

**Informing the Performance-Based  
Contract Between First 5 LA and  
LAUP: Child Progress in the 2011-  
2012 Program Year**

Final Report

September 25, 2012

Yange Xue  
Sally Atkins-Burnett  
Emily Moiduddin

**MATHEMATICA**  
**Policy Research**



Champions For Our Children



Contract Number:  
07110

Mathematica Reference Number:  
06940.314, 40079.201

Submitted to:  
First 5 LA  
750 N. Alameda St., Suite 300  
Los Angeles, CA 90012  
Project Officer: Sharon Murphy and  
Artineh Samkian

Submitted by:  
Mathematica Policy Research  
600 Maryland Avenue, SW  
Suite 550  
Washington, DC 20024-2512  
Telephone: (202) 484-9220  
Facsimile: (202) 863-1763  
Project Director: Emily Moiduddin

**Informing the Performance-  
Based Contract Between First 5  
LA and LAUP: Child Progress in  
the 2011–2012 Program Year**

Final Report

September 25, 2012

Yange Xue  
Sally Atkins-Burnett  
Emily Moiduddin

**MATHEMATICA**  
**Policy Research**



Champions For Our Children



## CONTENTS

EXECUTIVE SUMMARY .....	ix
<b>I</b> INTRODUCTION: SETTING TARGETS FOR CHILD PROGRESS .....	1
A. Domains of Child Development and Selected Measures .....	1
B. Target-Setting Procedures.....	3
C. Targets Established for the 2011–2012 Program Year.....	5
D. Evidence of Validity of the Targets for the 2011–2012 Program Year .....	5
<b>II</b> STUDY METHODS.....	7
A. Sample .....	7
B. Measures.....	7
C. Procedures .....	9
D. Scoring.....	10
<b>III</b> CHILD PROGRESS IN THE 2011–2012 PROGRAM YEAR .....	13
A. Approach to Analysis.....	13
B. Progress Across All Children.....	14
C. Progress by Language Group .....	21
D. Summary and Implications.....	29
<b>IV</b> CHILD PROGRESS RELATIVE TO TARGETS SET FOR THE 2011–2012 PROGRAM YEAR .....	31
A. Approach to Analysis.....	31
B. Progress Relative to Targets .....	31
C. Progress by Whether Level 1 Target Was Met .....	35
D. Summary and Implications.....	40
REFERENCES .....	45
APPENDIX A: PERCENTAGE OF CHILDREN WHO SWITCHED LANGUAGE OF ASSESSMENT FROM FALL TO SPRING .....	A.1

APPENDIX B:	<i>pre</i> LAS FOR CHILDREN IN THE OTHER LANGUAGE ONLY OR PRIMARILY GROUP: FALL 2011 AND SPRING 2012 .....	B.1
APPENDIX C:	FALL AND SPRING SCORES BY LANGUAGE GROUP AND PERFORMANCE RELATIVE TO TARGETS.....	C.1

## TABLES

<b>ES.1</b>	Child Outcomes Measures and Developmental Domains Addressed in the First 5 LA-LAUP Performance-Based Contract.....	x
<b>ES.2</b>	Summary of Child Progress in 2011–2012 Program Year in a Representative Sample of LAUP Programs and Children Relative to the Performance-Based Contract Targets.....	xiii
<b>I.1</b>	Child Outcomes Measures and Developmental Domains Addressed in the First 5 LA-LAUP Performance-Based Contract.....	2
<b>I.2</b>	Final Targets for Child Progress in LAUP Programs During the 2011–2012 Program Year, by Domain and Measure .....	6
<b>II.1</b>	Number of Children Assessed in Fall 2011 and Spring 2012, by Language Group (N = 586).....	7
<b>III.1</b>	Means and Standard Errors for UPCOS-5 Language, Literacy, and Math Measures: Fall 2011 and Spring 2012 .....	15
<b>III.2</b>	Percentage of Children Scoring in the Expected Range on the Leiter-R: Fall 2011 and Spring 2012.....	20
<b>III.3</b>	Means and Standard Errors for UPCOS-5 Language, Literacy, and Math Standardized Measures: Fall 2011 and Spring 2012, by Language Group .....	23
<b>III.4</b>	Change Scores for UPCOS-5 Language, Literacy, and Math Standardized Measures: Fall 2011–Spring 2012, by Language Group .....	24
<b>III.5</b>	Percentage of Children Scoring in the Expected Range on the Leiter-R: Fall 2011 and Spring 2012, by Language Group.....	25
<b>III.6</b>	Change in Percentage of Children Scoring in the Expected Range on the Leiter-R: Fall 2011–Spring 2012, by Language Group .....	26
<b>IV.1</b>	Summary of Child Progress in 2011–2012 in a Representative Sample of LAUP Programs and Children Relative to Performance-Based Contract Targets.....	32
<b>IV.2</b>	Percentage of Children Whose Progress Exceeded Level 1 Targets by Fall Quartile .....	34
<b>IV.3</b>	Means, Standard Errors, and Ranges for UPCOS-5 Language, Literacy, and Math Measures Change Scores, Child Level, by Whether Level 1 Target Was Met .....	36
<b>IV.4</b>	Means and Standard Errors for UPCOS-5 Language, Literacy, and Math Standardized Measures: Fall 2011 and Spring 2012 for	

	Children Whose Change Score Met or Exceeded Change Specified in Level 1 Targets, by Fall Quartile .....	37
<b>IV.5</b>	Means and Standard Errors for UPCOS-4 Language, Literacy, and Math Standardized Measures: Fall 2011 and Spring 2012 for Children Whose Change Score Fell Below Change Specified in Level 1 Targets, by Fall Quartile .....	38
<b>A.1</b>	Percentage of Children Who Were Assessed in English or Spanish in the Fall and Spring, by Language Group .....	A.3
<b>B.1</b>	Means and Standard Errors of <i>pre</i> LAS for Children in the Other Language Only or Primarily Group: Fall 2011 and Spring 2012 .....	B.3
<b>C.1</b>	Means and Standard Errors for UPCOS-5 Language, Literacy, and Math Standardized Measures: Fall 2011 and Spring 2012 for Children Whose Change Score Met or Exceeded Change Specified in Level 1 Targets, by Language Group .....	C.5
<b>C.2</b>	Means and Standard Errors for UPCOS-5 Language, Literacy, and Math Standardized Measures: Fall 2011 and Spring 2012 for Children Whose Change Score Fell Below Change Specified in Level 1 Targets, by Language Group .....	C.6

## FIGURES

III.1	Mean IRT Scores for Language Assessment (EOWPVT): Fall 2011 and Change to Spring 2012 .....	16
III.2	Mean Raw and IRT Scores for Literacy Assessment (Rapid Letter Naming): Fall 2011 and Change to Spring 2012 .....	17
III.3	Mean W Scores for Literacy and Mathematics Assessments: Fall 2011 and Change to Spring 2012 .....	18
III.4	Mean Standard Scores for Language, Literacy, and Mathematics Assessments: Fall 2011 and Change to Spring 2012 .....	19
III.5	Percentage of Children Scoring in Expected Range for Social-Emotional and Approaches to Learning Assessments: Fall 2011 and Change to Spring 2012 .....	21
IV.1	Percentage of Children Whose Progress Exceeded Level 1 Targets by Fall Quartile .....	35



## EXECUTIVE SUMMARY

A central objective of LAUP is to support early learning and development. 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 First 5 LA and LAUP sought to track, to identify appropriate measures, and to set targets for progress across the year to be applied in the 2011–2012 program year. During the 2011–2012 program year, as part of Phase 5 of the Universal Preschool Child Outcomes Study (UPCOS-5), Mathematica conducted direct child assessments to determine whether the agreed upon targets were met. In this report, we describe children’s progress from fall to spring and report on children’s progress relative to the targets for the 2011–2012 program year.

### Domains of Child Development and Selected Measures

The first step in designing the study of child progress and determining targets 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 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), and be similar to those in the California Preschool Learning Foundations (California Department of Education 2008). 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 available<sup>1</sup>
- 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 that are important for all children in ways that are not specific to a particular curriculum
- National norms available for comparisons to national samples

Table ES.1 outlines the measures selected and domains addressed.

---

<sup>1</sup> 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.

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

Measure	Domains Addressed
Expressive One-Word Picture Vocabulary Test (EOWPVT), English Edition and Spanish Bilingual Edition (EOWPVT-SBE) (Brownell 2000)	Language development: vocabulary
Rapid Letter Naming (RLN; Atkins-Burnett et al. 2007)	Literacy development
Woodcock-Johnson III (WJ-III) Spelling (Woodcock et al. 2001/2007)	Fine motor skills and literacy development
Woodcock-Johnson III (WJ-III) and Woodcock-Muñoz Batería III (WM-III) Applied Problems (Woodcock et al. 2001/2007; Woodcock et al. 2004/2007)	Mathematics development and reasoning
Leiter Examiner Rating Scale – Revised (Leiter-R), Attention, Activity Level, and Sociability subscales (Roid and Miller 1997)	Social-emotional development and approaches to learning

## Sample and Methods

UPCOS-5 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. In the fall, the final sample included 39 programs and 660 children. The response rates were 65 and 93 percent for programs and children, respectively. In the spring, 597 children (90 percent of the total assessed in the fall) were assessed in the 39 programs. Looking across the fall and spring, 586 children were assessed at both time points (82 percent of the selected child-level sample).

To minimize the burden on children, they were randomly assigned to complete one of two versions of the assessment, each of which included only half of the measures in the battery. Children were routed to the appropriate language of assessment (Spanish or English) based on a combination of parent reports of home language and performance on an English language screener at the start of the assessment.

## Analysis

All analyses are at the child level, account for clustering of children within programs,<sup>2</sup> and are weighted to represent LAUP as a whole. For four of the measures—the EOWPVT/SBE, the WJ-III Spelling subtest, and the WJ-III and WM-III Applied Problems subtests—we examined absolute progress as well as progress relative to a national sample of same-age peers. For the RLN task, we examined absolute progress. For the Leiter-R, we calculated the percentage of children scoring in the expected range, which is based on performance of a national sample of same-age peers; children scoring in this range are unlikely to be experiencing difficulties with social-emotional development and approaches to learning. Selected targets are based on children’s absolute progress over the year

<sup>2</sup> Children within a classroom are generally more similar than children from different classrooms. The estimates of children’s progress within a classroom are not independent of one another, violating assumptions for some statistical tests of significance. We used standard statistical methods to account for this clustering (or grouping) of children within classrooms. These 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 are correctly estimated and, thus, that any statistical tests based on those standard errors are accurate.

for all measures except the Leiter-R. For the Leiter-R, targets are based on the percentage of children scoring in the expected range in the spring.

### **Average Progress in LAUP**

On average, 5.6 months passed between the two assessments (range of 3.4 to 6.4 months). Between the fall and spring, children in LAUP made statistically significant progress in all areas when based on absolute progress. Relative to a national sample of same-age peers, children in LAUP made statistically significant progress in expressive vocabulary (based on both English and bilingual samples). Children assessed in Spanish also made significant progress in mathematics relative to a national sample of Spanish-speaking same age peers. LAUP children kept pace with but did not make significant progress relative to a national sample of peers in fine motor/literacy (as measured by the WJ-III Spelling and WM-III Ortografía subtests) or mathematics.

Note that for the WJ-III Spelling subtest and the Leiter-R, children's performance was already quite strong relative to national peers in the fall and remained strong in the spring. Scores for these measures exceeded a national sample of peers in both the fall and spring.<sup>3</sup> Children also exceeded the average performance of their bilingual peers and made significant progress on the EOWPVT-SBE relative to their same-age bilingual peers. On average, children's scores on WJ-III Applied Problems subtest kept up with the national mean in both the fall and spring, suggesting that children progressed at the same rate as their same-age English-speaking peers. Scores on the RLN task indicated that, on average, children made statistically significant progress in their letter-naming skills and the results were comparable to other preschool studies using similar measures (Early et al. 2005).

Although children made progress in most areas, there is still room to grow. LAUP children made significant progress in expressive vocabulary relative to a national sample of English-speaking peers during the program year, but their mean score remained about half of a standard deviation below that of their English-speaking peers nationally in the spring. Although for children in LAUP who took the English assessment, mathematics abilities were similar to those of a national sample of English-speaking peers and these children maintained their progress relative to peers, mathematics abilities as measured by the WM-III (Spanish version) were below those of their Spanish-speaking peers at both time points. Despite the fact that children who took the Spanish version at both time points made enough progress so as to not lose additional ground relative to a national sample of similarly achieving peers, their scores in both the fall and spring were one standard deviation or more below the national mean.

### **Progress Relative to Targets Set During the 2011–2012 Program Year**

The child progress targets for the 2011–2012 program year were set in a collaborative process between First 5 LA and LAUP facilitated by Mathematica. Together, the team decided to set targets based on a specified level of change (rather than mean change) and for the overall sample rather than subgroups (with the exception of social-emotional development and approaches to learning). In addition, targets were based on children's absolute progress over the year (IRT or W scores) rather than relative to a national sample of same-age peers (standard scores) and specify the percentage of

---

<sup>3</sup> For the Leiter-R, we calculated the percentage of children scoring in the expected range; children scoring in this range are unlikely to be experiencing difficulties with social-emotional development and approaches to learning. In the national sample, 84 percent of children score in this range.

children meeting two different levels of change.<sup>4</sup> LAUP is expected to meet targets at both levels. To inform the discussion (particularly for selecting the appropriate magnitude for targets), the team examined the distribution of the scores for each of the measures in a prior round of UPCOS child assessments. Table ES.2 outlines the targets associated with each measure as well as actual performance during the 2011–2012 year.

Level 1 and Level 2 targets were attained in language, fine motor/literacy, mathematics, and social-emotional development and approaches to learning. For literacy as measured by the RLN task, the percentage of children who made the required level of progress was less than the targeted percentage. For each measure, we also examined progress toward the target for subgroups defined by the quartile in which children scored in the fall (this analysis is specific to each measure). For the measures of language, fine motor/literacy (as measured by the WJ-III Spelling and the RLN), and mathematics (assessed in English), children who began the year with scores in the bottom or middle quartiles were more likely to have change scores in line with what was specified in the target for that measure than children who began the year with scores in the top quartile for that measure. Among children who took the mathematics assessment in Spanish in both the fall and the spring, all children scoring in the top quartile in the fall met the mathematics target; for children in the lower quartiles in the fall, 86 to 89 percent met the target.

We examined average change scores among the group of children whose scores exceeded Level 1 targets and the group whose scores did not exceed Level 1 targets.<sup>5</sup> Standard scores (scores relative to a national sample of peers that are adjusted for child age) were available for four of the measures: EOWPVT/SBE, WJ-III Spelling, WJ-III Applied Problems, and WM-III Problemas Aplicados. For children whose scores met or exceeded targets, on average, scores increased relative to a national sample of same-age peers. Among children whose scores fell below targets, on average, scores declined 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 letters, while children whose scores did not meet targets learned an average of one letter. Thus, the targets appear to be successfully differentiating the children who make meaningful gains during the year and the group of children who need additional support.

## Implications

Considering both average progress over the course of the 2011–2012 program year and performance relative to the selected targets highlights a number of strengths but also room for growth. In terms of strengths, children not only made progress relative to a national sample of peers but also ended the year performing at or above the national average in aspects of expressive bilingual language skills;<sup>6</sup> fine motor and literacy skills, particularly early writing; and social-emotional development and approaches to learning as rated by the independent assessors administering the

---

<sup>4</sup> 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 particular 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).

<sup>5</sup> The Leiter-R was excluded from this analysis. Because scores are truncated, we cannot assess the full range of progress among children whose scores met targets.

<sup>6</sup> Because the SBE norms were based on a sample that, on average, has more limited educational attainment, scores calculated relative to bilingual norms are higher than scores calculated relative to English norms.

**Table ES.2. Summary of Child Progress in 2011–2012 Program Year in a Representative Sample of LAUP Programs and Children Relative to the Performance-Based Contract Targets**

Domain: Measure	Level 1	Level 2
<b>Language: EOWPVT/SBE (N = 273)</b>		
Target	70% gain 2 points or more	45% gain 5 points or more
Actual	72% gained 2 points or more	51% gained 5 points or more
Target reached?	Yes	Yes
<b>Literacy: Rapid Letter Naming (N = 254)</b>		
Target	70% gain 7 points or more	45% gain 13 points or more
Actual	51% gained 7 points or more	20% gained 13 points or more
Target reached?	No	No
<b>Fine Motor/Literacy: WJ-III Spelling (N = 189)</b>		
Target	70% gain 7 points or more	45% gain 17 points or more
Actual	77% gained 7 points or more	51% gained 17 point or more
Target reached?	Yes	Yes
<b>Mathematics: WJ-III Applied Problems (N = 189)</b>		
Target	70% gain 4 points or more	45% gain 13 points or more
Actual	76% gained 4 points or more	49% gained 13 points or more
Target reached?	Yes	Yes
<b>Mathematics: WM-III Problemas Aplicados (N = 73)</b>		
Target	70% gain 4 points or more	45% gain 13 points or more
Actual	93% gained 4 points or more	72% gained 13 points or more
Target reached?	Yes	Yes
<b>Social-Emotional and Approaches to Learning: Leiter-R (N = 188 for Level 1 and 117 for Level 2)</b>		
Target	85% score in expected range in spring for English only and English primarily groups	75% score in expected range in spring for Spanish only or primarily and other language only and primarily groups
Attention		
Actual	98%	100%
Target reached?	Yes	Yes
Activity Level		
Actual	99%	100%
Target reached?	Yes	Yes
Sociability		
Actual	98%	98%
Target reached?	Yes	Yes

Source: UPCOS-5 Fall 2011 and Spring 2012 Direct Child Assessments.

Notes: Targets for all measures except the Leiter-R are based on W or IRT scores. Leiter-R targets refer to the proportion of children scoring in the expected range, which is based on standardized scaled scores.

All analyses are weighted to represent children attending LAUP in the 2011–2012 program year.

direct child assessments. For children in LAUP who took the English assessment, mathematics abilities were similar to those of a national sample of English-speaking peers, and these children maintained their progress relative to peers. In these four areas, progress exceeded the targets set for the 2011–2012 year.

The percentage of children who made the required level of progress was less than the targeted percentage in the area of literacy as measured by the RLN task. An analysis of overall progress (without attention to whether targets were met) indicates that, on average, children did in fact make statistically significant progress in this area. The target for this measure was based on data from

UPCOS-2, which also included a representative sample of children in LAUP centers. In UPCOS-2, children progressed from knowing an average of 10 letters in the fall to about 18 in the spring. In UPCOS-5, children entered program knowing more letters in the fall (an average of 13 letters) and knew an average of 18 in the spring. Note also that this level of progress was on par with that found in other large-scale preschool studies. For example, the Multi-State Study of Pre-Kindergarten and the Study of State-Wide Early Education Programs (SWEEP; Early et al. 2005)<sup>7</sup> found that children progressed from knowing 9.1 letters in the fall to 14.4 in the spring, on average.<sup>8</sup> In each subsequent phase of UPCOS, children have entered LAUP with stronger letter-naming skills than what was found in the previous cohort. UPCOS-5 demonstrated the greatest increase in fall scores, as indicated by both the mean and the median<sup>9</sup>. This has implications for the change we are likely to detect. First, that children are entering with higher skills indicates that it will be harder to make changes of the magnitude identified in the 2011–2012 targets based on UPCOS-2 data. Second, unlike the other measures of language, literacy, and mathematics in the battery, the RLN Task has a ceiling; children cannot progress beyond a certain point because the RLN Task presents children with only 30 letters (a combination of upper- and lowercase). An examination of the distribution of RLN scores indicates that approximately 5 percent of children named all 30 letters presented in fall 2011. Prior knowledge of letter names also differed by language group. Children in the English only group entered with the highest mean score, correctly naming an average of about 17 letters in the fall, whereas children in the Spanish only and Spanish primarily groups knew about half as many in the fall. Given these circumstances, LAUP and First 5 LA should consider revisiting this target for subsequent years. Recent research (Piasta et al. 2012) provides some guidance about the number of letters associated with later success in school. An alternative way to set targets might be to set an absolute number of letters expected and target the percentage of children in each language group expected to name that many letters.

Children assessed in Spanish at both time points continued to perform one standard deviation below the national mean in mathematics in the spring even though the level of progress was sufficient to meet the target. This suggests a strong need for increased attention to helping Spanish-speaking children develop number sense and reasoning skills.

Average performance indicates there is still room to grow for language as well. Although language abilities relative to a national sample of English or bilingual peers improved across the year (with scores surpassing the national bilingual sample by the spring<sup>10</sup>), scores relative to a national sample of English-speaking peers were still below the national mean in the spring. Although the percentage of preschool children in LAUP learning two or more languages is higher than the percentage found in national samples, continued attention to language development is important

---

<sup>7</sup> The Multi-State Study includes a random sample of state-funded prekindergarten programs in four states (Georgia, Illinois, Kentucky, and Ohio) and two large regions of two other states (Los Angeles and the Central Valley in California and New York City and Albany in New York); it began in 2001. The SWEEP study includes a random sample of state-funded prekindergarten programs in four states (Massachusetts, New Jersey, Washington, and Wisconsin) and a large region in Texas; it began in 2003. These studies focused on classrooms for 4-year-olds.

<sup>8</sup> The Multi-State Study/SWEEP version of the letter naming tasks included only 26 letters on each form, compared to 30 in the RLN task.

<sup>9</sup> The median provides a stronger indicator of the change across cohorts. The fall median number of letters in UPCOS-2 was 3, while in UPCOS-5 the fall median was 9.

<sup>10</sup> The national English sample is a more advantaged sample than the national bilingual sample. The bilingual sample includes a large percentage of children from families with low parental education (less than a high school diploma).

because language plays such an important role in academic learning and success. Knowledge of vocabulary and concepts reflects knowledge of the world that children draw on as they read and comprehend text and in other educational opportunities.

## **Conclusion**

Overall, progress in LAUP during UPCOS-5 points to some clear successes, but also room for growth. LAUP met the targets in four of the five domains identified in the performance-based contract: language, fine motor/literacy, mathematics, and social-emotional development and approaches to learning. The area in which the target was not met was literacy as measured by the RLN. However, the RLN task has a ceiling and the distribution of entering scores differs greatly across language groups. Each year children have entered LAUP with stronger letter-naming skills than in the previous cohorts. Recent research provides new guidance around rapid letter naming as an indicator of later success in first grade reading. These factors suggest that the RLN targets should be revisited for the 2012–2013 program year. In addition, it appears that children may be particularly well served if teachers and families provide additional support in the area of mathematics for ELL children because children assessed in Spanish remained one standard deviation below the mean for the national sample. Looking ahead, LAUP and First 5 LA can use the information presented here to determine priorities and consider whether and how targets need to be updated to align with those priorities.



## **I. INTRODUCTION: SETTING TARGETS FOR CHILD PROGRESS**

A central objective of LAUP is to support early learning and development. To meet this objective, First 5 LA and LAUP 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 First 5 LA and LAUP sought to track, to identify appropriate measures, and to set targets for child progress in the 2011–2012 year. During the 2010–2011 program year, as part of Phase 4 of the Universal Preschool Child Outcomes Study (UPCOS-4), Mathematica conducted direct child assessments to test the validity of the targets set for the 2011–2012 program year. Based on the findings, which suggest that the targets successfully differentiated children who made substantial gains during the year and the group of children who needed additional support, LAUP and First 5 LA agreed to include the child progress targets in First 5 LA’s performance-based contract with LAUP for the 2011–2012 program year. During the 2011–2012 program year, as part of Phase 5 of the Universal Preschool Child Outcomes Study (UPCOS-5), Mathematica conducted direct child assessments to determine whether the agreed upon targets were met.

In this chapter, we begin by presenting an overview of the rationale for the domains of development selected for assessment and the measures identified for this purpose. We then outline the 2011–2012 child progress targets and evidence of validity for those targets. In Chapter II, we describe the study sample and methods. Chapter III describes child progress during the 2011–2012 program year for the overall sample and by language group. Finally, Chapter IV reports on children’s progress relative to the targets for the 2011–2012 program year.

### **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 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 available<sup>11</sup>

---

<sup>11</sup> 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 in LA County in general.

- 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 that are important for all children in ways that are not specific to a particular curriculum
- National norms available for comparisons to national samples

Table I.1 summarizes the child outcome measures agreed upon by First 5 LA and LAUP and notes the domains of development that each addresses.<sup>12</sup> Additional detail regarding each of the measures is provided in Chapter II, which describes study methods. The selected measures were all used in prior phases of UPCOS and together address multiple dimensions of school readiness.

**Table I.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) and Expressive One-Word Picture Vocabulary Test – Spanish Bilingual Edition (EOWPVT-SBE) (Brownell 2000)	Language development: vocabulary	Standardized measure; children required to name pictures; conceptually scored <sup>a</sup>
Rapid Letter Naming (RLN; Atkins-Burnett et al. 2007)	Literacy development	Child must name upper- and lowercase letters; conceptually scored <sup>a</sup>
Woodcock-Johnson III (WJ-III) and Woodcock-Muñoz Bateria III (WM-III) Applied Problems (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 number; conceptually scored <sup>a</sup>
Woodcock-Johnson III (WJ-III) Spelling and Woodcock-Muñoz III (WM-III) Ortografía <sup>13</sup> (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

<sup>a</sup>In a measure that is conceptually scored, respondents receive credit for a correct response regardless of the language of the response.

<sup>12</sup> In addition to the measures shown in Table I.1, children completed the English and Spanish Preschool Language Assessment Survey (*pre*LAS; Duncan and DeVila 2002) as both a language screener and warm-up exercise. The team determined that these measures should not be used for setting targets.

<sup>13</sup> Most classrooms taught the English names for letters, so we deviated slightly from standard administration by allowing assessors to provide the names of letters in both Spanish and English on the Ortografía when asking children to write orally presented letters.

## B. Target-Setting Procedures

The child progress targets were set in a collaborative process between First 5 LA and LAUP that was facilitated by Mathematica. Together, the team considered 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. To inform the discussion, the team examined the distribution of the scores for each of the measures in prior rounds of UPCOS.<sup>14</sup> 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.<sup>15</sup>

### 1. Appropriate Metric (Type of Score)

Each of the selected assessments offers multiple types of scores. For most of the measures, the group agreed to 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 do not adjust for age. These scores are referred to as IRT scores or *W* scores, depending on the measure. 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 of them named an additional consonant whereas the second child moved from being able to name letters to being able to read a word. The children would show the same level of progress if you look at a raw score, but the child who read the word would show a greater gain in an IRT score than the child who named one more consonant, reflecting the developmental change represented by the items. The progress noted on a standard score would be dependent on the child's age group. Sometimes, a standard score for a child could indicate progress on one day, but if the child moves to a different age category a day or week later, the same score may not reflect change in a standard score. 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.<sup>16</sup> We can also look at change across age groups using IRT scores because the scores do not adjust for child age.

---

<sup>14</sup> For all measures except Rapid Letter Naming (RLN), targets were initially set 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 was to be based on data collected by independent assessors, the team initially set the RLN target based 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. (2010). For detail regarding UPCOS-2, see Love et al. (2009). All targets for 2011-2012 were revisited and finalized based on UPCOS-4 data (collected during 2010-2011); sampling and administration procedures were the same in UPCOS-4 and 5 for all measures.

<sup>15</sup> Additional information on the target-setting process as well as the rationale for decisions regarding the targets is available in the memo "Developing Child Progress Targets for the Performance-Based Contract—Revised Memo" (Xue et al. 2011).

<sup>16</sup> Although the age range of UPCOS children is limited (46 months to 60 months in the fall), the same raw score can mean different standard scores if a child falls in the next age grouping.

## 2. Specified Change versus Mean Change

The selected targets are based 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 make large changes even if many children do not change. Because First 5 LA and LAUP are interested in all children benefiting from their time in their LAUP program, the targets are based on percentages of children who make a specified change, with one exception. Because Leiter-R scores are truncated (that is, children cannot score above a certain level because the measure is intended to identify children who are at risk for developmental problems in this area rather than to assess child progress), the target refers to the percentage of children who score in a particular range in the spring rather than on 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). For measures on which the distributions of change scores did not drastically differ across the language groups, the team agreed to set an overall target; these include measures of language (EOWPVT/EOWPVT-SBE), literacy (RLN), fine motor and literacy (WJ-III Spelling), and cognitive ability/mathematics (WJ-III/WM-III Applied Problems).<sup>17</sup> For the social-emotional measure—the Leiter-R—the distribution of scores did differ across language groups. Thus, the team agreed to set separate targets for the groups that speak English only or primarily and those who speak Spanish or another language only or primarily (the latter groups are considered English language learners [ELLs]).

## 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 score (89) did not differ. 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) was not consistent with the letter names they had learned. Because LAUP programs do not typically teach children the names of letters in Spanish, we would not expect to see growth in this area. Therefore, the group agreed not to set targets based on the WM-III Ortografía. We fielded different procedures in spring 2010 (using both the English and Spanish names for the letters) and followed the same procedure for both fall 2011 and spring 2012. We describe children's progress on this measure in this report although no targets were set.

## 5. Appropriate Magnitude for Targets

Targets are based on the distribution of children's scores during prior phases of UPCOS—UPCOS-2 (the 2007–2008 program year) for the RLN and UPCOS-3 (the 2009–2010 program year)

---

<sup>17</sup> Note that for some of these measures, the distributions of the change scores differed in the other language group. However, the sample size for the other language group is small (around 20), and its inclusion in an overall target is unlikely to bias the results. In addition, the team agreed that because these children are part of LAUP, their performance should be addressed by the targets as well, despite the difference in the distribution of scores.

for all other measures. 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 that differed markedly from that seen in previous program years. As a reminder, children are in LAUP for one year, and teachers can change from year to year. Thus, the group agreed that targets for the 2011–2012 program year should specify a level of growth similar to that documented in prior phases of UPCOS. Examination of the item difficulties where the mean change occurred and the substantive change in the items (across the mean change) supported this decision.

The group also agreed that targets should be set at two levels. 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 particular 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.

### **C. Targets Established for the 2011–2012 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 I.2 summarizes the final Level 1 and Level 2 targets and summarizes the key decisions outlined above that First 5 LA and LAUP made in setting the targets.

### **D. Evidence of Validity of the Targets for the 2011–2012 Program Year**

In the 2010–2011 program year, Mathematica collected direct child assessment data to assess the validity of the targets. We examined average change scores for the group of children whose scores exceeded Level 1 targets and the group whose scores did not exceed Level 1 targets.<sup>18</sup> Standard scores (scores relative to a national sample peers that are adjusted for child age) are available for four of the measures: EOWPVT/EOWPVT-SBE, WJ-III Spelling, WJ-III Applied Problems, and WM-III Problemas Aplicados. For children whose scores reached or exceeded targets, on average, scores increased relative to a national sample of same-age peers. Among children whose scores fell below targets, on average, scores declined 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 letters, whereas children whose scores did not meet targets learned an average of one letter. Across the measures, the specified percentage of children in UPCOS-4 met the targets for the children in the lowest and middle quartiles with the exception of mathematics in English. As anticipated when setting the targets, the children in the top quartile made smaller gains than the other children but on average were performing at or above the national mean in both fall and spring, with the exception of Spanish mathematics (WM-III Problemas Aplicados). The evidence suggests that the targeted level of change is very meaningful and reasonable.

---

<sup>18</sup> The Leiter-R was excluded from this analysis. Because scores are truncated, we cannot assess the full range of progress among children whose scores met targets.

**Table I.2. Final Targets for Child Progress in LAUP Programs During the 2011–2012 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 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 do not adjust 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.
Literacy Development	Rapid Letter Naming (RLN)	70% children gain 7 points or more on the RLN IRT score 45% children gain 13 points or more on the RLN IRT score	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 a risk for developmental difficulties in the areas assessed by this measure).
Fine Motor and Literacy Development	WJ-III Spelling	70% gain 7 points or more on the WJ-III Spelling W score 45% gain 17 points or more on the WJ-III Spelling W score	<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.
Mathematics Development	WJ-III/WM-III Applied Problems	70% gain 4 points or more on the WJ-III/WM-III Applied Problems W score 45% gain 13 points or more on the WJ-III/WM-III Applied Problems W score	Because Leiter-R scores are truncated (children cannot score above a certain level) for this measure, 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. <i>Overall versus Language Subgroup Targets</i> Set overall targets for all measures on which the distributions of change scores from a prior phase of UPCOS did not drastically differ across the language groups. For the Leiter-R set separate targets by language group because the distribution of scores differed across language groups in a prior phase of UPCOS.
Social-Emotional Development and Approaches to Learning	Leiter-R Examiner Rating Scales	85% score in the expected range in the spring for English only and English primarily 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 and WM-III Applied Problems 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).
	Attention		
	Activity Level	75% score in the expected range in the spring for Spanish only, Spanish primarily, and other language only and primarily groups	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). For the Leiter-R, the two targets are for the different language groups. LAUP expects to meet the targets at both levels of difficulty/for both language groups.
	Sociability		

## II. STUDY METHODS

This chapter describes the study methods. We begin with a description of the sample for UPCOS-5 child assessments, followed by an overview of the child outcome measures and procedures for the administration of the child assessment battery. The last section of this chapter presents how we scored the child outcome measures.

### A. Sample

UPCOS-5 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. There were 60 eligible programs and 712 children in the original sample, and the goal was to recruit 40 programs. In the fall, the final sample included 39 programs and 660 children. The response rates were 65<sup>19</sup> and 93 percent for programs and children, respectively. In the spring, 597 children (90 percent of the total assessed in the fall) were assessed in the 39 programs. Looking across the fall and spring, 586 children were assessed at both time points (82 percent of the selected child-level sample). Table II.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.

**Table II.1. Number of Children Assessed in Fall 2011 and Spring 2012, by Language Group (N = 586)**

Language Group	Percentage
English only	35.80
English primarily	28.67
Spanish only	9.36
Spanish primarily	20.85
Other language only or primarily	5.32

Note: Analyses are weighted to represent children attending LAUP in the 2011-2012 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 oral 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*

---

<sup>19</sup> We compared characteristics of programs that did and did not participate in the study based on LAUP administrative data including: the percent minority, funding sources other than LAUP (California Department of Education, state preschool, Head Start), and program capacity. Differences were found only in program capacity; programs that participated had higher total capacity than programs that did not participate.

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, though the English version includes a few additional items). The SBE version allows probes to be offered in both English and Spanish while the English version probes only in English (“What is this?”, “What are these called?”). We used conceptual scoring for both versions; that is, children are given credit for correct responses in either Spanish or English. For the English version, standard scores were calculated relative to a 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 bilingual (Spanish and English). Note that the norming samples for the two versions were 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 a cultural and linguistic background similar to LAUP children. Children from California are overrepresented slightly in the standardization sample for the bilingual norms, and more than 50 percent of the children in the standardization sample come from families with a mother whose educational attainment is less than a high school diploma. The English norms were based on a standardization sample that represents a greater range of maternal education (approximately 10 percent of the children in the standardization sample for the English norms come from families with a mother who has less than a high school diploma) and cultural backgrounds than the standardization sample for bilingual norms. Thus, the linguistically diverse children in LAUP were more similar to the standardization sample for SBE norms than for English norms with regard to cultural and linguistic background. Because the SBE norms were based on a sample that, on average, has more limited maternal educational attainment, scores calculated relative to bilingual norms are likely to 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. The fall form has a slightly higher percentage of uppercase letters, which are typically easier for young children to identify.

**Woodcock-Johnson III Battery (WJ-III) and Woodcock-Muñoz Bateria III (WM-III) (Woodcock et al. 2001/2007; Woodcock et al. 2004/2007).** The WJ/WM assessments have been widely used in early childhood studies including Head Start’s Family and Child Experiences Survey (FACES), the Tulsa preschool study (Gormley et al. 2005), and the Study of State-Wide Early Education Programs (SWEEP; Early et al. 2005).

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, addition, and subtraction computations. 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.<sup>20</sup>

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 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<sup>21</sup> in both fall and spring in UPCOS-5. 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. Therefore, when children reached the items asking them to write a letter, we presented the names of letters in both Spanish and English.

In UPCOS-5, a greater percentage of the children who were assessed in Spanish in the fall (52 percent) switched to English in the spring<sup>22</sup> than in previous phases of UPCOS. The Spanish and English versions have normative samples that differ beyond the language, so it is not possible to compare standard scores on the Spanish version with standard scores on the English version and examine their progress. Therefore, when children changed to English in the spring, we report their mean score in the spring separately from those who were assessed in Spanish in both the fall and spring.

**Leiter Examiner Rating Scales—Revised (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, for children who were not routed out of the assessment entirely because they spoke neither English nor Spanish). The ratings were based on the unique assessment situation. Once the assessment is complete, the assessor provides ratings of the child's behavior on items that tap Attention, Activity Level, and Sociability. Because this scale is designed as a screening measure, the scores are truncated; that is, skill levels that are above average are not measured because the purpose is to identify children who are having difficulty in these areas and are thus at risk for developmental problems in the areas tapped by the measure.

### C. Procedures

To minimize the burden on children, we designed two versions of the child assessment battery, each approximately 15 minutes in length. 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:

---

<sup>20</sup> Accepting correct answers in English or Spanish is consistent with the norming procedures for the WM-III.

<sup>21</sup> In standard administration of the WM-III Ortografía, the names of the letters are presented only in Spanish.

<sup>22</sup> For Version 1 (see next section), 60 percent of the children switched from Spanish to English assessment in the spring. In Version 2, all measures are conceptually-scored (answers are accepted in either English or Spanish) and so the change in performance on the language screener was not an issue.

**Version 1:**

1. English *preLAS* (Simon Says)
2. English/Spanish *preLAS* (Art Show)
3. WJ-III and WM-III Applied Problems
4. WJ-III Spelling and WM-III Ortografía

**Version 2:**

1. English *preLAS* (Simon Says)
2. English/Spanish *preLAS* (Art Show)
3. Expressive One-Word Picture Vocabulary Test (EOWPVT/EOWPVT-SBE)
4. Rapid Letter Naming (RLN)
5. Leiter Examiner Rating Scale–Revised (Leiter-R)

Note that all children received the English *preLAS* as a warm-up for the assessment and for routing to the language of assessment. Note also that any 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.<sup>23</sup> All children who were from Spanish-speaking homes also received the *preLAS* 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.

**D. 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.<sup>24</sup> Raw scores simply summarize how many items a child answered correctly without accounting for factors such as the difficulty of the items. Raw scores are provided for the English and Spanish *PreLAS* and RLN.<sup>25</sup> For the English and Spanish *preLAS*, each subscale has a maximum possible score of 10 points. For the Spanish *preLAS* we only used one subscale. Because we used two subscales for the English *preLAS*, that assessment has a maximum total score of 20. The RLN has a possible raw score of 30.

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. 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

---

<sup>23</sup> Those who were routed out of the assessment received only the RLN Task.

<sup>24</sup> 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.

<sup>25</sup> 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.

of these assessments, the mean for the norming sample is 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/SBE and RLN using the item difficulties from the UPCOS-3 and UPCOS-2 samples, respectively.

For the Leiter-R, raw scores were converted to standardized scaled scores (using a table provided in the manual), which are truncated at 10 points.<sup>26</sup> The scaled scores were then used to calculate the proportion of children scoring in the expected range (also referred to as the acceptable range); children scoring in the expected range are unlikely to be experiencing difficulties with social-emotional development and approaches to learning.

---

<sup>26</sup> The scaled scores are truncated at 10 because the rating scales are designed to determine whether behaviors fall within a problematic or unexpected range.



### III. CHILD PROGRESS IN THE 2011–2012 PROGRAM YEAR

In this chapter we present child progress in LAUP during the 2011–2012 year. We begin by describing procedures used in the data analysis. We then discuss performance in the overall sample and by language subgroup.

#### A. Approach to Analysis

To look at how children enrolled in LAUP programs in fall 2011 developed during the preschool year, we examined children's scores for those children who were assessed in both fall and spring. On average, 5.6 months passed between the two assessments (range of 3.4 to 6.4 months).<sup>27</sup>

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 for each type of score available (raw, standard, and/or IRT or W score). We tested whether progress between the fall and spring was significant using *t*-tests for the overall sample and within subgroups. In addition, we conducted analysis of variance (ANOVA, *F*-tests) to determine whether progress differed across subgroups (shown in Tables III.4 and III.6). We report *p* levels at the .05, .01, and .001 levels and consider  $p < .05$  to indicate statistical significance.<sup>28</sup> For any test resulting in  $p > .05$ , we consider this as evidence of no change. 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 compared to their same-age peers, it is possible to identify significant change for raw and IRT/W scores but not for standard scores for the same measure. We conducted these analyses for the full sample of children (Tables III.1 and III.2) and for subgroups defined by children's home language (Tables III.3 through III.6). No additional demographic data were collected. 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 language subgroups were likely to have large standard errors relative to those for the whole sample and, as a consequence, larger confidence intervals.

We also present figures illustrating children's progress over the course of the year in the overall sample for all measures with the exception of the warm-up assessment (the *pre*LAS). Measures were grouped in figures according to the type and metric of the score. Figure III.1 presents EOWPVT IRT<sup>29</sup> scores. Figure III.2 presents raw and IRT scores for RLN. Figure III.3 presents W scores for WJ-III Spelling, WM-III Ortografía, WJ-III Applied Problems, and WM-III Problemas Aplicados.<sup>30</sup> Figure III.4 presents standard scores for the EOWPVT English Edition and SBE, WJ-III Spelling, WM-III Ortografía, WJ-III Applied Problems, and WM-III Problemas Aplicados because all of

---

<sup>27</sup> The majority (97 percent) of the children had 5.0 to 6.0 months between assessments. Only five children had fewer than 5.0 months between assessments.

<sup>28</sup> *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.

<sup>29</sup> EOWPVT IRT scores were estimated using both the EOWPVT and the EOWPVT-SBE. The EOWPVT has a few additional items which were treated as missing for the children who took the SBE version.

<sup>30</sup> Children who changed language of administration from fall to spring are not represented in the figures. We report the spring scores separately.

these measures have a mean of 100 and standard deviation of 15. Finally, Figure III.5 presents the percentage of children scoring in the expected range for the Leiter-R subscales. In all figures, we present children's performance in the fall and the change that occurred by the spring.

All analyses discussed in this chapter were at the child level and were 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 correctly estimated and, thus, that any statistical tests based on those standard errors are accurate.

## B. Progress Across All Children

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

The first section of Table III.1 shows average scores in fall 2011 and spring 2012, as well as change scores, for the full sample of children for the *preLAS*. Children's skills on the portions of the English *preLAS* measuring both receptive (listening comprehension) and expressive (oral vocabulary) skills increased significantly from fall to spring, from 6.3 to 8.0 for Simon Says and from 7.4 to 8.4 for Art Show. The total score across the two subscales also increased significantly. However, children did not make significant progress in Spanish expressive vocabulary skills; the fall and spring mean scores were not statistically different.

### 2. Language: EOWPVT

The second section of Table III.1 presents IRT scores (Figure III.1) as well as standard scores for the English and SBE versions of the EOWPVT (Figure III.4). As a reminder, the EOWPVT was conceptually scored. We analyzed the measure using both the English and the SBE norms.

Absolute progress in expressive skills based on the IRT scores is statistically significant. Mean scores increased from 44.5 in the fall to 50.2 in the spring, an increase of 5.7 points.

The mean English standard scores indicated that, on average, LAUP children lagged behind the national norms in both the fall (mean English standard score [SS] = 89.2) and spring (mean English SS = 91.4) in expressive vocabulary concepts. However, the mean score on expressive language skills based on the English norms was significantly higher in the spring than in the fall, suggesting the children in LAUP made some progress toward catching up with their national peers. The mean SBE standard scores indicated that, on average, children in LAUP performed better than their same-age Spanish-speaking peers nationally in expressive vocabulary at preschool entry (mean SBE SS = 105.5) and more so in the spring (mean SBE SS = 110.3). They also made statistically significant progress between the fall and spring relative to a national sample of bilingual peers.

**Table III.1. Means and Standard Errors for UPCOS-5 Language, Literacy, and Math Measures: Fall 2011 and Spring 2012**

Outcome	N	Fall 2011		Spring 2012		Fall-Spring Change	
		Mean	Standard Error	Mean	Standard Error	Mean	Standard Error
<b>preLAS English</b>							
Total Language Screener Score	586	13.63	0.43	16.44	0.31	2.81***	0.24
Simon Says	586	6.28	0.24	8.01	0.17	1.72***	0.17
Art Show	586	7.35	0.22	8.43	0.16	1.09***	0.10
<b>preLAS Spanish</b>							
Exposición de Arte	172	5.36	0.29	5.57	0.40	0.21	0.29
<b>Language: Expressive One-Word Picture Vocabulary Test</b>							
IRT Score	273	44.53	1.30	50.24	1.03	5.71***	0.53
English Edition Standard Score <sup>a</sup>	164	89.16	1.93	91.36	1.57	2.20**	0.80
Spanish Bilingual Edition Standard Score <sup>a, b</sup>	273	105.45	2.37	110.28	1.92	4.83***	0.99
<b>Literacy: Rapid Letter Naming</b>							
Raw Score	271	12.52	0.92	18.29	0.84	5.78***	0.47
IRT Score	271	22.49	1.14	29.76	0.98	7.26***	0.56
<b>Fine Motor and Literacy: Woodcock-Johnson III Spelling</b>							
W Score	189	381.62	2.21	399.05	2.39	17.43***	1.97
Standard Score <sup>a</sup>	189	104.54	1.26	106.47	1.44	1.92	1.09
<b>Fine Motor and Literacy: Woodcock-Muñoz III Ortografía</b>							
W Score	36	355.41	4.56	377.92	3.92	22.51***	3.99
Standard Score <sup>a</sup>	36	91.95	2.52	96.14	2.38	4.19	2.18
<b>Mathematics: Woodcock-Johnson III Applied Problems</b>							
W Score	189	396.45	1.68	408.94	1.77	12.49***	1.33
Standard Score <sup>a</sup>	189	100.11	1.03	101.66	1.19	1.55	0.76
<b>Mathematics: Woodcock-Muñoz III Problemas Aplicados</b>							
W Score	36	347.77	4.00	375.80	4.43	28.03***	3.91
Standard Score <sup>a</sup>	36	78.03	1.65	85.35	1.95	7.31***	1.73

Source: UPCOS-5 Fall 2011 and Spring 2012 Direct Child Assessments.

Note: All analyses are weighted to represent children attending LAUP in the 2011–2012 program year.

<sup>a</sup>National mean for standard scores is 100 with a standard deviation of 15.

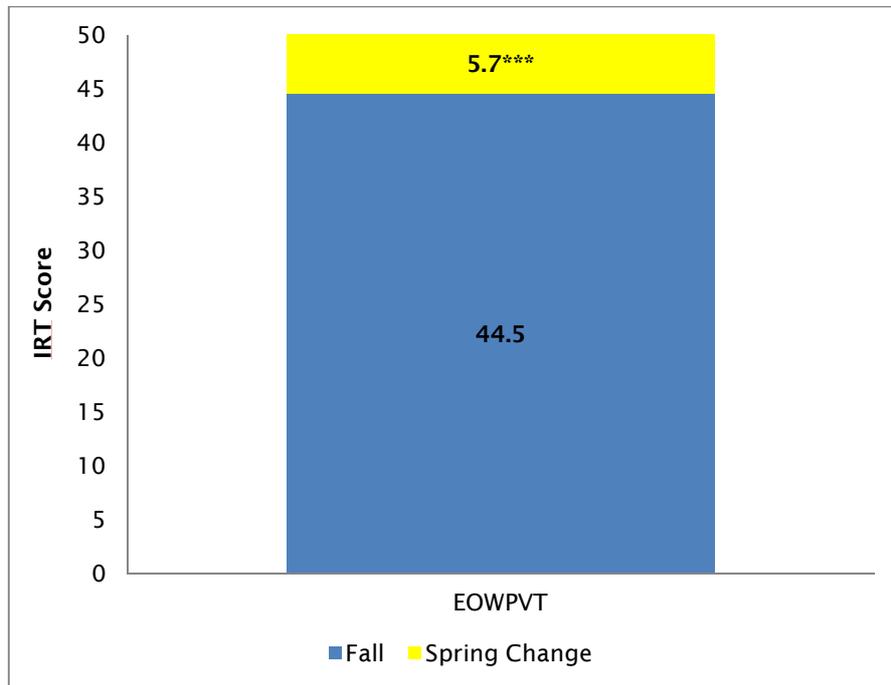
<sup>b</sup>This measure was conceptually scored, and the standard score was generated for all children in the sample.

\* 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.

**Figure III.1. Mean IRT Scores for Language Assessment (EOWPVT): Fall 2011 and Change to Spring 2012**



Source: UPCOS-5 Fall 2011 and Spring 2012 Direct Child Assessments.

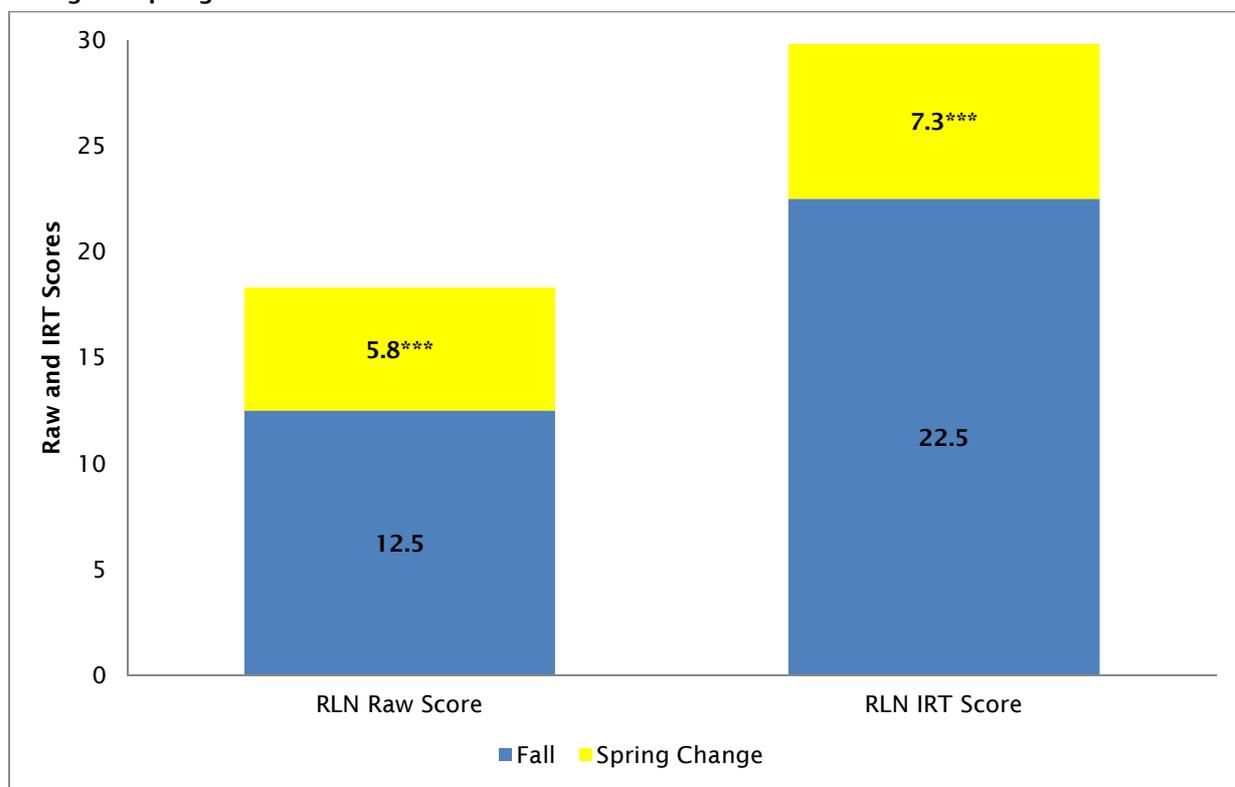
Notes: All analyses are weighted to represent children attending LAUP in the 2011–2012 program year.

Asterisks indicate that the change from fall to spring is statistically significant (\* $p < .05$ , \*\* $p < .01$ , \*\*\* $p < .001$ ).

### 3. Literacy: Rapid Letter Naming

Shifting to the third section of Table III.1 (Figure III.2) we see that children made statistically detectable progress in literacy on the RLN based on raw and IRT scores. Between the fall and spring, children progressed from being able to name 13 letters to being able to name 18, a statistically significant change (mean change = 5.8 letters). Similarly, IRT scores increased from 22.5 in the fall to 29.8 in the spring, also a statistically significant change of 7.3 points.

**Figure III.2. Mean Raw and IRT Scores for Literacy Assessment (Rapid Letter Naming): Fall 2011 and Change to Spring 2012**



Source: UPCOS-5 Fall 2011 and Spring 2012 Direct Child Assessments.

Notes: All analyses are weighted to represent children attending LAUP in the 2011–2012 program year.

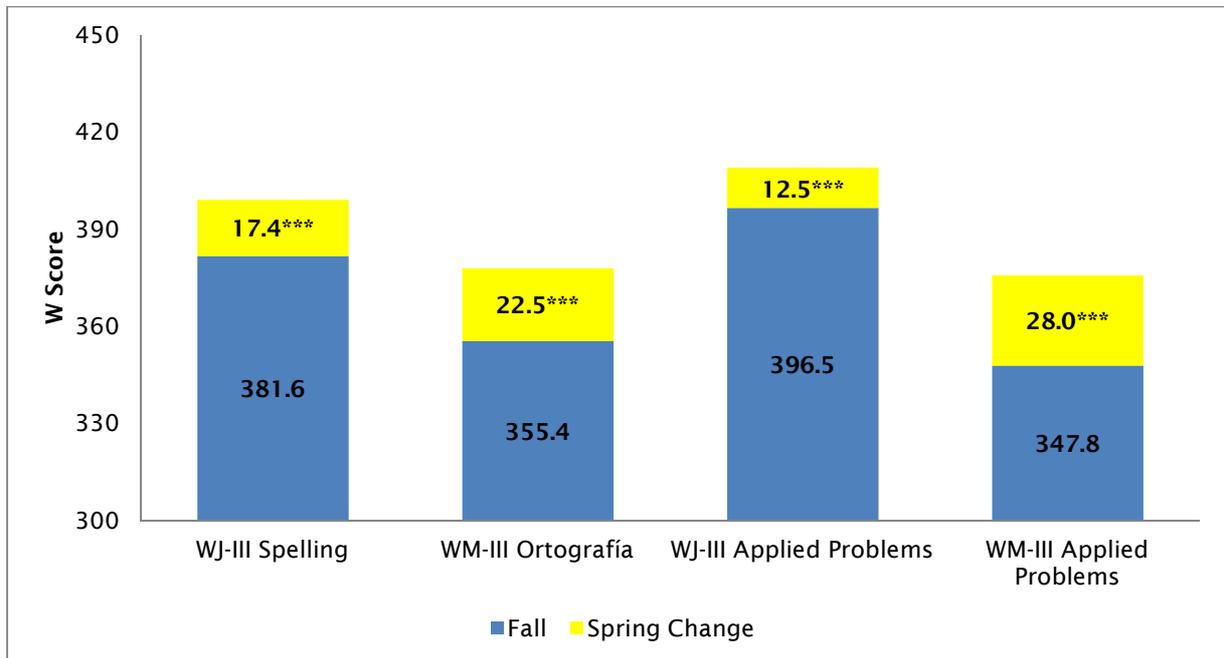
Asterisks indicate that the change from fall to spring is statistically significant (\* $p < .05$ , \*\* $p < .01$ , \*\*\* $p < .001$ ).

The possible range on the raw score is 0-30. The possible range for the IRT score is 0-52.

#### 4. Fine Motor and Literacy: WJ-III Spelling

The fourth section of Table III.1 presents progress in fine motor skills and literacy based on the WJ-III Spelling subtest. Beginning with absolute progress (Figure III.3), W scores showed a statistically significant increase from 381.6 points in the fall to 399.1 in the spring. Standard scores (Figure III.4) show that children scored about one-third of a standard deviation above a national sample of peers in the fall ( $SS = 104.5$ ) as well as in the spring ( $SS = 106.5$ ). Although children made significant progress in this area in absolute terms, their progress as measured by the standard scores is similar to their same-age peers nationally.

**Figure III.3. Mean W Scores for Literacy and Mathematics Assessments: Fall 2011 and Change to Spring 2012**



Source: UPCOS-5 Fall 2011 and Spring 2012 Direct Child Assessments.

Notes: All analyses are weighted to represent children attending LAUP in the 2011–2012 program year.

Asterisks indicate that the change from fall to spring is statistically significant ( $*p < .05$ ,  $**p < .01$ ,  $***p < .001$ ).

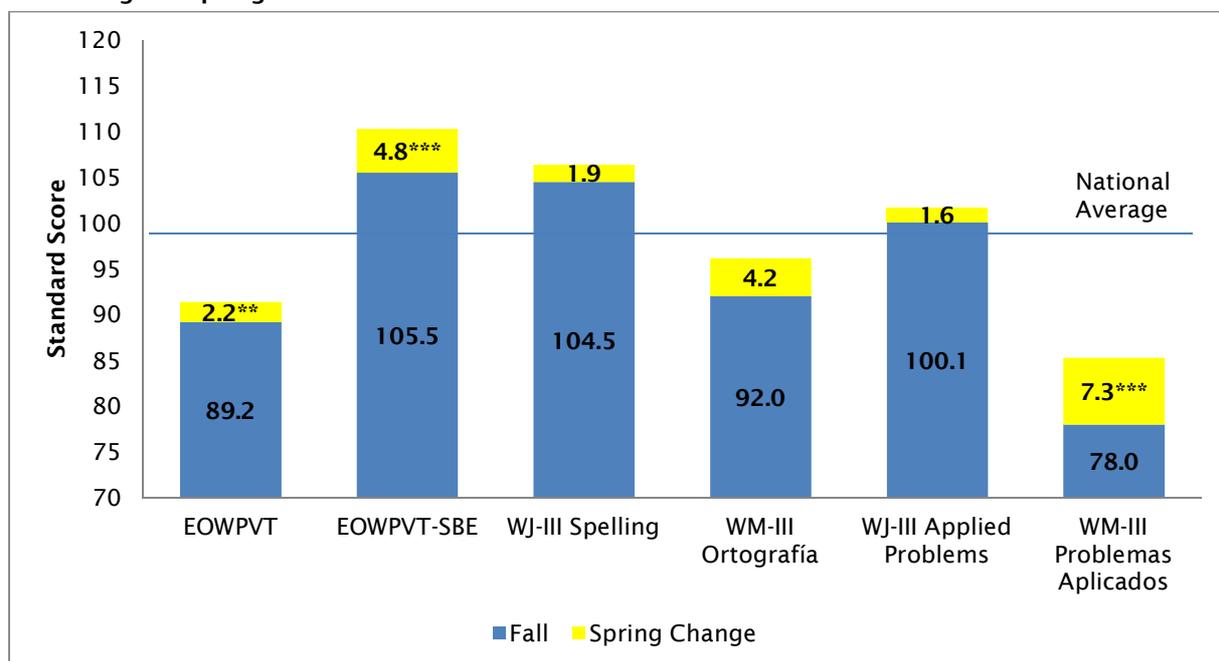
### 5. Fine Motor and Literacy: WM-III Ortografía

Spanish-speaking children who did not pass the English-language screener completed the Spanish version of the WJ-III spelling assessment—the WM-III Ortografía. As shown in the fifth section of Table III.1, like their LAUP peers who took the assessment of fine motor and literacy in English, children completing the WM-III Ortografía made statistically significant progress between the fall and spring in absolute terms, but their progress is not statistically significant relative to a national sample of Spanish-speaking peers. W scores (Figure III.3) increased from 355.4 in the fall to 377.9 in the spring, a gain of 22.5 points. Focusing on the standard scores (Figure III.4), we note that even though these Spanish-speaking children entered in the fall scoring more than one-half of a standard deviation below their national peers ( $SS = 92.0$ ), by the spring LAUP children were scoring at a level about one-fourth of a standard deviation below the national mean ( $SS = 96.1$ ).<sup>31</sup> It is important to note that these scores reflect the progress and status of children who took the assessment in Spanish in both fall and spring.

In UPCOS-5, a greater percentage of the children who were assessed in Spanish in the fall switched to English in the spring than in previous phases of UPCOS. For children who were assessed with WM-III Ortografía (Version 1 of the assessment) in the fall, 60 percent ( $n = 54$ ) switched to WJ-III Spelling in the spring. These children are not represented on the graphs. The

<sup>31</sup> The change is not statistically significant because of the small sample size.

**Figure III.4. Mean Standard Scores for Language, Literacy, and Mathematics Assessments: Fall 2011 and Change to Spring 2012**



Source: UPCOS-5 Fall 2011 and Spring 2012 Direct Child Assessments.

Notes: All analyses are weighted to represent children attending LAUP in the 2011-2012 program year.

Asterisks indicate that the change from fall to spring is statistically significant (\* $p < .05$ , \*\* $p < .01$ , \*\*\* $p < .001$ ).

mean spring score for the group that changed to an English administration in the spring was 102.3 on the WJ-III Spelling. Their mean score on the WM-III Ortografía in the fall was 91.0. On average, they scored similarly in the fall to children who were assessed in Spanish at both time points, but their mean score on the WJ-III Spelling in the spring was lower than those who were assessed in English at both time points.

## 6. Mathematics: WJ-III Applied Problems

The sixth section of Table III.1 summarizes progress in mathematics development as measured by the WJ-III Applied Problems subtest. Children’s W scores (Figure III.3) increased significantly from fall (396.5) to spring (408.9), an increase of 12.5 points. Shifting to the standard scores (Figure III.4), the mean scores are similar between the fall (100.1) and spring (101.7) and are about the same as the mean in the normative sample, indicating children’s progress kept pace with that of same-age peers nationally.

## 7. Mathematics: WM-III Problemas Aplicados

The final section of Table III.1 shows that Spanish-speaking children made significant progress in mathematics in absolute terms and relative to a national sample of same-age peers. W scores (Figure III.3) showed a statistically significant increase between the fall and spring, from 347.8 to 375.8. Standard scores also increased significantly (Figure III.4) from 78.0 in the fall to 85.4 in the spring (however, the means remained one standard deviation or more below the normative mean at

both time points). 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.

In UPCOS-5 a greater percentage of the children who were assessed in Spanish in the fall switched to English in the spring than in previous phases of UPCOS. For children who were assessed with WM-III Problemas Aplicados (Version 1 of the assessment) in the fall, 60 percent (N = 54) switched to WJ-III Applied Problems in the spring. These children are not represented on the graphs. The mean spring score for the group that changed to an English administration in the spring was 93.0 on the WJ-III Applied Problems. Their mean score on the WM-III Problemas Aplicados in the fall was 82.7. On average, they performed better in the fall than children who were assessed in Spanish at both time points, but their mean score on the WJ-III Applied Problems in the spring was lower than those who were assessed in English at both time points.

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

Table III.2 (Figure III.5) summarizes performance as measured by the Leiter-R subtests. As a reminder, Table III.2 shows the percentage of children scoring in the expected range; in the national sample, 84 percent of children score in the expected range. We see that the percentage of children scoring in the expected range for Attention, Activity Level, and Sociability was higher in the spring than in the fall, although the increase was significant only for Activity Level. In both fall and spring, the percentage of children scoring in this range exceeded the national average for all three subtests.

**Table III.2. Percentage of Children Scoring in the Expected Range on the Leiter-R: Fall 2011 and Spring 2012**

Outcome	N	Fall 2011		Spring 2012	
		Percentage	Standard Error	Percentage	Standard Error
<b>Social-Emotional and Approaches to Learning: Leiter Examiner Ratings Scaled Score in the Expected Range<sup>a</sup></b>					
Attention	273	94.77	1.74	98.70	0.78
Activity Level	273	92.05	2.03	99.37**	0.63
Sociability	273	95.23	1.93	98.42	0.82

Source: UPCOS-5 Fall 2011 and Spring 2012 Examiner Rating.

Note: All analyses are weighted to represent children attending LAUP in the 2011–2012 program year.

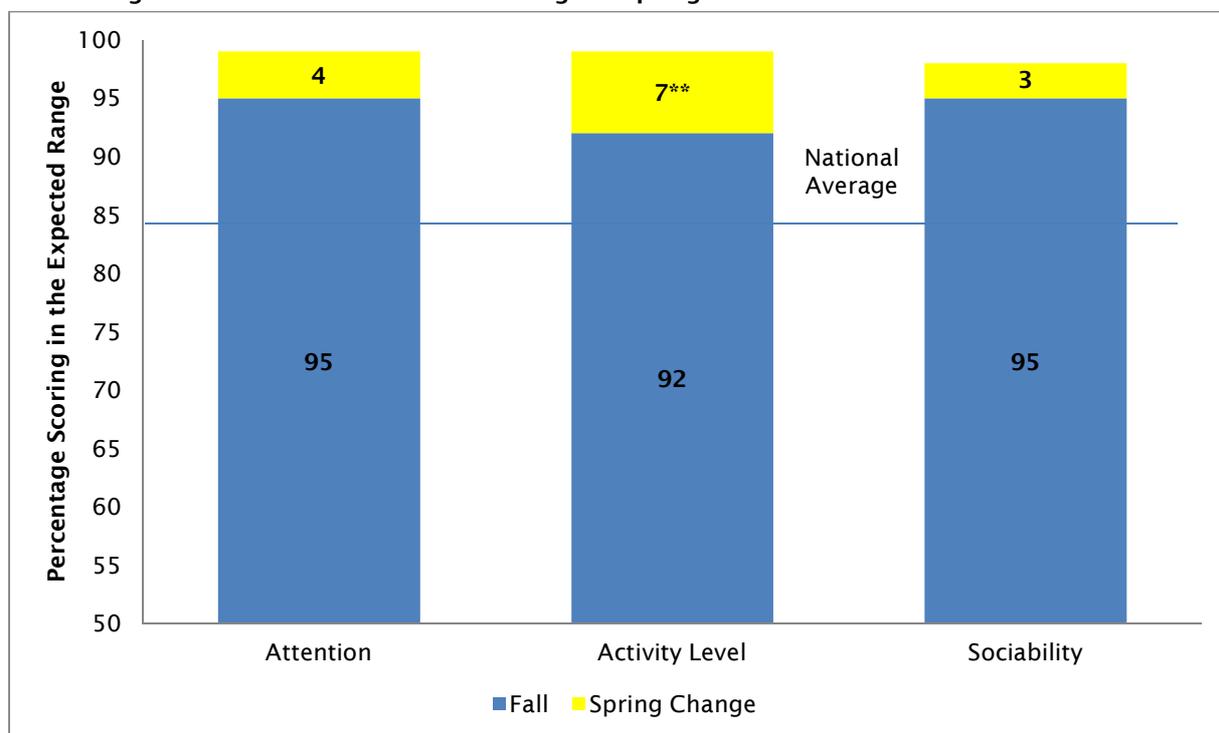
<sup>a</sup>Scores of 7 or greater.

\* 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.

**Figure III.5. Percentage of Children Scoring in Expected Range for Social-Emotional and Approaches to Learning Assessments: Fall 2011 and Change to Spring 2012**



Source: UPCOS-5 Fall 2011 and Spring 2012 Direct Child Assessments.

Notes: All analyses are weighted to represent children attending LAUP in the 2011–2012 program year.

Asterisks indicate that the change from fall to spring is statistically significant ( $*p < .05$ ,  $**p < .01$ ,  $***p < .001$ ).

### C. Progress by Language Group

For this analysis, we sorted children in four language groups based on parent report: (1) English only, (2) English primarily, (3) Spanish only, and (4) Spanish primarily. We considered the latter two groups to be ELLs. As a reminder, we used parent responses to a series of questions about the child’s home language use<sup>32</sup> to sort the children. Children identified as ELLs completed the assessment in English or Spanish depending on their performance on the language screener (English *pre*LAS 2000). Note that for the Spanish only and Spanish primarily groups, sample sizes for the assessments in English fell below 10 children; to avoid less stable or less precise estimation as well as to protect confidentiality, these children’s scores were deleted from the tables and are not discussed

<sup>32</sup> In the fall, 32 percent of children in the English primarily group were assessed in Spanish; 13 percent of children in the Spanish only group and 15 percent of children in the Spanish primarily group were assessed in English. By the spring, 8 percent of children in the English primarily group were assessed in Spanish; 39 percent of children in the Spanish only group and 58 percent of children in the Spanish primarily group were assessed in English. See Appendix A for the percentage of children who switched language of assessment from fall to spring by language group. For children who were assessed in both the fall and the spring, among those assessed English in the fall, 1 percent switched to Spanish in the spring. Among children who were assessed in Spanish in the fall, 52 percent switched to English test in the spring.

in the text. Similarly, small sample sizes led to the exclusion of all scores except for *pre*LAS for children speaking another language only or primarily.<sup>33</sup> Tables III.3 and III.5 present fall and spring scores for each assessment by language group, and Tables III.4 and III.6 present change scores. Statistical tests shown in Table III.3 and III.5 compare fall and spring scores within each language group using *t*-tests, and statistical tests shown in Tables III.4 and III.6 compare change scores across language groups using ANOVA.

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

Not surprisingly, children in the English only and primarily groups began and ended the program year with higher English language skills than did children in the Spanish only and primarily groups as well as children in the other language only or primarily as measured by the total English *pre*LAS and the Simon Says and Art Show subsets (Table III.3). However, all language groups showed statistically significant increases in skills between the fall and the spring. Shifting to Table III.4, 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 = 2.7 and 2.9 and Art Show mean change = 2.2 and 2.1 for Spanish only and Spanish primarily groups, respectively), and children in the English primarily group made the smallest gains (Simon Says mean change = 0.9, Art Show mean change = 0.3). Focusing on Spanish-language expressive skills (Exposición de Arte), neither the Spanish only group nor the Spanish primarily group made significant progress between the fall and spring assessments, suggesting that the majority of the instruction occurs in English.

### 2. Language: EOWPVT

Expressive language skills as measured by the EOWPVT are summarized in the second section of Tables III.3 (fall and spring scores) and III.4 (change scores). LAUP children made statistically significant progress in an absolute sense (based on EOWPVT IRT scores) between fall and spring across the language groups; for all but the English only group they also made statistically significant progress in expressive language skills relative to bilingual peers. For both the standard scores and the IRT scores, children in the English primarily group demonstrated the strongest gains (mean changes = 3.5, 7.4, and 7.3 for EOWPVT and EOWPVT-SBE standard scores and IRT scores, respectively), and children in the English only group had the weakest gains (mean changes = 1.8, 2.2 and 4.4 for EOWPVT and EOWPVT-SBE standard scores and IRT scores, respectively).

Focusing on standard scores, the mean English standard scores indicate 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 sample sizes of fewer than 10 children). At both time points, children in the English only group scored higher than those in the English primarily group on this measure. Children in the English primarily group made statistically significant progress relative to a national sample of English-speaking peers; while progress of children in the English only group was not significant suggesting that more can be done to support

---

<sup>33</sup> See Appendix B for *pre*LAS scores for children whose parents reported they spoke a language other than English or Spanish. Among the 34 children from this group who were in the sample in both fall and spring, 35 percent (N = 12) were assessed in English at both time points.

**Table III.3. Means and Standard Errors for UPCOS-5 Language, Literacy, and Math Standardized Measures: Fall 2011 and Spring 2012, by Language Group**

Outcome	English Only				English Primarily				Spanish Only				Spanish Primarily			
	Fall 2011		Spring 2012		Fall 2011		Spring 2012		Fall 2011		Spring 2012		Fall 2011		Spring 2012	
	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE
<b>preLAS English</b>																
Total Language Screener Score	17.22***	0.31	18.43	0.21	15.00***	0.37	17.43	0.29	6.41***	0.64	11.24	0.79	9.22***	0.60	14.20	0.66
Simon Says	8.11***	0.23	8.98	0.15	6.87***	0.26	8.50	0.19	3.17***	0.39	5.84	0.47	4.08***	0.32	6.94	0.39
Art Show	9.11***	0.11	9.45	0.10	8.14***	0.16	8.94	0.15	3.24***	0.38	5.40	0.36	5.14***	0.36	7.26	0.32
preLAS Spanish																
Exposición de Arte			NA				NA		6.18	0.42	6.21	0.49	4.99***	0.33	5.29	0.43
<b>Language: Expressive One-Word Picture Vocabulary Test</b>																
IRT Score	51.49***	1.64	55.89	1.33	41.66***	1.44	48.98	1.10	36.50***	1.72	41.93	1.53	39.96***	1.20	45.70	1.14
English Edition Standard Score <sup>a</sup>	92.25	2.57	94.00	2.02	84.22*	1.50	87.71	1.70	Sample < 10				Sample < 10			
SBE Standard Score <sup>a,b</sup>	118.89	2.64	121.11	2.24	100.65***	2.48	108.04	2.06	87.59*	3.98	92.30	3.52	96.66*	2.71	102.41	2.55
<b>Literacy: Rapid Letter Naming</b>																
Raw Score	16.53***	1.27	21.77	0.98	12.31***	1.20	17.02	1.28	5.72***	1.64	11.36	1.92	8.47***	0.98	16.81	1.13
IRT Score	27.25***	1.54	33.60	1.16	22.49***	1.61	28.12	1.58	13.65***	2.15	22.53	2.15	17.59***	1.37	28.20	1.22
<b>Fine Motor and Literacy: Woodcock-Johnson III Spelling</b>																
W Score	377.91***	3.20	395.93	3.41	382.96***	2.71	399.74	2.69	Sample < 10				Sample < 10			
Standard Score <sup>a</sup>	103.17	1.83	105.48	1.97	104.56	1.45	105.96	1.62	Sample < 10				Sample < 10			
<b>Fine Motor and Literacy: Woodcock-Muñoz III (WM-III) Ortografía</b>																
W Score			NA				Sample < 10		359.11	5.96	376.51	6.57	356.21***	5.66	379.38	5.68
Standard Score <sup>a</sup>			NA				Sample < 10		91.93	3.29	93.21	3.38	93.42	2.99	98.04	3.34
<b>Mathematics: Woodcock-Johnson III Applied Problems</b>																
W Score	393.81***	2.56	406.84	2.51	397.72***	2.24	408.40	2.23	Sample < 10				Sample < 10			
Standard Score <sup>a</sup>	99.06	1.36	100.88	1.43	100.26	1.30	100.59	1.34	Sample < 10				Sample < 10			
<b>Mathematics: Woodcock-Muñoz III Problemas Aplicados</b>																
W Score			NA				Sample < 10		350.10**	8.22	376.30	5.53	349.66***	4.30	378.45	5.39
Standard Score <sup>a</sup>			NA				Sample < 10		77.73*	3.67	84.09	2.28	79.46*	1.72	87.28	2.41

Source: UPCOS-5 Fall 2011 and Spring 2012 Direct Child Assessments.

Note: All analyses are weighted to represent children attending LAUP in the 2011–2012 program year.

<sup>a</sup>National mean for standard scores is 100 with a standard deviation of 15.

<sup>b</sup>This measure was conceptually scored, and 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.

**Table III.4. Change Scores for UPCOS-5 Language, Literacy, and Math Standardized Measures: Fall 2011–Spring 2012, by Language Group**

Outcome	English Only			English Primarily			Spanish Only			Spanish Primarily		
	N	Mean	Standard Error	N	Mean	Standard Error	N	Mean	Standard Error	N	Mean	Standard Error
<b>Pre-LAS English</b>												
Total Language Screener Score	205	1.21***	0.19	175	2.43	0.26	55	4.84	0.60	117	4.98	0.47
Simon Says	205	0.86***	0.16	175	1.63	0.19	55	2.67	0.44	117	2.86	0.36
Art Show	205	0.34***	0.07	175	0.80	0.14	55	2.18	0.34	117	2.12	0.20
<b>Pre-LAS Spanish</b>												
Exposición de Arte		NA			NA		55	0.02	0.35	117	0.30	0.37
<b>Language: Expressive One-Word Picture Vocabulary Test</b>												
IRT Score	100	4.40**	0.61	78	7.32	1.15	32	5.43	0.83	57	5.73	0.99
English Edition Standard Score <sup>a</sup>	100	1.75	0.89	48	3.49	1.44	Sample < 10			Sample < 10		
Spanish Bilingual Edition Standard Score <sup>a,b</sup>	100	2.21	1.18	78	7.39	1.60	32	4.72	2.04	57	5.75	2.35
<b>Literacy: Rapid Letter Naming</b>												
Raw Score	100	5.24***	0.81	78	4.72	0.74	32	5.64	1.00	57	8.34	1.11
IRT Score	100	6.35**	0.99	78	5.63	1.02	32	8.89	1.22	57	10.60	1.12
<b>Fine Motor and Literacy: Woodcock-Johnson III Spelling</b>												
W Score	100	18.02	2.84	67	16.78	2.51	Sample < 10			Sample < 10		
Standard Score <sup>a</sup>	100	2.31	1.58	67	1.41	1.40	Sample < 10			Sample < 10		
<b>Fine Motor and Literacy: Woodcock-Muñoz III (WM-III) Ortografía</b>												
W Score		NA			Sample < 10		12	17.40	9.63	21	23.17	4.35
Standard Score <sup>a</sup>		NA			Sample < 10		12	1.28	5.32	21	4.62	2.39
<b>Mathematics: Woodcock-Johnson III Applied Problems</b>												
W Score	100	13.03	1.90	67	10.68	1.89	Sample < 10			Sample < 10		
Standard Score <sup>a</sup>	100	1.82	1.01	67	0.34	1.17	Sample < 10			Sample < 10		
<b>Mathematics: Woodcock-Muñoz III Problemas Aplicados</b>												
W Score		NA			Sample < 10		12	26.20	6.51	21	28.79	6.18
Standard Score <sup>a</sup>		NA			Sample < 10		12	6.36	2.74	21	7.82	2.73

Source: UPCOS-5 Fall 2011 and Spring 2012 Direct Child Assessments.

Note: All analyses are weighted to represent children attending LAUP in the 2011–2012 program year.

<sup>a</sup>National mean for standard scores is 100 with a standard deviation of 15.

<sup>b</sup>This measure was conceptually scored and 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.

**Table III.5. Percentage of Children Scoring in the Expected Range on the Leiter-R: Fall 2011 and Spring 2012, by Language Group**

Outcome	English Only				English Primarily				Spanish Only				Spanish Primarily			
	Fall 2011		Spring 2012		Fall 2011		Spring 2012		Fall 2011		Spring 2012		Fall 2011		Spring 2012	
	%age	SE	%age	SE	%age	SE	%age	SE	%age	SE	%age	SE	%age	SE	%age	SE
<b>Social-Emotional and Approaches to Learning</b>																
Leiter Examiner Ratings Scaled Score in the Acceptable Range <sup>a</sup>																
Attention	96.88	1.78	97.29	1.95	91.36*	3.29	98.93	1.09	91.57	5.77	100.00	0.00	96.40	2.51	100.00	0.00
Activity	100.00	0.00	98.31	1.70	90.35**	3.12	100.00	0.00	81.56*	7.96	100.00	0.00	85.01*	5.54	100.00	0.00
Sociability	100.00	0.00	98.31	1.70	92.16	3.11	98.68	1.27	87.41	7.91	94.77	3.21	94.31	3.00	100.00	0.00
N	100				78				32				57			

Source: UPCOS-5 Fall 2011 and Spring 2012 Direct Child Assessment.

Note: All analyses are weighted to represent children attending LAUP in the 2011–2012 program year.

<sup>a</sup>Scores of 7 or greater.

\* 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 III.6. Change in Percentage of Children Scoring in the Expected Range on the Leiter-R: Fall 2011–Spring 2012, by Language Group**

Outcome	English Only			English Primarily			Spanish Only			Spanish Primarily		
	N	Percentage	Standard Error	N	Percentage	Standard Error	N	Percentage	Standard Error	N	Percentage	Standard Error
<b>Social-Emotional and Approaches to Learning</b>												
Leiter Examiner Ratings Scaled Score in the Acceptable Range <sup>a</sup>												
Attention	100	0.41	2.77	78	7.57	3.54	32	8.43	5.77	57	3.60	2.51
Activity	100	-1.69***	1.70	78	9.65	3.12	32	18.44	7.96	57	14.99	5.54
Sociability	100	-1.69	1.70	78	6.52	3.65	32	7.36	6.78	57	5.69	3.00

Source: UPCOS-4 Fall 2011 and Spring 2012 Direct Child Assessments.

Note: All analyses are weighted to represent children attending LAUP in the 2011–2012 program year.

<sup>a</sup>Scores of 7 or greater.

\* 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.

English vocabulary development of children entering with English. However, the sample size is smaller than in previous phases of UPCOS limiting the power to detect change.

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 one-half to a full standard deviation. Children in the Spanish only group scored below the national sample of bilingual peers at both time points. In the fall, the children in this group scored more than two-thirds of a standard deviation below the national average, and in the spring, the children in this group remained more than one-third of a standard deviation below the national average. For children in the Spanish primarily group, although their scores in the fall were slightly below the national average, they scored above the national average by spring. Children in all language groups except the English only group made statistically significant progress from the fall to the spring as measured by SBE standard scores. 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 demonstrated greater conceptual knowledge (as measured through expressive vocabulary) and that 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; children whose parents indicated they spoke Spanish primarily performed similarly to their same-age bilingual peers. However, the children in the English only group demonstrated the least amount of change when compared with the other groups. Note that even though parents of children in the English only and primarily groups indicated that most of their children's interactions were in English, these children often came from homes where Spanish and/or other languages were spoken.

### **3. Literacy: Rapid Letter Naming**

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. Children in the Spanish primarily group made the most progress (mean change score = 8.3), and children in the English primarily group made the least progress (mean change = 4.7). The pattern was similar for the IRT scores, which adjusted for the difficulty of the items; children who spoke Spanish primarily gained the most (mean change = 10.6), and those who spoke English primarily gained the least (mean change = 5.6).

### **4. Fine Motor and Literacy: WJ-III Spelling**

For fine motor and literacy development as measured by the WJ-III Spelling subtest, all children in both English only and English primarily groups made statistically significant absolute progress from fall to spring (with an average change of 18.0 and 16.8, respectively, based on *W* scores; we excluded scores for children in the Spanish only and Spanish primarily groups from the table because of sample sizes of fewer than 10 children). The progress of children who spoke English only or primarily was similar to that of a national sample of same-age peers (with average change of 2.3 and 1.4 based on standard scores, respectively, that was not statistically significant). In both fall and spring, children in both groups scored above the national average on the English Spelling subtest.

### 5. Fine Motor and Literacy: WM-III Ortografía

Shifting to the Spanish measure of fine motor and literacy (WM-III Ortografía), the Spanish only and the Spanish primarily groups made similar progress from fall to spring in absolute terms (with average change of 17.4 and 23.2, respectively, based on *W* scores). The gains were significant for the Spanish primarily group but not significant for the Spanish only group (the small sample size in the Spanish only group limits the power to detect differences). We excluded from the table the scores for children in the English primarily group because of a sample size of fewer than 10 children, and none of the children in the English only group received this measure. Neither group of children made significant progress relative to a national sample of same-age Spanish-speaking peers. By spring, on average, the performance of children in the Spanish only group was below that of their peers ( $SS = 93.2$ ), whereas the performance of children in the Spanish primarily group approached the national average ( $SS = 98.0$ ). 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. More than one-half (60 percent) of the children who completed Version 1 of the measure battery were assessed in Spanish in the fall and switched to English in the spring. The mean spring scores on the WJ-III Spelling for children who changed to an English administration in the spring was 103.6 for those in the Spanish primarily group and 99.9 for those in the English primarily group.<sup>34</sup>

### 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 (13.0 and 10.7, respectively; we excluded from the table scores for children in the Spanish only and Spanish primarily groups 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 both groups; their standard scores were around the national average of 100 in both fall and spring.

### 7. Mathematics: WM-III Problemas Aplicados (Applied Problems)

Shifting to the Spanish-language measure of mathematics (WM-III Problemas Aplicados), both groups of ELL children made statistically significant progress in absolute terms (with average change based on *W* scores of 26.2 for the Spanish only group and 28.8 for the Spanish primarily group; we excluded from the table the scores for children in the English primarily group because of a sample size of fewer than 10 children, and none of the children in the English only group received this measure). Although, on average, children had one-third of a standard deviation increase in standard scores, the estimates have greater error, so the progress relative to same-age Spanish speaking peers (mean change = 6.4 and 7.8 for Spanish only and Spanish primarily groups, respectively) does not represent a statistically significant change. Their scores remained nearly one standard deviation below the national mean by spring. 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. More than half (60 percent) of the children who completed Version 1 of the measure battery were assessed in Spanish in the fall and switched to English in the spring. The mean spring scores on the WJ-III Applied Problems for the group that changed to an English administration in the spring were 89.0 and 94.7, respectively, for the English primarily and Spanish primarily groups.

---

<sup>34</sup> The result was not presented for children in the Spanish only group because the sample size was fewer than 10 children.

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

There were also differences across the language groups in progress in social-emotional development and approaches to learning (Tables III.5 and III.6). Overall, fewer children in the English only group moved to the expected score range than did children in other groups for all of the Leiter-R subtests because almost all children in the English only group began the year in the expected range. In Attention, there were more children in the English only and Spanish primarily groups (97 and 96 percent, respectively) than in the English primarily and Spanish only groups (91 and 92 percent, respectively) who scored in the expected range in the fall; by the spring, almost all children across language groups scored in the expected range (ranging from 97 to 100 percent). In Activity Level, the Spanish only group experienced the largest increase (18 points) in the percentage of children scoring in the expected range, followed by the Spanish primarily group (15 percent), whereas children in the English only group had no change because all children in that group scored in the expected range in the fall. In Sociability, almost all children in the English only group scored in the expected range in both fall and spring; the percentage of children scoring in the expected range increased about 6 to 7 points from fall to spring for the other groups.

### D. Summary and Implications

On average, 5.6 months passed between the fall and spring assessments (range of 3.4 to 6.4 months). Between the fall and spring, children in LAUP made statistically significant progress in all areas when based on absolute progress. Relative to a national sample of same-age peers, children in LAUP made statistically significant progress in expressive vocabulary (based on both English and bilingual samples). Children assessed in Spanish also made significant progress in mathematics relative to a national sample of Spanish-speaking same age peers. LAUP children kept pace with but did not make significant progress relative to a national sample of peers in literacy (as measured by the WJ-III Spelling and WM-III Ortografía) and mathematics. Note that for the WJ-III Spelling subtest and the Leiter-R, children's performance was quite strong relative to national peers. Scores for these measures exceeded a national sample of peers in both the fall and spring,<sup>35</sup> and children also exceeded the average performance of their bilingual peers and made significant progress in EOWPVT-SBE relative to their same-age bilingual peers. On average, children's scores on WJ-III Applied Problems subtest kept up with the national mean in both the fall and spring, suggesting that children progressed at the same rate as their same-age English-speaking peers. Scores on the RLN task indicated that, on average, children made statistically significant progress in their letter-naming skills and the results were comparable to other preschool studies using similar measures (Early et al. 2005). Recent research indicates that children who are able to quickly and easily name 15 lowercase and 18 uppercase letters are successful in literacy in first grade (Piasta et al. 2012). Children name a mix of upper- and lowercase letters on the RLN. The uppercase letters are easier for children to name and on average, in the spring, LAUP children named the equivalent of 14 lowercase and 18 uppercase letters.

Although children made progress in most areas, there is still room to grow. LAUP children made significant progress in expressive vocabulary relative to a national sample of English-speaking peers during the program year, but their mean score remained about half of a standard deviation below that of their English-speaking peers nationally in the spring. Knowledge of vocabulary and

---

<sup>35</sup> For the Leiter-R, we calculated the percentage of children scoring in the expected range; children scoring in this range are unlikely to be experiencing difficulties with social-emotional development and approaches to learning. In the national sample, 84 percent of children score in this range.

concepts reflects knowledge of the world. Children draw on this background knowledge as they read and comprehend text and other educational opportunities in school. Children with more limited knowledge of the world and the words used to describe it are at a disadvantage when trying to comprehend. They are limited in their ability to draw on background knowledge and make connections among ideas. For example, a recent study (Grissmer et al. 2010) found general knowledge when entering kindergarten to be the strongest predictor of later “science and reading and also contributed significantly to predicting later mathematics” (p. 1008). Drawing on analyses of multiple longitudinal studies, the authors concluded, “Together, attention, fine motor skills, and general knowledge are much stronger overall predictors of later math, reading, and science scores than early math and reading scores alone” (p. 1008). More frequent reading to children and talking about books is one of the most widely used means of increasing vocabulary and exposure to world knowledge (De Temple and Snow 2003; Hargrave and Sénéchal 2000; Justice et al. 2009).

For children in LAUP who took the English assessment, mathematics abilities were similar to those of a national sample of English-speaking peers, and these children maintained their progress relative to peers. This represents an improvement over last year and may reflect greater attention to asking reasoning questions after LAUP’s focus on this area at the Teacher Institute and the inclusion of the Classroom Assessment Scoring System (CLASS, Pianta et al. 2008) in LAUP’s 5-Star Quality Assessment and Improvement Scale. The qualitative study of coaching also completed as part of UPCOS-5 suggested that in addition to more attention given to asking questions, some programs also implemented classroom organization systems that included asking children to reason about number as they determined how many additional children could play in a particular area during free choice time (Winston et al. 2012).

However, mathematics abilities as measured by the WM-III (Spanish version) were below those of their Spanish-speaking peers. Despite the fact that ELL children made enough progress so as to not lose additional ground relative to a national sample of similarly achieving peers, their scores in both the fall and spring were nearly one standard deviation below the national mean. Mathematics is an important area of cognition and is a strong predictor of success in school in both reading and mathematics (Claessens et al. 2009; Duncan et al. 2007). Thus, ELL children are in particular need additional support to make gains in the area of mathematics.

## IV. CHILD PROGRESS RELATIVE TO TARGETS SET FOR THE 2011–2012 PROGRAM YEAR

In this chapter we present child progress in LAUP during the 2011–2012 program year relative to the targets outlined in Chapter I. We begin with a discussion of performance relative to targets in all domains,<sup>36</sup> including an analysis of children’s average performance by whether targets were met. We then present additional detail about children’s performance in each domain to better understand the pattern of results.

### 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 (Table IV.1) and by fall quartile (Table IV.2). For each measure, we also examined (1) the mean and range of change scores (Table IV.3) and (2) fall and spring scores by fall quartile (Tables IV.4 and IV.5) for the group of children whose scores exceeded the Level 1 target and for 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. We conducted this analysis regardless of whether the overall target was met. For these five measures, we also calculated fall and spring scores by language group for the group of children whose scores exceeded the Level 1 target and for the group whose scores did not as a sensitivity analysis; those scores are presented in Appendix C (Tables C.1 and C.2). We excluded the Leiter-R from this analysis because virtually all children scored at a level required by the targets. The analysis was specific to each measure (a child could be in the group whose scores exceeded targets for one measure but in the group whose scores fell below targets for another measure). For the quartile analysis, groups were based on fall scores in the full sample of children; thus, whether children fell 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. In some instances, a particular subgroup fell below 10 observations; these results were excluded from the table 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 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.

As in the prior chapter, all analyses discussed in this chapter were at the child level and were weighted to represent LAUP children as a whole. Analyses were conducted using survey procedures that address the clustering of children within programs and classrooms. Specifically, the survey procedures accounted 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 correctly estimated and, thus, that any statistical tests based on those standard errors were accurate.

### B. Progress Relative to Targets

Table IV.1 outlines children’s performance relative to targets set for the 2011–2012 program year. Both Level 1 and Level 2 targets were met for language, fine motor/literacy, mathematics, and social-emotional development and approaches to learning. Literacy as measured by the RLN task

---

<sup>36</sup> For measures that differed by language of administration (that is, Applied Problems/Problemas Aplicados and Spelling/Ortografía), the analysis of targets was based on the sample of children who took the assessment in the same language in the fall and spring.

**Table IV.1. Summary of Child Progress in 2011–2012 in a Representative Sample of LAUP Programs and Children Relative to Performance-Based Contract Targets**

Domain: Measure	Level 1	Level 2
<b>Language: EOWPVT (N = 273)</b>		
Target	70% gain 2 points or more	45% gain 5 points or more
Actual	72% gained 2 points or more	51% gained 5 points or more
Target reached?	Yes	Yes
<b>Literacy: Rapid Letter Naming (N = 254)</b>		
Target	70% gain 7 points or more	45% gain 13 points or more
Actual	51% gained 7 points or more	20% gained 13 points or more
Target reached?	No	No
<b>Fine Motor/Literacy: WJ-III Spelling (N = 189)</b>		
Target	70% gain 7 points or more	45% gain 17 points or more
Actual	77% gained 7 points or more	51% gained 17 point or more
Target reached?	Yes	Yes
<b>Mathematics: WJ-III Applied Problems (N = 189)</b>		
Target	70% gain 4 points or more	45% gain 13 points or more
Actual	76% gained 4 points or more	49% gained 13 points or more
Target reached?	Yes	Yes
<b>Mathematics: WM-III Problemas Aplicados (N = 36)</b>		
Target	70% gain 4 points or more	45% gain 13 points or more
Actual	93% gained 4 points or more	72% gained 13 points or more
Target reached?	Yes	Yes
<b>Social-Emotional and Approaches to Learning: Leiter-R (N = 187 for Level 1 and 94 for Level 2)</b>		
Target	85% score in expected range in spring for English only and English primarily groups	75% score in expected range in spring for Spanish only or primarily and other language only and primarily groups
Attention		
Actual	98%	100%
Target reached?	Yes	Yes
Activity Level		
Actual	99%	100%
Target reached?	Yes	Yes
Sociability		
Actual	98%	98%
Target reached?	Yes	Yes

Source: UPCOS-5 Fall 2011 and Spring 2012 Direct Child Assessments.

Notes: Targets for all measures except the Leiter-R were based on W or IRT scores. Leiter-R targets refer to the proportion of children scoring in the expected range, which was based on standardized scaled scores.

All analyses are weighted to represent children attending LAUP in the 2011–2012 program year.

was the only area where the percentage of children who made the required level of progress was less than the targeted percentage at both Level 1 and Level 2.

For the EOWPVT, WJ-III, and WM-III measures, we repeated the target analysis excluding children who scored more than two standard deviations below the mean for that particular measure as a proxy for identifying children who might have special needs. Because the RLN is not a standardized measure (and does not have national norms), we based the exclusion on the EOWPVT. The pattern of results did not change when instituting this exclusion (that is, the literacy targets based on the RLN were not met). We now present additional detail on each target.

### **1. Language: EOWPVT**

The first section of Table IV.1 presents performance relative to the Level 1 and Level 2 targets for language development as measured by the EOWPVT. LAUP exceeded both Level 1 and Level 2 targets in this area. Beginning with Level 1, 72 percent of children gained 2 points or more on the EOWPVT, compared to the target of 70 percent of children who achieved this gain. For the Level 2 target, 51 percent of children gained 5 points or more (compared to a target of 45 percent).

### **2. Literacy: Rapid Letter Naming**

The second section of Table IV.1 presents performance relative to the Level 1 and Level 2 targets for literacy development as measured by the RLN. At both Level 1 and Level 2, the percentage of children who made the required level of progress was less than the targeted percentage in this area. Beginning with Level 1, 51 percent of children gained 7 points or more on the RLN; the target calls for 70 percent to achieve a gain of this magnitude. Similarly, the Level 2 target calls for 45 percent of children to gain 13 points or more, but only 20 percent of children made this gain.

### **3. Fine Motor and Literacy: WJ-III Spelling**

The third section of Table IV.1 summarizes performance relative to the targets for fine motor and literacy development as measured by the WJ-III Spelling subtest. LAUP surpassed both the Level 1 and Level 2 targets in this area. At Level 1, the target calls for 70 percent of children to gain 7 points or more, and 77 percent achieved a gain of this magnitude. At Level 2, the target calls for 45 percent of children to gain 17 points or more, and 51 percent achieved a gain of this magnitude.

### **4. Mathematics: WJ-III Applied Problems**

The fourth section of Table IV.1 summarizes performance relative to the targets for mathematics development as measured by the WJ-III Applied Problems subtest. LAUP exceeded both the Level 1 and Level 2 targets in this area. At Level 1, the target called for 70 percent of children to gain 4 points or more, and 76 percent achieved a gain of this magnitude. At Level 2, the target called for 45 percent of children to gain 13 points or more, and 49 percent achieved a gain of this magnitude.

### **5. Mathematics: WM-III Problemas Aplicados**

The fifth section of Table IV.1 summarizes performance relative to targets for mathematics development as measured by the WM-III Problemas Aplicados subtest. LAUP surpassed the targets in this area. At Level 1, the target calls for 70 percent of children to gain 4 points or more, and 93 percent achieved a gain of this magnitude. At Level 2, the target calls for 45 percent of children to gain 13 points or more, and 72 percent achieved a gain of this magnitude.

## 6. Social-Emotional and Approaches to Learning: Leiter-R

The final section of Table IV.1 summarizes performance relative to the targets for social-emotional development and approaches to learning as measured by the Leiter-R subtests for Attention, Activity Level, and Sociability. LAUP achieved targets at both Level 1 and Level 2. At Level 1, the target calls for 85 percent of children in the English only and English primarily groups to score in the expected range in the spring on each of the three subtests; 98 to 99 percent or more of children did so, depending on the subtest. At Level 2, the target calls for 75 percent of children in the Spanish only, Spanish primarily, and other language only or primarily groups to score in the expected range in the spring; 98 to 100 percent of children did so, depending on the subtest. Note that we did not separately summarize results for children whose scores were not within range of the target by quartile or by language group because, fortunately, too few children were in these subgroups.

## 7. Targets by Quartile

Table IV.2 (Figure IV.1) presents the percentage of children whose progress exceeded the Level 1 targets by fall quartile (the Leiter-R was excluded from this table because the approach to scoring—percentage of children in the expected range—does not lend itself to quartile analysis). For four of the five measures shown—the EOWPVT (language), RLN (literacy), the WJ-III Spelling (fine motor and literacy), and the WJ-III Applied Problems (mathematics in English)—children whose fall scores were in the bottom or middle quartiles were more likely to have made progress that exceeded the Level 1 target than were children in the top quartile. The opposite was true for WM-III Problemas Aplicados (mathematics in Spanish)—all children in the top quartile made progress that met the Level 1 target, whereas 86 to 89 percent of children in the bottom or middle quartiles did so.

In terms of targets, children in the bottom quartile had change scores that exceeded the Level 1 target of 70 percent for all measures. Children who scored in the middle 50 percent in the fall also progressed to a degree that exceeded the Level 1 target for all measures except in the case of the RLN task. With the exception of the WM-III Problemas Aplicados subtest, scores of children in the top quartile did not exceed targets for the measures shown.

**Table IV.2. Percentage of Children Whose Progress Exceeded Level 1 Targets by Fall Quartile**

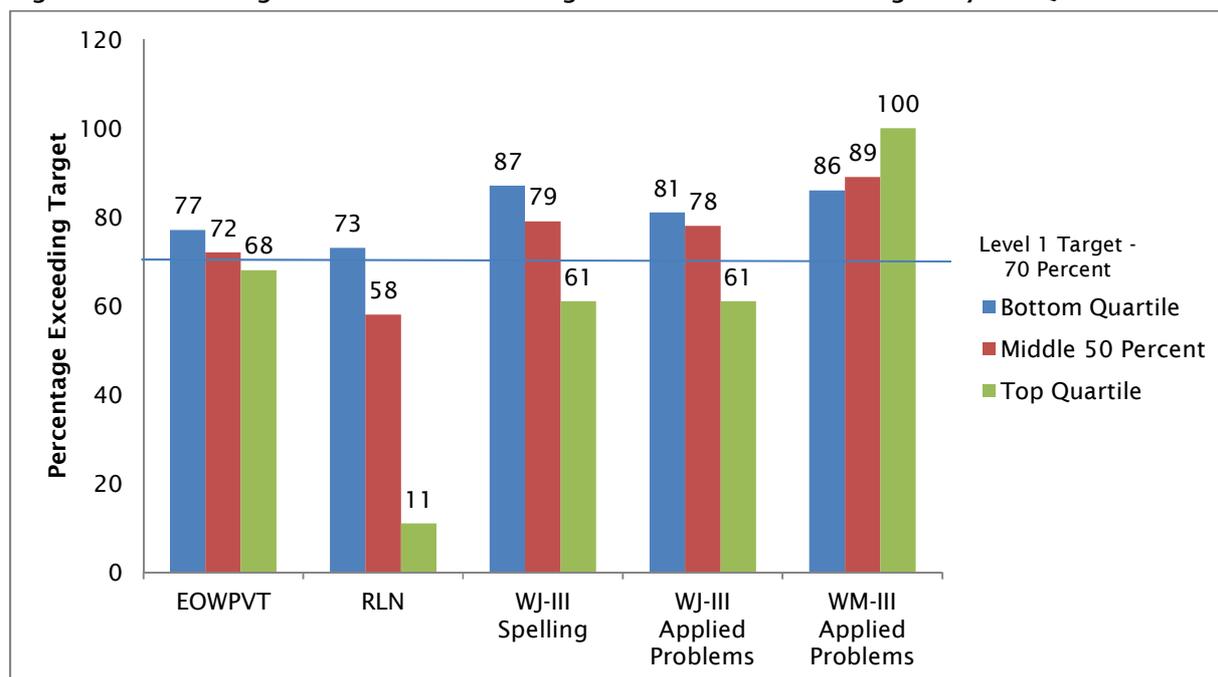
Domain: Measure	Bottom Quartile	Middle 50 Percent	Top Quartile
Language: EOWPVT	76.98	71.50	68.38
Literacy: Rapid Letter Naming	72.59	58.32	11.30
Fine Motor/Literacy: WJ-III Spelling	87.20	78.69	61.27
Mathematics: WJ-III Applied Problems	80.87	77.58	69.53
Mathematics: WM-III Problemas Aplicados	86.38	89.22	100.00

Source: UPCOS-5 Fall 2011 and Spring 2012 Direct Child Assessments.

Note: Targets for all measures were based on W or IRT scores.

All analyses are weighted to represent children attending LAUP in the 2011–2012 program year.

**Figure IV.1. Percentage of Children Whose Progress Exceeded Level 1 Targets by Fall Quartile**



### C. Progress by Whether Level 1 Target Was Met

In the remainder of this chapter, we provide additional detail on children’s performance in each domain of development addressed in UPCOS-5 (during the 2011–2012 program year) to understand the pattern of children’s progress and performance relative to targets. Table IV.3 presents the mean change and the range of change scores by whether the Level 1 target was met. In Tables IV.4 and IV.5, we present fall and spring scores for the children whose progress did and did not meet targets, respectively, by quartile. For all tables, standard scores follow the IRT or W score on which the target was based (as a reminder, the RLN task does not have standard scores).

**Language: EOWPVT.** Turning to the first section of Table IV.3, we see that for children whose scores met the target, because the target was based on the IRT score, by definition all children in this group made positive, absolute change (mean = 8.5, range = 2.1 to 55.2). For this group, average standard scores for the English and SBE editions increased; the expressive vocabulary of children in this group generally improved relative to samples of same-age peers. Note, however, that ranges for both editions include negative minimum values (English: range = -8 to 43; SBE: range = -12 to 63), indicating that although all children whose scores met targets made absolute progress, some lost ground relative to a national sample of same-age peers. For children whose score did not meet targets, most lost ground relative to same-age peers in English and bilingual expressive skills (English: mean = -7.0, range = -27 to 11; SBE: mean = -9.1, range = -26 to 0), and the mean change (loss) was equivalent to more than one-half of a standard deviation for the national sample.

**Literacy: Rapid Letter Naming.** As shown in the second section of Table IV.3, consistent with the finding that progress in the overall sample was positive, average progress in literacy as measured by the RLN task was positive based on raw or IRT scores. However, for the group whose scores did not meet targets, average progress was only about one point for both raw and IRT scores, and ranges indicate that some children lost ground during the year (raw score: range = -10 to 7; IRT

**Table IV.3. Means, Standard Errors, and Ranges for UPCOS-5 Language, Literacy, and Math Measures Change Scores, Child Level, by Whether Level 1 Target Was Met**

Outcome	Change Scores That Reached Target					Change Scores That Did Not Meet Target				
	N	Mean	Std Error	Min	Max	N	Mean	Std Error	Min	Max
<b>Language: Expressive One-Word Picture Vocabulary Test</b>										
IRT Score	197	8.45	0.58	2.11	55.23	76	-1.26	0.34	-21.99	1.91
English Edition Standard Score <sup>a</sup>	118	5.98	0.72	-8	43	46	-7.04	0.94	-27	11
Spanish Bilingual Edition Standard Score <sup>a, b</sup>	197	10.30	1.02	-12	63	76	-9.08	0.59	-26	0
<b>Literacy: Rapid Letter Naming</b>										
Raw Score	137	10.37	0.60	2	25	134	1.05	0.28	-10	7
IRT Score	137	13.41	0.54	7.09	32.83	134	0.94	0.47	-20.82	6.7
<b>Fine Motor and Literacy: Woodcock-Johnson III Spelling</b>										
W Score	144	25.53	1.28	7	70	45	-9.34	2.08	-52	6
Standard Score <sup>a</sup>	144	6.34	0.70	-5	30	45	-12.69	1.21	-38	-3
<b>Mathematics: Woodcock-Johnson III Applied Problems</b>										
W Score	142	18.77	1.16	4	80	47	-7.54	1.47	-54	0
Standard Score <sup>a</sup>	142	5.14	0.66	-5	31	47	-9.88	0.73	-27	-4
<b>Mathematics: Woodcock-Muñoz-III Problemas Aplicados</b>										
W Score	32	31.81	3.21	6	81	Sample < 10			-47	0
Standard Score <sup>a</sup>	32	9.03	1.38	-2	30	Sample < 10			-31	-5

Source: UPCOS-5 Fall 2011 and Spring 2012 Direct Child Assessments.

Note: All analyses are weighted to represent children attending LAUP in the 2011–2012 program year.

<sup>a</sup>National mean for standard scores is 100 with a standard deviation of 15.

<sup>b</sup>This measure was conceptually scored, and the standard score was generated for all children in the sample.

**Table IV.4. Means and Standard Errors for UPCOS-5 Language, Literacy, and Math Standardized Measures: Fall 2011 and Spring 2012 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 2011		Spring 2012		Fall 2011		Spring 2012		Fall 2011		Spring 2012	
	Mean	Standard Error	Mean	Standard Error	Mean	Standard Error	Mean	Standard Error	Mean	Standard Error	Mean	Standard Error
<b>Language: Expressive One-Word Picture Vocabulary Test</b>												
IRT Score	30.58	0.97	42.96	0.67	43.56	0.50	50.91	0.57	57.46	0.88	63.92	1.03
English Edition Standard Score <sup>a</sup>	71.12	1.06	80.83	1.58	88.48	0.95	93.13	1.20	106.31	1.80	109.69	2.48
SBE Standard Score <sup>a,b</sup>	76.04	1.09	95.77	2.09	104.74	1.23	112.72	1.15	130.16	1.48	134.74	1.37
<b>Literacy: Rapid Letter Naming</b>												
Raw Score	0.92	0.11	10.00	1.13	9.71	0.72	21.60	0.88	Sample < 10			
IRT Score	6.32	0.57	21.75	1.06	20.89	0.76	33.23	0.93	Sample < 10			
<b>Fine Motor and Literacy: Woodcock-Johnson III Spelling</b>												
W Score	355.96	1.85	389.66	2.81	383.99	1.17	407.55	1.22	403.95	1.03	419.54	0.99
Standard Score <sup>a</sup>	88.28	1.23	98.91	2.07	105.75	0.50	111.43	0.82	119.14	1.06	119.49	1.29
<b>Mathematics: Woodcock-Johnson III Applied Problems</b>												
W Score	373.12	2.96	398.91	1.97	397.71	0.67	414.38	1.31	414.51	2.18	429.36	3.06
Standard Score <sup>a</sup>	88.50	1.25	95.85	0.92	100.38	0.41	104.93	0.88	112.06	1.83	115.76	2.57
<b>Mathematics: Woodcock-Muñoz-III Problemas Aplicados</b>												
W Score	325.46	3.79	373.93	4.98	352.72	3.02	380.26	3.79	375.20	2.39	394.72	3.02
Standard Score <sup>a</sup>	Sample < 10				79.45	1.11	85.76	1.50	89.81	1.28	93.87	1.23
N (range)	13-64				16-104				10-48			

Source: UPCOS-5 Fall 2011 and Spring 2012 Direct Child Assessments.

Note: All analyses are weighted to represent children attending LAUP in the 2011–2012 program year.

<sup>a</sup>National mean for standard scores is 100 with a standard deviation of 15.

<sup>b</sup>This measure was conceptually scored, and the standard score was generated for all children in the sample.

**Table IV.5. Means and Standard Errors for UPCOS-4 Language, Literacy, and Math Standardized Measures: Fall 2011 and Spring 2012 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 2011		Spring 2012		Fall 2011		Spring 2012		Fall 2011		Spring 2012	
	Mean	Standard Error	Mean	Standard Error	Mean	Standard Error	Mean	Standard Error	Mean	Standard Error	Mean	Standard Error
<b>Language: Expressive One-Word Picture Vocabulary Test</b>												
IRT Score	32.35	1.04	31.83	1.92	44.94	0.73	43.71	0.61	60.46	1.78	58.69	1.61
English Edition Standard Score <sup>a</sup>	Sample <10				90.41	0.77	84.01	1.87	112.18	3.41	102.36	2.78
SBE Standard Score <sup>a,b</sup>	77.93	3.04	73.02	2.16	108.58	1.74	98.56	1.56	135.67	2.40	125.56	2.68
<b>Literacy: Rapid Letter Naming</b>												
Raw Score	0.90	0.16	1.54	0.31	13.11	1.31	15.53	1.37	27.24	0.25	27.15	0.25
IRT Score	6.14	0.80	8.50	1.11	23.89	1.27	26.42	1.28	40.46	0.74	39.63	0.59
<b>Fine Motor and Literacy: Woodcock-Johnson III Spelling</b>												
W Score	Sample < 10				383.07	2.28	371.10	4.15	410.99	2.53	403.39	4.10
Standard Score <sup>a</sup>	Sample < 10				106.72	1.45	93.16	2.47	121.81	1.55	110.76	2.75
<b>Mathematics: Woodcock-Johnson III Applied Problems</b>												
W Score	367.25	7.41	352.13	10.01	400.07	1.42	396.58	1.78	415.18	2.30	407.58	1.82
Standard Score <sup>a</sup>	86.32	3.12	74.24	3.81	100.23	0.58	93.21	0.69	112.87	1.61	101.12	1.38
<b>Mathematics: Woodcock-Muñoz-III Problemas Aplicados</b>												
W Score	Sample < 10				Sample < 10				Not Applicable			
Standard Score <sup>a</sup>	Sample < 10				Sample < 10				Not Applicable			
N (range)	10-23				21-62				16-61			

Source: UPCOS-5 Fall 2011 and Spring 2012 Direct Child Assessments.

Note: All analyses are weighted to represent children attending LAUP in the 2011-2012 program year.

<sup>a</sup>National mean for standard scores is 100 with a standard deviation of 15.

<sup>b</sup>This measure was conceptually scored, and the standard score was generated for all children in the sample.

score: range = -20.8 to 6.7), suggesting that more attention was given to letter names in the fall than in the spring of the year.

**Fine Motor and Literacy: WJ-III Spelling.** The third section of Table IV.3 shows that children whose progress exceeded targets also made progress relative to a nationally representative sample of peers (mean SS change = 6.3), although children who made the smallest amount of absolute progress lost some ground relative to peers (range = -5 to 30). The mean change is equivalent to more than one one-third of a standard deviation increase in the national sample. For children whose scores did not meet the Level 1 target, mean change scores indicate large declines both in absolute terms and relative to same-age peers (mean change = -9.3 and -12.7, respectively); all these children lost ground relative to peers (range = -38 to -3).

**Mathematics: WJ-III Applied Problems.** The fourth section of Table IV.3 presents means and ranges of change scores for the English mathematics assessment. On average, children whose progress exceeded targets made progress both in absolute terms and relative to peers (mean change = 18.8 and 5.1, respectively), although children who made the smallest amount of absolute progress lost some ground relative to peers (range = -5 to 31). For children whose scores did not meet the Level 1 target, scores indicate declines both in absolute terms (mean W score change = -7.5) and relative to peers (mean SS change = -9.9); all these children lost ground relative to peers (range = -27 to -4). The mean decrease in standard scores was equivalent to approximately two-thirds of a standard deviation for the national sample.

**Mathematics: WM-III Problemas Aplicados.** The final section of Table IV.3 presents means and ranges of change scores for the Spanish mathematics assessment. On average, children whose progress exceeded targets made change relative to a national sample of similarly achieving peers on the WM-III Problemas Aplicados test (mean SS change = 9.0), although children who made the smallest amount of absolute progress lost some ground relative to peers (range = -2 to 30). There were only four children whose scores did not meet the Level 1 target; thus, mean change scores are not shown in the table due to possible imprecise estimation. Ranges indicate large declines both in absolute terms (W score range = -47 to 0) and relative to peers (SS range = -31 to -5).

### 1. Fall and Spring Scores by Quartile and Performance Relative to Targets

Table IV.4 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 IV.5 presents the same information for children whose progress did not exceed targets. Beginning with Table IV.4, 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 children who began the year in the bottom quartile had the lowest spring scores.

Shifting to Table IV.5 (which presents scores by quartile for the group whose scores did not exceed Level 1 targets), we see that whether they began the year in the top quartile, middle 50 percent, or the bottom quartile, children scored lower in the spring than in the fall, on average, 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 target to be included in the table.

Looking across Tables IV.4 and IV.5, we see that for each measure and for each 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). All but a few of the children in the bottom quartile for the EOWPVT English standard scores, the bottom quartile for the WJ-III Spelling, and for all quartiles for the WM-III Problemas Aplicados subtest met the targets, whereas too few of the children in the top quartile for the RLN tasks met the target. 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 that did not within each quartile.

As a sensitivity analysis, we also examined the fall and spring scores of children, by language group, whose progress did and did not meet targets (Appendix C). Patterns were similar for the analysis by language group as for the analysis by fall quartile. 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 meet targets for each measure. Typically, scores in the spring were higher for the group whose progress met targets, with the exception of the RLN, where scores were often similar. Note that in a number of instances the sample size of the language group fell below 10; thus, some comparisons could not be made.

## 2. Summary

These analyses provide strong support for the importance and meaningfulness of most of the targets. On average, children who met the targets made gains relative to national samples of peers on each of the measures, and large gains for most of them. It is very unusual for groups of children to make progress that increases standard scores equivalent to half a standard deviation in the national sample. Alternatively, those children who did not meet the targets lost considerable ground relative to national samples of peers, suggesting that the targets were not set unreasonably high. This pattern of findings—large gains among children whose progress met targets and declines among children whose progress fell below targets—was found in the overall sample and in subgroups defined by fall quartile and language group. If children did not make the specified level of change, they fell further behind relative to other children across the nation, so it is important to continue to strive to help children make the specified levels of change.

## D. Summary and Implications

In the 5.6 months, on average, between assessments during the 2011–2012 program year, the targets were met for language, fine motor/literacy, mathematics, and social-emotional development and approaches to learning. For literacy as measured by the RLN task, the percentage of children who made the required level of progress was less than the targeted percentage. Additional examination of performance by fall quartile indicated that children who began the year with lower scores (in the bottom or middle quartiles) were more likely to have change scores in line with what

was specified in the targets than children who began the year with scores in the top quartile for the measures of language, fine motor/literacy (as measured by the WJ-III Spelling and the RLN), and mathematics (assessed in English). All of the children in the top quartile who took the mathematics assessment in Spanish in both fall and spring met the target, whereas about 10 to 15 percent of the children in the other quartiles who took the assessment in Spanish at both time points did not.

### 1. Meaningfulness of the Targets

We examined average change scores among the group of children whose scores exceeded Level 1 targets and the group whose scores did not exceed Level 1 targets.<sup>37</sup> Standard scores (scores relative to a national sample peers that are adjusted for child age) were available for four of the measures: EOWPVT, WJ-III Spelling, WJ-III Applied Problems, and WM-III Problemas Aplicados. For children whose scores met or exceeded targets, on average scores increased relative to a national sample of same-age peers. Among children whose scores fell below targets, on average scores declined 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 letters, whereas children whose scores did not meet targets learned an average of one letter. Thus, the targets appear to be successfully differentiating the children who make significant gains during the year from the group of children who need additional support.

### 2. Patterns of Progress Relative to Targets

A key question for LAUP is who was and who was not making progress. We conducted a series of subgroup analyses to shed some light on this issue. Note that home language was the only child or family background characteristic available for this analysis. Both quartile analyses based on fall scores and the analyses by children's home language by whether targets were met showed that the greatest progress occurred for children who began the year with the lowest scores. For the groups in the quartile analysis (top, middle, bottom), there was some indication that children whose fall scores were on the lower end of the distribution within their own group were more likely to meet targets (for example, among children in the lowest quartile, average fall scores were lower for those whose progress did meet targets than for children whose progress did not meet targets). Similarly, within each language group, children whose progress met the level identified in the targets started the year with lower average scores than children whose progress did not meet the targets.

Note that this pattern of progress—where children who initially had the lowest scores made the greatest gains—is not surprising. It is typically easier to support progress for the lowest performing children (assuming developmental disabilities were not present) than to maintain or raise the scores of children who enter programs performing above average. Children who perform at or above the national average in the fall and who maintain their position relative to the national mean were nonetheless making absolute progress. For children who fall further behind peers of the same age, additional support may be necessary.

### 3. Implications: Achieved Targets

As previously described, LAUP met the selected targets for language, fine motor/literacy, mathematics, and social-emotional development and approaches to learning. This indicates that

---

<sup>37</sup> The Leiter-R was excluded from this analysis. Because scores are truncated, we cannot assess the full range of progress among children whose scores met targets.

these targets are reasonable and achievable in LAUP. Our earlier discussion indicated the evidence of the meaningfulness of these targets for children. Overall, children who had greater needs at the beginning of the year (relative to those who were in the highest quartile) met the target, making significant progress as evidenced by the standard score gains.

#### 4. Implications: Missed Targets

LAUP did not meet targets in the area of literacy as measured by the RLN task. An analysis of overall progress (without attention to whether targets were met) indicates that children did in fact make statistically significant progress in this area. The target for this measure was based on data from UPCOS-2, which also included a representative sample of children in LAUP centers. In UPCOS-2, children progressed from knowing an average of 10 letters in the fall to an average of 18 letters in the spring. In UPCOS-5, children entered program knowing more letters in the fall (average of 13 letters) and knew an average of 18 in the spring. Note also that this level of progress was on par with that found in other large-scale preschool studies. For example, the Multi-State Study of Pre-Kindergarten and SWEEP (Early et al. 2005)<sup>38</sup> found that children progressed from knowing 9.1 letters in the fall to 14.4 in the spring, on average.<sup>39</sup> The distribution of the scores was not normal in either phase of UPCOS. In UPCOS-2, the distribution in the fall was very skewed; the median number of letters in the fall was only three, whereas in UPCOS-5 the fall median was nine. In each subsequent phase of UPCOS, children have entered LAUP with stronger letter-naming skills than what was found in the previous cohort. UPCOS-5 demonstrated the greatest increase in fall scores. This has implications for the change we are likely to detect. First, that children are entering with higher skills indicates that it will be harder to make changes of the magnitude identified in the 2011–2012 targets based on UPCOS-2 data. Second, unlike the other measures of language, literacy, and mathematics in the battery, the RLN Task has a ceiling; children cannot progress beyond a certain point because the RLN Task presents children with only 30 letters (a combination of upper- and lowercase). An examination of the distribution of RLN scores indicates that approximately 5 percent of children named all 30 letters presented in fall 2011. Prior knowledge of letter names also differed by language group. Children in the English only group entered with the highest mean score, correctly naming an average of about 17 letters in the fall, while children in the Spanish only and Spanish primarily groups knew about half as many in the fall. Given these circumstances, LAUP and First 5 LA should consider revisiting this target for subsequent years. Recent research (Justice et al. 2012) provides some guidance about the number of letters associated with later success in school. An alternative way to set targets might be to set an absolute number of letters expected and target the percentage of children in each language group expected to name that many letters.

Letter naming addresses only part of early literacy; literacy is comprehension of written language, so increases in vocabulary scores are also important for children's future success in reading. It will be important for LAUP to continue to introduce children to letters as building blocks to decoding and reading, but also important to continue to make greater progress in language and cognition.

---

<sup>38</sup> The Multi-State Study includes a random sample of state-funded prekindergarten programs in four states (Georgia, Illinois, Kentucky, and Ohio) and two large regions of two other states (Los Angeles and the Central Valley in California and New York City and Albany in New York); it began in 2001. The SWEEP study includes a random sample of state-funded prekindergarten programs in four states (Massachusetts, New Jersey, Washington, and Wisconsin) and a large region in Texas; it began in 2003. These studies focused on classrooms for 4-year-olds.

<sup>39</sup> The Multi-State Study/SWEEP version of the letter naming tasks included only 26 letters on each form, compared to 30 in the RLN task.

## **5. Conclusion**

Overall, children’s progress in LAUP during UPCOS-5 points to some clear successes but also room for growth. LAUP met the targets in four of the five domains identified in the performance-based contract: language, fine motor/literacy, mathematics, and social-emotional development and approaches to learning. The area in which the target was not met was literacy as measured by the RLN. However, the RLN Task has a ceiling, and the distribution of entering scores differs greatly across language groups. Each year children have entered LAUP with stronger letter-naming skills than in the previous cohorts. Recent research provides new guidance around rapid letter naming as an indicator of later success in first grade reading. These factors suggest that the RLN targets should be revisited for the 2012–2013 program year.

Overall, the evidence suggests that the targeted level of change was very meaningful. Across the measures, with the exception of the RLN, the specified percentage of children met the targets for the children in the lowest and middle quartiles. As anticipated when setting the targets, with the exception of mathematics in Spanish, the children in the top quartile typically made smaller gains than the other children<sup>40</sup> but on average were performing at or above the national mean in both fall and spring. Children assessed in Spanish at both time points continued to perform one standard deviation below the national mean in mathematics in the spring, suggesting a strong need for increased attention to helping ELLs develop number sense and reasoning skills. Looking ahead, LAUP and First 5 LA can use the information presented here to determine priorities and consider whether and how targets need to be updated to align with those priorities.

---

<sup>40</sup> All of the children in the top quartile on the fall Spanish mathematics assessment met the target.



## REFERENCES

- Brownell, R. "Expressive One-Word Picture Vocabulary Tests." San Antonio, TX: Harcourt Assessment, Inc., 2000.
- California Department of Education. "California Preschool Learning Foundations." Sacramento, CA: Author, 2008.
- Claessens, Amy, Greg Duncan, and Mimi Engel. "Kindergarten Skills and Fifth-Grade Achievement: Evidence from the ECLS-K." *Economics of Education Review*, vol. 28, no. 4, August 2009, pp. 415–427.
- De Temple, Jeanne and Catherine E Snow. "Learning Words from Books." In *On reading books to children: Parents and teachers*, edited by Anne van Kleeck, Steven A. Stahl, and Eurydice B. Bauer, Center for Improvement of Early Reading Achievement (CIERA). Mahwah, NJ: Lawrence Erlbaum Associates, 2003, pp. 16–36.
- Duncan, Greg J., Chantelle J. Dowsett, Amy Claessens, Katherine Magnuson, Aletha C. Huston, Pamela Klebanov, Linda S. Pagani, Leon Feinstein, Mimi Engel, Jeanne Brooks-Gunn, Holly Sexton, Kathryn Duckworth, and Crista Japel. "School Readiness and Later Achievement." *Developmental Psychology*, vol. 43, no. 6, November 2007, pp. 1428–1446.
- 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: NCEDE's Multi-State Study of Pre-Kindergarten and Study of State-Wide Early Education Programs (SWEEP)." Preliminary descriptive report, NCEDE working paper. Chapel Hill, NC: University of North Carolina, May 24, 2005.
- 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.
- Grissmer, David, Kevin J. Grimm, Sophie M. Aiyer, William M. Murrain, and Joel S. Steele. "Fine Motor Skills and Early Comprehension of the World: Two New School Readiness Indicators." *Developmental Psychology*, vol. 46, no. 5, September 2010, pp. 1008–1017.
- Hargrave, Anne C., and Monique Sénéchal. "A Book Reading Intervention with Preschool Children Who Have Limited Vocabularies: The Benefits of Regular Reading and Dialogic Reading." *Early Childhood Research Quarterly*, vol. 15, no. 1, 2000, pp. 75–90.
- Justice, Laura M., Joan N. Kaderavek, Xitao Fan, Amy Sofka, and Aileen Hunt. "Accelerating Preschoolers' Early Literacy Development Through Classroom-Based Teacher–Child Storybook Reading and Explicit Print Referencing." *Language, Speech, and Hearing Services in Schools*, vol.40, January 2009, pp. 67–85.

- 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.
- Roid, Gale H., and Lucy J. Miller. "Leiter-R Performance Scale—Revised." Wood Dale, IL: Stoelting Co., 1997.
- Atkins-Burnett, Sally, Kathy Sonnenfeld, and Susan Sprachman. "Rapid Letter Naming (RLN)." Washington, DC: Mathematica Policy Research, 2007.
- Pianta, R., K. La Paro, and B. Hamre. *Classroom Assessment Scoring System (CLASS) Manual*. Charlottesville, VA: Center for Advanced Study of Teaching and Learning, 2008.
- Piasta, S. B., Petscher, Y., and Justice, L. M. "How Many Letters Should Preschoolers in Public Programs Know? The Diagnostic Efficiency of Various Preschool Letter-Naming Benchmarks for Predicting First-Grade Literacy Achievement." *Journal of Educational Psychology*, April 2012. Advance online publication. doi: [10.1037/a0027757](https://doi.org/10.1037/a0027757)
- Winston, P., S. Atkins-Burnett, L. Akers, Y. Xue, E. Moiduddin, E. Smith, N. Aikens, J. Lyskawa, R. Mason, and S. Sprachman. "Quality Support Coaching in LAUP: Findings from the 2011–2012 Program Year." Final report submitted to First 5 LA. Washington, D.C.: Mathematica Policy Research, July 20, 2012.
- 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, 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, 2010.

**APPENDIX A**

**PERCENTAGE OF CHILDREN WHO SWITCHED LANGUAGE  
OF ASSESSMENT FROM FALL TO SPRING**



**Table A.1. Percentage of Children Who Were Assessed in English or Spanish in the Fall and Spring, by Language Group**

	Fall		Spring	
	Spanish	English	Spanish	English
English only	--	100	--	100
English primarily	31.6	68.4	8.2	91.8
Spanish only	86.7	13.3	61.6	38.9
Spanish primarily	84.8	15.2	42.3	57.7
Other language only or primarily <sup>a</sup>	--	100	--	100

Note: Analyses are weighted to represent children attending LAUP in the 2011–2012 program year.

<sup>a</sup>This table reflects only the children who actually completed the assessment. Because the assessment was only available in English or Spanish, children who spoke another language only or primarily are not reflected. For children in this group, 38 percent passed the language screener in the fall and 72 percent did so in the spring.



**APPENDIX B**

***pre*LAS FOR CHILDREN IN THE OTHER LANGUAGE ONLY OR PRIMARILY  
GROUP: FALL 2011 AND SPRING 2012**



**Table B.1. Means and Standard Errors of *preLAS* for Children in the Other Language Only or Primarily Group: Fall 2011 and Spring 2012**

Outcome	N	Fall 2011		Spring 2012		Fall-Spring Change	
		Mean	Standard Error	Mean	Standard Error	Mean	Standard Error
<b><i>preLAS</i> English</b>							
Total Language Screener Score	34	12.05	1.70	15.61	0.99	3.57**	0.81
Simon Says	34	4.97	1.21	6.87	0.84	1.90*	0.58
Art Show	34	7.07	0.54	8.74	0.21	1.67**	0.38

Source: UPCOS-5 Fall 2011 and Spring 2012 Direct Child Assessments.

Note: All analyses are weighted to represent children attending LAUP in the 2011–2012 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.



**APPENDIX C**

**FALL AND SPRING SCORES BY LANGUAGE GROUP AND PERFORMANCE  
RELATIVE TO TARGETS**



Table C.1 presents mean scores from the fall and spring by language group for children whose progress exceeded Level 1 targets; Table C.2 presents the same information for children whose progress did not exceed targets. Information is provided for five of the six measures: EOWPVT, RLN, WJ-III Spelling, WJ-III Applied Problems, and WM-III Problemas Aplicados. We conducted this analysis regardless of whether the overall target was met. The Leiter-R was excluded from this analysis because virtually all children scored at a level required by the 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 appendix were at the child level and were weighted to represent LAUP children as a whole. Analyses were conducted 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 correctly estimated.

**Language: EOWPVT.** Among children whose progress exceeded targets (Table C.1), 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 = 95.3), whereas scores for the English primarily group were somewhat lower (mean SS = 89.7). Based on the SBE version, children in the English only, English primarily, and Spanish primarily groups scored well above a national sample of peers in the spring, and children in the Spanish only group scored slightly below (mean SS = 96.9).

Table C.2 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 group lost ground relative to a national sample of peers between the fall and spring assessments. (Too few children in the English primarily and Spanish primarily 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.

Looking across Tables C.1 and C.2, 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 among those whose progress met targets than among those whose progress did not meet targets, with the exception of the Spanish primarily group.

**Literacy: Rapid Letter Naming.** For children whose scores did meet the target (Table C.1), progress for children in the Spanish primarily group was especially large (approximately 12 letters learned between the fall and spring). For children whose progress did not exceed targets (Table C.2), on average, letter naming abilities were flat over the course of the 5.6 months between assessments.

Looking across Tables C.1 and C.2, 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, that is,

within a language group, children whose change scores did and did not meet the targets had average spring scores within about 2 points of each other.

**Fine Motor and Literacy: WJ-III Spelling.** The third section of Table C.1 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 English only and English primarily groups (with sufficient number of children) began the year with scores similar to a national sample of same-age peers (average standard scores were near or above 100) and ended the year with scores exceeding a national sample of peers (mean SS = 107.5 and 108.3 for English only and English primarily group, respectively).

Table C.2 shows that children in both groups who did not meet the targets lost ground relative to a national sample of peers. They began the year with scores well above a national sample of peers (mean SS = 111.7 and 111.4 for English only and English primarily groups, respectively) but ended the year with scores near the national mean (mean SS = 98.6 and 98.9, respectively).

Looking across Tables C.1 and C.2, we see that average scores in the fall were lower in both language groups 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 C.1 shows mathematics scores by language group for the sample whose progress exceeded targets. Both language groups with 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 = 103.5 and 104.0 for English only and English primarily groups, respectively).

When examining scores for the group whose scores did not meet the target by language (Table C.2), for both English only and English primarily groups, children entered their programs with similar skills in the fall (average standard scores were above 100 for both groups) but lost ground relative to their peers in the spring. The decline in scores between the fall and spring was similar for both groups (mean SS = 90.7 and 93.1 for English only and English primarily, respectively, in the spring).

Looking across Tables C.1 and C.2, we see that average mathematics scores in the fall were lower in the English only and English primarily groups for the children whose change scores met the Level 1 target than for children whose change scores did not meet the target. By the spring, scores were higher for these two language groups among children whose progress did meet the target.

**Mathematics: WM-III Problemas Aplicados.** Shifting to the bottom section of Table C.1, we see that children in the Spanish only and primarily groups whose scores exceeded targets made progress relative to a national sample of peers in mathematics. However, scores for both groups were still two-thirds to one 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 C.2 is not discussed here.

**Table C.1. Means and Standard Errors for UPCOS-5 Language, Literacy, and Math Standardized Measures: Fall 2011 and Spring 2012 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 2011		Spring 2012		Fall 2011		Spring 2012		Fall 2011		Spring 2012		Fall 2011		Spring 2012	
	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE
<b>Language: Expressive One Word-Picture Vocabulary Test</b>																
IRT Score	50.12	1.55	57.14	1.36	40.76	1.64	50.62	0.99	36.32	2.12	43.81	1.78	38.90	0.93	47.91	0.76
English Edition Standard Score <sup>a</sup>	89.65	2.29	95.26	2.04	83.51	1.82	89.69	1.67	Sample < 10				Sample < 10			
Spanish Bilingual Edition Standard Score <sup>a,b</sup>	115.23	2.60	122.78	2.30	98.35	2.57	110.20	2.04	87.83	4.77	96.87	3.76	94.18	2.02	107.38	1.49
<b>Literacy: Rapid Letter Naming</b>																
Raw Score	11.38	1.26	20.94	1.13	6.35	1.11	16.95	1.47	3.08	1.08	11.64	2.02	5.11	0.92	17.24	1.67
IRT Score	21.36	1.54	33.28	1.34	15.15	1.59	28.60	1.51	9.72	2.00	23.53	1.93	13.41	1.57	28.83	1.62
<b>Fine Motor and Literacy: Woodcock-Johnson III Spelling</b>																
W Score	373.62	3.44	399.85	2.50	378.94	3.47	404.14	2.85	Sample < 10				Sample < 10			
Standard Score <sup>a</sup>	100.72	1.89	107.47	1.45	102.27	1.84	108.32	1.79	Sample < 10				Sample < 10			
<b>Mathematics: Woodcock-Johnson III Applied Problems</b>																
W Score	393.37	2.89	412.14	1.70	395.14	2.62	413.84	2.49	Sample < 10				Sample < 10			
Standard Score <sup>a</sup>	98.64	1.62	103.54	1.19	98.75	1.47	103.98	1.62	Sample < 10				Sample < 10			
<b>Mathematics: Woodcock-Muñoz III Problemas Aplicados</b>																
W Score	Not Applicable				Sample < 10				351.00	8.70	379.16	4.58	350.50	4.68	384.55	3.37
Standard Score <sup>a</sup>	Not Applicable				Sample < 10				78.16	3.91	85.37	1.96	79.79	1.87	90.02	1.40
N (range)	49-78				32-63				11-23				18-38			

Source: UPCOS-5 Fall 2011 and Spring 2012 Direct Child Assessments.

Note: All analyses are weighted to represent children attending LAUP in the 2011-2012 program year.

<sup>a</sup>National mean for standard scores is 100 with a standard deviation of 15.

<sup>b</sup>This measure was conceptually scored, and the standard score was generated for all children in the sample.

**Table C.2. Means and Standard Errors for UPCOS-5 Language, Literacy, and Math Standardized Measures: Fall 2011 and Spring 2012 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 2011		Spring 2012		Fall 2011		Spring 2012		Fall 2011		Spring 2012		Fall 2011		Spring 2012	
	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE
<b>Language: Expressive One-Word Picture Vocabulary Test</b>																
IRT Score	54.38	2.71	53.24	2.31	45.26	3.07	42.44	3.37	Sample < 10		42.15	2.46	41.14	2.35		
English Edition Standard Score <sup>a</sup>	97.74	4.29	91.33	3.43	Sample < 10				Not Applicable				Sample < 10			
Spanish Bilingual Edition Standard Score <sup>a, b</sup>	126.62	3.93	117.58	3.99	109.85	7.12	99.44	5.80	Sample < 10		101.77	5.64	92.20	5.41		
<b>Literacy: Rapid Letter Naming</b>																
Raw Score	21.83	1.40	22.62	1.33	16.48	1.52	17.08	1.53	8.78	3.07	11.04	3.27	14.19	1.53	16.07	1.49
IRT Score	33.32	1.75	33.94	1.57	27.63	1.96	27.78	2.01	18.19	3.75	21.38	3.86	24.72	2.02	27.11	1.89
<b>Fine Motor and Literacy: Woodcock-Johnson III Spelling</b>																
W Score	392.77	6.73	382.39	9.58	394.96	3.20	386.59	3.50	Not Applicable				Sample < 10			
Standard Score <sup>a</sup>	111.66	3.54	98.61	5.24	111.38	1.92	98.92	2.42	Not Applicable				Sample < 10			
<b>Mathematics: Woodcock-Johnson III Applied Problems</b>																
W Score	395.47	4.92	386.60	6.12	403.39	5.39	396.44	5.67	Sample < 10				Sample < 10			
Standard Score <sup>a</sup>	100.68	2.34	90.74	2.65	103.57	2.94	93.13	2.62	Sample < 10				Sample < 10			
<b>Mathematics: Woodcock-Muñoz III Problemas Aplicados</b>																
W Score	Not Applicable				Not Applicable				Sample < 10				Sample < 10			
Standard Score <sup>a</sup>	Not Applicable				Not Applicable				Sample < 10				Sample < 10			
N (range)	22-51				15-46				14				19-20			

Source: UPCOS-5 Fall 2011 and Spring 2012 Direct Child Assessments.

Note: All analyses are weighted to represent children attending LAUP in the 2011–2012 program year.

<sup>a</sup>National mean for standard scores is 100 with a standard deviation of 15.

<sup>b</sup>This measure was conceptually scored, and the standard score was generated for all children in the sample.

## UPCOS-5 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.

*Bateria 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.



---

# **MATHEMATICA** **Policy Research**

[www.mathematica-mpr.com](http://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