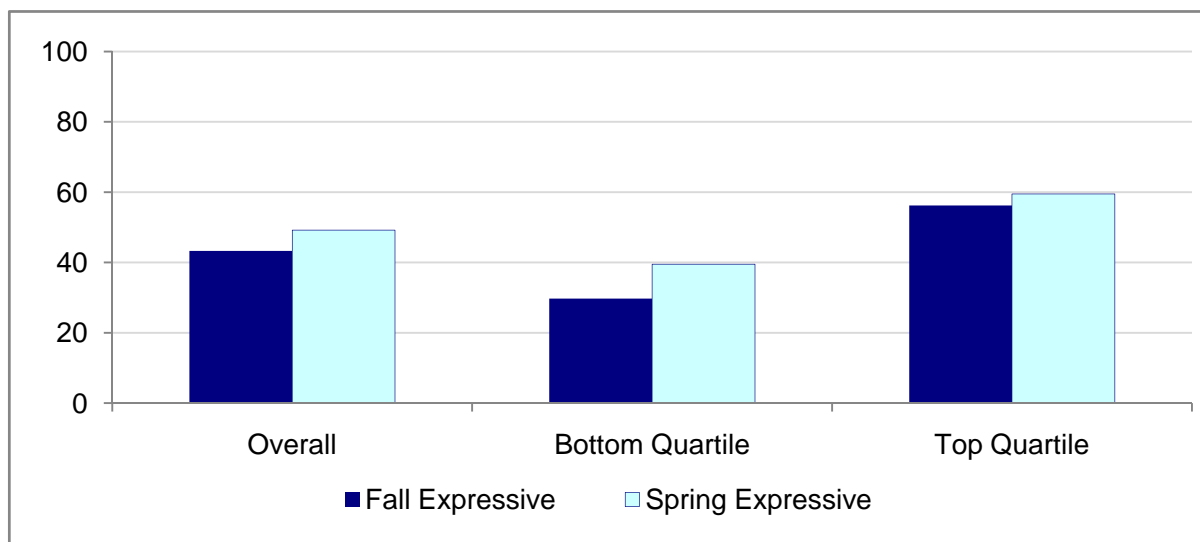
In the area of receptive vocabulary (ROWPVT), children in the lowest quartile who took the English version in the fall scored an IRT-derived average of 35.1 in the fall and gained 16 points, to 51.4, by spring. Fall-to-spring gains among children in the lowest quartile who took the Spanish bilingual edition in the fall showed a different pattern; they entered with a mean score of 28.9 and increased 11 points, to 40.1. Children in the top

---

[14] Standard scores reported in the text are based on English norms for both fall and spring. We cannot present scores for the spring relative to a bilingual sample because we administered the ROWPVT only in English in the spring. However, for this group who took the ROWPVT-SBE in the fall, their equivalent standard score in the fall was 99.4, based on Spanish-bilingual norms.

quartile taking the English version in the fall entered with an average score of 61.5 and gained 4 points by spring (to 65.7). Their counterparts who took the Spanish bilingual edition in the fall scored 62.9, but dropped to 57.5 in the spring. Compared to a nationally representative English-speaking sample, children demonstrated significantly stronger proficiency in receptive vocabulary than in expressive vocabulary.[15]

Mathematics Development. As with other developmental areas, LAUP children's mathematics knowledge and skills improved during the preschool year. The ECLS-B mathematics assessment taps children's number concepts, spatial abilities, and measurement proficiency. In the fall, children correctly answered an average of 8 of 23 mathematics items. At that time, the average child demonstrated proficiencies in such areas as ability to identify shapes, basic counting skills, and understanding of relative amounts. In the spring, the average child correctly answered 12.5 items, and children demonstrated skills such as identifying order (for example, "point to the last bicycle"), pattern matching and continuation, and counting to 10 with one-to-one correspondence.[16]

The amount of improvement during the preschool year was similar for children in the top and bottom quartiles. In the fall, children in the bottom quartile correctly answered an average of 4 of 23 mathematics items and those in the top quartile correctly answered approximately 14 items. By the spring, the average child in the bottom quartile correctly answered almost 10 items and top-quartile children correctly answered an average of 16 items. Thus, children in the bottom quartile improved at a faster rate than children in the top quartile, although their spring scores did not match the fall baseline scores of those in the top quartile (Figure III.6).

Figure III.6.   Number Correct ECLS-B Mathematics Fall and Spring
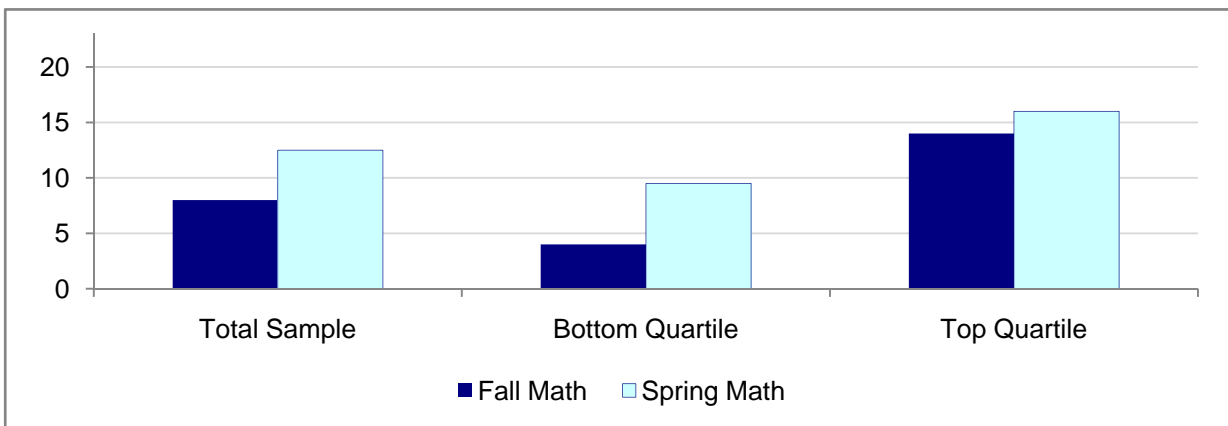
---

Social-Emotional Development and Approaches to Learning. In Table III.25, we present scores on each of the measures of children's prosocial and problem behaviors, based on reports from three separate raters: teachers, parents, and our assessors. We discuss each in turn.

Teacher ratings of children's Social Cooperation, Interaction, and Independence were higher in the spring than in the fall. Standard scores on these subscales of the Preschool and Kindergarten Behavior Scales-2 (PKBS-2) (Merrell 2002) are normed for a mean of 100 and standard deviation of 15; therefore, a standardized score of 100 signifies that the child performed at the average level for children his or her age. Average scores on the Social Cooperation subscale, which measures children's ability to share and to follow classroom rules, increased from 101.1 in the fall to 105.6 in the spring (Figure III.7). In the area of Social Interaction, which measures children's prosocial behavior with peers and in the classroom, children's development increased from 94.9 to 103.4, an average increase of more than 8 points. Similarly, teachers rated children as demonstrating greater Social Independence and Adjustment in the spring than in the fall. Thus, in the fall, teachers rated children's cooperative and independent behavior at about the mean compared to children nationally and then above the mean by the spring. They reported similar increases in children's positive interactions with peers. Teachers did not report marked improvements from the fall to the spring in children's ECLS-K Approaches to Learning. At both the fall and spring assessment points, teachers reported that children demonstrated behaviors such as attention, persistence, and eagerness to learn "sometimes" (mean = 1.9 and 2.2, respectively) which means that although the increase in scores reached statistical significance, the improvement in these aspects of children's behavior were not practically meaningful.[17]

Figure III.7. Preschool Kindergarten Behavior Total Scores, Fall and Spring

---
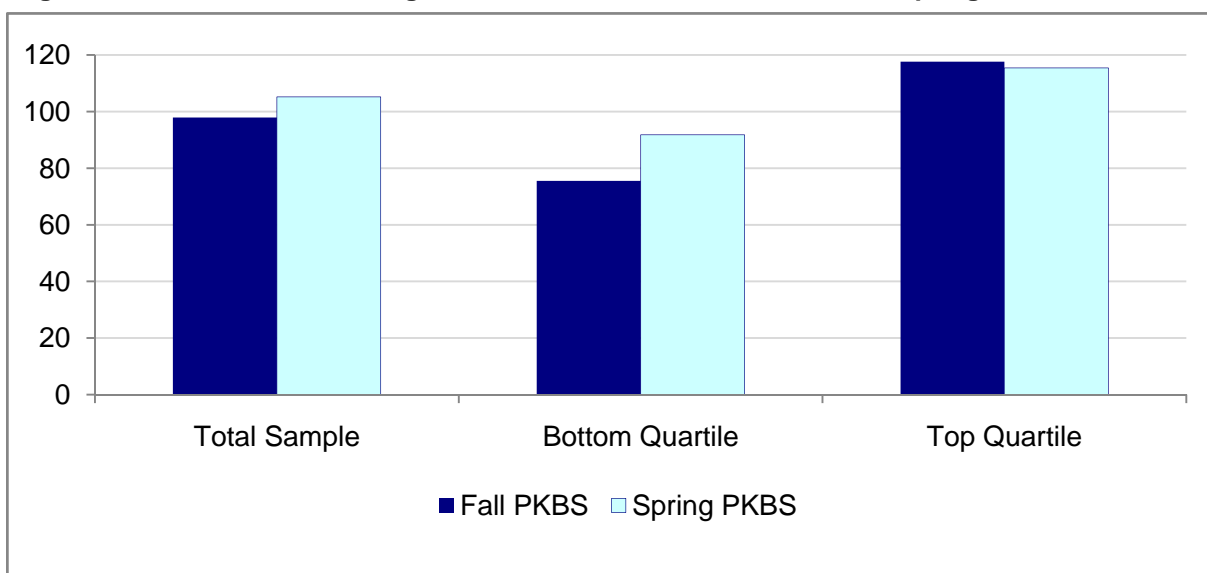
[17] The frequency scale for the teacher-reported approaches to learning was never (0), rarely (1), sometimes (2), or often (3). The scores derived from these data reflect the range of the response scale (0–3).

Among the bottom-quartile children, average scores on the Social Cooperation subscale increased from 77.6 in the fall to 91.3 in the spring (Table III.28). In the area of Social Interaction, bottom-quartile children's development increased from 71.8 to 90.3, an average increase of nearly 19 points. Similarly, teachers rated these children as demonstrating greater Social Independence and Adjustment in the spring (94.1) than in the fall (78.2). Although they began more than one standard deviation below the mean in the fall on all three subscales of the PKBS (Social Cooperation, Social Interaction, and Social Independence), by the spring, teacher ratings of the behavior of children in the bottom quartile gained between 13 and 19 points. However, compared to children nationally, children in the bottom quartile remained about two-thirds of a standard deviation below national norms. For the top quartile, teachers' ratings fell slightly from fall to spring in Social Cooperation (119.1 to 116.2), Social Interaction (115.9 to 113.4), and Social Independence (115.5 to 111.9). However, in both fall and spring, teachers rated children in the top quartile at or near one standard deviation above the national mean on these scales. For the top quartile group, we observed little fall-to-spring change in teachers' ratings of Positive Approaches to Learning. At both the fall and spring assessment points, teachers reported that top-quartile children "often" demonstrated behaviors such as attention, persistence, and eagerness to learn. Teachers rated bottom-quartile children as "rarely" exhibiting these behaviors in the fall; by spring, teachers rated these children between "sometimes" and "often."

As measured by the Social Skills Rating System (SSRS) (Gresham and Elliott 1990), teachers tended to rate children favorably on Problem Behaviors in the fall and in the spring. Average fall scores were about one-third of a standard deviation below the norm, indicating less frequent problem behaviors, thus indicating low levels of problem behavior compared with children nationally.

Similarly, we observed little difference between the average fall and spring SSRS scores for the bottom-quartile children, as teachers did not change their already favorable ratings of children's problem behaviors in the spring compared with their fall ratings. In both the fall and spring, average scores for children with the fewest problem behaviors were about one standard deviation below the national mean. Teachers rated children with the most problem behaviors (the top quartile) about one standard deviation above the mean nationally in the fall. By the spring, the frequency of problem behaviors, although higher than the national mean, had reduced by two-thirds of a standard deviation.

Parents did not report discernible improvements from fall to spring in children's social skills and positive approaches to learning (based on the adapted PKBS), such as paying attention and persisting in difficult tasks. In the fall, parents rated children's Social Cooperation/Approaches to Learning at 2.9 and Social Interaction/Independence at 3.1 (equivalent to "sometimes"). By the spring, parent ratings on both subscales remained virtually unchanged (0.1 points on each scale). Behavior problems (Internalizing and

Externalizing) were rated by parents as occurring "rarely" (2.0 and 2.2 respectively) and remained unchanged by the spring.[18]

Assessors used the Leiter-R Examiner Rating Scale (Roid and Miller 1997) to rate aspects of children's behavior observed during the direct assessment. The ratings provide additional information on children's social skills and behaviors. Assessors rated the extent to which children attended to the assessment tasks, their activity level during the assessment, and their level of sociability during the assessment. Assessors rated children favorably on the Leiter-R subscales in the fall and the spring, respectively: 7.8 and 8.9 on Attention, 8.2 and 8.9 on Activity Level, and 8.2 and 9.0 on Sociability. These scales are all coded in the positive direction and are normed to have a mean of 10 and a standard deviation of 2.[19] Children moved toward national norms during the preschool year; assessors rated children as showing higher levels of Attention, more appropriate Activity Levels, and greater Sociability in the spring compared with the fall ratings. Fall-spring improvements corroborate teacher ratings of improved social skills during the year. The gains may reflect greater comfort and familiarity with the assessment process by the spring.

Finally, we administered direct assessment measures of executive functioning. These executive functioning tasks required children to inhibit a natural response and keep in mind the rule that they were to follow. In the Pencil Tapping task, children were asked to inhibit the natural response to imitate the adult assessor exactly (or to tap repeatedly) and instead to keep in mind that the rule was to do the opposite of what the assessor did; that is, if the assessor tapped twice, the child was to tap once and vice versa. In the fall, 43 percent of the time, or less than chance, children effectively inhibited their initial impulse and responded correctly on the Pencil Tapping task. In the spring, children effectively remembered the rule and inhibited the initial response across a greater percentage (67 percent) of trials, indicating improved self-regulation and impulse control.[20] By way of comparison, a smaller percentage of children responded correctly across trials (59 percent) in a sample of low-income 3- and 4-year-olds in Head Start (Smith-Donald et al. 2007).

The Walk-a-Line Slowly task (Murray and Kochanska 2002; Smith-Donald et al. 2007) also was administered to children to measure their ability to inhibit the impulse to speed up their rate of walking. Assessors asked children to walk along a six-foot line and then to walk the line again as slowly as possible. The difference between the first attempt and the slow attempt was the child's score. In the fall, children in LAUP programs took an average of 1.2 seconds longer, with a range from a minimum score of -11.3 seconds to a maximum of

---

[18] The frequency scale for the parent-reported PKBS-2 scales was never (1), rarely (2), sometimes (3), or often (4). The IRT scores derived from these data were scaled to the range of the response scale (1–4).

[19] We do not show the fall-spring scores for the Leiter quartile groups.

[20] We also do not show the graphs for the Pencil Tapping quartile groups. The lowest quartile increased from 6.25 percent in the fall to 18.75 in the spring; the upper quartile group went from 87.5 to 100 percent, thus reaching the ceiling in the spring.

140.9 seconds.[21] The negative minimum score indicates that some children were not able to control the impulse to go faster the second time. Children took an average of 2.3 seconds longer in the spring. When used in a study in Chicago Head Start classrooms (Smith-Donald et al. 2007), children took approximately 2 seconds longer when asked to walk the line slowly.

Physical Health and Development. We also used the Walk-a-Line Slowly task as an indicator of children's balance or gross motor skills. On the child's first attempt at the task, examiners observed whether the child was able to stay on the line. In the fall, 71 percent of the children were able to stay on the line during their first attempt "almost all of the time," and 81 percent were able to do so in the spring (Table III.26).

Children's height, weight, and the ratio of the two (body mass index [BMI]) are a reflection of children's general health status and well-being. The Institute of Medicine of the National Academies considers children to be overweight when their gender-specific BMI-for-age is between the 85th and 94th percentile and obese when their BMI is at or above the 95th percentile for their age and gender (Centers for Disease Control and Prevention 2008; Whitaker 2005). Using this criterion, 15 percent of LAUP children were overweight and 22 percent were obese in the fall. The percentages decreased slightly to 14 percent overweight and 20 percent obese in the spring. Although we have not found comparable data for 4-year-olds in LA County or the state, the County of Los Angeles Department of Public Health (2008) reported that 22.9 percent of fifth, seventh, and ninth graders were obese. The rate of obesity in similarly aged children in the FACES study was 16 percent (West et al. 2008). Regardless of the comparisons, the finding that one-third of children are overweight or obese suggests there are important issues related to obesity in the LAUP population.

Finally, reports of children's general health did not change from the fall to the spring. On a scale of 1 ("poor") to 5 ("excellent"), 96 percent of parents rated their children's overall health as excellent or very good at both time points. For comparison, parents rated 81 percent of children in Head Start programs as being in very good or excellent health (ACF 2006). These high favorable parent ratings at the beginning and end of the preschool year indicate low levels of perceived serious health problems among children served by LAUP.

3. Development of Children Within Language Subgroups

Given that many children in LAUP are speaking and learning more than one language, typically Spanish and English, we examined fall-to-spring development within the four main language groups that we classified children into using parent's reports at baseline: (1) English, (2) Spanish-only, (3) Spanish-primarily, and (4) other-language. In this section, we describe how these groups of children developed in their language and literacy skills,

---

[21] Because of the skewed distribution, we did not compute quartile-group weighted percentages for Walk-a-Line Slowly.

mathematics skills, social-emotional development, approaches to learning, and physical health and motor development.

Language and Literacy Development. In Table III.29, we show the performance of the four language groups on measures of language and literacy. The table may be used to compare the fall and the spring scores of children in each language group and to see how the language groups differed at each time point.[22] In the area of receptive English vocabulary, the skills and proficiency of children in English households improved from a mean of 51.5 in the fall to 59.5 in the spring, for an 8-point gain.[23] Children in other-language households, however, demonstrated the greatest increase in receptive English vocabulary from the fall (37.8) to the spring (50.1). In contrast, for children in Spanish-only and Spanish-primarily groups, receptive vocabulary in English was less highly developed than for the other groups.

Children's performance on conceptually scored assessments of expressive vocabulary from the fall to the spring reflected similar progress, with children in other-language households, who were assessed only in English at both time points, showing the largest average increase from the fall to the spring (8.4-point gain), followed by those in English (6.4-point gain), Spanish-primarily (5.5-point gain), and Spanish-only households (4.2-point gain) (Figure III.8).

With respect to the Rapid Letter Naming measure (conceptually scored with correct responses accepted in both English and Spanish), children in Spanish-only households showed the greatest gain in letter knowledge from the fall to the spring, a 15-point increase, but they also scored lower than the other groups on average at both times (10.3 and 25.0) (Figure III.9). Children in other-language households could correctly identify the most letters in the fall (24.8) and the spring (36.8), followed by children in English-language households (20.5 and 32.7) and children in Spanish-primarily homes (13.0 and 26.5).[24]

---

[22] In Section III.D, we report on analyses that examine the relationships of child and family characteristics (including language groups) to child developmental outcomes. In this section, we report only descriptive information.

[23] The ROWPVT was conceptually scored in Spanish and English in the fall but, in the spring, was administered in English only; thus, the fall and spring scores are from measures of different, though related, constructs. The more limited gains for Spanish-only and Spanish-primarily children may reflect differences in the language of administration at the two assessment points. In other words, children who were routed to Spanish in the spring were still assessed in English on the ROWPVT. We did not graph this outcome because of the difference in construct measurement at each time point.

[24] Scores are based on IRT analysis and are not literally the number of letters named.

Figure III.8. Conceptually Scored Expressive Vocabulary Fall and Spring, by Language Group
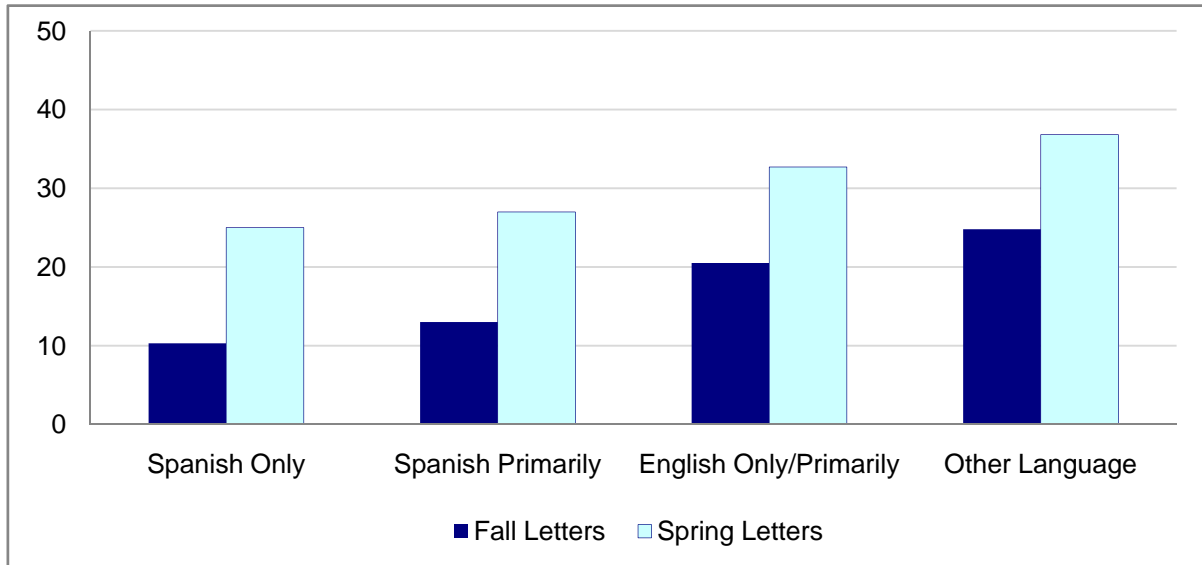
Figure III.9. Letters Named Correctly Fall and Spring, by Language
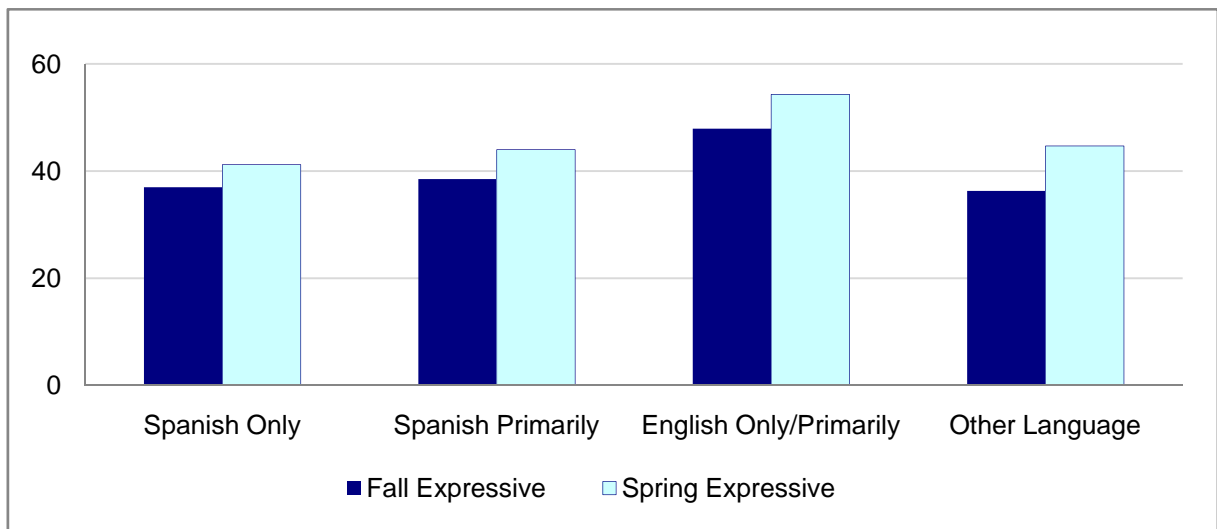
Mathematics Development. The mathematics skills and knowledge of all children improved during the preschool year. English language and other-language children demonstrated the highest-level skills and knowledge in mathematics in both the fall and spring (Table III.29 and Figure III.10). Children from other-language households increased their scores from 491 to 504, correctly answering an average of 8 of 23 mathematics items on the ECLS-B mathematics assessment in the fall and approximately 12 items in the spring;

scores for English-language children were similar (9 and 11 items for fall and spring, respectively).[25] The means for children from Spanish-only and Spanish-primarily households increased by 10 scale score points, and the children's average scores did not quite reach the normative mean difficulty for the fall preschoolers in ECLS-B. Spanish-only children correctly answered a mean of 7 items in the fall and 9 in the spring. Similarly, the mean number of correct items increased from 7 to 10 from the fall to the spring assessments for the Spanish-primarily group. Across language groups, children had fairly similar average gains in mathematics from the fall to the spring.[26]

Figure III.10. Number Correct ECLS-B Mathematics Fall and Spring, by Language Group
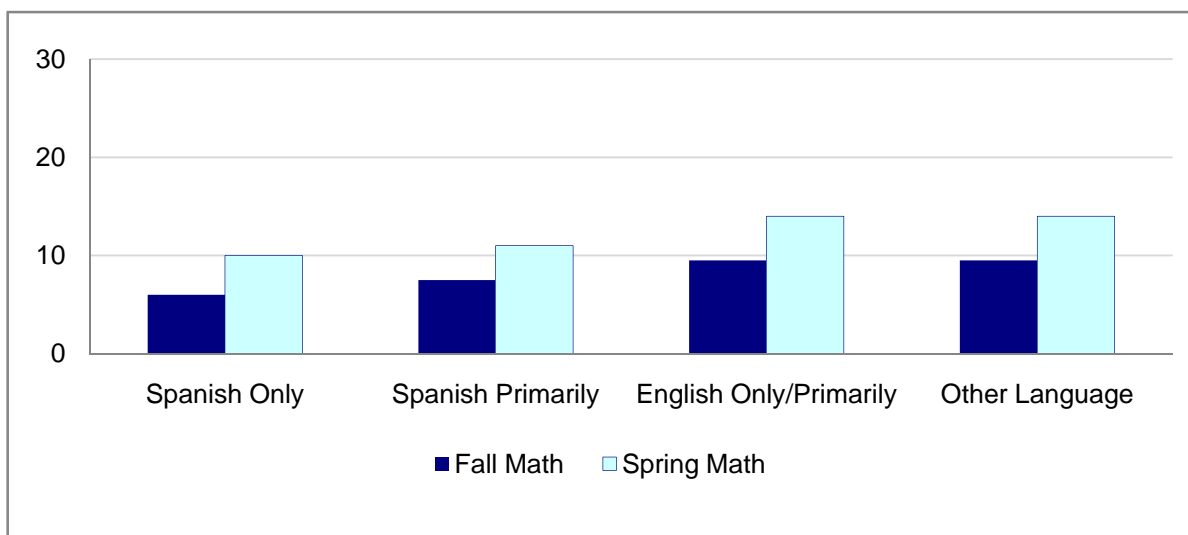
Social-Emotional Development and Approaches to Learning. Across language groups, lead teacher ratings of children's overall Social Skills, Social Cooperation, Social Interaction, and Social Independence were higher in the spring than in the fall (Table III.30). On the PKBS in the fall, teachers rated children at or just under the mean compared to children nationally on the Social Cooperation subscale, and they rated all children above the national mean in the spring. Children experienced the greatest increases or gains relative to peers in the area of Social Interaction, with fall-to-spring standard score gains for each of the language groups ranging from 7.5 to 10.3 points. Children from English-speaking households performed close to the mean in Social Interaction in the fall; however, by the spring, all children scored at or above the national mean, except for the Spanish-only group, which fell just under the mean (98.6). Similarly, teachers rated children as demonstrating

---

[25] The table presents W scores for the ECLS-B. We anchored the item difficulties on the estimates obtained from the ECLS-B, then arithmetically transformed the scores (Woodcock 1999, p. 111) so that the mean difficulty of the items was 500. In the text, we also present the number of correct items to show concretely improvements in children's performance over time.

[26] The fall-spring change was not statistically significant for the Spanish-only and Spanish-primarily groups.

greater Social Independence and Adjustment in the spring than in the fall, and all children scored above the national mean by the spring. Figure III.11 compares overall Positive Behavior standard scores across language groups. Children from Spanish-only and other-language households made the greatest gains relative to peers in overall Social skills, with average scores increasing 8.5 points between the fall and the spring.

All children demonstrated greater self-regulation and impulse control in the spring than in the fall. However, a larger percentage of children from English- (70 percent) and other-language households (80 percent) effectively inhibited their initial impulse and responded correctly on the Pencil Tapping task in the spring. More than half (58 percent) of the time, children from Spanish-only households effectively remembered the rule and inhibited their initial response across trials in the spring. On the Walk-a-Line Slowly task, all children, regardless of household language, took longer on the second "slow" trial in the spring than in the fall. In other words, they demonstrated greater self-regulation in the spring. At that time, children from other-language households took an average of 4.0 seconds longer, followed by children from English-speaking (3.1 seconds), Spanish-primarily (1.4 seconds), and Spanish-only (0.9 seconds) households. In comparison, in the fall, children from other-language households took an average of 2.1 seconds longer, followed by children from English-language (1.7 seconds), Spanish-primarily (0.5 seconds), and Spanish-only (0.1 seconds) households. A study of Chicago Head Start programs found that children were able to walk 2 seconds slower on their second attempt (Smith-Donald et al. 2007).

Figure III.11. Preschool Kindergarten Behavior Scores Fall and Spring, by Language Group
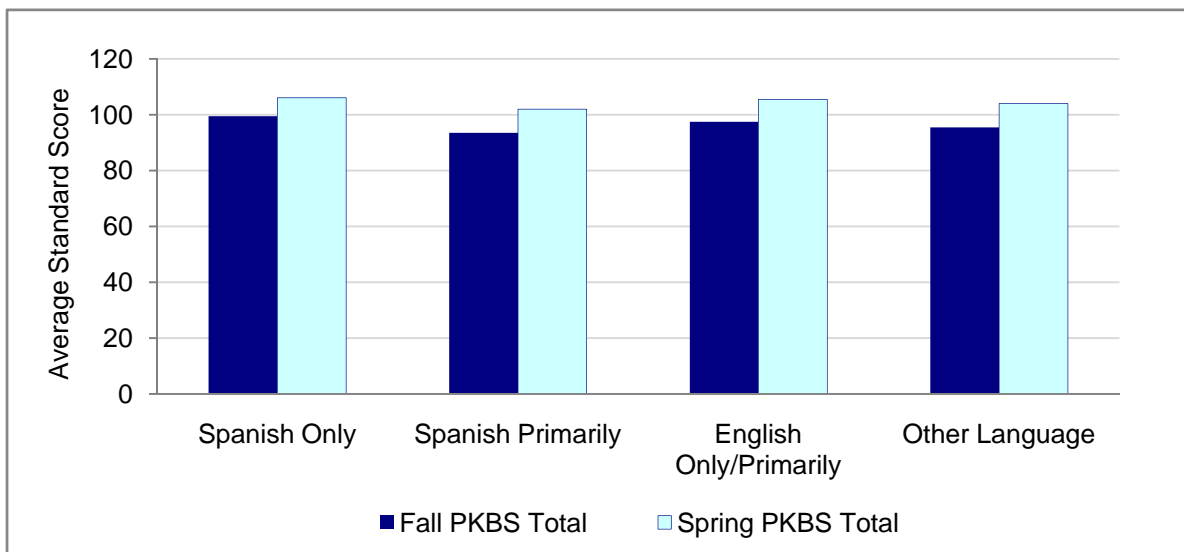
Physical Health and Motor Development. All groups of children demonstrated improvement in balance and gross motor skills from fall to spring (Table III.31). Across language groups in the spring, a similar percentage of children were able to stay on the line on their first attempt "almost all of the time." The change in the percentage of children successful in staying on the line was greatest for children from other-language households

(16.4 percentage point increase), followed by children from English-language (11.5 percentage point increase), Spanish-only (7.8 percentage point increase), and Spanish-primarily (5.3 percentage point increase) households.

Across children, the rate of obesity (defined as BMI ≥ 95th percentile in CDC's national profiles) was high in both the fall and the spring, but the incidence of obesity decreased from fall to spring (Figure III.12). Compared to other children, a larger percentage of Spanish-only (24.9 and 22.9 percent) and Spanish-primarily (25.1 and 22.9 percent) children were obese at the two assessments. Rates of overweight (85th to 94th percentile for age and gender) were fairly constant across language groups, with 14 percent of children classified as overweight in the English group, 15 percent in the Spanish-only, 16 percent in Spanish-primarily, and 11 percent in other-language groups. Finally, reports of children's general health did not change markedly from fall to spring, with parents reporting that nearly all children were in excellent or very good health at both time points.

Figure III.12. Percentage Obese (BMI > 95th Percentile) Fall and Spring, by Language Group
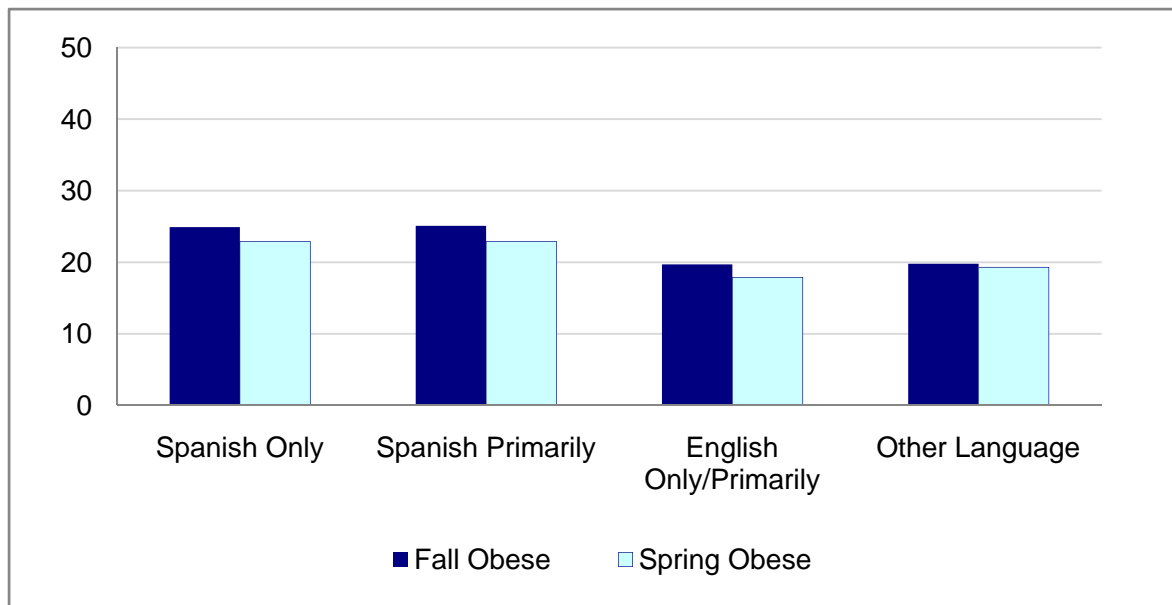
We also examined rates of obesity by racial/ethnic group and found rates were similar across racial/ethnic groups in the fall, except for much lower rates for Asians. By the spring, white children showed the largest decrease in the rate of obesity compared to the other groups, and Asians had almost doubled their rate. In the spring, children from African American, and Latino groups looked similar in their rates of obesity, whites had decreased somewhat, and Asians had increased markedly, although they continued to show much lower rates of obesity overall (Figure III.13).

Figure III.13. Percentage Obese (BMI > 95th Percentile) Fall and Spring, by Racial/Ethnic Group
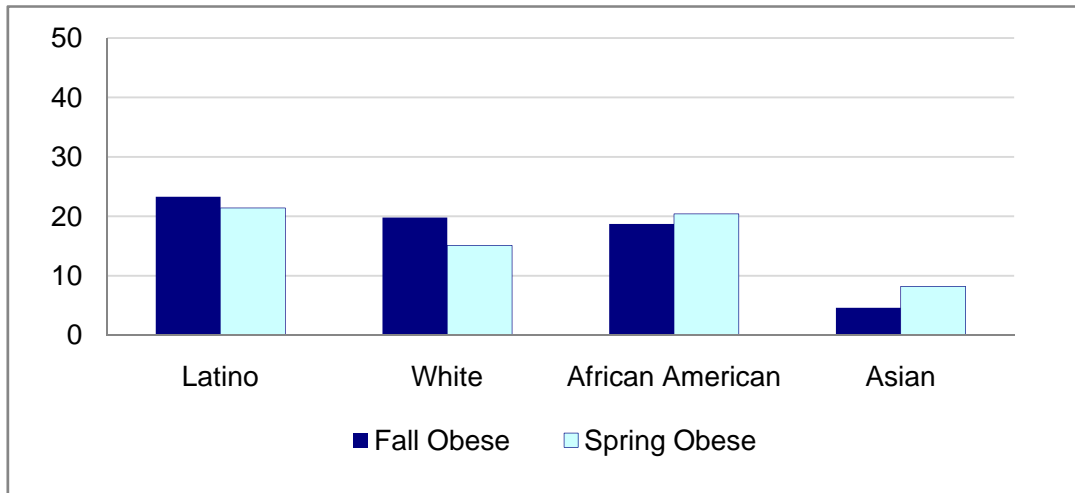
4.   Summary of Key Findings

Family Challenges and Resources. Children and families enrolled in LAUP programs faced a number of challenges but also had strengths and resources to draw upon.

- Children were diverse in race/ethnicity, language, and culture. Children were primarily Latino (75 percent), with approximately equal proportions who were white (7 percent), African American (8 percent), and Asian (7 percent).

- Although nearly all children were born in the United States, a substantial minority had parents born outside the country (63 percent of mothers and 68 percent of fathers were born elsewhere).

- About 47 percent of mothers had at least some college education; more than one-third lacked a high school diploma.

- Employment was high, with 89 percent of fathers and 43 percent of mothers employed full-time.

- Nearly three-quarters of children lived with both parents, many of whom were married.

- Families spanned a range of income levels, with nearly 40 percent under the federal poverty level.

- Family resources were generally good, including low rates of maternal depressive symptoms (73 were not depressed) and high rates of health

insurance (93 percent had coverage) and receipt of medical care (97 percent had a regular checkup within the past year).

Fall-Spring Gains in School Readiness. Children's development in school readiness skills was generally below national norms, on average, but improved between the fall and spring. Key findings included:

- A greater percentage of children were able to take the assessments in English in the spring than in the fall (79 percent compared to 62 percent).

- Language and literacy progressed over the program year. For example, children increased by 13 scale score points ("letters") on the letter naming task in the spring compared to the fall.

- Children's average social-emotional development showed improvement (according to teacher reports) on the Social Cooperation, Social Interaction, and Social Independence and Adjustment subscales of the PKBS, with children scoring above national means by spring. An exception was Approaches to Learning, on which children did not make substantial gains, according to their teachers' ratings.

- Despite absolute gains made in expressive vocabulary, LAUP children made less than expected growth relative to national norms; they achieved equivalent to a standard score of 72 in the spring.

- There were few fall-spring changes in Problem Behaviors as rated by teachers on the SSRS: children's problem behavior was generally well below national norms at both time points (mean spring scores were one-third of a standard deviation below national norms).

- Assessor ratings of children on the Leiter-R Examiner Rating Scale showed improvement between fall and spring on Attention, Activity, and Sociability; these corroborated teacher ratings of improved social skills during the year.

- Children's executive functioning improved as measured by the Pencil Tapping and Walk-a-Line Slowly tasks. By the spring, children exhibited improved self-regulation and impulse control.

- Children generally were in excellent or very good health according to parent reports, although about one-third would be considered overweight or obese in the spring.

High- and Low-Performing Children. Children who entered the program with lower skill levels made greater progress than children who entered at higher skill levels. However, these gains typically did not bring low-performing children up to national norms by the time of the spring assessments, and never brought them to the levels that children in the highest

quartile had in the fall. Children with the highest skills either made small rates of progress or showed slight declines by the spring, possibly due to regression to the mean.

- For the "Spelling" test, children in the lowest quartile who took the Spanish version continued to score more than one standard deviation below the national mean, indicating possible educational risk. In contrast, children in the lowest quartile who took the English version scored also more than one standard deviation below the mean in the fall but scored nearly at the mean by the spring.

- Improvement in expressive vocabulary was about three times as large for the lowest quartile, but average scores in the spring did not reach the fall means for the highest quartile.

- Similarly, receptive vocabulary for children in the lowest quartile showed relatively large gains compared to children in the highest quartile, but did not reach the fall levels for those children (the largest gains occurred among those who took the English version in the fall).

- Children's mathematics fall-spring development was greater for the bottom quartile compared to the top, but as in other domains, the lowest-quartile children did not reach the level that the highest quartile group had attained in the fall.

- Children in the highest quartile did not show any improvement in social skills from fall to spring, and in fact were rated as having declined slightly; however, they remained at or near one standard deviation above the national mean on Social Cooperation, Social Interaction, and Social Independence. Children in the bottom quartile improved quite a bit more, and by the spring had nearly reached national norms.

Performance Within Language Groups. We saw persistent differences in school readiness outcomes by language group:

- Children in the Spanish-only and Spanish-primarily groups generally had greater difficulties, particularly in the areas of receptive vocabulary and early writing.

- Children in the Spanish-primarily group made greater gains than other groups in letter naming, although they lagged behind their peers in the English-language group in absolute terms. Children in the Spanish-only group began at the lowest level in the fall and made relatively smaller gains by the spring, compared to other groups.

- Children's social-emotional development did not differ by language group, but executive functioning did. Although all children improved in executive functioning by the spring, children in the other-language and English-language

groups performed better on the Pencil Tapping and Walk-a-Line tasks than children in Spanish-only and Spanish-primarily groups.

- Children's health ratings did not differ by language group; however, rates of obesity were somewhat higher for the Spanish-only and Spanish-primarily groups.

C. PARENT INVOLVEMENT IN LAUP PROGRAMS

1. Parent Involvement from the Parents' Perspective

In consultation with First 5 LA and LAUP, we developed a comprehensive list of 16 parent involvement activities to ask about in the winter parent interview (not all programs may have offered each activity, however). During the preschool year, parents reported that they participated with varying frequency in the range of LAUP program activities (Table III.32). Children's families reported engaging in five activities most frequently, that is, "several times a year," "about once a month," or "at least once a week": (1) attendance at parent-teacher conferences (50 percent of parents reported), (2) participation in at-home activities suggested by the LAUP program (36 percent), (3) attendance at a school or class event (32 percent), (4) volunteering in the classroom (30 percent), and (5) participation in social activities (29 percent). In contrast, 75 percent or more of the parents reported never participating in 6 of the 16 activities: (1) attending any other workshops sponsored by LAUP (91 percent), (2) attending adult education classes (79 percent), (3) attending an LAUP expo and conference (78 percent), (4) serving on a committee or volunteering outside of class (77 percent), (5) participating in a parent support group (76 percent), and (6) attending a kindergarten transition workshop (75 percent).

Parents reported a variety of barriers to program involvement. The most common barrier was their work, school, or training schedule (cited by 43 percent of parents). One-fifth of parents also indicated that the need for child care (23 percent) and other commitments (20 percent) affected their ability to participate in program activities. About 10 percent of parents reported that they were not sure how to participate in the program or that the program did not provide ample participation opportunities. Only six percent said that staff at the program did not speak their language. A similar percentage indicated that they did not know other parents, which they felt interfered with their involvement (not shown).

2. How Teachers Involved Parents

Teachers reported on the family involvement opportunities that they offered during the preschool year and how they tried to engage families (Table III.33).[27] Teachers reported that

---

[27] The teacher and parent reports do not provide comparable information. Teachers' reports of family involvement practices are not specific to sampled children's families. Rather, teachers reported on their practices with parents in the program in general, and the parents in our sample reported on the activities in which they themselves participated.

a number of the practices we asked about occurred quite frequently. We found that five teacher activities that would encourage parent involvement occurred for at least 49 percent of the teachers one or more times a week: (1) 72 percent of children had teachers who reported that they encouraged a parent to ask his or her child about his or her day, (2) 66 percent of children's teachers said they encouraged parents to read to the child, (3) 64 percent contacted a parent if the child did something well, (4) 58 percent asked a parent to visit the classroom, and (5) 49 percent assigned at-home activities to promote parent-child interaction.

A number of activities occurred less often, but still a sizable proportion of children had teachers who also reported that one or more times a week they (1) involved a parent as a classroom volunteer (41 percent), (2) provided at-home activities for a parent to help improve the child's skills (38 percent), (3) told a parent about the skills the child needed for kindergarten (35 percent), (4) asked a parent to take the child to library or community events (33 percent), (5) contacted a parent if the child had problems (33 percent), and (6) gave a parent ideas to help the parent become an effective advocate for the child (33 percent). Home visits are not an LAUP requirement and teachers did not commonly conduct home visits—only 15 percent of children had a teacher who had made a home visit by the spring. On the other hand, 83 percent of children had teachers who held a conference with parents once or twice during the year and nearly 10 percent did so more frequently.

Parents were very satisfied with their child's experiences in the LAUP program during the year (not shown). Most parents were very satisfied with how the program helped their child develop English-language skills (88 percent), maintained a safe program (86 percent), prepared the child for kindergarten (84 percent), helped the child grow and develop (83 percent), supported and respected the family's culture and background (81 percent), and identified and provided services for the child (76 percent). Smaller but sizable proportions of parents were very satisfied with the program's openness to their ideas and participation (71 percent) and the program's efforts to help children's development in their home language (64 percent). We examined satisfaction reports by race/ethnicity and found that white parents were typically the most satisfied with the program. For most items we asked about, a larger proportion of white parents indicated that they were "very satisfied." Unlike other parents, white parents never reported that they were "very dissatisfied" with the program. The smallest percentage of parents reporting that they were very satisfied with the program's openness to their ideas and participation were African American (62 percent) and Asian (65 percent). African American parents were also the least likely to indicate that they were very satisfied with the program's identification and provision of services for the child (63 percent).

D. RELATIONSHIPS OF CHILD AND FAMILY CHARACTERISTICS (AND PROGRAM COMPOSITION) TO CHILDREN'S SCHOOL READINESS OUTCOMES

In this section, we describe the relationships between children's development in school readiness outcomes and a variety of child and family characteristics and measures of program composition. We estimated multivariate HLM models of the relationships between children's school readiness development in LAUP programs and child, family, and program

composition characteristics as measured in this study. We summarize the main findings at the conclusion of this section.

1. Approach to the Analysis

We used hierarchical linear modeling (HLM) to examine associations among children's development in school readiness outcomes and a variety of child and family characteristics and measures of program composition, while accounting for the clustering of children within programs. The use of HLM recognizes that children in the same program had a common set of preschool experiences and thus are not independent of each other.

We used covariate-adjusted fixed effects models to estimate children's progress toward school readiness between the fall and spring. In other words, we included children's fall scores in the models to help explain the variance in their respective school readiness outcomes in the spring. In examining the data, we noted that some children with greater initial skills or abilities did not make gains in at least one area of development (such as in social-emotional development). Therefore, we included an indicator to test whether, after controlling for demographic characteristics, children in our sample with different initial ability made the same rates of progress over time; that is, we included both linear and quadratic terms of initial skills and abilities in the models. We dropped the quadratic term from any models in which it was not significant.

We examined other factors that might be related to children's outcomes in the spring and controlled for those that were related. Even though we had attempted to maintain about the same time interval in terms of the number of weeks between fall and spring assessments, the intervals ranged from 22 to 33 weeks, with a median of 30 weeks, largely because of variations in child and program availability. Therefore, we added a variable to the models indicating the number of weeks between the fall and spring assessments for each child, controlling for the variation in time.[28]

We examined child and family characteristics and program composition in relation to the children's school readiness outcomes listed in Chapter II and discussed earlier in this chapter. These school readiness outcomes included language, literacy, mathematics, social-emotional development, and approaches to learning. In this descriptive look at the children, we report the significance level of the coefficients without adjusting for multiple comparisons, and we looked for patterns in relationships across outcomes.

---

[28] We also assessed whether the interval between program start date and the fall assessment was associated with any of the outcomes. In fact, it was not significant in any of the models; therefore, we omitted the variable in favor of a more parsimonious model. Finally, given the range in the amount of time that parents reported that their child attended the program each week, we entered the number of hours per week that the child attended into the child level of the HLM for the EOWPVT. After controlling for all other child and family characteristics and program composition, we observed that the number of hours per week that a child was involved in the program was not significant; therefore, we omitted the number of hours per week from the models.

As we described in Chapter II, because we did not randomly assign families and children to programs or centers, a number of non-random factors determined why children enrolled in particular centers. Thus, factors beyond our control could affect the composition of families within a given program, and these factors may be associated with children's development over time. To understand whether factors associated with the mix of children in programs were related to outcomes and to control for these factors in our analyses, we included program composition variables in the second level of the HLM models.[29] The program composition variables included (for those in our sample) the average baseline ability of children in the program, the baseline variation of children's ability within the program, the mean income of families in the program, and concentration of recent immigrants in the program. As with all the findings from the study, the associations should not be interpreted as causal relationships. Instead, the associations suggest relationships that could be tested in a more rigorous research design.

For each outcome, we estimated a series of five models by using a sequential approach to show the proportion of variance explained by each set of variables. (Box III.2 lists the categories of variables included in each model.) In Model 1, we included child characteristics (gender, age in months, race/ethnicity, and

---

Box III.2: Categories of Variables Included in HLM Models

Model 1: Includes child characteristics

Model 2: Includes child characteristics and the risk factor indicator

Model 3: Includes child characteristics and family characteristics

Model 4: Includes child characteristics, the risk factor indicator, and family characteristics

Model 5: Includes child characteristics, the risk factor indicator, family characteristics, and program composition.

Box III.3: Factors Comprising the Risk Index (High Risk Defined as Three or More):

- Family income below poverty threshold
- Maternal educational attainment less than high school
- Teenage mother
- Single-parent household
- Maternal depression
- Parent born outside the United States
- Parent living in the United States for five or fewer years
- High exposure to violence, witnessed by child
- Language other than English spoken in household
- Household size greater than five

---

language group) in level 1. In Model 2, we added an indicator of risk status (in which three or more of nine risk factors defined "at-risk" children; see Box III.3).[30] Then, in Model 3, we added a set of individual family characteristics (family income-to-poverty ratio, mother's educational level, maternal employment, whether one or both parents were born outside the United States, recency of immigration, household dependency ratio, and the mother's depressive symptoms; Chapter II lists the variables in the models). We also investigated at-risk and individual family

---

[29] Program composition variables are aggregates of information gathered from sample families within each program.

[30] We examined characteristics of families' risk factors using a cumulative risk index based on earlier research (for example, Sameroff et al. 1987) that indicates that it is the accumulation of risk that is detrimental to children's development. We also examined a cut point of 5 risk factors, but it was not significant.

characteristics together in the level 1 model to help us and First 5 LA understand whether individual characteristics contributed over and above the risk index (Model 4). Finally, in Model 5, we added program composition to level 2 of the models, explaining the intercept of the level 1 model. We estimated the models using the sampling weights that account for sampling probabilities and non-response at both the program and individual levels. We present the results from the HLM analyses in Tables III.34 to III.49 in Appendix C.

The outcomes are standardized (z scores) so that the coefficients may be interpreted as the change in the outcome in standard deviation units for each 1 point increase in the respective variable. In the case of nominal variables, the coefficients should be interpreted as a difference in the outcome in standard deviation units relative to the reference group. In Table II.6 (Chapter II), we show the reference group for all nominal or categorical variables in the models.

Models specified with only child characteristics and time between assessments explained 16 to 63 percent of the variance in the spring scores. When we added family characteristics and program composition to the models, the percentage of explained variance increased by roughly 1 to 3 percentage points for language and cognitive outcomes and by up to 4 percentage points for some noncognitive outcomes, suggesting that although child- and family-level characteristics contributed the most to predicting children's development, a combination of program composition variables, including the mix of peers in classrooms, also contributed to children's learning and development.[31]

2. Relationships of Child Initial Ability to School Readiness Outcomes (Model 1)

Controlling for children's demographic characteristics and the assessment time interval, we found that children's initial ability in the fall was significantly associated with non-linear growth in most school readiness outcomes, suggesting that children with highest ability in the fall gained less over the preschool year than their peers with lower initial ability. In other words, children at the lower end of the distribution made larger gains than their peers who were performing well in the fall. An exception occurred in English receptive vocabulary, for which we measured children's initial ability conceptually[32] and observed a positive linear predictive relationship with performance in the spring. Teacher ratings of children's initial social-emotional skills and approaches to learning also were associated with children's non-linear growth in the same skills in the spring. The exception was Social Cooperation, for which children's improvement increased at the same rate at all levels of initial ability. Children improved in a consistent, linear fashion for executive functioning skills and for parent reports of social-emotional outcomes (except for Social Interaction/Independence, for which children with higher initial ratings gained less than peers with lower initial ratings).

---

[31] For ease of discussion, in this section we use the term "cognitive" outcomes to encompass receptive and expressive vocabulary, spelling, and mathematics; the "noncognitive" outcomes include the various social-emotional and approaches to learning measures.

[32] The ROWPVT-SBE was administered in Spanish and/or English in the fall and conceptually scored. Children received credit for correct answers in either language.

Child Characteristics and Language and Literacy Development. Consistent with the findings from subgroup descriptive analysis, we note that household language was significantly associated with children's development in language and literacy from fall to spring. In the spring, children in the Spanish-only and Spanish-primarily groups scored significantly lower in Expressive Vocabulary, Receptive English Vocabulary, and Letter Naming than their peers in the English-language group. Children in the Other-language group scored lower only in Receptive English Vocabulary compared to their English-speaking counterparts.

Gender, age, and race/ethnicity were also associated with children's development in language and literacy. Boys named fewer letters of the alphabet than girls. Older children identified more letters and scored higher in Receptive Vocabulary than younger children. Latino children scored lower in Receptive Vocabulary than children in other racial/ethnic groups; they also identified fewer letters than Asian children.

Child Characteristics and Mathematics Development. Children classified as Spanish-only and Spanish-primarily scored lower in Mathematics than the English group. Age and race/ethnicity were also significantly associated with children's Mathematics development. Older children improved more than younger children. Latino children gained less than white and Asian children.

Child Characteristics and Social-Emotional Development. We examined both teacher and parent ratings of children's social-emotional behaviors and skills. Controlling for other child characteristics and initial ability, household language was only marginally associated ($p < .10$) with teacher reports of children's overall positive social skills, with children in the Spanish-only group rated lower than those in the English group. We obtained similar results when we examined subscales of children's positive social skills (Social Cooperation, Interaction, and Independence). Teachers of children in the other-language group rated those children lower on the SSRS Problem Behaviors than did teachers of children in the English group. Compared to the English group, teacher ratings of Problem Behaviors in the Spanish-only and Spanish-primarily group did not differ.

Parents of children in the other-language group reported more Social Cooperation and Approaches to Learning skills than those in the English-speaking group. In contrast to teacher ratings, parents of children in the Spanish-only and Spanish-primarily groups reported fewer Externalizing Problem Behaviors.

On average, both parents and teachers rated girls and older children more positively than boys and younger children in their social-emotional development. Teachers rated Latino children as exhibiting more Problem Behaviors than Asian children. Parents rated Latino children as evidencing fewer Problem Behaviors than white children and fewer Internalizing Problems than white and Asian children. Parents of Latino children rated their children more positively on Social Cooperation and Approaches to Learning than parents of African American children rated their children and more positively than parents of Asian children rated their children on Social Interaction and Independence.

Child Characteristics and Development of Approaches to Learning. Gender and age were significantly associated with teacher ratings of children's Positive Approaches to Learning, while household language was not. On average, teachers rated boys as lower in Approaches to Learning than girls and rated older children higher than younger children. Teachers rated African American children as lower in Approaches to Learning skills than Latino children. Household language, however, was significantly associated with children's Self-Regulation and Impulse Control on the executive functioning tasks. On the Pencil Tapping task, children in the Spanish-only group responded correctly across a lower percentage of trials in the spring than those in the English group. On the Walk-a-Line Slowly task, children classified as Spanish-only and Spanish-primarily went faster rather than slower as instructed than those in the English group.[33] Girls correctly passed a higher percentage of the trials on the Pencil Tapping task in the spring than did boys. Older children performed better than younger children on both executive functioning tasks. White children and children in the Other racial group performed better on the Pencil Tapping task than did Latino children.

3. Relationships of Family Characteristics to School Readiness Outcomes (Models 2, 3, and 4)

We added to the models the indicator for risk (Model 2), then the variables for individual family characteristics separately (Model 3), and then both together (Model 4) to show their unique contributions in explaining the variance in children's school readiness outcomes (we note in the text only the significant relationships). When added to the model without the other individual family characteristics, risk was significantly associated with weaker Expressive Vocabulary, Receptive English Vocabulary, Letter Naming, Mathematics development, and performance on the Walk-a-Line Slowly task and with lower teacher ratings of Positive Approaches to Learning. Counter to expectations, in these analyses, parents of children with greater risk rated their children as having significantly fewer Internalizing and Externalizing Problem Behaviors relative to parents of children with fewer risk indicators. When examined in combination with individual family characteristics, risk remained significantly associated with poorer performance only for Letter Naming and the Walk-a-Line Slowly task and was not associated with parent ratings of children's Internalizing and Externalizing Problem Behaviors.

Family Characteristics and Language and Literacy Development. When we examined individual family characteristics variables without the indicator for risk, we saw that family poverty level, maternal education, parent birthplace, immigration recency, and maternal depressive symptoms all were associated with children's language and literacy development as expected. Children whose mothers had at least a high school education scored higher in Expressive Vocabulary and Receptive English Vocabulary than did those whose mothers were without a high school diploma. Children whose parents were both born

---

[33] It is important to keep in mind that all assessors of Spanish-speaking children were bilingual Spanish speakers and the language of administration was in keeping with the language routing procedures described in Appendix B.

outside the United States scored lower in Expressive and Receptive Vocabulary than those whose parents were both born in the United States. Family income was positively associated with Letter Naming. Children whose parents lived in the United States for fewer than five years or whose mothers had more depressive symptoms identified fewer letters.

Family Characteristics and Mathematics Development. When we examined individual family characteristics without the indicator for risk, we noted that family poverty level, maternal education, employment, and maternal depressive symptoms were associated with children's mathematics development. Children from higher-income families gained more in Mathematics. Maternal education was positively associated with children's gains, while maternal employment and maternal depressive symptoms were negatively associated with children's progress in Mathematics. In the model that added both the indicator for risk and individual family characteristics, maternal education was no longer significant.

Family Characteristics and Social-Emotional Development. When we considered individual family characteristics (without the indicator for risk added to the model), we saw that maternal employment was associated with lower teacher ratings of children's Social Interaction. None of the other family characteristics was significantly associated with teacher ratings of children's social-emotional development.

In models with individual family characteristics (without the risk indicator; Model 3), we found some unexpected associations: higher maternal education was associated with lower parent ratings of children's Social Cooperation and Approaches to Learning, and smaller family size was associated with lower Social Cooperation and Approaches to Learning and fewer Internalizing Problems. Children with either or both parents born outside the United States were rated as exhibiting fewer Externalizing Problem Behaviors compared to children with both parents born in the United States.

Family Characteristics and Approaches to Learning Development. In contrast to parent ratings, teachers rated children from larger families and children with working mothers less positively on Approaches to Learning than children from smaller families and children with nonworking mothers. Children of recent immigrants responded correctly across a smaller percentage of trials on the Pencil Tapping task than children with parents who had lived in the United States for more than five years.

4. Relationships of Program Composition to School Readiness Outcomes (Model 5)

In this section, we report associations between average program outcomes, controlling for child and family characteristics, and program composition (because families self-selected into programs). We controlled for some of the observable aspects of program composition and clustering at level 2 of our HLM models and report on which aspects were associated with the outcomes. We included the average entering ability (program-level mean fall scores) of sample children for each outcome, the program-level variation (standard deviation) of

children's entering ability, the program-level concentration of Spanish-speaking children, program socioeconomic status (measured by aggregating individual sample family poverty levels to the program level), and concentration of the program's recent immigrants.[34]

In the area of language and literacy, the concentration of children of recent immigrants and the variation in children's ability were significantly associated with program means in Expressive Vocabulary. Greater variation in children's Expressive Vocabulary scores within program was associated with higher program mean scores in Expressive Vocabulary. A higher concentration of recent immigrants in programs was associated with lower program mean scores in Expressive Vocabulary. Programs with a small percentage of Spanish speakers had lower letter naming scores than programs with a mix of Spanish and English speakers (although programs may have enrolled children who spoke languages other than English).[35] The average program-level socioeconomic status of sample families was positively associated with the program mean score in Receptive Vocabulary. None of the program composition characteristics was significantly associated with children's development in Mathematics.

In the social-emotional domain, teacher ratings of children's social-emotional development were strongly associated with children's average ability within program, program socioeconomic status, and the concentration of Spanish speakers. Teachers in programs with higher socioeconomic status tended to rate children lower on total Positive Social Skills and all PKBS subscales and higher on the SSRS problem behaviors than did teachers in programs with lower socioeconomic status. Teachers in programs with higher average cognitive ability scores rated children higher on all PKBS subscales except for social interaction and lower on the SSRS problems behaviors as compared to children in programs with lower average ability. Children in programs with a high concentration of Spanish speakers were rated as demonstrating fewer Social Interaction skills than children in programs that enrolled a mix of English and Spanish speakers. The only program composition characteristic associated with parent ratings of children's social-emotional development was the concentration of recent immigrants in the program, which was associated with lower average scores of Social Cooperation and Approaches to Learning within programs.

Teachers in programs with higher socioeconomic status rated children as having fewer positive Approaches to Learning skills than did teachers in programs with lower socioeconomic status. Program composition was not significantly associated with children's performance on the Pencil Tapping task. However, for the Walk-a-Line Slowly task, higher program average cognitive ability and greater variation of children's ability within program were significantly associated with better performance.

---

[34] In the models for social-emotional and approaches to learning outcomes, we used the mean and standard deviation of the children's expressive language (EOWPVT-SBE scores) as measures of average ability and variation in ability.

[35] Programs with a majority of Spanish speakers did not differ significantly from those with a mix of Spanish- and English-speaking children.

5. Summary of Findings Pertaining to Relationships of Child and Family Characteristics (and Program Composition) to Children's School Readiness Outcomes

- Across school readiness outcomes, household language was associated consistently with children's progress from fall to spring. Children in the Spanish-only and Spanish-primarily groups did not gain as much as their English-speaking peers.

- The number of challenges faced by families was associated with less positive cognitive outcomes. Children in families with more than three factors that challenged their ability to support children's development were not as successful as children with fewer risk factors.

- At the program level, greater variation within program in children's initial skills was associated with stronger expressive vocabulary and executive functioning skills in the spring.

- Teachers in programs serving children from families of higher socioeconomic status rated children less positively than did teachers in programs serving children from lower socioeconomic status families. Teachers in programs whose children entered with higher average skills rated children more positively on social skills and problem behaviors than did teachers whose children entered at lower average skill levels.

- Higher concentration of recent immigrants within a program was associated with less child progress in expressive vocabulary and lower parent-rated social cooperation and approaches to learning.

CHAPTER IV

CONCLUSIONS AND IMPLICATIONS

In this report on the Universal Preschool Child Outcomes Study, we have provided a wealth of information about programs funded by the Los Angeles Universal Preschool, the characteristics and quality of its center-based programs, the development of LAUP children from fall to spring, and the child and family characteristics that we found to be related to the children's school readiness outcomes. In this chapter, we summarize the major conclusions from all the analyses presented in Chapter III and then, in keeping with the formative nature of the study, we suggest key implications for LAUP's preschool programs.

A. PROGRAM CHARACTERISTICS AND QUALITY

- In 2007-2008, LAUP center-based programs were small, each serving an average of 37 children (in slightly fewer than 2 classrooms on average). Classrooms had an average child:staff ratio of 6:4, although the ratio ranged from 3:4 to 13:4. Ninety-one percent of classrooms had child:staff ratios of 10:1 or better; 83 percent had ratios of 8:1 or better (the requirement for a 5-Star rating).

- Two published preschool curricula predominated in LAUP classrooms. About 31 percent of LAUP children had teachers who used either High/Scope or the Creative Curriculum. No other curriculum was used with more than 10 percent of children. More than half the curricula that teachers reported using addressed multiple developmental domains. However, one-third of the children had teachers who used a nonspecific curriculum.

- Lead teachers in LAUP had considerable experience, averaging more than 10 years teaching preschool. They were educated about child development: almost all children in LAUP classrooms (99 percent) had teachers who had taken six or more classes in early childhood or child development, and 87 percent of children had lead teachers who held at least an AA degree.[1] Lead teacher

---

[1] If a center has at least one teacher who has completed an AA in Child Development, with a minimum of 24 ECE units in child development that include core courses listed under the Child Development Permit Matrix, it fulfills one of the requirements for a four-Star rating.

education level was associated with classroom quality ratings, such that observed quality in the Instructional Support domain of the Classroom Assessment Scoring System (CLASS) was higher in classrooms where teachers had at least a two-year degree.

- LAUP teachers reported that they frequently carried out activities that support children's language and mathematics learning: at least 75 percent of the children had teachers who reported engaging in important language and mathematics activities every day.

- The overall quality of the LAUP programs, as measured by the CLASS, compares favorably with the quality levels reported in other studies of preschool programs. Nevertheless, the observed level of one domain—Instructional Support—was low and at a level generally considered less than ideal for supporting children's school readiness, according to developers of the CLASS. (The other CLASS domains were Emotional Support and Classroom Organization.)

- The low level of observed quality in the Instructional Support domain in LAUP classes is consistent with findings from other studies of preschool programs. The low quality was particularly apparent for the dimensions of Instructional Support that measure Concept Development and Quality of Feedback by teachers.

- One component of the Star rating—the ECERS-R total score—was consistently associated with observed quality as measured by each of the CLASS domains.

- Across LAUP programs, class size ranged from 9 to 28 children with an average of 17.7 children. Class size was associated with quality ratings, with smaller classrooms scoring higher on the CLASS.

- Classrooms in both the Early Launch zip code areas (Areas of Greatest Service Need [AGSN] and non-AGSN areas scored higher on the CLASS Instructional Support domain than those in Second Round AGSN zip codes (controlling for other classroom characteristics). LAUP classrooms located in schools scored higher than non-school-based classrooms on Instructional Support.

B. CHILDREN AND FAMILIES AND THEIR LAUP EXPERIENCES

- Consistent with the diversity of LA County, most LAUP children (75 percent) in fall 2007 were Latino. Nearly two-thirds of the children had parents who were immigrants, but almost all LAUP children were born in the U.S. About 90 percent of the parents have lived in this country for more than five years.

- Families reported that they experienced a variety of challenges (for example, 40 percent lived below the federal poverty level) but also demonstrated strengths

that might help them in meeting some of these challenges. Most children (73 percent) lived in two-parent families. Most families reported maintaining routines such as eating together and having a regular bedtime for their child. Most families said they were able to locate resources to support their children's physical and mental health. In fact, by the spring, virtually all children had medical insurance coverage.

- Families reported engaging in activities to support their children's learning. Approximately 40 percent of families reported reading to their child every day in the past week, and 79 percent read at least three times a week.

- Almost all parents (more than 90 percent) reported that their children experienced good or excellent health. Although most parents said their child had a dental checkup in the past year, nearly 18 percent reported that their child had never been to a dentist. We found that one-fifth of the children were obese according to CDC guidelines, with a body mass index (BMI) at or above the 95th percentile for their age and gender.

- LAUP funds half-day programs, but the programs often operate for the full day using other resources. Children's attendance, as reported by parents, ranged from 1 day to 5 days a week and from 3 hours to more than 40 hours per week. According to parents, children averaged nearly 24 hours per week in "LAUP programs," although parents may not have distinguished between LAUP and non-LAUP programs operating in the same facility.

- More than 20 percent of the LAUP families reported enrolling their child in child care in addition to an LAUP program. On average, children in those families spent an additional 15 hours per week in the care of someone other than their parents (mostly relatives).

- Parents reported not being involved with their LAUP program very often. Five activities occurred most frequently (that is, "several times a year" or more): attending parent-teacher conferences, participating in activities at home that LAUP suggested attending a school or class event, volunteering in the classroom, and, participating in social activities.

- LAUP teachers reached out to parents in a number of ways. Forty-nine percent or more of the teachers reported five practices that often occurred (one or more times a week): encouraging parents to ask their child about his or her day, asking parents to read to their child, contacting parents if their child did something well, inviting parents to visit the classroom, and assigning at-home activities for parent-child interaction.

C. CHILDREN'S PERFORMANCE RELATED TO SCHOOL READINESS

- On average, children's skills increased from fall to spring on most of the direct assessments, teacher behavior rating scales, and parent behavior rating scales

IV: Conclusions and Implications

that we administered. Children showed some of their largest gains in letter knowledge, early writing, social cooperation, and executive functioning.

- For some of the measures of school readiness, a national comparison is available. For these measures in the spring of preschool, LAUP children performed better than the national mean in early writing and problem behavior (that is teachers rated them as displaying fewer problem behaviors). Children's performance was also better than the national mean in social cooperation, social interaction, social independence, and overall prosocial behavior, but below the national mean in receptive and expressive English and Spanish vocabulary, attention, activity level, and sociability.

- Children's performance in the spring was related to how well they performed when they entered the program in the fall, with children who came into the program with higher performance levels also scoring higher in the spring.

- Not all children learned the same amount or at the same rate over the year. Although all groups of children progressed during the year, children who performed more poorly in the fall (that is, scored in the lowest quartile on a measure) learned as much or more during the year as the higher-performing (top quartile) children did, but were not performing sufficiently well in the spring to meet the national norms when they enter kindergarten. In fact, children in the lowest quartile in expressive vocabulary and social skills had lower mean scores in spring 2008 than children performing at the overall average had in fall 2007.

D. HOW CHILD AND FAMILY CHARACTERISTICS WERE RELATED TO SCHOOL READINESS OUTCOMES

- Having more than 3 risk factors (out of the 10 we measured) was a consistent predictor of lower spring scores even after controlling for other factors (including the child's fall scores, age, language grouping, and number of days between fall and spring assessments).[2]

- Membership in the Spanish-only or the Spanish-primarily language group was the most consistent, and often the strongest, predictor of spring scores even after controlling for fall scores, risk factors, and other child and family characteristics. For example, on the conceptually scored measure of expressive vocabulary, the difference between children in the English-language group and those in the Spanish-only group was approximately one-half a standard deviation.

---

[2] Risk factors included (1) Family income below the poverty threshold, (2) Maternal educational attainment less than high school, (3) Teenage mother, (4) Single-parent household, (5) Maternal depression, (6) Parent born outside the United States, (7) Parent living in the United States for five or fewer years, (8) Exposure to violence witnessed by child, (9) Language other than English spoken in household, and (10) Household size greater than five.

- Most preschool classrooms include children whose developmental levels vary. LAUP classrooms differed in how much variation existed, and greater variation in children's expressive language skills when children entered the program in the fall was associated with higher spring scores in expressive language. This suggests that having children with a range of language skills in the classroom supports growth in language development and that children in part learn language from one another.

E. IMPLICATIONS FOR PROGRAMS

- Some of these findings suggest areas in which LAUP programs might consider various strategies that could increase the benefits they provide for enrolled children and families.

- The relatively low levels of parent involvement that parents reported suggest that there are opportunities for LAUP programs to engage parents more.

- The relatively lower observer ratings on the Instructional Support domain of the CLASS suggest that LAUP teachers could benefit from additional professional development in the areas of concept development and quality of teachers' feedback to the children. Other research has found that teachers' performance on these dimensions of instructional practice is associated with children's academic development.

- Classrooms located in Early Launch programs in AGSN zip codes had significantly higher scores on the CLASS Instructional Support scale compared with Second Round programs in AGSN zip codes. LAUP staff may be able to suggest ways in which these two rounds of programs differ, such as how programs in the two rounds were recruited or perhaps differences in professional development opportunities.

- The difference between the performance of English-speaking children and Spanish-only children was large. Although the reasons are not completely clear, some of the group differences could be addressed through strengthening teaching strategies, learning why Spanish-only and Spanish-primarily children attend LAUP for fewer hours per week than other groups, increasing the duration of children's program participation, and/or enhancing Spanish-only parents' program involvement and participation in activities with their children.

- The importance of cumulative risk factors (at least as defined in this study) in predicting children's development beyond the individual elements of risk suggests there may be approaches for targeting families most in need and developing outreach strategies to encourage their enrollment.

- Given the finding that children seem to benefit from classroom environments that include children with a variety of skills and abilities, it may be useful to develop policies that will maintain within-classroom diversity of children from

more- and less-advantaged households (such as reserving a percentage of the spaces for children who are not in the targeted enrollment group)

APPENDIX A

MEASURES USED IN PHASE 2 OF THE UNIVERSAL PRESCHOOL CHILD OUTCOMES STUDY

One of the goals of the First 5 LA Universal Preschool Child Outcomes Study (UPCOS) was to learn how children who were enrolled in the Los Angeles Universal Preschool (LAUP) and Power of Preschool (PoP) programs fared in the fall and spring of one program year, 2007-2008. Phase 1 was a pilot study in which we examined the feasibility, reliability, and validity of various child development measures in the large, culturally and linguistically diverse population of children served by the LAUP programs in Los Angeles County. Based on that experience, and psychometric analyses of the child outcome measures (as reported in Vogel et al. 2008), we held discussions with First 5 LA and its Research Advisory Committee (RAC).[1] Through this process we decided on the battery of measures for Phase 2 listed in Table II.3 in Chapter II. In some cases, we found that measures piloted in Phase 1 worked well with the LAUP enrollees and could be used in Phase 2; in other cases, we created adaptations of the piloted measures; in yet other cases, we decided not to use a measure we had piloted but selected an alternative for use in Phase 2. In this appendix we provide descriptions of each of the measures used in Phase 2 to provide the data reported in this volume.

A. DEVELOPMENTAL DOMAINS ASSESSED BY DIRECT CHILD ASSESSMENTS, TEACHER RATINGS, AND PARENT RATINGS

1. Language and Literacy Development

Preschool Language Assessment Survey (Pre-LAS). For the initial language screening in English, we administered two scales of the English Pre-LAS Simon Says and Art Show. Spanish-speaking children also took the Spanish Pre-LAS Tío Simón Dice and Exposición de Arte. (Duncan and DeAvila 2002). The scales measure receptive and

---

[1] The First 5 LA Universal Preschool Research Advisory Committee (RAC) is made up of research experts representing varied depth and breadth of local and national early childhood research and related expertise who provide input and advice to First 5 LA on matters relating to research and evaluation.

expressive language proficiency, respectively. For determining the language in which the other assessments would be administered, we combined the Pre-LAS with parent reports about the language(s) children spoke at home and to other children. Based on the results, child received the English or the Spanish version of the child assessment battery—or were not tested at all. See Appendix B for details about the language routing protocol.

Rapid Letter Naming Task. Letter naming (Moats 1998; Snow, Burns, and Griffin 1998) and, in particular, rapid letter naming (O'Connor and Jenkins 1999; Rouse and Fantuzzo 2006; Uhry 2002) are among the strongest predictors of later success in school. We developed two different forms with some overlap of items. Each form had some uppercase and some lowercase items. If an item was uppercase on one form, it was lowercase on the other form. Both forms presented very common letters (such as vowels) in both upper and lower case to help provide overlap in scaling the measure. We randomly assigned children in each classroom to one of these forms. Each child was asked to name only 30 uppercase or lowercase letters and the examiner moved to the next letter after three seconds. We allowed both English and Spanish letter names to be accepted as correct. We learned in the pilot study that the 30 letters can be administered in under two minutes.

Woodcock-Johnson–III Spelling (WJ-III) subtest. After the Phase 1 pilot test, we added the WJ-III subtest from the Woodcock-Johnson Psycho-Educational Battery (and Batería III Woodcock-Muñoz) Third Edition (Woodcock, McGrew, and Mather 2001). Following the pilot, the RAC pointed to the need to collect more data about literacy than letter naming. We considered a norm-referenced measure of concepts of print or spelling (written literacy), but standardized measures of concepts of print are not available in both English and Spanish. Therefore, we added environmental print words to the letter-naming task (described next) and used the WJ-III Spelling Subtest. The subtest is a nationally standardized assessment that was used in FACES 2006 and in earlier FACES studies and has well-established evidence of it reliability and validity.

Environmental Print Awareness. As noted, after the pilot testing, we decided to add both Spanish and English words to the end of the Rapid Letter Naming assessment. Each word was presented separately in English and Spanish. To create this assessment, we generated a list of potential words and obtained feedback from the LAUP coaches on the words children are most likely to see in their classrooms. We then added to the list four common sight words (is, me, *es*, and *mi*) that children would see if teachers model sentence writing and take children's dictation. We selected the test words in accordance with the frequency with which they occur in English and Spanish sentences. Assessors credited children with the correct Spanish or English naming of the words or with partial credit if they named a word that begins with the same sound but is not the test word (for example, if a child said "top" for "table" he or she received partial credit for recognizing the sound of the initial consonant).[2]

---

[2] Because this was an experimental measure, we administered it to only half the sample. We also instituted a stop rule that turned out to end the task too early for most children. Thus, we are not able to report results with this measure in this report.

Receptive One-Word Picture Vocabulary Test: English and Spanish–Bilingual Edition (ROWPVT and ROWPVT-SBE). We used the ROWPVT (Brownell 2001) as our measure of receptive vocabulary rather than the Peabody Picture Vocabulary Test (PPVT-4) that had been piloted.[3] The ROWPVT is designed to be administered in English, Spanish, or as a bilingual measure. In fall 20007, we administered it in both languages; in the spring, we administered it in English only as a measure of English-language vocabulary. The ROWPVT corresponds to the EOWPVT (described next) and provides a strong basis of comparison between receptive and expressive vocabulary development since the two measures were standardized on the same sample.

Expressive One-Word Picture Vocabulary Test: English and Spanish–Bilingual Edition (EOWPVT and EOWPVT-SBE). We used the English EOWPVT and the Spanish-Bilingual edition (EOWPVT-SBE; Brownell 2000) to measure expressive vocabulary. The scoring guide provides prompts for both English and Spanish, and indicates acceptable responses in each language (using conceptual scoring). The same stimuli (color pictures) are used for both the English and Spanish-bilingual editions. We administered the EOWPVT-SBE to all children and scored it conceptually according to the standardized scoring procedures. In scoring, assessors indicated the language in which the child responded. The EOWPVT-SBE has been used with diverse populations, and its item development has earned acclaim as "impeccably and responsibly created, analyzed, reformulated, and researched" (Jenkins 2006). At least one study has found it to be sensitive to interventions in Head Start classrooms (Wasik, Bond, and Hindman 2006). Our analysis of the spring 2007 pilot study data found high internal-consistency reliability (alpha = .96), but identified four misfitting items and three items that favored English speakers in the smaller pilot sample. We were concerned that these results might have been spurious due to the large number of comparisons made. Therefore, we asked several Spanish-speaking early childhood professionals if they could identify any bias in the items, and no one identified any. Furthermore, these items had not shown bias in the normative sample. We decided to use the measure in Phase 2, but examined the psychometrics of these items again and found acceptable item functioning

Desired Results Developmental Profile-Revised (DRDP-R) Language.[4] In the area of language, the DRDP-R (California Department of Education 2006) assesses comprehension of increasingly complex language (attributes, time, and causality relationships), ability to follow instructions, conversational skills, and expressive language (including semantic and syntactic complexity). Although predictive and concurrent validity

---

[3] Differential item functioning and misfit in the PPVT, combined with the floor problems in the Spanish version (the Test de Vocabulario en Imagenes Peabody or TVIP) led us to adopt the ROWPVT and ROWPVT-SBE (Brownell 2001) for Phase 2. Given the large percentage of Spanish-speaking children in LAUP programs, we wanted to use a recently developed measure of Spanish language.

[4] The DRDP-R is implemented statewide in child development programs funded by the California Department of Education, Child Development Division, and is required by California State First 5 Commission in Power of Preschool (PoP) demonstration programs. We will prepare a special report on the psychometric characteristics of the DRDP-R found in UPCOS Phase 2 in spring 2009.

and sensitivity to intervention have not been established, the DRDP-R is one of the few teacher report measures of children's development that has established reliability and some evidence of construct validity. The developmental trend in skills was evident in the preliminary research on this measure. Based on our experience with this measure in the pilot study, questions arose as to the independence of the scales. Therefore, during the Phase 2 data collection, we provided teachers with enhanced training in understanding the definitions and examples of the constructs before they completed the DRDP-R forms in fall 2007. We collected DRDP-R data only in the PoP programs.

2. Mathematics Development

Early Childhood Longitudinal Study-Birth Cohort (ECLS-B) Mathematics. Preschool mathematics curricula typically include skills in geometry, data analysis, spatial abilities, and measurement (Clements, Sarama, and DiBiase 2004; National Council of Teachers of Mathematics 2006), all of which are areas tapped by the U.S. Department of Education's Early Childhood Longitudinal Study, Birth Cohort (ECLS-B). The ECLS-B developed the math items for the preschool age group, targeting the range of mathematical concepts that 4-year-olds are typically developing. The measure has been used with children from diverse backgrounds, is available in English and Spanish, and has evidence of reliability and validity. Items in Spanish and English are identical, and we trained assessors to accept responses in either language. In the spring 2007 pretest, the scale's reliability was adequate (Rasch scale reliability = 0.80), and we found no differential item functioning (DIF). We anchored the scale on the national item difficulties to provide a national comparison in the fall. The ECLS-B study did not administer the mathematics scale in the spring of preschool, and the fall kindergarten norms are not available so we did not have a set of national spring norms for comparison. Therefore, we anchored the fall and spring UPCOS data to the ECLS-B item difficulties and created W scores so that we could analyze the scores longitudinally fall to spring.

Desired Results Developmental Profile-Revised (DRDP-R) Mathematics. In the area of mathematics, the DRDP-R assesses number sense and operations, shapes, classification, patterns, time, and measurement. As with all the DRDP-R scales, we provided teachers with enhanced training in understanding the definitions and examples of the constructs before they completed the DRDP-R forms in fall 2007, and only PoP programs completed the DRDP-R.

3. Social-Emotional Development and Approaches to Learning

We selected measures that obtain information from three sources: parents', teachers', and the assessors' ratings of children's social-emotional development and behavior.

Preschool Kindergarten Behavior Scales–2 (PKBS-2) (Merrell 2002). The teachers rated children's social-emotional competence by using the positive social scales from the PKBS-2. The scales include social cooperation (children's ability to share and follow classroom rules), social independence, positive social interactions.. In interviews with the teachers in Phase 1, we found that they were positive about the PKBS-2 social scales, and some teachers commented that the scales captured more typical behaviors, noting that the

scales were more closely aligned with common preschool behaviors than other scales that we piloted. We also found that teachers preferred a frequency rating ("how often" rather than "how true") and expressed a preference for more points on a scale (4- to 5-point scales rather than 3-point scales). The PKBS-2 met these criteria.

During the pilot study, we selected some of the PKBS-2 items on the basis of the parents' responses in focus groups. We then used cognitive interviewing to ask the parents about the items and learned that parents interpreted and responded to some of the items in very different ways, indicating difficulty in understanding some of the terms used in the items (for examples, "clings"). During the cognitive interviewing, we revised the items iteratively until we found that parents from a variety of linguistic and cultural backgrounds were interpreting the items in similar ways. We administered the adapted version of the PKBS-2 in the Phase 2 parent interview.

Social Skills Rating System (SSRS) (Gresham and Elliott 1990). Given that teachers in the pilot phase were concerned about the high number of problem behaviors included on both the PKBS and the SSRS scales they used for rating children's behavior, for Phase 2 we responded to the teachers (as well as to our concerns about teacher burden) by asking teachers to complete the short, 10-item scale of problem behavior ratings on the SSRS.

ECLS-K Approaches to Learning. Teachers rated each child on the six items comprising the Approaches to Learning Scale from the Early Childhood Longitudinal Study—Kindergarten Class of 1998–1999 (ECLS-K) study. The reliability of the scale in our sample was strong in Phase 1 ($\alpha = 0.91$), although we identified a ceiling problem (27 percent of the children scored at the assessment's maximum).

Leiter-R Examiner Rating Scales. At the conclusion of the direct assessment, assessors rated children on subtests from the Leiter-R Examiner Rating Scales (Roid and Miller 1997) to obtain a more comprehensive picture of each child's attention, activity level, and sociability during the assessment task. The items on these three subscales are part of the Leiter cognitive/social scale.[5] The scales have been used successfully in three large-scale studies—FACES 2006 (ACF 2006a), Early Head Start prekindergarten follow-up study (ACF 2006b), and Home Visiting 2000 (Olds et al. 2004)—and demonstrated good reliability and predictive validity.

Executive Functioning Tasks: Pencil Tapping and Walk-a-Line Slowly (Balance Beam). Executive functioning (EF) involves "processes that are integral to the emerging self-regulation of behavior and developing social and cognitive competence in young children" (Blair, Zelazo, and Greenberg 2005, p. 561). These important skills develop during the preschool years and are strong indicators of children's future social and academic success. EF tasks may be categorized as either "hot" or "cool," with the "hot" tasks

---

[5] We had included the self-regulation scale in the pilot, but since our assessments did not include manipulatives that children would have had to refrain from playing with during the assessment, assessors were unable to complete these items.

assessing social-emotional aspects that have an affective component and the "cool" tasks assessing aspects of control of cognition (Hongwanishkul, Happaney, Lee, and Zelazo 2005). The "hot" aspects are more difficult to assess and may be more contextual (Blair, et al., 2005), and the emotional arousal of a task may differ for children from diverse cultural backgrounds. The hot tasks typically involve food, gifts, or frustration (Kochanska and Murray 2002). In addition, it may be difficult to select tasks that are emotionally evocative for diverse groups of children. Therefore, we did not include a hot EF task. Each of the measures that we used assesses inhibitory control. These tasks have some evidence of reliability and validity in small studies, and developmental sensitivity has been established. The measures we selected have been used with diverse samples of children. They have all been used with 4-year-olds and may be administered in both English and Spanish.

The Pencil-Tapping task (Smith-Donald, Raver, Hayes, and Richardson 2007) is an adaptation of a peg-tapping task (Blair 2002; Diamond and Taylor 1996) that assesses working memory, attention, and inhibitory control. Evidence of reliability of the 16 trials in this task is strong ($\alpha = 0.82$ with preschoolers in Blair and Razza 2007 and $\alpha = .93$ in our pilot test). With a sample of low-income 3- to 4-year-olds, the peg-tapping task has demonstrated a relationship to later kindergarten outcomes in literacy and mathematics ($r = 0.18$ to 0.47) (Blair and Razza 2007). On the other hand, Smith-Donald and colleagues combined Peg Tapping with other "cool" tasks and did not find a relationship to academic outcomes but did find a significant relationship with a measure of social competence ($r = 0.53$).

The Walk-a-Line Slowly task (Murray and Kochanska 2002; Smith-Donald et al. 2007) is a measure of inhibitory control of large muscle (gross motor) movement. . The child is first asked to walk on a 6 foot long line, and then to walk the line as slowly as possible. Each of the child's attempts is timed, and the child's score for impulse control is the time difference between the first and second trial. We followed the Smith-Donald et al. protocol, using a line that is only 1-inch wide (rather than the 6-inch-wide line used by Murray and Kochanska 2002). In the pilot test, we found internal consistency reliability to be strong ($\alpha = .96$). When used in a study in Chicago Head Start classrooms, the task loaded on a compliance and executive control factor with the Pencil-Tapping task, and that factor was correlated with social competence (Smith-Donald et al. 2007).

Desired Results Developmental Profile-Revised (DRDP-R). In the area of social-emotional development, the DRDP-R scales include measures of empathy, self-identity, positive self-regard, cooperation with adults, friendship skills, cooperative play, and conflict negotiation. Its executive functioning scales include self-regulation and approaches to learning. As already noted the DRDP-R has been used with diverse populations and has established evidence of internal-consistency reliability. It meets all of our selection criteria, although its predictive and concurrent validity and sensitivity to intervention have not been established. Only PoP teachers completed the DRDP-R.

4. Physical Health and Motor Development

Walk-a-Line Slowly Motor competence has been associated with both academic and social outcomes (Frisk 1995; Seitz, Jenni, Molinari, Caflisch, Largo, and Latal Hajnal 2006;

Wolff, Gunnoe, and Cohen 1985). In the Walk-a-Line Slowly task, the assessor (in addition to impulse-control scoring) records the child's motor control. Assessors rate the child's success in staying on the line on his or her first attempt, using a 4-point scale from "hardly at all" to "almost all the time." Maintaining the balance needed to walk on a line has a long history of use in early childhood studies. The measure has been used with diverse populations, is age appropriate for 4-year-olds, and has instructions that may be administered in both English and Spanish.

Height and Weight. Based on the recommendations of the RAC, we weighed the children and measured their height in the fall and spring. From these measurements, we computed and reported a BMI for each child.

Parent Ratings of Health. The RAC recommended that we use scales from the health survey developed by Frances Glascoe in order to have a standardized measure of health and physical well-being (Los Angeles County Department of Health Services 2005). To provide a comparison with the overall population in LA County we included some questions from the LA County Health Survey, which adapted the Glascoe survey.

Desired Results Developmental Profile-Revised (DRDP-R) Motor Development. The DRDP-R includes scales that allow the teacher to assess the child's fine and gross motor development, and knowledge of health and safety. Only teachers in PoP programs completed the DRDP-R.

B. CLASSROOM OBSERVATION MEASURES

Classroom Assessment Scoring System (CLASS). The CLASS, which examines classroom interactions, is a measure of classroom climate and teacher instructional practice (Pianta, La Paro, and Hamre 2006). It does not depend on the presence of any particular materials or curriculum. The "focus is on what teachers do with the materials they have and in the interactions they have with students" (Pianta et al. 2008). The preschool CLASS assesses three domains: (1) emotional support for children (including class climate, teacher sensitivity, and regard for student perspectives), (2) organization of the classroom (including the learning formats used, time use and productivity, and behavior management), and (3) support for instruction (including the level of concept development, the quality of the feedback to children, and the modeling of language). The CLASS was developed based on extensive literature reviews and on data from classroom observations in both the NICHD studies (ORCE and the COS-K; NICHD ECCRN 2002; Pianta et al. 2002) and the NCEDL Multi-State Pre-K Study (Early et al. 2005). The CLASS is used widely (several thousand classrooms) and has shown relationships to important child outcomes such as performance on tests of literacy and increased child engagement in learning (Pianta 2003; Pianta, La Paro, and Hamre 2006; Pianta et al. 2005). Concurrent validity has been demonstrated with the ECERS-R, the Snapshot (Ritchie et al. 2002), and the percentage of time children who are not engaged (Pianta, La Paro, and Hamre 2006).

Language Interaction Snapshot (LISn). In consideration of First 5 LA's goal to examine the language support for English Language Learners, particularly in the absence of

an existing scale that would capture the extent to which English or other languages are used in conversations between adults and children in the classroom, we developed a time-sampling observation tool, which we called the Language Interaction Snapshot (LISn). We used it in Phase 2 on a pilot basis to capture information about the language(s) teachers use with children in the classroom, the types of language used in different settings, and the adults who interact with ELL children. We present a complete report on the data obtained from the LISn in Appendix D.

C. REFERENCES

Administration for Children and Families. "Head Start Family and Child Experiences Survey (FACES) 2006 Cohort. Revision of a Currently Approved Collection. Clearance Package Supporting Statement and Data Collection Instruments." Princeton, NJ: Mathematica Policy Research, Inc., February 20, 2006. (a)

Administration for Children and Families. "Research to Practice: Preliminary Findings from the Early Head Start Prekindergarten Followup." U.S. Department of Health and Human Services, April 2006. (b)

Blair, Clancy. "School Readiness: Integrating Cognition and Emotion in a Neurobiological Conceptualization of Children's Functioning at School Entry." American Psychologist, vol. 57, 2002, pp. 111-127.

Blair, Clancy, and Rachel Peters Razza. "Relating Effortful Control, Executive Function, and False Belief Understanding to Emerging Math and Literacy Ability in Kindergarten." Child Development, vol. 78, 2007, pp. 647-663.

Blair, Clancy, Philip David Zelazo, and Mark Greenberg. "The Measurement of Executive Function in Early Childhood." Developmental Neuropsychology, vol. 28, 2005, pp. 561-571.

Brownell, R. "Expressive One-Word Picture Vocabulary Tests." San Antonio, TX: Harcourt Assessment, Inc., 2000.

Brownell, R. "Receptive One-Word Picture Vocabulary Tests." San Antonio, TX: Harcourt Assessment, Inc., 2001.

California Department of Education, Child Development Division. "Desired Results for Children and Families—Revised." Sacramento, CA: California Department of Education, 2006.

Clements, Douglas H., Julie Sarama, and Ann-Marie DiBiase (eds.). *Engaging Young Children in Mathematics: Standards for Early Childhood Mathematics Education.* Mahwah, NJ: Lawrence Erlbaum Associates, 2004.

Diamond, Adele, and Colleen Taylor. "Development of an Aspect of Executive Control: Development of the Abilities to Remember What I Said and to Do as I Say, Not as I Do." *Developmental Psychobiology*, vol. 29, 1996, pp. 315-334.

Duncan, Sharon, and Edward DeAvila. "Pre-Language Assessment Scales [PreLAS2000]." Monterey, CA: CTB-McGraw Hill, 2002.

Frisk, Max. "Mental and Somatic Health and Social Adjustment in Ordinary School Children During Childhood and Adolescence Related to Central Nervous Functions as Expressed by a Complex Reaction Time." *European Child & Adolescent Psychiatry*, vol. 4, 1995, pp. 197-208.

Gresham, Frank M., and Stephen N. Elliott. "Social Skills Rating System." Circle Pines, MN: American Guidance Service, 1990.

Harvard University and Center for Applied Linguistics. "Development of English Literacy in Spanish-Speaking Children—Phonological Awareness I." Cambridge, MA: author, 2002.

Hongwasnishkul, Donaya, Keith Happaney, Wendy Lee, and Philip David Zelazo. "Assessment of Hot and Cool Executive Function in Young Children: Age-Related Changes and Individual Differences." *Developmental Neuropsychology*, vol. 28, 2005, pp. 617-644.

Jenkins. Jill A.. "Review of the Expressive One-Word Picture Vocabulary Test: Spanish-Bilingual Edition." *Mental Measurements Yearbook*, 2006.

Kochanska, Grazyna, and Kathleen Murray. "Parent-Child Study: Effortful Control Battery." Iowa City, IA: University of Iowa, 2002.

Los Angeles County Department of Health Services, *Children with Special Health Care Needs.* Los Angeles: LA Health; June 2005.

Merrell, Kenneth. "Preschool Kindergarten Behavior Scale-2." Austin, TX: PRO-ED, 2002.

Moats, Louisa. "Teaching Reading Is Rocket Science: What Expert Teachers of Reading Should Know and Be Able to Do." Washington, DC: American Federation of Teachers, 1998.

Murray, Kathleen T., and Grazyna Kochanska. "Effortful Control: Factor Structure and Relation to Externalizing and Internalizing Behaviors." *Journal of Abnormal Child Psychology,* vol. 30, 2002, pp. 503-514.

National Council of Teachers of Mathematics. "Curriculum Focal Points for Prekindergarten through Grade 8 Mathematics: A Quest for Coherence." Reston, VA: NCTM, April 2006.

Olds, David L., JoAnn Robinson, Lisa Pettitt, Dennis W. Luckey, John Holmberg, Rosanna K. Ng, Kathy Isacks, Karen Sheff, and Charles R. Henderson. "Effects of Home Visits by Paraprofessionals and by Nurses: Age 4 Follow-Up Results of a Randomized Trial." *Pediatrics*, vol. 114, 2004, pp. 1560-1568.

O'Connor, Rollanda E., and Joseph R. Jenkins. "Prediction of Reading Disabilities in Kindergarten and First Grade." *Scientific Studies of Reading*, vol. 3, no 2, 1999, pp. 159-197.

Pianta, R.C. "Teacher-Child Interactions: The Implications of Observational Research for Re-designing Professional Development." Presentation to the Science and Ecology of Early Development (SEED), National Institute of Child Health and Human Development, Washington, DC, February 2003.

Pianta, R.C., K.M. La Paro, C. Payne, M.J. Cox, and R. Bradley. "The Relation of Kindergarten Classroom Environment to Teacher, Family, and School Characteristics and Child Outcomes. *The Elementary School Journal*, vol. 102, 2002, pp. 225–238.

Pianta, R.C., Howes, C., Burchinal, M., Bryant, D., Clifford, R., Early, C., et al. "Features of Pre-kindergarten Programs, Classrooms, and Teachers: Do They Predict Observed Classroom Quality and Child-Teacher Interactions? *Applied Developmental Science,* vol. 9, 2005, pp. 144-159.

Pianta, Robert C., Karen M. LaParo, and Bridget K. Hamre. *Classroom Assessment Scoring System Manual, Pre-K.* Baltimore, MD: Paul H. Brookes Publishing Co., 2006.

Ritchie, S., C. Howes, M. Kraft-Sayre, and B. Weiser. *Emerging Academic Snapshot.* Los Angeles: University of California at Los Angeles, 2002.

Roid, Gale H., and Lucy J. Miller. "Leiter-R Performance Scale—Revised." Wood Dale, IL: Stoelting Co., 1997.

Rouse, Heather L., and John W. Fantuzzo. "Validity of the Dynamic Indicators for Basic Early Literacy for Urban Kindergarten Children." *School Psychology Review,* vol. 35, no 3, 2006, pp. 341–355.

Seitz, J., O.G. Jenni, L. Molinari, J. Caflisch, R.H. Largo, and B. Latal Hajnal. "Correlations Between Motor Performance and Cognitive Functions in Children Born <1250 G at School Age." *Neuropediatrics*, vol. 37, 2006. pp. 6-12.

Smith-Donald, Radiah, C. Cybele Raver, Tiffany Hayes, and Breeze Richardson. "Preliminary Construct and Concurrent Validity of the Preschool Self-Regulation Assessment (PSRA) for Field-Based Research." *Early Childhood Research Quarterly*, vol. 22, 2007, pp. 173-187.

Snow, Catherine E., M. Susan Burns, and Peg Griffin (eds.).*Preventing Reading Difficulties in Young Children.* Washington, DC: National Academy Press, 1998.

Uhry, J.K. "Finger-Point Reading in Kindergarten: The Role of Phonemic Awareness, One-to-One Correspondence, and Rapid Serial Naming." *Scientific Studies of Reading,* vol. 6, no. 4, 2002, pp. 319-342.

Vogel, Cheri, Nikki Aikens, Sally Atkins-Burnett, Emily Sama Martin, Margaret Caspe, Susan Sprachman, and John M. Love. "Reliability and Validity of Child Outcome Measures with Culturally and Linguistically Diverse Preschoolers: The First 5 LA Universal Preschool Child Outcomes Study Spring 2007 Pilot Study." Princeton, NJ: Mathematica Policy Research, Inc., March 2008.

Wasik, Barbara A., Mary Alice Bond, and Annemarie Hindman. "The Effects of a Language and Literacy Intervention of Head Start Children and Teachers." *Journal of Educational Psychology*, vol. 98, 2006, pp. 63-74.

Wolff, P.H., C. Gunnoe, and C. Cohen. "Neuromotor Maturation and Psychological Performance: a Developmental Study." *Developmental Medicine and Child Neurology*, vol. 27, 1985, pp. 344-355.

Woodcock, Richard W., Kevin S. McGrew, and Nancy Mather. *Woodcock-Johnson III*. Itasca, IL: Riverside Publishing, 2001.

APPENDIX B

PROCEDURES FOR ROUTING CHILDREN INTO THE MOST APPROPRIATE LANGUAGE FOR ASSESSMENT

---

1. INTRODUCTION

A primary concern for the UPCOS was to devise a way of assessing the development of children from multiple language backgrounds as accurately as possible. In what follows, we provide details on how MPR determined whether it was most appropriate to assess children in English or in Spanish. We used a combination of parental reports, language screeners, empirically derived rules for proceeding in a given language (or switching to the other) based on performance on the screener, and conceptual scoring of most assessments. Although we began with six groupings of children's language based on their parents' reports, for analysis we narrowed these to four policy relevant groupings: (1) English only or primarily, (2) Spanish only, (3) Spanish primarily, and (4) other language only or primarily.

2. LANGUAGE ROUTING PROTOCOL DEVELOPED FOR THE FIRST 5 LA UPCOS, 2007-2008[1]

The language routing protocol developed for UPCOS included four steps to determine the most appropriate language for assessing a child: (1) obtain a parent report of child's language use in the home and with peers; (2) assign child to one of six language groups and either an English-only or bilingual (Spanish/English) assessment, based on the parent report; (3) assess child's performance on two subtests from the Preschool Language Assessment Survey 2000 (Pre-LAS 2000; Duncan and DeAvila 2002); and (4) refine the routing based on a conceptually scored receptive vocabulary test.

---

[1] Mathematica Policy Research, Inc. (MPR) developed this protocol in collaboration with First 5 LA and the National Center for Latino Child & Family Research under contract 07110 for conducting UPCOS for First 5 LA. Those who adopt or adapt this protocol are asked to acknowledge MPR, First 5 LA, and the UPCOS project in any printings or electronic reproductions of assessment manuals, project reports to funders, presentations, and publications.

Step 1: On the consent form, parents reported their home language. Parents then reported the language they use in speaking to the child, the language the child uses with the parents, and the language the child uses with other children (see Figure 1).[2] This set of questions provides a concise picture of the child's sociolinguistic context. The parents responded to each question using one of five categories, ranging from "only English" to "only my primary language."

Figure B.1. Scale for Obtaining Parents' Report of Children's Use of Language in the Home and with Peers

	CHECK ONE BOX PER LINE				
	Only English	Mostly English, but sometimes my primary language	Both languages about equally	Mostly my primary language, but some English	Only my primary language
1. What language do you speak most often with your child?	☐ 5	☐ 4	☐ 3	☐ 2	☐ 1
2. What language does your child speak most often with you?	☐ 5	☐ 4	☐ 3	☐ 2	☐ 1
3. What language does your child speak most often with other children?	☐ 5	☐ 4	☐ 3	☐ 2	☐ 1

Step 2: We coded parents' responses on a scale ranging from 1 to 5 and then tabulated the total values across the three questions (ranging from 3 to 15). Along with reported home language, we used this information to assign children to one of six language groups and to an English-only or a bilingual (Spanish/English) assessment path. We used six language groups to more accurately reflect the variability in a child's language proficiencies across languages than the more typical English, Spanish, and other grouping.

Group 1: "English only" if the child received 14 or 15 points

Group 2: "Spanish only" if the child received 3 or 4 points and had Spanish as his or her home language

Group 3: "Spanish Primarily" if the child received 5 to 9 points and had Spanish as his or her home language

Group 4: "Other language only or primarily" was assigned to children who received 3 to 9 points and had a language other than English or Spanish as his or her home language

---

[2] We based the questions used to help determine language routing on the research of Vera F. Gutiérrez-Clellen (2003) and Head Start Impact Study procedures. In UPCOS, however, we asked these questions of the child's parent or guardian rather than the child care provider.

Group 5: "English Primarily with Spanish home language" if the child received 10 to 13 points and had Spanish as his or her home language

Group 6: "English Primarily with other home language" if the child received 10 to 13 points and had a language other than English or Spanish as his or her home language

Language Groups 1, 4, and 6 were administered an English-only assessment, and Groups 2, 3, and 5 a bilingual assessment (Spanish/English). Figure 2 illustrates the language and assessment group routing.

Figure B.2 Language and Assessment Group Routing Protocol (Fall 2007 Administration)[3]
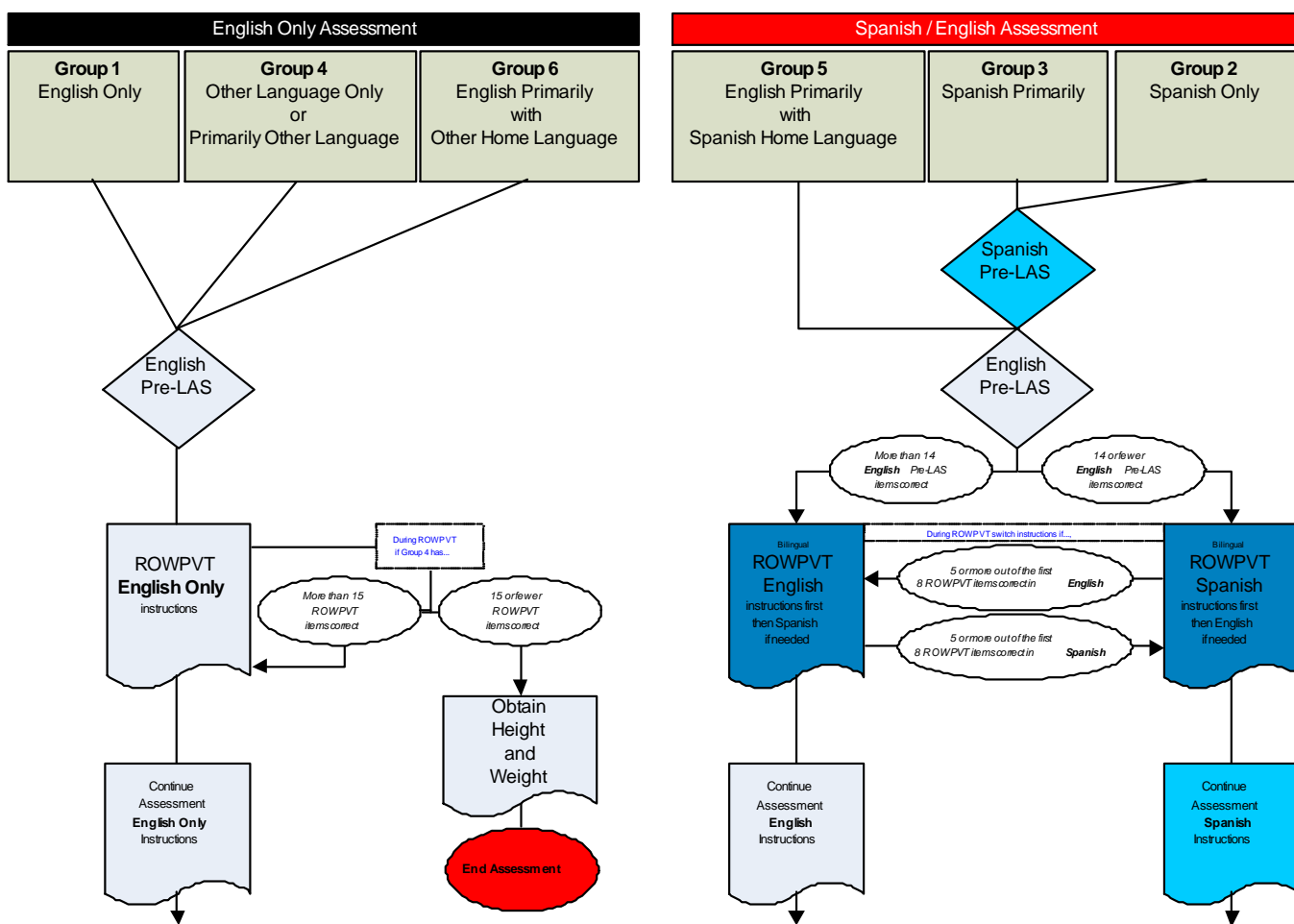
Step 3: All children were administered the English version of two subtests of the Pre-LAS 2000, Simon Says, and Art Show. For children who did not come from Spanish-speaking homes, this administration was used as a warm-up and not for routing purposes.[4] The two Pre-LAS 2000 subtests assessed the receptive and expressive language of children classified into the Spanish only or primarily groups (2 and 3). In combination, these tasks helped determine whether children should receive the rest of the assessment battery in English or Spanish. In an effort to determine an optimal cutoff point of these subtests, we estimated several receiver operating characteristic curves using pilot data. The analyses suggested that using children's performance on both Pre-LAS subtests as a whole, rather than individually on the Art Show or Simon Says, was the optimal method for routing children. As a result, the empirically derived stop rule for the Pre-LAS screener is such that children who had 14 or fewer correct responses across both of the English Pre-LAS subtests and were from Spanish-speaking households were routed into the bilingual Receptive One-Word Picture Vocabulary Test, Spanish Bilingual Edition (ROWPVT-SBE; Brownell 2001) and received the Spanish instructions first. For children who came from homes where a language other than English or Spanish is spoken, the Pre-LAS subtests were not used to route children out of the assessment.

Step 4: We used the children's performance on the ROWPVT-SBE assessment to determine if they should continue the assessment in Spanish or English. The ROWPVT is conceptually scored so that it takes into account the fact that bilingual children's vocabulary typically is distributed across two languages. Thus, we first asked children to identify an object in their dominant language, and if they respond incorrectly, we asked the item again in their non-dominant language. If a child routed into the Spanish assessment answered five or more of the first eight items correctly in Spanish, he or she continued or switched to the administration that leads with Spanish questions. Conversely, if a child answered five or more of the first eight items correctly in English, he or she continued or switched to the administration that leads with English questions. The assessment then continued with a variety of conceptually scored measures.[5]

3. References

Brownell, R. "Receptive One-Word Picture Vocabulary Tests: Spanish Bilingual Edition." San Antonio, TX: Harcourt Assessment, Inc., 2001.

Duncan, S.E., and E. DeAvila. "Preschool Language Assessment Survey 2000 Examiner's Manual." Monterey, CA: CTB/McGraw-Hill, 1998.

---

[4] Children in the Spanish-only and Spanish-primarily groups were also administered the two Spanish Pre-LAS subtests as a warm-up before they are given any assessments in English.

[5] Although the ROWPVT's publisher specifically designed this assessment to be conceptually scored (and also normed it as such), we administered the other conceptually scored measures following a conceptual scoring approach we developed for the UPCOS study.

Gutierrez-Clellen, V.F., and J. Kreiter. "Understanding Child Bilingual Acquisition Using Parent and Teacher Reports." *Applied Psycholinguistics*, vol. 24, no. 2, 2003, pp. 267-88.

López, M.L., S. Barrueco, and J. Miles. "Latino Infants and Their Families: A National Perspective of Protective and Risk Factors for Development." Commissioned Report for the National Task Force on Early Childhood Education for Hispanics. Available at [www.ecehispanic.org/work/Latino_Infants.pdf]. 2006.

APPENDIX C

TABLES OF RESULTS PRESENTED IN CHAPTER III

Table III.5.   Distribution of Programs Served by LAUP Lead Teachers: Demographic Characteristics and Language Use (Weighted)

Teacher Characteristic	Number of Teachers	Percentage of Programs[b]
Gender		
Female	52	97.4
Male	1	2.6
Race/Ethnicity[a]		
Latino	70	49.0
African American, non–Latino	24	17.1
White, non–Latino	23	17.4
Asian, non–Latino	8	7.8
Other	13	8.6
Language Spoken at Home[a]		
English	58	45.5
Spanish	3	2.3
English and Spanish	53	41.1
Other language	5	3.7
English and other language	9	7.4
Language Used to Read to Children		
English	48	39.1
English and Spanish	71	58.0
English and another language	4	2.9
Language Used for Presentations[a]		
English	57	44.5
English and Spanish	71	51.4
English and another language	6	4.2
Sample Size	**138**	

Source: Teacher interviews, fall 2007.

Note: Data are weighted by program.

[a]Categories are mutually exclusive.

[b]Categories with missing data do not sum to 100 percent.

Appendix C

Table III.6.   Experience of Children's Lead Teachers (Weighted)

	Mean (Standard Deviation)	Median
Years working with children	14.7 (1.1)	12
Years teaching preschool	12.0 (0.8)	11
Sample Size	**138**	

Source: Teacher interview, fall 2007.

Note: Data are weighted by program.

Table III.7.   Education and Credentials of Children's Lead Teachers (Weighted)

Teacher's Education and Credentials	Percentage of Children[b]
Highest Grade of School Completed	
High school diploma/equivalent	0.7
Some college but no degree	13.1
Associate's degree	30.3
Bachelor's degree	40.1
Graduate or professional school, no degree	5.9
Graduate or professional degree	10.0
Field in Which Obtained Highest Degree[a]	
Child development/developmental psychology	37.3
Early childhood education	31.3
Other	31.4
College Courses Included 6 or More Classes in Early Childhood or Child Development	
Has a Child Development Associate Credential	69.1
Has a State-Awarded Preschool Certificate	86.0
Has a Teaching Certificate or License	72.6
Sample Size	**138**

Source: Teacher interview, fall 2007.

Note: Data are weighted by program.

[a] Includes only those teachers with an associate's degree or higher (N=114).

[b] Categories with missing data do not sum to 100 percent.

Appendix C

Table III.9.   Frequency of Reading and Language Activities (Weighted)

Reading and Language Activity	Never		Monthly		Weekly		Daily or Almost Daily	
	Percentage	Standard Error	Percentage	Standard Error	Percentage	Standard Error	Percentage	Standard Error
Work on letter naming	1.6	1.6	0.0	0.0	6.1	2.1	92.3	2.6
Practice writing letters	0.5	0.5	1.1	1.1	20.9	4.1	77.5	4.2
Discuss new words	0.0	0.0	1.3	1.3	14.7	3.6	84.0	3.8
Dictate stories to an adult	1.1	0.8	3.1	1.6	29.9	4.6	66.0	4.9
Work on phonics	2.1	1.5	5.5	2.2	15.6	3.4	73.8	4.2
Listen to teacher read stories where they see the print	0.0	0.0	0.5	0.5	0.0	0.0	99.5	0.5
Listen to teacher read stories where they do not see the print	37.3	4.7	12.4	3.4	14.8	3.9	35.6	4.8
Retell stories	1.3	1.3	10.3	3.2	23.8	4.3	64.7	5.0
Learn about conventions of print	0.2	0.2	5.3	2.1	13.9	3.4	80.6	3.8
Write own name	0.0	0.0	0.0	0.0	10.3	3.2	89.7	3.2
Learn about rhyming words and word families	1.4	1.1	9.3	2.7	27.5	4.5	61.8	4.9
Learn about common prepositions	0.0	0.0	9.4	3.0	22.6	4.3	68.0	4.8

Source: Teacher interview, winter 2008.

Table III.10.   Frequency of Mathematics Activities (Weighted)

Mathematics Activity	Never		Monthly		Weekly		Daily or Almost Daily	
	Percentage	Standard Error	Percentage	Standard Error	Percentage	Standard Error	Percentage	Standard Error
Count out loud	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0
Work with geometric manipulatives	0.0	0.0	0.4	0.4	8.5	3.0	91.2	3.0
Work with counting manipulatives	2.1	1.4	1.5	1.0	7.7	2.6	88.6	3.1
Play mathematics-related games	0.0	0.0	1.2	1.0	23.2	4.2	75.6	4.2
Use music to understand mathematics concepts	2.5	1.7	8.8	2.9	20.3	4.2	68.3	5.0
Work with rulers or other measuring instruments	1.5	1.3	4.3	1.7	23.0	4.7	71.2	4.8
Engage in calendar-related activities	2.1	1.4	2.0	1.4	2.9	1.5	93.1	2.5
Engage in activities related to telling time	3.1	1.4	13.3	3.6	15.6	3.7	68.0	4.6
Engage in activities involving shapes and patterns	0.0	0.0	4.2	2.0	9.2	2.8	86.6	3.3

Source: Teacher interview, winter 2008.

Table III.11.   Children Screened for Health and Developmental Problems (Weighted)

	Children	
	Number	Percentage
Developmental Screener Name		
Ages and Stages Questionnaires	263	33.6 (6.3)
Level 2 Health Screening	90	9.4 (3.6)
Child Development Inventory	47	7.1 (3.6)
Parent Evaluation of Developmental Status	27	4.0 (2.5)
Developmental Indicators for Assessment of Learning	25	3.2 (2.2)
Brigance Preschool Screen for Three and Four Year Old Children	11	0.8 (0.6)
Other	357	41.9 (6.1)
Sample Size	**839**	

Source: Teacher interview, winter 2008.

Appendix C

Table III.12   Observed Classroom Quality Scores in UPCOS Classrooms Compared with Studies of Other Preschool Programs (Winter 2008, weighted)

Domains and Dimensions	UPCOS Classrooms	"Prepared to Learn": RAND Study of Early Care and Education in California	Multi-State Study of Prekindergartens and Study of State-Wide Early Education Programs (MS/SWEEP)	My Teaching-Partner	Tulsa County Head Start	Tulsa's Public Schools' Pre-K Programs
CLASS Emotional Support	5.9	5.6	NR	NR	NR	NR
Positive climate	5.9	5.6	5.3	5.2	5.0	5.1
Negative climate	1.2	1.4	1.6	1.6	1.5	1.4
Teacher sensitivity	5.4	5.0	4.7	4.3	4.8	4.8
Regard for student perspectives	5.2	5.0	NR	4.4	4.6	4.4
CLASS Classroom Organization	5.4	5.0				
Behavior management	5.5	5.3	5.0	4.9	4.4	5.0
Productivity	5.6	5.1	4.5	5.4	5.1	5.2
Instructional learning formats	5.1	4.5	3.9	4.6	4.9	4.6
CLASS Instructional Support	2.6	2.7				
Concept development	2.1	2.4	2.1	2.7	2.6	2.8
Quality of feedback	2.5	2.8	2.0	2.9	3.5	3.3
Language modeling	3.4	3.0	NR	2.9	3.7	3.5
Number of Classrooms	**79-83**	**384**	**694**	**164**	**28**	**77**

Source: For MS/SWEEP, and My TeachingPartner: Pianta, LaParo, and Hamre (2006). For Tulsa Early Childhood Programs: Phillips, Gormley, and Lowenstein (2007). For RAND study: Karoly et al. 2008, p. 103 (we show data from the 4-year-old cohort only).

NR = not reported

Table III.13   Observed Classroom Quality Scores: Winter 2008 (Weighted)

Domains and Dimensions	N	Mean (Standard Error)	Reported Response Range	Possible Response Range	Percentage Low (1 to 2)	Percentage Mid (3 to 5)	Percentage High (6 to 7)
CLASS Emotional Support*	76	5.9 (0.1)	3–7	1–7	0.0	45.9	54.0
Positive climate	81	5.9 (0.1)	3–7	1–7	0.0	27.9	72.1
Negative climate	82	1.2 (0.1)	1–4	1–7	97.3	2.6	0.0
Teacher sensitivity	83	5.4 (0.1)	3–7	1–7	0.0	49.9	50.1
Regard for student perspectives	79	5.4 (0.1)	3–7	1–7	0.0	50.5	49.5
CLASS Classroom Organization	79	5.4 (0.1)	2.33–6.67	1–7	0.7	72.7	26.6
Behavior management	83	5.5 (0.1)	2–7	1–7	0.6	41.0	58.4
Productivity	81	5.6 (0.1)	2–7	1–7	1.4	32.9	65.8
Instructional learning formats	80	5.1 (0.1)	2–7	1–7	1.3	59.3	39.3
CLASS Instructional Support	81	2.6 (0.1)	1–5.67	1–7	63.2	36.7	0.0
Concept development	82	2.1 (0.1)	1–5	1–7	73.6	26.3	0.0
Quality of feedback	81	2.5 (0.1)	1–6	1–7	60.0	39.3	0.7
Language modeling	82	3.4 (0.2)	1–7	1–7	34.5	51.6	13.8
Child/Adult Ratio	74	6.4 (0.3)	3.36–13.39	NA	NA	NA	NA
Class Size	85	17.7 (0.4)	9–28	NA	NA	NA	NA

Source: Classroom observations, winter 2008.

Note: We report here all classroom observations, including three conducted with assistant teachers.

* When constructing the emotional support scale, we reverse-coded the negative climate score (as instructed in the manual).

Table III.14. Relationship Between Class Scales and Teacher and Program Characteristics: Weighted OLS Results

CLASS Scale	Emotional Support			Instructional Support			Classroom Organization		
Model	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Star ratings	.28 (.14)*			.08 (.15)			.19 (.14)		
Teacher education (<AA)		.30 (.34)	.34 (.31)		-.99 (.36)**	-1.03 (.35)**		.21 (.34)	.22 (.34)
Teacher education (≥BA)		.39 (.19)*	.33 (.19)+		-.20 (.21)	-.25 (.21)		.14 (.19)	.04 (.20)
Teacher education (=AA) (reference)									
Years working with young children		.01 (.01)	.01 (.01)		.00 (.01)	-.01 (.01)		.02 (.01)*	.03 (.01)**
Group size		-.08 (.02)***	-.08 (.02)***		-.04 (.03)+	-.07 (.03)**		-.06 (.02)**	-.05 (.03)*
Child/adult ratio		.06 (.04)	.03 (.04)		-.02 (.04)	-.04 (.04)		.09 (.04)*	.06 (.04)
ECERS-R		.36 (.15)*	.46 (.16)**		.34 (.17)*	.29 (.17)+		.52 (.16)***	.63 (.17)***
Non-AGSN			-.25 (.21)			.60 (.24)*			-.39 (.22)+
Early Launch, AGSN			.37 (.21)+			.52 (.23)*			.10 (.22)
Second Round, AGSN (reference)									
Program SES			.00 (.10)			-.04 (.11)			-.11 (.10)
Auspice (School)			.27 (.19)			.42 (.21)*			.18 (.19)
R^2	.03*	.16***	.26***	.00	.14**	.22***	.01	.15***	.21***
Weighted N	151	129	129	161	136	136	158	137	137

Note: Model results reported as unstandardized regression coefficients and standard errors (in parentheses). CLASS scales are z-scored. We omitted scores for three classroom observations with assistant teachers in these analyses.

+p<.10; *p<.05. **p<.01; ***p≤.001.

Analyses were done at the classroom level.

Table III.15. Mother and Father Characteristics (Weighted)

Characteristic	Mother		Father	
	N	Mean/ Percentage (Standard Error)	N	Mean/ Percentage (Standard Error)
Mean Age	1,379	32.0 (0.3)	1,315	34.8 (0.3)
Race/Ethnicity				
White, non–Latino	119	8.6 (2.0)	132	9.5 (2.1)
African American, non–Latino	107	7.8 (2.0)	111	8.1 (1.9)
Hispanic/Latino	1,024	73.0 (3.1)	1,005	72.0 (3.1)
Asian/Pacific Islander	105	7.5 (1.3)	100	7.3 (1.4)
Multiple race/other	47	3.1 (0.7)	48	3.2 (0.6)
Education[a]				
Less than high school diploma/GED	478	35.2 (2.6)	529	39.6 (2.7)
High school diploma or GED	260	18.2 (1.2)	331	25.2 (1.5)
Some college/associate's degree	424	29.8 (1.8)	267	19.4 (1.1)
Bachelor's degree or more	233	16.8 (2.1)	214	15.9 (2.2)
Employment Status[a]				
Full-time	504	42.5 (2.7)	868	89.3 (1.0)
Part-time	225	18.3 (1.3)	82	7.8 (1.1)
Not employed	490	39.2 (2.6)	30	2.9 (0.5)
Born in the United States	531	36.7 (2.5)	466	32.5 (2.4)
Time in the United States if Born Elsewhere				
5 years or fewer	90	9.4 (1.1)	71	7.8 (1.1)
6 to 10 years	270	32.5 (2.0)	175	21.6 (1.7)
More than 10 years	509	58.1 (2.3)	642	70.6 (1.9)
Country of (non–United States) Origin				
Mexico	579	64.9 (3.1)	621	64.9 (3.1)
Central America	95	12.0 (1.8)	104	12.6 (1.6)
Asia	49	5.0 (1.6)	42	3.9 (1.4)
Southeast Asia	47	5.8 (1.4)	48	5.7 (1.4)
Armenia	19	2.1 (1.5)	24	2.7 (2.0)
Other	83	10.2 (1.7)	91	10.2 (1.5)

Source: Parent interview, fall 2007.

[a] Due to skip patterns, we asked about father's education and employment only if he lived in the household.

Table III.16. Child and Household Characteristics (Weighted)

Characteristic	N	Mean/Percentage (Standard Error)
Female	677	48.3 (1.3)
Mean Child Age in Spring (months)	1,438	59.0 (0.1)
Child Race/Ethnicity		
White, non–Latino	92	6.7 (1.8)
African American, non–Latino	107	7.6 (1.9)
Hispanic/Latino	1,058	74.9 (3.0)
Asian/Pacific Islander	94	6.8 (1.4)
Multiple race/other	53	4.0 (0.8)
Child Born in the United States	1,357	96.7 (0.5)
Child's First Language		
English only/primarily	588	41.3 (2.5)
Another language only or primarily	769	55.3 (2.5)
English and another language equally	49	3.4 (0.6)
Mean Number of Adults in Household	1,346	2.5 (0.0)
Mean Number of Children in Household	1,346	2.5 (0.0)
Mean Number of Persons in Household	1,346	4.9 (0.1)
Mean Household Dependency Ratio[a]	1,346	1.2 (0.0)
Child Living with		
Both parents	991	73.4 (1.8)
Mother only	317	23.6 (1.7)
Father only	19	1.5 (0.4)
Other	20	1.6 (0.4)
Child's Parents Are[b]		
Married	819	57.6 (2.3)
Divorced/separated	132	9.7 (1.1)
Not married	446	32.6 (2.1)
Both of Child's Parents Born in United States	370	25.8 (2.3)
Household Income as Percentage of Federal Poverty Level		
Below 50 percent	78	7.1 (1.0)
50–99 percent	348	31.8 (2.6)
100–129 percent	156	14.3 (1.5)
130–184 percent	174	15.0 (1.2)
185–239 percent	107	8.7 (1.0)
240 percent or more	265	23.2 (2.8)
Family Moved at Least Once in Last 12 Months	280	20.3 (1.4)
Family Moved More than Once in Last 12 Months	54	4.0 (0.6)

Source: Parent interview, fall 2007.

[a]Household dependency ratio is the ratio of children to adults in the household.
[b]Marital status reflects the marital status of the child's mother and father.

Table III.17.   Home Environment, Activities, and Routines (Weighted)

Characteristic	N	Mean/Percentage (Standard Error)
At Home, Child Is Usually Spoken to in		
English	666	46.5 (2.6)
Spanish	621	45.1 (2.9)
Another language	118	8.4 (1.7)
Parent Understands English		
Not at all	81	9.9 (1.5)
Not well	298	32.9 (2.2)
Well	246	29.6 (1.8)
Very well/native	249	27.6 (2.1)
When Parent Reads to Child, Language Is Usually		
English	740	56.4 (2.6)
Another language	351	27.2 (2.2)
English and another language equally	201	16.4 (1.3)
Frequency of Reading to Child in Past Week		
Never	39	2.8 (0.5)
Once or twice	247	18.3 (1.4)
Three or more times, but not every day	527	38.6 (1.5)
Every day	533	40.2 (2.0)
Number of Children's Books in Home		
0–10	362	27.2 (1.9)
11–25	354	25.7 (1.6)
26–50	392	29.9 (1.6)
51–100	183	13.1 (1.4)
101+	54	4.1 (0.7)
Mean Number of Days per Week Family Eats Dinner Together	1,346	5.5 (0.1)
Mean Number of Days per Week Child Goes to Bed at Regular Bedtime	1,339	4.5 (0.0)
Mean Number of Hours Child Sleeps per Night	1,339	10.4 (0.0)
Current Child Care Outside LAUP		
Attends child care center or formal program	61	4.6 (0.6)
Receives child care from relative	196	14.2 (1.1)
Receives child care from non-relative	45	3.4 (0.5)
Does not receive care outside LAUP	1,025	77.8 (1.3)
Mean Hours per Week in LAUP Program	1,395	23.6 (1.1)
Mean Hours per Week in non–LAUP Child Care[a]	319	14.9 (0.8)
Mean Hours in Out-of-Home Care (LAUP and non-LAUP)	1,395	27.0 (1.1)

Source: Parent interview, fall 2007.

[a]Only parents reporting child care outside LAUP indicated the number of hours per week in non–LAUP care.

Appendix C

Table III.18.   Family and Parent Well-Being Characteristics (Weighted)

Characteristic	N	Mean/Percentage (Standard Error)
Degree of Parent's Depressive Symptoms[a]		
Not depressed	973	73.3 (1.2)
Mildly depressed	233	17.2 (1.1)
Moderately depressed	72	5.3 (0.6)
Severely depressed	61	4.3 (0.6)
Mean Exposure to Crime and Violence Index[b]	1,403	0.8 (0.1)
Child's Health Insurance Status		
Private health insurance plan	557	41.3 (3.0)
Public/government insurance plan[c]	739	55.8 (2.7)
No health insurance	105	7.3 (0.9)
Child's Last Regular Doctor Checkup Was Less than 1 Year ago	1,309	97.4 (0.5)
Last Time Child Saw Dentist for Regular Checkup		
6 months ago or less	803	60.9 (1.9)
Between 6 months and a year	215	15.6 (1.3)
More than 1 year ago	82	5.9 (0.9)
Never	242	17.6 (1.5)

Source: Parent interview, fall 2007.

[a] The short version of the Center for Epidemiological Studies--Depression Short Form ([CES-D] Radloff 1977; Ross et al. 1983) measures levels of depressive symptoms by using 12 of the original 20 items from the full CES-D.   Four threshold scores are constructed: (1) not depressed--Short Form scores 0–4; (2) mildly depressed--Short Form scores 5–9; (3) moderately depressed--Short Form scores 10–14; and (4) severely depressed--Short Form scores 15 or greater. In the table, parent's depressive symptoms reflect the symptoms of the parent respondent, typically the child's mother.

[b] Exposure to crime and violence is an index reflecting the severity of families' exposure to crime and violence. Higher scores indicate that household members had been a victim of or were acquainted with someone who had been a victim of violent crime while lower scores reflect exposure to or witnessing non-violent crime. Scores greater than 0 indicate that household members had witnessed or had been a victim of any type of crime (i.e., violent and/or non-violent) within the last year. Scores range from 0–5.

[c] Public/government insurance includes Medi-Cal or Medicaid, Healthy Families, Healthy Kids, and/or military insurance.

Table III.19.   Mother Characteristics, By Language Group (Weighted)

Characteristic	English N	English Mean/ Percentage (Standard Error)	Spanish Only N	Spanish Only Mean/ Percentage (Standard Error)	Spanish Primarily N	Spanish Primarily Mean/ Percentage (Standard Error)	Other Language N	Other Language Mean/ Percentage (Standard Error)
Mean Age	736	32.0 (0.4)	194	31.5 (0.4)	383	31.9 (0.4)	66	34.3 (0.7)
Race/Ethnicity								
White, non-Latino	100	13.2 (2.7)	0	0.0 (0.0)	2	0.8 (0.5)	17	30.5 (15.6)
African American, non-Latino	107	14.7 (3.5)	0	0.0 (0.0)	0	0.0 (0.0)	0	0.0 (0.0)
Hispanic/Latino	439	58.3 (3.7)	198	100.0 (0.0)	387	99.2 (0.5)	0	0.0 (0.0)
Asian/Pacific Islander	72	10.1 (1.4)	0	0.0 (0.0)	0	0.0 (0.0)	33	46.3 (17.2)
Multiple race/other	30	3.7 (0.8)	0	0.0 (0.0)	0	0.0 (0.0)	17	23.2 (6.1)
Education								
Less than high school diploma/GED	114	16.9 (2.1)	151	76.4 (3.1)	210	53.9 (3.0)	3	5.7 (2.8)
High school diploma or GED	149	19.3 (1.7)	22	11.1 (2.2)	83	21.0 (2.3)	6	8.7 (4.6)
Some college/Assoc. degree	303	40.3 (2.2)	18	9.2 (2.0)	78	19.9 (2.7)	25	34.1 (5.6)
Bachelor's degree or more	174	23.5 (2.7)	7	3.3 (1.3)	19	5.3 (1.5)	33	51.5 (8.4)
Employment Status								
Full-time	338	52.8 (3.4)	48	30.0 (4.4)	105	32.7 (3.8)	13	25.6 (7.5)
Part-time	127	19.7 (1.6)	20	11.4 (2.1)	63	17.7 (2.6)	15	25.7 (4.2)
Not employed	176	27.5 (2.8)	108	58.6 (4.5)	177	49.5 (3.9)	29	48.7 (9.4)
Born in the United States	477	62.1 (2.6)	5	2.3 (1.2)	48	12.2 (1.5)	1	2.1 (1.6)
Time in the United States if Born Elsewhere								
5 years or fewer	16	5.3 (1.4)	36	18.4 (2.9)	24	6.4 (1.3)	14	18.4 (6.6)
6 to 10 years	42	16.8 (2.6)	85	45.0 (4.5)	121	37.5 (3.4)	22	36.5 (5.2)
More than 10 years	210	77.9 (2.9)	72	36.6 (4.6)	197	56.1 (3.0)	30	45.1 (8.5)
Country of (Non-US) Origin								
Mexico	121	43.8 (3.6)	169	86.4 (3.2)	289	82.1 (2.6)	0	0.0 (0.0)
Central America	30	11.3 (2.4)	21	11.9 (3.1)	44	14.8 (2.5)	0	0.0 (0.0)
Asia	21	7.3 (1.7)	0	0.0 (0.0)	0	0.0 (0.0)	28	37.0 (16.7)
Southeast Asia	41	15.7 (3.1)	0	0.0 (0.0)	0	0.0 (0.0)	6	11.1 (7.8)
Armenia	0	0.0 (0.0)	0	0.0 (0.0)	0	0.0 (0.0)	19	29.6 (14.9)
Other	58	22.0 (3.4)	3	1.7 (1.0)	9	3.1 (1.1)	13	22.4 (6.6)

Source: Parent interview, fall 2007.
Note: All comparisons across language groups are significant at *p* < .001

Table III.20.   Father Characteristics, By Language Group (Weighted)

Characteristic	English N	English Mean/Percentage (Standard Error)	Spanish Only N	Spanish Only Mean/Percentage (Standard Error)	Spanish Primarily N	Spanish Primarily Mean/Percentage (Standard Error)	Other Language N	Other Language Mean/Percentage (Standard Error)
Mean Age	704	34.9 (0.4)	183	34.3 (0.5)	364	34.3 (0.4)	64	38.5 (0.7)
Race/Ethnicity								
White, non-Latino	115	15.2 (2.8)	0	0.0 (0.0)	0	0.0 (0.0)	17	31.0 (15.8)
African American, non-Latino	111	15.3 (3.5)	0	0.0 (0.0)	0	0.0 (0.0)	0	0.0 (0.0)
Hispanic/Latino	418	55.8 (3.7)	198	100.0 (0.0)	389	99.7 (0.3)	0	0.0 (0.0)
Asian/Pacific Islander	66	9.4 (1.6)	0	0.0 (0.0)	0	0.0 (0.0)	34	49.1 (18.0)
Multiple race/other	32	4.2 (0.7)	0	0.0 (0.0)	1	0.3 (0.3)	15	19.9 (5.8)
Education[a]								
Less than high school diploma/GED	158	21.5 (2.3)	138	74.2 (3.6)	230	63.1 (3.3)	3	4.3 (3.0)
High school diploma or GED	197	27.9 (2.0)	30	16.3 (3.2)	91	24.6 (2.5)	13	22.4 (10.1)
Some college/Assoc. degree	202	27.2 (1.7)	16	8.8 (2.1)	36	9.6 (1.8)	13	20.9 (3.9)
Bachelor's degree or more	163	23.4 (3.0)	2	0.8 (0.5)	12	2.7 (0.8)	37	52.3 (12.2)
Employment Status[a]								
Full-time	421	89.8 (1.4)	141	87.9 (2.7)	270	91.7 (1.7)	36	74.2 (5.7)
Part-time	38	7.3 (1.4)	16	9.8 (2.6)	22	7.3 (1.4)	6	9.6 (4.8)
Not employed	14	3.0 (0.7)	5	2.2 (1.0)	3	1.0 (0.6)	8	16.1 (4.1)
Born in the United States	424	56.1 (2.7)	8	4.2 (1.5)	34	8.0 (1.4)	0	0.0 (0.0)
Time in the United States if Born Elsewhere								
5 years or fewer	14	4.6 (1.2)	26	14.2 (2.9)	19	5.7 (1.5)	12	16.6 (6.5)
6 to 10 years	36	14.3 (2.6)	46	24.6 (3.0)	78	25.7 (2.8)	15	26.0 (6.1)
More than 10 years	252	81.1 (2.7)	111	61.2 (3.9)	241	68.6 (2.7)	38	57.4 (9.1)
Country of (Non-US) Origin								
Mexico	160	48.6 (3.8)	161	82.3 (4.4)	300	81.8 (2.6)	0	0.0 (0.0)
Central America	33	10.9 (2.1)	25	16.4 (4.2)	46	14.6 (2.3)	0	0.0 (0.0)
Asia	14	0.0 (0.0)	0	0.0 (0.0)	0	0.0 (0.0)	28	36.5 (16.9)
Southeast Asia	41	0.0 (0.0)	0	0.0 (0.0)	0	0.0 (0.0)	7	13.0 (8.2)
Armenia	2	0.6 (0.5)	0	0.0 (0.0)	0	0.0 (0.0)	22	35.7 (18.8)
Other	69	21.9 (2.8)	3	1.4 (0.8)	10	3.6 (1.2)	9	14.7 (5.9)

Source: Parent interview, fall 2007.

Note: All comparisons across language groups are significant at $p < .001$, except employment status ($p < .05$)

[a]Due to skip patterns, we asked about father's education and employment only if he lived in the household.

Table III.21. Child and Household Characteristics, By Language Group (Weighted)

Characteristic	English N	English Mean/Percentage (Standard Error)	Spanish Only N	Spanish Only Mean/Percentage (Standard Error)	Spanish Primarily N	Spanish Primarily Mean/Percentage (Standard Error)	Other Language N	Other Language Mean/Percentage (Standard Error)
Female	352	47.6 (1.8)	102	51.5 (3.7)	191	47.9 (2.2)	32	50.2 (9.3)
Mean Child Age in Spring (Months)	772	59.1 (0.2)	199	58.9 (0.2)	394	59.0 (0.2)	73	58.9 (0.6)
Child Race/Ethnicity								
White, non-Latino	75	9.9 (2.5)	0	0.0 (0.0)	1	0.4 (0.4)	16	28.0 (14.0)
African American, non-Latino	105	14.2 (3.5)	0	0.0 (0.0)	0	0.0 (0.0)	2	2.4 (2.1)
Hispanic/Latino	473	62.4 (3.7)	198	100.0 (0.0)	387	98.5 (0.9)	0	0.0 (0.0)
Asian/Pacific Islander	60	8.5 (1.5)	0	0.0 (0.0)	0	0.0 (0.0)	34	47.9 (16.8)
Multiple race/other	35	5.0 (1.0)	0	0.0 (0.0)	2	1.1 (0.8)	16	21.7 (6.9)
Child Born in the United States	739	98.6 (0.4)	176	89.2 (2.3)	384	98.5 (0.6)	58	86.5 (5.6)
Mean Number of Adults in Household	718	2.3 (0.1)	190	2.8 (0.1)	376	2.5 (0.1)	62	2.2 (0.1)
Mean Number of Children in Household	718	2.4 (0.0)	190	2.6 (0.1)	376	2.6 (0.1)	62	2.1 (0.1)
Mean Number of Persons in Household	718	4.7 (0.1)	190	5.4 (0.2)	376	5.2 (0.1)	62	4.3 (0.1)
Mean Household Dependency Ratio[a]	718	1.2 (0.0)	190	1.1 (0.0)	376	1.2 (0.0)	62	1.0 (0.1)
Child Is Living with								
Both parents	470	65.6 (2.9)	165	86.6 (2.7)	304	80.3 (2.2)	52	80.4 (6.1)
Mother only	218	30.1 (2.8)	24	12.8 (2.7)	67	18.4 (2.2)	8	13.1 (4.1)
Father only	17	2.5 (0.7)	1	0.5 (0.5)	1	0.3 (0.3)	0	0.0 (0.0)
Other	13	1.9 (0.6)	0	0.0 (0.0)	4	1.0 (0.5)	3	6.5 (3.0)
Child's Parents Are[b]								
Married	425	57.5 (3.1)	111	52.4 (4.3)	225	55.8 (2.7)	58	85.7 (4.4)
Divorced/separated	92	12.5 (1.7)	9	5.2 (1.9)	27	7.6 (1.6)	4	4.5 (1.9)
Not married	225	30.0 (2.7)	78	42.4 (4.1)	138	36.5 (2.7)	5	9.8 (3.7)

Table III.21 *(continued)*

Characteristic	English N	English Mean/Percentage (Standard Error)	Spanish Only N	Spanish Only Mean/Percentage (Standard Error)	Spanish Primarily N	Spanish Primarily Mean/Percentage (Standard Error)	Other Language N	Other Language Mean/Percentage (Standard Error)
Both of Child's Parents Born in US	359	47.4 (2.8)	1	0.4 (0.4)	10	2.7 (0.9)	0	0.0 (0.0)
Household Income as Percentage of Federal Poverty Level								
Below 50 percent	25	4.4 (1.1)	19	11.6 (2.6)	32	10.7 (2.0)	2	2.3 (1.4)
50 to 99 percent	101	17.1 (2.0)	89	59.9 (4.4)	144	46.5 (3.8)	14	24.6 (9.8)
100 to 129 percent	65	12.3 (1.9)	18	11.0 (2.5)	65	19.1 (2.6)	8	14.8 (5.2)
130 to 184 percent	99	15.7 (1.7)	19	12.2 (2.8)	48	14.9 (2.6)	8	14.8 (5.4)
185 to 239 percent	77	11.6 (1.4)	6	3.5 (1.3)	17	4.6 (1.2)	7	16.4 (5.2)
240 percent or more	233	38.9 (3.9)	3	1.8 (1.0)	14	4.3 (1.2)	15	27.1 (5.5)
Family Moved at Least Once in Last 12 Months	144	19.7 (1.7)	46	26.1 (3.6)	79	19.3 (2.2)	11	17.7 (7.2)
Family Moved More than Once in Last 12 Months	24	3.0 (0.6)	11	7.5 (2.1)	19	4.8 (1.3)	0	0.0 (0.0)

Source: Parent interview, fall 2007.

Note: All comparisons across language groups are significant at *p* < .001, except "family moved at least once in the past 12 months."

[a]Household dependency ratio is the ratio of children to adults in the household.
[b]Marital status reflects the marital status of the child's mother and father.

Table III.22.   Home Environment, Activities, and Routines, By Language Group (Weighted)

Characteristic	English N	English Mean/ Percentage (Standard Error)	Spanish Only N	Spanish Only Mean/ Percentage (Standard Error)	Spanish Primarily N	Spanish Primarily Mean/ Percentage (Standard Error)	Other Language N	Other Language Mean/ Percentage (Standard Error)
At Home, Child Is Usually Spoken to in								
English	611	81.0 (2.0)	6	2.5 (0.9)	44	10.4 (1.6)	5	7.0 (3.8)
Spanish	94	12.9 (1.7)	191	96.9 (1.1)	336	87.1 (1.7)	0	0.0 (0.0)
Another language	44	6.1 (1.2)	1	0.6 (0.6)	11	2.5 (1.0)	62	93.0 (3.8)
Parent Understands English								
Not at all	2	0.7 (0.5)	47	25.2 (3.5)	32	11.3 (2.5)	0	0.0 (0.0)
Not well	25	10.2 (2.6)	107	58.0 (3.6)	153	40.8 (3.1)	13	18.4 (10.7)
Well	84	30.6 (2.9)	27	13.9 (2.9)	111	33.5 (2.8)	24	49.3 (7.9)
Very well/native	173	58.5 (3.4)	5	3.0 (1.3)	53	14.3 (2.1)	18	32.3 (8.9)
When Parent Reads to Child, Language Is Usually								
English	569	85.7 (1.6)	20	10.2 (2.2)	125	31.2 (2.6)	26	37.8 (5.8)
Another language	20	3.1 (0.9)	150	77.8 (2.8)	162	44.2 (2.6)	19	29.9 (4.5)
English and another language equally	69	11.3 (1.5)	21	12.0 (2.7)	89	24.6 (2.2)	22	32.4 (6.3)
Frequency of Reading to Child in Past Week								
Never	9	1.2 (0.5)	18	8.2 (2.0)	12	3.7 (1.1)	0	0.0 (0.0)
Once or twice	97	14.3 (1.7)	60	30.0 (4.0)	77	20.2 (2.5)	13	19.2 (3.8)
Three or more times, but not every day	284	38.2 (2.2)	64	35.6 (3.4)	160	42.6 (3.2)	19	27.0 (5.7)
Every day	328	46.3 (2.6)	48	26.3 (4.4)	127	33.5 (3.7)	30	53.8 (7.4)
Number of Children's Books in Home								
0 – 10	104	15.2 (1.7)	111	59.1 (2.7)	140	36.8 (2.6)	7	10.6 (2.7)
11-25	156	20.9 (1.9)	54	26.9 (2.7)	126	33.8 (3.1)	18	25.3 (5.2)
26-50	269	38.2 (1.8)	21	10.7 (2.2)	84	22.8 (2.7)	18	35.6 (6.9)
51-100	143	19.1 (2.0)	3	2.9 (1.8)	21	5.3 (1.2)	16	24.6 (6.0)
101+	45	6.6 (1.1)	1	0.4 (0.4)	5	1.3 (0.7)	3	3.9 (6.0)
Mean Number of Days per Week Family Eats Dinner Together	718	5.4 (0.1)	190	5.7 (0.1)	376	5.8 (0.1)	62	5.0 (0.3)
Mean Number of Days per Week Child Goes to Bed at Regular Bedtime	713	4.5 (0.0)	189	4.6 (0.1)	375	4.6 (0.0)	62	4.7 (0.1)
Mean Number of Hours Child Sleeps per Night	713	10.4 (0.0)	190	10.5 (0.1)	374	10.4 (0.1)	62	10.4 (0.2)

Table III.22 *(continued)*

Characteristic	English		Spanish Only		Spanish Primarily		Other Language	
	N	Mean/ Percentage (Standard Error)	N	Mean/ Percentage (Standard Error)	N	Mean/ Percentage (Standard Error)	N	Mean/ Percentage (Standard Error)
Current Child Care Outside of LAUP								
Attends child care center or formal program	36	5.0 (0.9)	8	4.8 (1.6)	11	3.1 (1.1)	6	8.4 (5.9)
Receives child care from relative	124	16.6 (1.5)	18	10.9 (2.9)	46	11.1 (1.5)	8	16.0 (6.6)
Receives child care from nonrelative	24	3.4 (0.8)	5	2.4 (1.1)	15	3.8 (0.9)	1	2.6 (2.6)
Does not receive care outside of LAUP	522	74.9 (1.8)	156	81.8 (3.0)	302	81.9 (2.3)	45	73.0 (5.8)
Mean Hours per Week in LAUP Program	739	25.7 (1.3)	198	19.6 (1.3)	390	21.7 (1.4)	68	24.0 (1.6)
Mean Hours per Week in non–LAUP Child Care[a]	194	15.5 (1.0)	34	14.4 (2.3)	74	13.8 (1.2)	17	13.6 (3.8)
Mean Hours in Out-of-Home Care (LAUP and non-LAUP)	739	29.7 (1.2)	198	22.2 (1.4)	390	24.1 (1.3)	68	27.6 (1.6)

Source: Parent interview, fall 2007.

Note: All comparisons across language groups are significant at $p < .001$, except days family eats dinner together ($p < .01$), child care outside LAUP ($p < .05$) and days with regular bedtime, hours child sleeps per night, and hours per week in non-LAUP care (nonsignificant).

[a]Only parents who reported having child care outside of LAUP indicated the number of hours per week in non-LAUP care.

Table III.23.   Family and Parent Well-Being Characteristics, by Language Group (Weighted)

	English		Spanish-Only		Spanish-Primarily		Other Language	
Scales	N	Mean/Percentage (Standard Error)	N	Mean/Percentage (Standard Error)	N	Mean/Percentage (Standard Error)	N	Mean/Percentage (Standard Error)
Degree of Parent's Depressive Symptoms[a]								
Not depressed	509	71.4 (1.7)	138	72.2 (1.7)	283	77.2 (2.3)	43	74.2 (7.8)
Mildly depressed	139	19.8 (1.7)	29	15.4 (2.4)	55	13.4 (1.6)	10	15.2 (3.7)
Moderately depressed	38	5.3 (1.0)	9	4.1 (1.4)	20	5.3 (1.3)	5	8.0 (4.9)
Severely depressed	27	3.5 (0.8)	14	8.2 (2.3)	17	4.1 (1.1)	3	2.6 (1.6)
Mean Exposure to Crime and Violence Index[b]	748	0.9 (0.1)	198	0.6 (0.1)	390	0.9 (0.1)	67	0.2 (0.1)
Child's Health Insurance Status								
A private health insurance plan	441	62.1 (3.3)	16	8.1 (2.1)	73	18.1 (2.4)	27	44.7 (4.6)
A public/government insurance plan[c]	265	37.8 (2.9)	157	83.7 (2.6)	283	76.1 (3.2)	34	52.4 (7.8)
No health insurance	48	5.9 (1.0)	20	9.5 (2.6)	31	8.3 (1.8)	6	9.8 (4.9)
Child's Last Regular Doctor Checkup Was Less than 1 Year Ago	698	97.6 (0.7)	183	96.9 (1.1)	368	97.4 (1.0)	60	96.1 (2.6)
Last Time Child Saw Dentist for Regular Checkup								
6 months ago or less	397	57.1 (2.4)	126	68.0 (3.7)	242	64.2 (3.0)	38	64.0 (8.4)
Between 6 months and a year	133	17.7 (1.8)	17	8.1 (2.0)	59	16.1 (2.4)	6	9.7 (3.4)
More than 1 year ago	37	5.0 (0.9)	18	9.2 (2.3)	22	5.5 (1.6)	5	8.8 (4.6)
Never	148	20.3 (1.9)	28	14.6 (2.6)	53	14.2 (2.2)	13	17.5 (6.2)

Source: Parent interview, fall 2007.

Note: All fall to spring comparisons within language group are significant at $p < .001$, except parent depression, and doctor visit.

[a]The short version of the Center for Epidemiological Studies-Depression Short Form ([CES-D] Radloff 1977; Ross et al. 1983) measures levels of depressive symptoms using 12 of the original 20 items from the full CES-D. Four threshold scores are constructed: (1) not depressed—Short Form scores from 0-4, (2) mildly depressed—Short Form scores from 5-9, (3) moderately depressed—Short Form scores from 10-14, and (4) severely depressed—Short Form scores of 15 or greater. In the table, parent's depressive symptoms reflects the symptoms of the parent respondent, typically the child's mother.

[b]Exposure to crime and violence is an index that reflects the severity of families' exposure to crime and violence. Higher scores indicate that household members had been a victim of or were acquainted with someone who had been a victim of violent crime, while lower scores reflect exposure or witnessing of nonviolent crime. Scores greater than 0 indicate that household members had witnessed or been a victim of any sort of crime (i.e., violent and/or non-violent) within the last year. Scores range from 0 to 5.

[c]Public/government insurance includes Medi-Cal or Medicaid, Healthy Families, Healthy Kids, and/or military insurance..

Table III.24. Language, Literacy, and Mathematics Development Fall 2007 and Spring 2008 (Weighted)

Developmental Domain	Data Source	Fall 2007			Spring 2008		
		N	Mean	Standard Error	N	Mean	Standard Error
Language and Literacy							
	Pre–LAS total language screener score English[a] ***	1,438	12.9	0.4	483	12.8	0.3
	Pre–LAS total language screener score Spanish[a] ***	593	11.1	0.3	410	12.7	0.6
	Expressive One Word Picture Vocabulary Test-- Spanish Bilingual Edition IRT Scale Score ***	1,427	43.3	0.6	1,436	49.2	0.6
	Receptive One Word Picture Vocabulary Test-- English Edition IRT Scale Score[b] ***	575	49.8	1.0	579	58.3	0.8
	Receptive One Word Picture Vocabulary Test-- Spanish Bilingual Edition IRT Scale Score[c] ***	856	47.0	0.8	858	49.6	0.6
	Rapid Letter Naming IRT Scale Score ***	1,424	17.2	0.6	1,435	30.1	0.6
	Woodcock-Johnson III Spelling Standard Score[d] ***	892	101.1	0.7	892	108.1	0.7
	Woodcock-Muñoz-III Spelling Standard Score[d] ***	247	86.2	0.9	247	87.9	0.8
Mathematics							
	ECLS-B Mathematics W Score ***	1,426	489.1	0.6	1,436	499.7	0.6

Source: Direct child assessments, fall 2007 and spring 2008.

[a] These scores are raw counts of children's correct responses on Simon Says (Tío Simón Dice) and Art Show (Exposición de Arte).

[b] The means reported here are for the children who took the English version in the fall and their corresponding spring scores (all children took the English version in the spring).

[c] This measure was conceptually scored in the fall and administered in English in the spring only. The means reported here are for children who took the Spanish version in the fall and their corresponding spring scores on the English version.

[d] National mean for standard scores is 100 with a standard deviation of 15. Some children were administered the Spelling subtest in Spanish in the fall and in English in the spring. Sixty-six percent of the children were tested in English, using the Woodcock-Johnson, both fall and spring, 18 percent were tested in Spanish with the Woodcock-Muñoz at both times, and 15 percent switched languages between fall and spring. The means reported here are for children taking the same version fall and spring. This excludes 206 children who took the Spanish version in the fall, but were able to take the English version in the spring. The average spring score for these children was 105.2.

*** $p < .001$

Table III.25   Social-Emotional and Approaches to Learning Development Fall 2007 and Spring 2008 (Weighted)

Developmental Domain	Data Source	Fall 2007			Spring 2008		
		N	Mean or Percentage	Standard Error	N	Mean or Percentage	Standard Error
Social-Emotional							
	Teacher Report						
	Preschool Kindergarten Behavior Scale (PKBS)						
	Total Positive Behavior Standard Score[a] ***	1,388	97.9	0.9	1,331	105.2	1.0
	Social Cooperation ***	1,388	101.1	0.8	1,331	105.6	0.9
	Social Interaction ***	1,388	94.9	1.2	1,331	103.4	1.0
	Social Independence ***	1,388	99.9	0.8	1,331	104.9	0.8
	Social Skills Rating System Problem Behaviors						
	Standard Score***	1,386	96.8	0.7	1,329	95.4	0.7
	Parent Report						
	Preschool Kindergarten Behavior Scale						
	IRT Scale Score[b]						
	Social Cooperation/Approaches to Learning***	1,346	2.9	0.0	1,301	3.0	0.0
	Social Interaction/Independence ***	1,346	3.1	0.0	1,301	3.2	0.0
	Externalizing Problem Behaviors ***	1,346	2.0	0.0	1,301	2.0	0.0
	Internalizing Problem Behaviors ***	1,346	2.2	0.0	1,301	2.2	0.0
Approaches to Learning							
	Direct Child Assessment						
	Executive Functioning						
	Pencil tapping[c] ***	1,435	43.3%	1.3	1,435	66.9%	1.3
	Walk-a-Line Slowly--difference						
	between first attempt and slow attempt[d] ***	1,349	1.2 sec.	0.2	1,434	2.3 sec.	0.2
	Leiter Examiner Ratings Scaled Score						
	Attention[e] ***	1,438	7.8	0.1	1,438	8.9	0.1
	Activity[e] ***	1,438	8.2	0.1	1,438	8.9	0.1
	Sociability[e] ***	1,438	8.2	0.1	1,438	9.0	0.1
	ECLS-K Approaches to Learning Scale						
	Teacher-Reported Raw Score[f] ***	1,387	1.9	0.0	1,330	2.2	0.0

Source: Teacher and parent interviews, direct child assessments and observer ratings, fall 2007 and spring 2008.

[a] National mean for standard scores is 100 with a standard deviation of 15.
[b] IRT score rescaled to match the 1–4 response scale.
[c] Pencil Tapping is percentage of the time the child responded correctly.
[d] Mean score for Walk a Line task is in seconds.
[e] National mean for scaled scores is 10 with a standard deviation of 2.
[f] Raw score scaled to reflect 0–3 response scale

*** $p<.001$

Table III.26. **Physical/Motor Development Fall 2007 and Spring 2008 (Weighted)**

Developmental Domain	Data Source	Fall 2007		Spring 2008	
		N	Percentage (Standard Error)	N	Percentage (Standard Error)
Motor Development					
	Walk-a-Line–stayed on line***	1,026	71.3 (2.4)	1,158	80.9 (2.0)
Physical Health/ Development					
	Direct Child Assessment				
	Body mass index (obesity indicator)***	311	21.9 (1.6)	252	20.1 (1.3)
	Parent Report				
	Child good or excellent health***	1,290	96.0 (0.5)	1,248	96.3 (0.6)

Source: Direct child assessments and parent interviews, fall 2007 and spring 2008.

*** $p<.001$

Developmental Domain	Data Source	Lowest Quartile						Highest Quartile					
		Fall			Spring			Fall			Spring		
		N	Mean	Standard Error	N	Mean	Standard Error	N	Mean	Standard Error	N	Mean	Standard Error
Language and Literacy													
	Pre-LAS total language screener score English[a] ***	370	4.0	0.2	304	11.1	0.3	410	18.9	0.1	1	17.0	0.0
	Pre-LAS total language screener score Spanish[a] ***	167	5.2	0.2	119	10.2	0.6	197	15.9	0.2	124	15.1	0.4
	Expressive One Word Picture Vocabulary Test–Spanish Bilingual Edition IRT Scale Score ***	363	29.7	0.4	363	39.5	0.7	366	56.2	0.4	366	59.5	0.7
	Receptive One Word Picture Vocabulary Test–English Edition IRT Scale Score[b] ***	144	35.1	1.6	144	51.4	1.5	151	61.5	0.4	151	65.7	0.9
	Receptive One Word Picture Vocabulary Test– Spanish Bilingual Edition IRT Scale Score[b] ***	215	28.9	1.0	215	40.1	1.0	215	62.9	0.4	215	57.5	1.0
	Rapid Letter Naming IRT Scale Score ***	404	2.5	0.1	403	19.8	0.7	360	35.3	0.3	360	42.0	0.4
	Woodcock-Johnson III Spelling Standard Score[c] ***	223	80.5	0.7	223	99.2	1.2	238	120.5	0.4	238	117.8	0.8
	Woodcock-Muñoz-III Spelling Standard Score[c] ***	66	69.2	1.6	66	81.5	1.4	67	98.5	0.6	67	94.3	1.7
Mathematics													
	ECLS-B Mathematics IRT Scale Score ***	429	476.5	0.3	429	491.5	0.5	397	502.6	0.5	397	510.0	0.8

Source: Direct child assessments, fall 2007 and spring 2008.

[a] These scores are raw counts of children's correct responses on Simon Says (Tío Simón Dice) and Art Show (Exposición de Arte). Those who passed the two subscales of the Pre-LAS in the fall, did not take it again in the spring.

[b] The means reported here are for the children who took the English version in the fall and their corresponding spring scores (all children took the English version in the spring).

[c] National mean for standard scores is 100 with a standard deviation of 15. Some children were administered the Spelling subtest in Spanish in the fall and in English in the spring. The means reported here are for children taking the same version fall and spring. This excludes 55 children in the lowest quartile and 53 children in the highest quartile who took the Spanish version in the fall, but were able to take the English version in the spring. The average spring score for these children was 98.7 for the lowest quartile and 110.6 for the highest quartile.

*** $p<.001$

Table III.28. Social-Emotional and Approaches to Learning Development Scores, by Quartile Group Fall 2007–Spring 2008

Developmental Domain	Data Source	Lowest Quartile						Highest Quartile					
		Fall			Spring			Fall			Spring		
		N	Mean	Standard Error	N	Mean	Standard Error	N	Mean	Standard Error	N	Mean	Standard Error
Social-Emotional													
Teacher Report Preschool Kindergarten Behavior Scale (PKBS)[a] Total Positive Social Skills		362	75.5	(0.6)	345	91.8	(1.3)	357	117.6	(0.4)	322	115.4	(1.1)
Social Cooperation–Standard Score		364	77.6	(0.9)	339	91.3	(1.5)	364	119.1	(0.1)	331	116.2	(0.6)
Social Interaction–Standard Score		372	71.8	(1.1)	356	90.3	(1.6)	350	115.9	(0.4)	307	113.4	(1.1)
Social Independence– Standard Score		363	78.2	(0.6)	341	94.1	(1.4)	400	115.5	(0.2)	366	111.9	(0.8)
Social Skills Rating System Problem Behaviors–Standard Score		470	85.0	(0.0)	417	88.3	(0.5)	429	113.3	(0.6)	405	105.1	(1.1)
Approaches to Learning													
Direct Child Assessment Executive Functioning Pencil Tapping[b]		408	6.0	(0.3)	405	53.4	(2.3)	384	89.1	(0.5)	384	87.6	(1.1)
Leiter Examiner Ratings Attention—scaled score[c]		415	4.5	(0.1)	415	8.1	(0.1)	648	10.0	(0.0)	648	9.4	(0.1)
Activity—scaled score[c]		399	5.4	(0.1)	399	8.4	(0.1)	643	10.0	(0.0)	643	9.2	(0.1)
Social—scaled score[c]		367	4.8	(0.1)	367	8.5	(0.1)	749	10.0	(0.0)	749	9.4	(0.1)
ECLS-K Approaches to Learning Scale teacher-reported raw score[d]		367	0.9	(0.0)	341	1.5	(0.1)	280	2.9	(0.0)	252	2.7	(0.0)

Source: Teacher and parent interviews, direct child assessments and observer ratings, fall 2007 and spring 2008.

Note: All fall to spring comparisons are significant at *p* < .01

[a] National mean for standard scores is 100 with a standard deviation of 15.

[b] Pencil Tapping is percentage of the time the child responded correctly.

[c] National mean for scaled scores is 10 with a standard deviation of 2.

[d] Raw score scaled to reflect 0–3 response scale.

Table III.29.   Language, Literacy, and Mathematics Scores, by Language Group Fall 2007–Spring 2008 (Weighted)

Developmental Domain	Data Source	English Only or Primarily		Spanish Only		Spanish Primarily		Other Language Only or Primarily	
		Fall Mean (Standard Error)	Spring Mean (Standard Error)	Fall Mean (Standard Error)	Spring Mean (Standard Error)	Fall Mean (Standard Error)	Spring Mean (Standard Error)	Fall Mean (Standard Error)	Spring Mean (Standard Error)
Language and Literacy									
Direct Child Assessment									
Pre-LAS total language screener score English[a]		16.2 (0.2)	15.9 (0.5)	5.3 (0.4)	11.4 (0.5)	10.6 (0.4)	13.2 (0.3)	10.6 (1.3)	9.1 (2.0)
Pre-LAS total language screener score Spanish[a]		n.a.	n.a.	11.5 (0.4)	13.4 (0.3)	10.9 (0.4)	12.2 (0.5)	n.a.	n.a.
Expressive One Word Picture Vocabulary Test–SBE–IRT Scale Score		47.9 (0.5)	54.3 (0.6)	37.0 (0.7)	41.2 (0.7)	38.5 (0.7)	44.0 (0.7)	36.3 (1.9)	44.7 (1.5)
Receptive One Word Picture Vocabulary Test–English Edition IRT Scale Score[b]		51.5 (0.7)	59.5 (0.7)	n.a.	41.9 (1.2)	n.a.	48.6 (0.7)	37.8 (3.1)	50.1 (1.5)
Receptive One Word Picture Vocabulary Test–Spanish Bilingual Edition IRT Scale Score[c]		51.7 (0.9)	56.8 (0.7)	41.1 (1.2)	n.a.	46.6 (1.0)	n.a.	n.a.	n.a.
Rapid Letter Naming–IRT Score[d]		20.5 (0.7)	32.7 (0.6)	10.3 (0.9)	25.0 (1.2)	13.0 (0.7)	26.5 (0.8)	24.8 (2.8)	36.8 (1.5)
Woodcock-Johnson III Spelling Standard Score[e]		100.5 (0.8)	108.3 (0.8)	99.7 (4.9)	103.3 (4.2)	100.7 (1.2)	106.1 (1.3)	107.0 (3.1)	112.2 (3.2)
Woodcock-Muñoz-III Spelling Standard Score[e]		84.3 (3.7)	91.3 (3.9)	86.4 (1.2)	87.4 (1.3)	86.1 (1.4)	87.9 (1.2)	n.a.	n.a.
Mathematics									
Direct Child Assessment									
ECLS-B Mathematics W Score		492.1 (0.7)	502.0 (0.7)	483.0 (0.6)	493.7 (0.7)	486.2 (0.6)	495.9 (0.7)	491.4 (1.3)	504.3 (1.3)
Sample Size		**59–769**	**14–772**	**21–199**	**80–199**	**153–394**	**127–394**	**61–73**	**11–73**

Source: Direct child assessments, fall 2007 and spring 2008.

Note: All fall and spring comparisons are statistically significant at $p < 0.001$, except the ECLS-B mathematics score for the Spanish-only and Spanish-primarily groups

[a] These scores are raw counts of children's correct responses on Simon Says (Tío Simón Dice) and Art Show (Exposición de Arte).
[b] The means reported here are for the children who took the English version in the fall and their corresponding spring scores (all children took the English version in the spring).
[c] Receptive One Word Picture Vocabulary Test was administered only in English in the spring. Some children who were classified as speaking English primarily, but had Spanish language in the home took the Spanish Bilingual edition of the ROWPVT in the fall, and, as did all the children, the English version in the spring.
[d] Means are average number of letters named correctly out of 52. Children did not see all 52 letters. IRT was used to estimate the number of letters.
[e] National mean for standard scores is 100 with a standard deviation of 15. Some children were administered the Spelling subtest in Spanish in the fall and in English in the spring. Sixty-six percent of the children were tested in English, using the Woodcock-Johnson both fall and spring, 18 percent were tested in Spanish with the Woodcock-Muñoz at both times, and 15 percent switched languages between fall and spring. The means reported here are for children taking the same version fall and spring. This excludes 206 children who took the Spanish version in the fall, but were able to take the English version in the spring.

Table III.30.   Social-Emotional and Approaches to Learning Development Scores, by Language Group Fall 2007–Spring 2008 (Weighted)

Developmental Domain	Data Source	English		Spanish Only		Spanish Primarily		Other Language	
		Fall Mean (Standard Error)	Spring Mean (Standard Error)	Fall Mean (Standard Error)	Spring Mean (Standard Error)	Fall Mean (Standard Error)	Spring Mean (Standard Error)	Fall Mean (Standard Error)	Spring Mean (Standard Error)
Social-Emotional									
	Teacher Report								
	Preschool Kindergarten Behavior Scale (PKBS)[a]								
	Total Positive Social Skills	99.5 (1.1)	106.1 (1.3)	93.5 (1.9)	102.0 (1.6)	97.5 (1.1)	105.5 (1.1)	95.5 (2.0)	104.0 (3.0)
	Social Cooperation–standard score	101.5 (1.0)	105.4 (1.2)	99.2 (1.8)	104.6 (1.5)	101.4 (1.1)	106.1 (1.1)	99.7 (2.4)	106.5 (2.5)
	Social Interaction–standard score	97.7 (1.1)	105.2 (1.2)	88.7 (2.6)	98.6 (1.8)	93.4 (1.5)	102.8 (1.2)	91.5 (2.4)	101.8 (3.1)
	Social Independence– standard score	100.2 (1.0)	105.5 (1.1)	95.5 (1.5)	102.0 (1.3)	99.0 (0.9)	105.8 (1.0)	97.0 (1.9)	102.5 (2.6)
	Social Skills Rating System Problem Behaviors– standard score	98.0 (0.9)	96.2 (0.9)	94.9 (1.0)	93.9 (1.0)	95.1 (0.8)	94.6 (0.9)	98.9 (1.1)	94.6 (2.3)
	Parent Report								
	Preschool Kindergarten Behavior Scale[b]								
	Social Cooperation/Approaches to Learning–IRT score	2.9 (0.0)	3.0 (0.0)	2.8 (0.0)	3.0 (0.0)	2.8 (0.0)	3.0 (0.0)	3.0 (0.0)	3.1 (0.1)
	Social Interaction/Independence– IRT score	3.1 (0.0)	3.2 (0.0)	2.9 (0.0)	3.1 (0.0)	3.0 (0.0)	3.2 (0.0)	3.0 (0.1)	3.2 (0.1)
	Externalizing Problem Behaviors–IRT score	2.1 (0.0)	2.1 (0.0)	2.0 (0.0)	1.8 (0.0)	2.0 (0.0)	1.9 (0.0)	2.0 (0.0)	2.0 (0.0)
	Internalizing Problem Behaviors–IRT score	2.2 (0.0)	2.2 (0.0)	2.2 (0.0)	2.2 (0.0)	2.2 (0.0)	2.2 (0.0)	2.1 (0.0)	2.2 (0.0)

Table III.30. *(continued)*

Developmental Domain	Data Source	English		Spanish Only		Spanish Primarily		Other Language	
		Fall Mean (Standard Error)	Spring Mean (Standard Error)	Fall Mean (Standard Error)	Spring Mean (Standard Error)	Fall Mean (Standard Error)	Spring Mean (Standard Error)	Fall Mean (Standard Error)	Spring Mean (Standard Error)
Approaches to Learning									
	Direct Child Assessment								
	Executive Functioning								
	Pencil Tapping [c]	47.5	70.0	37.0	58.4	37.8	62.8	45.9	79.8
	Walk-a-Line Slowly–difference between first attempt and slow attempt [d]	1.7 sec.	3.1 sec.	0.1 sec.	0.9 sec.	0.5 sec.	1.4 sec.	2.1 sec.	4.0 sec.
	Leiter Examiner Ratings								
	Attention–scaled score [e]	8.1 (0.1)	9.1 (0.1)	7.4 (0.3)	8.5 (0.2)	7.6 (0.2)	8.9 (0.1)	7.2 (0.3)	9.1 (0.3)
	Activity–scaled score [e]	8.2 (0.1)	9.0 (0.1)	8.0 (0.2)	8.6 (0.2)	8.2 (0.1)	8.9 (0.1)	8.2 (0.3)	8.9 (0.3)
	Social–scaled score [e]	8.5 (0.1)	9.1 (0.1)	7.6 (0.2)	8.8 (0.2)	7.9 (0.2)	9.0 (0.1)	8.4 (0.2)	9.1 (0.2)
	ECLS-K Approaches to Learning Scale teacher-reported raw score [f]	1.9 (0.1)	2.2 (0.1)	1.8 (0.1)	2.1 (0.1)	1.9 (0.1)	2.2 (0.1)	2.0 (0.1)	2.3 (0.2)
Sample Size		**718–772**	**691–772**	**183–199**	**184–199**	**373–392**	**364–394**	**62–73**	**57–73**

Source: Teacher and parent interviews, direct child assessments, and observer ratings, fall 2007 and spring 2008.

Note: All fall to spring comparisons are statistically significant at $p < 0.001$.

[a] National mean for standard scores is 100 with a standard deviation of 15.

[b] IRT score rescaled to match the 1–4 response scale.

[c] Pencil Tapping is percentage of time the child responded correctly.

[d] Mean score for Walk-a-Line task is in seconds.

[e] National mean for scaled scores is 10 with a standard deviation of 2.

[f] Raw score scaled to reflect 0–3 response scale.

Table III.31.   Physical/Motor Development Scores, by Language Group Fall 2007– Spring 2008 (Weighted)

Developmental Domain	Data Source	English		Spanish Only		Spanish Primarily		Other Language	
		Fall Percent-age	Spring Percent-age	Fall Percent-age	Spring Percent-age	Fall Percent-age	Spring Percent-age	Fall Percent-age	Spring Percent-age
	Direct Child Assessment								
	Walk-a-Line Slowly–stayed on line	70.7	82.2	70.6	78.4	74.1	79.4	63.0	79.4
Physical Health/Development									
	Direct Child Assessment								
	Body mass index (obesity indicator)	19.7	17.9	24.9	22.9	25.1	22.9	19.8	19.3
	Parent Report								
	Child good or excellent health	97.5	96.9	90.7	93.8	95.8	95.9	94.7	100.0
Sample Size		**148–698**	**177–667**	**47–172**	**51–171**	**105–360**	**104–353**	**11–60**	**15–59**

Source: Teacher and parent interviews, direct child assessments, and observer ratings, fall 2007 and spring 2008.

Note: All fall to spring comparisons are statistically significant at *p* < .001.

Table III.32.   Frequency of Parent Involvement with LAUP, Parent Report (Weighted)

	Never	Once or Twice a Year	Several Times a Year	About Once a Month	At Least Once a Week	Not Offered
	Percentage (Standard Error)	Percentage (Standard Error)	Percentage (Standard Error)	Percentage (Standard Error)	Percentage (Standard Error)	Percentage (Standard Error)
Attended parent-teacher conference	8.3 (1.2)	41.0 (1.9)	36.9 (1.9)	9.1 (0.9)	4.1 (0.7)	0.6 (0.3)
Volunteered in classroom	43.4 (2.9)	24.7 (1.6)	21.5 (1.9)	5.5 (1.3)	2.8 (0.7)	2.1 (0.7)
Participated in parent and family social activities	33.8 (2.0)	34.3 (1.8)	23.2 (1.6)	4.6 (0.7)	0.9 (0.3)	3.3 (0.8)
Been on class trips	60.1 (2.5)	24.2 (2.0)	8.8 (1.2)	0.8 (0.3)	0.4 (0.2)	5.7 (1.2)
Participated in fundraising activities	56.7 (3.0)	19.8 (2.1)	13.4 (1.8)	0.8 (0.3)	0.3 (0.1)	9.1 (1.7)
Attended a general school meeting	30.4 (1.8)	40.9 (1.9)	21.3 (1.3)	2.5 (0.5)	0.4 (0.2)	4.5 (0.8)
Attended a school or class event	26.0 (1.8)	39.8 (1.9)	29.1 (1.7)	3.0 (0.6)	0.2 (0.1)	2.0 (0.4)
Served on a school committee or volunteered outside classroom	76.6 (1.6)	9.0 (0.9)	6.2 (0.8)	2.3 (0.4)	0.2 (0.1)	5.7 (0.9)
Participated in parenting education classes	65.6 (1.9)	16.0 (1.4)	8.4 (1.1)	1.5 (0.4)	0.6 (0.3)	7.9 (1.2)
Participated in a parent support group	76.2 (1.7)	9.2 (1.2)	4.3 (0.7)	0.5 (0.2)	0.6 (0.4)	9.1 (1.4)
Attended adult education classes	79.2 (1.8)	3.4 (0.6)	2.8 (0.6)	0.5 (0.2)	1.2 (0.4)	13.0 (1.9)
Participated in activities at home that were suggested by LAUP	38.7 (2.0)	21.3 (1.4)	29.4 (1.6)	3.2 (0.8)	3.8 (0.7)	3.6 (0.9)
Attended LAUP presentations	60.1 (2.1)	22.2 (1.4)	8.2 (1.1)	1.1 (0.3)	0.4 (0.2)	8.1 (1.5)
Attended an LAUP Expo and Conference	78.0 (1.9)	9.2 (1.1)	1.7 (0.4)	0.2 (0.1)	0.3 (0.2)	10.7 (1.8)
Attended a kindergarten transition workshop	75.2 (2.2)	22.0 (2.1)	2.3 (0.2)	0.3 (0.2)	0.1 (0.1)	n.a.
Attended any other workshops sponsored by LAUP	91.0 (1.0)	6.7 (0.8)	1.9 (0.4)	0.4 (0.2)	0.0 (0.0)	n.a.

Source: Parent interviews, spring 2008.

n.a. = not available. This response category was not offered for these items.

Table III.33   Frequency of Teacher-Reported Family Involvement Practices (Weighted)

	Never	Once or Twice This Year	Once Every Two to Three Months	Once a Month	Once Every One to Two Weeks	One or More Times a Week
	Percentage (Standard Error)	Percentage (Standard Error)	Percentage (Standard Error)	Percentage (Standard Error)	Percentage (Standard Error)	Percentage (Standard Error)
Had a conference with a parent	6.4 (2.7)	82.9 (3.7)	5.3 (2.1)	3.9 (1.5)	0.5 (0.5)	1.0 (0.8)
Contacted a parent if child has problems	11.2 (3.4)	15.7 (3.7)	11.5 (3.1)	15.8 (3.4)	13.0 (3.3)	32.7 (5.0)
Contacted a parent if child does something well	3.1 (1.9)	5.2 (2.4)	3.3 (1.5)	9.4 (2.8)	14.6 (3.2)	64.4 (4.9)
Involved a parent as a classroom volunteer	12.0 (3.4)	9.8 (3.1)	5.5 (2.2)	16.7 (3.6)	15.6 (3.6)	40.5 (5.2)
Told a parent about skills child needs for kindergarten	1.5 (1.1)	27.1 (4.4)	8.6 (2.7)	21.9 (4.5)	5.8 (1.8)	35.2 (5.1)
Provided at-home activities for a parent to do to improve child's skills	2.4 (1.3)	8.5 (2.8)	9.0 (2.9)	17.0 (3.7)	25.4 (4.4)	37.7 (5.1)
Given a parent ideas about discussing TV shows with child	41.8 (5.2)	19.4 (3.9)	6.6 (2.5)	16.3 (3.8)	8.5 (2.7)	7.4 (2.2)
Assigned at-home activities for parent-child interaction	2.5 (1.5)	6.3 (2.3)	4.6 (1.8)	14.4 (3.8)	23.2 (4.2)	49.0 (5.3)
Asked a parent to read to child	1.3 (1.3)	6.4 (2.5)	5.9 (2.5)	7.6 (2.6)	13.2 (3.3)	65.6 (5.0)
Encouraged a parent to ask child about his/her day	3.8 (2.1)	8.3 (2.9)	1.5 (1.1)	6.2 (2.6)	8.0 (2.7)	72.1 (4.8)
Asked a parent to visit classroom	4.4 (2.1)	7.0 (2.5)	4.9 (2.1)	11.3 (3.2)	14.5 (3.4)	58.0 (5.0)
Asked a parent to take child to library or community events	5.6 (2.5)	14.6 (3.5)	7.2 (2.6)	23.2 (4.3)	16.9 (3.5)	32.5 (4.6)
Given a parent ideas to help him/her become an effective advocate for child	4.5 (2.3)	17.1 (3.7)	9.6 (3.0)	17.3 (3.9)	18.8 (3.5)	32.7 (4.3)
Sent home letters telling parents what children have been learning and doing in class	16.0 (3.8)	8.1 (2.5)	5.4 (2.3)	42.6 (5.0)	10.1 (3.0)	17.8 (3.6)
Conducted home visits	84.6 (3.8)	14.5 (3.7)	0.0 (0.0)	0.0 (0.0)	0.7 (0.7)	0.2 (0.2)
Provided opportunities for parents to get to know one another	8.4 (2.9)	27.7 (4.6)	11.4 (2.9)	32.3 (4.5)	5.2 (2.4)	15.0 (3.7)
Given parents information related to kindergarten	2.7 (1.4)	62.2 (4.8)	8.7 (2.8)	13.5 (3.1)	4.0 (1.6)	8.9 (3.2)

Source: Teacher interviews, winter 2008.

Appendix C

Table III.34.   Summary of HLM Analyses of Expressive Vocabulary (EOWPVT)

	Model 1	Model 2	Model 3	Model 4	Model 5
Intercept					
Child Characteristics					
Female (reference)					
Male	-0.005	-0.002	-0.004	-0.004	-0.004
	(0.031)	(0.031)	(0.031)	(0.031)	(0.031)
Age	0.006	0.006	0.007	0.007	0.007
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Race/ethnicity					
Latino (reference)					
African American	0.088	0.061	0.062	0.056	0.048
	(0.059)	(0.059)	(0.059)	(0.061)	(0.063)
White	0.027	-0.005	-0.013	-0.024	-0.021
	(0.089)	(0.089)	(0.089)	(0.090)	(0.091)
Asian	0.055	0.068	0.094	0.096	0.096
	(0.074)	(0.076)	(0.078)	(0.077)	(0.078)
Other	0.123	0.113	0.129	0.124	0.127
	(0.108)	(0.111)	(0.115)	(0.115)	(0.117)
Language group					
English (reference)					
Spanish only	-0.546***	-0.517***	-0.435***	-0.438***	-0.426***
	(0.055)	(0.056)	(0.065)	(0.066)	(0.064)
Spanish primarily	-0.360***	-0.338***	-0.280***	-0.281***	-0.267***
	(0.048)	(0.048)	(0.055)	(0.056)	(0.055)
Other language	-0.141	-0.123	-0.115	-0.112	-0.113
	(0.088)	(0.089)	(0.092)	(0.092)	(0.095)
Initial ability	0.656***	0.648***	0.640***	0.638***	0.643***
	(0.027)	(0.027)	(0.027)	(0.027)	(0.028)
Squared initial ability	0.057***	0.054***	0.055***	0.054***	0.053***
	(0.013)	(0.013)	(0.012)	(0.012)	(0.013)
Assessment time interval	0.012+	0.012+	0.013+	0.012+	0.016*
	(0.007)	(0.012)	(0.007)	(0.007)	(0.007)
Family Characteristics					
Risk factors (3 or more)		-0.108**		-0.060	-0.056
		(0.040)		(0.050)	(0.050)
Family poverty ratio			0.015	0.010	0.009
			(0.016)	(0.017)	(0.017)
Maternal education (at least a high school diploma)			0.097**	0.084**	0.086*
			(0.035)	(0.036)	(0.036)
Maternal employment			-0.019	-0.018	-0.025
			(0.043)	(0.043)	(0.043)
Parent immigration status					
No immigrant parent (reference)					
One immigrant parent			-0.057	-0.040	-0.041
			(0.055)	(0.059)	(0.060)
Two immigrant parents			-0.126**	-0.108+	-0.109+
			(0.056)	(0.061)	(0.060)
Parent immigration recency			0.033	0.043	0.061
			(0.066)	(0.067)	(0.071)
Household dependency ratio			-0.004	-0.003	-0.002
			(0.020)	(0.020)	(0.020)
Parent depressive symptoms			-0.000	0.000	0.000
			(0.003)	(0.003)	(0.003)

Table III.34 *(continued)*

	Model 1	Model 2	Model 3	Model 4	Model 5
Program Characteristics					
Percentage Spanish speakers					
≤ 25					-0.027
					(0.054)
26–74 (reference)					
≥75					-0.115
					(0.074)
Program socioeconomic status					0.009
					(0.044)
Percentage recent immigrants					-0.773*
					(0.387)
Mean ability					-0.035
					(0.041)
Variation of ability					0.053*
					(0.025)
Percentage of variance explained at level 1	57.04	57.23	57.76	57.80	
Percentage of variance explained at level 2					15.60
Percentage of total variance explained	63.00	63.20	63.50	63.55	64.19

Note: ICC=0.23. $+p < .10$; $*p < .05$; $**p < .01$; $***p < .001$.

Table III.35.   Summary of HLM Analyses of Receptive Vocabulary (ROWPVT)

	Model 1	Model 2	Model 3	Model 4	Model 5
Intercept	0.270***	0.376***	0.330***	0.391***	0.396***
	(0.042)	(0.045)	(0.062)	(0.075)	(0.079)
Child Characteristics					
Female (reference)					
Male	-0.60	-0.058	-0.059	-0.059	-0.060
	(0.039)	(0.038)	(0.385)	(0.039)	(0.040)
Age	0.023***	0.022***	0.024***	0.024***	0.024***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Race/ethnicity					
Latino (reference)					
African American	0.180**	0.131*	0.143*	0.132*	0.113+
	(0.058)	(0.055)	(0.059)	(0.059)	(0.059)
White	0.212**	0.152*	0.167*	0.148*	0.105
	(0.065)	(0.062)	(0.065)	(0.064)	(0.065)
Asian	0.247**	0.256**	0.297***	0.297***	0.254**
	(0.080)	(0.082)	(0.085)	(0.086)	(0.083)
Other	0.189*	0.164*	0.190*	0.178+	0.149
	(0.085)	(0.083)	(0.095)	(0.094)	(0.094)
Language group					
English (reference)					
Spanish only	-0.898***	-0.856***	-0.734***	-0.739***	-0.726***
	(0.083)	(0.082)	(0.087)	(0.087)	0.087
Spanish primarily	-0.487***	-0.453***	-0.383***	-0.385***	-0.367***
	(0.058)	(0.060)	(0.064)	(0.063)	0.063
Other language	-0.422***	-0.383***	-0.343**	-0.337**	-0.350**
	(0.118)	(0.117)	(0.120)	(0.120)	(0.131)
Initial ability	0.403***	0.397***	0.392***	0.391***	0.386***
	(0.032)	(0.032)	(0.032)	(0.032)	(0.032)
Squared initial ability	--	--	--	--	--
Assessment time interval	-0.0004	-0.001	0.001	0.000	-0.001
	(0.012)	(0.012)	(0.011)	(0.011)	(0.011)
Family Characteristics					
Risk factors (3 or more)		-0.165***		-0.099	-0.091
		(0.046)		0.062	(0.063)
Family poverty ratio			0.024	0.016	0.003
			(0.015)	(0.017)	(0.018)
Maternal education (at least a high school diploma)			0.098+	0.076	0.070
			(0.054)	(0.058)	(0.059)
Maternal employment			-0.070	-0.070	-0.075
			(0.047)	(0.047)	(0.048)
Parent immigration status					
No immigrant parent (reference)					
One immigrant parent			-0.027	0.003	0.006
			(0.044)	(0.049)	(0.049)
Two immigrant parents			-0.159**	-0.128*	-0.118+
			(0.061)	(0.061)	(0.062)
Parent immigration recency			-0.128+	-0.111	-0.104
			(0.076)	0.076	(0.073)
Household dependency ratio			-0.038	-0.037	-0.038
			(0.027)	(0.026)	(0.027)
Parent depressive symptoms			-0.001	0.000	0.000
			(0.004)	(0.004)	(0.004)

Table III.35 *(continued)*

	Model 1	Model 2	Model 3	Model 4	Model 5
Program Characteristics					
Percentage Spanish speakers					
≤ 25					0.023
					(0.068)
26–74 (reference)					
≥75					-0.079
					(0.109)
Program socioeconomic status					0.092*
					(0.047)
Percentage recent immigrants					-0.297
					(0.374)
Mean ability					0.003
					(0.043)
Variation of ability					0.022
					(0.036)
Percentage of variance explained at level 1	36.92	37.29	38.52	38.63	
Percentage of variance explained at level 2					14.43
Percentage of total variance explained	45.78	46.42	47.99	48.10	48.93

Note: ICC=0.24. $+p < .10$; $*p < .05$; $**p < .01$; $***p < .001$.

Table III.36.   Summary of HLM Analyses of Rapid Letter Naming

	Model 1	Model 2	Model 3	Model 4	Model 5
Intercept					
Child Characteristics					
Female (reference)					
Male	-0.072*	-0.066*	-0.079*	-0.078*	-0.077*
	(0.031)	(0.031)	(0.031)	(0.031)	(0.031)
Age	0.009*	0.009+	0.010*	0.009+	0.008+
	(0.004)	(0.005)	(0.004)	(0.005)	(0.005)
Race/ethnicity					
Latino (reference)					
African American	-0.051	-0.082	-0.017	-0.031	-0.007
	(0.057)	(0.055)	(0.055)	(0.053)	(0.054)
White	-0.020	-0.060	-0.021	-0.044	-0.000
	(0.095)	(0.086)	(0.093)	(0.087)	(0.088)
Asian	0.181*	0.210**	0.181*	0.183*	0.209**
	(0.072)	(0.073)	(0.076)	(0.075)	(0.080)
Other	0.007	-0.002	0.016	0.003	0.031
	(0.085)	(0.090)	(0.085)	(0.088)	(0.090)
Language group					
English (reference)					
Spanish only	-0.155*	-0.102	-0.109	-0.116	-0.127+
	(0.067)	(0.069)	(0.075)	(0.075)	(0.074)
Spanish primarily	-0.129**	-0.086+	-0.115*	-0.117*	-0.126**
	(0.041)	(0.044)	(0.047)	(0.047)	(0.047)
Other language	0.079	0.112	0.085	0.096	0.127
	(0.088)	(0.091)	(0.094)	(0.093)	(0.093)
Initial ability	0.709***	0.705***	0.694***	0.694***	0.700***
	(0.023)	(0.024)	(0.025)	(0.025)	(0.025)
Squared initial ability	-0.092***	-0.093***	-0.092***	-0.092***	-0.092***
	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)
Assessment time interval	0.007	0.008	0.008	0.007	0.004
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
Family Characteristics					
Risk factors (3 or more)		-0.161**		-0.130*	-0.138*
		(0.053)		(0.058)	(0.058)
Family poverty ratio			0.030+	0.019	0.020
			(0.015)	(0.016)	(0.016)
Maternal education (at least a high school diploma)			0.036	0.008	0.016
			(0.054)	(0.054)	(0.054)
Maternal employment			-0.048	-0.047	-0.043
			(0.038)	(0.038)	(0.040)
Parent immigration status					
No immigrant parent (reference)					
One immigrant parent			0.041	0.079	0.078
			(0.055)	(0.055)	(0.056)
Two immigrant parents			0.040	0.081	0.075
			(0.051)	(0.055)	(0.056)
Parent immigration recency			-0.158*	-0.137+	-0.133+
			(0.074)	(0.075)	(0.075)
Household dependency ratio			-0.039	-0.037	-0.037
			(0.031)	(0.030)	(0.029)
Parent depressive symptoms			-0.011**	-0.010**	-0.010**
			(0.003)	(0.004)	(0.004)

Table III.36 *(continued)*

	Model 1	Model 2	Model 3	Model 4	Model 5
Program Characteristics					
Percentage Spanish speakers					
≤ 25					-0.242**
					(0.084)
26–74 (reference)					
≥75					0.018
					(0.109)
Program socioeconomic status					0.069
					(0.056)
Percentage recent immigrants					-0.184
					(0.344)
Mean ability					-0.036
					(0.044)
Variation of ability					0.011
					(0.038)
Percentage of variance explained at level 1	49.14	49.81	50.20	50.45	
Percentage of variance explained at level 2					12.19
Percentage of total variance explained	50.70	50.99	51.88	51.98	52.71

Note: ICC=0.14. +p <.10; *p <.05; **p <.01; ***p <.001.

Table III.37.   Summary of HLM Analyses of Mathematics (ECLS-B Mathematics)

	Model 1	Model 2	Model 3	Model 4	Model 5
Intercept	0.058 (0.049)	0.127* (0.048)	0.091 (0.082)	0.101 (0.089)	-0.099 (0.096)
Child Characteristics					
Female (reference)					
Male	-0.024 (0.042)	-0.025 (0.042)	-0.029 (0.042)	-0.032 (0.042)	-0.034 (0.043)
Age	0.019** (0.006)	0.019* (0.006)	0.020** (0.007)	0.020** (0.007)	0.020** (0.007)
Race/ethnicity					
Latino (reference)					
African American	-0.134 (0.088)	-0.166+ (0.090)	-0.154 (0.095)	-0.148 (0.096)	-0.153 (0.095)
White	0.232* (0.096)	0.192* (0.096)	0.185* (0.093)	0.182* (0.092)	0.176+ (0.093)
Asian	0.337** (0.119)	0.344** (0.118)	0.305* (0.129)	0.305* (0.129)	0.292* (0.129)
Other	0.085 (0.097)	0.070 (0.094)	0.064 (0.097)	0.069 (0.096)	0.061 (0.097)
Language group					
English (reference)					
Spanish only	-0.297*** (0.071)	-.272*** (0.071)	-0.247** (0.085)	-0.248** (0.085)	-0.257** (.087)
Spanish primarily	-0.261*** (0.053)	-0.242*** (0.056)	-0.240*** (0.060)	-0.240*** (0.060)	-0.243*** (0.060)
Other language	-0.039 (0.110)	-0.012 (0.111)	-0.035 (0.105)	-0.037 (0.106)	-0.032 (0.111)
Initial ability	0.553*** (0.030)	0.545*** (0.030)	0.536*** (0.030)	0.534*** (0.030)	0.537*** (0.029)
Squared initial ability	0.082*** (0.016)	0.080*** (0.016)	0.078*** (0.015)	0.078*** (0.016)	0.079*** (0.015)
Assessment time interval	0.004 (0.008)	0.004 (0.008)	0.003 (0.008)	0.003 (0.008)	0.002 (0.008)
Family Characteristics					
Risk factors (3 or more)		-0.103* (0.046)		-0.014 (0.056)	-0.014 (0.057)
Family poverty ratio			0.034* (0.014)	0.033** (0.015)	0.025 (0.016)
Maternal education (at least a high school diploma)			0.083+ (0.050)	0.081 (0.053)	0.082 (0.053)
Maternal employment			-0.093+ (0.050)	-0.093+ (0.049)	-0.088+ (0.050)
Parent immigration status					
No immigrant parent (reference)					
One immigrant parent			-0.045 (0.050)	-0.043 (0.054)	-0.040 (0.054)
Two immigrant parents			-0.004 (0.060)	-0.001 (0.059)	0.005 (0.061)
Parent immigration recency			-0.099 (0.070)	-0.096 (0.072)	-0.096 (0.074)
Household dependency ratio			-0.025 (0.027)	-0.025 (0.027)	-0.024 (0.027)

Table III.37 *(continued)*

	Model 1	Model 2	Model 3	Model 4	Model 5
Parent depressive symptoms			-0.008+ (0.004)	-0.008+ (0.004)	-0.008+ (0.004)
Program Characteristics					
Percentage Spanish speakers					
≤ 25					-0.022 (0.089)
26–74 (reference)					
≥75					0.076 (0.094)
Program socioeconomic status					0.079 (0.056)
Percentage recent immigrants					-0.022 (0.364)
Mean ability					-0.032 (0.044)
Variation of ability					-0.003 (0.030)
Percentage of variance explained at level 1	39.09	39.22	39.85	39.92	
Percentage of variance explained at level 2					6.92
Percentage of total variance explained	45.93	46.15	47.06	47.09	47.34

Note: ICC=0.18. $+p < .10$; $*p < .05$; $**p < .01$; $***p < .001$.

Table III.38. Summary of HLM Analyses of Positive Social Skills (Teacher PKBS)

	Model 1	Model 2	Model 3	Model 4	Model 5
Intercept	0.125	0.141	0.112	0.132	0.198*
	(0.079)	(0.091)	(0.086)	(0.099)	(0.086)
Child Characteristics					
Female (reference)					
Male	-0.202***	-0.196***	-0.204***	-0.202***	-0.199***
	(0.045)	(0.046)	(0.046)	(0.046)	(0.046)
Age	0.018**	0.017**	0.017**	0.017**	0.016**
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Race/ethnicity					
Latino (reference)					
African American	-0.036	-0.036	-0.015	-0.016	-0.009
	(0.098)	(0.101)	(0.104)	(0.105)	(0.108)
White	0.087	0.087	0.087	0.082	0.096
	(0.116)	(0.117)	(0.112)	(0.113)	(0.122)
Asian	0.028	0.046	-0.008	-0.006	0.018
	(0.102)	(0.105)	(0.103)	(0.103)	(0.101)
Other	0.073	0.077	0.082	0.074	0.085
	(0.076)	(0.077)	(0.083)	(0.083)	(0.088)
Language group					
English (reference)					
Spanish only	-0.107+	-0.082	-0.103	-0.106+	-0.099
	(0.057)	(0.059)	(0.063)	(0.063)	(0.062)
Spanish primarily	-0.016	0.002	-0.019	-0.020	-0.021
	(0.052)	(0.056)	(0.061)	(0.061)	(0.062)
Other language	-0.036	-0.030	-0.075	-0.069	-0.045
	(0.099)	(0.103)	(0.104)	(0.105)	(0.106)
Initial ability	0.482***	0.485***	0.485***	0.485***	0.484***
	(0.034)	(0.035)	(0.034)	(0.034)	(0.034)
Squared initial ability	-0.096***	-0.096***	-0.094***	-0.093***	-0.093***
	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)
Assessment time interval	0.012	0.013	0.013	0.012	0.012
	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)
Family Characteristics					
Risk factors (3 or more)		-0.055		-0.030	-0.030
		(0.052)		(0.065)	(0.063)
Family poverty ratio			0.016	0.014	0.021
			(0.019)	(0.019)	(0.020)
Maternal education (at least a high school diploma)			0.063	0.056	0.057
			(0.061)	(0.065)	(0.063)
Maternal employment			-0.060	-0.062	-0.068
			(0.050)	(0.050)	(0.050)
Parent immigration status					
No immigrant parent (reference)					
One immigrant parent			-0.031	-0.022	-0.021
			(0.066)	(0.073)	(0.073)
Two immigrant parents			0.035	0.044	0.040
			(0.064)	(0.070)	(0.069)
Parent immigration recency			-0.003	0.002	-0.001
			(0.077)	(0.078)	(0.079)
Household dependency ratio			-0.021	-0.021	-0.022
			(0.030)	(0.030)	(0.030)
Parent depressive symptoms			-0.004	-0.004	-0.004
			(0.004)	(0.004)	(0.004)

Appendix C

Table III.38 *(continued)*

	Model 1	Model 2	Model 3	Model 4	Model 5
Program Characteristics					
Percentage Spanish speakers					
≤ 25					-0.130
					(0.179)
26–74 (reference)					
≥75					-0.208
					(0.148)
Program socioeconomic status					-0.239+
					(0.127)
Percentage recent immigrants					-0.213
					(0.625)
Mean ability					0.172+
					(0.101)
Variation of ability					0.037
					(0.071)
Percentage of variance explained at level 1	37.34	37.66	38.10	38.16	
Percentage of variance explained at level 2					14.51
Percentage of total variance explained	39.15	38.83	39.34	39.32	42.43

Note: ICC=0.36. +$p < .10$; *$p < .05$; **$p < .01$; ***$p < .001$.

Table III.39. Summary of HLM Analyses of Social Cooperation (Teacher PKBS)

	Model 1	Model 2	Model 3	Model 4	Model 5
Intercept	0.065	0.088	-0.012	0.028	0.074
	(0.065)	(0.076)	(0.087)	(0.092)	(0.083)
Child Characteristics					
Female (reference)					
Male	-0.151***	-0.144**	-0.154***	-0.152***	-0.149**
	(0.045)	(0.046)	(0.046)	(0.046)	(0.047)
Age	0.011+	0.011+	0.011+	0.011+	0.010
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Race/ethnicity					
Latino (reference)					
African American	-0.109	-0.111	-0.076	-0.083	-0.078
	(0.102)	(0.106)	(0.108)	(0.110)	(0.112)
White	0.103	0.100	0.126	0.116	0.126
	(0.091)	(0.097)	(0.095)	(0.097)	(0.102)
Asian	0.062	0.084	0.017	0.018	0.044
	(0.088)	(0.091)	(0.089)	(0.090)	(0.090)
Other	0.001	0.005	0.005	-0.004	0.005
	(0.089)	(0.089)	(0.091)	(0.092)	(0.094)
Language group					
English (reference)					
Spanish only	-0.035	-0.003	-0.007	-0.011	-0.001
	(0.065)	(0.065)	(0.063)	(0.062)	(0.062)
Spanish primarily	-0.001	0.023	-0.007	-0.008	-0.008
	(0.049)	(0.051)	(0.057)	(0.057)	(0.058)
Other language	0.096	0.103	0.043	0.050	0.082
	(0.093)	(0.100)	(0.107)	(0.108)	(0.110)
Initial ability	0.580***	0.582***	0.579***	0.578***	0.575***
	(0.032)	(0.032)	(0.031)	(0.031)	(0.030)
Squared initial ability	--	--	--	--	--
Assessment time interval	0.003	0.005	0.004	0.003	0.003
	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)
Family Characteristics					
Risk factors (3 or more)		-0.073		-0.064	-0.064
		(0.054)		(0.064)	(0.062)
Family poverty ratio			0.015	0.010	0.018
			(0.020)	(0.020)	(0.021)
Maternal education (at least a high school diploma)			0.106	0.091	0.090
			(0.066)	(0.066)	(0.065)
Maternal employment			-0.069	-0.069	-0.077
			(0.060)	(0.060)	(0.061)
Parent immigration status					
No immigrant parent (reference)					
One immigrant parent			0.063	0.082	0.084
			(0.066)	(0.071)	(0.070)
Two immigrant parents			0.093	0.114	0.111
			(0.067)	(0.073)	(0.072)
Parent immigration recency			-0.057	-0.046	-0.049
			(0.073)	(0.076)	(0.077)
Household dependency ratio			0.013	0.014	0.013
			(0.033)	(0.033)	(0.033)
Parent depressive symptoms			-0.004	-0.003	-0.003
			(0.004)	(0.004)	(0.004)

Appendix C

Table III.39 *(continued*

	Model 1	Model 2	Model 3	Model 4	Model 5
Program Characteristics					
Percentage Spanish speakers					
≤ 25					-0.100
					(0.157)
26–74 (reference)					
≥75					-0.145
					(0.158)
Program socioeconomic status					-0.217+
					(0.115)
Percentage recent immigrants					-0.251
					(0.535)
Mean ability					0.172*
					(0.085)
Variation of ability					0.030
					(0.060)
Percentage of variance explained at level 1	38.00	38.49	38.94	39.04	
Percentage of variance explained at level 2					15.72
Percentage of total variance explained	40.83	40.40	41.57	41.50	43.89

Note: ICC=0.29. +$p < .10$; *$p < .05$; **$p < .01$; ***$p < .001$.

Table III.40.   Summary of HLM Analyses of Social Interaction (Teacher PKBS)

	Model 1	Model 2	Model 3	Model 4	Model 5
Intercept	0.153+ (0.082)"	0.150+ (0.090)	0.204* (0.085)	0.195+ (0.099)	0.269** (0.096)
Child Characteristics					
Female (reference)					
Male	-0.227*** (0.046)	-0.225*** (0.047)	-0.230*** (0.046)	-0.229*** (0.046)	-0.227*** (0.046)
Age	0.017*** (0.005)	0.017*** (0.005)	0.016** (0.005)	0.016** (0.005)	0.016** (0.005)
Race/ethnicity					
Latino (reference)					
African American	0.039 (0.086)	0.042 (0.088)	0.048 (0.091)	0.055 (0.093)	0.061 (0.097)
White	0.065 (0.133)	0.069 (0.134)	0.035 (0.127)	0.037 (0.127)	0.056 (0.137)
Asian	0.012 (0.099)	0.017 (0.101)	-0.034 (0.099)	-0.032 (0.099)	-0.008 (0.098)
Other	0.134+ (0.070)	0.137+ (0.071)	0.138+ (0.073)	0.135+ (0.074)	0.150+ (0.080)
Language group					
English (reference)					
Spanish only	-0.126* (0.058)	-0.121* (0.060)	-0.157* (0.063)	-0.158* (0.063)	-0.151* (0.062)
Spanish primarily	-0.019 (0.054)	-0.015 (0.057)	-0.033 (0.060)	-0.033 (0.059)	-0.033 (0.060)
Other language	-0.058 (0.109)	-0.059 (0.108)	-0.100 (0.111)	-0.097 (0.112)	-0.083 (0.113)
Initial ability	0.475*** (0.038)	0.476*** (0.038)	0.479*** (0.039)	0.479*** (0.039)	0.477*** (0.039)
Squared initial ability	-0.128*** (0.018)	-0.128*** (0.018)	-0.128*** (0.018)	-0.128*** (0.017)	-0.128*** (0.017)
Assessment time interval	0.020 (0.015)	0.020 (0.015)	0.019 (0.015)	0.019 (0.015)	0.020 (0.016)
Family Characteristics					
Risk factors (3 or more)		-0.006 (0.046)		0.018 (0.062)	0.019 (0.061)
Family poverty ratio			0.038 (0.058)	0.014 (0.015)	0.021 (0.016)
Maternal education (at least a high school diploma)			0.052 (0.056)	0.056 (0.060)	0.058 (0.058)
Maternal employment			-0.100* (0.048)	-0.103* (0.048)	-0.108* (0.047)
Parent immigration status					
No immigrant parent (reference)					
One immigrant parent			-0.087 (0.060)	-0.094 (0.065)	-0.094 (0.064)
Two immigrant parents			-0.003 (0.061)	-0.010 (0.066)	-0.015 (0.065)
Parent immigration recency			0.087 (0.082)	0.084 (0.083)	0.085 (0.084)
Household dependency ratio			-0.022 (0.029)	-0.023 (0.028)	-0.024 (0.028)
Parent depressive symptoms			-0.002 (0.005)	-0.002 (0.005)	-0.002 (0.005)

Table III.40 *(continued)*

	Model 1	Model 2	Model 3	Model 4	Model 5
Program Characteristics					
Percentage Spanish speakers					
≤ 25					-0.112
					(0.171)
26–74 (reference)					
≥75					-0.284+
					(0.145)
Program socioeconomic status					-0.213+
					(0.112)
Percentage recent immigrants					-0.473
					(0.665)
Mean ability					0.098
					(0.094)
Variation of ability					0.039
					(0.071)
Percentage of variance explained at level 1	36.88	36.92	38.07	38.14	
Percentage of variance explained at level 2					12.89
Percentage of total variance explained	39.71	39.64	40.12	40.16	43.00

Note: ICC=0.38. +$p < .10$; *$p < .05$; **$p < .01$; ***$p < .001$.

Table III.41.   Summary of HLM Analyses of Social Independence (Teacher PKBS)

	Model 1	Model 2	Model 3	Model 4	Model 5
Intercept	0.125	0.141	0.087	0.091	0.167+
	(0.078)	(0.093)	(0.090)	(0.106)	(0.097)
Child Characteristics					
Female (reference)					
Male	-0.161***	-0.156**	-0.162**	-0.160**	-0.155**
	(0.049)	(0.050)	(0.050)	(0.051)	(0.051)
Age	0.022**	0.021**	0.021**	0.021**	0.020**
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Race/ethnicity					
Latino (reference)					
African American	-0.112	-0.115	-0.109	-0.102	-0.089
	(0.112)	(0.112)	(0.116)	(0.116)	(0.118)
White	0.057	0.054	0.053	0.052	0.070
	(0.126)	(0.128)	(0.125)	(0.125)	(0.133)
Asian	-0.048	-0.034	-0.061	-0.058	-0.030
	(0.123)	(0.125)	(0.124)	(0.125)	(0.119)
Other	0.063	0.065	0.067	0.059	0.074
	(0.091)	(0.090)	(0.095)	(0.095)	(0.100)
Language group					
English (reference)					
Spanish only	-0.163*	-0.142*	-0.133+	-0.135+	-0.130+
	(0.067)	(0.070)	(0.077)	(0.077)	(0.076)
Spanish primarily	0.002	0.017	0.019	0.019	0.015
	(0.060)	(0.063)	(0.069)	(0.068)	(0.070)
Other language	-0.067	-0.060	-0.065	-0.057	-0.025
	(0.107)	(0.112)	(0.115)	(0.116)	(0.116)
Initial ability	0.423***	0.425***	0.423***	0.424***	0.423***
	(0.045)	(0.046)	(0.045)	(0.045)	(0.045)
Squared initial ability	-0.050+	-0.050+	-0.049+	-0.048+	-0.048+
	(0.028)	(0.027)	(0.027)	(0.027)	(0.027)
Assessment time interval	0.012	0.013	0.014	0.014	0.012
	(0.015)	(0.015)	(0.014)	(0.014)	(0.014)
Family Characteristics					
Risk factors (3 or more)		-0.048		0.001	-0.001
		(0.057)		(0.067)	(0.064)
Family poverty ratio			0.022	0.022	0.030
			(0.020)	(0.019)	(0.020)
Maternal education (at least a high school diploma)			0.032	0.033	0.032
			(0.066)	(0.071)	(0.069)
Maternal employment			0.032	0.027	0.022
			(0.044)	(0.044)	(0.043)
Parent immigration status					
No immigrant parent (reference)					
One immigrant parent			-0.053	-0.055	-0.053
			(0.077)	(0.083)	(0.083)
Two immigrant parents			0.013	0.012	0.008
			(0.067)	(0.070)	(0.069)
Parent immigration recency			-0.048	-0.049	-0.057
			(0.081)	(0.082)	(0.083)
Household dependency ratio			-0.034	-0.035	-0.036
			(0.034)	(0.034)	(0.034)
Parent depressive symptoms			-0.005	-0.004	-0.004
			(0.004)	(0.004)	(0.004)

Table III.41 *(continued)*

	Model 1	Model 2	Model 3	Model 4	Model 5
Program Characteristics					
Percentage Spanish speakers					
≤ 25					-0.195
					(0.192)
26–74 (reference)					
≥75					-0.130
					(0.131)
Program socioeconomic status					-0.230+
					(0.135)
Percentage recent immigrants					0.173
					(0.599)
Mean ability					0.210*
					(0.106)
Variation of ability					0.031
					(0.072)
Percentage of total variance explained at level 1	27.14	27.34	27.85	27.99	
Percentage of variance explained at level 2					16.17
Percentage of total variance explained	30.71	30.54	30.57	30.69	34.21

Note: ICC=0.34. +$p < .10$; *$p < .05$; **$p < .01$; ***$p < .001$.

Table III.42. Summary of HLM Analyses of Problem Behaviors (Teacher SSRS)

	Model 1	Model 2	Model 3	Model 4	Model 5
Intercept	0.032	0.011	0.063	0.060	0.010
	(0.071)	(0.077)	(0.098)	(0.102)	(0.107)
Child Characteristics					
Female (reference)					
Male	0.003	0.001	0.009	0.007	0.001
	(0.044)	(0.044)	(0.045)	(0.045)	(0.046)
Age	-0.010*	0.010+	-0.010+	-0.010+	-0.009+
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Race/ethnicity					
Latino (reference)					
African American	0.042	0.049	0.028	0.022	0.003
	(0.116)	(0.118)	(0.120)	(0.121)	(0.124)
White	0.046	0.053	0.047	0.048	0.047
	(0.099)	(0.099)	(0.098)	(0.097)	(0.102)
Asian	-0.186+	-0.199+	-0.139	-0.142	-0.161
	(0.103)	(0.104)	(0.102)	(0.102)	(0.102)
Other	-0.107	-0.106	-0.105	-0.099	-0.106
	(0.110)	(0.108)	(0.105)	(0.104)	(0.107)
Language group					
English (reference)					
Spanish only	-0.038	0.058	-0.068	-0.067	-0.070
	(0.063)	(0.066)	(0.075)	(0.075)	(0.074)
Spanish primarily	-0.021	0.035	-0.017	-0.017	-0.011
	(0.061)	(0.060)	(0.069)	(0.069)	(0.070)
Other language	-0.202+	-0.211+	-0.166	-0.171	-0.207
	(0.118)	(0.120)	(0.126)	(0.125)	(0.129)
Initial ability	0.561***	0.561***	0.556***	0.555***	0.553***
	(0.038)	(0.038)	(0.038)	(0.038)	(0.038)
Squared initial ability	--	--	--	--	--
Assessment time interval	-0.002	-0.003	-0.002	-0.002	-0.001
	(0.014)	(0.015)	(0.014)	(0.014)	(0.013)
Family Characteristics					
Risk factors (3 or more)		0.052		-0.002	0.001
		(0.051)		(0.066)	(0.064)
Family poverty ratio			-0.022	-0.023	-0.030
			(0.017)	(0.018)	(0.019)
Maternal education (at least a high school diploma)			-0.097	-0.098	-0.096
			(0.070)	(0.075)	(0.073)
Maternal employment			0.064	0.067	0.072
			(0.057)	(0.057)	(0.057)
Parent immigration status					
No immigrant parent (reference)					
One immigrant parent			-0.012	-0.010	-0.013
			(0.079)	(0.085)	(0.085)
Two immigrant parents			-0.067	-0.065	-0.066
			(0.074)	(0.080)	(0.080)
Parent immigration recency			0.096	0.096	0.107
			(0.094)	(0.096)	(0.096)
Household dependency ratio			0.009	0.010	0.010
			(0.032)	(0.032)	(0.032)
Parent depressive symptoms			0.006	0.006	0.006
			(0.006)	(0.006)	(0.006)

Appendix C

Table III.42 *(continued)*

	Model 1	Model 2	Model 3	Model 4	Model 5
Program Characteristics					
Percentage Spanish speakers					
≤ 25					0.179 (0.170)
26–74 (reference)					
≥75					0.004 (0.117)
Program socioeconomic status					0.185+ (0.107)
Percentage recent immigrants					-0.273 (0.614)
Mean ability					-0.222** (0.081)
Variation of ability					-0.035 (0.044)
Percentage of variance explained at level 1	32.98	33.19	34.11	34.17	
Percentage of variance explained at level 2					17.48
Percentage of total variance explained	35.83	35.52	36.75	36.85	39.64

Note: ICC=0.27. $+p < .10$; $*p < .05$; $**p < .01$; $***p < .001$.

Table III.43. Summary of HLM Analyses of Social Cooperation/Approaches to Learning (Parent PKBS)

	Model 1	Model 2	Model 3	Model 4	Model 5
Intercept					
Child Characteristics					
Female (reference)					
Male	-0.178***	-0.179***	-0.194***	-0.194***	-0.193***
	(0.049)	(0.050)	(0.049)	(0.049)	(0.049)
Age	0.026***	0.026***	0.025***	0.025***	0.024***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Race/ethnicity					
Latino (reference)					
African American	-0.191+	-0.185+	-0.152	-0.157	-0.143
	(0.107)	(0.108)	(0.111)	(0.111)	(0.112)
White	-0.052	-0.042	-0.033	-0.039	-0.012
	(0.102)	(0.106)	(0.107)	(0.0108)	(0.113)
Asian	-0.031	-0.032	-0.041	-0.042	-0.008
	(0.120)	(0.120)	(0.125)	(0.125)	(0.128)
Other	-0.068	-0.064	-0.033	-0.039	-0.007
	(0.137)	(0.137)	(0.137)	(0.137)	(0.140)
Language group					
English (reference)					
Spanish only	0.096	0.086	-0.013	-0.014	-0.0158
	(0.080)	(0.083)	(0.096)	(0.096)	(0.098)
Spanish primarily	0.056	0.048	-0.020	-0.020	-0.024
	(0.063)	(0.065)	(0.075)	(0.075)	(0.076)
Other language	0.281*	0.273+	0.232	0.235	0.289*
	(0.139)	(0.141)	(0.144)	(0.144)	(0.148)
Initial ability	0.409***	0.411***	0.405***	0.406***	0.404***
	(0.025)	(0.025)	(0.025)	(0.025)	(0.025)
Squared initial ability	--	--	--	--	--
Assessment time interval	0.021*	0.021*	0.020*	0.020*	0.020*
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
Family Characteristics					
Risk factors (3 or more)		0.026		-0.022	-0.023
		(0.063)		(0.768)	(0.076)
Family poverty ratio			0.016	0.014	0.020
			(0.020)	(0.021)	(0.023)
Maternal education (at least a high school diploma)			-0.137*	-0.141*	-0.131*
			(0.064)	(0.066)	(0.066)
Maternal employment			-0.001	-0.001	-0.000
			(0.058)	(0.058)	(0.058)
Parent immigration status					
No immigrant parent (reference)					
One immigrant parent			-0.007	-0.000	-0.007
			(0.081)	(0.084)	(0.084)
Two immigrant parents			0.119	0.127	0.116
			(0.081)	(0.084)	(0.085)
Parent immigration recency			-0.074	-0.070	-0.037
			(0.096)	(0.097)	(0.100)
Household dependency ratio			0.078*	0.079*	0.078*
			(0.036)	(0.036)	(0.036)
Parent depressive symptoms			-0.008	-0.008	-0.007
			(0.005)	(0.005)	(0.005)

Appendix C

Table III.43 *(continued)*

	Model 1	Model 2	Model 3	Model 4	Model 5
Program Characteristics					
Percentage Spanish speakers					
≤ 25					-0.109+ (0.095)
26–74 (reference)					
≥75					0.048 (0.086)
Program socioeconomic status					-0.021 (0.055)
Percentage recent immigrants					-0.677+ (0.407)
Mean ability					0.031 (0.043)
Variation of ability					0.004 (0.032)
Percentage of variance explained at level 1	19.84	19.85	21.46	21.47	
Percentage of variance explained at level 2					15.74
Percentage of total variance explained	20.88	20.90	22.43	22.44	22.75

Note: ICC=0.03. $+p < .10$; $*p < .05$; $**p < .01$; $***p < .001$.

Table III.44.   Summary of HLM Analyses of Social Interaction/Independence (Parent PKBS)

	Model 1	Model 2	Model 3	Model 4	Model 5
Intercept	0.060 (0.051)	0.027 (0.063)	0.128 (0.092)	0.106 (0.100)	0.108 (0.109)
Child Characteristics					
Female (reference)					
Male	-0.067 (0.048)	-0.064 (0.048)	-0.075 (0.048)	-0.073 (0.056)	-0.074 (0.048)
Age	0.009 (0.007)	0.009 (0.007)	0.008 (0.007)	0.008 (0.007)	0.008 (0.007)
Race/ethnicity					
Latino (reference)					
African American	-0.002 (0.106)	0.014 (0.107)	-0.010 (0.109)	0.009 (0.116)	0.003 (0.110)
White	-0.159 (0.101)	-0.136 (0.104)	-0.172 (0.105)	-0.159+ (0.096)	-0.153 (0.111)
Asian	-0.236* (0.118)	-0.232* (0.117)	-0.215+ (0.123)	-0.208 (0.149)	-0.200 (0.126)
Other	-0.106 (0.133)	-0.094 (0.134)	-0.086 (0.134)	-0.065 (0.117)	-0.056 (0.136)
Language group					
English (reference)					
Spanish only	-0.089 (0.080)	-0.100 (0.082)	-0.117 (0.095)	-0.111 (0.111)	-0.102 (0.096)
Spanish primarily	0.047 (0.062)	0.039 (0.064)	0.035 (0.073)	0.041 (0.083)	0.051 (0.074)
Other language	0.139 (0.137)	0.124 (0.138)	0.121 (0.142)	0.115 (0.137)	0.114 (0.145)
Initial ability	0.494*** (0.028)	0.496*** (0.028)	0.497*** (0.028)	0.498*** (0.027)	0.494*** (0.028)
Squared initial ability	-0.090*** (0.018)	-0.089*** (0.018)	-0.090*** (0.018)	-0.089*** (0.019)	-0.088*** (0.018)
Assessment time interval	0.025** (0.009)	0.025** (0.009)	0.025** (0.009)	0.025** (0.009)	0.027** (0.009)
Family Characteristics					
Risk factors (3 or more)		0.046 (0.062)		0.041 (0.077)	0.046 (0.074)
Family poverty ratio			-0.005 (0.019)	-0.001 (0.022)	0.000 (0.022)
Maternal education (at least a high school diploma)			-0.036 (0.063)	-0.029 (0.064)	-0.025 (0.065)
Maternal employment			-0.041 (0.057)	-0.042 (0.059)	-0.046 (0.057)
Parent immigration status					
No immigrant parent (reference)					
One immigrant parent			-0.063 (0.079)	-0.078 (0.077)	-0.081 (0.082)
Two immigrant parents			-0.025 (0.080)	-0.045 (0.076)	-0.049 (0.083)
Parent immigration recency			0.026 (0.094)	0.021 (0.109)	0.047 (0.097)
Household dependency ratio			0.049 (0.035)	0.047 (0.038)	0.047 (0.035)
Parent depressive symptoms			-0.006 (0.005)	-0.006 (0.005)	-0.006 (0.005)

Table III.44 *(continued)*

	Model 1	Model 2	Model 3	Model 4	Model 5
Program Characteristics					
Percentage Spanish speakers					
≤ 25					-0.007
					(0.098)
26–74 (reference)					
≥75					-0.050
					(0.089)
Program socioeconomic status					-0.014
					(0.057)
Percentage recent immigrants					-0.585
					(0.416)
Mean ability					-0.009
					(0.044)
Variation of ability					0.024
					(0.033)
Percentage of variance explained at level 1	22.53	22.82	23.09	23.43	
Percentage of variance explained at level 2					8.76
Percentage of total variance explained	23.47	23.63	24.23	24.43	24.65

Note: ICC=0.04. +$p < .10$; *$p < .05$; **$p < .01$; ***$p < .001$.

Table III.45.   Summary of HLM Analyses of Externalizing Problem Behaviors (Parent PKBS)

	Model 1	Model 2	Model 3	Model 4	Model 5
Intercept	-0.007	0.084	0.197*	0.231*	0.198+
	(0.049)	(0.060)	(0.087)	(0.097)	(0.103)
Child Characteristics					
Female (reference)					
Male	0.095*	0.096*	0.115*	0.114*	0.113*
	(0.046)	(0.046)	(0.046)	(0.046)	(0.046)
Age	-0.015*	-0.016*	-0.014*	-0.014*	-0.013*
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Race/ethnicity					
Latino (reference)					
African American	-0.009	-0.052	-0.115	-0.131	-0.149
	(0.102)	(0.103)	(0.104)	(0.104)	(0.105)
White	0.211*	0.148	0.115	0.101	0.066
	(0.097)	(0.100)	(0.100)	(0.101)	(0.106)
Asian	0.000	0.000	0.154	0.149	0.113
	(0.113)	(0.112)	(0.117)	(0.117)	(0.119)
Other	0.348**	0.315*	0.367**	0.349**	0.311*
	(0.128)	(0.128)	(0.127)	(0.128)	(0.130)
Language group					
English (reference)					
Spanish only	-0.245**	-0.200*	-0.030	-0.035	-0.023
	(0.076)	(0.078)	(0.090)	(0.090)	(0.091)
Spanish primarily	-0.098+	-0.064	0.075	0.071	0.080
	(0.060)	(0.061)	(0.070)	(0.070)	(0.070)
Other language	-0.188	-0.143	-0.007	0.000	-0.045
	(0.132)	(0.133)	(0.135)	(0.135)	(0.138)
Initial ability	0.529***	0.525***	0.515***	0.514***	0.511***
	(0.024)	(0.024)	(0.024)	(0.024)	(0.024)
Squared initial ability	--	--	--	--	--
Assessment time interval	-0.022*	-0.022*	-0.020*	-0.020*	-0.021*
	(0.009)	(0.009)	(0.008)	(0.008)	(0.009)
Family Characteristics					
Risk factors (3 or more)		-0.146*		-0.057	-0.051
		(0.059)		(0.070)	(0.071)
Family poverty ratio			0.012	0.007	0.002
			(0.018)	(0.019)	(0.021)
Maternal education (at least a high school diploma)			-0.001	-0.011	-0.020
			(0.059)	(0.061)	(0.061)
Maternal employment			-0.036	-0.035	-0.039
			(0.054)	(0.054)	(0.054)
Parent immigration status					
No immigrant parent (reference)					
One immigrant parent			-0.171*	-0.151+	-0.143+
			(0.075)	(0.078)	(0.078)
Two immigrant parents			-0.398***	-0.376***	-0.365***
			(0.076)	(0.079)	(0.080)
Parent immigration recency			-0.029	-0.019	-0.041
			(0.090)	(0.091)	(0.093)
Household dependency ratio			-0.005	-0.003	-0.003
			(0.033)	(0.033)	(0.033)
Parent depressive symptoms			0.001	0.002	0.001
			(0.005)	(0.005)	(0.005)

Table III.45 *(continued)*

	Model 1	Model 2	Model 3	Model 4	Model 5
Program Characteristics					
Percentage Spanish speakers					
≤ 25					0.144
					(0.091)
26–74 (reference)					
≥75					-0.057
					(0.082)
Program socioeconomic status					0.007
					(0.053)
Percentage recent immigrants					0.500
					(0.388)
Mean ability					-0.021
					(0.041)
Variation of ability					-0.013
					(0.030)
Percentage of variance explained at level 1	29.83	30.18	31.45	31.59	
Percentage of variance explained at level 2					12.00
Percentage of total variance explained	32.71	33.16	34.69	34.81	35.09

Note: ICC=0.08. +p < .10; *p < .05; **p < .01; ***p < .001.

Table III.46.   Summary of HLM analyses of Internalizing problem behaviors (Parent PKBS)

	Model 1	Model 2	Model 3	Model 4	Model 5
Intercept	-0.028 (0.052)	0.055 (0.065)	0.023 (0.095)	0.086 (0.106)	0.050 (0.112)
Child Characteristics					
Female (reference)					
Male	-0.014 (0.050)	-0.013 (0.050)	-0.004 (0.050)	-0.003 (0.050)	-0.005 (0.050)
Age	-0.005 (0.007)	-0.005 (0.007)	-0.004 (0.007)	-0.005 (0.007)	-0.004 (0.007)
Race/ethnicity					
Latino (reference)					
African American	-0.078 (0.110)	-0.118 (0.111)	-0.038 (0.112)	-0.054 (0.112)	-0.074 (0.113)
White	0.328** (0.105)	0.269* (0.108)	0.296** (0.108)	0.273* (0.109)	0.215+ (0.115)
Asian	0.212+ (0.123)	0.212+ (0.123)	0.170 (0.127)	0.167 (0.127)	0.114 (0.131)
Other	0.182 (0.140)	0.154 (0.140)	0.157 (0.139)	0.138 (0.140)	0.091 (0.143)
Language group					
English (reference)					
Spanish only	0.020 (0.082)	0.064 (0.085)	0.066 (0.098)	0.063 (0.098)	0.058 (0.100)
Spanish primarily	0.015 (0.064)	0.049 (0.067)	0.057 (0.076)	0.056 (0.076)	0.059 (0.077)
Other language	-0.141 (0.143)	-0.099 (0.144)	-0.111 (0.146)	-0.100 (0.146)	-0.124 (0.151)
Initial ability	0.364*** (0.025)	0.364*** (0.025)	0.357*** (0.025)	0.358*** (0.025)	0.359*** (0.025)
Squared initial ability	--	--	--	--	--
Assessment time interval	-0.002 (0.009)	-0.002 (0.009)	-0.002 (0.009)	-0.002 (0.008)	0.000 (0.009)
Family Characteristics					
Risk factors (3 or more)		-0.134* (0.065)		-0.103 (0.077)	-0.097 (0.078)
Family poverty ratio			0.028 (0.020)	0.019 (0.021)	0.006 (0.023)
Maternal education (at least a high school diploma)			0.026 (0.065)	0.005 (0.067)	-0.001 (0.067)
Maternal employment			-0.055 (0.059)	-0.056 (0.059)	-0.052 (0.059)
Parent immigration status					
No immigrant parent (reference)					
One immigrant parent			-0.001 (0.083)	0.031 (0.086)	0.042 (0.086)
Two immigrant parents			-0.059 (0.083)	-0.027 (0.086)	-0.004 (0.087)
Parent immigration recency			0.005 (0.099)	0.024 (0.100)	0.012 (0.102)
Household dependency ratio			-0.154*** (0.036)	-0.151*** (0.036)	-0.151*** (0.036)
Parent depressive symptoms			0.005 (0.005)	0.006 (0.006)	0.006 (0.006)

Appendix C

Table III.46 *(continued)*

	Model 1	Model 2	Model 3	Model 4	Model 5
Program Characteristics					
Percentage Spanish speakers					
≤ 25					0.072
					(0.092)
26–74 (reference)					
≥75					0.060
					(0.083)
Program socioeconomic status					0.043
					(0.054)
Percentage recent immigrants					0.136
					(0.396)
Mean ability					0.001
					(0.042)
Variation of ability					0.015
					(0.031)
Percentage of variance explained at level 1	15.82	16.09	17.12	17.21	
Percentage of variance explained at level 2					11.58
Percentage of total variance explained	16.00	16.39	18.20	18.33	18.57

Note: ICC=0.02. $+p < .10$; $*p < .05$; $**p < .01$; $***p < .001$.

Table III.47. Summary of HLM Analyses of Approaches to Learning (Teacher Ratings ECLS-K Approaches to Learning)

	Model 1	Model 2	Model 3	Model 4	Model 5
Intercept	0.166** (0.057)	0.219** 0.069)	0.244** (0.083)	0.274** (0.095)	0.325** (0.097)
Child Characteristics					
Female (reference)					
Male	-0.201*** (0.041)	-0.197*** (0.041)	-0.208*** (0.041)	-0.206*** (0.041)	-0.204*** (0.041)
Age	0.020 (0.006)	0.020*** (0.006)	0.020*** (0.006)	0.020*** (0.006)	0.019*** (0.006)
Race/ethnicity					
Latino (reference)					
African American	-0.158 (0.083)	-0.181* (0.089)	-0.168+ (0.087)	-0.173+ (0.089)	-0.160+ (0.090)
White	-0.072 (0.098)	-0.098 (0.098)	-0.113 (0.084)	-0.121 (0.087)	-0.092 (0.091)
Asian	0.052 (0.112)	0.067 (0.113)	0.046 (0.121)	0.047 (0.121)	0.086 (0.112)
Other	-0.052 (0.109)	-0.061 (0.110)	-0.057 (0.105)	-0.064 (0.108)	-0.042 (0.109)
Language group					
English (reference)					
Spanish only	-0.033 (0.066)	-0.001 (0.064)	0.029 (0.070)	0.026 (0.070)	0.024 (0.069)
Spanish primarily	-0.065 (0.041)	-0.040 (0.041)	-0.028 (0.052)	-0.029 (0.052)	-0.036 (0.052)
Other language	0.130 (0.117)	0.153 (0.115)	0.142 (0.109)	0.148 (0.110)	0.169 (0.105)
Initial ability	0.522*** (0.028)	0.521*** (0.028)	0.517*** (0.028)	0.517*** (0.028)	0.513*** (0.029)
Squared initial ability	-0.097*** (0.022)	-0.098*** (0.022)	-0.092*** (0.022)	-0.093*** (0.022)	-0.095*** (0.022)
Assessment time interval	0.016 (0.018)	0.016 (0.018)	0.016 (0.017)	0.015 (0.017)	0.018 (0.017)
Family Characteristics					
Risk factors (3 or more)		-0.102+ (0.058)		-0.047 (0.064)	-0.049 (0.063)
Family poverty ratio			0.023 (0.016)	0.020 (0.016)	0.031* (0.016)
Maternal education (at least a high school diploma)			0.055 (0.055)	0.045 (0.058)	0.048 (0.058)
Maternal employment			-0.097* (0.045)	-0.097* (0.046)	-0.102* (0.045)
Parent immigration status					
No immigrant parent (reference)					
One immigrant parent			-0.067 (0.070)	-0.053 (0.078)	-0.053 (0.077)
Two immigrant parents			-0.063 (0.078)	-0.047 (0.082)	-0.054 (0.082)
Parent immigration recency			-0.106 (0.078)	-0.098 (0.076)	-0.103 (0.077)
Household dependency ratio			-0.040+ (0.024)	-0.039 (0.024)	-0.040+ (0.024)
Parent depressive symptoms			-0.007 (0.004)	-0.006 (0.004)	-0.006 (0.004)

Table III.47 *(continued)*

	Model 1	Model 2	Model 3	Model 4	Model 5
Program Characteristics					
Percentage Spanish speakers					
≤ 25					-0.128
					(0.145)
26–74 (reference)					
≥75					-0.092
					(0.123)
Program socioeconomic status					-0.215*
					(0.099)
Percentage recent immigrants					-0.036
					(0.510)
Mean ability					0.119
					(0.080)
Variation of ability					0.053
					(0.059)
Percentage of variance explained at level 1	35.50	35.87	36.94	37.01	
Percentage of variance explained at level 2					14.76
Percentage of total variance explained	38.44	38.15	39.27	39.20	41.54

Note: ICC=0.28. $+p < .10$; $*p < .05$; $**p < .01$; $***p < .001$.

Table III.48. Summary of HLM Analyses of Pencil Tapping

	Model 1	Model 2	Model 3	Model 4	Model 5
Intercept	0.050 (0.053)	0.079 (0.069)	0.095 (0.101)	0.109 (0.107)	0.087 (0.114)
Child Characteristics					
Female (reference)					
Male	-0.115* (0.057)	-0.116* (0.057)	-0.117* (0.054)	-0.121* (0.054)	-0.123* (0.054)
Age	0.043*** (0.006)	0.043*** (0.006)	0.044*** (0.006)	0.044*** (0.006)	0.045*** (0.006)
Race/ethnicity					
Latino (reference)					
African American	-0.076 (0.109)	-0.090 (0.116)	-0.093 (0.115)	-0.087 (0.117)	-0.089 (0.119)
White	0.160* (0.065)	0.140+ (0.074)	0.148* (0.072)	0.143+ (0.076)	0.149+ (0.082)
Asian	0.218 (0.134)	0.217 (0.134)	0.210 (0.144)	0.208 (0.143)	0.212 (0.147)
Other	0.245+ (0.134)	0.237+ (0.136)	0.241+ (0.133)	0.248+ (0.128)	0.258* (0.131)
Language group					
English (reference)					
Spanish only	-0.222* (0.091)	-0.210* (0.094)	-0.216* (0.109)	-0.216* (0.110)	-0.222* (0.113)
Spanish primarily	-0.077 (0.058)	-0.067 (0.060)	-0.084 (0.073)	-0.083 (0.073)	-0.086 (0.073)
Other language	0.186 (0.127)	0.200 (0.127)	0.190 (0.133)	0.187 (0.133)	0.200 (0.130)
Initial ability	0.395*** (0.024)	0.393*** (0.024)	0.394*** (0.025)	0.394*** (0.025)	0.396*** (0.025)
Squared initial ability	--	--	--	--	--
Assessment time interval	0.006 (0.008)	0.005 (0.008)	0.006 (0.008)	0.006 (0.009)	0.005 (0.009)
Family Characteristics					
Risk factors (3 or more)		-0.043 (0.056)		-0.019 (0.061)	-0.014 (0.061)
Family poverty ratio			0.000 (0.017)	-0.001 (0.017)	-0.003 (0.018)
Maternal education (at least a high school diploma)			0.023 (0.066)	0.020 (0.068)	0.025 (0.068)
Maternal employment			-0.047 (0.051)	-0.048 (0.051)	-0.040 (0.050)
Parent immigration status					
No immigrant parent (reference)					
One immigrant parent			-0.073 (0.077)	-0.070 (0.078)	-0.073 (0.078)
Two immigrant parents			0.010 (0.079)	0.015 (0.084)	0.010 (0.084)
Parent immigration recency			-0.169* (0.081)	-0.166* (0.082)	-0.146+ (0.085)
Household dependency ratio			-0.021 (0.033)	-0.021 (0.033)	-0.020 (0.033)
Parent depressive symptoms			0.003 (0.005)	0.003 (0.005)	0.003 (0.005)

Table III.48 *(continued)*

	Model 1	Model 2	Model 3	Model 4	Model 5
Program Characteristics					
Percentage Spanish speakers					
≤ 25					0.013
					(0.117)
26–74 (reference)					
≥75					0.068
					(0.082)
Program socioeconomic status					0.024
					(0.061)
Percentage recent immigrants					-0.277
					(0.369)
Mean ability					-0.022
					(0.047)
Variation of ability					-0.030
					(0.031)
Percentage of variance explained at level 1	23.42	23.43	24.03	24.11	
Percentage of variance explained at level 2					9.77
Percentage of total variance explained	25.28	25.32	25.79	25.89	26.08

Note: ICC=0.50. $+p < .10$; $*p < .05$; $**p < .01$; $***p < .001$.

Table III.49.   Summary of HLM Analyses of Walk-a-Line

	Model 1	Model 2	Model 3	Model 4	Model 5
Intercept	0.039 (0.040)	0.152* (0.066)	0.164 (0.123)	0.261+ (0.145)	0.220 (0.151)
Child Characteristics					
Female (reference)					
Male	0.021 (0.038)	0.024 (0.037)	0.024 (0.038)	0.024 (0.038)	0.033 (0.038)
Age	0.021** (0.008)	0.020** (0.007)	0.022** (0.008)	0.021** (0.008)	0.020** (0.008)
Race/ethnicity					
Latino (reference)					
African American	-0.029 (0.091)	-0.083 (0.093)	-0.089 (0.101)	-0.107 (0.103)	-0.127 (0.105)
White	-0.039 (0.128	-0.114 (0.130)	-0.099 (0.119)	-0.130 (0.124)	-0.216+ (0.129)
Asian	-0.015 (0.147)	-0.011 (0.145)	0.044 (0.129)	0.044 (0.130)	-0.010 (0.126)
Other	0.606 (0.445)	0.575 (0.440)	0.614 (0.465)	0.597 (0.465)	0.532 (0.454)
Language group					
English (reference)					
Spanish only	-0.184* (0.084)	-0.129 (0.090)	-0.094 (0.110)	-0.100 (0.108)	-0.056 (0.108)
Spanish primarily	-0.140** (0.051)	-0.096+ (0.054)	-0.067 (0.071)	-0.069 (0.070)	-0.039 (0.068)
Other language	-0.047 (0.202)	0.002 (0.191)	0.020 (0.182)	0.031 (0.180)	0.042 (0.207)
Initial ability	0.293+ (0.177)	0.291+ (0.175)	0.292+ (0.175)	0.293+ (0.175)	0.288+ (0.174)
Squared initial ability	--	--	--	--	--
Assessment time interval	0.007 (0.008)	0.007 (0.008)	0.008 (0.008)	0.008 (0.008)	0.013+ (0.008)
Family Characteristics					
Risk factors (3 or more)		-0.182* (0.072)		-0.155* (0.074)	-0.144+ (0.075)
Family poverty ratio			-0.004 (0.018)	-0.017 (0.021)	-0.022 (0.018)
Maternal education (at least a high school diploma)			0.018 (0.047)	-0.015 (0.047)	-0.034 (0.053)
Maternal employment			-0.002 (0.045)	-0.002 (0.045)	-0.023 (0.043)
Parent immigration status					
No immigrant parent (reference)					
One immigrant parent			-0.188 (0.130)	-0.142 (0.121)	-0.126 (0.120)
Two immigrant parents			-0.198 (0.140)	-0.149 (0.130)	-0.123 (0.130)
Parent immigration recency			-0.050	-0.023	-0.034

APPENDIX D

DESCRIPTIVE RESULTS FROM THE PILOT STUDY OF THE LANGUAGE INTERACTION SNAPSHOT (LISn)

A LISn DEVELOPMENT AND PILOT TESTING

1. LISn Development

The main purpose of the classroom observations in the First 5 LA Universal Preschool Child Outcomes Study (UPCOS) was to describe the quality and characteristics of Los Angeles Universal Preschool (LAUP) center-based programs. After discussions with First 5 LA and the Research Advisory Committee (RAC), we reached a consensus to use the Classroom Assessment Scoring System (CLASS; Pianta, LaParo, and Hamre 2006) to capture the overall quality of the classroom environment.[1] In addition, given the makeup of the Los Angeles population and the high proportion of English language learners (ELLs) (or dual language learners [DLLs]) in LAUP classrooms, it was important for First 5 LA to understand the language and literacy environments of ELL children. However, locating measures to assess how programs specifically support language and literacy for linguistically diverse children proved more difficult. For example, although the Instructional Support domain in the CLASS includes three dimensions that rate the quality of language feedback and language support, the CLASS requires observers to provide global ratings of the experiences of all the children in the classroom. No indicators capture the language used and the specific experiences of ELLs. For UPCOS, our goal was to capture the types of language children hear in English and in their primary home language, the different roles of the speakers, and the links between individual children and their teachers' actions. In other

---

[1] We did not administer the Early Childhood Environment Rating Scale-Revised (ECERS-R) because LAUP was already using it as a key component of the Star rating system. The general review of classroom observation measures we conducted before selecting the CLASS included measures of environment, instruction, interaction, and teacher self-reports that have been used to assess various qualities of early childhood settings serving children from 3 to 6 years old (preschool or kindergarten). We did not consider measures that require videotaping, such as the Teacher-Child Verbal Interaction Profile (TCVI; Dickinson, Haine, and Howard 1996), or measures that would substantially increase the burden on teachers, such as the Student-Teacher Relationship Scale (STRS; Pianta 2001), which requires teachers to rate aspects of their relationship with individual children.

words, how was the classroom (particularly the teacher and adult-child interaction) supporting the language development of children learning two or more languages?

We reviewed instruments that could be used to enrich the CLASS observation, with a particular focus on language and literacy and bilingual children, including the Observation Measures of Language and Literacy Instruction (OMLIT; Goodson, Layzer, and Smith 2004), the Support for ELL Classroom Assessment (SELLCA; National Institute for Early Education Research [NIEER] 2005), the Early Language and Literacy Classroom Observation (ELLCO; Smith and Dickinson 2002), the Child Care Assessment Tool for Relatives (CCAT-R; Porter et al. 2005), and the Child-Caregiver Observation System (C-COS; Boller, Sprachman, and the Early Head Start Research Consortium 1998). The CCAT-R and the C-COS are both observation instruments that MPR has developed (or participated in developing) for evaluating language experiences of individual children in caregiving environments. These instruments have a longer history than the other measures. Both employ time sampling procedures. These measures have demonstrated high reliability when used in a variety of studies. For example, the original C-COS was used in the Early Head Start Research and Evaluation Project (Administration for Children and Families 2004), as well as in two other studies, "Growing Up in Poverty" (Fuller et al. 2004) and "Who Leaves? Who Stays"? (Whitebook et al. 2004). These studies found that the C-COS caregiver talk measure was correlated with the Arnett Caregiver Interaction Scale in center-based child care settings ($r = 0.33$ to 0.34). In center-based settings in Growing Up in Poverty, C-COS caregiver talk was correlated with the total ECERS-R score ($r = 0.24$; Fuller et al. 2004). In family child care settings, C-COS caregiver talk was associated with caregiver sensitivity as measured by the Arnett ($r = 0.22 - 0.29$). The C-COS also was correlated with structural features of the child care setting, such as caregiver education ($r = 0.25$). Analysis of the C-COS items by Fuller and colleagues (2004) suggests that the measure has high internal reliability (alpha = .90). In "Who Leaves? Who Stays?," statistically significant positive correlations were found between environmental quality (total ECERS-R, FDCRS, and/or ITERS scores) and items on the C-COS dealing with language interaction between providers/teachers and children (Whitebook et al. 2004).

Based on past studies' use of these measures, input from experts (including the RAC), and the literature on language and literacy development among ELLs, we concluded that none met the needs of UPCOS. Therefore, we created the Language Interaction Snapshot (LISn) by modifying the C-COS to capture the experiences of ELLs in early childhood classrooms serving 4-year-olds. The main adaptation was to expand the original C-COS coding categories to capture additional differences in interactions (for example, designation of the language used and provision of decontextualized versus contextualized information). We describe the details of the LISn coding categories later in this appendix; further details are available in the training manual (Sprachman et al. 2008).

2. LISn Pilot Testing

Using the new coding sheet and code definitions, our trained observers visited a bilingual early childhood classroom in New Jersey and videotaped nearly two hours of uninterrupted morning activities. We transcribed the videotapes at the utterance level. From transcripts, we refined the codes, dropping some and collapsing or adding others. We then

wrote clear objective descriptions to ensure reliable coding by our observers. We established the reliability of the definitions and coding with observers who were coding the videotapes and conducting classrooms observations. Observers focused on an individual child for a five-minute time period, coding interactions that occurred in 10 30-second cycles. For each 30-second time sample, the observer noted the presence of different types of talk in English, Spanish, or mixed utterances (a mix of English and Spanish in a single utterance). Observers coded individual verbal interactions with the focus child, as well as interactions directed toward a group that included the focus child. At the end of the five-minute period, the observer noted the context of the interaction and rated the level of engagement. Then, the observer switched focus to a different child. In the pilot study, three children in each classroom were observed a minimum of three five-minute time periods. Observers alternated between making CLASS observations and conducting a set of LISn observations for each child (Table D.1). In the next sections, we provide more information about the codes, the training, and the reliability.

3.   Description of LISn Codes

The LISn documents four specific aspects of early childhood classroom language environments: (1) language spoken (for example, whether the utterances spoken by the child, or to the child by the lead teacher or another adult in the room are English, Spanish, or mixed utterances; (2) focus child verbal communication (for example, whether the focus child is talking or silent during the observation period); (3) type of teacher verbal communication (such as ways the lead teacher, assistant teacher, or other adult[2] is communicating and scaffolding language with the focus child); and (4) global classroom setting (that is, the structure of the activity in which the observation occurred, the content addressed in the activity, and the level of engagement). Each is further described below. For specific examples and ways to use each code and the LISn coding sheet, see the LISn training manual (Sprachman et al. 2008).

1. **Language Spoken.** Observers use this set of codes every time a child or a teacher in the classroom speaks. Observers determine whether the speaker has spoken a complete utterance in English, in another language (such as Spanish or Chinese), or in a mixture of languages (for example, in both English and Spanish – "My mom está feliz" [My mom is happy]).

2. **Focus Child Verbal Communication. This set of codes identifies the language spoken by the focus child (English, another language, or in mixed utt**erances) to the lead teacher, another adult in the classroom, another child, or a group of children.

---

[2] The "other adult" refers to adults acting as assistant teachers in the room. In some cases, this may include parents who volunteered in the classroom. We wanted to describe the language experiences of the child in the classroom and chose not to exclude any interactions with adults. The lead teacher is coded separately, but the utterances of all other adults toward the focus child are coded together.

Table D.1.   Structure of the LISn and CLASS Used Together

	Observation Instrument	Time in Minutes
Cycle 1	CLASS observe	20
	CLASS record	10
	LISn Child A	7 *(5 minutes observe; 2 minutes global coding)*
	LISn Child B	7
	LISn Child C	7
Cycle 2	CLASS observe	20
	CLASS record	10
	LISn Child A	7
	LISn Child B	7
	LISn Child C	7
Cycle 3	CLASS observe	20
	CLASS record	10
	LISn Child A	7
	LISn Child B	7
	LISn Child C	7
Cycle 4	CLASS observe	20
	CLASS record	10
	Additional LISn (if time permitted)	

1. **Type of Teacher Verbal Communication.** This set of codes captures whether the lead teacher or another adult in the classroom communicates with the focus child alone or with the focus child as part of a small- or whole-group instruction. Because it was not always clear whether the additional adults in the room were assistant teachers or volunteers, we differentiate them from the lead teacher, referring to them throughout this document as "Other Adult." The nine codes for adult verbal utterances fall into three broad domains: (1) response to child, which includes direct teacher/adult responses to children's utterances; (2) types of teacher/adult language, including the types of talk that teachers either produce spontaneously or in response to children; and (3) classroom talk, including reading and singing. Table D.2 provides the codes and definitions.

Table D.2.   LISn Codes for Teacher or Other Adult Verbal Communication

Code	Definition
Response to child language	
Repeats or confirms	This code is used to represent occasions in which the teacher repeats or confirms the focus child's utterance (for example, when the child says, "milk" and the teacher says "Yes, it's milk.").
Elaborates or builds	This code is used when the teacher responds to what the focus child says by building on the child's comment (for example, if the child says, "milk" and the teacher says, "You opened the carton of milk yourself."). *Note: This code must be used with one of the four codes below to indicate how the teacher elaborates.*
Types of teacher language	
Requests language	This code is used to capture a teacher's eliciting a response from a child, usually in the form of a question (for example, "What is this called?").
Gives directions	This code is used to capture a teacher's making a statement that prompts the child to do something that does not require a verbal response (for example, "Jump up and down like a frog." "Come over here." "Put your crayons away.").
Provides information (contextualized—objects present)	This code is used when the teacher provides information to the child that is connected to a visual or physical cue in the environment, a facial expression, or physical movement. The teacher is providing information about things or events that are present at that time and giving the child contextual cues about what he or she is talking about (for example, "The apple and the banana [in the child's snack] are both fruits." "This is a cotton ball. I am gluing it to the construction paper.").
Provides information (decontextualized—objects not present)	This code is used when the teacher provides information in which the meaning is conveyed solely by language. This code is most clearly recognized in the form of telling a story, recounting past events (for example, "This weekend I went to McDonalds. First I waited in line. Then, I ate a salad."), or anticipating future events ("Next week we will go to the zoo."). This information might also be about a feeling or preference when the object is not present (for example, "I love cold days." without any contextual cues).
Classroom talk	
Reading	This code is used when the teacher is reading a book with the focus child either individually or in a group (including whole group).
Singing	This code is used when the teacher sings with the focus child either individually or in a group.
Other	This code is used to capture any type of language the teacher might use that falls outside one of the codes already discussed (for example, "please," "thank you," or "wow").

Source: Sprachman et al, 2008.

2. **Global Classroom Setting.** After completing each five-minute coding period, observers capture the focus of the activity or activities that have taken place in global codes. The three categories of general codes are (1) classroom content, (2) classroom structure, and (3) classroom ratings. *Classroom content* captures whether during the five minutes children were involved in print-related or nonprint-related content, sounds, writing/copying, mathematics, singing, aesthetics, science/nature, social studies, fine motor, or gross motor activities. *Classroom structure* captures whether during the five-minute observation the class was involved in whole-group instruction, small-group instruction, individual work, routines, meals/snacks, free choice/play activities, or recess/outside activities. Whole group, small group, and individual time are teacher-organized activities. Individual time may occur without the presence of a teacher, such as when a child is assigned to work on a computer. Free play and recess involve child-selected activities. *Classroom ratings* allow the observer to note the extent to which the focus child engaged with materials, participated in activities, and was engaged in activities. The observer also rates the extent to which the teacher directed his or her attention to the focus child and how the focus child reacted to others in the class. Finally, the observer indicates whether the child participated in any sustained interactions during the five-minute period. These are defined as interactions in which there were more than two conversational turns.

B. SPRING 2008 PILOT STUDY METHODS

1. Observer Training

Training for the LISn pilot took place in Los Angeles over a three-day period in March 2008 as part of the larger training for the UPCOS winter/spring observation protocol. All observers were bilingual. In the first four days of the training, observers were trained on the CLASS and how to conduct teacher interviews. Four of the most reliable observers were then selected for LISn training. This group of four met for three half-day sessions. In addition, observers worked independently at home coding transcripts and videotapes on two evenings for additional practice. On the final morning of training, the team met in groups of two at an LAUP center in downtown Los Angeles to establish field reliability with the lead trainer (this is commonly referred to as establishing reliability with the "gold standard").[3]

We trained the LISn observers to observe three focus children three times by using the discrete LISn micro-level language coding scheme. They coded language interactions for Child A (a predominantly English speaker), Child B (a predominantly or exclusively Spanish speaker), and Child C (a bilingual child) for 20 seconds and then took 10 seconds to review their notes. The process continued over 10 cycles for a total of 5 minutes (Table D.1). To

---

[3] The gold standard observer for the LISn was an MPR researcher who was involved in the development of the LISn and led the development of the training materials. Through experience and training, she was considered to be more expert in using the LISn codes than any other observer.

keep track of time, observers relied on an ear bud and an MP3 player with the prerecorded time sequence indicating when they should observe and when they should review. Observers were instructed to talk to the classroom teacher upon arrival at the center and to record identifying information about each child selected to be a focal child. The observers also collected identifying information on three children selected ahead of time as "substitutes" for the focal children in the event that one of the focal children was absent from the classroom.

2. Reliability of the Observations

We established LISn observer inter-rater reliability by using both videotape and field methods and calculated percent agreement between each observer and the gold standard observer separately for the child and teacher components and the English, Spanish, and mixed-language verbal interactions. Thus, for each of the 30-second cycles[4] of a LISn snapshot, a percentage agreement score was calculated in nine different areas.[5] A total inter-rater reliability score was then calculated in each of the nine areas for each LISn snapshot (5 minutes) by averaging across the 10 cycles.

We examined inter-rater agreement several ways. For the training reliability using videos, inter-rater agreement was high for the overall coding (96 percent). Due to concern that the low incidence of other language or mixed-language utterances inflated the rates of agreement, we computed the reliability separately for the incidences of English. For the child component, inter-rater agreement on the English codes ranged from 72 to 98 percent, with an average inter-rater agreement across the observers of 89 percent. For the teacher component, one video highlighted "lead teacher" communication, whereas the other underscored "other adult" communication. Thus, inter-rater reliability was calculated using the average of these two scores. Agreement in English ranged from 73 to 90 percent, with an average of 82 percent across the observers. In the lead teacher video, the lead teacher switched back and forth between English and Spanish, and observers showed strong agreement in capturing the language and categories (92 percent agreement).

We based field reliability on the averages obtained from the six snapshots completed, (observations of three children across two observation cycles). Each observer was given an average percentage agreement score in the nine areas for the entire observation period (that is, six snapshots). Across the child, teacher, and other adult categories, correct coding for English utterances was 92.0 percent, for Spanish utterances 99.0 percent, and for mixed utterances 99.6 percent agreement. Because the classroom observed was dominated by an English-speaking lead teacher, we recalculated agreement using only the child and lead teacher variables in English. For the child, inter-rater agreement ranged from 85 to 95

---

[4] "Cycle" refers to the 30-second time samples within each five-minute observation. "Snapshot" refers to each five-minute time period of observation for a single child.

[5] The nine areas were Child English, Child Spanish/Other, Child Mixed Utterances, Teacher English, Teacher Spanish/Other, Teacher Mixed Utterances, Other Adult English, Other Adult Spanish/Other, and other Adult Mixed Utterances.

percent, with an average across observers of 88 percent. Similarly, for the lead teacher, scores ranged between 85 and 96 percent, with an average across observers of 89 percent. Reliability also was monitored in the initial field observations using paired observations (described below).

Data collection began the Monday following training (with only the weekend in between) to reduce possibility of rater drift. For the first three days of data collection, LISn observers were sent to centers in pairs so that they could answer each other's questions and come to consensus on scores. The lead presenter followed up with each observer after each of these three days of data collection and asked the observer to provide feedback on what was successful or difficult. They were also asked to express any concerns they had about themselves or their partners. Stronger observers, as indicated from inter-rater reliability scores with the gold standard, were paired with those who were weaker. The observer with the lowest reliability scores in the field continued to have difficulty during the first three days of pilot data collection. We therefore decided to exclude her scores and have her conduct CLASS observations only. Thus, three observers were responsible for conducting the observations in the 18 LAUP classrooms. The LISn was also piloted in San Francisco County. The San Francisco observers, who also attended the training in Los Angeles, followed the same protocol for inter-rater reliability. We report the results of the LAUP sample here; American Institutes for Research is reporting the San Francisco LISn data separately.

3. Pilot Study Sampling and Observation Procedures

We purposively sampled the classrooms from among those selected for CLASS observations. We identified 20 classrooms that included a mix of English- and Spanish-speaking children and completed LISn observations in 18 classrooms. However, in three of the classrooms we were able to complete only two LISn observations. As noted earlier, the LISn observations were conducted in conjunction with the CLASS.

We selected three children in each classroom using stratified random sampling. The children in each classroom were divided into three strata based on their membership in a language group in the fall. Stratum A included children from English-only or English-primarily with other-language groups. Stratum B included children from Spanish-primarily or English-primarily with Spanish home language. Stratum C included children who were in the Spanish-only group. One child from each stratum was randomly selected. A second child from each stratum was then randomly selected to serve as an alternate in the event that the initially selected child was absent the day of the observation. Children all spoke either English or Spanish or both. The final analytic sample included 18 children from the English routing category, 13 children from the Spanish-primarily language group, and 9 children from the Spanish-only language group.

As noted earlier, LISn observers observed three different focus children each for three five-minute snapshots. Thus, observers conducted nine snapshots during the visit—three snapshots for each of the three children. Child A was designated as the first child to observe, Child B was designated as the second, and Child C as the third. Observers were instructed to talk to the classroom teacher upon arrival at the center and record identifying information about each child selected to be the focal child. Additional information was collected on three

Appendix D

children selected ahead of time to be "substitutes" for the focal child if the focal child was absent from the classroom.

We constructed variables that are a count of the number of interactions in a particular category across the 30 cycles of observation for observations in which all 30 cycles had been completed.[6] We also constructed summary scores for the number of cycles in which each type of talk was observed (maximum of 30 per type of utterance). We constructed a variable that added the child talk to other children, the lead teacher, and the other adults. Since a major goal was to look at the types of talk occurring by the language used, each talk variable was constructed separately by language used and by the speaker. For example, there was a variable for "Lead Teacher repeats or confirms in English," "Lead Teacher repeats or confirms in Spanish," "Other Adult repeats or confirms in English," and "Other Adult repeats or confirms in Spanish." We then examined the reliability of "total talk" variables for each speaker and language (see Table D.3), but only the English variables had enough variance to construct a reliable total talk scale (teacher $\alpha = 0.77$; other adult teacher $\alpha = 0.72$). The "Lead Teacher Talk in English" and "Other Adult Talk in English" excluded singing and elaboration. Elaboration of child language seldom occurred.

Table D.3.   Constructed Variables of Total Talk

	Number of Times Observers Coded:	Potential Range
Total Lead Teacher Talk in English	B1. Repeats or confirms child language	0-210[a]
	B3. Requests language	
	B4. Gives directions	
	B5. Provides information, names, labels (objects present)	
	B6. Provides information, names, labels (objects not present)	
	B7. Reads	
	B9. Other Talk	
Total Other Adult Talk in English	C1. Repeats or confirms child language	0-210[a]
	C3 Requests language	
	C4. Gives directions	
	C5. Provides information, names, labels (objects present)	
	C6. Provides information (objects not present)	
	C7. Reads	
	C9. Other Talk	
Any Talk by Adults	Count of the number of observation cycles in which there was any talk (in any language) attributed to the lead teacher or other adults.	0-30

[a]Although it is theoretically possible to check that many categories, it is unlikely that all of the categories could be checked in each 20-second time period.

---

[6] Twelve children (three classrooms) had fewer than 30 cycles and so were excluded from the analysis.

Appendix D

Observers noted the lead teacher elaborating on child language one time (0.03 percent of time) for only four children and it was not noted at all in Spanish or other language by the lead teacher, although the other adult was observed elaborating in Spanish. Singing was more frequent, but the variance was not related to variance in the other forms of talk. We observed very few occurrences of mixed utterances and, with one exception, those were in child talk; therefore, we did not create separate variables for mixed utterances. To examine the proportion of talk in different languages descriptively, we created a variable for overall talk in Spanish by different speakers, using a count of the total talk in Spanish by the lead teacher and by the other adult, recognizing that these variables are not reliable scales.

Given the pilot nature of the study with its small sample size and overall low frequency of talk, in either English or especially Spanish, our analyses were exploratory and aimed at understanding what is captured by these observations. We examined the mean, median, and range of these variables at the child level. We explored the differences among language groups in the amount of talk observed with children. We looked at the proportion of each type of talk by language and speaker, comparing how the adults in the classroom used language. We explored differences in mean amount of talk across activity structures, looking at the total talk when a child's observations did or did not include that type of activity. We looked for any differences in interaction based on the language of the child across classrooms. We aggregated to a classroom level looking for relationships between the ratings on the CLASS scales and the findings on the LISn.

C. RESULTS

In this section, we first discuss the context for the observations, the activities occurring, and the content addressed during the observations. We next look at the frequency of child and adult talk in English and Spanish. We compare the types of talk used by the lead teacher and the other adults in English and in Spanish. We examine differences in the amount of mean amount of talk by different speakers (child, lead teacher, and other adults) across the different activity settings. Finally, we examine the relationship between the LISn and CLASS domains and dimensions.

1. Context for the Observations

The context differed across observations for each child. Observers noted what was occurring in the classroom during the five-minute observation periods for each child. After the observers coded all language interactions that occurred, they noted the activities and content included during those five-minute periods. More than one classroom activity structure or content area may have occurred in a single observation period. (See Tables D.4 and D.5 for list of activity structures and content areas.) For example, children may have been washing hands (routine) and then eating a snack (meals) in one five-minute period. The time sampling resulted in more than half of the children being observed at least once during a small-group time, recess or outside time, or a routine time. Standard administration of the CLASS excludes observations during outside or recess time, but during these times observers could conduct LISn observations; thus, it is not surprising that more than half of the sample had at least one LISn observation period that included recess or outside time.

Had observers not also been completing the CLASS, they may have captured more instructional situations with the LISn.

The context for the three observations for each child differed across time periods. None of the children was observed for even part of the snapshot in the same structure in all three observations. When the observations occurred in a similar activity structure across time periods, it was most likely to be a routine or small-group setting. Table D.4 indicates the types of activities or structure of the class time. It shows, for example, that 70 percent of the children were observed at least one time while in small-group activities, 25 percent during free play or center time, and 15 percent during individual time. Of the children observed in small groups, 19 were observed once and 9 were observed twice. Of the six children observed during individual time, all had a single observation.

Table D.4. Number of Time Periods and Children Observed in Different Activity Structures

Activity Structure	Percentage of Children with Any Observations in This Structure	Number of Observations per Child			
		0	1	2	3
Small Group	70.0	12	19	9	0
Routine	57.5	17	13	10	0
Recess or Outside	55.0	18	20	2	0
Meals or snacks	42.5	23	10	7	0
Free play	25.0	30	8	2	0
Individual Time	15.0	34	6	0	0
Whole Group	15.0	34	6	0	0

The content addressed during the observation periods also varied, as shown in Table D.5. Across the observation periods, more than 70 percent of the children were observed at least once when involved in fine motor, print-related content, and/or gross motor activities. More than half of the children were observed while engaged in an activity that involved mathematics, colors, or numbers. Observers did not observe any of the children during a social studies lesson or activity.

Frequency of Verbal Interactions at the Child Level. As noted earlier, we sampled three children in each classroom and observed each child during three five-minute snapshots using the LISn. For each snapshot, observers coded the interactions that occurred in each of 10 30-second cycles. Thus, the observers noted presence or absence of specific verbal interactions involving the target child across a total of 30 observation cycles. Observers also recorded the language used by the target child during that time period, indicating whether the talk was directed toward another child or toward an adult in the classroom.

For the child-to-child talk, child-to-lead-teacher talk, and child-to-other-adult talk categories in each language, the maximum possible frequency was 30. The majority of the child interactions were directed toward other children (Table D.6). On average, more than twice as much child talk was directed toward other children when compared with talk directed toward adults. More than 10 percent of the children did not talk to adults at all in English, and 80 percent of the children did not talk with adults in Spanish. Child talk toward adults included choral responses—that is, if the children responded as a group and the target child was a member of that group, the talk was counted as talk directed toward an adult. Thus, the low mean in the number of cycles with any child talk to an adult is surprising.

Table D.5.   Content During Observations

Content	Percentage of Children with Any Observation in This Content Area	Number of Observations per Child			
		0	1	2	3
Fine motor	80.0	8	20	12	0
Print-related content	80.0	8	20	11	1
Gross motor	70.0	18	22	13	1
Mathematics, colors, or numbers	55.0	18	17	4	1
Aesthetics	42.5	23	11	6	0
Singing	40.0	24	15	1	0
Sounds	27.5	29	10	1	0
Writing or copying	25.0	30	10	0	0
Science or nature	15.0	34	6	0	0
Not print-related	15.0	34	6	0	0
Social studies	0.0	40	0	0	0
Other	7.5	37	3	0	0

Table D.6.   Mean and Median Frequency of Child Talk Across 30 Cycles

|	English			Spanish		
	Observed Range	Mean	Median	Observed Range	Mean	Median
Child-to-Child Talk	0-24	10.4	9.5	0-11	1.13	0.0
Child-to-Adult Talk	0-18	5.0	4.5	0-5	0.40	0.0
Total Child Talk*	0-27	15.4	15.5	0-11	1.53	0.0

*Sum of Child-to-Child Talk and Child-to-Adult Talk.

2. Descriptive Statistics

Child talk directed toward other children occurred more frequently in English than in Spanish. In this sample, the maximum observed frequency of talk to other children in English was 24 cycles, with a mean of a little more than a third of the observation cycles including child utterances in English to another child. Our observers coded fewer instances of child talk in Spanish: a maximum of 11 cycles included child-to-child talk in Spanish, with a mean of only a little over one cycle including talk in Spanish. There were only four instances in which mixed utterances were observed in children's talk, and these occurred as single incidences (recorded in only one cycle) for four of the children when talking with other children.

The observers coded teacher and other adult talk whenever an adult's communication was directed toward the target child, whether individually or toward the target child's group (during large- or small-group instruction). Because the adult talk could be concentrated in a few cycles or the talk could be distributed throughout the observation cycles, we used the "Any Talk by Adults" variable to examine the distribution of time in which children experienced verbal interactions with adults (including singing as well as any talk). Out of a possible 30 observation cycles, the range is from 5 to 26 observation cycles in which there was any talk by any adult in any language to the child. More than half of the children were spoken to by an adult (at least one time) in at least 15 of the 30 observation cycles (see Figure D.1).

Figure D.1. Number of Children by Number of Observation Cycles with Any Adult Talk
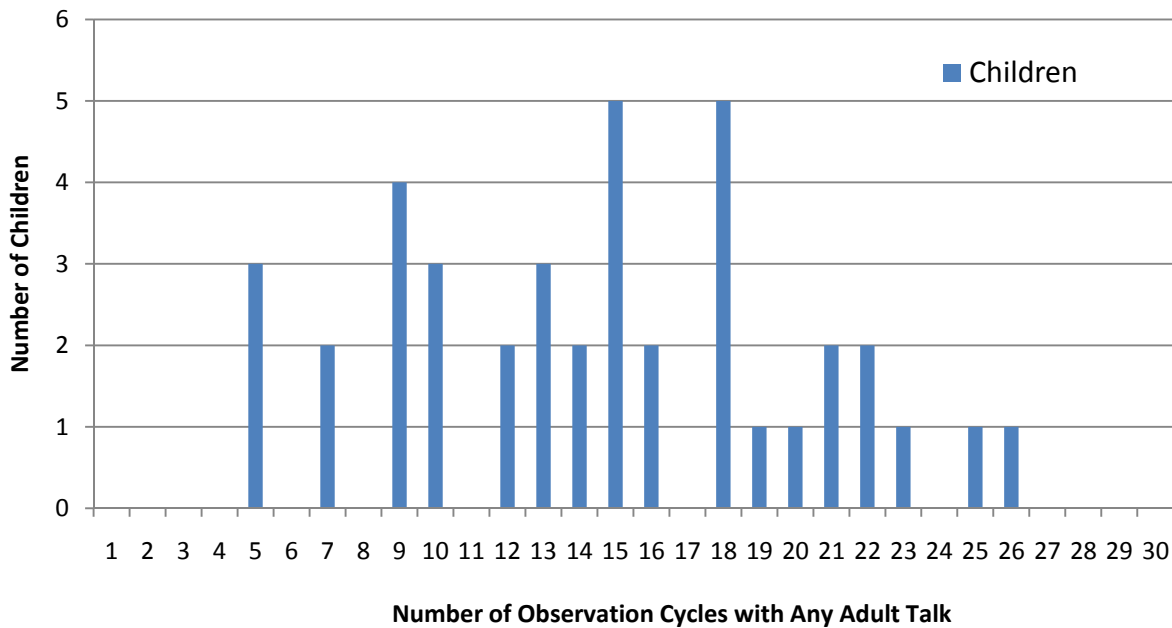

For each of the three five-minute periods, observers indicated whether the child participated in an extended or sustained conversation (defined as more than two conversational turns). In the LAUP sample, four children (10 percent) were observed in at

least one sustained conversation. Only one of those children was observed in a sustained conversation in more than one observation cycle. None of the children participated in a sustained conversation in all three observation cycles.

The amount of adult talk observed within each cycle with individual children (or children in groups) was limited, as was the variety of types of talk that children experienced (Figure D.2). The child talk in Figure D.2 includes talk directed toward both other children and talk directed toward the lead teacher and other adult. The teacher total talk in English and other adult total talk in English are based on the constructed variables that include all talk except singing and elaborates (see Table D.3). The total talk in Spanish includes singing as well as all other talk.

Figure D.2.   Total Talk by Language by Speaker
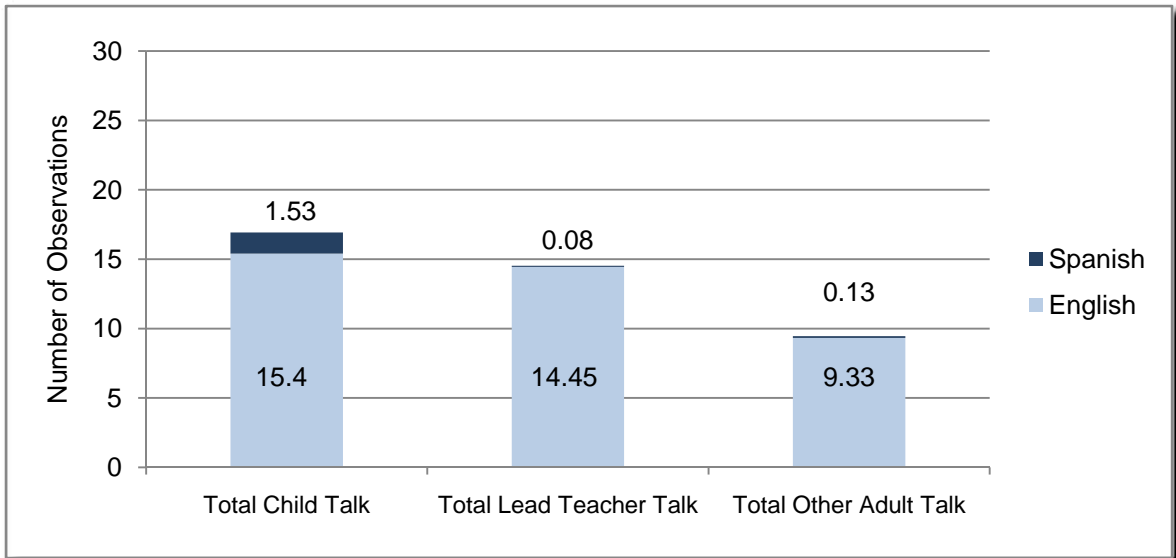
At the individual child level, the total number of the different types of adult talk was limited even when we counted all the different kinds of talk that occurred in each 30-second cycle. If we add together the lead teacher and the other adult talk across all 30 observation cycles and all categories of talk, including elaborates and singing, children experienced a mean of 24 (median = 22) categories of talk from adult interactions. As Figure D.2 shows, the majority of that talk was in English.

Figures D.3 and D.4 illustrate the relative proportion of different categories of talk in English by the lead teacher and the other adult, respectively. The proportions in each category were similar for the lead teacher and the other adult in English. Both the teacher and the lead adult used English most often to give directions and to ask questions or request language from the child, followed by labeling or other ways of providing information in context. The other adults in the room were observed reading in English more often, and it reading represented a greater proportion of their talk in English.

Fewer categories of talk were evident in Spanish, and the lead teacher and other adult varied more in their use of Spanish. Figures D.5 and D.6 illustrate the relative proportion of

the different categories of talk in Spanish by the lead teacher and the other adults. Requesting child language (for example, asking questions) and singing were the most frequent categories of talk for the lead teacher and represented 75 percent of the talk in Spanish by the lead teacher. The lead teacher also gave some directions in Spanish and repeated or confirmed child talk in Spanish. Although the other adults also often requested child language, they were not observed using Spanish for singing. The other adults elaborated, repeated, or confirmed child talk, and provided information in Spanish.

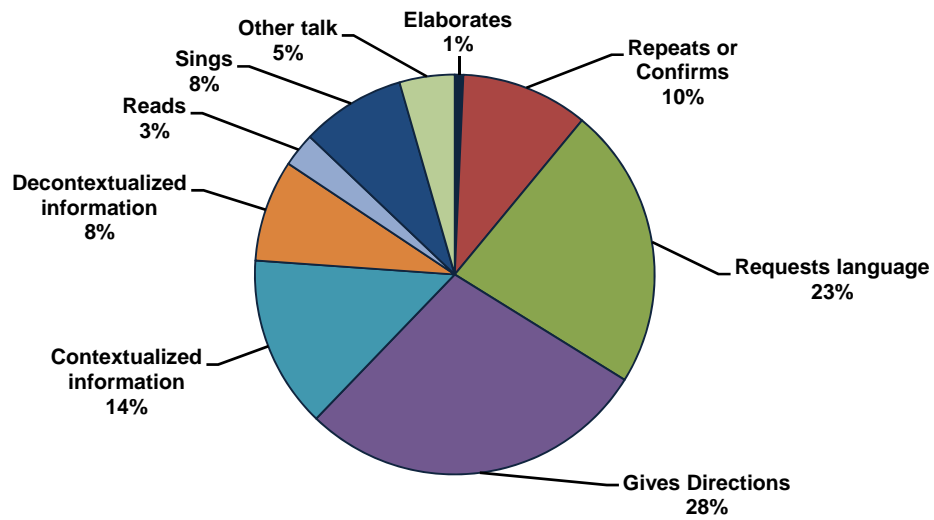**Figure D.3.   Types of Lead Teacher Talk in English**
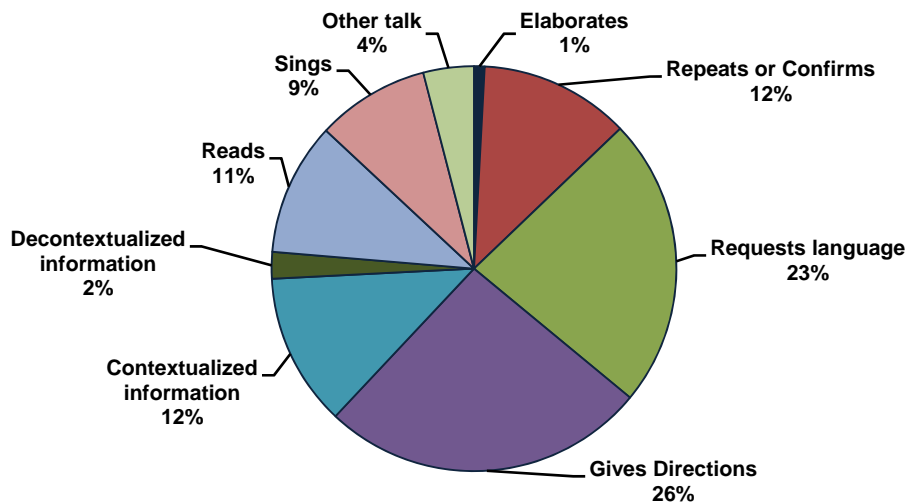

Figure D.4.   Types of Other Adult Talk in English

Figure D.5.   Types of Lead Teacher Talk in Spanish
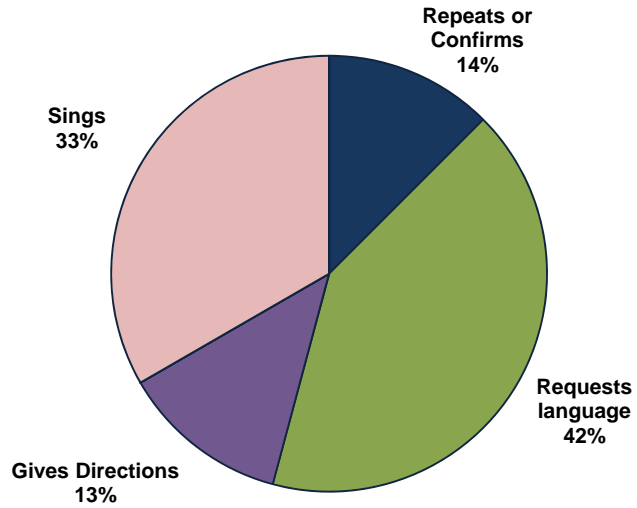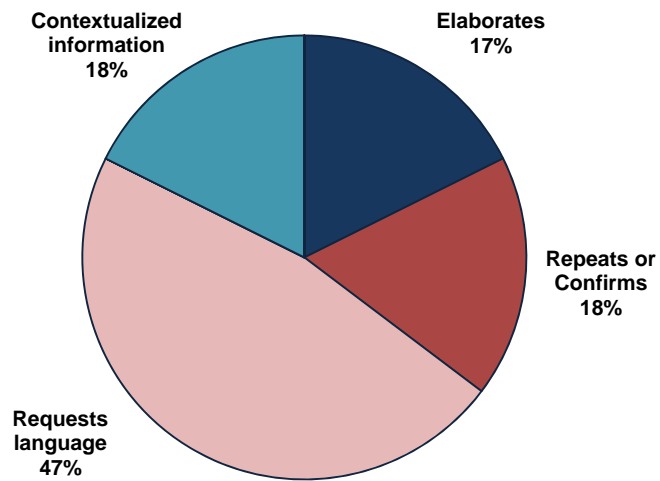
Repeats or
Confirms
14%

Sings
33%

Requests
language
42%

Gives Directions
13%

Figure D.6.   Types of Other Adult Talk in Spanish

Contextualized
information
18%

Elaborates
17%

Repeats or
Confirms
18%

Requests
language
47%

Although the differences appear very large when examined as proportions, it is important to keep in mind that the mean number of cycles with utterances in any category was less than 5 out of the 30 cycles in English and 1 out of the 30 cycles in Spanish. Table D.7 indicates the mean prevalence of English and of Spanish used in talking with children in the various categories of interactions.

Table D.7.   Range and Mean Amount of Lead Teacher and Other Adult Talk by Language

	English		Spanish	
Lead Teacher:	Range	Mean	Range	Mean
Elaborates	0-1	0.1	0	0
Repeats or confirms	0-7	1.5	0-1	0.03
Requests language	0-12	3.3	0-2	0.10
Gives directions	0-19	4.1	0-1	0.03
Give information (in context)	0-13	2.0	0	0
Gives information (decontextualized)	0-9	1.2	0	0
Reads	0-9	0.4	0	0
Sings	0-9	1.2	0-2	0.08
Other Talk	0-3	0.7	0	0
Teacher Total Talk	0-45	13.2		n.a.
Other Adult:				
Elaborates	0-2	0.1	0-1	0.03
Repeats or confirms	0-5	1.1	0-1	0.03
Requests language	0-9	2.2	0-2	0.08
Gives directions	0-8	2.5	0	0
Give information (in context)	0-10	1.2	0-3	0.03
Gives information (decontextualized)	0-3	0.2	0	0
Reads	0-9	1.0	0	0
Sings	0-10	0.9	0	0
Other Talk	0-2	0.4	0	0
Other Adult Total Talk	0-29	8.5		n.a.

Frequency of Talk by Activity. We examined the frequency of talk by the number of observations in each type of activity (Tables D.8 through D.14; and Figures D.7 through D.8). Because observers could code more than one activity in each five-minute period, we cannot isolate the category of talk by specific activity structures. Thus, although Table D.8, for example, shows the mean number of instances of each type of talk during a five-minute period that included children in small-group time, some of the talk could have occurred during a routine or snack time that also happened during those five minutes. The tables indicate the mean amount of talk in the different categories when a child had no observations that included that activity, one observation that included that activity, or two observations that included that activity. The figures illustrate the mean amount of talk when children had one or more observations in that activity structure, and compare the mean amount of talk across observations that include those activities.

During small-group time (see Table D.8), the most frequently observed activity period, child-to-adult talk in English occurred more frequently when children had one or more

observations during a small-group time (5.0 and 6.7 times when there were one or two observations, respectively, versus 3.8 times if children had no small-group time). More teacher talk in both English and Spanish was observed among the groups of children who had one observation occurring during small-group time when compared to the children who did not have any observations during that time. When two observations occurred during small-group time, the mean for other adult talk in English and in Spanish was greater.

Table D.8.   Mean Amount of Talk Occurring by Number of Observations That Included Small-Group Time

	Number of LISn Observations During Small-Group Time		
	0	1	2
	(N = 12)	(N = 19)	(N = 9)
Total Child-to-Child Talk in English	9.3	11.8	9.0
Total Child-to-Child Talk in Spanish	1.2	1.4	0.6
Total Child-to-Adult Talk in English	3.8	5.0	6.7
Total Child-to-Adult Talk in Spanish	0.7	0.4	0.1
Total Teacher Talk in English	13.4	15.4	8.3
Total Teacher Talk in Spanish	0.3	0.3	0.0
Total Other Adult Talk in English	5.4	7.8	13.9
Total Other Adult Talk in Spanish	0.1	0.1	0.3
Total Talk by All Adults	20.8	26.7	23.1

Children who had at least one observation during recess or outside time (Table D.9) were observed talking to other children in English more frequently than children without any recess observations. Mean amount of lead teacher talk in both Spanish and English was less among children observed during two recess periods than among children who had fewer observations during recess, but more frequent verbal interactions in English occurred with the assistant teacher or other adult.

Table D.9.   Mean Amount of Talk Occurring by Number of Observations That Included Recess or Outside Time

	Number of LISn Observations During Recess or Outside Time		
	0 (N = 18)	1 (N = 20)	2 (N = 2)
Total Child-to-Child Talk in English	8.1	12.1	14.5
Total Child-to-Child Talk in Spanish	2.0	0.4	1.0
Total Child-to-Adult Talk in English	4.7	5.1	7.5
Total Child-to-Adult Talk in Spanish	0.4	0.5	0.0
Total Teacher Talk in English	13.6	13.5	7.5
Total Teacher Talk in Spanish	0.2	0.3	0.0
Total Other Adult Talk in English	6.8	9.9	9.0
Total Other Adult Talk in Spanish	0.2	0.1	0.0
Total Talk by All Adults	23.0	25.5	21.5

Children who were observed more frequently during routine times experienced more lead teacher talk in English than children who were not observed during a routine time (Table D.10).

Table D.10.   Mean Amount of Talk Occurring by Number of Observations That Included Routine Time

	Number of LISn Observations During Routine Time		
	0 (N = 17)	1 (N = 13)	2 (N = 10)
Total Child-to-Child Talk in English	9.7	10.9	10.9
Total Child-to-Child Talk in Spanish	1.1	1.7	0.4
Total Child-to-Adult Talk in English	6.4	5.0	2.7
Total Child-to-Adult Talk in Spanish	0.4	0.2	0.7
Total Teacher Talk in English	9.9	12.1	20.4
Total Teacher Talk in Spanish	0.2	0.1	0.5
Total Other Adult Talk in English	10.2	8.7	5.3
Total Other Adult Talk in Spanish	0.2	0.1	0.0
Total Talk by All Adults	20.8	24.2	29.8

Table D.11. Mean Amount of Talk Occurring by Number of Observations That Included Meals or Snack Time

	Number of LISn Observations During Meals or Snack Time		
	0	1	2
	(N = 23)	(N = 10)	(N = 7)
Total Child-to-Child Talk in English	13.2	6.4	7.0
Total Child-to-Child Talk in Spanish	0.8	0.2	3.6
Total Child-to-Adult Talk in English	3.7	7.2	6.3
Total Child-to-Adult Talk in Spanish	0.3	0.3	0.9
Total Teacher Talk in English	11.0	19.5	11.7
Total Teacher Talk in Spanish	0.2	0.3	0.3
Total Other Adult Talk in English	7.3	11.9	7.4
Total Other Adult Talk in Spanish	0.0	0.3	0.1
Total Talk by All Adults	21.2	33.5	20.6

Children observed during one free play period experienced more lead teacher talk in English and in Spanish than children with no observations in free play (Table D.12). However, the majority of children were not observed during a free play period, so the pattern of the frequency of talk during this activity may not be indicative of what typically happens during free play. In a similar way, fewer children had observations occurring during individual activities or during whole-group time, so it is difficult to see any patterns (Tables D.13 and D.14).

Table D.12. Mean Amount of Talk Occurring by Number of Observations That Included Free Play

	Number of LISn Observations During Free Play		
	0	1	2
	(N = 30)	(N = 8)	(N = 2)
Total Child-to-Child Talk in English	11.2	6.6	14.0
Total Child-to-Child Talk in Spanish	1.0	0.6	5.5
Total Child-to-Adult Talk in English	5.3	4.5	2.5
Total Child-to-Adult Talk in Spanish	0.3	0.9	0.0
Total Teacher Talk in English	12.5	16.6	10.5
Total Teacher Talk in Spanish	0.2	0.5	0.0
Total Other Adult Talk in English	8.6	9.3	3.5
Total Other Adult Talk in Spanish	0.1	0.4	0.0
Total Talk by All Adults	23.3	28.6	18.5

**Table D.13. Mean Amount of Talk Occurring by Number of Observations That Included Individual Time**

	Number of LISn Observations During Individual Time		
	0	1	2
	(N = 34)	(N = 6)	(N = 0)
Total Child-to-Child Talk in English	11.2	5.8	
Total Child-to-Child Talk in Spanish	0.8	3.0	
Total Child-to-Adult Talk in English	5.3	3.3	
Total Child-to-Adult Talk in Spanish	0.4	0.2	
Total Teacher Talk in English	12.7	16.2	
Total Teacher Talk in Spanish	0.2	0.5	
Total Other Adult Talk in English	9.5	2.7	
Total Other Adult Talk in Spanish	0.2	0.0	
Total Talk by All Adults	24.2	23.7	

Table D.14. Mean Amount of Talk Occurring by Number of Observations That Included Whole-Group Time

|	Number of LISn Observations During Whole-Group Time		
	0	1	2
	(N = 34)	(N = 6)	(N = 0)
Total Child-to-Child Talk in English	10.8	8.0	
Total Child-to-Child Talk in Spanish	1.3	0.0	
Total Child-to-Adult Talk in English	4.5	7.7	
Total Child-to-Adult Talk in Spanish	0.4	0.2	
Total Teacher Talk in English	12.3	18.7	
Total Teacher Talk in Spanish	0.2	0.2	
Total Other Adult Talk in English	8.3	9.7	
Total Other Adult Talk in Spanish	0.1	0.5	
Total Talk by All Adults	22.7	32.7	

When examined across the mean amount of talk when children had at least one observation in that activity structure, the pattern of talk suggests that teacher talk in English dominated individual, large-group, free play, routine, and meal times. The mean amount of child talk in English was greatest during outside or recess time. The other adult spoke less frequently in English than the child or lead teacher across all activity structures (Figure D.7).

Figure D.7.   Mean Amount of Talk in English by Activity Structure
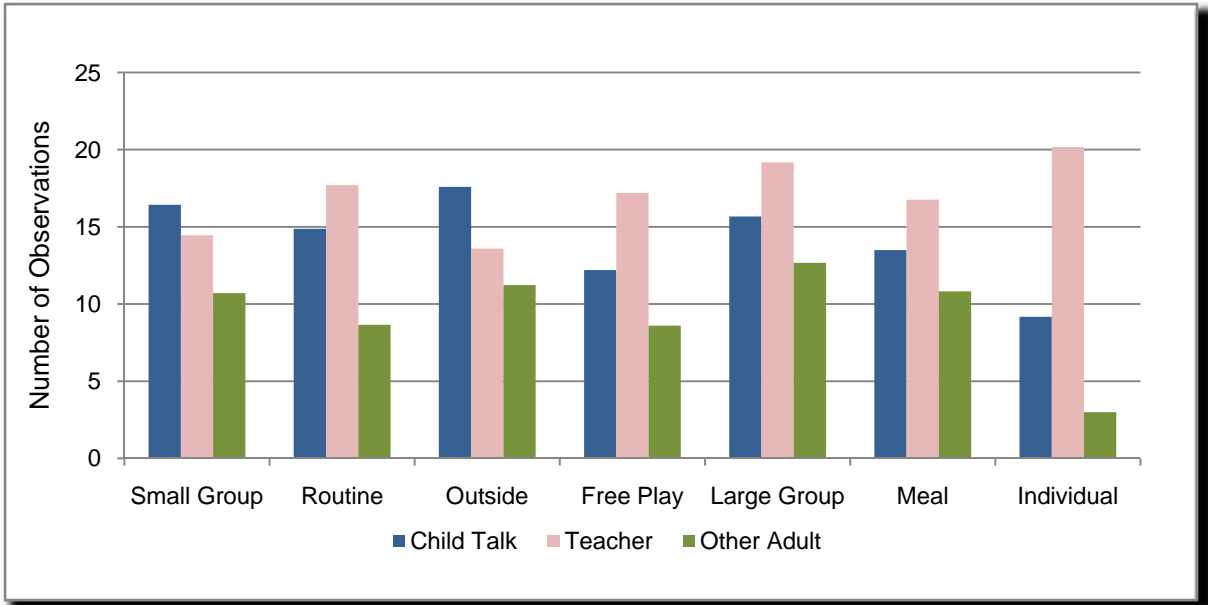
As illustrated in Figure D.8, the pattern of the mean amount of talk suggests that, with the exception of large-group time, talk in Spanish was more frequently found in child talk. During large-group time, the other adult used Spanish more frequently than the child or lead teacher.

Figure D.8.   Mean Amount of Talk in Spanish by Activity Structure[a]
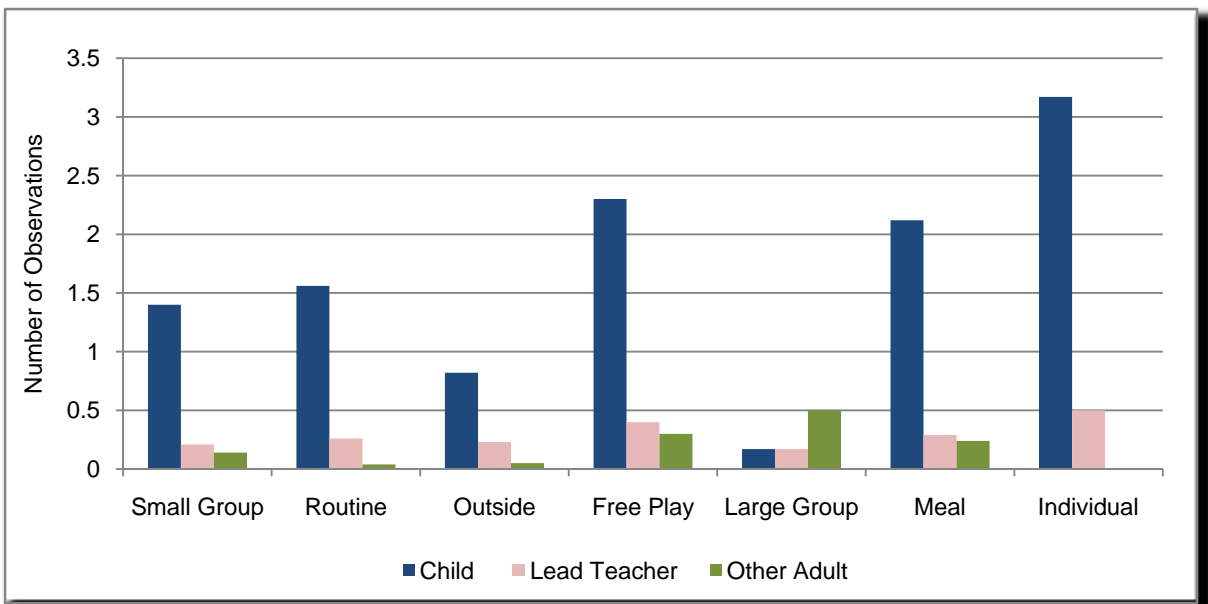
[a]Child-adult talk includes choral responding.

Child-to-child talk in English was more frequent than child-to-child talk in Spanish, and was most frequent among those children who had at least one observation that occurred during recess or outside time (Figure D.9). Child-to-child talk in Spanish was more frequent among those that had observations that occurred during a meal time or free play. With the exception of talk in English during meal time, child-to-child language was more frequent in both languages than child-to-adult language.

Figure D.9.   Child Talk by Conversational Partner Across Activity Structures[a]
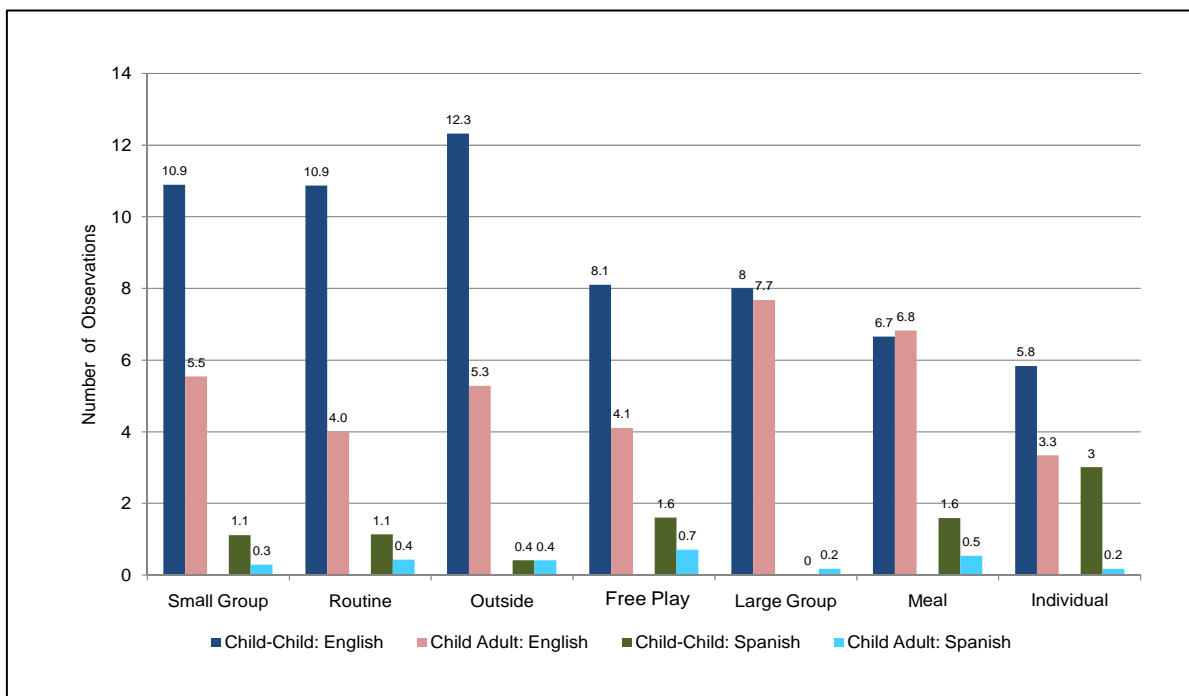
[a]Child-adult talk includes choral responding.

Bivariate Correlations. We examined the bivariate correlations among the count of the total number of cycles for each category, language, and speaker. Using the child-level data, few significant relationships were detected across speakers. Child-adult talk in English was positively correlated with total Other Adult Talk in English ($r = 0.39$; $p<.05$), but showed no significant relationship to teacher total talk in English. However, the lead teacher total talk in Spanish was significantly associated with child-to-adult talk in Spanish ($r = 0.48$; $p<.01$). Total teacher talk in Spanish was correlated with total teacher talk in English ($r = .33$; $p<.05$). Total other adult talk in Spanish was not related to total other adult talk in English, nor was it related to total talk by the lead teacher or child in either language.

When looking at the data aggregated to the classroom level, total lead teacher talk in Spanish had a stronger relationship with total lead teacher talk in English when singing was

included, but the significance was marginal with the smaller sample size (r = 0.49; *p*<.10).[7] The total other adult talk in Spanish was also marginally associated with the total other adult talk in English (r = 0.45; *p*<.10). However, this was true only if we considered the other adult talk in English variable that excludes singing. Child-to-adult talk in English was marginally associated with total teacher talk in Spanish (r = 0.44; *p*<.10). Child-to-adult talk in Spanish was not associated with the total talk by teacher or by other adult in either language.

2. Classroom-Level Analysis: Relationships Between the LISn and CLASS Scores

To examine evidence of the convergent validity of the LISn, we examined the relationship between the CLASS and the LISn. As noted earlier, the observations for both measures were conducted by the same observer on the same day. We aggregated the constructed LISn variables to the classroom level and examined the relationship with the CLASS Emotional Support, Classroom Organization, and Instructional Support scales (N = 14 classrooms).[8] We also looked at the dimensions of the Instructional Support scale to see whether the LISn variables were more strongly related to the dimensions such as language modeling than to other dimensions. Similarly, we looked to see whether the LISn variables that reflected more-responsive teaching ("repeats and confirms," "elaborates") showed a stronger relationship to positive climate.

The majority of the significant relationships were found with the lead teacher talk in English variables (see Table D.15 and D.16). The focus of the CLASS observations was on the teacher, so it is not surprising that few of the other adult talk variables were related to the CLASS scores. The majority of the relationships were with the Instructional Support scale, particularly the dimensions measuring quality of feedback and language modeling. Quality of feedback had the strongest relationship with the LISn variables, particularly the Total Talk in English and Gives Information in Context. Reading, singing, giving directions, and other talk were not related to the Instructional Support scale or related dimensions. Elaborates was related to Language Modeling. Repeats or confirms was related to Quality of Feedback. Request for Language and Gives Information–Decontextualized were both related to Quality of Feedback and Language Modeling.

The expected relationship with responsive language (elaborating and repeats or confirms) was not evident. However, Total Talk in English and Gives Information in Context showed moderate relationships to Positive Climate.

Although the lead teacher talk in Spanish was limited and there was no variance in the scores on most types of language, a few relationships were found between lead teacher talk

---

[7] The child-level correlations are based on a sample size of 40, whereas the classroom-level correlations are based on 15 classrooms.

[8] Observations for one LISn did not have complete CLASS data, and another classroom was missing an Emotional Support score.

Table D.15. Relationship Between CLASS Scales and Lead Teacher Talk in English

	CLASS Emotional Support	CLASS Classroom Organization	CLASS Instructional Support
Total talk[a]	0.48	0.34	0.55*
Repeats or Confirms	0.32	0.23	0.45
Elaborates	0.35	0.45	0.36
Requests Language	0.43	0.24	0.49†
Directions	0.32	0.47†	0.35
Gives Information in Context	0.53†	0.29	0.63*
Gives Information Decontextualized	0.39	0.05	0.45
Reads	0.11	0.27	0.04
Sings	-0.05	0.31	-0.13
Other Talk	0.35	0.27	0.39

Note: N=14 classrooms.

†*p*<.10 * *p*<.05 ***p*<.01

[a]Total talk includes Repeats or Confirms, Requests Language, Gives Directions, Provides Information (contextualized and decontextualized), Reading, and Other.

Table D.16. Relationship Between Select Dimensions of CLASS and Lead Teacher Talk in English[b]

	CLASS Quality of Feedback	CLASS Concept Development	CLASS Language Modeling	CLASS Positive Climate
Total Talk[a]	0.72*	-0.10	0.59*	0.56*
Repeats or Confirms	0.64*	-0.03	0.41	0.32
Elaborates	0.14	0.17	0.55*	0.22
Requests Language	0.69**	-0.22	0.59*	0.51†
Directions	0.47	-0.05	0.35	0.39
Gives Information in Context	0.71**	0.03	0.69**	0.60*
Gives Information Decontextualized	0.64*	-0.20	0.52†	0.47†
Reads	0.10	0.14	-0.10	0.13
Sings	-0.27	0.05	-0.07	-0.18
Other Talk	0.40	0.06	0.45	0.50†

Note: N=14 classrooms.

†*p*<.10 * *p*<.05 ***p*<.01

[a]Total Talk includes Repeats or Confirms, Requests Language, Gives Directions, Provides Information (contextualized and decontextualized), Reading, and Other.

[b] We include all the dimensions in the Instructional Support scale as well as the Positive Climate dimension from the Emotional Support scale, which had a moderate relationship with some of the LISn categories.

in Spanish and the CLASS. The majority of the lead teacher talk in Spanish was giving directions, singing, repeats or confirms, or requests for language. The lead teacher giving directions in Spanish was related to the Instructional Support scale ($r = 0.56$; $p<.05$), Quality of Feedback ($r = 0.68$; $p<.05$), and Language Modeling ($r = 0.54$; $p<.05$). Lead teacher sings, repeats or confirms, and requests for language in Spanish were not related to any CLASS scores.

The CLASS and the LISn observations were completed by the same observer on the same day, so we examined whether the use of the LISn in combination with the CLASS biased the ratings on the CLASS in any way. The focus on individual interactions between children and teachers had the potential to influence how the observers rated teachers on the CLASS. We examined mean scores on the CLASS for classrooms that had LISn observations and those that did not. No significant differences were found between the two groups (Table D.17). Given the small sample of LISn classrooms and limited power to detect differences, we looked at the unweighted descriptives to see the size of the raw differences. The greatest difference was found on Language Modeling and Instructional Support. For those classrooms with a LISn observation, the mean Language Modeling scores were lower than for those without a LISn observation.[9] A similar pattern is evident for the Instructional Support scale, perhaps because the Language Modeling dimension is part of Instructional Support.

Table D.17. Comparison of Mean CLASS Scores for Classrooms With and Without the LISn

CLASS Domains and Dimensions	Classrooms Without LISn (N = 68)		LISn Classrooms (N = 16)	
	Mean	Standard Deviation	Mean	Standard Deviation
Emotional Support	5.8	0.68	6.0	0.74
Positive Climate	5.8	0.96	6.0	0.93
Negative Climate	1.2	0.63	1.1	0.34
Classroom Organization	5.4	0.80	5.2	0.85
Instructional Support	2.6	1.00	2.2	0.54
Quality of Feedback	2.3	1.11	2.3	0.70
Concept Development	2.0	0.86	2.0	0.73
Language Modeling	3.3	1.67	2.7	0.79

---

[9] All estimates are unweighted.

D. SUMMARY AND CONCLUSIONS

This pilot study shows promise for a new observational measure of teacher-child interactions focused on language (whether English or Spanish). In this context, it is important to remember that we piloted the LISn with a small sample that was not representative of LAUP classes in the same way the child assessments and CLASS observations were. In addition, the low overall frequencies of talk, especially in Spanish, further limited what we could learn about this observational measure. Although more than half of the children experienced at least some talk from an adult, the overall experience in language was not as rich as one might expect to observe in a preschool classroom. Most of the talk heard by the children was in English. The most frequent type of adult talk in English was giving directions. Our observations included few instances of reading or providing information, and almost no instances of elaborating on language. Nevertheless, the information gained from this pilot suggests the potential of the LISn to contribute to understanding how teachers and children interact in classrooms that support children from a variety of language backgrounds.

The LISn allowed us to examine the different types of talk that lead teachers and other adults use in the classroom and to look at the proportion of talk by lead teachers, other adults, and children. We were able to analyze the relative frequency with which children spoke with other children and with adults. Although the time sampling method did not allow an exact match of talk to activity, we could begin to look for patterns in when lead teachers, other adults, and children talk most frequently.

The associations between the LISn and the CLASS provide initial evidence of convergent validity. Expected relationships were moderate to high. In addition to the evidence of validity, these associations suggest that the LISn may be helpful in understanding scores on the various dimensions of the CLASS by providing additional information about the categories of talk that teachers use more and less frequently. Overall, the LISn holds promise for supporting research about the learning interactions for ELL or DLL children. There are several avenues for further research that will enable us to draw stronger conclusions about the promise of the LISn. First among these is work with larger samples of both English- and Spanish-speaking children and more observational cycles completed in each classroom. Other possibilities for future research include further refining definitions for language categories, examining training guidance in rating low frequency interactions, and evaluating whether time sampling is the best way to reliably capture language interactions with children.

E. REFERENCES

Administration for Children and Families. "The Role of Early Head Start Programs in Addressing the Child Care Needs of Low-Income Families with Infants and Toddlers: Influences on Child Care Use and Quality." Princeton, NJ: Mathematica Policy Research, Inc., February 2004.

Beck, Isabella L., and Margaret G. McKeown. "Text Talk: Capturing the Benefits of Read-Aloud Experiences for Young Children." *The Reading Teacher,* vol. 55, no. 1, 2001, pp. 10-20.

Boller, K., S. Sprachman, and the Early Head Start Research Consortium. "Child-Caregiver Observation System." Princeton, NJ: Mathematica Policy Research, Inc., 1998.

Dickinson, David K., R. A. Haine, and C. Howard. "Teacher-Child Verbal Interaction Profile." Newton, MA: Education Development Center, Center for Children and Families, 1996.

Dickinson, David K., and Patton O. Tabors (eds.). *Beginning Literacy with Language.* Baltimore, MD: Brookes Publishing, 2001.

Fuller, Bruce., Sharon L. Kagan, Susan Loeb, and Y. Chang. "Child Care Quality: Centers and Home Settings That Serve Poor Families." *Early Childhood Research Quarterly*, vol. 19, 2004, pp. 505-527.

Goodson, B. D., C. J. Layzer, and W. C. Smith. "Observation Measures of Language Literacy Instruction in Early Childhood (OMLIT)." Cambridge, MA: Abt Associates Inc., 2004.

National Institute for Early Education Research (NIEER). "Support for DLL Classroom Assessment." New Brunswick, NJ: Rutgers University, 2005.

Pearson, Barbara Zurer, and Rebecca Burns. "Promoting Language Development from Birth to 5: In One Language or Two." Washington, DC: Presentation at the National Head Start Dual Language Institute, October 29, 2008.

Pianta, Robert C. "Student-Teacher Relationships Scale." Lutz, FL: Psychological Assessment Resources, 2001.

Pianta, Robert C., Karen M. La Paro, and Bridget K. Hamre. "CLASS Classroom Assessment Scoring System Manual." Charlottesville, VA: Center for Advanced Study of Teaching and Learning, 2006.

Porter, Toni, Rena Rice, Shannon Kearns, Sally Mabon, and Susan Sprachman. "The Child-Caregiver Assessment Tool for Relatives (CCAT-R)." New York: Bank Street College of Education, 2005.

Smith, M., and D. Dickinson. "Early Language and Literacy Classroom Observation (ELLCO)." Baltimore, MD: Brookes, 2002.

Sprachman, Susan, Margaret Caspe, Sally Atkins-Burnett, and Barbara Kennen. "Language Interaction Snapshot (LISn) Coding Manual." Princeton, NJ: Mathematica Policy Research, Inc., March 2008.

Whitebook, M., D. Phillips, D. Bell, and M. Almaraz. "Two Years in Early Care and Education: A Community Portrait of Quality and Workforce Stability." Berkeley, CA: University of California, Institute for Industrial Relations, 2004.

APPENDIX E

DESCRIPTIVE INFORMATION ON LAUP PROGRAMS PARTICIPATING IN THE POWER OF PRESCHOOL DEMONSTRATION PROGRAM

A. INTRODUCTION TO THE UPCOS POWER OF PRESCHOOL SUBSTUDY

First 5 LA received funding from the California First 5 Commission to be one of nine counties to operate a special demonstration program, the Power of Preschool (PoP) program.[1] Jointly funded by the state and First 5 LA, LAUP began implementing the programs in fall 2006 in 13 cities within Los Angeles County's areas of greatest service needs (AGSN). The cities are Azusa, Baldwin Park, Bell, Bell Gardens, Cudahy, Bellflower, Hawthorne, Huntington Park, Lynwood, Montebello, Norwalk, Paramount, and South Gate. First 5 LA selected locations for implementing PoP programs based on such factors as the number of 4-year-old children residing in each zip code, current supply of licensed preschool providers, and the percentage of elementary schools with low Academic Performance Index scores. Approximately 17,337 4-year-olds live in the communities. The goal, as outlined by First 5 CA, is that, by 2010, at least 70 percent of children living in these demonstration cities will have access to and participate in high-quality preschool programs staffed by a highly trained, diverse, and well-compensated workforce that understands and can meet the developmental needs of children and their families. The children in these underserved areas of the county are primarily Latino (First 5 LA estimated that 83 percent are Latino and that about half are English Language Learners [ELLs]). All 4-year-olds in the AGSN communities are eligible for services. To receive funding, all PoP-funded preschool programs were required to meet the same LAUP high-quality standards as the other LAUP programs included in this study.

In this report, we examine the quality and implementation of PoP programs, the characteristics of the representative sample of children and families enrolled in the programs, and children's behavior and development as they progressed from fall 2007 to spring 2008 with respect to the range of domains related to school readiness. Given that we selected all

---

[1] These PoP demonstration programs initially were referred to as Preschool for All (PFA) programs; some counties still use the PFA designation.

26 PoP programs for inclusion in the study sample, by using weighted data from the 23 programs that agreed to participate in UPCOS we are able to generalize the results to all PoP programs, classrooms, and children. One special feature of the PoP substudy is that we administered the Desired Results Developmental Profile-Revised (DRDP-R), a unique teacher observation measure used across California in all counties with PoP demonstration grants. Using the DRDP-R in the PoP programs allowed us to describe how the children's skills and knowledge as observed by teachers within the familiar environment of the daily classroom activities changed from fall to spring. Its inclusion also allowed us to examine the psychometric properties of the DRDP-R, which will be reported separately.

In this appendix, we report the findings from our analysis of the UPCOS data collected in PoP programs in the fall and spring of 2007–2008. We first provide a descriptive overview of the PoP programs, their teachers, and classroom activities. We then describe the children and families enrolled in the programs. Next, we report on the children's development from fall to spring in all the domains of school readiness that we measured.[2] We conclude with a description of children's performance on the DRDP-R.

B. OVERVIEW OF PROGRAMS, CLASSROOMS, AND TEACHERS

1. PoP Program Characteristics

The study sample included 23 PoP programs that enrolled 339 children who were randomly sampled within each program; they were in 42 classrooms. The programs were small: 21.7 percent of them operated just a single classroom, 73.9 percent operated two classrooms, and one program had three classrooms. Class size ranged from 9 to 28 children, with an average 18.5 children. The average child:adult ratio was slightly more than 6:1 but ranged from just 3.6 in the smallest class to a high of 12.6.

2. PoP Teachers

The study included 22 lead teachers of 339 children.[3] Nearly half of the children received instruction from Latino teachers, and more than half of the children received instruction from teachers who reported speaking both Spanish and English at home (Table E.1). Two-thirds of the children had teachers who reported reading to children in both English and Spanish, and one-third of children had teachers who read to children in English only. Similarly, more than half of children had teachers who made classroom instructional presentations in both languages, and 45 percent of children had teachers who made classroom presentations in English only.

On average, children's lead teachers reported 15.1 years' experience in working with children and a median experience level of 12 years (Table E.2). One-quarter of children's

---

[2] With the smaller sample size and the small number of non–Latino children in the PoP program sample, we do not present data on fall-spring growth for the various language groups as we did in the main study.

[3] The results in this section represent the number or percentage of children served by teachers with a given characteristic.

teachers held an associate's degree, and 64 percent held a bachelor's or higher degree (Table E.3). Among classrooms with teachers holding a college degree (AA or higher), more than one-quarter of the children had a teacher whose highest degree was in child development or developmental psychology. An additional 31 percent of children's teachers with a degree obtained their highest degree in early childhood education. All of the children's teachers had taken six or more courses in early childhood education or child development, and 94 percent had obtained a state-issued preschool certificate. Six in 10 children had a teacher who held a teaching certificate or license.

Table E.1. Demographic Characteristics and Language Use by Children's Teachers (Weighted)

Teacher Characteristic	Number of Children	Percentage of Children[b]
Race/Ethnicity[a]		
Latino	149	49.5
Black, non–Latino	88	27.7
White, non–Latino	63	16.7
Asian, non–Latino	0	0.0
Other	30	6.1
Language Spoken at Home[a]		
English	133	45.1
Spanish	0	0.0
English and Spanish	149	53.3
Other language	9	1.6
English and other language	0	0.0
Language Used to Read to Children		
English	83	32.3
English and Spanish	163	67.7
English and another language	0	0.0
Language Used for Presentations[a]		
English	141	44.1
English and Spanish	171	55.9
English and another language	0	0.0
Sample Size	**330**	

Source: Teacher interview, fall 2007.

[a]Categories are mutually exclusive.

[b]Categories with missing data do not sum to 100 percent.

Table E.2.   Experience of Children's Teachers (Weighted)

	Mean (standard error)	Median
Years working with children	15.1 (2.5)	12
Years teaching preschool	12.5 (2.2)	10
Sample Size	**339**	

Source: Teacher interview, fall 2007

Table E.3.   Education and Credentials of Children's Teachers (Weighted)

Teacher's Education and Credentials	Percentage of Children[b]
Highest Grade of School Completed	
High school diploma/equivalent	0.0
Some college but no degree	8.3
Associate's degree (AA)	24.0
Bachelor's degree (BA)	64.0
Graduate or professional degree	3.7
Field in Which Obtained Highest Degree[a]	
Child development/developmental psychology	26.9
Early childhood education	30.9
Other	42.2
College Courses Included 6 or More Classes in Early Childhood Education or Child Development	100.0
Holds a Child Development Associate Credential	60.5
Holds a State-Awarded Preschool Certificate	94.2
Holds a Teaching Certificate or License	59.0
Sample Size	**334**

Source: Teacher interview, fall 2007.

[a]Includes only those teachers holding an associate's degree or higher (N=305).

[b]Categories with missing data do not sum to 100 percent.

3.   Classroom Curricula and Activities

Teachers reported the curricula that they used with the children.[4] The most popular published curriculum was High/Scope, used by 19 percent of the programs and 9 percent of programs used Doors to Discovery (Table E.4). Fewer than 5 percent of programs used any other single published curriculum, but 71 percent of teachers named some other curriculum

---

[4] Estimates represent the primary curriculum used by teachers in the classroom regardless of whether the teacher reported the use of only one curriculum or a combination of curricula.

that they used. About 17 percent of teachers reported the use of a nonspecific curriculum, including those described as "thematic," "Armenian," and "my own." In all, more than half of the curricula reported by programs and classrooms were comprehensive, addressing several domains of development. Almost all teachers reported receiving training and support in curriculum use (90 percent), and reported receiving a median of 13 hours of training and support in the past 12 months. Frequently, the training and support came from the curricula developers (58 percent), other staff at the center (14 percent), or LAUP coaches (14 percent), or other sources (12 percent). Few staff noted receiving training or support from faculty from the school of education (3 percent).

Table E.4.   Primary Curriculum, by Program (Weighted)

	Programs	
Curriculum Name	Number	Percentage
High/Scope	5	19.3
Doors to Discovery	2	8.5
Creative Curriculum	1	4.2
Open Court PreK	1	3.6
Other, non-specific curriculum	17	64.4
Sample Size	**18**	

Source: Teacher interview, winter 2008.

Most of the children's teachers reported that they undertook reading and language activities daily or almost daily (Table E.5). The most common reading and language activities (reported as daily or almost daily by teachers of 90 percent or more of the children's teachers) included letter naming (99 percent), listening to the teacher read stories where children could see the print (98 percent), writing their own name (94 percent), discussing new words (92 percent), and learning about conventions of print (90 percent). The less-common classroom activities (reported by 70 percent or fewer of the children's teachers) included listening to stories in which children could not see the print (23 percent), dictating stories to an adult (61 percent), learning about rhyming words and word families (62 percent), practicing letter writing (68 percent), retelling stories (69 percent), and working on phonics (70 percent).

Ninety percent or more of children's teachers reported that four mathematics activities occurred daily or almost daily (Table E.5): counting out loud (100 percent), working with geometric manipulatives (100 percent), engaging in calendar-related activities (99 percent), and working with counting manipulatives (94 percent). Of the mathematics activities we asked about, only two were reported as occurring daily or almost daily by fewer than 70 of the children's teachers: using music to understand mathematics concepts (66 percent) and working with rulers or other measuring instruments (68 percent).

Table E.5   Frequency of Reading, Language, and Mathematics Classroom Activities (Weighted)

	Never		Monthly		Weekly		Daily or Almost Daily	
	Percentage	Standard Error	Percentage	Standard Error	Percentage	Standard Error	Percent	Standard Error
Reading and Language Activity								
Work on letter naming	0	-	0	-	1.5	1.6	98.5	1.6
Practice writing letters	0	-	0	-	30.6	10.3	68.4	10.3
Discuss new words	0	-	0	-	7.8	5.7	92.2	5.7
Dictate stories to an adult	2.5	2.5	3.4	3.5	33.3	9.8	60.8	9.7
Work on phonics	5.9	5.9	5.8	4.2	18.2	8.2	70.0	10.0
Listen to teacher read stories where they see the print	0	-	2.5	2.5	0	-	97.5	2.5
Listen to teacher read stories where they do not see the print	45.1	10.9	18.9	8.9	12.6	7.5	23.3	9.4
Retell stories	0	-	15.2	8.5	16.1	8.5	68.7	10.8
Learn about conventions of print	0	-	1.3	1.3	8.5	6.2	90.3	6.3
Write own name	0	-	0	-	6.0	5.9	94.0	5.9
Learn about rhyming words and word families	0	-	2.4	2.4	35.3	10.0	62.3	10.1
Learn about common prepositions	0	-	5.5	5.5	10.8	6.1	83.6	7.7
Mathematics Activity								
Count out loud	0	-	0	-	0	-	100.0	0
Work with geometric manipulatives	0	-	0	-	0	-	100.0	0
Work with counting manipulatives	0	-	1.3	1.4	4.6	3.4	94.1	3.7
Play mathematics-related games	0	-	1.3	1.4	27.4	9.5	71.2	9.5
Use music to understand mathematics concepts	0	-	9.9	6.7	23.9	9.0	66.2	10.2
Work with rulers or other measuring instruments	0	-	0	-	31.6	10.1	68.4	10.1
Engage in calendar-related activities	0	-	0	-	1.3	1.4	98.7	1.4
Engage in activities related to telling time	0	-	12.5	8.2	12.1	6.8	75.5	9.8
Engage in activities involving shapes and patterns	0	-	2.6	2.5	16.1	8.2	81.3	8.4

4. Observed Classroom Quality

To obtain information on classroom instructional practices and teacher-child interactions in PoP program classrooms, we randomly sampled one classroom in each program and trained observers to observe an entire class session using the Pre-K Classroom Assessment Scoring System (CLASS) (Pianta, La Paro, and Hamre 2006). The CLASS measures classroom quality across four domains of interaction (see Chapter II and Appendix A), three of which we used in this study: (1) Emotional Support, (2) Classroom Organization, and (3) Instructional Support (Table E.6). The Emotional Support domain assesses the emotional supportiveness of the classroom environment across four dimensions (Positive Climate, Negative Climate, Teacher Sensitivity, and Regard for Student Perspectives) while the three dimensions of Classroom Organization (Behavior Management, Productivity, and Instructional Learning Formats) capture the teacher's ability to organize the classroom and manage behavior to make efficient use of class time for instruction. Instructional Support measures the quality of instructional practices used in the classroom (Concept Development, Quality of Feedback, and Language Modeling). Observers rated each dimension on a 7-point scale, with anchor behavioral descriptions for low (1 - 2), middle (3 - 5), and high (6 - 7) scores. See Box E.1 for a summary of the CLASS domains and their component dimensions.

Although the CLASS does not rely on normative data, the CLASS Technical Appendix (Pianta et al. 2006) reports mean scores from several large-scale studies that used CLASS or its precursor, the Classroom Observation System (COS). COS instruments rated more than 1,500 preschool classrooms in studies such as the National Center for Early Development and Learning's Multi-State Study of Prekindergarten and State-Wide Early Education Programs (MS/SWEEP) and the "My TeachingPartner Study." We also compared PoP data with a recent RAND study of child care quality in California (Karoly et al. 2008) In Table E.7 we present the mean CLASS dimension scores of PoP classrooms alongside scores obtained in other studies of preschool programs.

Box E.1 CLASS Domains and Dimensions

Domain: Emotional Support
  Positive Climate
  Negative Climate
  Teacher Sensitivity
  Regard for Student Perspectives

Domain: Classroom Organization
  Behavior Management
  Productivity
  Instructional Learning Formats

Domain: Instructional Support
  Concept Development
  Quality of Feedback
  Language Modeling

Table E.6.   Observed Classroom Quality Scores, Winter 2008

Domains and Dimensions	N	Mean (standard error)	Reported Response Range	Possible Response Range	Percentage Low (1–2)	Percentage Middle (3–5)	Percentage High (6–7)
CLASS Emotional Support*	18	5.9 (0.1)	4–7	1–7	0.0	41.8	58.2
Positive Climate	19	6.0 (0.2)	5–7	1–7	0.0	29.9	70.1
Negative Climate	19	1.1 (0.1)	1–2	1–7	100.0	0.0	0.0
Teacher Sensitivity	20	5.4 (0.2)	4–7	1–7	0.0	45.3	54.7
Regard for Student Perspectives	19	5.4 (0.2)	4–7	1– 7	0.0	51.2	48.8
CLASS Classroom Organization	19	5.4 (0.2)	4–6	1–7	0.0	85.0	15.1
Behavior Management	20	5.7 (0.2)	4–7	1–7	0.0	37.0	63.0
Productivity	19	5.6 (0.2)	3–7	0.0	0.0	40.9	59.1
Instructional Learning Formats	19	4.9 (0.3)	2–7	1–7	2.9	62.0	35.2
CLASS Instructional Support	20	2.3 (0.2)	1–3	1–7	71.0	29.0	0.0
Concept Development	20	1.8 (0.2)	1–3	1–7	85.9	14.1	0.0
Quality of Feedback	20	2.0 (0.2)	1–3	1–7	71.7	28.3	0.0
Language Modeling	20	3.0 (0.4)	1–6	1–7	57.1	25.4	17.4
Child:Adult Ratio	16	6.5 (0.6)	3.6–12.6	NA	NA	NA	NA
Class Size	21	18.5 (0.8)	9–28	NA	NA	NA	NA

Source: Classroom observations, winter 2008.

*To construct the Emotional Support scale, we reverse-coded the Negative Climate score, according to CLASS scoring instructions. Therefore, low scores on Negative Climate are desirable.

Mean PoP classroom Emotional Support scores on the Positive Climate, Regard for Student Perspectives, and Teacher Sensitivity dimensions were all higher (by 0.7 – 1.2 points) than the averages reported in the CLASS Technical Appendix, and the Negative Climate dimension was up to about 0.5 point lower (Table E.7).[5] Given that the individual dimension scores were all more positive than the CLASS averages reported in the CLASS Technical

---

[5] With few exceptions, the mean and median scores were similar on each dimension and for overall domain scale scores. Therefore, we report the means here for ease of comparison with previously published results from the CLASS Technical Appendix.

Appendix, PoP classrooms also scored higher on the aggregate Emotional Support domain than the averages reported in the CLASS Technical Appendix. Similarly, the PoP Classroom Organization domain ratings were higher than the averages listed in the CLASS Technical Appendix for all three dimensions: Behavior Management, Productivity, and Instructional Learning Formats (with PoP ratings 0.2 – 1.3 points higher).[6] Finally, in the Instructional Support domain, PoP classrooms received average ratings on Concept Development, Quality of Feedback, and Language Modeling dimensions that were lower in all but one instance than those in the CLASS Technical Appendix by 0.1 – 1.2 points. In sum, classroom quality in PoP classrooms observed during the study was somewhat higher than the averages reported elsewhere for the measure in Classroom Organization and Emotional Support but lower in Instructional Support. Similarly, PoP classrooms scored more favorably than classrooms in the RAND study on all dimensions in Emotional Support and Classroom Organization,

Table E.7.   Observed Classroom Quality Scores in PoP Classrooms, Winter 2008, Compared with Scores Found in Studies of Other Preschool Programs

Domains and Dimensions	All PoP Classrooms	"Prepared to Learn": RAND Study of Early Care and Education in California	Multi-State Study of Prekindergarten and Study of State-Wide Early Education Programs Preschool (MS/SWEEP)	My Teaching-Partner	Tulsa County Head Start	Tulsa's Public Schools Pre-K Programs
CLASS Emotional Support	5.9	5.6	NR	NR	NR	NR
Positive Climate	6.2	5.6	5.3	5.2	5.0	5.1
Negative Climate	1.1	1.4	1.6	1.6	1.5	1.4
Teacher Sensitivity	5.5	5.0	4.7	4.3	4.8	4.8
Regard for Student Perspectives	5.4	5.0	NR	4.4	4.6	4.4
CLASS Classroom Organization	5.4	5.0	NR	NR	NR	NR
Behavior Management	5.7	5.3	5.0	4.9	4.4	5.0
Productivity	5.6	5.1	4.5	5.4	5.1	5.2
Instructional Learning Formats	4.8	4.5	3.9	4.6	4.9	4.6
CLASS Instructional Support	2.3	2.7	NR	NR	NR	NR
Concept Development	1.7	2.4	2.1	2.7	2.6	2.8
Quality of Feedback	2.1	2.8	2.0	2.9	3.5	3.3
Language Modeling	2.8	3.0	NR	2.9	3.7	3.5
Number of Classrooms	**20**	**384**	**694**	**164**	**28**	**77**

Source: For MS/SWEEP, and My TeachingPartner, Hamre et al. 2008. For Tulsa Early Childhood Programs, Phillips, Gormley, and Lowenstein 2007. For RAND study: Karoly et al. 2008, p. 103 (we show data from the 4-year-old cohort only).

N = not reported

---

[6] Tulsa County Head Start Study reported a score 0.1 points higher on Instructional Learning Formats.

ranging between .3 points lower on Negative Climate and .3 points higher on Instructional Learning Formats to .6 points higher on Positive Climate. However, scores on each of the dimensions within the Instructional Support domain were lower (by .2 to .7 points)

C. POP CHILDREN AND FAMILIES AND THEIR PROGRAM EXPERIENCE

In this section, we present information on the children's and their families' characteristics as they entered the program in fall 2007. We then describe children's development from fall to spring on the measures of school readiness that we administered.

1. Characteristics of Children and Families Served by PoP Programs

Nearly all children (93 percent) in PoP programs were Latino. About 5 percent were Asian/Pacific Islander, 2 percent multiple race/other, 1 percent white, and no African American children (Table E.8). English and Spanish languages predominated, with 19 percent of children speaking English only, 18 percent Spanish only, 36 percent Spanish primarily, and 25 percent English primarily (not shown, based on language routing). Close to 3 percent primarily spoke languages other than English and Spanish. Nearly 67 percent of children had mothers who were immigrants, and 74 percent had fathers who were immigrants, primarily from Mexico (Table E.9); 81 percent of children lived in households in which at least one parent was born outside the United States (Table E.8), although, nearly all children (96 percent) were born in the United States.

The PoP children were split about evenly between boys and girls. At the fall assessments, children averaged about 52 months of age, or almost 4.5 years. By the spring assessments, children averaged about 59 months, or 5 years (not shown).

Families faced a range of challenges but also exhibited strengths in the face of those challenges at the start of children's PoP experience. Forty-one percent of mothers and 42 percent of fathers lacked a high school diploma or GED, although nearly all families had at least one parent employed full time, usually the father (Table E.9). Despite the high rate of employment among PoP families, just over 40 percent lived below the federal poverty line; 18 percent of families had moved at least once during the past year, and 4 percent had moved at least twice during that time (Table E.8). In the face of poverty and some instability in housing, we also found that 79 percent of children lived with both parents and 62 percent of parents were married. Households were on average small, and children lived in households with 2.7 children and 2.7 adults.

Table E.8. Child and Household Characteristics, Power of Preschool (PoP) Programs (Weighted)

Characteristic	N	Mean/Percentage (Standard Error)
Female	172	52.9 (2.5)
Mean Child Age (Months)	339	51.8 (0.3)
Child Race/Ethnicity		
White, non-Latino	6	1.1 (0.7)
African American, non-Latino	0	0.0 (0.0)
Hispanic/Latino	308	92.5 (2.7)
Asian/Pacific Islander	11	4.8 (2.4)
Multiple race/other	8	1.5 (1.0)
Child Born in the United States	323	96.4 (1.4)
Child's First Language		
English only/primarily	87	28.1 (3.5)
Another language only or primarily	238	69.3 (3.6)
English and another language equally	10	2.6 (1.2)
Mean Number of Adults in Household	317	2.7 (0.1)
Mean Number of Children in Household	317	2.7 (0.1)
Mean Number of Persons in Household	317	5.3 (0.1)
Mean Household Dependency Ratio[a]	317	1.2 (0.0)
Child is Living with		
Both parents	252	79.1 (2.4)
Mother only	61	19.5 (2.4)
Father only	2	0.8 (0.5)
Other	2	0.7 (0.4)
Child's Parents Are[b]		
Married	210	62.3 (2.7)
Divorced/separated	30	10.5 (3.1)
Not married	93	27.2 (3.0)
Both of Child's Parents Born in US	58	18.6 (3.1)
Household Income as Percentage of Federal Poverty Level		
Below 50 percent	25	9.2 (2.2)
50 to 99 percent	88	31.2 (4.8)
100 to 129 percent	43	15.2 (2.7)
130 to 184 percent	44	16.1 (2.7)
185 to 239 percent	17	5.7 (1.3)
240 percent or more	49	22.6 (6.4)
Family Moved at Least Once in Last 12 Months	57	17.9 (1.9)
Family Moved More than Once in Last 12 Months	12	3.6 (1.0)

Source: Parent interview data, fall 2007.

[a]Household dependency ratio is the ratio of children to adults in the household.
[b]Marital status reflects the marital status of the child's mother and father.

Children's home environments provided support for their development (Table E.10). For example, PoP parents reported eating dinner together as a family an average of more than five days a week, and children maintained regular bedtimes more than four days a week (Table E.10). One challenge parents faced was limited understanding of English, with about half of mothers reporting that they did not understand English well or at all. However, about 39 percent of parents reported reading daily to their child in the fall, and about 40 percent of children had more than 25 books in the home. Most families had resources to support their own and their children's physical and mental health (Table E.11). All but about 7 percent of the children had either public or private health insurance. Virtually all children had received a medical checkup in the past year (99 percent), and 77 percent visited a dentist during that time. The majority of parents reported few depressive symptoms, suggesting positive mental health, although 10 percent reported symptoms that could signal moderate or severe depression. Ninety-five percent of children had parents who described their child's physical health as excellent or very good, and 50 percent described their own health as the same (not shown). Parents also reported on sources of social support when faced with six typical emotional, financial, and parenting problems: 77 percent reported having support for all six types of problems (not shown). About 25 percent of parents reported that they received community services, such as English as a Second Language classes, mental health services, and help with medical care. Less than 1 percent of parents and children were exposed to crime or violence in the home or neighborhood within the last year (Table E.11). About 10 percent of children had parents who had witnessed a nonviolent crime in the past year, and about 13 percent had witnessed a violent crime. Six percent of children witnessed domestic violence, and less than 1 percent of children were victims of violence in the home (not shown).

About 25 percent of children were cared for by someone other than their parents before or after LAUP (Table E.10). Excluding the time they spent in LAUP, children spent an average of nearly 17 hours per week in the care of someone other than their parents. Relative care was the most common type of non–LAUP child care (19 percent); 3 percent received before- or after-care in a center-based program, and 3 percent were cared for by a nonrelative in a private home.

Table E.9. **Mother and Father Characteristics, Power of Preschool (PoP) Programs (Weighted)**

Characteristic	Mother		Father	
	N	Mean/Percentage (Standard Error)	N	Mean/Percentage (Standard Error)
Mean Age	329	31.8 (0.3)	314	34.0 (0.3)
Race/Ethnicity				
White, non-Latino	12	4.5 (1.7)	8	2.1 (1.0)
African American, non-Latino	0	0.0 (0.0)	1	0.2 (0.2)
Hispanic/Latino	299	88.8 (3.9)	297	90.0 (2.7)
Asian/Pacific Islander	10	4.5 (2.1)	13	5.4 (2.4)
Multiple race/other	12	2.3 (1.1)	13	2.3 (1.4)
Education[a]				
Less than high school diploma/GED	143	41.1 (4.7)	144	42.3 (4.8)
High school diploma or GED	66	19.7 (2.3)	89	30.0 (3.1)
Some college/Assoc. degree	105	32.7 (4.1)	62	18.7 (2.3)
Bachelor's degree or more	19	6.5 (2.2)	27	8.9 (2.7)
Employment Status[a]				
Full-time	109	39.4 (6.0)	229	92.1 (1.9)
Part-time	46	14.5 (1.9)	14	4.3 (1.4)
Not employed	143	46.1 (5.6)	9	3.6 (1.4)
Born in the US	105	33.2 (4.0)	83	26.3 (2.9)
Time in the US if Born Elsewhere				
5 years or fewer	26	11.0 (2.3)	18	7.9 (1.8)
6 to 10 years	60	28.2 (3.8)	47	21.7 (2.9)
More than 10 years	142	60.8 (3.9)	170	70.5 (3.3)
Country of (non-US) Origin				
Mexico	189	82.6 (4.3)	196	78.0 (5.2)
Central America	11	4.5 (1.6)	15	6.0 (1.2)
Asia	5	2.6 (1.3)	5	2.4 (1.2)
Southeast Asia	4	3.9 (2.9)	5	3.8 (2.6)
Armenia	5	1.0 (1.1)	5	0.9 (1.0)
Other	15	5.3 (2.0)	23	8.9 (2.5)

Source: Parent interview data, fall 2007.

[a]Due to skip patterns, we asked about father's education and employment only if he lived in the household.

Table E.10. Home Environment, Activities, and Routines, Power of Preschool (PoP) Programs (Weighted)

Characteristic	N	Mean/Percentage (Standard Error)
At Home, Child is Usually Spoken to in		
English	123	39.9 (5.3)
Spanish	187	54.1 (5.3)
Another language	25	6.1 (2.5)
Parent Understands English		
Not at all	23	9.8 (2.2)
Not well	99	39.8 (3.8)
Well	53	21.3 (3.1)
Very well/native	70	29.1 (4.1)
When Parent Reads to Child, Language is Usually		
English	147	49.2 (4.8)
Another language	116	35.6 (4.6)
English and another language equally	45	15.2 (2.1)
Frequency of Reading to Child in Past Week		
Never	12	3.8 (1.0)
Once or twice	67	19.8 (2.6)
Three or more times, but not every day	124	37.9 (3.3)
Every day	114	38.6 (4.2)
Number of Children's Books in Home		
0-10	102	31.2 (3.6)
11-25	95	29.7 (3.2)
26-50	80	26.3 (2.2)
51-100	32	10.2 (2.1)
101+	8	2.6 (0.8)
Mean Number of Days per Week Family Eats Dinner Together	317	5.4 (0.1)
Mean Number of Days per Week Child Goes to Bed at Regular Bedtime	313	4.6 (0.1)
Mean Number of Hours Child Sleeps per Night	313	10.6 (0.1)
Current Child Care Outside of LAUP		
Attends child care center or formal program	10	3.3 (1.2)
Receives child care from relative	63	19.1 (3.0)
Receives child care from non-relative	10	2.9 (1.0)
Does not receive care outside of LAUP	227	74.7 (3.7)
Mean hours in LAUP care	335	18.3 (1.9)
Mean Hours per Week in non-LAUP Child Care[a]	90	16.7 (1.6)
Mean hours in out-of-home care (LAUP and non-LAUP)	335	22.8 (2.1)

Source: Parent interview data, fall 2007.

[a]Only parents who reported having child care outside of LAUP indicated the number of hours per week in non-LAUP care.

Appendix E

Table E.11. Family and Parent Well-Being Characteristics Power of Preschool (PoP) Programs (Weighted)

Characteristic	N	Mean/Percentage (Standard Error)
Degree of Parent's Depressive Symptoms[a]		
Not depressed	229	73.3 (2.3)
Mildly depressed	53	16.4 (2.4)
Moderately depressed	15	5.7 (1.6)
Severely depressed	15	4.6 (1.2)
Mean Exposure to Crime and Violence Index[b]	334	0.8 (0.1)
Child's Health Insurance Status		
A private health insurance plan	115	39.0 (5.6)
A public/government insurance plan[c]	192	59.4 (4.2)
No health insurance	23	6.6 (1.7)
Child's Last Regular Doctor Checkup was Less than 1 Year ago	312	98.5 (0.8)
Last Time Child Saw Dentist for Regular Checkup		
6 months ago or less	203	64.2 (2.8)
Between 6 months and a year	40	12.6 (2.2)
More than 1 year ago	20	6.0 (1.3)
Never	54	17.1 (2.5)

Source: Parent interview data, fall 2007.

[a]The short version of the Center for Epidemiological Studies-Depression Short Form ([CES-D] Radloff 1977; Ross et al. 1983) measures levels of depressive symptoms using 12 of the original 20 items from the full CES-D. Four threshold scores are constructed: (1) not depressed—Short Form scores from 0-4, (2) mildly depressed—Short Form scores from 5-9, (3) moderately depressed—Short Form scores from 10-14, and (4) severely depressed—Short Form scores of 15 or greater. In the table, parent's depressive symptoms reflect the symptoms of the parent respondent, typically the child's mother.

[b]Exposure to crime and violence is an index that reflects the severity of families' exposure to crime and violence. Higher scores indicate that household members had been a victim of or were acquainted with someone who had been a victim of violent crime, while lower scores reflect exposure or witnessing of nonviolent crime. Scores greater than 0 indicate that household members had witnessed or been a victim of any sort of crime (i.e., violent and/or non-violent) within the last year. Scores range from 0 to 5.

[c]Public/government insurance includes Medi-Cal or Medicaid, Healthy Families, Healthy Kids, and/or military insurance.

2. PoP Children's Development from Fall 2007 to Spring 2008

We present results from the information we collected on children's development based on the direct assessments, teacher ratings, and parent ratings. (See Appendix A for information on the specific measures used in the UPCOS PoP study.) We organize the findings according to the domains of school readiness that we measured—language and literacy, mathematics, social-emotional development and approaches to learning, and

physical health and motor development. Acknowledging the importance of the DRDP-R for PoP programs, we report at the end of Section 2 on the results from the teachers' ratings.

Language and Literacy Development. Children in PoP programs made progress over the course of the preschool year in language and literacy in all the areas we measured— letter naming, receptive vocabulary, expressive vocabulary, and early writing skills. In Table E.12, we present scores for each language and literacy measure for the children in PoP programs.

Of the 52 upper- and lower-case letters of the alphabet, children scored a mean of 14 in the fall, which increased to 29 in the spring (raw scores were 6 letters named in the fall and 17 in the spring). The average standard score on the Woodcock-Johnson III Spelling subtest increased from fall to spring, for children who took the English version. Standard scores compare each child's performance to that of a national sample of children of the same age. The scores are normed to have a mean of 100 and a standard deviation of 15. An increase in the standardized score for children from fall to spring indicates that children increased their early writing skills by more than the average for children of the same age nationally. In fact, children who took the English version performed at the national mean in the fall (99.7) and one-third of a standard deviation above the national mean in the spring (105.9). Children who were administered the Spelling subtest in Spanish in both fall and spring however, did not perform as strongly relative to peers in either the fall or the spring and scored approximately four-fifths of a standard deviation below the mean in both the fall and spring on the Woodcock-Muñoz III subtest.[7] Scoring one standard deviation below the mean is an important threshold for indicating delayed performance or educational risk.

In the area of English receptive vocabulary (measured with the ROWPVT), LAUP children's skills and proficiency also improved. Children who took the English version of the ROWPVT both fall and spring scored 46.7 and 56.7, respectively. Those who took the Spanish version in the fall entered with the same score as those who took the English version (46.7), but did not score as highly in the spring on the English version (49.0).[8] Children's performance on an assessment of expressive vocabulary (measured with the EOWPVT-SBE) in the fall to spring reflects somewhat less progress. The average child's proficiency in expressive vocabulary was 40.9 in the fall and 45.9 in the spring.

---

[7] Some of the children who took the Spanish version of the subtest in the fall were able to take the English version in the spring; therefore, the spring estimates on the Spanish version reflect the performance of a smaller (and different) sample of children.

[8] The ROWPVT was conceptually scored in Spanish and English in the fall but administered only in English in the spring.

Table E.12. **Language, Literacy, and Mathematics Development, Power of Preschool (PoP) Programs Fall 2007 and Spring 2008 (Weighted)**

Developmental Domain	Data Source	Fall 2007			Spring 2008		
		N	Mean	Standard Error	N	Mean	Standard Error
Language and Literacy							
	Direct Child Assessment						
	PreLAS Total Language Screener score English [a]	339	11.3	0.5	166	13.5	0.5
	PreLAS Total Language Screener score Spanish [a]	186	10.4	0.5	140	11.7	0.5
	Expressive One Word Picture Vocabulary Test – Spanish Bilingual Edition IRT Scale Score	339	40.9	1.0	339	45.9	0.9
	Receptive One Word Picture Vocabulary Test– English Edition IRT Scale Score[b]	82	46.7	1.2	83	56.7	1.8
	Receptive One Word Picture Vocabulary Test– Spanish Bilingual Edition IRT Scale Score [c]	255	46.7	1.1	256	49.0	1.0
	Rapid Letter Naming IRT Scale Score	339	13.7	1.1	339	29.0	1.0
	Woodcock-Johnson III Spelling Standard Score [d]	158	99.7	1.3	158	105.9	1.3
	Woodcock-Muñoz-III Spelling Standard Score [d]	82	87.8	1.6	82	88.7	1.2
Mathematics							
	Direct Child Assessment						
	ECLS-B Mathematics W Score	339	486.6	0.9	339	497.0	0.9

Source: Direct child assessments, fall 2007 and spring 2008.

[a] These scores are raw counts of children's correct responses on Simon Says (Tío Simón Dice) and Art Show (Exposición de Arte).
[b] The means reported here are for the children who took the English version in the fall and their corresponding spring scores (all children took the English version in the spring).
[c] This measure was conceptually scored in the fall and administered in English only in the spring.
[d] National mean for standard scores is 100 with a standard deviation of 15. The means reported here are for the children taking the same version fall and spring. This excludes 76 children who took the Spanish version in the fall but were able to take the English version in the spring. The average spring score for these children was 108.8.

Mathematics Development. As with other developmental areas, children's knowledge and skills in mathematics improved during the preschool year (Table E.12), with the average W score increasing from fall to spring.[9] The ECLS-B mathematics assessment taps children's number concepts, spatial abilities, and measurement proficiency. In the fall, children scored an average of 486.6 on the ECLS-B mathematics assessment.[10] At that time, the average child demonstrated proficiencies in areas such as basic counting skills and understanding relative amounts. In the spring, the average child scored 497.0 and demonstrated skills such as identifying order (for example, "point to the last bicycle"), shapes, numbers, pattern matching, and counting to 10 with one-to-one correspondence.

Social-Emotional Development and Approaches to Learning. In Table E.13, we present scores on each measure of children's prosocial and problem behaviors based on reports from three separate raters: teachers, parents, and our assessors.

Teacher ratings of children's social cooperation, interaction, and independence were higher in the spring than in the fall. Standard scores on these subscales of the Preschool and Kindergarten Behavior Scales-2 [PKBS-2; Merrell 2002] are normed to have a mean of 100 and standard deviation of 15 so that a standardized score of 100 signifies that the child performed at the average level for children his or her age. Average scores on the total positive social skills scale increased from 92.1 in the fall to 102.3 in the spring. On the social cooperation subscale, which measures children's ability to share and to follow classroom rules, scores increased from 97.3 in the fall to 105.0 in the spring. In the area of social interaction, which measures children's prosocial behavior with peers and in the classroom, children's development increased from 87.9 to 98.9, an average increase of 11 points. Similarly, teachers rated children as demonstrating greater social independence and adjustment in the spring than in the fall. Thus, in the fall, teachers rated the cooperative and independent behavior of children in PoP programs below the mean compared to children nationally and at or above the mean by the spring. Teachers did not report marked improvements from fall to spring in children's positive approaches to learning. At both the fall and spring assessment points, teachers reported that children demonstrated behaviors such as attention, persistence, and eagerness to learn "sometimes" (means=1.7 and 2.0, respectively).[11]

---

[9] We report W scores for the ECLS-B mathematics measure. We anchored the item difficulties on the estimates obtained from the ECLS-B and then arithmetically transformed the scores (Woodcock 1999, p. 111) so that the mean difficulty of the items was 500. W scores allow for measurement of change or growth in performance on the same scale over time and are an indicator of absolute rather than relative performance.

[10] The 17-point increase indicates that in the spring the children had about an 87 percent probability of correctly answering something that they had only a 50 percent probability of answering correctly in the fall.

[11] The frequency scale for the teacher-reported approaches to learning was never (0), rarely (1), sometimes (2), or often (3). The scores derived from these data reflect the range of the response scale (0–3).

Table E.13. **Social-Emotional and Approaches to Learning Development, Power of Preschool (PoP) Programs Fall 2007 and Spring 2008 (Weighted)**

Developmental Domain	Data Source	Fall 2007			Spring 2008		
		N	Mean	Standard Error	N	Mean	Standard Error
Social-Emotional							
	Teacher Report						
	Preschool Kindergarten Behavior Scale (PKBS) Standard Score [a]						
	Total Positive Social Skills	326	92.1	2.3	327	102.3	1.5
	Social Cooperation	326	97.3	2.1	327	105.0	1.0
	Social Interaction	326	87.9	3.2	327	98.9	2.3
	Social Independence	326	94.9	1.7	327	102.3	1.1
	Social Skills Rating System Problem Behaviors Standard Score [a]	326	98.1	1.6	326	94.9	1.3
	Parent Report						
	Preschool Kindergarten Behavior Scale IRT Scale Score [b]						
	Social Cooperation/Approaches to Learning	317	2.9	0.0	302	3.0	0.0
	Social Interaction/Independence	317	3.0	0.0	302	3.1	0.0
	Externalizing Problem Behaviors	317	2.0	0.0	302	2.0	0.0
	Internalizing Problem Behaviors	317	2.2	0.0	302	2.2	0.0
Approaches to Learning							
	Direct Child Assessment						
	Executive Functioning						
	Pencil Tapping [c]	338	39.6	2.0	339	65.4	2.5
	Walk a Line Slowly–difference between first attempt and slow attempt [d]	304	0.6 s	0.2	339	1.3 s	0.2
	Leiter Examiner Ratings Scaled Score						
	Attention [e]	339	7.7	0.2	339	9.0	0.1
	Activity [e]	339	8.1	0.2	339	8.8	0.1
	Sociability [e]	339	8.0	0.2	339	9.1	0.1
	ECLS-K Approaches to Learning Scale Teacher-Reported Raw Score [f]	326	1.7	0.1	327	2.0	0.1

Source: Direct child assessments, parent reports, and teacher reports fall 2007 and spring 2008.

[a] National mean for Standard Scores is 100 with a standard deviation of 15.
[b] IRT score rescaled to match the 1-4 response scale.
[c] Pencil tapping is percent of the time the child responded correctly.
[d] Mean score for Walk a Line task is in seconds.
[e] National mean for Scaled Scores is 10 with a standard deviation of 2.
[f] Raw score scaled to reflect 0-3 response scale.

As measured by the Social Skills Rating System (SSRS; Gresham and Elliott 1990), teachers tended to rate children favorably on problem behaviors in the fall and in the spring. Average fall scores were just below the mean. Teachers rated children as having fewer problem behaviors in the spring, with the average spring score about 3 points below the fall mean. Thus, teachers tended to rate children somewhat favorably in this developmental area at the beginning of the preschool year and even more favorably by the spring, indicating lower levels of problem behavior compared with children nationally.

Parents did not report discernible improvements from fall to spring in children's social skills and positive approaches to learning, such as paying attention and persisting in difficult tasks. Children did not show improvements in parent-reported problem behaviors in the spring. However, as with teachers, parents' reports suggest that LAUP children in PoP programs were more likely to demonstrate positive social skills than problem behaviors. For example, in both the fall and spring according to parents' reports, the average child demonstrated positive social skills and approaches to learning "sometimes" (mean ranging from 2.9 to 3.1) while the average child exhibited externalizing and internalizing problem behaviors "rarely" (mean ranging from 2.0 to 2.2).[12]

Assessors used the Leiter-R Examiner Rating Scale (ERS; Roid and Miller 1997) to rate aspects of children's behavior observed during the direct assessment. The ratings provide additional information on children's social skills and behaviors. Assessors rated the extent to which children attended to the assessment tasks, their activity level, and their level of sociability. Assessors rated children favorably on the ERS subscales in the fall and spring, respectively, with mean ratings of 7.7 and 9.0 on attention, 8.1 and 8.8 on activity level, and 8.0 and 9.1 on sociability (Table E.13). These scales are all coded in the positive direction and normed to have a mean of 10 and a standard deviation of 2. Children moved toward national norms during the preschool year; assessors rated children as showing better attention, more appropriate activity levels, and greater sociability in the spring compared with the fall. Fall-spring improvements in assessor ratings corroborated teacher ratings of improved social skills during the year and may reflect children's greater comfort and familiarity with the assessment situation by the spring.

Finally, children were administered direct assessment measures of executive functioning. Executive functioning tasks require children to inhibit a natural response and keep in mind the rules that they are to follow. In the Pencil Tapping task, children were asked to inhibit the natural response to imitate the adult exactly (or to tap repeatedly) and instead to keep in mind that the rule was to do the opposite of what the assessor did; that is, if the assessor tapped twice, the child was to tap once and vice versa. In the fall, 40 percent of the time or less than chance, children in PoP programs effectively inhibited their initial impulse and responded correctly on the Pencil Tapping task. Children effectively remembered the rule and inhibited the initial response across a greater percentage (65 percent) of trials in the spring, indicating that they demonstrated greater self-regulation and impulse control with the

---

[12] The frequency scale for the parent-reported PKBS-2 scales was never (1), rarely (2), sometimes (3), or often (4). The IRT scores derived from these data were scaled to the range of the response scale (1–4).

progression of the year. By way of comparison, a smaller percentage of children responded correctly across trials (59 percent) in a sample of low-income 3- and 4-year-olds in Head Start (Smith-Donald et al. 2007).

The Walk-a-Line Slowly task (Murray and Kochanska 2002; Smith-Donald et al, 2007) was also administered to children. Children were asked to walk on a 6-foot-long line and then to walk the line again as slowly as possible. The difference between their first attempt and the slow attempt is their score. In the fall, children in PoP programs took an average of 0.6 seconds longer on the second attempt. Children took an average of 1.3 seconds longer on the second attempt in the spring. The increased time between attempts indicates greater control of children's impulse to walk the line quickly. In a study in Chicago Head Start classrooms, children averaged approximately 2 seconds longer when asked to walk the line slowly (Smith-Donald et al. 2007).

Children's Physical Health and Motor Development. The Walk-a-Line Slowly task is also an indicator of children's balance or gross motor skills. On the child's first attempt on the task, assessors observed the extent to which the child was able to stay on the line. In the fall, 75 percent of the children were able to stay on the line on their first attempt "almost all of the time" while only 70 percent were able to do so in the spring (Table E.14).

Table E.14. Physical/Motor Development, Power of Preschool (PoP) Programs Fall 2007 and Spring 2008 (Weighted)

Developmental Domain	Data Source	Fall 2007		Spring 2008	
		N	Percentage (SE)	N	Percentage (SE)
Motor Development					
	Direct Child Assessment				
	Walk-a-Line Slowly–Stayed on Line	261	75.0 (4.2)	242	69.5 (4.7)
Physical Health/ Development					
	Direct Child Assessment				
	Body Mass Index (Obesity Indicator)	87	26.3 (5.5)	64	19.7 (2.6)
	Parent Report				
	Child Good or Excellent Health	301	95.1 (1.1)	283	94.2 (1.6)

Source: Direct child assessments and parent interviews, fall 2007 and spring 2008.

Children's height, weight, and the ratio of the two (body mass index [BMI]) are a reflection of children's general health status and well-being. Children are considered to be overweight when their BMI is at or above the 85th percentile for their age and gender and obese if at or above the 95th percentile (Centers for Disease Control and Prevention 2008). According to these criteria, 26 percent of children in PoP programs were obese in the fall, but the percentage decreased to 20 percent in the spring. For comparison, the rate of obesity in similarly aged children in the FACES study was 16 percent (West et al. 2008). The County of Los Angeles Department of Public Health (2008) reported that 22.9 percent of 5th, 7th,

and 9th graders were obese. Although the decline from fall to spring is positive, the fact that one-fifth of the children in PoP programs are obese by the CDC criterion suggests an important issue for LAUP to address.

Finally, reports of children's general health did not change from fall to spring. On a scale of 1 (poor) to 5 (excellent), 95 percent of parents rated their child's overall health as excellent or very good in the fall, and a similar percentage (94 percent) rated their child's health as such in the spring. Thus, parents tended to rate children favorably at the beginning and end of the preschool year, indicating very low levels of perceived serious health problems among children in PoP programs.

DRDP-R Domains. We asked teachers in PoP classrooms to assess their children by using the DRDP-R (California Department of Education, Child Development Division 2006). Teachers rated children by using rubrics and their ongoing assessments and observations of children in the classroom. Thus, the results represent the performance of children in the context of everyday classroom activities and instruction. Here, we present the fall and spring mean scores on the DRDP-R indicators for the sample of children with ratings on all indicators at both time points (N=321).[13]

Teachers rated the children higher in all areas in the spring (Table E.15). The average rating changed from a mean of about 2.0 or "developing" in the fall to a mean greater than 3 ("building") by the spring. Teachers rated the children as strongest in Motor Development and the weakest (in both the fall and spring) in Literacy and Mathematics. However, the change in mean scores from the fall to the spring was greatest for Literacy (+1.38) and for Mathematics (+1.32).

Table E.15. Mean Ratings (with standard errors) on the DRDP-R Scales in the Fall and Spring, 2007–2008

Indicator	Fall		Spring	
	Mean	Standard Error	Mean	Standard Error
Self/Social	1.90	.04	3.11	.04
Regulation	2.10	.05	3.17	.05
Language	1.92	.05	3.10	.05
Approaches to Learning/ Cognition	1.85	.04	3.11	.04
Mathematics	1.69	.05	3.01	.05
Literacy	1.64	.04	3.02	.05
Motor	2.51	.05	3.60	.03
Safety and Health	2.13	.05	3.28	.04

Source: Teacher child reports, fall 2007 and spring 2008.

---

[13] In reviewing the results, we want to consider the potential presence of rater effects. First 5 LA commissioned a separate substudy to investigate the psychometric properties of the DRDP-R. The findings are forthcoming and will be reported separately.

D. SUMMARY OF THE PoP SUBSTUDY

In this appendix, we examined the quality and implementation of a representative sample of PoP programs, the characteristics of the children and families enrolled, and children's behavior and development as they progressed from fall 2007 to spring 2008 with respect to the range of domains related to school readiness. The PoP study sample included 23 PoP programs, 42 classrooms, and 339 children. By using weighted data from the 23 programs that agreed to participate in UPCOS, we can generalize the findings to all PoP programs, classrooms, and children.

Class size in PoP programs ranged from 9 to 28 children, with an average of 18.5 and an average child:adult ratio of 6:5. More than half the children received instruction from bilingual teachers, and two-thirds had teachers who reported reading to them in both English and Spanish. Sixty-four percent of PoP children had lead teachers with a BA degree or higher. Ninety percent of more of the teachers reported that they engaged children in a wide range of literacy and mathematics classroom activities every day or almost daily. Based on the observational measure we used, classroom quality was slightly higher than the averages reported elsewhere in the domains of classroom organization and emotional support but lower in the instructional support domain.

Ninety-three percent of the PoP children were Latino; 54 percent spoke Spanish either primarily or only, and 81 percent lived in households where one or both of the parents was born outside the United States.

In only a few areas did children not show much progression from fall to spring: positive approaches to learning, as reported by teachers, and social skills as reported by parents (including paying attention and persistence). In most domains of school readiness, however, children in PoP programs showed substantial gains from fall to spring. Most notably, the important fall-spring gains occurred in the following areas:

1. Rapid letter naming and spelling (early writing)

2. Both receptive and expressive vocabulary

3. Mathematics (counting, understanding relative amounts, identifying order)

4. Positive social skills as reported by teachers (cooperation, prosocial interaction with peers, social independence, lower problem behavior ratings)

5. Executive functioning as measured in child assessments (self-regulation and impulse control) and positive approaches to learning as reported by the assessors (attention)

6. Positive social skills as reported by assessors (appropriate activity level, sociability)

7. Motor skills and overall health (reduction in rates of obese children)

8. All areas measured by the DRDP-R, with greatest increases in literacy, mathematics, and self/social

E. REFERENCES

California Department of Education, Child Development Division. "Desired Results for Children and Families—Revised." Sacramento, CA: California Department of Education, 2006.

Centers for Disease Control and Prevention. "About BMI for Children and Teens." Retrieved December 2008 from http://www.cdc.gov/nccdphp/dnpa/healthyweight/assessing/bmi/childrens_BMI/about_childrens_BMI.htm#What%20is%20BMI%20percentile, September 2008.

Early, Diane, Oscar Barbarin, Donna Bryant, Margaret Burchinal, F. Chang, Richard Clifford, et al.. *Prekindergarten in Eleven States: NCEDL's Multi-State Study of Prekindergarten and Study of State-Wide Early Education Programs.* Retrieved November 2008, from http://www.fpg.unc.edu/NCEDL/pdfs/SWEEP_MS_summary_final.pdf , 2005.

Gresham, Frank M., and Stephen N. Elliott. "Social Skills Rating System." Circle Pines, MN: American Guidance Service, 1990.

Karoly, Lynn, Bonnie Ghosh-Dastidar, Gail L. Zellman, Michal Perlman, and Lynda Fernyhough. *Prepared to Learn: The Nature and Quality of Early Care and Education for Preschool-Age Children in California.* Santa Monica: RAND, June 2008.

Merrell, Kenneth W. "Preschool and Kindergarten Behavior Scales–Second Edition." Austin, TX: Pro-ED, 2002.

Murray, Kathleen T., and Grazyna Kochanska. "Effortful Control: Factor Structure and Relation to Externalizing and Internalizing Behaviors." *Journal of Abnormal Child Psychology*, vol. 30, 2002, pp. 503-514.

Phillips, Deborah, William T. Gormley, and Amy Lowenstein. "Classroom Quality and Time Allocation in Tulsa's Early Childhood Programs." Paper presented at the biennial meetings of the Society for Research in Child Development, Boston, MA, March 30, 2007.

Pianta, Robert C., Karen M. LaParo, and Bridget K. Hamre. *Classroom Assessment Scoring System Manual, Pre-K.* Baltimore, MD: Paul H. Brookes Publishing Co., 2006.

Roid, Gale H., and Lucy J. Miller. "Leiter-R Performance Scale–Revised." Wood Dale, IL: Stoelting Co., 1997.

Smith-Donald, Radiah, C. Cybele Raver, Tiffany Hayes, and Breeze Richardson. "Preliminary Construct and Concurrent Validity of the Preschool Self-Regulation Assessment (PSRA) for Field-Based Research." *Early Childhood Research Quarterly,* vol. 22, 2007, pp. 173-187.

West, J., L. Tarullo, N. Aikens, and L. Hulsey. "FACES 2006: A Portrait of Head Start Children, Families, and Programs at Program Entry." Presented at Head Start's Ninth National Research Conference, Washington, DC, June 23, 2008.

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

APPENDIX F

DEVELOPMENT OF A RESPECT FOR DIFFERENCES SCALE

In its Master Plan (Hill-Scott 2004), Los Angeles Universal Preschool (LAUP) discusses the importance of respecting diversity and being responsive to different cultural and language needs of families and children. In addition, in one of its first years of operation, LAUP provided training for providers in respecting the family's home culture. This appendix describes the development and results of a new questionnaire designed by Mathematica Policy Research, Inc. (MPR) to assess parent perceptions of how the LAUP goal of respecting home culture and diversity was affecting children and families in the program year 2007-2008. We first describe what we learned in the development of the measure in the spring 2007 pilot and then present the descriptive findings from the Universal Preschool Child Outcomes Study (UPCOS).

We began by searching for existing measures that might yield information about how families perceived the implementation of the principle of respect for, and responsiveness to, different cultures. Unfortunately, we found no measures or scales that were specifically designed to assess program support for home cultures. We considered how to develop survey questions that would provide information from families about this important area. This was a more challenging task than might be anticipated in that it was necessary to take into account families from a variety of cultural and socioeconomic backgrounds. The questions needed to have the same meaning for different groups. Therefore, we used focus groups and cognitive interviews to develop and refine a measure that would be meaningful to linguistically and ethnically diverse parents. Seven focus groups with parents representing five racial/ethnic groups revealed common themes across families regarding the importance of children showing respect for adults, developing knowledge about cultures, and appreciating language differences. These themes were integrated into a measure that was fine-tuned by using cognitive interviewing and then tested with a diverse sample.

Data were collected in three phases. In the first phase, families from diverse racial/ethnic backgrounds were invited to a series of focus groups that enabled researchers to learn about the aspects of home culture that parents wanted and expected the early childhood programs to foster. Two months later, we conducted cognitive interviews with a different group of families to assess their interpretations of survey questions we developed from the focus group findings. In the third phase, a diverse group of parents completed the

survey by phone. Using the information obtained from the focus groups, cognitive interviews, and telephone surveys, we created the final Respect for Differences Scale that we then used in the UPCOS parent interviews in fall 2007 and spring 2008.

1. Focus Groups, Spring 2007

Focus Group Process. During the pilot phase of this study in spring 2007, 57 parents representing 5 racial/ethnic groups—Filipino (n=13), Korean (n=8), Chinese (n=10), African American (n=9), Latino monolingual Spanish-speaking (n=10), and Latino bilingual or monolingual English-speaking (n=7)—were invited to take part in one-hour focus groups at their children's preschool sites. At that point in the year, children had experienced at least eight months of preschool activities. Each separate focus group was composed of parents from one of these racial/ethnic groups. Except for the Latino Spanish-speaking monolingual group, all focus group participants spoke some English. Interpreters were provided for members of the Filipino, Korean, and Chinese focus groups to help clarify some concepts and to enable the parents to express themselves in their native language if they wished. The Latino monolingual Spanish-speaking focus group was conducted in Spanish, and the other Latino focus group was conducted in English. The majority (63.2 percent) of parents participating in the focus group discussions had at least some college education. Nearly one-quarter (24.6 percent) were born in the United States. Most parents spoke English (50.9 percent) and/or Spanish (22.8 percent) as the main language at home.

The purpose of the focus groups was to enable parents from diverse racial/ethnic backgrounds to highlight and discuss how preschools fostered or conflicted with their family's values. Experienced moderators who were native speakers of English or Spanish asked parents about traditions and customs that are important to share with their child, challenges to passing these on when their child was in preschool, and how preschool programs could support home culture(s). Specifically, parents answered questions such as:

- What traditions and customs do you think are important for you and your family to share with your child?

- How can the preschool program support you in sharing your culture with your child?

- What challenges are there to passing on the important parts of your culture to your child?

- What traditions or customs do you have at home that you think will be more difficult to continue now that your child is spending more time with children from other cultural backgrounds?

Focus group discussions were audiotaped and then transcribed for content analysis. MPR researchers independently reviewed the transcripts and developed specific, inductive codes with which to analyze the data (see Strauss and Corbin 1998). We integrated the themes emerging from the focus group findings into a draft scale designed to explore the

ways in which parents perceived preschool programs as fostering home culture and the values that they considered important to preserve through the preschool environment.[1]

Focus Group Themes. Three main themes emerged across all focus groups that highlighted the cultural practices parents found important to share with their children and their feelings about how preschool programs could support home culture.

- **Respect for adults.** Parents from each of the racial/ethnic groups asserted that an important part of their heritage was that children learn to be respectful of adults and talk with adults respectfully. This was especially predominant for the Filipino and Latino parents, who spoke at length about the importance of respecting elders and talking respectfully to adults in the home and in the community (including teachers and doctors).

- **Knowledge of and curiosity about cultural traditions.** Parents believed that developing knowledge of and curiosity about their own and other cultures was an important dimension of the preschool experience. For example, parents wanted children to have opportunities at the preschool to eat foods that were familiar to them but also have opportunities to become interested in foods from new cultures. Parents, particularly African American parents, believed it was important that children not only celebrate holidays from other cultures at school but also develop an understanding for why they are celebrated.

- **Appreciation and use of home language.** Parents reported that programs should promote their home language as well as English. Parents, particularly Spanish-speaking parents, believed that one of the major drawbacks of their children's preschool program was that children were less likely to interact at home in the language of their culture.

2. Cognitive Interviews, Summer 2007

Cognitive Interview Process. After the development of a set of questions based on the issues raised in the focus groups, 53 parents from diverse racial/ethnic groups participated in cognitive interviews that were conducted in either English or Spanish. (See Table F.1 for the beta version of the scale.) The parents from LAUP programs in various parts of Los Angeles County were recruited to participate. Most parents spoke English (66.0 percent) or Spanish (24.5 percent) as the main language at home. The majority of the respondents (60.4 percent) had at least some college education, and many were born in the United States (43.4 percent).

The cognitive interviews were used to assess parental interpretation of the items and the rationale for their responses (see Tourangeau, Rips, and Rasinski 2000). Cognitive

---

[1] See Table F.1 at the end of the document.

interviewing with individual parents proceeded over a period of two weeks. Interviewers recorded each session and took notes on each item, indicating any difficulties that the respondent mentioned. Interviewers also noted the importance or relevance for the parent of the different aspects of the activities addressed in the questions. Parents were asked to provide examples of how their own experiences led to their ratings. The form and content of the instrument changed iteratively over the course of the cognitive interviewing period as the research team coded responses and revised the survey to reflect respondents' reactions to different items.

Cognitive Interview Findings. Cognitive interviewing identified two main problems with the original survey but also presented ways to improve it, including focusing on the child as the level of measurement and reinforcing the need to avoid the use of American idioms.

- **Child outcomes as level of measurement.** The original draft survey consisted of 15 items that asked parents to determine how often their children's early childhood programs promoted children's respect for adults, cultural knowledge, and appreciation for home languages. That is, the first draft of these questions focused on how often programs had certain activities or emphasized certain behaviors. Early cognitive interviews suggested that many parents were unfamiliar with whether or how frequently programs were engaging in specific activities to support the family's home culture. Parents would rate the program positively (that is, parents rated practices as occurring often) but then were unable to provide an example of when such practices occurred. In addition, the variability in parents' frequency of reports of these items was limited (skewed distribution). In a second iteration of the survey, we asked parents to report on how important it was for programs to engage in these activities. Once again, showed limited variability, with most parents citing all activities as very important for programs to do.

  After probing parents in the initial days of cognitive interviewing, the research team created alternative questions focused on observable child behaviors that would result from positive practices in diverse cultural settings. After considerable probing, it became clear that the survey could elicit greater variability by asking questions at the level of child outcomes. For example, in response to the question, "Compared to when [CHILD] started preschool, how interested is [CHILD] in words from other languages now?," 66 percent of parents felt their child was "a lot more willing," 17 percent said "a little more willing," 9 percent said "less willing," and 8 percent said there had been "no change." This result enriched focus group findings by shedding light on the positive nature of diverse language environments, whereas in the focus groups, English classrooms were seen primarily as inhibiting home language development.

  To enable the research team to look at change over the course of the preschool year and to understand what "no change" means (that is, whether it is "no change" because the child had always been willing to play with children from different cultural backgrounds or because she or he did not do so at the start of preschool

and still does not), the team developed a final pretest version of these questions that uses a frequency response scale. This scale is presented in Tables F.2 and F.3. In spring 2008, we asked parents to report on the change over the course of the preschool year (see English version of this in Table F.4).

- **Clarification of unclear terms.** Cognitive interviews also highlighted the need to avoid the use of idioms and terms that were not easily translatable. Based on specific comments from parents in the focus groups, we asked parents in the interviews how frequently children now "talked back" to adults. "Talking back," however, was difficult to translate into Spanish, and many of the Asian parents found the question confusing. One mother perceived "talking back" as a positive (that is, as sticking up for oneself)—a reminder that idioms are culturally bound. Therefore, we changed the question to use the phrase "talk disrespectfully."

3. Fall 2007 Parent Responses

As part of the parent interview, 1,201 parents responded to the questions on the final Respect for Differences scale in the fall of their children's preschool year. The final scale consisted of 12 items that measured the frequency with which parents observed different behaviors with response categories: *always, very often, sometimes, rarely, never* (see Table F.5).[2] A score of 4 on the scale is high and a score of 0 is low.

In spring 2008, we asked parents about any change in the child's respect for differences.[3] The areas in which more than half of the parents reported that children showed a lot of positive change included interest in words from other languages, willingness to play with children from other backgrounds, and showing respect to adults in the family (Table F.6).

We developed scales based on the mean ratings of nine items. The items that we planned to reverse code did not scale well with the other items, and the question about willingness to speak the home language had too much missing data, so these items (d, h, and i) were dropped from the scales. The final scale with nine items had adequate internal consistency (fall α=.73; spring α=.87). On average, parents reported in the fall that children very often showed interest and willingness to engage in activities showing respect for differences (mean=3.2) (Table F.7). In the spring, parents reported that children increased "a little more" in interest and willingness to engage in these activities (mean=3.2). However, some differences were noted across groups. In the fall, the parents of African American children and children in the "other" ethnic group reported that their children more frequently showed interest or willingness to engage in these activities. In the spring, greater change was reported by parents of Hispanic and Asian children (Table F.7).

---

[2] The complete scale in English is shown in Table F.2, and in Spanish in Table F.3.

[3] See Table F.4 for the complete set of questions in English.

To compare the results for the fall and spring Respect for Differences scales, we constructed a variable using both the fall rating and the parent's reported change. We subtracted one from the fall rating if parents said that a child did something less in the spring, added one if the parent said it occurred a "little more," and added two if the parent said that it happened a "lot more." When parents reported no change in the spring, the spring rating was the same as the fall rating. The resulting spring variable ranged from -1 to 6. This spring Respect for Differences variable indicates that parents of children in the "other" ethnic group report that their children most frequently show a willingness to engage in activities showing respect for differences, and parents of white children report less frequent willingness to engage in these activities. However, it is important to keep in mind that parents used a different response scale when reporting in fall and spring.

4. Limitations

A few limitations to the findings from this measure are important to note. First, highly educated parents were overrepresented in the focus groups and cognitive interviews. It is possible that the areas highlighted by these parents might overlook important areas of concern for parents with more limited educational opportunities. In addition, the presence of a translator for the groups that did not speak English or Spanish might have influenced the openness with which parents shared information. The type of self-disclosure in focus groups is not culturally acceptable for some groups, so there might be under-representation of something that is important to one of the groups.

The Respect for Differences scale is based on parent ratings. Parents from different cultural backgrounds might be interpreting the scale in different ways that were not uncovered through cognitive interviewing. In addition, the spring response scale is different from the fall measure and asks parents to make retrospective ratings of change rather than report frequency. This might result in a tendency to over-represent children's likelihood of change over time.

5. Conclusions

The Respect for Differences scale provides a unique look at children's willingness to participate in behaviors that parents from diverse cultural groups reported to be important for their children. Although parents were already positive in the fall about children's interest and willingness to engage in activities that show respect for diverse groups of people, they reported additional positive change from fall to spring in children's interest and willingness. We found differences across ethnic groups. Further exploration of this scale could include asking parents to use the same frequency scale in the spring to look at change over the year, and looking at parents' use of other response scales to determine if some of the differences among ethnic groups are due to the propensity to be more lenient or stringent in using response scales (Fox and Jones 1998).

6. References

Fox, C.M., and J.A. Jones, "Uses of Rasch Modeling in Counseling Psychology Research." *Journal of Counseling Psychology,* Vol. 45, 1998, pp. 30-45.

Hill-Scott, Karen. "Los Angeles Universal Preschool Master Plan." Retrieved March 29, 2008 from http://www.laup.net/downloads/Proj_UPK_MasterPlanFinalDraft.pdf.

Strauss, Anselm, and Juliet M. Corbin. *Basics of Qualitative Research, Techniques and Procedures for Developing Grounded Theory, Second Edition.* Thousand Oaks, CA: Sage Publications, 1998.

Tourangeau, Roger, Lance J. Rips, and Kenneth Rasinski. *The Psychology of Survey Response.* Cambridge, UK: Cambridge University Press, 2000.

Table F.1.   Beta Version of Respect for Differences in Preschool

Please tell us how often your child's preschool has the following activities. If you do not know if an activity takes place in your child's preschool, you can check "don't know." Try to answer as many items as you can.

My child's preschool . . .	Never	Rarely	Sometimes	Often	DON'T KNOW
a. Celebrates holidays in meaningful ways, for example, explains why we celebrate the holiday	1 ☐	2 ☐	3 ☐	4 ☐	d ☐
b. Shows respect for all different kinds of families	1 ☐	2 ☐	3 ☐	4 ☐	d ☐
c. Helps my child learn about a variety of different cultures	1 ☐	2 ☐	3 ☐	4 ☐	d ☐
d. Teaches my child how to speak in ways that show respect for others	1 ☐	2 ☐	3 ☐	4 ☐	d ☐
e. Gives my child opportunities to express him/herself in our home language (lets him/her talk in our home language)	1 ☐	2 ☐	3 ☐	4 ☐	d ☐
f. Teaches my child to respect people from other backgrounds or cultures	1 ☐	2 ☐	3 ☐	4 ☐	d ☐
g. Has pictures or books that reflect my family's culture. For example, has books that show families like mine doing the kinds of things my family does......................................	1 ☐	2 ☐	3 ☐	4 ☐	d ☐
h. Sings songs in different languages	1 ☐	2 ☐	3 ☐	4 ☐	d ☐
i. Asks children to talk about or share food from different cultures	1 ☐	2 ☐	3 ☐	4 ☐	d ☐
j. Has pretend food in the play area that is familiar to my child (similar to what we eat at home)	1 ☐	2 ☐	3 ☐	4 ☐	d ☐
k. Teaches my child to show respect to adults	1 ☐	2 ☐	3 ☐	4 ☐	d ☐

Please tell me how important it is that you child's classroom does the following. Is it not at all important, a little important, somewhat important, or very important? Try to answer as many items as you can.

How important is it that your child's classroom . . .	Not at all Important	A Little Important	Somewhat Important	Very Important	DON'T KNOW
a. Celebrates holidays in meaningful ways, for example, explains why we celebrate the holiday	1 ☐	2 ☐	3 ☐	4 ☐	d ☐
b. Shows respect for all different kinds of families	1 ☐	2 ☐	3 ☐	4 ☐	d ☐
c. Helps my child learn about a variety of different cultures	1 ☐	2 ☐	3 ☐	4 ☐	d ☐
d. Teaches my child how to speak in ways that show respect for others	1 ☐	2 ☐	3 ☐	4 ☐	d ☐
e. Gives my child opportunities to express him/herself in our home language (lets him/her talk in our home language)	1 ☐	2 ☐	3 ☐	4 ☐	d ☐
f. Teaches my child to respect people from other backgrounds or cultures	1 ☐	2 ☐	3 ☐	4 ☐	d ☐
g. Has pictures or books that reflect my family's culture. For example, has books that show families like mine doing the kinds of things my family does......	1 ☐	2 ☐	3 ☐	4 ☐	d ☐
h. Sings songs in different languages	1 ☐	2 ☐	3 ☐	4 ☐	d ☐
i. Asks children to talk about or share food from different cultures	1 ☐	2 ☐	3 ☐	4 ☐	d ☐
j. Has pretend food in the play area that is familiar to my child (similar to what we eat at home)	1 ☐	2 ☐	3 ☐	4 ☐	d ☐
k. Teaches my child to show respect to adults	1 ☐	2 ☐	3 ☐	4 ☐	d ☐

Table F.2.   The Respect for Differences Scale—English Version Fall

Please tell me how often each of these statements is true of your child.

	Always	Very often	Sometimes	Rarely	Never	DON'T KNOW	NA
a. is willing to try foods from different cultures Is (he/she) always, very often, sometimes, rarely, or never willing to try foods from different cultures?	4	3	2	1	0	d	
b. is respectful to adults in your family Is (he/she) always, very often, sometimes, rarely, or never respectful to adults?	4	3	2	1	0	d	
c. is willing to play with children from different backgrounds Is (he/she) always, very often, sometimes, rarely, or never willing?	4	3	2	1	0	d	
d. talks disrespectfully to adults in your home Does (he/she) talk disrespectfully always, very often, sometimes, rarely, or never?	4	3	2	1	0	d	
e. is willing to play with children who have special needs or physical or mental disabilities Is (he/she) always, very often, sometimes, rarely, or never willing?	4	3	2	1	0	d	N/A
f. is interested in different cultures Is (he/she) always, very often, sometimes, rarely, or never interested in different cultures?	4	3	2	1	0	d	

		Always	Very often	Sometimes	Rarely	Never	DON'T KNOW	NA
g.	is interested in why different people celebrate different holidays Is (he/she) always, very often, sometimes, rarely, or never interested in why different people celebrate different holidays?	4	3	2	1	0	d	
h.	avoids playing with children from different cultural backgrounds Does (he/she) avoid children from different cultural backgrounds always, very often, sometimes, rarely, or never?	4	3	2	1	0	d	
i.	is willing to speak your home language Is (he/she) always, very often, sometimes, rarely, or never willing to speak your home language?	4	3	2	1	0	d	N/A
j.	is willing to take part in events that are important to your family's cultural traditions Is (he/she) always, very often, sometimes, rarely, or never willing to take part in these events?	4	3	2	1	0	d	
k.	is interested in words from other languages Is (he/she) always, very often, sometimes, rarely, or never interested in words from other languages?	4	3	2	1	0	d	
l.	is respectful of people from different cultures or backgrounds Is (he/she) always, very often, sometimes, rarely, or never respectful of people from different cultures or backgrounds?	4	3	2	1	0	d	

-

Table F.3.   The Respect for Differences Scale –Spanish Version Fall

Por favor dígame con qué frecuencia cada una de las siguientes declaraciones es cierta en relación a su niño/a.

	SIEMPRE	MUCHAS VECES	A VECES	CASI NUNCA	NUNCA	DON'T KNOW	NA
a. está dispuesto(a) a probar comidas de diferentes culturas. ¿Está (él/ella) siempre, muchas veces, a veces, casi nunca o nunca está dispuesto(a) a probar comidas de diferentes culturas?	4	3	2	1	0	d	
b. es respetuoso(a) a adultos en su familia. ¿Es (él/ella) siempre, muchas veces, a veces, casi nunca o nunca respetuoso(a) a adultos en su familia?	4	3	2	1	0	d	
c. está dispuesto(a) a jugar con niños de diferentes orígenes ¿Está (él/ella) siempre, muchas veces, a veces, casi nunca o nunca está dispuesto(a)?	4	3	2	1	0	d	
d. habla faltándoles el respeto a adultos en su hogar ¿(Él/Ella) habla sin respeto siempre, muchas veces, a veces, casi nunca o nunca?	4	3	2	1	0	d	
e. está dispuesto(a) a jugar con niños con necesidades especiales o con incapacidades físicas o mentales ¿Está (él/ella) siempre, muchas veces, a veces, casi nunca o nunca dispuesto(a)?	4	3	2	1	0	d	N/A
f. está interesado(a) en diferentes culturas ¿Está (él/ella) siempre, muchas veces, a veces, casi nunca o nunca interesado(a) en diferentes culturas?	4	3	2	1	0	d	

Appendix F

		SIEMPRE	MUCHAS VECES	A VECES	CASI NUNCA	NUNCA	DON'T KNOW	NA
g.	está interesado(a) en por qué diferentes personas celebran diferentes días de fiesta o feriados ¿(Él/Ella) siempre, muchas veces, a veces, casi nunca o nunca sabe acerca del por qué diferentes personas celebran diferentes días de fiesta o feriados?	4	3	2	1	0	d	
h.	evita (CHILD) jugar con niños de diferentes orígenes culturales ¿(Él/Ella) evita estar con niños de diferentes orígenes siempre, muchas veces, a veces, casi nunca o nunca?	4	3	2	1	0	d	
i.	está dispuesto(a) a hablar su lengua materna ¿Está (él/ella) siempre, muchas veces, a veces, casi nunca o nunca dispuesto(a) a hablar su lengua materna?	4	3	2	1	0	d	N/A
j.	está dispuesto(a) a tomar parte en eventos que son importantes en las tradiciones culturales de su familia ¿Está (él/ella) siempre, muchas veces, a veces, casi nunca o nunca dispuesto(a) a tomar parte en estos eventos?	4	3	2	1	0	d	
k.	está interesado(a) en palabras de otros idiomas ¿Está (él/ella) siempre, muchas veces, a veces, casi nunca o nunca interesado(a) en palabras de otros idiomas?	4	3	2	1	0	d	
l.	es respetuoso(a) hacia personas de diferentes culturas u orígenes ¿Es (él/ella) siempre, muchas veces, a veces, casi nunca o nunca respetuoso(a) hacia personas de diferentes culturas u orígenes?	4	3	2	1	0	d	

Appendix F

Table F.4.   The Respect for Differences Scale –English Version Spring

Please think of how [CHILD] is now compared to how (he/she) was when (he/she) started attending this preschool. Compared to when [CHILD] started [PRESCHOOL], how (INSERT ITEM)?

	A Lot More	A Little More	Less[4]	No Change	DON'T KNOW	NA
a. willing is [CHILD] to try foods from different cultures now? Is (he/she) a lot more willing, a little more willing, less willing, or has there been no change (since (he/she) started attending this preschool)?	4	3	2	1	d	
b. respectful is [CHILD] to adults in your family now? Is (he/she) a lot more respectful, little more respectful, less respectful to adults or has there been no change?	4	3	2	1	d	
c. willing is [CHILD] to play with children from different cultures and backgrounds now? ... Is (he/she) a lot more willing, a little more willing, less willing, or has there been no change?	4	3	2	1	d	
d. much does [CHILD] talk disrespectfully to adults in your home now? Does (he/she) talk disrespectfully a lot more often, a little more often, less often, or has there been no change?	4	3	2	1	d	
e. willing is [CHILD] to play with children who have special needs or physical or mental disabilities now? Is (he/she) a lot more willing, a little more willing, less willing, or has there been no change?	4	3	2	1	d	N/A
f. interested is [CHILD] in different cultures now? Is (he/she) a lot more interested, a little more interested, less interested, or has there been no change?	4	3	2	1	d	
g. interested is [CHILD] about why people celebrate different holidays now? Is (he/she) a little more interested, less knowledgeable, or has there been no change?	4	3	2	1	d	

---

[4] "No Change" was recoded to be 2, "Less" was recoded to be 1.

		A Lot More	A Little More	Less[5]	No Change	DON'T KNOW	NA
h.	much does [CHILD] avoid playing with children from different cultural backgrounds now? Does (he/she) avoid children from different cultural backgrounds a lot more often, a little more often, less often, or has there been no change?	4	3	2	1	d	
i.	willing is [CHILD] to speak your home language now? Is (he/she) a lot more willing, a little more willing, less willing, or has there been no change?	4	3	2	1	d	N/A
j.	willing is [CHILD] to take part in events that are important to your family's cultural traditions now? Is (he/she) a lot more willing, a little more willing, less willing, or has there been no change?	4	3	2	1	d	
k.	interested is [CHILD] in words from other languages now? Is (he/she) a lot more interested, a little more interested, less interested, or has there been no change?	4	3	2	1	d	
l.	respectful is [CHILD] of people from different cultures or backgrounds now? Is (he/she) a lot more respectful, a little more respectful, less respectful, or has there been no change?	4	3	2	1	d	

---

[5] "No Change" was recoded to be 2, "Less" was recoded to be 1.

Appendix F

Table F.5: **Percentage of Responses for the Fall 2007 Administration of the Respect for Differences Scale (Weighted)**

Please tell me how often each of these statements is true of your child: [CHILD]	Rarely or Never Percentage (Standard Error)	Sometimes Percentage (Standard Error)	Very Often Percentage (Standard Error)	Always Percentage (Standard Error)
a. Is willing to try foods from different cultures	20.0 (1.2)	37.5 (1.4)	15.2 (1.4)	27.3 (1.7)
b. Is respectful to adults in your family	1.5 (0.3)	13.7 (1.1)	24.2 (1.6)	60.6 (2.1)
c. Is willing to play with children from different backgrounds	1.6 (0.4)	8.7 (1.1)	14.4 (1.4)	75.3 (1.8)
d. Talks disrespectfully to adults in your home (*not reverse coded*)	82.6 (1.6)	11.0 (1.1)	2.0 (0.4)	4.4 (1.0)
e. Is willing to play with children who have special needs or physical or mental disabilities	7.4 (0.8)	18.5 (1.7)	16.5 (1.6)	57.5 (2.4)
f. Is interested in different cultures	10.5 (1.0)	31.1 (1.7)	18.5 (1.4)	39.9 (2.1)
g. Is interested about why different people celebrate different holidays	17.4 (1.3)	26.1 (1.5)	16.7 (1.5)	39.8 (2.1)
h. avoids playing with children from different cultural backgrounds (*not reverse coded*)	85.7 (1.4)	7.8 (1.0)	2.0 (0.5)	4.5 (0.7)
i. Is willing to speak your home language	3.7 (0.8)	8.5 (1.3)	9.3 (1.3)	78.5 (2.0)
j. Is willing to take part in events that are important to your family's cultural traditions	3.4 (0.5)	12.6 (1.2)	14.0 (1.3)	69.9 (1.6)
k. Is interested in words from other languages	8.7 (0.9)	20.7 (1.4)	16.5 (1.4)	54.1 (1.8)
l. Is respectful of people from different cultures or background?	1.6 (0.4)	7.7 (0.9)	17.0 (1.6)	73.6 (1.9)

Table F.6: **Percentage of Responses for the Spring 2008 Administration of the Respect for Differences Scale (Weighted)**

Please think of how [CHILD] is now compared to how (he/she) was when (he/she) started attending this preschool. Compared to when [CHILD] started [PRESCHOOL], how (INSERT ITEM)?	Less Percentage (Standard Error)	No Change Percentage (Standard Error)	A Little More Percentage (Standard Error)	A lot More Percentage (Standard Error)
a. Willing to try foods from different cultures	2.3 (0.4)	19.5 (1.6)	33.0 (1.6)	45.2 (2.1)
b. Respectful to adults in your family	2.5 (0.5)	18.6 (1.6)	24.4 (1.6)	54.6 (2.2)
c. Willing to play with children from different backgrounds	0.6 (0.2)	21.1 (2.1)	18.9 (1.4)	59.4 (2.2)
d. Talk disrespectfully to adults in your home (*not reverse coded*)	37.8 (2.2)	47.2 (2.2)	8.4 (1.2)	6.6 (0.9)
e. Willing to play with children who have special needs or physical or mental disabilities	1.5 (0.4)	43.1 (2.6)	19.6 (1.6)	35.8 (2.7)
f. Interested in different cultures	1.1 (0.3)	24.2 (1.9)	33.2 (1.7)	41.5 (2.3)
g. Interested about why different people celebrate different holidays	1.5 (0.4)	21.7 (1.7)	31.5 (1.5)	45.2 (2.1)
h. Avoid playing with children from different cultural backgrounds (*not reverse coded*)	24.3 (1.9)	57.8 (2.4)	8.9 (1.0)	9.0 (1.0)
i. Willing to speak your home language?	4.6 (1.3)	91.3 (1.9)	1.8 (0.8)	2.3 (0.8)
j. Willing to take part in events that are important to your family's cultural traditions	2.2 (0.4)	30.7 (2.4)	24.7 (1.6)	42.5 (2.4)
k. Interested in words from other languages	1.4 (0.3)	15.0 (1.7)	23.5 (1.3)	60.2 (1.9)
l. Respectful of people from different cultures or backgrounds	1.0 (0.3)	28.4 (2.2)	21.5 (1.5)	49.1 (2.4)

Table F.7. **Mean Parent Ratings by Ethnic Background (Weighted)**

	Asian (N=67) Mean (Standard Error)	African-American (N=79) Mean (Standard Error)	Hispanic (N=862) Mean (Standard Error)	White (N=76) Mean (Standard Error)	Other (N=45) Mean (Standard Error)	Total (N=1131) Mean (Standard Error)
Fall Respect for Differences (possible range 0 to 4)	3.2 (0.1)	3.3 (0.0)	3.1 (0.0)	3.1 (0.1)	3.5 (0.1)	3.2 (0.0)
Spring Respect for Differences (possible range -1 to 6)	4.5 (0.1)	4.4 (0.2)	4.4 (0.0)	4.1 (0.1)	4.6 (0.2)	4.4 (0.0)
Spring Parent Rating of Change (possible range 0 to 4)	3.3 (0.1)	3.0 (0.2)	3.3 (0.0)	2.9 (0.1)	3.1 (0.1)	3.2 (0.0)