LESSONS FROM THE UNIVERSAL PRESCHOOL CHILD OUTCOMES STUDY

Looking at Classroom and Program Quality Through Multiple Lenses:
First 5 LA is using multiple strategies to improve children’s access to affordable, high quality early care and education (ECE) in Los Angeles (L.A.) County. These strategies include supporting the development of a uniform quality rating and improvement system (QRIS). A QRIS is an approach to measuring the quality of ECE programs. A uniform QRIS in L.A. County would enable parents to make more informed decisions about their children’s care and would help First 5 LA and other stakeholders to identify the improvements that are needed in ECE programs.

To inform First 5 LA’s efforts in this area, this brief presents several findings from the Universal Preschool Child Outcomes Study (UPCOS), a multiyear study of Los Angeles Universal Preschool (LAUP) programs. These analyses drew on data collected through UPCOS and LAUP’s own administrative data. The findings are particularly valuable because they are specific to L.A. County and include not only center-based programs but also family child care (FCC) homes. Though all the analyses are descriptive in nature, we observed some important associations. In this brief, we present findings that shed light on key considerations for practitioners and policymakers as they design and implement a QRIS in L.A. County.

### Finding Key Considerations

<table>
<thead>
<tr>
<th>Finding</th>
<th>Key Considerations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teachers are the key to quality.</td>
<td>Findings based on both UPCOS and LAUP data confirm that observed quality tends to be stronger when lead teachers have more education and higher levels of teaching permits. While we focused on lead teachers in our analyses of the relationship between permit levels and education, the assistant teachers and aides in a classroom contribute to quality as well. <em>This suggests the importance of taking the qualifications and contributions of all classroom teachers into account in a QRIS.</em></td>
</tr>
<tr>
<td>Classroom quality needs to be better than “good enough.”</td>
<td>Using UPCOS data, we found that observed classroom quality that surpasses a certain threshold is associated with better outcomes for children. When we analyzed LAUP data, however, we saw few associations between quality ratings and children’s outcomes, although in some cases, those associations were observed in the higher ranges of program quality ratings. <em>These findings argue for an alternative to measuring just the average quality of all classrooms in a program. A QRIS rating could instead (or also) account for the proportion of classrooms in a program that attain a certain level of quality.</em></td>
</tr>
<tr>
<td>All the children in a classroom do not experience or benefit from aspects of the classroom’s quality in the same way.</td>
<td>Analyses showed that children with English as their primary language were spoken to more frequently, on average, than Spanish-speaking dual-language learners (DLLs) were. In addition, DLLs benefited especially from stronger emotional support, rather than stronger instructional support. <em>Professional development provided through a QRIS could consider the population that a program serves and support goals aligned with child needs.</em></td>
</tr>
<tr>
<td>Quality ratings appear to be appropriate for programs with different characteristics.</td>
<td>We found few associations between program characteristics (center versus FCC, funding source) and quality. <em>These findings suggest that these indicators are appropriate in different types of settings.</em></td>
</tr>
</tbody>
</table>
First 5 LA’s 2015–2020 strategic plan lays out a path to strengthen families and improve outcomes for the greatest possible number of young children in L.A. County. Improving access to affordable, quality early care and education (ECE) for the children of L.A. County is a central goal of the strategic plan because the quality of services offered to children and families in early childhood settings makes a difference in children’s future success.\(^1\) First 5 LA is using multiple strategies to achieve this goal. One strategy provides support for the development of a uniform quality rating and improvement system (QRIS) in L.A. County. A QRIS is an approach to measuring the quality of ECE programs. A uniform QRIS in L.A. County would enable parents to make more informed decisions about their children’s care and would help First 5 LA and other stakeholders identify the improvements that are needed in ECE programs.

To inform First 5 LA’s efforts, we present several findings from the Universal Preschool Child Outcomes Study (UPCOS), a multiyear study of Los Angeles Universal Preschool (LAUP) programs that was conducted from 2007 - 2015. LAUP is a nonprofit organization funded by First 5 LA; it has provided high quality preschool in L.A. County since 2005. Between 2005 and 2015, LAUP programs served over one hundred thousand children and families who reflect the cultural and linguistic diversity of L.A. County. Note that while LAUP no longer provides preschool slots as of July 2016, they will continue their mission of advancing early education program quality and capacity by supporting development of the whole child, growing a qualified and diverse workforce, strengthening family engagement and advocating for policies that promote access and program excellence.

Mathematica Policy Research conducted UPCOS with First 5 LA to learn more about the children and families of LAUP and the programs that serve them. The findings presented in this brief are especially valuable because they are specific to L.A. County and include not only center-based programs, but also family child care (FCC) homes. FCCs serve a substantial proportion of L.A. County children\(^2\) yet are often excluded from studies on the quality of classrooms and programs.

In the rest of this brief, we present analyses that shed light on key considerations for practitioners and policymakers as they design and implement a uniform QRIS in L.A. County. Table 1 shows the rich set of data from UPCOS and LAUP that were used for these analyses, and it identifies those data elements from UPCOS that were linked to LAUP data. Table 2 shows the observed classroom quality measures available from UPCOS and/or LAUP, and Box 1 includes details on the samples and limitations of the analyses. Notably, LAUP’s administrative data include ratings from its own QRIS—the 5-Star Quality Assessment and Improvement System. The Star rating has four components: (1) qualifications of the lead teacher, (2) qualifications of other teaching staff, (3) ECERS-R or FCCERS-R scores, and (4) CLASS scores. Note that the Star rating is no longer in use as of the 2015-16 program year. LAUP now uses the Quality Continuum Framework Rating Matrix to conduct quality ratings of its programs.

### Table 1. Information used from UPCOS and LAUP

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Star ratings</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Observed classroom quality</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Teacher education and/or permits</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Children’s and families’ characteristics</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Children’s outcomes</td>
<td>X</td>
<td>X</td>
<td>Linked UPCOS data(^c)</td>
</tr>
</tbody>
</table>

\(^a\)Additional data are available in both UPCOS-2 and 3. This table focuses on data used for this brief.

\(^b\)Not all information from LAUP is available in all program years.

\(^c\)UPCOS data on children’s outcomes are available for 2007–2008, and for the 2009-2010 through 2013-2014 program years.

### Table 2. Observed classroom quality measures used from UPCOS and LAUP\(^d\)

<table>
<thead>
<tr>
<th>Measure: source</th>
<th>Areas assessed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Classroom Assessment Scoring System (CLASS): available from UPCOS-2, UPCOS-3, and LAUP</td>
<td>Measures teachers’ interactions with children in three domains: Emotional Support (the social and emotional functioning in the setting), Classroom Organization (the teachers’ ability to organize the environment to make efficient use of time), and Instructional Support (the quality of instructional practices used in the setting)</td>
</tr>
<tr>
<td>Early Childhood Environment Rating Scale-Revised (ECERS-R)/ Family Child Care Environment Rating Scale-Revised (FCCERS-R): available from LAUP</td>
<td>Measures structural features of the environment and process quality (interactions among staff, children, parents and other adults in a classroom; children’s interactions with materials and activities in the environment)</td>
</tr>
<tr>
<td>Language Interaction Snapshot (LISn): available from UPCOS-3</td>
<td>Measures the support for language development in the classroom, including languages used in the classroom by teachers and children and instructional strategies for supporting language</td>
</tr>
</tbody>
</table>
Box 1. Overview of UPCOS and LAUP samples and study limitations

This is a brief description of the UPCOS-2, UPCOS-3, and LAUP samples and study limitations. Extensive detail on these topics and the analysis approach is available in the study reports.6

- UPCOS-2 includes a sample of 88 center-based classrooms and 1,555 children representative of LAUP programs.

- UPCOS-3 includes a sample of 72 center-based classrooms and 52 FCCs representative of LAUP. LISn data were available for a representative sample of 38 center-based classrooms and 26 FCCs. The study included 854 children, but the child-level sample was not representative of LAUP programs.

- LAUP administrative data include all classrooms in the network in a given year, but not all data are available for all classrooms in each year. Star rating, CLASS, and ECERS-R data were available for 827 to 992 classrooms over time. FCCERS-R data were available for only 141 classrooms over time. Depending on the measure and program year, child outcome data from UPCOS were available for 612 to 807 children.

Study limitations. The analyses in this report are descriptive and the findings should not be interpreted as demonstrating that one thing caused another. Notably, LAUP data were collected for the purpose of monitoring and not for research, and these data had numerous issues, most notably large amounts of missing data (particularly on teacher education and permits). For analyses examining associations between UPCOS data on children’s outcomes and LAUP data on quality, we had to aggregate data to a program level because it was difficult to match data to specific classrooms. Associations between the factors we studied in the LAUP data may have been obscured by these and other issues. For both the UPCOS and LAUP data, it is possible that some of the associations we did observe in the data happened due to chance. In addition, DLL research findings are specific to children whose home language was Spanish. Children from other language backgrounds were not included in these analyses because the assessment tools were only available in English and Spanish.

Teachers are key to quality

When teachers had more advanced qualifications, classroom quality was stronger.

In the analyses based on UPCOS and LAUP data, our findings confirm that program and classroom quality tend to be stronger when teachers have more education and higher levels of permits (Box 2). In UPCOS-2, we found that the teacher’s educational level was associated with CLASS Emotional Support and Instructional Support ratings. Using LAUP administrative data, we found that teacher education and permit level were associated with both Star ratings and observed quality.

In the LAUP data, the teacher’s educational level was related to CLASS scores and overall Star ratings, but not to ECERS-R scores. Given that many of the ECERS-R items are not under the control of the teacher (for example, the availability of resources, the condition of materials, and the environment), there simply may not have been a strong association between ECERS-R scores and a teacher’s educational level. The teacher’s permit level was related to the ECERS-R/FCCERS-R scores, however. Specifically, teachers with a master teacher, program director or site supervisor permit had higher ECERS-R and FCCERS-R scores. This suggests that some of the differences in ERS scores may be related to teachers’ awareness and understanding of important indicators of environmental quality. This is not surprising given that teachers with master teacher, program director or site supervisor permits may have more influence over environmental quality.

Box 2. Associations between teacher education and permit and classroom or program quality based on UPCOS-2 and LAUP administrative data

Teacher educational level

- Teachers holding a B.A. (rather than an A.A.) had higher CLASS Emotional Support Scores.
- Teachers holding an A.A. had higher CLASS Instructional Support Scores than teachers with lower levels of education did.
- Teachers holding a B.A. (rather than an A.A.) had higher Star ratings, FCCERS-R scores, and CLASS scores in all three domains (Emotional Support, Classroom Organization, and Instructional Support).

Teacher permit level

- Teachers holding a master teacher, program director, or site supervisor permit had higher Star ratings, ECERS-R scores, and FCCERS-R scores.

Note: All associations described in the table were statistically significant at the p<.05 level.
The quality of a classroom depends on all the teachers in the classroom.

In the preceding analyses on the relationship between teacher education and permits and quality ratings, we relied on information about lead teachers. However, a classroom’s quality depends on all the teachers in the classroom. For example, CLASS scores are based on the average experience in the classroom, and for the ECERS-R, the behavior of all adults in the classroom is noted.

In UPCOS-2 and 3, we used the LISn data to examine language interactions of individual children with adults (and peers) in the classroom; language use by lead teachers and other adults or teachers in the classroom was rated separately. We found differences in how teachers in the same classroom interacted verbally with children. Overall, verbal interaction with children was limited, but in the classroom, lead teachers talked more to children and provided more types of verbal interactions compared to other adults in the classroom.

Classroom quality needs to be better than “good enough.”

Classroom quality that surpassed a certain threshold was associated with better outcomes.

Recent meta-analyses point to consistent, yet modest relationships between observed classroom quality and outcomes for children. The findings of these studies suggest that a certain threshold of quality must be reached before classroom interactions can influence outcomes for children. In other words, the evidence suggests that lower levels of quality are not associated with growth in children’s development. Classroom quality must reach at least a level of low-moderate (“good enough”) quality before children begin to benefit from their preschool experiences. Once that threshold has been met, the higher a classroom’s quality, the greater the children’s outcomes.

Analyses based on UPCOS-3 and LAUP administrative data have findings that are consistent with the findings of these studies. UPCOS-3 analyses focused on associations between the CLASS domains and children’s outcomes in two areas: expressive vocabulary and executive functioning. We found that low levels of classroom quality did not support children’s development. Further, associations between classroom quality and children’s outcomes were stronger in higher quality classrooms than in lower quality classrooms. Once classrooms reached a certain threshold of quality, children continued to benefit more as classroom quality increased.

In analyses based on LAUP data, however, we found very few associations between quality (Star ratings, CLASS and ECERS-R scores) and children’s outcomes. This is not surprising given the limitations of the analysis of the administrative data (Box 1). In some cases, those associations were observed in the higher ranges of program quality ratings.

All the children in a classroom do not experience or benefit from aspects of the classroom’s quality in the same way.

Not all children experienced the same levels of quality in the classroom.

Hart and Risley wrote that the “most important aspect to evaluate in child care settings for very young children is the amount of talk actually going on, moment by moment between children and their caregivers” (p. xxi). Through the LISn, we examined the variability of verbal interactions experienced by different children in the same classroom. For each child, we recorded and counted the number of language interactions with all adults in the classroom during the observation period. The number of language interactions ranged from zero (9.3 percent of children) to 81 (for one child). Children with English as their primary language were spoken to more frequently, on average, than Spanish-speaking DLLs were.

In addition to the number of interactions, the sophistication of the language adults used also differed across children. For example, the adults in the classroom were less likely to speak with Spanish-speaking DLLs about things or events that took place outside of the classroom (for example, what they did at home that morning or a field trip they had taken in the past) than they were with children who primarily spoke English. Such “decontextualized” language interactions during preschool are important for literacy (including reading comprehension) during elementary school; understanding decontextualized language is the essence of school literacy since academic text often communicates about things not actually present.
Emotional support in a classroom is important for all children, but especially important for dual language learners.

Analyses based on UPCOS data show that progress for DLLs appears to be greatly influenced—perhaps more so than for other children—by the level of emotional support in a classroom. Figure 1 broke out the association between CLASS Emotional Support and expressive vocabulary separately for DLL children. At the higher range of emotional support, the expressive vocabulary of DLLs improved rather quickly. For DLLs, in contrast with the group of LAUP children as a whole, the association between Emotional Support and expressive vocabulary is present even in classrooms with quality in the lower range, but the associations are stronger in the higher range; that is, the difference between the association in the higher and lower range is statistically significant.

UPCOS analyses show other differences for DLLs as well. For their expressive vocabulary, the association with scores on CLASS Classroom Organization was significant in classrooms scoring in both the low and high range of quality (that is, above and below the threshold). Scores on Instructional Support were not associated with any outcomes for DLLs, though there is an association between the level of CLASS Instructional Support and the outcomes of LAUP children as a whole. This suggests that the instructional needs of DLLs may be different from those of non-DLLs.

Quality ratings appear to be appropriate for programs with different characteristics.

A uniform QRIS in L.A. County would likely include programs of different types (center-based versus FCC) that have different funding sources. We looked at the associations between program characteristics and program quality (both Star ratings and quality observations) to better understand whether these characteristics matter in terms of how we measure quality. Using LAUP data, we observed few associations between program characteristics and program quality.

Using UPCOS data, we also examined CLASS scores for centers and FCCs. Although the CLASS does not rely on normative data, the CLASS Technical Appendix reports mean scores from several large-scale studies that used the CLASS. Classroom quality scores in LAUP were, for the most part, on a par with or higher than those in other studies in both centers and FCCs. Consistent with other studies, Instructional Support scores were the lowest of any in the three domains (see Figure 2).

These findings suggest that these indicators are appropriate in different types of settings. However, it is also important to keep in mind that LAUP offers coaching in support of quality (along with other resources) to the programs in its network, both centers and FCCs. In L.A. County as a whole, such supports may be more readily available to center-based programs.
Summary and implications

In this brief, we presented the lessons learned in our analyses of UPCOS and LAUP data. These lessons shed light on the factors that contribute to classroom and program quality and the ways that different children can benefit from quality. Though all analyses are descriptive in nature, we observed some important associations.

First, teachers are the key to quality. Our analyses, which are based on UPCOS and LAUP data, confirm that quality tends to be stronger when lead teachers have more years of education and higher levels of teaching permits. Although our analyses on education and permit levels focused on lead teachers, other teachers and adults in a classroom contribute to quality as well. For example, we found differences in the way that teachers in the same classroom interact verbally with the children in the classroom. Together, these findings highlight the potential importance of taking the qualifications and contributions of all teachers, not just the lead teacher, into account when designing a QRIS.

Second, using UPCOS data we found that once classroom quality surpasses a certain minimum threshold, it is associated with better outcomes for children. In contrast, our analysis of LAUP data uncovered few associations between quality ratings and children’s outcomes; in some cases, those associations were observed in the higher ranges of program quality ratings. The lack of associations in the LAUP data is likely due at least in part to limitations in the data; notably, we had to use program-level averages for quality (rather than the classroom-level rating). These findings argue for an alternative to measuring just the average quality of all classrooms in a program. A QRIS rating could instead (or also) account for the proportion of classrooms in a program that attain a certain level of quality.

Third, we found that children in the same classroom do not experience or benefit from quality in the same way. For example, UPCOS observations revealed that children whose only language was English were spoken to by teachers and other adults in the classroom more frequently, on average, than Spanish-speaking DLL children were. Analyses also showed that DLLs may benefit from different aspects of quality in different ways than other children do; they especially benefited from stronger emotional support. Professional development provided through a QRIS could consider the population that a program serves and support goals aligned with child needs.

Finally, we found that there were few associations between program characteristics (center versus FCC, funding source) and quality. Because First 5 LA and other stakeholders are working toward a uniform QRIS, it is important that these quality ratings appear to be appropriate for different kinds of programs in L.A. County. A uniform QRIS for L.A. County would help parents as they make decisions about child care and preschool.

Figure 2. Average CLASS scores in LAUP center-based classroom and FCCs

Source: UPCOS-2 spring classroom observation; UPCOS-3 winter classroom and FCC observation.
Note: Analyses are weighted to represent LAUP classrooms in spring 2008 and winter 2010, respectively.


6 Atkins-Burnett et al. 2010.


9 Analyses focused on expressive vocabulary and executive functioning, because children whose primary language was English or Spanish completed assessments in these two areas. Therefore, these assessments were available for the maximum number of children in the sample.

10 The thresholds were developed empirically and were set at 5.5, 5, and 3 for Emotional Support, Classroom Organization, and Instructional Support, respectively.


15 Pianta et al. 2008.