

**Informing the Performance-Based
Contract Between First 5 LA and
LAUP – Volume 2: Background
and Supporting Analyses for the
2012–2013 Study of Child
Progress**

Final Report

May 27, 2014

Sally Atkins-Burnett

Yange Xue

Emily Moiduddin

Nikki Aikens

Judy Cannon

MATHEMATICA
Policy Research



Champions For Our Children

Contract Number:
07110

Mathematica Reference Number:
40079.212

Submitted to:
First 5 LA
750 N. Alameda Street, Suite 300
Los Angeles, CA 90012
Project Officer: Kelly Goods

Submitted by:
Mathematica Policy Research
1100 1st Street, NE
12th Floor
Washington, DC 20002-4221
Telephone: (202) 484-9220
Facsimile: (202) 863-1763
Project Director: Emily Moiduddin

**Informing the Performance-
Based Contract Between First 5
LA and LAUP – Volume 2:
Background and Supporting
Analyses for the 2012–2013
Study of Child Progress**

Final Report

May 27, 2014

Sally Atkins-Burnett
Yange Xue
Emily Moiduddin
Nikki Aikens
Judy Cannon

MATHEMATICA
Policy Research

UPCOS-6 COPYRIGHT PERMISSIONS

Woodcock-Johnson® III (WJ III®), WJ III® Tests of Achievement. Copyright © 2001, 2007, The Riverside Publishing Company. Reproduced with permission of the publisher. All rights reserved.

No part of this work may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying and recording or by any information storage or retrieval system without the proper written permission of The Riverside Publishing Company unless such copying is expressly permitted by federal copyright law. Address inquiries to Contracts and Permissions Department, The Riverside Publishing Company, 3800 Golf Road, Rolling Meadows, Illinois 60008-4015.

Batería III Woodcock-Muñoz®. Copyright © 2004, 2007, The Riverside Publishing Company. Reproduced with permission of the publisher. All rights reserved.

No part of this work may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying and recording or by any information storage or retrieval system without the proper written permission of The Riverside Publishing Company unless such copying is expressly permitted by federal copyright law. Address inquiries to Contracts and Permissions Department, The Riverside Publishing Company, 425 Spring Lake Drive, Itasca, Illinois 60143-2079.

preLAS 2000, by Sharon E. Duncan, Ph.D., and Edward A. DeAvila, Ph.D. Copyright © 1998 CTB/McGraw-Hill LLC, a subsidiary of The McGraw-Hill Companies, Inc. Reproduced by permission of CTB/McGraw-Hill LLC.

Leiter International Performance Scale-Revised Examiner Ratings. Copyright © 1997, 2002 Stoelting Co., 620 Wheat Lane, Wood Dale, IL 60191. All rights reserved. The material may not be distributed to countries who have no copyright relations with the United States of America.

Expressive One-Word Picture Vocabulary Test (EOWPVT). Copyright © 2000, Academic Therapy Publications, 20 Commercial Boulevard, Novato, CA, 94949-6191. All rights reserved. Reproduced by permission of Academic Therapy Publications.

ACKNOWLEDGMENTS

The authors of this report have many people to thank for making this study possible. Our project officers, Sharon Murphy and Artineh Samkian, worked closely with us at every step along the way. Staff at the Los Angeles Universal Preschool (LAUP) program—particularly Schellee Rocher and Rosa Valdes—kept us focused on what matters for children, families, and their programs, and contributed tremendously to the substance of the project.

Susan Sprachman, as survey director, and Rebecca Mason, as assistant survey director, coordinated all aspects of the project related to data collection and made sure that all field activities were carried out with great skill. Anne Self partnered with them valiantly in all of their efforts. The survey team could not have carried out its work without the efforts of the LAUP program directors, who provided tremendous support. The LAUP teachers gave generously of their class time, which we know was not always convenient, allowing us to assess the children. Especially central to the success of the study was the participation of the LAUP children and their parents. We are deeply indebted to them for their time and cooperation.

Enrolling programs into the study, arranging the schedules for data collection, working with the programs to establish the eligible sample and obtain parental consent, and conducting quality assurance activities are critical tasks on a study like UPCOS; we are grateful to the Mathematica Policy Research and Juárez & Associates (J&A) staff who coordinated this process. From Mathematica, this includes Anna Situ, Chrissie Grover-Roybal, Daniel Mendieta, Flor Garay, Elva Lopez, and Laura Gonzalez-Orellana. From J&A, this includes Elizabeth Perez-LoPresti, Sandy Figueroa, Claudia Benavides, and Sara Montolla. We thank all of the field staff who carried out assessments with great respect for the research process, the children being assessed, and their families.

Our information services team was led by Scott Reid. Finally, we give special thanks to Donna Dorsey and Alfreda Holmes for the tremendous effort that went into producing this report.

CONTENTS

I	INTRODUCTION	1
II	STUDY BACKGROUND	3
	A. Domains of Child Development and Selected Measures.....	3
	B. Target-Setting Procedures	4
	C. Targets Established for the 2012–2013 Program Year.....	9
	D. Evidence of Validity of the Targets.....	9
III	STUDY METHODS.....	11
	A. Sample	11
	B. Measures.....	12
	C. Documenting Family Backgrounds.....	15
	D. Procedures	15
	E. Scoring	16
IV	FALL-SPRING PROGRESS BY LANGUAGE GROUP	19
	A. Approach to Analysis	19
	B. Progress by Language Group	20
V	FALL-SPRING PROGRESS FOR CHILDREN WHO SWITCHED LANGUAGE OF ASSESSMENT: PRELIMINARY ANALYSIS.....	29
	A. Approach to Analysis	29
	B. Progress Among Children Who Switched Language of Assessment.....	30
	C. Summary and Implications.....	35
VI	FALL AND SPRING SCORES BY PERFORMANCE RELATIVE TO TARGETS: AN EXAMINATION OF QUARTILES AND LANGUAGE GROUPS	37
	A. Approach to Analysis	37
	B. Quartile Analysis.....	37
	C. Language Group Analysis.....	38
	D. Summary and Implications.....	44

VII AN EXPLORATORY ANALYSIS OF CHILD AND FAMILY CHARACTERISTICS ASSOCIATED WITH CHILDREN'S PROGRESS45

A. Approach to Analysis45

B. Factors Associated with Child Progress46

C. Summary48

REFERENCES53

TABLES

II.1	Child Outcomes Measures and Developmental Domains Addressed in the First 5 LA-LAUP Performance-Based Contract	4
II.2	Final Targets for Child Progress in LAUP Programs During the 2012–2013 Program Year, by Domain and Measure	10
III.1	Number of Children Assessed in Fall 2012 and Spring 2013, by Language Group (N = 573).....	12
IV.1	Means and Standard Errors of <i>preLAS</i> : Fall 2012 and Spring 2013, by Language Group.....	21
IV.2	Means and Standard Errors of EOWPVT: Fall 2012 and Spring 2013, by Language Group.....	22
IV.3	Means and Standard Errors of RLN: Fall 2012 and Spring 2013, by Language Group.....	23
IV.4	Means and Standard Errors of WJ-III Spelling: Fall 2012 and Spring 2013, by Language Group	24
IV.5	Means and Standard Errors of WM-III Ortografía: Fall 2012 and Spring 2013, by Language Group	25
IV.6	Means and Standard Errors of WJ-III Applied Problems: Fall 2012 and Spring 2013, by Language Group	25
IV.7	Means and Standard Errors of WM-III Problemas Aplicados: Fall 2012 and Spring 2013, by Language Group	26
IV.8	Percentage of Children with Leiter Examiner Ratings Scaled Score in the Acceptable Range on the Leiter-R: Fall 2012 and Spring 2013, by Language Group.....	27
IV.9	Percentage of Children Who Were Overweight or Obese: Fall 2012 and Spring 2013, by Language Group	28
V.1	<i>preLAS</i> Scores for Children Who Switched Language in the Spring and Those Who Did Not.....	31
V.2	Means and Standard Errors for WM-III and WJ-III Subtest W Scores, by Whether Children Switched Language of Assessment: Fall 2012 and Spring 2013	32
V.3	Means and Standard Errors for EOWPVT and RLN, by Whether Children Switched Language of Assessment: Fall 2012 and Spring 2013.....	34
V.4	Likelihood That Children Switch from Spanish to English Assessment Path	35

VI.1	Means and Standard Errors for UPCOS-6 Language, Literacy, and Math Standardized Measures: Fall 2012 and Spring 2013 for Children Whose Change Score Met or Exceeded Change Specified in Level 1 Targets, by Fall Quartile	39
VI.2	Means and Standard Errors for UPCOS-6 Language, Literacy, and Math Standardized Measures: Fall 2012 and Spring 2013 for Children Whose Change Score Fell Below Change Specified in Level 1 Targets, by Fall Quartile	40
VI.3	Means and Standard Errors for UPCOS-6 Language, Literacy, and Math Standardized Measures: Fall 2012 and Spring 2013 for Children Whose Change Score Met or Exceeded Change Specified in Level 1 Targets, by Language Group	41
VI.4	Means and Standard Errors for UPCOS-6 Language, Literacy, and Math Standardized Measures: Fall 2012 and Spring 2013 for Children Whose Change Score Fell Below Change Specified in Level 1 Targets, by Language Group	42
VII.1	Independent Variables Included in HLM Models	46
VII.2	Relationships of Child and Family Characteristics with Children’s Progress on the EOWPVT	50
VII.3	Relationships of Child and Family Characteristics with Children’s Progress on the RLN	51
VII.4	Relationships of Child and Family Characteristics with the Likelihood of Meeting the RLN Target.....	52

FIGURES

V.1	Mean W Scores for the WM-III and WJ-III Subtests, by Whether Children Switched Language of Assessment: Fall 2012 and Change to Spring 2013.....	33
-----	-------------------------------------------------------------------------------------------------------------------------------------------------	----

I. INTRODUCTION

In February 2004, First 5 LA adopted a 10-year universal preschool master plan to increase the availability of high quality preschool spaces in Los Angeles County and created Los Angeles Universal Preschool (LAUP) to implement the plan. A shared objective of First 5 LA and LAUP is to support early learning and development for the four-year-old children that have an opportunity to participate in LAUP programs. To meet this objective, First 5 LA and LAUP have begun to track the progress of children during their year in LAUP programs as an element of the performance-based contract between the two organizations. Beginning in the 2009–2010 program year, Mathematica Policy Research worked with First 5 LA and LAUP to identify the domains of development they sought to track, identify appropriate measures, and set targets for progress across the year. Those targets were applied for the first time in the 2011–2012 program year based on data collected by Mathematica as part of the fifth phase of the Universal Preschool Child Outcomes Study (UPCOS).

Because LAUP has a new population of children every year, First 5 LA and LAUP have agreed to revisit the targets each year to consider whether the most recent data regarding LAUP children signal a need to change the targets or whether there are any factors within or external to LAUP that might influence what First 5 LA and LAUP want prioritized within the targets. Thus, the targets were revised for the 2012–2013 program year in a collaborative process between First 5 LA and LAUP that was facilitated by Mathematica. As part of UPCOS-6, Mathematica conducted direct child assessments to determine whether the agreed-upon targets for 2012–2013 were met.

Purpose and Summary of Volume 2. This report is split into two volumes. In Volume 1, we answered the study’s primary research questions: (1) what was the progress of LAUP children during the 2012–2013 program year, and (2) did LAUP meet the child progress performance targets included in its contract with First 5 LA? We also described LAUP families based on parent responses to a brief questionnaire. The primary purpose of Volume 2 is to provide information that can help extend understanding of the purpose of the study and the results presented in Volume 1—both background information about the study and results of exploratory and technical analyses based on study data. In Chapter II, we include the study background with detailed information on how the measures were selected, target-setting procedures, and the final 2012-2013 targets. In Chapter III, we describe the study methods, including the sample, measures, procedures, and scoring. In Chapter IV, we include fall-spring progress for subgroups of children defined by parent reports of home language use. In Chapter V, we present outcomes among children who switched from a Spanish assessment in the fall to an English assessment in the spring relative to the outcomes of children who remained in the same language of assessment in fall and spring. In Chapter VI, we present sensitivity analyses conducted for the purpose of exploring whether the targets are valid and meaningful. Finally, in chapter VII, we present exploratory multivariate analyses of child, family, and program factors associated with child progress.

In general, findings based on exploratory and technical analyses presented Chapters IV and VI of this volume underscore conclusions reached in Volume I: overall, progress in LAUP during 2012–2013 points to some clear successes but they also indicate that there is room for growth. Regardless of the subgroup and measure, children typically made absolute progress in language, literacy, fine motor, and mathematics skills during their time in LAUP. Among measures with national norms, standard scores show that children also grew as much as their national peers (in other words, their fall and spring standard scores were the same); in a few instances, they progressed beyond what would have been expected by the spring given their fall scores. Despite this progress, many children began and ended the year scoring well below national peers in language skills

(vocabulary concepts) and mathematics skills—especially those children who were assessed in Spanish.

At the same time, findings highlight that children who entered performing at high levels did not necessarily broaden or deepen their skills over the year—patterns indicate they made less progress than peers who came in with lower scores and in some instances lost ground relative to national peers. Although we expected that children who entered with the highest skills would show less change, it is a concern that they lost so much ground relative to peers. Particularly in vocabulary, focusing on increasing the skills of those children likely will benefit all children in the classroom, as they all will be exposed to more sophisticated words.

In Chapter V, where we examine spring fine motor/literacy and math scores based on language-specific assessments (English and Spanish) for children who failed the English-language screener in the fall (and were assessed in Spanish), we see evidence that they may perform better in the spring on an English assessment than a Spanish assessment. In other words, the results suggest that scores that capture absolute progress are lower in the spring on the Spanish assessment than the English assessment for the same ability level for those subtests addressing both fine motor/literacy and mathematics. This exploratory finding is important to consider when determining how best to capture progress among children with a home language other than English and to include their outcomes in the determination of whether progress targets are met.¹

Focusing specifically on the child progress targets included in LAUP's contract with First 5 LA, analyses presented in chapter VI generally confirm the validity of the targets. Children whose progress met or exceeded targets progressed relative to a national sample of peers and in most cases had higher scores in the spring than children whose progress did not meet targets. Typically, mean scores in the spring were higher for the group whose progress met targets across almost all subgroups and tests.

Finally, the exploratory analyses addressing the relationships between children's language and literacy progress and characteristics of the child, family, and program presented in Chapter VII. We found few relationships with progress, which may reflect limited power to detect associations due to our small sample size, or alternatively may indicate that factors other than what we were able to measure are important for children's outcomes.

¹ In UPCOS-7 (2013–2104), children from Spanish-speaking homes will complete the language-specific assessments of fine motor/literacy and math in both Spanish and English in both the fall and spring. This will allow us to more fully address the question of which assessments are better able to capture progress in this group.

II. STUDY BACKGROUND

In this chapter, we present an overview of the rationale for the domains of development selected for assessment and the measures identified for this purpose, followed by an overview of the target-setting process and the final targets for 2012–2013.

A. Domains of Child Development and Selected Measures

The first step in designing the studies of child progress in each phase of UPCOS was to identify the domains of development and associated outcome measures on which to focus. The team agreed that, although it is impossible for a brief assessment to encompass all of the elements of school readiness, the assessment tools should collectively tap the important developmental domains, including those identified by the National Education Goals Panel (Kagan et al. 1995), which are similar to those in the California Preschool Learning Foundations (California Department of Education 2008): language and literacy; cognition (mathematics, thinking and reasoning); social-emotional development; approaches to learning; and motor development. With these key domains in mind, we sought assessments with the following characteristics:

- Evidence of reliability and validity
- Minimal burden on children (fewer than 15 minutes per child to complete the assessment)
- Relatively low cost to administer
- History of use with Spanish-speaking children
- Spanish or bilingual version available²
- Sensitivity to change (that is, able to detect change) in development over a six-month period
- Minimal problems with floor and ceiling effects
- Assessment of skills, knowledge, and behaviors important for all children in ways not specific to a particular curriculum
- National norms available for comparisons to national samples

Table II.1 summarizes the child outcome measures agreed upon by First 5 LA and LAUP, and notes the domains of development that each addresses.³ We provide additional detail regarding each of the measures in Chapter II of this volume. The selected measures all were used in prior phases of UPCOS and together address multiple dimensions of school readiness.

² To the best of our knowledge, there are no direct assessments (monolingual or bilingual with English) in the selected domains with evidence of reliability and validity in any non-English, non-Spanish language. Spanish is the predominant non-English language spoken in LAUP and LA County in general.

³ In addition to the measures shown in Table I.1, children completed the English and Spanish Preschool Language Assessment Survey (*pre*LAS) (Duncan and DeAvila 2002) as both a language screener and warm-up exercise. The team determined that these measures should not be used for setting targets.

Table II.1. Child Outcomes Measures and Developmental Domains Addressed in the First 5 LA-LAUP Performance-Based Contract

Measure	Purpose	Brief Description
Expressive One-Word Picture Vocabulary Test (EOWPVT) – English Edition and Spanish Bilingual Edition (EOWPVT-SBE) (Brownell 2000)	Language development: vocabulary	Standardized measure; children required to name pictures; conceptually scored ^a
Rapid Letter Naming (RLN) (Atkins-Burnett et al. 2007)	Literacy development	Child must name upper- and lowercase letters; conceptually scored ^a
Woodcock-Johnson III (WJ-III) Applied Problems and Woodcock-Muñoz Bateria III (WM-III) Problemas Aplicados (Woodcock et al. 2001/2007; Woodcock et al. 2004/2007)	Mathematics development and reasoning	Standardized measure; children must perform simple counting, solve brief story problems involving numbers; conceptually scored ^a
Woodcock-Johnson III (WJ-III) Spelling and Woodcock-Muñoz III (WM-III) Ortografía ^b (Woodcock et al. 2001/2007; Woodcock et al. 2004/2007)	Fine motor skills and literacy development	Standardized measure; children required to copy shapes and letters, and write orally presented letters
Leiter Examiner Rating Scale – Revised (Leiter-R), Attention, Activity Level, and Sociability subscales (Roid and Miller 1997)	Social-emotional development and approaches to learning	Standardized measure; assessor completes ratings based on observations made during administration ^c

^aIn a measure that is conceptually scored, respondents receive credit for a correct response regardless of the language of the response.

^bMost classrooms taught the English names for letters, so we deviated slightly from standard administration on the Ortografía, allowing assessors to provide the names of letters in both Spanish and English when asking children to write orally presented letters.

^cThese ratings are based on the 25- to 30-minute observation of the child during the UPCOS assessment, rather than the full Leiter measure or classroom performance.

B. Target-Setting Procedures

The child progress targets for the 2012–2013 program year were revisited in a collaborative process between First 5 LA and LAUP, facilitated by Mathematica. First, the team considered whether the existing targets should change due to any new LAUP initiatives or initiatives either in Los Angeles County or California more broadly; would any existing targets need updating due to the potential influence of these initiatives, or would any targets be needed in new domains? The team agreed that no ongoing initiatives would necessitate a change in existing targets but did discuss whether it would be desirable to add a target in the area of nutrition and physical activity. First 5 LA’s strategic plan includes a goal of children maintaining a healthy weight, and LAUP has a program focused on training providers and parents on nutrition and physical fitness. However, funding for the program for 2012–2013 year was uncertain at the time of target revisiting, and LAUP had not yet determined what outcomes should be addressed for this program. The team agreed that no target would be set in the area of nutrition and physical activity at this time, but we collected data on relevant indicators (for example, child overweight or obesity, physical activity in programs) as part of Phase 6 of UPCOS during the 2012–2013 year without setting a target.

Second, the team considered whether the key decisions made in the process of setting the 2011–2012 targets were still the “right” decisions for the 2012–2013 program year. When initially

setting targets, the team addressed questions regarding the appropriate metric (type of score) to use; whether targets should be based on a specified change (for example, the percentage of children achieving a particular score) or mean change; whether targets should be set for the overall sample or for specific language subgroups; whether any of the selected measures were ultimately inappropriate for setting targets; and the appropriate magnitude for each target. The decisions made and any exceptions to them have been reviewed each year. To inform the discussion, the team examined the distribution of the scores for each of the measures in prior rounds of UPCOS⁴ and newly available UPCOS data to determine if the magnitude of each target for the 2011–2012 year was appropriate for the 2012–2013 year. In addition to examining the distribution of scores for the overall sample for each measure, the team also examined scores by child language group (as determined by parent report) to ensure that the targets were valid for children from diverse language backgrounds. In this section, we summarize decisions reached by the team.⁵

1. Appropriate Metric (Type of Score)

Each of the selected assessments offers multiple types of scores: raw scores, standard scores, and item response theory (IRT) or W scores. Raw scores simply summarize how many items a child answered correctly, without accounting for factors such as the difficulty of the items. Standard scores adjust for age and compare local scores to a nationally representative sample of same-age peers (referred to as a norming sample). Two children might have the same raw score, but if one of them is a little older and falls into a different age grouping, the standard score for the slightly older child would be lower than for the younger child. In contrast to a standard score, the IRT or W score is an equal-interval score—that is, a score in which a change at one point on the scale is equal to a change at another—that reflects the relative difficulty of the items. With this type of score, children who correctly respond to more difficult items receive credit for knowing more challenging words or information. For example, two children may answer one additional item on a measure of letters and words, but one names an additional consonant while the other moves from being able to name letters to being able to read a word. A raw score would show the same level of progress for the children, but an IRT score would show a greater gain for the one who read the word than for the one who named one more consonant, reflecting the developmental change represented by the items. The progress noted by a standard score, on the other hand, is dependent on the child’s age group. For instance, a standard score for a child indicating progress on one day may not reflect progress a day or a week later if the child falls into a different age category. That is, the same raw score change may not reflect change in a standard score, depending on the child’s age. The IRT score/W score is not affected by the differences in age group and provides the best estimate of growth over time. For most of the measures, the team agreed to use equal interval scores (that is, IRT or W scores in which changes at one point on the scale are equal to changes at another point on the scale) that adjust for

⁴ For all measures except RLN, targets were based on data from Phase 3 of UPCOS. During UPCOS-3, teachers administered the RLN task, whereas all other measures were administered by independent assessors. During UPCOS-2, independent assessors administered the RLN task. Because the determination of whether targets were met during the 2010–2011 program year was to be based on data collected by independent assessors, the team based the RLN target on UPCOS-2 data. UPCOS-2 included a representative sample of programs and children. Although children who participated in UPCOS-3 came from a representative sample of classrooms and programs, the children who participated in the study were not selected randomly and thus were not representative of all LAUP children. A detailed description of UPCOS-3 and the results of the analyses are presented in Moiduddin et al. (2010) and Xue et al. (2010b). For details regarding UPCOS-2, see Love et al. (2009).

⁵ Additional information on the target-setting process, as well as the rationale for decisions regarding the targets, is available in the memos “Developing Child Progress Targets for the Performance-Based Contract—Revised Memo” (Xue et al. 2011) and “Confirmation of the 2012–2013 Program Year Child Progress Targets” (Moiduddin et al. 2012).

item difficulty but do not adjust for age. Because the IRT scale is equal interval, even though children will be making changes at different points on the scale, we can still compare how far they are moving along the scale.⁶

There are two exceptions to the use of IRT or W scores in the targets. First, the literacy target based on the RLN measure offers the option of a proficiency level target based on the raw score (the number of letters named) or progress on the IRT score. Second, the Leiter-R does not offer an IRT or raw score; rather, the target is based on the percentage of children scoring in the expected range in the spring (that is, the percentage not at risk for developmental difficulties in the areas assessed by this measure). These targets and associated scoring are discussed in more detail below.

2. Specified Change Versus Mean Change

The selected targets are based on the percentage of children who make a specified change from fall to spring, rather than on mean change scores. A mean score can increase when a small number of children make large changes, even if many others do not change. LAUP aims to help *all* children progress.

The preference for targets based on progress has exceptions due to ceiling problems (that is, limitations in the test that prevent the scoring from showing progress beyond a certain point). First, since there are only 26 letters in the English alphabet, the amount of progress that can be measured in the RLN is limited, particularly for children who enter with strong skills. Thus, this target takes into account progress from fall to spring and offers the option of a percentage of children naming the specified number of letters in the spring. Second, the Leiter-R is a screening measure designed to identify children at risk for developmental difficulties, rather than document progress across the full spectrum of social-emotional development, so the scores are truncated (that is, children cannot show progress beyond the expected range or proficiency level). The target based on this measure identifies the percentage of children who score in the expected range in the spring, rather than a level of growth from fall to spring.

3. Overall Versus Language Subgroup Targets

The team examined the distribution of scores for five subgroups of children, based on parents' reports of children's home language use. Groups included children who speak (1) English only; (2) English primarily; (3) Spanish only; (4), Spanish primarily; and (5) another non-Spanish, non-English language (referred to as the "other language" group). The team agreed to set overall targets for all measures for the 2012–2013 program year. For the targets based on progress, change scores from all prior phases of UPCOS did not drastically differ across the language groups. For the target based on the Leiter-R, spring scores from UPCOS-4 and UPCOS-5 did not differ drastically.

4. Exclusion of Measures from Targets

UPCOS-3 data showed that about 70 percent of children who completed the WM-III Ortografía (that is, they completed the Spelling assessment in Spanish) did not show any growth from fall to spring. Statistical tests indicated that the fall standard score (91) and spring standard

⁶ Although the age range of UPCOS children is limited (46 to 60 months in the fall), the same raw score can mean different standard scores if a child falls in the next age grouping.

score (89) did not differ⁷. We hypothesized that these children likely were learning letters in English in the program but were assessed in Spanish. Thus, the prompt in the Spanish test (the orally presented name of a letter that a child must write to receive credit) might not have been consistent with the letter names they had learned. For this reason, the group agreed not to set targets based on the WM-III Ortografía for the original target setting.

Since that original decision, we have adapted the way we administer this test. The spelling assessment asks children to copy or write lines and letters (and eventually words). In the most recent phases of UPCOS, however, we departed from the standard administration to allow assessors to give the letter names in both Spanish and English. UPCOS-5 (2011–2012) was the first year in which these procedures were followed in both the fall and spring. Children made statistically significant progress, and a target equivalent in magnitude to that based on the WJ-III Spelling would have been met. The team agreed to gather additional information about WM-III Ortografía during UPCOS-6 and examine whether a target of the same in magnitude as that for the WJ-III Spelling is appropriate for 2013–2014.

5. Appropriate Magnitude for Targets

The magnitude of the targets is based on the distribution of children’s scores during prior phases of UPCOS. Although LAUP is committed to ongoing improvement and constantly devising new ways to support teachers and improve practice, there was no expectation that support received up to this point would result in child progress differing markedly from that observed in previous program years. As a reminder, children are in LAUP for one year, and teachers can change from year to year. Further, the analysis of UPCOS-4 and UPCOS-5 data supported the substantive importance of these targets—when children met the targeted growth, they demonstrated significantly greater progress compared to their peers nationally.⁸ Reviewing all available data allowed the team to conclude that the magnitude of the targets for 2011–2012 would likely be appropriate for 2012–2013.

The group also agreed that targets should be set at two levels of progress or performance (referred to as Level 1 and Level 2). Children who enter programs with strong skills (above average) may not make as much progress as children who enter with weaker skills. However, LAUP and First 5 LA are interested in seeing progress for all children. Thus, for each of the measures, targets were selected at two levels of difficulty. At the first level, the target indicates that 70 percent of children are to make a specified gain. At the second, more difficult level, the target indicates that 45 percent of children are to make a larger gain. The magnitude of the gains is specific to each measure (and the scale in that measure). LAUP expects to meet the targets at both levels of difficulty. See Box II.1 for an overview of why 70 percent was the proportion of children selected for the targets.

⁷ Because there is error in group estimates, we report only statistically significant differences unless the sample size would make it difficult to detect the differences. In those cases of smaller sample size, we sometimes highlight progress or change that is not statistically significant (and describe it as such).

⁸ A detailed description of UPCOS-4 and the results of the analyses are presented in Moiduddin et al. (2011). For UPCOS-5 results, see Xue et al. (2012).

Box II.1. Identifying the Appropriate Percentage of Children to Include in the Targets

For the language, literacy, fine motor, and mathematics targets, why are gains specified for 70 percent rather than a higher proportion of children?

Unlike targets in public schools, most of the current LAUP targets (language, literacy, fine motor, and mathematics) are based on growth rather than level of proficiency. Children enter preschool with a wide range of abilities. First 5 LA and LAUP are interested in seeing progress for all LAUP children. However, brief assessments have limitations for assessing progress, particularly when children score above average.

The WJ-III, WM-III, and EOWPVT assess a sampling of skills from preschool through old age. This wide age span means children will not reach a ceiling (that is, get all items correct), but it also means some will reach questions on the assessments that reflect content beyond their opportunity to learn in a preschool setting, thus limiting the amount of growth (level of change) that can be measured. In a sense, this is a “glass ceiling”—children ultimately will break through it as they learn and grow during kindergarten and beyond.

Children who enter LAUP with below-average skills (relative to what would be expected for preschool-age children) can show a great deal of growth on these brief assessments. However, while those who enter with above-average skills can deepen or extend their skills in important ways, such assessments cannot necessarily capture those changes, as they reflect only a sample of skills for each age and grade level. Furthermore, in any preschool setting, an additional small percentage of children with disabilities may not be able to make strong gains.

When initially developing the targets, the team wanted to set a percentage that could be used across assessments. The targets thus specify that at least 70 percent of children make gains in each of these domains during their LAUP year.

This general approach to setting the magnitude of the targets has two exceptions. First, the social-emotional target based on the Leiter-R has only one level and, in keeping with the low ceiling on screening measures, it specifies a greater proportion (85 percent) of children. Second, for the literacy target based on the RLN, the target deviates from the approach taken for the other measures in the cognitive battery in several ways: (1) Level 1 of the target specifies that 45 percent (rather than 70 percent) of children are to gain 7 points or more; (2) Level 2 of the target is based on raw scores rather than IRT scores; and (3) the target specifies that LAUP meet at least one of the two levels, rather than both, allowing for the possibility of change in the entering skill level of the children in LAUP in subsequent years.

6. Language Switching

As compared to earlier phases of UPCOS, more Spanish-speaking children who had not passed the English-language screener in the fall were able to do so in the spring in UPCOS-5. Thus, these children switched from an assessment administered in Spanish in the fall to English in the spring. The change in language of administration has implications for who is included in the calculation of whether targets are met. Up through the 2012–2013 program year, children who switched from Spanish to English were not reflected in targets based on the WJ-III and WM-III measures, which are language specific (scores are calculated separately for the English and Spanish measures).

To inform our discussion about how this group is best reflected in the targets, we conducted an experiment in UPCOS-6: for the group that screened into Spanish in the fall, we assessed skills and knowledge in both English and Spanish in the spring, regardless of whether they passed the English screener in the spring. Half of the children started with the English assessment and half with the Spanish assessment, to counterbalance any effects of the order of language of administration. We will compare scores by language. We then will examine the comparability of the scores that capture children’s absolute progress (*W* scores), thereby helping us consider how best to interpret scores for these children. Preliminary results from this experiment are presented in Chapter V of this volume and will be used as part of the target-setting process for 2013–2014.

C. Targets Established for the 2012–2013 Program Year

At the end of the process, LAUP and First 5 LA agreed on targets in each of six domains of child development: language, literacy, mathematics, fine motor, social-emotional, and approaches to learning. Table II.2 summarizes the final Level 1 and Level 2 targets and the key decisions described above that First 5 LA and LAUP made in revisiting the targets for the 2012–2013 program year.

D. Evidence of Validity of the Targets

Analyses of UPCOS-4 (2010–2011 program year) data demonstrated the validity of the targets set for the next year. Analyses of child progress relative to targets in UPCOS-5 again supported the substantive importance of these targets.⁹ For both UPCOS-4 and UPCOS-5, we examined average change scores among the group of children whose scores exceeded Level 1 targets and the group whose scores did not.¹⁰ Standard scores (scores relative to a national sample of peers, adjusted for child age) are available for four of the measures: EOWPVT, WJ-III Spelling, WJ-III Applied Problems, and WM-III Problemas Aplicados. For children whose scores reached or exceeded targets, scores increased, on average, relative to a national sample of same-age peers. Among children whose scores fell below targets, scores declined, on average, relative to a national sample of peers. For the RLN task, raw scores indicate that children whose scores met targets learned an average of 10–11 letters, whereas children whose scores did not meet targets learned an average of one letter. The evidence suggests that the targeted level of change is meaningful and reasonable.

⁹ A detailed description of UPCOS-4 and the results of the analyses are presented in Moiduddin et al. (2011). For UPCOS-5 results, see Xue et al. (2012).

¹⁰ We excluded the Leiter-R from this analysis. Because scores are truncated, we cannot assess the full range of progress among children whose scores met targets.

Table II.2. Final Targets for Child Progress in LAUP Programs During the 2012–2013 Program Year, by Domain and Measure

Domain	Measure	Targets	Key Decisions Underlying Targets
Language Development	EOWPVT/ EOWPVT-SBE	70% of children gain 2 points or more on the EOWPVT IRT score AND 45% of children gain 5 points or more on the EOWPVT IRT score	<i>Appropriate Metric (Type of Score)</i> – For most targets, use equal interval scores (that is, scores in which changes at one point on the scale are equal to changes at another point on the scale) that adjust for item difficulty but not for age. These scores are referred to as IRT scores or W scores. With this type of score, children who correctly respond to more difficult items receive credit for knowing more challenging words or information. There are two exceptions to the use of IRT or W scores in the targets. First, for the literacy target based on the RLN measure, the target includes the raw score (the number of letters named) in addition to the IRT score. Second, the Leiter-R does not offer this type of score; thus, base the target on the percentage of children scoring in the expected range (that is, the percentage of children not at risk for developmental difficulties in the areas assessed by this measure). <i>Specified Change Versus Mean Change</i> – Base targets on the percentage of children who make a specified change rather than on mean change scores from fall to spring. A mean score can increase when a small group of children makes large changes, even if many children do not change.
Literacy Development	RLN	70% of children name 14 letters or more in the spring OR 45% children gain 7 points or more on the RLN IRT score	There are two exceptions. First, because the RLN measure has a ceiling (there are only 26 letters in the English alphabet), the amount of progress that can be measured is limited, particularly for children who enter with strong skills. Thus, this target takes into account both progress from fall to spring and the number of letters named in the spring. Second, because Leiter-R scores are truncated (children cannot score above a certain level), base the target on the percentage of children who score in the expected range in the spring, rather than on growth from fall to spring.
Fine Motor and Literacy Development	WJ-III Spelling	70% gain 7 points or more on the WJ-III Spelling W score AND 45% gain 17 points or more on the WJ-III Spelling W score	<i>Overall Versus Language Subgroup Targets</i> – Set overall targets for all measures because the distributions of change scores from prior phases of UPCOS did not drastically differ across the language groups. <i>Appropriate Magnitude for Targets</i> – Base targets on the distribution of children’s scores during a prior phase of UPCOS. Set targets that specify a level of growth (or on spring performance, in the case of the Leiter-R) similar to that documented in a prior phase of UPCOS. For the WJ-III Applied Problems and WM-III Problemas Aplicados subtests, specify the same level of change for both subtests based on the distribution of WJ-III scores despite differences in the distribution of scores for each measure during a prior phase of UPCOS (change was greater based on the WM-III).
Mathematics Development	WJ-III Applied Problems/ WM-III Problemas Aplicados	70% gain 4 points or more on the WJ-III Applied Problems/WM-III Problemas Aplicados W score AND 45% gain 13 points or more on the WJ-III Applied Problems/WM-III Problemas Aplicados W score	Set targets at two levels to reflect that children who enter programs with weaker skills may make more progress than children who enter programs with stronger skills (above average). The magnitude of the gains at each level is specific to each measure (and the scale in that measure). LAUP expects to meet the targets at both levels of difficulty. There are two exceptions to this approach to setting the magnitude of the targets. First, the social-emotional target based on the Leiter-R has only one level and it specifies a greater proportion (85 percent) of children. Second, for the literacy target based on the RLN, LAUP must meet at least one of the levels, rather than both.
Social-Emotional Development and Approaches to Learning	Leiter-R Examiner Rating Scales Attention Activity Level Sociability	85% score in the expected range on the Leiter-R in the spring	

III. STUDY METHODS

In this chapter, we describe the study methods. We begin with a description of the sample for UPCOS-6 child assessments, followed by an overview of the child outcome measures and procedures for the administration of the child assessment battery. In the last section of this chapter, we present how we scored the child outcome measures.

A. Sample

UPCOS-6 includes a stratified random sample of center-based programs and family child care homes (FCCs). The total number of FCCs and centers in the sample is proportional to their overall numbers in LAUP. We randomly selected one classroom from each sampled program. We then selected all children from each classroom; thus, when weighted, the sample of children is representative of all LAUP children.¹¹ The sample frame received from LAUP contained 304 unique programs. After removing one program that was terminated and 12 programs that were ineligible by merit of being included in another study, we were left with 291 programs: 202 centers and 89 FCCs. We selected an initial sample of 66 programs, 53 of which were eligible to participate in the study. Reasons for ineligibility in the selected sample included (1) not passing the LAUP's quality rating and thus not being considered part of regular operations, (2) leaving the network, or (3) transitioning from FCC to center status and thus being considered new providers this year (and working with Starting Points coaches rather than Quality Support coaches).

There were 53 programs in the original sample, and the goal was to recruit 40 programs (proportionally allocated between center-based programs and FCCs—28 and 12, respectively). In the fall, the sample included 40 programs, reflecting a response rate of 76 percent. There were 641 eligible children in the 40 programs; 627 were assessed in the fall, reflecting a response rate of 98 percent. In the spring, 573 children were assessed in the 40 programs—96 percent of the children still eligible to participate.¹² In the fall, 564 parents completed a brief parent questionnaire, a response rate of 88 percent. Among those who completed the questionnaire, 501 had children who were assessed in both the fall and spring.¹³

Table III.1 presents the distribution of the sample assessed in both the fall and spring for language groups defined by parents' reports of children's home language use (the language the parent speaks most often with the child, the language the child speaks most often with parents, and the language the child speaks most often with other children). Among the children in the sample, more than one-third (36 percent) speak only English, about one-third (33 percent) speak primarily English, 8 percent speak only Spanish, 17 percent speak primarily Spanish, and the remaining 6 percent speak another language only or primarily. About 29 percent responded to the Spanish or bilingual version of the assessments based on their language proficiency as measured by the English *pre*LAS.

¹¹ The analysis weights at the program level account for the probability of selection, eligibility, and participation, as well as a ratio adjustment to poststratify to frame totals. The group (classroom) level weight accounts for the random selection of one group per program. The child weights account for eligibility, consent, random assignment to version 1 or version 2 of the assessment battery, and assessment completion.

¹² By the spring, 46 additional children became ineligible by merit of having left the sampled program or classroom. The total eligible sample in the spring included 595 children.

¹³ In 2012-2013, LAUP programs served a total of 11,620 children.

Table III.1. Number of Children Assessed in Fall 2012 and Spring 2013, by Language Group (N = 573)

Parent-Reported Language Group	Percentage
English only	36.93
English primarily	32.69
Spanish only	7.77
Spanish primarily	16.56
Other language only or primarily	6.05

Note: Analyses are weighted to represent children attending LAUP in the 2012–2013 program year.

B. Measures

Preschool Language Assessment Survey 2000 (*preLAS* 2000). Simon Says and Art Show are two subtests of the *preLAS* 2000 (Duncan and DeAvila 2002) available in both English and Spanish. Simon Says assesses a child’s listening comprehension of basic instructions. Art Show is a picture vocabulary test that measures a child’s expressive vocabulary. We used Simon Says and Art Show in combination with parent reports of the child’s primary home language to determine whether children should receive the WJ-III (English) or WM-III (Spanish) versions of the Applied Problems and Spelling assessments. Children who had fewer than six items incorrect on the English *preLAS* received the assessment in English. The Spanish version of the Art Show subtest (Exposición de Arte) was used as a warm-up activity for Spanish-speaking children. As a reminder, there is no target based on the *preLAS*. It is included in the assessment battery only as a warm-up and to route children to the appropriate language of assessment.

Expressive One-Word Picture Vocabulary Test, English edition (EOWPVT) and Spanish-Bilingual Edition (EOWPVT-SBE) (Brownell 2000). The EOWPVT English and SBE editions are measures of expressive vocabulary. (In the remainder of this report, we refer to the combination of the assessments as the EOWPVT.) For each version, children were asked to name pictures (the same set of pictures for both versions). The SBE version is conceptually scored; that is, children are given credit for correct responses in either Spanish or English, whereas the English version accepts responses only in English. For the English version, standard scores were calculated relative to a nationally representative sample of same-age peers who spoke English. For the SBE version, standard scores were calculated relative to a sample of same-age peers who were nationally representative of children who are bilingual in Spanish and English.

Notably, the nationally representative samples used for developing the standard scores (norming samples) for the two versions of the EOWPVT are different in ways beyond language. The SBE standard scores compare children in LAUP to a nationally representative sample of same-age Spanish-speaking children in the United States who come from similar cultural and linguistic backgrounds. Children from California are slightly overrepresented in the standardization sample for the bilingual norms, and more than 50 percent of children in the sample come from families with mothers whose educational attainment is less than a high school diploma. By comparison, the English norms are based on a standardization sample representing a greater range of maternal education (approximately 10 percent of the children come from families with mothers having less than a high school diploma) and cultural backgrounds. Thus, the linguistically diverse children in LAUP are more similar to the standardization sample for SBE norms than for English norms with regard to cultural and linguistic background and maternal education. Because the SBE norms are based on a sample that, on average, has more limited maternal educational attainment, which

strongly influences the child's exposure to and attainment of words, scores calculated relative to these bilingual norms will be higher than scores calculated relative to English norms.

Rapid Letter Naming (RLN) Task (Atkins-Burnett et al. 2007). For UPCOS-2, we developed a criterion-referenced measure of the number of letters that a child could name quickly and easily. All 26 letters of the English alphabet were assessed in either upper- or lowercase, or both. There were two different forms, each with 30 items, so that there is some overlap of items across forms. Except for the overlapping items, if an item is uppercase on one form, it is lowercase on the other form. We used one form in the fall and the other in the spring. The child receives credit for correctly naming the letter in either English or Spanish. Although relatively parallel in difficulty, the fall form has a slightly higher percentage of uppercase letters, which typically are easier for young children to identify.

Woodcock-Johnson III Battery (WJ-III) and Woodcock-Muñoz Batería III (WM-III) (Woodcock et al. 2001/2007; Woodcock et al. 2004/2007). The WJ/WM assessments are used widely in early childhood studies, including Head Start's Family and Child Experiences Survey (FACES) (West et al. 2011); the Tulsa preschool study (Gormley et al. 2005); the Study of State-Wide Early Education Programs (SWEEP) (Early et al. 2005); Evaluation of the North Carolina More at Four Pre-Kindergarten Program (Peisner-Feinberg and Schaff 2008); the National Institute of Child Health and Human Development (NICHD) Study of Early Child Care and Youth Development (SECCYD) (NICHD Early Child Care Research Network 2005); and the Preschool Curriculum Evaluation Research initiative (PCER) (Early et al. 2007).

The Applied Problems subtest of the WJ-III and the corresponding Problemas Aplicados subtest of the WM-III assess a child's ability to analyze and solve practical problems in mathematics. To solve the problems, the child must perform simple counting and the equivalents of addition and subtraction (understanding different combinations of numbers of objects less than 10, such as, "If I put three flowers in the vase and you put in one more, how many flowers are in the vase?") Although children received instructions in the language identified during the screening process, responses were accepted in either English or Spanish on the WM-III version of the subtest. However, it is important to note that the WM-III version includes more complex language (for example, if-then statements) than the WJ-III version; thus, the WM-III is an indicator of language comprehension at a higher level.

The Spelling subtest from the WJ-III and the corresponding Ortografía subtest from WM-III provide a measure of children's early writing skills. The first six items in the WM-III subtest and the first eight in the WJ-III subtest tap fine motor coordination and pre-writing skills, such as drawing a line and copying letters. The remaining items measure a child's ability to provide written responses when asked to write specific uppercase or lowercase letters of the alphabet or words. It is important to note that we veered slightly from standard administration of the WM-III Ortografía in both fall and spring in UPCOS-6. In standard administration of the WM-III Ortografía, the names of the letters are presented only in Spanish. Because many of the children in LAUP learned the names of the letters only in English, we determined that naming the letters only in Spanish was not a fair representation of their ability to write a letter from memory. For this reason, when children reached the items asking them to write a letter, we presented the names of letters in both Spanish and English.

Spanish-speaking children in LAUP may be disadvantaged in ways that children in the WM-III norming sample are not. For example, the norming process for the WM-III included children from Mexico whose socioeconomic status (SES) is more diverse; higher-SES children are likely to have

stronger skills. In addition, children from Mexico receive all of their primary instruction in Spanish. Thus, Spanish-speaking children in LAUP may have lower mean scores than children included in the WM-III norming process.¹⁴

Leiter Examiner Rating Scales–Revised (Leiter-R) (Roid and Miller 1997). The Leiter-Examiner Ratings assess overall social-emotional development and approaches to learning. Assessors complete the Leiter Examiner Ratings on Attention, Activity Level, and Sociability for all children, regardless of language, as long as the children complete the assessment battery (that is, they complete them for those children not routed out of the assessment entirely because they speak neither English nor Spanish). The ratings are based on the unique assessment situation. Once the assessment is complete, the assessors provide ratings on a number of items that together tap attention, activity level, and sociability. Because this scale is designed as a screening measure, the scores are truncated—that is, above-average skill levels are not measured, as the purpose is to identify children who are having difficulty in these areas and thus at-risk for developmental problems in them. Previous phases of UPCOS identified social-emotional development as an area of strength for most children in LAUP.

It is important to note that the ratings were designed for use with the Leiter-R assessment, a lengthier nonverbal assessment of cognitive abilities that includes materials for children to handle and use in responding to items, thus presenting additional distraction from the task for some children. For example, some children will choose to play with materials rather than attend to the task at hand, which could have implications for the assessor ratings. The UPCOS assessments require only verbal or pointing responses and have a briefer administration time, so they may produce higher scores than would be obtained under different circumstances. However, this year we weighed and measured children, offering an additional opportunity to observe children’s ability to attend and follow directions in more challenging interactions.¹⁵ Despite the lengthier assessment, Leiter-R results are similar to prior years. Alternative methods for assessing social-emotional development would add burden for teachers (or parents). The Leiter-R provides an assessment of how children interact with adults when responding to potentially challenging tasks without adding burden for LAUP teachers.

Body Mass Index (BMI). In addition to assessing child progress, UPCOS-6 includes a new component: providing information about efforts and policies in LAUP programs related to physical activity and nutrition. To that end, we have integrated a measure of children’s physical health—body mass index (BMI)—into the assessment battery. BMI, which is the ratio between height and weight, reflects children’s general health status and physical well-being. The Institute of Medicine of the National Academies considers children to be overweight when their BMI is between the 85th and 94th percentile for their age and gender, and obese when it is at or above the 95th percentile (Centers for Disease Control and Prevention [CDC] 2011). Obesity is a risk factor for many different health concerns, such as type 2 diabetes, stroke, and heart disease.

¹⁴ As shown in the presentation of results, children assessed with the WJ-III in UPCOS have higher mean scores than those assessed with the WM-III. This pattern is evident in national studies as well. The entering mean scores for four-year-olds in FACES 2009 ranged from 81 to 91 for Spanish-speaking children on the subtests of the WM-III, while the mean scores for English-speaking children on the subtests of the WJ-III ranged from 90 to 95 (Hulsey et al. 2011).

¹⁵ Weighing and measuring involved asking children to move to a different location, remove and later put on shoes, and step on an unfamiliar scale.

C. Documenting Family Backgrounds

To provide context for what we learned about child progress, in the fall we implemented a brief, self-administered questionnaire for parents. The 20-item questionnaire included questions on household routines (reading/looking at books with the child) and parents' demographic and background characteristics (for example, race-ethnicity, education, household income, and family structure).

D. Procedures

Child Assessments. To minimize the burden on children, we designed two versions of the child assessment battery. As a result, no children were assessed in all domains of development. Each child within a program (classroom) was randomly assigned to either Version 1 or Version 2. Thus, we have estimates of progress for each domain for every classroom. As in prior years, the combination of the language screener and measures addressing development in the cognitive and fine motor domains took approximately 15 minutes to complete for both versions. The inclusion of height and weight measures this year added another 10 minutes to both versions (3 to 4 minutes for the actual measurements and 5 to 6 minutes for helping children take their shoes off, move to the scale, and put their shoes back on).

Version 1:

1. English *pre*LAS (Simon Says)
2. English/Spanish *pre*LAS (Art Show)
3. WJ-III Applied Problems and WM-III Problemas Aplicados
4. WJ-III Spelling and WM-III Ortografía
5. Height and weight

Version 2:

1. English *pre*LAS (Simon Says)
2. English/Spanish *pre*LAS (Art Show)
3. EOWPVT/EOWPVT-SBE
4. RLN
5. Height and weight
6. Leiter-R

Note that all children received the English *pre*LAS as a warm-up for the assessment and routing to the language of assessment. Children whose parents reported they speak only English at home completed the assessment in English no matter what score they received on the screener. Children from homes where a language other than English or Spanish is spoken were routed out of the assessment entirely if they did not pass the English screener.¹⁶ All children who were from Spanish-

¹⁶ Those who were routed out of the assessment received only the RLN Task.

speaking homes also received the *pre*LAS Spanish Exposición de Arte, both as a warm-up and to communicate to children that we value both languages. The remainder of the measures were split into two versions and distributed across two groups of children within each classroom. Because of this split, no single child was assessed in all domains of development.

In UPCOS-5, more than half of children (52 percent) who were assessed in Spanish in the fall switched to English in the spring, a larger percentage than in prior cohorts. We typically have not presented progress based on the WJ-III and WM-III measures for this group. Although the WJ-III and WM-III tests are equated according to the publisher's technical manual, the Spanish and English versions have normative samples that differ beyond the language, so we were not confident about the validity of comparing scores on the Spanish and English versions to examine their progress. Thus, as part of UPCOS-6, we conducted a small experiment: for the children who screened into Spanish in the fall and were randomly assigned to receive Version 1 of the assessment, we assessed them in both English and Spanish in the spring (i.e., they completed both the WJ-III and WM-III measures). Half of the students were randomly assigned to start with the English assessment and the other half with the Spanish assessment, to counterbalance any effects of the order of language of administration. In addition to the time noted above, this experiment required about 10 additional minutes in the spring assessment for children who were assigned to complete Version 1 of the assessment and screened into Spanish in the fall.

In UPCOS-6, 43 percent of the children who completed the WM-III measures in the fall passed the English language screener in the spring. Results of this experiment will inform the target-setting process for 2014–2015 and help determine whether it is appropriate to present on progress based on the difference between fall WM-III scores and spring WJ-III scores for the group of children who switch language of assessment.

Brief Parent Questionnaire. The questionnaire was distributed to parents with the consent form during the consent process. We distributed English and Spanish versions of the questionnaire. If requested by parents, a study representative who was on site at the program assisted parents in completing the questionnaire. Parents could choose to return the questionnaire at the time they returned the consent form or complete and mail it in at a later time (postage-paid envelope provided).

E. Scoring

We followed the publisher guidelines for scoring all standardized measures. Depending on the assessment, multiple types of scores are available: raw scores, standard scores, or IRT/W scores.¹⁷ Raw scores are provided for the English and Spanish *pre*LAS and RLN.¹⁸ For the English and Spanish *pre*LAS, each subscale has a maximum possible score of 10 points. For the Spanish *pre*LAS, we used only one subscale. Because we used two subscales for the English *pre*LAS, that assessment has a maximum total score of 20. The RLN has a possible raw score of 30.

We provide the standard scores for the English and SBE versions of the EOWPVT, and all of the WJ-III and WM-III subtests. For all of these assessments, the mean for the norming sample is

¹⁷ A W score is a type of IRT score. Scores are referred to as W scores, IRT scores, or growth scores, depending on the particular assessment.

¹⁸ The RLN was conceptually scored; that is, the child received credit for correct answers in English or Spanish and the raw score represents the number of letters named in either language.

100 points (with a standard deviation of 15); standard scores for the LAUP sample are described relative to this mean.

We provide *W* scores for the Woodcock-Johnson and Woodcock-Muñoz tests and estimated IRT scores for the EOWPVT and RLN, using the item difficulties from the UPCOS-3 and UPCOS-2 samples, respectively.

For the Leiter-R, we used a table provided in the manual to convert the raw scores to standardized scale scores, which are truncated at 10 points.¹⁹ We then used the scale scores to calculate the proportion of children scoring in the expected range (also referred to as the acceptable range), the clinical range, and the possible clinical range. Children scoring in the expected range are unlikely to be experiencing difficulties with social-emotional development and approaches to learning.

As a reminder, we diverged from standard administration procedures for both the WM-III Ortografía (by giving children the letter names in both Spanish and English, rather than just English) and the Leiter-R (children do not complete the Leiter-R assessment tasks, and ratings are instead based on behavior during the UPCOS assessment), so the norms may not be valid national comparisons for this sample, but do offer a point of comparison to other samples, including earlier rounds of UPCOS, and other studies that diverge from standard administration procedures, such as FACES.

¹⁹ The scale scores are truncated at 10 because the rating scales are designed to determine whether behaviors fall within a problematic or unexpected range.

IV. FALL-SPRING PROGRESS BY LANGUAGE GROUP

In chapter II of the first volume of this report, we present fall-spring progress during 2012–2013 for the overall sample. In this chapter, we examine the progress of children based on parent reports of their language use. While these analyses are exploratory—the sample was selected to represent LAUP as a whole, not subgroups—they provide information to First 5 LA and LAUP about the skills and well-being within LAUP’s linguistically diverse population that may be important for further work, particularly if noted consistently across time. For this analysis, we sorted children into five language groups: (1) English only, (2) English primarily, (3) Spanish only, (4) Spanish primarily, and (5) other language only or primarily. We consider the latter three groups to be Dual-language learners (DLLs). We used parent responses to a series of questions about the child’s home language use²⁰ to sort the children. Children completed the assessment in English or Spanish, depending on their performance on the language screener (English *pre*LAS 2000).

A. Approach to Analysis

We examined the mean of spring scores in our sample relative to the fall baseline and tested the statistical significance of the difference between fall and spring scores. We tested whether progress between the fall and spring was significant, using *t*-tests. In addition, we conducted analysis of variance (ANOVA, *F*-tests) to determine whether progress differed across subgroups. We report *p* levels at the .05, .01, and .001 levels, and consider $p < .05$ to indicate statistical significance.²¹ For any test resulting in $p > .05$, we consider this as evidence of no change.

Note that the present sample was designed for the purpose of understanding patterns of progress in LAUP as a whole, not for particular subgroups. Thus, patterns of progress in subgroups were estimated with less precision than for the sample as a whole; estimates of fall and spring performance and change across the program year in the subgroups were likely to have large standard errors relative to those for the whole sample and, as a consequence, larger confidence intervals with lower likelihood of detecting significant change.

We examined fall-spring differences for each type of score available (raw, standard, and/or IRT or W score). Because raw and IRT/W scores address children’s absolute progress along a continuum of skill, whereas standard scores are adjusted for the developmental progress children are expected to make based on performance of their same-age peers nationally, it is possible to identify significant change for raw and IRT/W scores but not for standard scores for the same measure.

As a reminder, for the EOWPVT, EOWPVT-SBE, RLN, and Leiter-R, the content of the measures is the same regardless of whether the child completed the assessment in Spanish or English; thus, scores for the total sample reflect performance of all children, regardless of the

²⁰ In the fall, 26 percent of children in the English-primarily group were assessed in Spanish; 4 percent of children in the Spanish-only group and 24 percent of children in the Spanish-primarily group were assessed in English. By the spring, 12 percent of children in the English-primarily group were assessed in Spanish; 27 percent of children in the Spanish-only group and 48 percent of children in the Spanish-primarily group were assessed in English. For children who were assessed in both the fall and the spring, among those assessed in English in the fall, 1 percent switched to Spanish in the spring. Among children who were assessed in Spanish in the fall, 41 percent passed the English screener in the spring.

²¹ *p* is the probability of making an error in the inference about the change. If $p < .05$, the likelihood of finding a difference by chance is less than 5 percent.

language of assessment. For the WJ-III and WM-III subtests, the English and Spanish versions of the measures have different items; thus, WJ-III scores reflect performance of children who followed the English path and WM-III scores reflect performance of children who followed the Spanish path. For the WJ-III and WM-III measures, scores of children who switched from the Spanish assessment in the fall to the English assessment in the spring (43 percent of children who completed the Spanish assessment in the fall) are not reflected in the means presented in this chapter.

All analyses discussed in this chapter were at the child level and weighted to represent LAUP children as a whole. Analyses were conducted using statistical survey procedures that address the clustering of children within programs and classrooms. Specifically, the survey procedures account for the design of the sample (multiple children from the same classroom in each program) in the data analysis to ensure that standard errors were estimated correctly and thus that any statistical tests based on those standard errors are accurate.

Note that for the Spanish-only group, sample sizes for the assessments in English fell below 10 children, and for the group of children speaking another language only or primarily, sample sizes were below 10 on all measures except the *preLAS* and BMI. To avoid less stable or precise estimation, these children's scores were deleted from the tables and are not discussed in the text.²²

B. Progress by Language Group

Tables IV.1 through IV.7 present fall and spring scores and the change in scores for each language, literacy, and math assessment by language group; Tables IV.8 presents information on children's social-emotional development, and information on children's BMI is shown in Table IV.9.

1. Language Screener and Warm-Up: English and Spanish *preLAS*

Not surprisingly, children in the English-only and English-primarily groups began and ended the program year with higher English language skills than did children in the Spanish-only and Spanish-primarily groups, as measured by the total English *preLAS* and the Simon Says and Art Show subsets (Table IV.1). However, all language groups showed statistically significant increases in skills between the fall and the spring. We see that the magnitude of progress in both receptive (Simon Says) and expressive (Art Show) English skills differed across language groups. Children in the Spanish-only and Spanish-primarily groups made the greatest gains (Simon Says mean change = 2.9 and 2.3 and Art Show mean change = 2.2 for both the Spanish-only and Spanish-primarily groups), and children in the English-only group made the smallest gains (Simon Says mean change = 0.9, Art Show mean change = 0.5). Focusing on Spanish-language expressive skills (Exposición de Arte), neither the Spanish-only nor the Spanish-primarily groups made significant progress between the fall and spring assessments, suggesting that the majority of the instruction occurs in English.

²² Among the 25 children who speak a language other than English or Spanish at home who were in the sample in both fall and the spring, 26 percent ($n = 7$) passed the language screener at both time points.

Table IV.1. Means and Standard Errors of preLAS: Fall 2012 and Spring 2013, by Language Group

Outcome	N	Fall 2012		Spring 2013		Fall-Spring Change	
		Mean	Standard Error	Mean	Standard Error	Mean	Standard Error
English Only							
<i>preLAS</i> English							
Total Language Screener Score	217	17.20	0.28	18.51	0.18	1.32***	0.2
Simon Says	217	8.07	0.16	8.94	0.12	0.87***	0.13
Art Show	217	9.12	0.13	9.57	0.08	0.45***	0.1
English Primarily							
<i>preLAS</i> English							
Total Language Screener Score	185	15.19	0.43	17.24	0.26	2.06***	0.39
Simon Says	185	6.76	0.33	8.14	0.19	1.38***	0.31
Art Show	185	8.42	0.14	9.1	0.12	0.68***	0.14
Spanish Only							
<i>preLAS</i> English							
Total Language Screener Score	47	5.09	0.61	10.2	0.72	5.11***	0.66
Simon Says	47	2.31	0.4	5.21	0.49	2.90***	0.5
Art Show	47	2.78	0.34	4.99	0.32	2.21***	0.4
<i>preLAS</i> Spanish							
Exposición de Arte	47	5.90	0.4	6.09	0.45	0.19	0.24
Spanish Primarily							
<i>preLAS</i> English							
Total Language Screener Score	99	9.43	0.67	13.89	0.61	4.46***	0.48
Simon Says	99	4.48	0.4	6.79	0.38	2.31***	0.26
Art Show	99	4.95	0.34	7.1	0.28	2.15***	0.29
<i>preLAS</i> Spanish							
Exposición de Arte	99	4.60	0.34	4.78	0.26	0.18	0.38
Other Language Only or Primarily							
<i>preLAS</i> English							
Total Language Screener Score	25	7.84	0.76	12.58	1.58	4.73**	1.06
Simon Says	25	3.09	0.33	5.49	0.79	2.40**	0.67
Art Show	25	4.76	0.48	7.09	0.81	2.33***	0.42

Sources: UPCOS-6 Fall 2012 and Spring 2013 Direct Child Assessments.

Note: All analyses are weighted to represent children attending LAUP in the 2012–2013 program year. Additional analyses show that *preLAS* English scores (the total language screener score and the Simon Says and Art Show subtests) differ across language groups ($p < .001$).

*Significantly different from zero at the .05 level, two-tailed test.

**Significantly different from zero at the .01 level, two-tailed test.

***Significantly different from zero at the .001 level, two-tailed test.

2. Language: EOWPVT

LAUP children made statistically significant progress in expressive language skills in an absolute sense (based on EOWPVT IRT scores) between fall and spring in all language groups (Table IV.2). The magnitude of the progress is similar across language groups, with mean change scores ranging from 3.3 to 5.8.

Table IV.2. Means and Standard Errors of EOWPVT: Fall 2012 and Spring 2013, by Language Group

Outcome	N	Fall 2012		Spring 2013		Fall-Spring Change	
		Mean	Standard Error	Mean	Standard Error	Mean	Standard Error
English Only							
IRT Score	101	51.64	1.34	54.94	1.65	3.30**	1.06
English Edition Standard Score ^{a,b}	99	93.59	2.36	94.57	2.56	0.98	1.68
SBE Standard Score ^{a,b}	101	121.02	2.85	119.58	3.05	-1.45	1.92
English Primarily							
IRT Score	88	46.89	0.69	52.64	1.01	5.75***	0.91
English Edition Standard Score ^{a,b}	63	87.33	1.55	91.49	1.86	4.16*	1.69
SBE Standard Score ^{a,b}	88	110.20	1.4	113.12	1.71	2.93	1.8
Spanish Only							
IRT Score	25	36.05	2.27	40.28	1.09	4.23**	1.42
English Edition Standard Score ^{a,b}				Sample<10			
SBE Standard Score ^{a,b}	25	85.83	4.9	88.52	2.24	2.69	3.24
Spanish Primarily							
IRT Score	44	37.74	1.47	43.19	1.61	5.44***	0.81
English Edition Standard Score ^{a,b}				Sample<10			
SBE Standard Score ^{a,b}	44	89.60	3.48	94.2	4.00	4.6*	2.04

Sources: UPCOS-6 Fall 2012 and Spring 2013 Direct Child Assessments.

Note: All analyses are weighted to represent children attending LAUP in the 2012–2013 program year.

^aNational mean for standard scores is 100, with a standard deviation of 15.

^bThis measure was conceptually scored. For the English edition, the standard score was generated only for children assessed in English. For the Spanish bilingual edition, the standard score was generated for all children in the sample.

*Significantly different from zero at the .05 level, two-tailed test.

**Significantly different from zero at the .01 level, two-tailed test.

***Significantly different from zero at the .001 level, two-tailed test.

The mean English standard scores indicated that children in English-only and English-primarily groups were scoring below the mean of a national sample of same-age English-speaking peers in the fall and spring (scores for children in the Spanish-only and Spanish-primarily groups were excluded from the table because of a sample size of fewer than 10 children). Children in the English-primarily group made statistically significant progress relative to a national sample of English-speaking peers (mean change score = 4.2); progress of children in the English-only group was not statistically significant (mean change score = 1.0), suggesting that more can be done to support English vocabulary development for children in this group.

SBE standard scores, estimated for all children, indicated that children in the English-only and English-primarily groups outperformed their same-age peers in the national bilingual sample in both the fall and spring; in the spring, average scores in these two groups exceeded the national bilingual average by approximately a full standard deviation. Children in the Spanish-only and Spanish-primarily groups scored below the national sample of bilingual peers at both time points. In the fall, children in the Spanish-only group scored about one standard deviation below the national average; in the spring, children in this group remained more than two-thirds of a standard deviation below the national average. Children in the Spanish-primarily group scored two-thirds of a standard deviation below the national average in the fall, and their scores were still one-third of a standard deviation below the national average in the spring. Children in the Spanish-primarily group made statistically significant progress from the fall to the spring, as measured by SBE standard scores, while the progress of children in other groups kept pace with but did not surpass their peers

nationally. As a reminder, the EOWPVT is conceptually scored; answers are accepted in either English or Spanish. Thus, this pattern indicated that, by spring, children in LAUP whose parents indicated they spoke English only or primarily demonstrated greater conceptual knowledge (as measured through expressive vocabulary) and children whose parents indicated they spoke Spanish only or primarily demonstrated less conceptual knowledge, on average, than did a national sample of same-age bilingual peers. However, children in the English-only group demonstrated the least amount of change when compared with the other groups (although the differences in progress were not statistically significant). Note that, even though parents of children in the English-only and English-primarily groups indicated that most of their children’s interactions were in English, nearly three-quarters of these children came from homes where Spanish and/or other languages were spoken.

3. Literacy: RLN

Children made statistically significant progress from fall to spring on the RLN, based on both raw scores (the number of letters) and IRT scores across language groups (Table IV.3). The magnitude of the progress is similar across language groups, ranging from 4.9 to 6.9 for raw scores and 6.3 to 8.5 for IRT scores.

Table IV.3. Means and Standard Errors of RLN: Fall 2012 and Spring 2013, by Language Group

Outcome	N	Fall 2012		Spring 2013		Fall-Spring Change	
		Mean	Standard Error	Mean	Standard Error	Mean	Standard Error
English Only							
Raw Score	101	11.61	.54	18.08	1.09	6.47***	0.94
IRT Score	101	21.89	.80	29.37	1.29	7.49***	1.03
English Primarily							
Raw Score	88	10.02	1.19	16.93	1.08	6.9***	0.81
IRT Score	88	19.95	1.43	28.42	1.16	8.47***	1.01
Spanish Only							
Raw Score	25	4.86	1.44	9.79	2.22	4.93***	1.29
IRT Score	25	12.60	2.29	18.85	2.58	6.25**	1.89
Spanish Primarily							
Raw Score	44	7.43	1.11	14.28	1.46	6.85***	1.48
IRT Score	44	16.57	1.64	24.51	1.65	7.94***	1.89

Sources: UPCOS-6 Fall 2012 and Spring 2013 Direct Child Assessments.

Note: All analyses are weighted to represent children attending LAUP in the 2012–2013 program year.

*Significantly different from zero at the .05 level, two-tailed test.

**Significantly different from zero at the .01 level, two-tailed test.

***Significantly different from zero at the .001 level, two-tailed test.

4. Fine Motor and Literacy: WJ-III Spelling

For fine motor and literacy development as measured by the WJ-III Spelling subtest, children in both English-only and English-primarily groups made statistically significant absolute progress from fall to spring (with an average change of 18.6 and 17.4, respectively, based on W scores; Table IV.4). In contrast, children in the Spanish-primarily group did not make significant progress in WJ-III Spelling in an absolute sense. We excluded scores for children in the Spanish-only group from the table because of sample sizes of fewer than 10 children. The progress of children in the three language groups kept pace with that of a national sample of same-age peers (with average change

ranging from -2.0 to 2.1, based on standard scores). In both fall and spring, children in all three groups scored above the national average on the English Spelling subtest.

Table IV.4. Means and Standard Errors of WJ-III Spelling: Fall 2012 and Spring 2013, by Language Group

Outcome	N	Fall 2012		Spring 2013		Fall-Spring Change	
		Mean	Standard Error	Mean	Standard Error	Mean	Standard Error
English Only							
W Score	116	379.55	2.91	398.1	3.13	18.55***	2.68
Standard Score ^a	116	104.05	1.37	106.13	1.54	2.08	1.44
English Primarily							
W Score	67	384.68	2.46	402.04	2.30	17.36***	2.1
Standard Score ^a	67	105.62	1.36	107.24	1.28	1.62	1.16
Spanish Only							
W Score				Sample<10			
Standard Score ^a				Sample<10			
Spanish Primarily							
W Score	12	391.05	5.41	401.26	5.70	10.21	5.57
Standard Score ^a	12	110.75	2.65	108.72	2.95	-2.03	2.9

Sources: UPCOS-6 Fall 2012 and Spring 2013 Direct Child Assessments.

Note: All analyses are weighted to represent children attending LAUP in the 2012–2013 program year.

^aNational mean for standard scores is 100, with a standard deviation of 15.

*Significantly different from zero at the .05 level, two-tailed test.

**Significantly different from zero at the .01 level, two-tailed test.

***Significantly different from zero at the .001 level, two-tailed test.

5. Fine Motor and Literacy: WM-III Ortografía

Shifting to the Spanish measure of fine motor and literacy—the WM-III Ortografía—children in the English-primarily, Spanish-only, and Spanish-primarily groups all made significant progress from fall to spring in absolute terms (with average change ranging from 16.3 to 29.9, based on W scores; Table IV.5). The differences in gains across groups were not statistically detectable (the small sample sizes limit the power to detect difference). None of the children in the English-only group received this measure. Children in the English-primarily group made significant progress relative to a national sample of same-age Spanish-speaking peers, while neither the Spanish-only nor the Spanish-primarily group made significant progress relative to a national sample of peers.²³ By spring, on average, the performance of children in the Spanish-only or Spanish-primarily groups was below that of their peers (SS = 96.9 and 93.1, respectively), whereas the performance of children in the English-primarily group approached the national average (SS = 98.3). It is important to remember that these scores reflect the progress and status of children who took the assessment in Spanish in both fall and spring.

6. Mathematics: WJ-III Applied Problems

For mathematics development as measured by the WJ-III Applied Problems subtest, both English-only and English-primarily groups made significant progress in absolute terms (10.1 and 16.4, respectively; Table IV.6). In contrast, children in the Spanish-primarily group did not make

²³ The sample size was small, with limited power to detect differences.

Table IV.5. Means and Standard Errors of WM-III Ortografia: Fall 2012 and Spring 2013, by Language Group

Outcome	N	Fall 2012		Spring 2013		Fall-Spring Change	
		Mean	Standard Error	Mean	Standard Error	Mean	Standard Error
English Primarily							
W Score	11	351.97	12.27	381.90	10.91	29.93***	5.38
Standard Score ^a	11	90.31	6.96	98.34	6.12	8.02*	3.10
Spanish Only							
W Score	16	364.10	3.98	380.40	6.33	16.31**	4.96
Standard Score ^a	16	96.53	1.82	96.86	3.34	0.33	2.95
Spanish Primarily							
W Score	26	354.87	4.82	376.13	4.69	21.26***	4.58
Standard Score ^a	26	90.11	2.37	93.07	2.28	2.96	2.49

Sources: UPCOS-6 Fall 2012 and Spring 2013 Direct Child Assessments.

Note: All analyses are weighted to represent children attending LAUP in the 2012–2013 program year.

^aNational mean for standard scores is 100, with a standard deviation of 15.

*Significantly different from zero at the .05 level, two-tailed test.

**Significantly different from zero at the .01 level, two-tailed test.

***Significantly different from zero at the .001 level, two-tailed test.

Table IV.6. Means and Standard Errors of WJ-III Applied Problems: Fall 2012 and Spring 2013, by Language Group

Outcome	N	Fall 2012		Spring 2013		Fall-Spring Change	
		Mean	Standard Error	Mean	Standard Error	Mean	Standard Error
English Only							
W Score	116	397.07	2.67	407.18	2.69	10.12***	1.84
Standard Score ^a	116	101.00	1.45	100.88	1.43	-0.12	0.95
English Primarily							
W Score	67	394.90	3	411.24	1.87	16.35***	1.76
Standard Score ^a	67	98.76	1.57	102.07	1.09	3.30**	0.98
Spanish Only							
W Score				Sample<10			
Standard Score ^a				Sample<10			
Spanish Primarily							
W Score	12	400.22	4.19	408.24	5.46	8.02	3.98
Standard Score ^a	12	102.66	2.26	101.89	3.00	-0.77	2.23

Sources: UPCOS-6 Fall 2012 and Spring 2013 Direct Child Assessments.

Note: All analyses are weighted to represent children attending LAUP in the 2012–2013 program year. Additional analyses show that WJ-III W and standard scores differ across language groups ($p < .05$).

^aNational mean for standard scores is 100, with a standard deviation of 15.

*Significantly different from zero at the .05 level, two-tailed test.

**Significantly different from zero at the .01 level, two-tailed test.

***Significantly different from zero at the .001 level, two-tailed test.

significant progress in mathematics in English in absolute terms. We excluded from the table scores for children in the Spanish-only group because of sample sizes of fewer than 10 children. Children's progress kept pace with that of a national sample of same-age peers in the English-only and the Spanish-primarily groups, and exceeded the national peers in the English-primarily group; the standard scores in all three groups were around the national average of 100 in both fall and spring.

7. Mathematics: WM-III Problemas Aplicados

Shifting to the Spanish-language measure of mathematics (WM-III Problemas Aplicados), children in the English-primarily, Spanish-only, and Spanish-primarily groups made statistically significant progress in absolute terms (with average change based on W scores ranging from 16.0 to 27.4; Table IV.7). None of the children in the English-only group received this measure. Children in the Spanish-primarily group made significant progress from fall to spring relative to same-age Spanish-speaking peers, with nearly half of a standard deviation increase in standard scores (mean change = 6.6), whereas the progress for children in the English-primarily or Spanish-only groups was not statistically significant. By spring, the scores remained about one standard deviation below the national mean for the Spanish-only and the Spanish-primarily groups, and one-and-a-half of a standard deviation below the national mean for the English-primarily group. It is important to remember that these scores reflect the progress and status of children who took the assessment in Spanish in both fall and spring.

Table IV.7. Means and Standard Errors of WM-III Problemas Aplicados: Fall 2012 and Spring 2013, by Language Group

Outcome	N	Fall 2012		Spring 2013		Fall-Spring Change	
		Mean	Standard Error	Mean	Standard Error	Mean	Standard Error
English Primarily							
W Score	11	339.89	10.22	363.02	10.01	23.13*	8.86
Standard Score ^a	11	74.52	4.54	79.98	4.81	5.46	4.03
Spanish Only							
W Score	16	360.82	7.39	376.81	7.8	15.99**	4.81
Standard Score ^a	16	83.85	3.26	85.77	3.5	1.92	2.26
Spanish Primarily							
W Score	26	347.93	6.5	375.34	4.73	27.41***	5.81
Standard Score ^a	26	77.17	2.56	83.82	1.87	6.64*	2.55

Sources: UPCOS-6 Fall 2012 and Spring 2013 Direct Child Assessments.

Note: All analyses are weighted to represent children attending LAUP in the 2012–2013 program year.

^aNational mean for standard scores is 100, with a standard deviation of 15.

*Significantly different from zero at the .05 level, two-tailed test.

**Significantly different from zero at the .01 level, two-tailed test.

***Significantly different from zero at the .001 level, two-tailed test.

8. Social-Emotional and Approaches to Learning: Leiter-R

With limited room to improve, there were no statistically significant differences across the language groups in progress in social-emotional development and approaches to learning (Table IV.8). Almost all children across the language groups scored in the expected range in the fall (97 to 100 percent) and in the spring (95 to 100 percent) on the Leiter-R Attention subscale. In Activity Level, the Spanish-only group experienced the greatest increase (8 points) in the percentage of children scoring in the expected range, although it was not statistically significant.²⁴ By spring, 98 to 100 percent of children scored in the expected range for Activity Level. In Sociability, 89 to 97 percent of children scored in the expected range in the fall across language groups; by spring, 94 to

²⁴ The sample size was small, with limited power to detect differences.

100 percent of children did so. The English-primarily group had the greatest increase (8 points) in the percentage of children scoring in the expected range.

Table IV.8. Percentage of Children with Leiter Examiner Ratings Scaled Score in the Acceptable Range^a on the Leiter-R: Fall 2012 and Spring 2013, by Language Group

Outcome	N	Fall 2012		Spring 2013		Fall-Spring Change	
		Percentage	Standard Error	Percentage	Standard Error	Percentage	Standard Error
English Only							
Attention	101	97.05	2.88	99.10	0.92	2.05	3.06
Activity	101	96.75	2.25	100	0	3.25	2.25
Sociability	101	97.06	1.54	94.48	3.10	-2.58	3.80
English Primarily							
Attention	88	97.02	1.73	100	0	2.98	1.73
Activity	88	98.15	1.32	98.54	1.46	0.39	2.00
Sociability	88	90.97*	3.83	98.97	1.03	8.00	3.76
Spanish Only							
Attention	25	100	0	96.02	3.77	-3.98	3.77
Activity	25	91.79	4.85	100	0	8.21	4.85
Sociability	25	89.24	5.78	96.02	3.77	6.79	7.57
Spanish Primarily							
Attention	44	98.74	1.26	95.06	3.27	-3.68	3.55
Activity	44	95.29	3.38	97.81	2.23	2.52	2.51
Sociability	44	96.32	2.76	100	0	3.68	2.76

Sources: UPCOS-6 Fall 2012 and Spring 2013 Direct Child Assessments.

Note: All analyses are weighted to represent children attending LAUP in the 2012–2013 program year.

^aScores of 7 or greater out of 10.

*Significantly different from zero at the .05 level, two-tailed test.

**Significantly different from zero at the .01 level, two-tailed test.

***Significantly different from zero at the .001 level, two-tailed test.

9. BMI: Levels of Obesity and Overweight Status

Change in the rate of obesity and being overweight appeared to differ substantially across language groups, in part due to differences in sample size and the initial prevalence of obesity in different groups (Table IV.9). Children in the Spanish-only group had the highest rate of obesity in both fall and spring (35 percent in the fall and 30 percent in the spring), followed by children in the Spanish-primarily group (24 percent in both fall and spring). Children in the English-only and other language-only or primarily groups had lower rates of obesity in both fall and spring (about 10 percent at both time points). The Spanish-only group experienced the greatest decrease in the rate of obesity (from 35 percent in the fall to 30 percent in the spring, representing a net change for two children); the Spanish-only group also experienced the greatest increase in the rate of being overweight (from 8 percent in the fall to 21 percent in the spring, an increase of six children in this category). However, these changes were not statistically detectable.

Table IV.9. Percentage of Children Who Were Overweight or Obese: Fall 2012 and Spring 2013, by Language Group

Outcome	N	Fall 2012		Spring 2013		Fall-Spring Change	
		Percentage	Standard Error	Percentage	Standard Error	Percentage	Standard Error
English Only							
Child is overweight ^a	214	18.96	2.22	16.76	2.35	-2.20	2.78
Child is obese ^b	214	9.83	2.46	11.21	3.12	1.37	1.40
English Primarily							
Child is overweight ^a	184	16.42	4.01	15.94	3.05	-0.48	2.49
Child is obese ^b	184	18.02	3.07	18.65	3.42	0.63	2.81
Spanish Only							
Child is overweight ^a	46	8.01	4.35	20.67	5.35	12.65	6.43
Child is obese ^b	46	34.86	6.64	30.06	5.68	-4.80	8.70
Spanish Primarily							
Child is overweight ^a	99	15.48	2.88	19.07	3.12	3.59	2.54
Child is obese ^b	99	23.67	3.12	23.82	3.95	0.15	2.65
Other Language Only or Primarily							
Child is overweight ^a	24	6.28	3.12	6.28	3.12	0.00	0.00
Child is obese ^b	24	9.53	3.49	6.00	4.83	-3.53	5.61

Sources: UPCOS-6 Fall 2012 and Spring 2013 Direct Child Assessments.

Note: All analyses are weighted to represent children attending LAUP in the 2012–2013 program year.

^aWhen child's gender-specific BMI-for-age is between the 85th and 94th percentiles.

^bWhen child's gender-specific BMI-for-age is at or above the 95th percentile.

*Significantly different from zero at the .05 level, two-tailed test.

**Significantly different from zero at the .01 level, two-tailed test.

***Significantly different from zero at the .001 level, two-tailed test.

V. FALL-SPRING PROGRESS FOR CHILDREN WHO SWITCHED LANGUAGE OF ASSESSMENT: PRELIMINARY ANALYSIS

In the first volume of this report, we present child progress in 2012–2013 (Volume 1, Chapter II) and whether the performance targets set by First 5 LA and LAUP were met (Volume 1, Chapter III). Children from Spanish-speaking homes who did not pass the English language screener in the fall but did do so in the spring are not reflected in analyses based on the language-specific measures in the UPCOS battery: the WJ-III (English) and WM-III (Spanish). Scores are calculated separately for the English and Spanish versions of these measures; even though the corresponding subtests in the WJ-III and WM-III address the same set of skills, the items differ.²⁵

The group of children who progress enough in their English-language skills during the course of the program year to switch from a Spanish assessment path to an English assessment path is small in number but significant, both for UPCOS and LAUP. Children’s performance on an English-language screener determines the language path they follow through the assessment.²⁶ In 2012–2013, 29 percent of children did not pass the English-language screener in the fall and were assessed in Spanish. However, only half of the children participating in UPCOS complete the WJ-III and WM-III subtests included in the battery. (In Chapter II of this volume, we explain how measures were distributed to minimize the burden on children.) Among children who completed the version of the assessment that includes the WJ-III and WM-III measures, 31 percent did not pass the screener (and thus completed the WM-III subtests); by the spring, 43 percent of these children passed the English-language screener, shifting them to the WJ-III (English) versions.

To inform discussions about how this group is best reflected in analyses of child progress and targets, we conducted an experiment during UPCOS-6: for the group that screened into Spanish in the fall and was routed to the WM-III version of the battery, we assessed skills and knowledge in both English and Spanish in the spring, regardless of whether they passed the English screener in the spring. Half of the children started with the English assessment and half with the Spanish assessment, to counterbalance any effects of the order of language of administration.

We compare scores by language of administration examining the comparability of the scores that capture children’s absolute progress (W scores). The comparisons will help us better understand and interpret scores for these children. The information from this experiment, particularly the results pertaining to absolute progress, will be used as part of the target-setting process for 2013–2014.

A. Approach to Analysis

In this chapter, we focus on children’s scores on the English-language screener (the Simon Says and Art Show subtests of the English *pre*LAS), the Spanish warm-up (the Exposición de Arte subtest of the Spanish *pre*LAS), and the WJ-III/WM-III subtests: Spelling/Ortografía (literacy and fine motor skills) and Applied Problems/Problemas Aplicados (mathematics). To further understand the

²⁵ Neither version was developed with a bilingual sample. The children in the development and standardization samples for each test were fully fluent in the respective languages. In addition, the Spanish-speaking preschool sample from the United States was small.

²⁶ Children from homes in which a language other than English or Spanish was spoken were routed out of the assessment entirely if they did not pass the English screener.

patterns, we also describe the progress of the children who completed the conceptually scored assessments (EOWPVT and RLN) by their performance on the English language screener.

We present fall and spring means for all measures and examine fall-to-spring progress. For the WJ-III and WM-III, we present W scores only; the standard scores developed for each of these measures were based on different norming samples and thus are not comparable. However, the measure developers aimed to place the W scores on a comparable scale. For comparison purposes, we also present scores for children who did not switch the language of assessment from fall to spring. Note that we did not conduct statistical tests (comparing fall and spring scores or progress across groups) because we are concerned with what patterns indicate about the appropriate approach for including children in analyses of progress and performance relative to targets for LAUP as a whole.

Finally, we used logistic regression to examine what family and demographic differences are associated with increased proficiency in English. For children who screened into Spanish in the fall, we looked at whether socioeconomic risk factors and parent reported use of English at home predict whether the child screens into English in the spring.

All analyses we discuss in this chapter were at the child level and weighted to represent LAUP children as a whole. We conducted analyses using statistical survey procedures that address the clustering of children within programs and classrooms. Specifically, the survey procedures account for the design of the sample (multiple children from the same classroom in each program) in the data analysis to ensure that standard errors were estimated correctly and thus that any statistical tests based on those standard errors are accurate.

It is important to remember that we randomly selected to represent all LAUP children rather than the different subsamples within LAUP, and so these analyses are exploratory. In addition, some subsamples such as the children who switch language of assessment are very small making it difficult to detect significant differences and reliable estimates.

B. Progress Among Children Who Switched Language of Assessment

Table V.1 shows the *pre*LAS scores for children who switched language in the spring and those who did not. Table V.2 and Figure V.1 present children's fall and spring W scores and change scores on the WM-III/WJ-III subtests for children who switched language and those who did not. Table V.3 presents children's fall and spring scores and change scores on the EOWPVT and RLN for children who switched language and those who did not. Table V.4 presents the logistic regression models examining what characteristics are associated with increased proficiency in English.

1. Language Screener and Warm-Up: English and Spanish *pre*LAS

Compared to children who completed the assessment in Spanish in both the fall and spring, children who scored high enough on the screener to switch languages performed similarly on the Spanish warm-up (*pre*LAS Spanish, Exposición de Arte) in the fall and spring; mean scores averaged to 5 on a 10-point scale in both groups at both times. However, children who switched²⁷ outperformed children who did not on the English screener (*pre*LAS total scores) in both the fall

²⁷ For ease of discussion, we will refer to those children who took Spanish in the fall and passed the screener in the spring as children who switched.

and spring. On the 20-point scale, scores in the two groups were 10.4 (for those who switched) versus 5.2 (for those who did not) in the fall and 16.5 versus 9.6 in the spring. Thus, in the fall, children who switched the language of assessment scored higher than those who did not and made slightly more progress between the fall and spring (6.0 versus 4.4 points; not shown in the table).

Table V.1. preLAS Scores for Children Who Switched Language in the Spring and Those Who Did Not

	Children Who Switched			Children Who Did Not Switch		
	n	Fall Mean (SE)	Spring Mean (SE)	n	Fall Mean (SE)	Spring Mean (SE)
preLAS Spanish Exposición de Arte	21	4.89 (0.53)	4.96 (0.72)	42	4.93 (0.51)	5.14 (0.62)
preLAS English Total Score	38	10.44 (0.48)	16.45 (0.26)	53	5.16 (0.59)	9.58 (0.72)

Sources: UPCOS-6 Fall 2012 and Spring 2013 Direct Child Assessments.

Note: All analyses are weighted to represent children attending LAUP in the 2012–2013 program year.

2. WJ-III and WM-III Subtests: W Scores

Table V.2 and Figure V.1 include WJ-III and WM-III W scores for the Spelling/Ortografía and Applied Problems/Problemas Aplicados subtests. The following scores are available:

- For children who screened into Spanish in the fall and spring, fall WM-III scores and spring WJ-III and WM-III scores
- For children who screened into Spanish in the fall and English in the spring, fall WM-III scores and spring WJ-III and WM-III scores
- For children who screened into English in both the fall and spring, fall and spring WJ-III scores

Literacy and fine motor skills (WJ-III Spelling/WM-III Ortografía). Children who screened into the Spanish assessment in both the fall and spring scored lower than children in all other groups at both times; this pattern holds when comparing WM-III scores or WJ-III scores. Children who screened into English at both times had the highest scores at all time points.

For children who screened into Spanish in the fall, a key question is whether gains are larger when spring skills are measured by the WJ-III or WM-III. For children who did switch from Spanish to English, gains are higher when comparing fall WM-III scores to spring WJ-III scores than to spring WM-III scores (24.8 versus 18.5). For children who screened into Spanish at both time points, gains are similar, though slightly higher, when comparing fall WM-III scores to spring WJ-III scores than to spring WM-III scores (22.3 versus 21.4). Note that, when based on the fall and spring WM-III scores, the mean change score for children who switched is lower than the mean change for children who screened into Spanish at both time points. The mean change score for children who screened into English at both points is lower (mean change = 17.8) than the gains for the other two groups of children. Looking across groups, children who switched from the Spanish assessment in the fall to the English in the spring made the greatest gains when calculated based on fall WM-III and spring WJ-III scores. Their gains are lower (2.9 points) than gains for children who screened into Spanish at both points when they are calculated based on the spring WM-III scores.

Table V.2. Means and Standard Errors for WM-III and WJ-III Subtest W Scores, by Whether Children Switched Language of Assessment: Fall 2012 and Spring 2013

	Mean (Standard Error)				
	Fall	Spring WJ-III	Spring WM-III	Fall-Spring Change (with WJ)	Fall-Spring Change (with WM)
Children Who Switched: Screened into Spanish in Fall and English in Spring (n = 38)					
WJ-III Spelling or WM-III Ortografía	366.35 (4.63)	391.18 (3.58)	384.86 (6.54)	24.84 (4.04)	18.51 (4.10)
WJ-III Applied Problems or WM-III Problemas Aplicados	367.08 (5.92)	400.86 (3.44)	388.25 (6.86)	33.78 (3.97)	21.17 (3.35)
Children Who Did Not Switch: Screened into Spanish in the Fall and Spring (n = 53)					
WJ-III Spelling or WM-III Ortografía	357.27 (3.51)	379.72 (3.27)	378.69 (3.86)	22.33 (3.37)	21.42 (2.78)
WJ-III Applied Problems or WM-III Problemas Aplicados	350.47 (4.67)	380.02 (3.49)	373.30 (2.98)	29.60 (6.38)	22.83 (3.83)
Children Who Did Not Switch: Screened into English in the Fall and Spring (n = 198)					
WJ-III Spelling	382.22 (2.16)	400.03 (2.29)	NA	17.81 (1.74)	NA
WJ-III Applied Problems	396.54 (2.20)	408.95 (1.99)	NA	12.41 (1.10)	NA

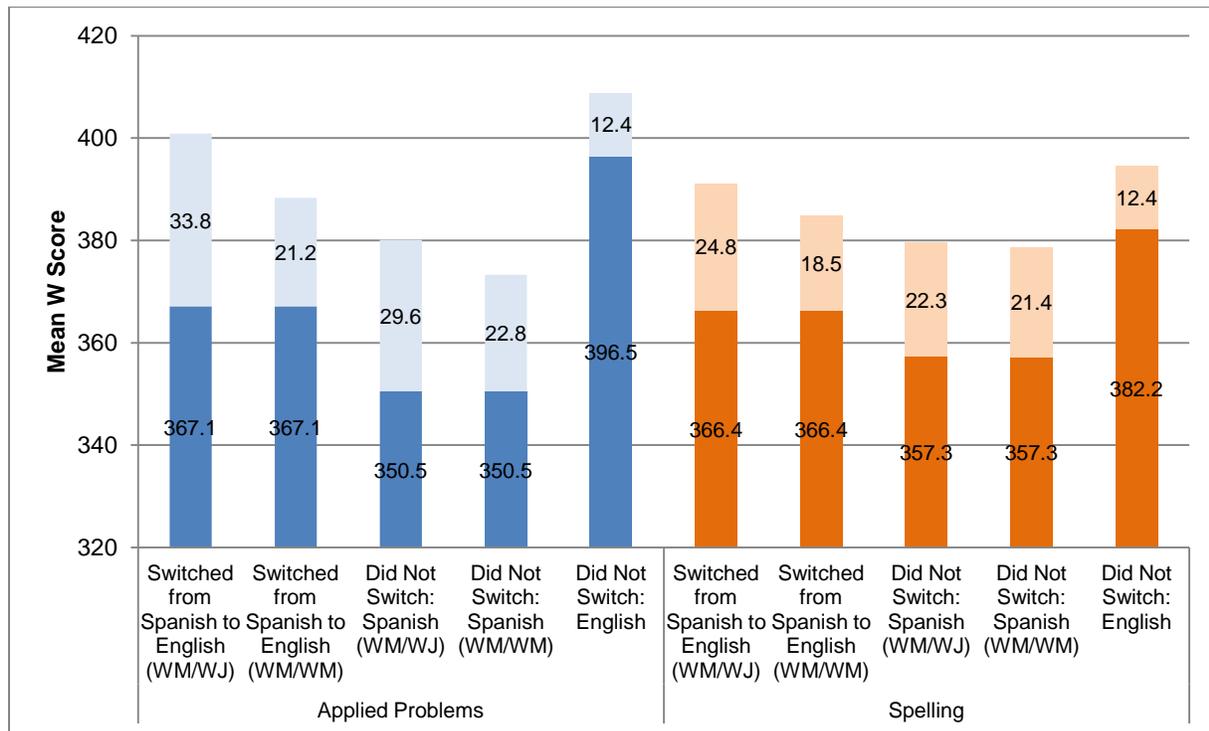
Sources: UPCOS-6 Fall 2012 and Spring 2013 Direct Child Assessments.

Note: All analyses are weighted to represent children attending LAUP in the 2012–2013 program year.

Mathematics (WJ-III Mathematics and WM-III Problemas Aplicados). Mathematics scores follow the same pattern as scores for literacy and fine motor skills. Children who screened into the Spanish assessment at both points scored lowest, children who switched from Spanish to English assessment based on the screener scored somewhat higher, and children who screened into the English assessment at both points scored highest.

From fall to spring, gains are higher when calculated relative to spring WJ-III scores than relative to spring WM-III scores for both groups of children (29.6 versus 22.8, respectively, for those screened into Spanish at both time points, and 33.8 versus 21.2, respectively, for children who switched). The gains for children who screened into English at both points were lower (mean change = 12.4) than the gains for the other two groups of children. Looking across groups, children who switched language of assessment in the spring made the greatest gains when these are calculated based on the fall WM-III and spring WJ-III scores. Their gains are slightly lower (1.6 points) than children who screened into Spanish at both points when they are calculated based on the spring WM-III scores.

Figure V.1. Mean W Scores for the WM-III and WJ-III Subtests, by Whether Children Switched Language of Assessment: Fall 2012 and Change to Spring 2013



Sources: UPCOS-6 Fall 2012 and Spring 2013 Direct Child Assessments.

Notes: All analyses are weighted to represent children attending LAUP in the 2012–2013 program year. Dark colors represent fall scores; light colors represent fall-spring change.

Switched from Spanish to English (WM/WJ) = spring WJ-III score – fall WM-III score for children who passed the screener in the spring but not in the fall; Switched from Spanish to English (WM/MM) = spring WM-III score – fall WM-III score for children who passed the screener in the spring but not in the fall; Did not switch: Spanish (WM/WJ) = spring WJ-III score – fall WM-III score for children who did not pass English screener in fall or spring; Did not switch: Spanish (WM/MM) = spring WM-III score – fall WM-III score for children who did not pass English screener in fall or spring.

3. EOWPVT and RLN

Similar patterns are noted for the conceptually-scored assessments as for the WJ-III and WM-III subtests. Children who screened into English in both fall and spring scored the highest at both time points. Children who switched had the next highest scores across all measures. Children who did not pass the English screener in fall or spring had the lowest scores on all assessments. Children who switched consistently had the largest change scores. With only one exception (RLN raw scores), children who screened into Spanish in both fall and spring had stronger change scores than the children who passed the screener at both time points.

Table V.3. Means and Standard Errors for EOWPVT and RLN, by Whether Children Switched Language of Assessment: Fall 2012 and Spring 2013

	Mean (Standard Error)		
	Fall	Spring	Fall-Spring Change
Children Who Switched: Screened into Spanish in Fall and English in Spring (n = 52-57)			
EOWPVT IRT Score	42.09 (1.49)	48.05 (1.42)	5.96 (0.86)
EOWPVT English Edition Standard Score ^a	--	--	--
EOWPVT-SBE Standard Score ^{a, b}	101.42 (3.57)	106.33 (3.26)	4.91 (1.87)
RLN Raw Score	8.85 (1.46)	15.63 (1.74)	6.78 (0.95)
RLN IRT Score	18.46 (1.88)	26.79 (2.10)	8.33 (1.02)
Children Who Did Not Switch: Screened into Spanish in the Fall and Spring (n = 49-55)			
EOWPVT IRT Score	34.92 (1.27)	39.70 (1.25)	4.78 (0.72)
EOWPVT English Edition Standard Score ^{a, b}	--	--	--
EOWPVT-SBE Standard Score ^{a, b}	84.12 (2.75)	87.79 (2.80)	3.44 (1.79)
RLN Raw Score	5.73 (0.96)	12.03 (1.43)	6.30 (1.26)
RLN IRT Score	13.79 (1.55)	21.98 (1.71)	8.19 (1.70)
Children Who Did Not Switch: Screened into English in the Fall and Spring (n = 154-155)			
EOWPVT IRT Score	51.09 (0.90)	55.50 (1.16)	4.42 (0.81)
EOWPVT English Edition Standard Score ^a	91.40 (1.58)	93.14 (1.81)	1.89 (1.25)
EOWPVT-SBE Standard Score ^{a, b}	118.71 (1.95)	118.94 (2.08)	0.23 (1.38)
RLN Raw Score	12.00 (1.19)	18.76 (0.81)	6.76 (0.80)
RLN IRT Score	22.55 (1.39)	30.10 (0.94)	7.55 (0.86)

Source: UPCOS-6 Fall 2012 and Spring 2013 Direct Child Assessments.

Note: All analyses are weighted to represent children attending LAUP in the 2012–2013 program year.

^aNational mean for standard scores is 100 with a standard deviation of 15.

^bThis measure was conceptually scored. For the English edition, the standard score was generated only for children assessed in English. For the Spanish bilingual edition, the standard score was generated for all children in the sample.

4. Factors Associated with Switching Language of Assessment

We also examined whether home language use and family risk were related to the likelihood children will switch language groups. For the examination of home language use, we sorted Spanish-speaking DLLs into three groups based on what parents reported about language use at home and with peers: (1) English primarily, (2) Spanish only, and (3) Spanish primarily. To address family risk we created an indicator of cumulative socioeconomic risk using the family risk factors addressed in the brief parent questionnaire (described in Chapter II): maternal education less than high school, single-parent household, teen parent household, either parent born outside of the United States, and household size of five or more. Children with such risk factors are more likely to exhibit other risk factors, and research has shown that the accumulation of risks can have negative consequences for children's development and school readiness (Downey et al. 2004; Rathbun and West 2004).

Home language use and family risk factors were both related to the likelihood that children would switch language of assessment from Spanish in the fall to English in the spring (Table V.4). In

Model 1, we examined whether home language and having high family risk (three or more of the risk indicators shown in the table) were associated with the likelihood of switching from a Spanish assessment path in the fall to English in the spring. In Model 2, we considered individual risk factors along with home language. Not surprisingly, children whose parents reported they spoke Spanish only or primarily at home were less likely to switch than children who spoke English primarily. Family socioeconomic risk is also related to switching. Model 1 shows that having high family risk is associated with a lower likelihood of switching. When considering individual risk indicators in Model 2, we see that of the children who screened into Spanish assessments in the fall, those whose mothers had less than high school education or who came from single-parent families were less likely to pass the English screener in the spring than those whose mothers had higher educational levels or who came from two-parent families. After controlling for home language use and other risk factors, children with parents born outside the United States (who used some English at home) and children in larger households were more likely to switch to English in the spring. That household size is positively associated with switching may reflect that 84 percent of the children in this analysis with a household size of five or more were from two-parent families. The finding regarding parents' country of origin should be interpreted with caution; only three of the children included in this analysis did not have a parent born outside of the U.S.

Table V.4. Likelihood That Children Switch from Spanish to English Assessment Path

	Model 1	Model 2
Gender (Boy)	.68 (.09)***	.58 (.10)***
Age	.01 (.01)	-.01 (.01)
Home Language		
English primarily (reference)		
Spanish only	-1.49 (.12)***	-1.71 (.14)***
Spanish primarily	-.67 (.10)***	-.54 (.11)***
High Family Risk	-.53 (.11)***	
Risk Indicators		
Maternal education less than high school		-.65 (.11)***
Single parent		-.82 (.16)***
Teen parent		-.16 (.33)
Parent born outside of the U.S.		1.39 (.25)***
Household size of 5 or greater		.79 (.11)***
n	67	67

Source: UPCOS-6 Fall 2012 and Spring 2013 Direct Child Assessments and Parent Questionnaire.

*Significantly different from zero at the .10 level, two-tailed test.

**Significantly different from zero at the .05 level, two-tailed test.

***Significantly different from zero at the .01 level, two-tailed test.

C. Summary and Implications

Children who switched languages performed similarly on the Spanish warm-up (*pre*LAS Spanish Exposición de Arte) in the fall and spring compared to children who screened into Spanish at both time points. However, on average, children who switched languages outperformed those children who did not on the English screener in both the fall and spring.

Children made greater gains from fall to spring in both literacy and mathematics when looking at the gains based on fall WM-III and spring WJ-III scores, with those who scored high enough on the English screener to switch in the spring making the greatest gains. However, when the gains are calculated based on fall WM-III to spring WM-III scores, the gains are slightly lower for Spanish-speaking children who switched languages than for those who did not pass the screener in the spring. When examining conceptually scored measures, the group that switched (did not pass the screener in the fall, but did in the spring) made stronger gains than children who screened into English at both time points or who screened into Spanish at both time points. Overall, the children who switched appear to have fewer risks than children who screened into Spanish at both time points.

The results suggest that the W scores for the WM-III are lower than for the WJ-III for the same ability level for those subtests addressing both fine motor/literacy and mathematics. The Spelling/Ortografía items are very similar across forms and the directions are very simple. Thus, we would expect that children who switched languages would perform at the same level across forms, yet there is an average of about a 6-point difference. For children who screened into Spanish at both points, we would expect higher WM-III scores than WJ-III scores, yet their spring scores are similar on both assessments (the WJ-III scores are slightly higher than WM-III scores).

The Applied Problems/Problemas Aplicados subtest relies more heavily on comprehension of language to respond to the questions. In light of this reliance on language, it is not surprising that a much greater difference is observed between the results of the spring WJ-III and the spring WM-III—close to a 13-point difference in mean W scores for children who scored high enough on the English screener to switch languages and a 7-point difference for children who did not score high enough to switch. These results raise concerns about the equating applied to this sample and reinforce the importance of language to the measurement of mathematics on this measure, particularly the Spanish assessment.

Overall, the results of the experiment suggest a need for additional analysis at the item level (looking at the differences in item difficulties relative to the type of information in the item). In addition, the strong influence of language on the mathematics assessment has implications for the measurement of children who do not pass the screener and are not receiving Spanish instruction. Their ability to demonstrate their reasoning and knowledge of mathematics may be impeded by limited proficiency in language.

VI. FALL AND SPRING SCORES BY PERFORMANCE RELATIVE TO TARGETS: AN EXAMINATION OF QUARTILES AND LANGUAGE GROUPS

In chapter III of the first volume of this report, we present child progress in LAUP during the 2012–2013 program year relative to the performance targets set in a collaborative process between First 5 LA and LAUP. In this chapter, we present additional detail about children’s performance in each domain by subgroups defined by fall quartile and child language as a sensitivity analysis. The data are presented for the purpose of determining whether the pattern of scores can help us understand the validity of the targets and which children were (or were not) meeting targets.

A. Approach to Analysis

We first calculated the percentage of children whose scores exceeded or fell below Level 1 and Level 2 targets for the overall sample and by fall quartile (Volume 1). For each measure, we also examined (1) the mean and range of change scores, and (2) fall and spring scores by fall quartile for the group of children whose scores exceeded the Level 1 target and the group whose scores did not for five of the six measures: EOWPVT, RLN, WJ-III Spelling, WJ-III Applied Problems, and WM-III Problemas Aplicados (also in Volume 1). We conducted this analysis regardless of whether the overall target was met. We excluded the Leiter-R from this analysis because virtually all children scored at a level required by the targets.

In this chapter, we present the sensitivity analysis among subgroups defined by fall quartile and language groups for the same five measures noted above: EOWPVT, RLN, WJ-III Spelling, WJ-III Applied Problems, and WM-III Problemas Aplicados. Specifically, we calculated fall and spring mean scores by language group and quartile for the group of children whose scores exceeded the Level 1 target and the group whose scores did not. For the quartile analysis, groups were based on fall scores in the full sample of children; thus, whether children fall in the bottom quartile, the middle 50 percent, or the top quartile was based on performance relative to all children in the sample, not just those with scores that did or did not exceed targets.

The analysis was specific to each measure (a child may be in the group whose scores exceeded targets for one measure but in the group whose scores fell below targets for another measure). In some instances, a particular subgroup fell below 10 observations; these results were excluded from the tables because they were likely to produce imprecise estimates of children’s performance. Note that we did not conduct statistical tests comparing fall and spring scores for these groups. The data were presented for the purpose of determining whether the pattern of scores can help us understand the validity of the targets and which children were (or were not) meeting targets.

All analyses discussed in this chapter were at the child level and weighted to represent LAUP children. We conducted all analyses using survey procedures that address the clustering of children within programs and classrooms. Specifically, the survey procedures account for the design of the sample (multiple children from the same classroom in each program) in the data analysis to ensure that standard errors were estimated correctly.

B. Quartile Analysis

Table VI.1 presents mean scores from the fall and spring for children who began the year in the bottom, middle, or top quartile of scorers for each measure in the fall and whose progress exceeded Level 1 targets. Table VI.2 presents the same information for children whose progress did not

exceed targets. Beginning with Table VI.1, looking across the measures included in the table, we noted two patterns. First, as anticipated, it appears that progress of children in the lower quartiles was greater than that of children in the top quartile. Second, spring scores indicate that, on average, children who began the year with scores in the bottom, middle, or top quartiles had scores in the spring that put them in the same relative order; that is, among the children whose scores exceeded the Level 1 target, children who began the year in the top quartile had the highest spring scores, and those who began the year in the bottom quartile had the lowest spring scores.

Table VI.2 presents scores by quartile for the group whose scores did not exceed Level 1 targets and indicates that whether children began the year in the top quartile, middle 50 percent, or the bottom quartile, they scored lower, on average, in the spring than in the fall for the EOWPVT and the WJ-III Spelling and WJ-III Applied Problems subtests. For the RLN Task, it appears that children at either extreme, on average, knew about the same number of letters by the spring as they knew in the fall, whereas children in the middle knew slightly more letters. Note that for one measure—the WM-III Problemas Aplicados—too few children had scores that did not meet the target to be included in the table. Similarly, for the bottom quartile on the English EOWPVT and the bottom quartile on the WJ-III Spelling, the number of children who had scores that did not meet the target was too small to be presented in the table.

Looking across Tables VI.1 and VI.2, we see that, for each measure and quartile, the fall means for children whose progress met targets were lower than or similar to the fall means for children whose performance did not meet targets. Note that this comparison cannot be made for some of the groups because of insufficient data (sample sizes fewer than 10 children in the relevant groups). Regardless of the relative position of fall scores, by the spring, standard scores were higher for the group that met targets than for the groups within each quartile that did not.

C. Language Group Analysis

Table VI.3 presents mean scores from the fall and spring by language group for children whose progress exceeded Level 1 targets. Table VI.4 presents the same information for children whose progress did not exceed targets.

Language: EOWPVT. Among children whose progress exceeded targets (Table VI.3), scores indicate that by the spring, the scores of children in the English-only group were approaching the mean for a national sample of English-speaking peers (mean SS = 97.7) and were below the mean for the English-primarily groups (93.2). Based on the SBE version, children in the English-only and English-primarily groups scored well above a national sample of bilingual peers in the spring; in contrast, children in the Spanish-primarily group scored slightly below the national means (mean SS = 97.4), and children in the Spanish-only group scored nearly one standard deviation below the national mean (mean SS = 85.1).

Table VI.4 shows average fall and spring scores for the children whose scores did not exceed the Level 1 targets by language group. Beginning with the English edition, children in the English-only and English-primarily groups lost ground relative to a national sample of peers between the fall and spring assessments. (Too few children in the other language groups whose progress did not meet targets were scored relative to English norms to present the data here.) Similarly, standard scores for the SBE version of the EOWPVT indicate that, on average, children who did not meet the targets lost ground relative to a national sample of bilingual peers.

Table VI.1. Means and Standard Errors for UPCOS-6 Language, Literacy, and Math Standardized Measures: Fall 2012 and Spring 2013 for Children Whose Change Score Met or Exceeded Change Specified in Level 1 Targets, by Fall Quartile

Outcome	Bottom Quartile				Middle 50 Percent				Top Quartile			
	Fall 2012		Spring 2013		Fall 2012		Spring 2013		Fall 2012		Spring 2013	
	Mean	Standard Error	Mean	Standard Error	Mean	Standard Error	Mean	Standard Error	Mean	Standard Error	Mean	Standard Error
Language: EOWPVT												
IRT Score	31.98	0.88	40.35	0.96	45.53	0.31	53.51	0.58	59.03	1.00	66.46	1.26
English Edition Standard Score ^{a,b}	72.19	1.25	81.11	1.04	87.48	0.62	94.7	1.23	107.74	1.50	114.76	2.15
SBE Standard Score ^{a,b}	76.62	2.08	88.82	2.41	107.81	0.69	115.05	1.14	135.11	1.79	137.26	1.57
Literacy: RLN												
Raw Score	0.82	0.1	10.95	0.94	6.98	0.59	20.16	0.83	22.58	0.73	28.88	0.4
IRT Score	5.54	0.51	22.44	0.95	18.24	0.62	31.16	0.79	32.87	0.91	44.83	1.68
Fine Motor and Literacy: WJ-III Spelling												
W Score	359.05	2.24	389.19	2.46	382.35	0.99	403.61	1.26	407.52	2.59	422.23	2.78
Standard Score ^a	89.15	1.49	99.03	1.58	105.82	0.61	108.46	0.95	120.68	1.31	121.85	1.68
Mathematics: WJ-III Applied Problems												
W Score	372.24	2.56	401.93	2.2	397.64	0.64	413.98	1.25	414.02	1.43	427.17	1.74
Standard Score ^a	87.4	1.17	97.06	1.31	99.28	0.3	103.18	0.72	112.5	1.16	114.39	1.62
Mathematics: WM-III Problemas Aplicados												
W Score	302.99	2.03	359.89	10.45	354.48	3.95	377.58	5.41	383.73	1.37	403.89	3.5
Standard Score ^a	57.62	1.32	77.22	4.34	80.32	1.6	85.98	2.94	93.35	0.73	96.92	1.63
N (range)	11–75				20–94				11–46			

Sources: UPCOS-6 Fall 2012 and Spring 2013 Direct Child Assessments.

Note: All analyses are weighted to represent children attending LAUP in the 2012–2013 program year.

^aNational mean for standard scores is 100, with a standard deviation of 15.

^bThis measure was conceptually scored. For the English edition, the standard score was generated only for children assessed in English. For the Spanish bilingual edition, the standard score was generated for all children in the sample.

Table VI.2. Means and Standard Errors for UPCOS-6 Language, Literacy, and Math Standardized Measures: Fall 2012 and Spring 2013 for Children Whose Change Score Fell Below Change Specified in Level 1 Targets, by Fall Quartile

Outcome	Bottom Quartile				Middle 50 Percent				Top Quartile			
	Fall 2012		Spring 2013		Fall 2012		Spring 2013		Fall 2012		Spring 2013	
	Mean	Standard Error	Mean	Standard Error	Mean	Standard Error	Mean	Standard Error	Mean	Standard Error	Mean	Standard Error
Language: EOWPVT												
IRT Score	35.83	0.73	35.74	0.89	45.74	0.62	43.73	0.96	60.96	0.98	55.47	1.97
English Edition Standard Score ^{a,b}	Sample < 10				90.51	0.84	83.7	1.4	111.19	2.33	96.84	3.97
SBE Standard Score ^{a,b}	86.45	2.16	77.93	2.46	109.79	1.48	98.2	2.58	136.16	1.27	121.02	3.54
Literacy: RLN												
Raw Score	1.03	0.22	1.56	0.44	8.75	1.00	11.44	1.30	25.56	0.54	25.84	0.49
IRT Score	6.1	1.18	6.8	1.52	19.34	0.84	21.5	1.24	37.71	0.97	37.64	0.78
Fine Motor and Literacy: WJ-III Spelling												
W Score	Sample < 10				385.47	2.10	381.05	2.45	415.13	2.91	410.42	2.57
Standard Score ^a	Sample < 10				108.11	1.16	97.85	1.44	122.95	1.82	112.47	2.47
Mathematics: WJ-III Applied Problems												
W Score	376.01	5.37	365.69	6.63	397.57	1.52	394.73	1.51	418.68	2.63	412.25	3.09
Standard Score ^a	88.68	1.74	78.39	3.22	100.3	0.69	93.68	0.66	116.01	2.22	104.33	2.41
Mathematics: WM-III Problemas Aplicados												
W Score	Sample < 10				Sample < 10				Sample < 10			
Standard Score ^a	Sample < 10				Sample < 10				Sample < 10			
N (range)	10–22				15–67				19–56			

Sources: UPCOS-6 Fall 2012 and Spring 2013 Direct Child Assessments.

Note: All analyses are weighted to represent children attending LAUP in the 2012–2013 program year.

^aNational mean for standard scores is 100, with a standard deviation of 15.

^bThis measure was conceptually scored. For the English edition, the standard score was generated only for children assessed in English. For the Spanish bilingual edition, the standard score was generated for all children in the sample.

Table VI.3. Means and Standard Errors for UPCOS-6 Language, Literacy, and Math Standardized Measures: Fall 2012 and Spring 2013 for Children Whose Change Score Met or Exceeded Change Specified in Level 1 Targets, by Language Group

Outcome	English Only				English Primarily				Spanish Only				Spanish Primarily			
	Fall 2012		Spring 2013		Fall 2012		Spring 2013		Fall 2012		Spring 2013		Fall 2012		Spring 2013	
	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE
Language: EOWPVT																
IRT Score	49.02	1.69	57.58	1.57	46.86	0.78	54.56	0.95	30.65	1.55	38.55	1.31	36.64	2	44.6	2.12
English Edition Standard Score ^{a,b}	89.11	2.59	97.69	2.32	86.81	1.75	93.24	2.06	Sample < 10				Sample < 10			
SBE Standard Score ^{a,b}	116.74	3.84	124.49	2.59	110.07	1.58	116.27	1.63	74.25	3.41	85.07	3.07	86.77	4.52	97.35	4.93
Literacy: RLN																
Raw Score	6.57	0.81	18.32	0.87	6	1.29	17.37	1.45	4.81	1.5	14.95	2.52	4.17	0.96	16.55	2.02
IRT Score	16.24	1.03	29.8	0.85	14.54	1.7	29.25	1.66	11.95	2.4	26.28	2.53	12.12	1.81	27.22	1.93
Fine Motor and Literacy: WJ-III Spelling																
W Score	377.72	3.03	400.76	2.71	378.92	3.03	402.19	2.94	Sample < 10				Sample < 10			
Standard Score ^a	103.04	1.44	107.62	1.36	102.79	1.55	107.66	1.43	Sample < 10				Sample < 10			
Mathematics: WJ-III Applied Problems																
W Score	395.22	2.86	412.71	1.86	392.5	3.53	413.78	2.23	Sample < 10				Sample < 10			
Standard Score ^a	99.78	1.39	103.77	1.04	97.54	1.8	103.57	1.38	Sample < 10				Sample < 10			
Mathematics: WM-III Problemas Aplicados																
W Score	NA			Sample <10				356.98	7.83	378.72	8.18	346.82	5.54	382.39	4.09	
Standard Score ^a	NA			Sample < 10				81.83	3.28	86.5	3.60	76.87	2.29	87.15	1.67	
N (range)	53–96				44–71				12–15				21–32			

Sources: UPCOS-6 Fall 2012 and Spring 2013 Direct Child Assessments.

Note: All analyses are weighted to represent children attending LAUP in the 2012–2013 program year.

^aNational mean for standard scores is 100, with a standard deviation of 15.

^bThis measure was conceptually scored. For the English edition, the standard score was generated only for children assessed in English. For the Spanish bilingual edition, the standard score was generated for all children in the sample.

Table VI.4. Means and Standard Errors for UPCOS-6 Language, Literacy, and Math Standardized Measures: Fall 2012 and Spring 2013 for Children Whose Change Score Fell Below Change Specified in Level 1 Targets, by Language Group

Outcome	English Only				English Primarily				Spanish Only				Spanish Primarily			
	Fall 2012		Spring 2013		Fall 2012		Spring 2013		Fall 2012		Spring 2013		Fall 2012		Spring 2013	
	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE
Language: EOWPVT																
IRT Score	55.65	1.47	50.9	2.02	47.02	1.35	44.63	1.63	43.45	1.88	42.64	1.13	40.35	1.2	39.83	1.34
English Edition Standard Score ^{a,b}	100.89	2.91	89.48	3.46	89.76	1.83	83.37	1.95	Sample < 10				Sample < 10			
SBE Standard Score ^{a,b}	127.57	2.56	112.06	4.15	110.74	2.51	99.96	3.7	101.7	4.81	93.26	2.61	96.32	3.35	86.72	3.61
Literacy: RLN																
Raw Score	17.35	1.98	17.82	1.87	14.25	1.78	16.46	1.64	4.9	2.28	5.07	2.31	10.32	1.46	12.27	1.87
IRT Score	28.31	2.47	28.89	2.24	25.62	1.93	27.55	1.82	13.19	2.97	12.07	2.93	20.51	1.87	22.11	2.39
Fine Motor and Literacy: WJ-III Spelling																
W Score	389.09	5.42	384.16	7.37	405.72	7.44	401.5	6.24	Sample < 10				Sample < 10			
Standard Score ^a	109.31	2.65	98.36	3.84	115.97	4.08	105.72	3.66	Sample < 10				Sample < 10			
Mathematics: WJ-III Applied Problems																
W Score	401.28	4.98	394.54	5.22	405.65	3.59	399.86	3.52	Sample < 10				Sample < 10			
Standard Score ^a	103.79	2.93	94.27	2.69	104.28	2.17	95.32	1.74	Sample < 10				Sample < 10			
Mathematics: WM-III Problemas Aplicados																
W Score	NA			Sample < 10				Sample < 10				Sample < 10				
Standard Score ^a	NA			Sample < 10				Sample < 10				Sample < 10				
N (range)	20–48				12–44				10–13				12–23			

Sources: UPCOS-6 Fall 2012 and Spring 2013 Direct Child Assessments.

Note: All analyses are weighted to represent children attending LAUP in the 2012–2013 program year.

^aNational mean for standard scores is 100, with a standard deviation of 15.

^bThis measure was conceptually scored. For the English edition, the standard score was generated only for children assessed in English. For the Spanish bilingual edition, the standard score was generated for all children in the sample.

Looking across Tables VI.3 and VI.4, we see that average expressive language scores in the fall were lower in each language group for the children whose change scores met targets than for children whose change scores did not. Despite the fall pattern, by the spring, expressive language scores were higher for all language groups (with the exception of the Spanish-only group) among those whose progress met targets than among those whose progress did not.

Literacy: RLN. Children whose scores did meet the target (Table VI.3) were able to name 11 to 12 more letters in the spring than in the fall. For children whose progress did not exceed targets (Table VI.4), on average, letter naming abilities were flat over the course of the 5.7 months between assessments.

Looking across Tables VI.3 and VI.4, we see that average letter naming scores in the fall were lower in each language group for the children whose change scores met targets than for children whose change scores did not. By the spring, scores were similar for each language group, except for the Spanish-only group.

Fine Motor and Literacy: WJ-III Spelling. The third section of Table VI.3 shows fall and spring scores on the WJ-III Spelling by language group for children whose progress exceeded the Level 1 target. Children in both the English-only and English-primarily groups began the year with scores similar to a national sample of same-age peers (average standard scores were 103.0 and 102.8, respectively), and ended the year with scores exceeding a national sample of peers (mean SS = 107.6 and 107.7 for the English-only and English-primarily groups, respectively).

Table VI.4 shows that children in the English-only and English-primarily groups who did not meet the targets lost ground relative to a national sample of peers. They began the year with scores above a national sample of peers (mean SS = 103.8 and 104.3, respectively) but ended the year with scores below the national mean (mean SS = 94.3 and 95.3, respectively). Too few children in the other language groups whose progress did not meet targets were scored relative to English norms to present the data here.

Looking across Tables VI.3 and VI.4, we see that average scores in the fall were lower for the children whose change scores met targets than for children whose change scores did not. The pattern reversed in the spring.

Mathematics: WJ-III Applied Problems. The fourth section of Table VI.3 shows mathematics scores by language group for the sample whose progress exceeded targets. Both language groups with a sufficient number of children that completed this assessment (English only and English primarily) began the year with mathematics skills below but close to the national mean and ended above the national mean in the spring (mean SS = 103.8 and 103.6 for the English-only and English-primarily groups, respectively).

When examining scores for the group whose scores did not meet the target by language (Table VI.4), children in the English-only group entered their programs with average standard scores above the national mean in the fall but lost ground to their peers in the spring. The decline in scores between the fall and spring was similar for both language groups (mean SS = 103.8 and 104.3 in the fall, and 94.3 and 95.3 in the spring, respectively). Too few children in the other language groups whose progress did not meet targets were scored relative to English norms to present the data here.

Looking across Tables VI.3 and VI.4, we see that average mathematics scores in the fall were lower for the children whose change scores met the Level 1 target than for children whose change scores did not. By the spring, scores were higher among children whose progress met the target.

Mathematics: WM-III Problemas Aplicados. Shifting to the bottom section of Table VI.3, we see that children in the English-primarily, Spanish-only, and Spanish-primarily groups whose scores exceeded targets made progress relative to a national sample of peers in mathematics. However, scores for both groups were still one-half to two-thirds of a standard deviation below the national mean in the spring.

Although the change in WM-III Problemas Aplicados scores for a subset of children did not surpass the Level 1 target, the subgroups defined by language include fewer than 10 children. Thus, Table VI.4 is not discussed here.

D. Summary and Implications

In these sensitivity analyses, patterns were similar for the analysis by language group and the analysis by fall quartile. In the quartile analysis, we see that, for each measure and quartile, the fall means for children whose progress met targets were lower than or similar to the fall means for children whose performance did not meet targets. In general, within each language group, scores in the fall were lower for the children whose scores met targets than for children whose scores did not for each measure. Typically, scores in the spring were higher for the group whose progress met targets across almost all subgroups and tests. The one exception is the Spanish-only group on the EOWPVT-SBE. While the children in this group scored above the Spanish-only group who met the targets in the spring, on average, they lost considerable ground (more than one-half of a standard deviation) in relation to their bilingual peers. Note that, in a number of instances, the sample size of the quartile or language group fell below 10; thus, some comparisons could not be made.

Looking across the sensitivity analyses presented in this chapter, findings generally confirm the validity of the targets. Children whose progress met or exceeded targets progressed relative to a national sample of peers and in most cases had higher scores in the spring than children whose progress did not meet targets. However, the findings also highlight that children who enter performing at high levels do not necessarily broaden or deepen their skills over the year. Although we expected that children who entered with the highest skills would show less change, it is a concern that they lost so much ground relative to peers. Particularly in vocabulary, focusing on increasing the skills of those children will benefit all children in the classroom, as they all will be exposed to more sophisticated words. Analyses of UPCOS-3 data (Xue et al. 2010a) indicated that variation in vocabulary in the fall in the classroom is related to growth in vocabulary for individual children, even when controlling for many other factors. This suggests that children are learning vocabulary from each other, and that advancing the knowledge of the highest-performing children will benefit other children.

VII. AN EXPLORATORY ANALYSIS OF CHILD AND FAMILY CHARACTERISTICS ASSOCIATED WITH CHILDREN'S PROGRESS

In the first volume of this report, we present child progress in 2012–2013 (Volume 1, Chapter II) and provide an introduction to the children and families of LAUP (Volume 1, Chapter IV). In this chapter, we describe exploratory analyses addressing the relationships between children's language and literacy progress and characteristics of the child, family, and program. We conducted multivariate HLM models and drew on data from fall and spring direct child assessments and the fall parent questionnaire.

A. Approach to Analysis

We used hierarchical linear modeling (HLM) to examine associations between children's progress in language and literacy and child and family characteristics and program type, accounting for the clustering of children and families 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 focused on the EOWPVT and RLN because they are conceptually scored and, thus, available for more children than the language-specific subtests from the WJ III and WM III.

We used covariate-adjusted fixed effects models to estimate children's progress in language and literacy outcomes between the fall and spring of the LAUP year.²⁸ In addition to using children's spring IRT scores on the EOWPVT and RLN as the dependent variables, we also examined what factors were associated with whether children knew 14 or more letters by spring (whether the second RLN target was met).²⁹ We included children's fall scores in the models to help explain the variance in the respective spring outcomes. Even though we had attempted to maintain about the same time interval in terms of the number of months between fall and spring assessments, the intervals ranged from 4.1 to 6.3, with a mean of 5.7 months, largely because of variations in child and program availability. Therefore, we added a variable to the models indicating the number of months between the fall and spring assessments for each child, controlling for the variation in assessment interval. We examined other child and family factors that might be related to children's outcomes in the spring and controlled for those that were related.

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 the average baseline ability of children in the classroom (mean peer ability) and the baseline variation of children's ability within the classroom (variation of peer ability) in the HLM models. Fall EOWPVT scores in the classroom were used for models examining progress on the EOWPVT, and fall RLN scores in the classroom were used for models examining progress on the RLN and whether the RLN target was met. It should be noted that the associations found in this

²⁸ To include more children in the analyses, we included missing data indicators for variables with missing data in the model.

²⁹ See Chapter 2 in this volume for additional detail on LAUP's performance based contract targets.

study should not be interpreted as causal relationships but just descriptive of characteristics that tend to increase or decrease together (correlated) in this sample of children and programs.

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. Table VII.1 identifies the categories of variables included in each model. In Model 1, we included child characteristics (gender, age in months, race/ethnicity, and language group in addition to assessment interval and fall scores). In Model 2, we added an indicator for high family risk (in which having three or more of five socioeconomic risk factors defined one has having “high” family risk; see Box VII.1 for the indicators that comprise the risk index).³⁰ In Model 3, we added the individual risk indicators shown in Box VII.1 but removed the high family risk indicator. In Model 4, we investigated at-risk and individual family characteristics together to help understand whether having specific socioeconomic risks contributed over and above the family risk index. Finally, in Model 5, we added mean peer abilities, variation of peer abilities, and program type to the models. We estimated the models using the sampling weights that account for sampling probabilities and non-response at the individual level. We present the results from the HLM analyses in Tables 1 to 3.

Table VII.1. Independent Variables Included in HLM Models

	Child Characteristics	High Family Risk (Based on Index)	Family Risk Indicators	Peer Abilities	Program Type
Model 1	X				
Model 2	X	X			
Model 3	X		X		
Model 4	X	X	X		
Model 5	X	X	X	X	X

Box VII.1. Indicators Comprising the Family Risk Index

- Maternal educational attainment less than high school
- Single-parent household
- Teenage parent
- Either parent born outside the United States
- Household size of five or greater

“High risk” defined as three or more

B. Factors Associated with Child Progress

Results are presented in Tables VII.2 (EOWPVT), VII.3 (RLN), VII.4 (RLN target). The outcomes in tables VII.2 and VII.3 are standardized ($\hat{\alpha}$ scores); coefficients may be interpreted as the change in the outcome in standard deviation units for each one 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. For models in table VII.4, we used a logistic regression model; coefficients should be interpreted as the change in the likelihood of meeting the target (with positive coefficients indicating a higher likelihood). Depending on the

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. The number of risk factors available in our study was more limited than the number investigated in some other research including UPCOS-2 (Love et al. 2009). It is possible that some of the families designated as low risk in our analysis actually experience more socioeconomic risks. Thus, our risk index is likely an underestimate of family risk.

outcomes, models include 261 to 272 children; this small sample size should be kept in mind when interpreting results.

Models specified with only child characteristics and time interval between assessments explained more than half (52 to 63 percent) of the variance in the spring scores. The family characteristics and classroom/program variables addressed in Model 2 to 5 did not add much to the explained variance, suggesting that the some other family and program characteristics that were not measured in our study are more important in predicting children's development. In the rest of this section, we describe what child, family, and program factors are associated with outcomes model-by-model. In section C, we highlight key findings based on Model 5 for each outcome (i.e., those factors that are statistically significant in the fully controlled models).

Relationships of Child Initial Ability and Child Characteristics to Progress in Language and Literacy (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 higher spring scores for both the EOWPVT and RLN. Children with higher initial ability in the fall were also more likely to meet the second RLN target (naming at least 14 letters in the spring).

Consistent with the findings from the subgroup descriptive analysis (Volume 2, Chapter IV), we found that household language was not significantly associated with children's EOWPVT scores in the spring controlling for race/ethnicity, assessment interval, and fall scores. Children's progress is similar across language groups. For RLN, HLM results show that children in the Spanish only group performed lower in the spring than children in the English only group controlling for race/ethnicity, assessment interval, and fall scores. Children in the Spanish only group were also less likely to meet the RLN target than children in the English only group.

Child gender and age were not significantly associated with children's progress in language and literacy, but were significantly associated with the likelihood of naming at least 14 letters in the spring. Boys and younger children were more likely to meet this RLN target than girls and older children after controlling for initial ability, assessment time interval, and other child characteristics. Race/ethnicity was associated with children's progress on the EOWPVT, but was not significantly associated with progress on the RLN. Asian children scored higher on the EOWPVT in the spring than Hispanic children, controlling for other child characteristics in the model. Race/ethnicity was significantly associated with the likelihood of naming at least 14 letters in the spring: African American children were more likely to meet this target than Hispanic children.

Relationships of Family Characteristics to Progress in Language and Literacy (Models 2, 3, and 4)

We added to the models the high family risk indicator based on the index (Model 2), then the variables for individual family risk factors (Model 3), and then both together (Model 4) to show their unique contributions in explaining the variance in children's progress in language and literacy. The family risk index was not significantly associated with progress on the EOWPVT or RLN when added to the model without the other individual family characteristics or examined in combination with individual family characteristics. However, the risk index was significantly associated with the likelihood of meeting the RLN target: children with high family risk were less likely to meet the RLN target than children with low risk. However, the statistical significance of high family risk disappeared when family risk indicators were included in the model as well.

When we examined individual family risk indicators without the family risk index, we found that having a single parent was associated less progress on and the RLN; the coefficient also approached significance for the EOWPVT. Controlling for other child characteristics, children whose mothers had less than a high school education, who had a single parent, or whose parents were born outside the United States were less likely to meet the RLN target; children with a household size of five or greater were more likely to meet the RLN target than those in a smaller household.

Relationships of Peer Ability and Program Type to Progress in Language and Literacy (Model 5)

In Model 5, we examined associations of peer ability and program type with children's progress in language and literacy, controlling for child and family characteristics. We included the average entering ability of sample children within a classroom (classroom mean fall scores) for each outcome, the within-classroom variation (standard deviation) of children's entering ability, and whether the program was a center or FCC. Classroom mean entering ability and variation of entering ability were not significantly associated with children's progress on the EOWPVT or RLN.

Peer abilities were related to the likelihood of meeting the RLN target. The greater the mean peer abilities in the classroom in the fall (that is, the higher the classroom average entering score on the RLN), the less likely children were to have met the RLN target in the spring. Similarly, the greater the variation in peer abilities in the fall, the less likely children were to have met the RLN target.

Program type was not associated with progress on the EOWPVT, but was significantly associated with progress on the RLN: children in FCCs made greater progress than children in centers on the RLN; children in FCCs were also more likely to meet the RLN target than children in centers.

C. Summary

In these exploratory analyses of what child, family, and program factors are associated with language and literacy outcomes, we found few relationships with progress, which may reflect limited power to detect associations due to our small sample size, or alternatively may indicate that factors other than what we were able to measure are important for children's outcomes. However, we found that a number of factors were associated with the likelihood of meeting one of LAUP's performance-based contract targets—knowing 14 or more letters in the spring. Focusing on Model 5 for each outcome (i.e., the model that includes all child, family, and program factors), we see that children's progress in language (as measured by the EOWPVT) and letter-naming (as measured by the RLN) is explained largely by their skills in each of these areas at the start of the program year. Children across language groups made similar progress from fall to spring in language and letter naming. However, children in the English primarily, Spanish primarily, and other language groups were more likely to meet the RLN target of easily naming 14 letters in the spring than children in the English only group. Race/ethnicity was also associated with outcomes. Asian children scored higher on the EOWPVT in the spring than Hispanic children. African American and Asian children were more likely and white children were less likely to meet the RLN target than Hispanic children.

Family factors were associated with both letter-naming progress and the likelihood of meeting the RLN target, but not progress in language. Having a single parent was associated with less success in learning letter names. Children with a single parent and those with at least one parent born

outside of the U.S. were less likely to meet the RLN target, while children in larger households (5 or more household members) were more likely to meet the target.

In terms of program factors, children in FCCs made more progress than children in centers on the RLN, and they were also more likely to meet the RLN target. The pattern of findings regarding the level and variation in baseline abilities was somewhat surprising—higher mean ability and greater variation in ability among classroom peers in the fall were associated with a lower likelihood of meeting the RLN target of easily naming 14 or more letters in the spring. This may suggest that when some children in the classroom are already naming letters, teachers are not practicing letter names as frequently or may only be practicing with children who know very few letters, potentially limiting the range of letters that children practice. When mean ability in letter-naming is high, teachers also may be focusing on other child outcomes. For LAUP, continued monitoring of children’s knowledge and progress by teachers will help teachers to understand which children may need more support in learning in different areas including letter-naming. Fluency in naming letters supports children in memory skills in a variety of domains as well as in early literacy.

Table VII.2. Relationships of Child and Family Characteristics with Children's Progress on the EOWPVT

	Model 1	Model 2	Model 3	Model 4	Model 5
Gender (male)	0.04 (0.08)	0.04 (0.08)	0.02 (0.08)	0.02 (0.08)	0.02 (0.08)
Age	0.00 (0.01)	0.00 (0.01)	0.00 (0.01)	0.00 (0.01)	0.00 (0.01)
Home language					
English only (reference)					
English primarily	0.11 (0.10)	0.11 (0.10)	0.10 (0.11)	0.09 (0.11)	0.11 (0.11)
Spanish only	-0.19 (0.16)	-0.18 (0.16)	-0.17 (0.17)	-0.17 (0.17)	-0.15 (0.18)
Spanish primarily	-0.10 (0.13)	-0.11 (0.13)	-0.07 (0.14)	-0.07 (0.14)	-0.05 (0.15)
Other language only or primarily	-0.15 (0.29)	-0.13 (0.29)	-0.24 (0.30)	-0.25 (0.31)	-0.26 (0.31)
Race/ethnicity					
Hispanic (reference)					
White	0.22 (0.19)	0.21 (0.19)	0.18 (0.19)	0.18 (0.19)	0.18 (0.20)
African American	0.00 (0.18)	0.00 (0.18)	0.01 (0.18)	0.01 (0.18)	0.00 (0.18)
Asian/Pacific Islander	0.31 (0.15)*	0.29 (0.15)	0.34 (0.15)*	0.35 (0.16)*	0.34 (0.16)*
Initial ability	0.75 (0.04)***	0.74 (0.05)***	0.73 (0.05)***	0.72 (0.05)***	0.72 (0.05)***
Assessment time interval	0.12 (0.14)	0.11 (0.14)	0.05 (0.15)	0.04 (0.15)	0.04 (0.15)
High family risk		-0.09 (0.12)		-0.04 (0.16)	-0.06 (0.17)
Family risk indicators					
Less than a high school education			-0.05 (0.11)	-0.04 (0.14)	-0.04 (0.14)
Single parent			-0.29 (0.14)	-0.30 (0.15)	-0.30 (0.15)
Teenage parent			0.11 (0.21)	0.10 (0.21)	0.12 (0.21)
Parent born outside of the U.S.			-0.10 (0.13)	-0.10 (0.13)	-0.08 (0.13)
Household size of 5 or greater			0.10 (0.08)	0.11 (0.09)	0.13 (0.09)
Mean peer abilities					0.02 (0.12)
Variation in peer abilities					-0.35 (0.23)
Program type (FCC)					0.09 (0.14)

Source: UPCOS-6 Fall 2012 and Spring 2013 Direct Child Assessments and Fall 2012 UPCOS-6 Brief Parent Questionnaire.

Note: Table presents regression coefficients and standard errors (in parentheses) from HLM models.

*p<.05; **p<.01; ***p<.001.

Table VII.3. Relationships of Child and Family Characteristics with Children's Progress on the RLN

	Model 1	Model 2	Model 3	Model 4	Model 5
Gender (male)	0.03 (0.08)	0.04 (0.08)	0.04 (0.08)	0.05 (0.08)	0.05 (0.08)
Age	0.00 (0.01)	-0.01 (0.01)	0.00 (0.01)	0.00 (0.01)	0.00 (0.01)
Home language					
English only (reference)					
English primarily	0.04 (0.10)	0.03 (0.10)	0.02 (0.11)	0.03 (0.11)	0.04 (0.12)
Spanish only	-0.35 (0.16)*	-0.36 (0.17)*	-0.35 (0.18)	-0.36 (0.18)	-0.36 (0.18)
Spanish primarily	-0.11 (0.13)	-0.12 (0.13)	-0.09 (0.15)	-0.10 (0.15)	-0.08 (0.15)
Other language only or primarily	0.09 (0.20)	0.11 (0.20)	0.09 (0.21)	0.09 (0.21)	0.13 (0.21)
Race/ethnicity					
Hispanic (reference)					
White	-0.27 (0.22)	-0.28 (0.22)	-0.30 (0.22)	-0.30 (0.22)	-0.28 (0.22)
African American	0.37 (0.19)	0.39 (0.19)*	0.36 (0.20)	0.37 (0.20)	0.35 (0.20)
Asian/Pacific Islander	0.28 (0.16)	0.23 (0.16)	0.27 (0.17)	0.25 (0.17)	0.23 (0.17)
Initial ability	0.69 (0.05)***	0.68 (0.05)***	0.68 (0.05)***	0.68 (0.05)***	0.69 (0.05)***
Assessment time interval	0.21 (0.16)	0.21 (0.16)	0.21 (0.17)	0.21 (0.17)	0.22 (0.17)
High family risk		-0.03 (0.13)		0.11 (0.18)	0.09 (0.18)
Family risk indicators					
Less than a high school education			0.02 (0.13)	-0.03 (0.16)	-0.01 (0.16)
Single parent			-0.43 (0.15)*	-0.43 (0.17)*	-0.42 (0.17)*
Teenage parent			-0.05 (0.24)	-0.04 (0.24)	-0.01 (0.24)
Parent born outside of the U.S.			-0.09 (0.14)	-0.09 (0.14)	-0.09 (0.14)
Household size of 5 or greater			0.00 (0.09)	-0.03 (0.10)	-0.03 (0.10)
Mean peer abilities					-0.05 (0.13)
Variation in peer abilities					-0.06 (0.26)
Program type (FCC)					0.34 (0.16)*

Source: UPCOS-6 Fall 2012 and Spring 2013 Direct Child Assessments and Fall 2012 UPCOS-6 Brief Parent Questionnaire.

Note: Table presents regression coefficients and standard errors (in parentheses) from HLM models.

*p<.05; **p<.01; ***p<.001.

Table VII.4. Relationships of Child and Family Characteristics with the Likelihood of Meeting the RLN Target

	Model 1	Model 2	Model 3	Model 4	Model 5
Gender (male)	0.29 (0.06)***	0.28 (0.06)***	0.26 (0.07)***	0.27 (0.07)***	0.31 (0.07)***
Age	-0.05 (0.01)***	-0.05 (0.01)***	-0.04 (0.01)***	-0.04 (0.01)***	-0.04 (0.01)***
Home language					
English only (reference)					
English primarily	0.11 (0.08)	0.12 (0.08)	0.28 (0.09)**	0.29 (0.09)**	0.32 (0.09)***
Spanish only	-0.23 (0.11)*	-0.18 (0.12)	0.33 (0.13)*	0.31 (0.13)***	0.21 (0.13)
Spanish primarily	0.46 (0.09)***	0.44 (0.09)***	0.92 (0.11)***	0.92 (0.11)***	0.92 (0.11)***
Other language only or primarily	0.29 (0.18)	0.26 (0.18)	0.43 (0.18)*	0.45 (0.18)*	0.48 (0.19)*
Race/ethnicity					
Hispanic (reference)					
White	-0.29 (0.18)	-0.28 (0.18)	-0.53 (0.19)**	-0.52 (0.19)**	-0.42 (0.20)*
African American	1.78 (0.20)***	1.77 (0.20)***	1.80 (0.21)***	1.82 (0.21)***	1.69 (0.20)***
Asian/Pacific Islander	0.06 (0.15)	0.02 (0.15)	0.37 (0.15)*	0.34 (0.16)*	0.60 (0.16)***
Initial ability	0.31 (0.01)***	0.31 (0.01)***	0.31 (0.01)***	0.31 (0.01)***	0.34 (0.01)***
Assessment time interval	1.44 (0.11)***	1.42 (0.11)***	1.50 (0.12)***	1.51 (0.12)***	1.77 (0.12)***
High family risk		-0.19 (0.09)*		0.26 (0.14)	0.02 (0.14)
Family risk indicators					
Less than a high school education			-0.29 (0.09)**	-0.40 (0.12)***	-0.20 (0.12)
Single parent			-0.32 (0.13)*	-0.34 (0.14)*	-0.37 (0.14)*
Teenage parent			-0.01 (0.16)	0.01 (0.16)	0.09 (0.16)
Parent born outside of the U.S.			-0.99 (0.11)***	-0.99 (0.11)***	-1.08 (0.11)***
Household size of 5 or greater			0.33 (0.07)***	0.26 (0.08)**	0.42 (0.08)***
Mean peer abilities					-0.05 (0.01)***
Variation in peer abilities					-0.08 (0.02)***
Program type (FCC)					1.48 (0.13)***

Source: UPCOS-6 Fall 2012 and Spring 2013 Direct Child Assessments and Fall 2012 UPCOS-6 Brief Parent Questionnaire.

Note: Table presents regression coefficients and standard errors (in parentheses) from HLM models.

*p<.05; **p<.01; ***p<.001.

REFERENCES

- Atkins-Burnett, Sally, Kathy Sonnenfeld, and Susan Sprachman. "Rapid Letter Naming (RLN)." Washington, DC: Mathematica Policy Research, 2007.
- Brownell, R. "Expressive One-Word Picture Vocabulary Tests." San Antonio, TX: Harcourt Assessment, Inc., 2000.
- Burchinal, Margaret, S. Field, Michael L. Lopez, Carollee Howes, and Robert Pianta. "Instruction in Spanish and Outcomes for Pre-Kindergarten English Language Learners." University of Virginia CASTL Research Brief, 2013. Available at [[http://curry.virginia.edu/uploads/resourceLibrary/CASTL_Research_Brief-Burchinal_et_al._\(2012\)_ECRQ.pdf](http://curry.virginia.edu/uploads/resourceLibrary/CASTL_Research_Brief-Burchinal_et_al._(2012)_ECRQ.pdf)]. Accessed July 29, 2013.
- Burchinal, Margaret, S. Field, Michael L. Lopez, Carollee Howes, and Robert Pianta. "Instruction in Spanish in Pre-Kindergarten Classrooms and Child Outcomes for English Language Learners." *Early Childhood Research Quarterly*, vol. 27, 2011, pp.188–197.
- California Department of Education. "California Preschool Learning Foundations." Sacramento, CA: 2008.
- Centers for Disease Control and Prevention (CDC). "About BMI for Children and Teens." Available at [http://www.cdc.gov/healthyweight/assessing/bmi/childrens_bmi/about_childrens_bmi.html]. Last updated September 13, 2011. Accessed January 18, 2013.
- Centers for Disease Control and Prevention (CDC). "Obesity Prevalence Among Low-Income, Preschool-Aged Children—New York City and Los Angeles County, 2003–2011." *Morbidity and Mortality Weekly Report*, vol. 62, no. 2, 2013, pp. 17–22.
- Downey, D.B., P.T. von Hippel, and B. Broh. "Are Schools the Great Equalizer? Cognitive Inequality During the Summer Months and the School Year." *American Sociological Review*, vol. 69, no. 5, 2004, pp. 613–635.
- Duncan, Sharon, and Edward DeAvila. "Pre-Language Assessment Scales [*preLAS* 2000]." Monterey, CA: CTB-McGraw Hill, 2002.
- Early, Diane, Oscar Barbarin, Donna Bryant, Margaret Burchinal, Florence Chang, Richard Clifford, Gisele Crawford, Wanda Weaver, Carollee Howes, Sharon Ritchie, Marcia Kraft-Sayre, Robert Pianta, and W. Steven Barnett. "Pre-Kindergarten in Eleven States: NCEDL's Multi-State Study of Pre-Kindergarten and Study of State-Wide Early Education Programs (SWEEP)." Preliminary descriptive report, NCEDL working paper. Chapel Hill, NC: University of North Carolina, May 24, 2005.
- First 5 LA. "Strengthening Families and Communities in LA County." First 5 LA Strategic Plan FY 2009–2015. Approved 6-11-09. Available at [http://www.first5la.org/files/F5LA_STRATEGIC_PLAN_2009-15.pdf]. Accessed July 25, 2013.

- Gormley, Jr., William T., Ted Gayer, Deborah Phillips, and Brittany Dawson. "The Effects of Universal Pre-K on Cognitive Development." *Developmental Psychology*, vol. 41, no. 6, 2005, pp. 872–884.
- Hulsey, Lara, Nikki Aikens, Ashley Kopack, Jerry West, Emily Moiduddin, and Louisa Tarullo. "Head Start Children, Families, and Programs: Present and Past Data from FACES." OPRE report 2011-33a. Washington, DC: U.S. Department of Health and Human Services, December 2011.
- Kagan, S.L., E. Moore, and S. Bredekamp (eds.). "Reconsidering Children's Early Development and Learning: Toward Shared Beliefs and Vocabulary." Washington, DC: National Education Goals Panel, 1995.
- Love, John M., S. Atkins-Burnett, C. Vogel, N. Aikens, Y. Xue, M. Mabutas, B.L. Carlson, E.S. Martin, N. Paxton, M. Caspe, S. Sprachman, and K. Sonnenfeld. "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." Report submitted to First 5 LA. Princeton, NJ: Mathematica Policy Research, June 2009.
- Moiduddin, Emily, Sally Atkins-Burnett, Yange Xue, Pia Caronongan, Elisha Smith, and Marta Induni. "Results of Activities Informing the Performance-Based Contract Between First 5 LA and LAUP." Final report submitted to First 5 LA. Washington, DC: Mathematica Policy Research, June 30, 2010.
- Moiduddin, Emily, Sally Atkins-Burnett, and Yange Xue. "Confirmation of the 2012–2013 Child Progress Targets." Memo submitted to First 5 LA. Washington, DC: Mathematica Policy Research, November 5, 2012.
- Moiduddin, Emily, Yange Xue, and Sally Atkins-Burnett. "Informing the Performance-Based Contract Between First 5 LA and LAUP: Child Progress in the 2010-2011 Program Year." Final report submitted to First 5 LA. Washington, DC: Mathematica Policy Research, November 30, 2011.
- NICHD Early Child Care Research Network (ed.). *Child Care and Child Development: Results of the NICHD Study of Early Child Care and Youth Development*. New York: Guilford Press, 2005.
- Peisner-Feinberg, E.S., and J.M. Schaff. "Evaluation of the North Carolina More at Four Pre-Kindergarten Program: Performance and Progress in the Seventh Year (2007–2008)." Chapel Hill, NC: FPG Child Development Institute, 2008.
- Rathbun, A., and J. West. *From Kindergarten Through Third Grade: Children's Beginning School Experiences*. NCES 2004-007. Washington, DC: National Center for Education Statistics, 2004.
- Roid, Gale H., and Lucy J. Miller. "Leiter-R Performance Scale—Revised." Wood Dale, IL: Stoelting Co., 1997.
- West, Jerry, Louisa Tarullo, Nikki Aikens, Lizabeth Malone, and Barbara Lepidus Carlson. "FACES 2009 Study Design." OPRE report 2011-9. Washington, DC: U.S. Department of Health and Human Services, Administration for Children and Families, Office of Planning, Research and Evaluation, 2011.

- Woodcock, R.W., K. McGrew, and N. Mather. “Woodcock-Johnson III Tests of Achievement.” Itasca, IL: Riverside Publishing, 2001/2007.
- Woodcock, R.W., A.F. Muñoz-Sandoval, K. McGrew, N. Mather, and F. Schrank. “Bateria III Woodcock-Muñoz.” Itasca, IL: Riverside Publishing, 2004/2007.
- Xue, Yange, Sally Atkins-Burnett, and Emily Moiduddin. “Children’s Progress During the 2011-2012 LAUP Year.” Final report submitted to First 5 LA. Washington, DC: Mathematica Policy Research, 2012.
- Xue, Yange, Sally Atkins-Burnett, and Emily Moiduddin. “Developing Child Progress Targets for the Performance-Based Contract, Revised Memo.” Memo submitted to First 5 LA. Washington, DC: Mathematica Policy Research, April 12, 2011.
- Xue, Yange, Sally Atkins-Burnett, and Nikki Aikens. “A Study of Peer Effects on Children’s Language Development in Linguistically Diverse Preschools.” Presented at the Head Start National Research Conference, Washington, DC, June 2010a.
- Xue, Yange, Sally Atkins-Burnett, Pia Caronongan, and Emily Moiduddin. “Informing the Performance-Based Contract Between First 5 LA and LAUP: Assessing Child Progress.” Final report submitted to First 5 LA. Washington, DC: Mathematica Policy Research, December 10, 2010b.



MATHEMATICA
Policy Research

www.mathematica-mpr.com



Improving public well-being by conducting high quality, objective research and surveys

Princeton, NJ ■ Ann Arbor, MI ■ Cambridge, MA ■ Chicago, IL ■ Oakland, CA ■ Washington, DC

Mathematica® is a registered trademark of Mathematica Policy Research

