



Systemwide Changes in the Quality of Early Childhood Education: Trends in Louisiana from 2015-16 to 2018-19

Daphna Bassok, Preston Magouirk, Anna Markowitz

Overview

High quality early care and education (ECE) can improve children’s school readiness and support their lifelong wellbeing (e.g. Campbell et al., 2002; Deming, 2009; Phillips et al., 2017). Children benefit from ECE settings where they learn about new things, interact with peers, and practice the skills they will need to be successful in school. Unfortunately, many of the available ECE sites in the United States are not of sufficient quality for children to experience long-term benefits (e.g. Burchinal, 2017; Friedman-Krauss & Barnett, 2016; Helburn, 1995). This is particularly true for low-income children served in publicly-funded settings—the very children for whom high-quality ECE may be the most beneficial.

In 2012, the Louisiana Department of Education (LDOE) began an ambitious effort to ensure high-quality ECE options for all children statewide. With the passage of the Early Childhood Education Act (known as Act 3), LDOE began uniting school-based pre-kindergarten (pre-K),¹ Head Start, and child care sites receiving public dollars into a single, more cohesive system.

At the heart of this effort was a unique Quality Rating and Improvement System (QRIS). QRIS are accountability systems that define and measure ECE quality, incentivize sites to meet quality standards, and publicize ratings to help parents identify high quality options. Louisiana’s QRIS is unique in that participation is required for all sites receiving public funds and that ratings for each site are based solely on scores from a widely used observational measure of teacher-child interactions known as the Classroom Assessment Scoring System, or CLASS (Pianta, La Paro, & Hamre, 2008). Since the 2015-16 school year, Louisiana has required that at least two CLASS observations are conducted per year in every ECE classroom serving publicly-funded toddlers (one- and two-year-olds) and preschool-aged children (three- to five-year-olds) in the state.

This report uses CLASS data from the last four years (2015-16 to 2018-19) to explore the extent to which Louisiana has been successful in improving the quality of ECE available to children and families statewide. The report presents trends in CLASS scores and ratings (1) statewide, (2) by community network, (3) by site type (i.e. school-based pre-K, Head Start, and child care), and (4) by age (toddler versus pre-K).

¹ Schools-based public pre-kindergarten includes both public and nonpublic schools.

The findings indicate that, on average, CLASS scores and ratings in Louisiana’s publicly-funded ECE sites improved significantly between 2015-16 and 2018-19. However, patterns of improvement varied across the state; several communities demonstrated rapid increases in CLASS performance while in others, quality was stagnant. Trends also varied across site types, as child care sites, which generally earned lower scores than Head Start or school sites, experienced the largest increases in scores between 2015-16 and 2018-19. Finally, CLASS scores and ratings improved in both toddler and pre-K classrooms. The report closes with a discussion of the implications of these improvements for children’s school readiness and for the future of ECE in Louisiana.

Louisiana’s Efforts to Improve Quality Statewide

In most states, QRIS ratings are based on a large set of factors, including structural quality features (e.g. class size, teacher credentials), classroom observations (e.g. CLASS or Early Childhood Environment Rating Scales (Harms, Clifford, & Cryer, 2014)), and a host of other measures (e.g. family engagement checklists, administration and business practices, measures of curriculum and assessment use, etc.). These systems were designed to help ECE sites improve incrementally across a variety of features. Unfortunately, research suggests that many of these specific factors are not strong predictors of children’s learning (Early et al., 2007; Mashburn et al., 2008). Moreover, recent research shows that when the multiple measures in QRIS are combined to form site ratings, these ratings are not strongly related to child outcomes (Cannon et al., 2017). In other words, a site that is rated a 5 on a 5-point scale is not necessarily doing a better job supporting children’s learning than one rated a 2 or a 3.

When Louisiana was designing their QRIS, they wanted to ensure that the quality features that were measured, publicized, and incentivized were ones that mattered for child outcomes. The quality of teacher-child interactions has been linked to children’s development across hundreds of studies. Young children learn and grow primarily through their interactions with caring adults (e.g. Burchinal, 2017; Hamre, 2014). CLASS is a widely used measure of teacher-child interactions that has been repeatedly, albeit modestly, linked to young children’s learning in ECE classrooms across the United States, and in Louisiana specifically (Burchinal, 2017; Hamre, 2014; Vitiello et al., 2018). In an effort to capture the quality of those interactions statewide, Louisiana decided to base their QRIS ratings on the CLASS.

All publicly-funded classrooms are observed at least twice a year using the CLASS. These observations are then aggregated into a site-level rating. The ratings, in turn, are linked to a variety of supports and resources for sites.

In addition to basing its QRIS solely on CLASS ratings, Louisiana's approach to early childhood quality improvement is different from other states' in that Act 3 required universal participation from all publicly-funded ECE sites, including all school-based pre-K, Head Start, and subsidized child care sites. In most states, QRIS participation is voluntary or sites can opt to participate only superficially. In contrast, in Louisiana CLASS observations are conducted, and scores are collected and publicized, for every publicly-funded ECE site in the state.

What CLASS Measures

The CLASS is an observational measure in which a trained observer spends 120 minutes in a classroom watching teachers interact with children, taking notes, and ultimately scoring the quality of the interactions they observe. It measures the warmth and positivity of the classroom, and the existence of classroom organizational structures, like schedules and routines, that help children behave and succeed. It also captures whether teachers use complex language, and whether they scaffold children's learning through questions and conversations that foster learning. The tool is designed to align with children's developmental stages, thus raters assessing pre-K classrooms and classrooms serving toddlers score on slightly different scales. In the Toddler CLASS, classrooms are assessed using the Engaged Supports for Learning and Emotional and Behavioral Support domains.² In the Pre-K CLASS, assessors score classrooms on three domains: Instructional Support, Emotional Support, and Classroom Organization.³

² The Engaged Supports for Learning domain measures support for language development and academic concept development. Observers assess the facilitation of learning and development, the quality of feedback, and language modeling. The Emotional and Behavioral Support domain measures the warmth and sensitivity of the classroom, and how teachers manage challenging situations. Observers assess positive and negative climate, regard for child perspectives, behavior guidance, and teacher sensitivity.

³ Instructional Support measures support for language development and academic concept development. Observers assess concept development, quality of feedback, and language modeling. Emotional support measures the warmth and sensitivity of the classroom. Observers assess positive and negative climate, teacher sensitivity, and regard for student perspectives. Classroom Organization measures how the classroom is set up to facilitate learning, including routines, and how teachers handle challenging situations. Observers assess behavior management, productivity, and instructional learning formats.

Data

Each year, local raters conduct two CLASS observations in every classroom in every publicly-funded site in the state, and upload those scores to a centralized state portal. This report uses CLASS data from the four years for which state-wide data are currently available: 2015-16, 2016-17, 2017-18, and 2018-19. The data include overall and domain-level CLASS scores, as well as ratings based on these scores, from a total of 1,871 unique sites. The number of publicly-funded sites with data varied across years from 1,640 in 2015-16 to 1,522 in 2018-19.

Method

This section describes how CLASS assessments are conducted in Louisiana, and outlines the analyses conducted for this report.

CLASS Observation Protocols

As part of Act 3, Louisiana created 64 Early Childhood Networks to oversee ECE activities locally. These networks typically encompass a parish (a geographic unit similar to a county). The lead agency in each network (typically, though not always, the parish school board) manages local CLASS data collection. Procedures for training and ongoing calibration vary from network to network, however observers must be trained and certified in the CLASS. To become certified, observers must demonstrate that their coding is aligned with master coders. Observers are typically school principals, site directors, or other individuals with a history of working in ECE in the network. Most networks “double code” a portion of their observations, sending two CLASS observers into a classroom to ensure they reach consensus on scores. In addition, to ensure the validity and reliability of local ratings, the LDOE also contracts “third party” observers to conduct CLASS observations in half of classrooms; scores from these observations are compared to local raters’ scores for the same classrooms, and replace local ratings if the scores are meaningfully divergent.

CLASS observations consist of four cycles in which observers take notes in the classroom for 15-20 minutes, and then score the classroom using the Toddler or Pre-K CLASS rubric. Scores for each domain range from 1-7. Domain scores are averaged across all four cycles, and the final CLASS observation score is the average across all domains.

A site's QRIS score is the average of all overall CLASS scores from all classroom observations conducted over the course of the year. This report uses the final overall and domain-level scores, including replacement by third party raters as applicable. It also presents trends in the site ratings that correspond with these scores.

Analyzing CLASS Growth

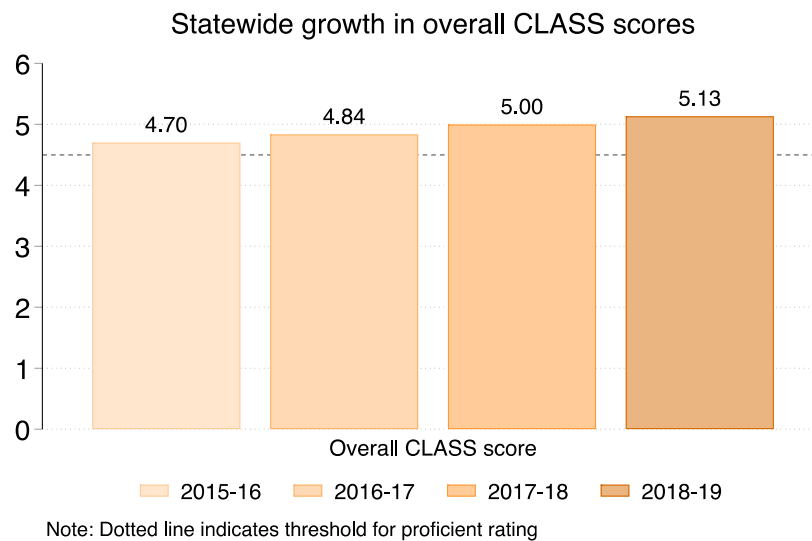
The goal of this report is to describe changes in the quality of ECE available to Louisiana children since 2015. To do this, the report examines two measures of growth: (1) change in average CLASS scores between 2015-16 and 2018-19 and (2) change in the percentage of sites rated as proficient or higher on Louisiana's QRIS. Since 2015-16, Louisiana has classified each site's overall quality based on where their average CLASS score fell relative to a set of cut-off scores. Sites earned ratings of "Unsatisfactory" for scores below 3 points out of a possible 7. They earned "Approaching Proficient" ratings for scores at or above 3 but below 4.5. "Proficient" sites earned total scores at or above 4.5 but below 6, which was the threshold for "Excellent" ratings. Beginning in 2018-19, Louisiana assigned the rating of "High Proficient" to sites achieving ratings at or above 5.5 overall. Note that this did not change the cutoff score for reaching proficiency (4.5).

The report presents trends at the state level, as well as by network, site type, and domains specific to age groups (e.g., toddlers or preschoolers, see Appendix Tables A1-A6).

CLASS Growth from 2015-2019

As shown in Figure 1 below (as well as in Appendix Table A1), the average statewide CLASS score in 2015-16 was 4.70. Statewide, average CLASS scores increased each year and reached 5.13 in 2018-19.

Figure 1.



The overall increase in CLASS reflects growth across all domains considered—that is, sites became more emotionally warm, well-structured places, and more supportive of children’s academic development. Appendix Tables A2-A6 show trends disaggregated by domain.

In 2015-16, 62% of publicly-funded sites in Louisiana achieved ratings of “Proficient” or above (in 2015-16 this was “Proficient” or “Excellent”). Figure 2 indicates that statewide site proficiency grew to 85% by 2018-19. In 2018-19, this 85% included sites earning “Proficient” (40%), “Excellent” (7%), and a new category, “High Proficient” (38%). As noted above, the addition of the “High Proficient” category did not change the threshold score for proficiency.

Figure 2.

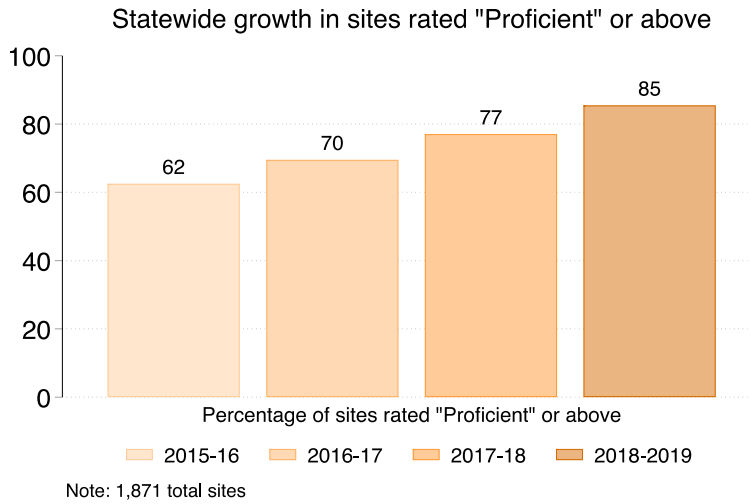
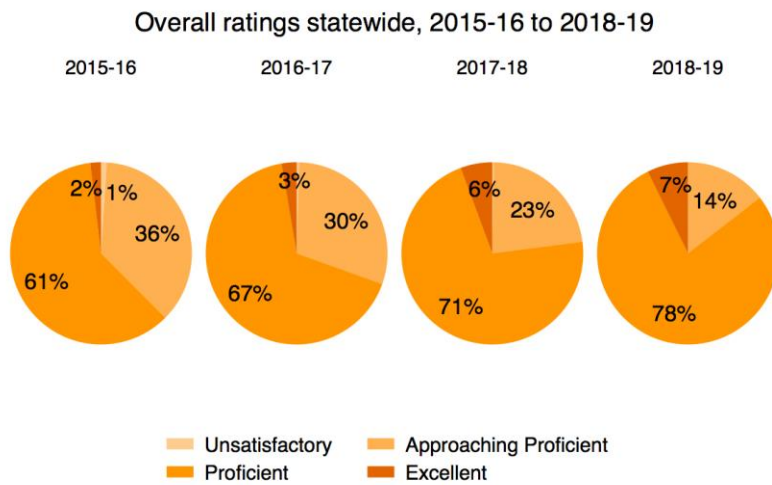


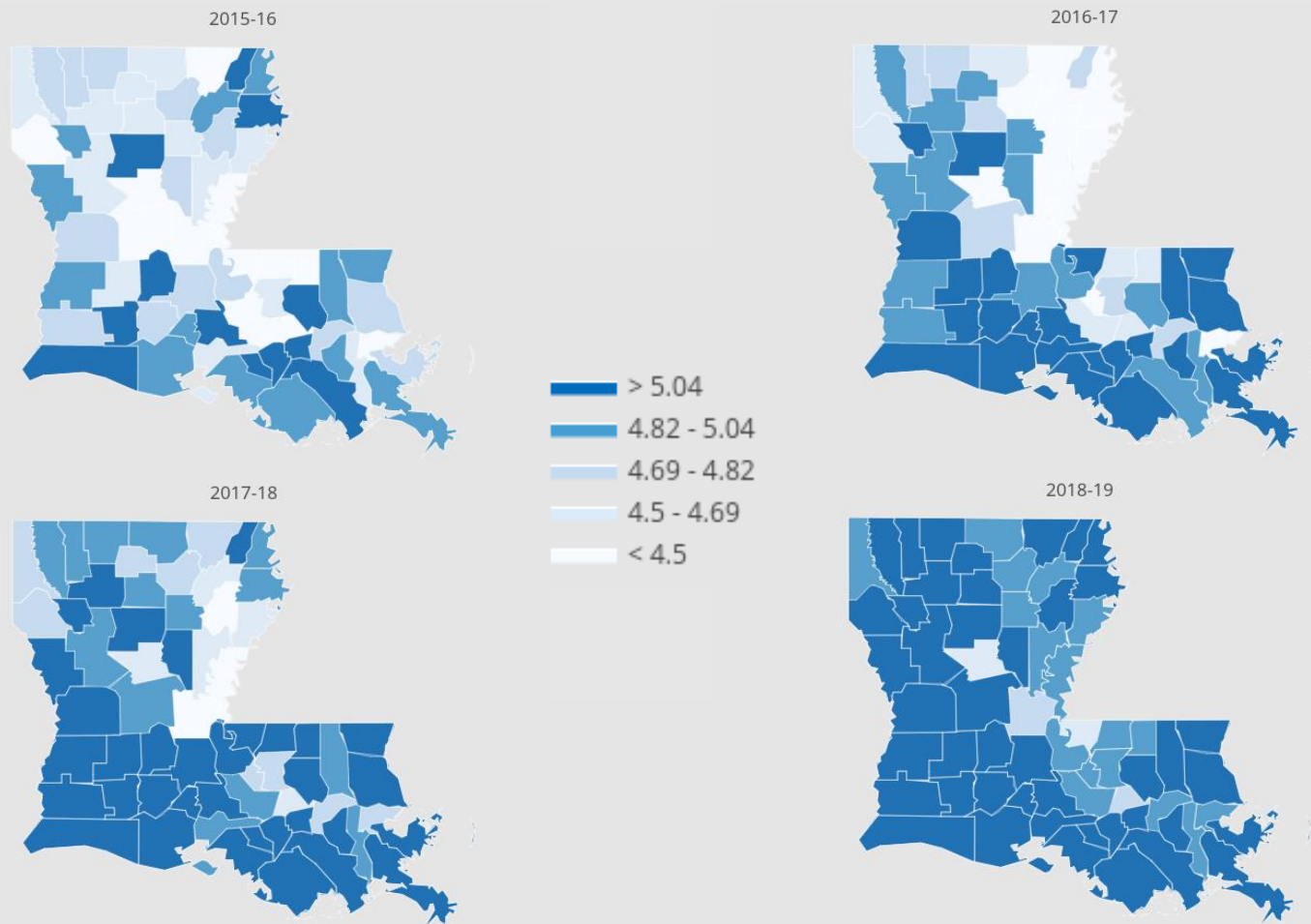
Figure 3 presents the distribution of ratings in each year. For comparability, we combine "Proficient" and "High Proficient" for 2018-19. The share of sites rated "Excellent" (indicated by darkest shading) more than tripled from 2015-16 to 2018-19, from 2% to 7%. The share of sites rated "Approaching Proficient" declined by more than half from 36% to 14%.

Figure 3.



Variation Across Early Childhood Networks

Figure 4. Network-level average CLASS scores, 2015-16 to 2018-19



The average gains in CLASS statewide mask considerable differences in patterns across communities. Figure 4 maps network-level average CLASS scores from 2015-16 to 2018-19. The color groupings, which are consistent across maps, are based on the five quintiles of 2015-16 scores; that is, white represents the bottom fifth (20%) of network-level CLASS scores in 2015-16 (i.e. networks with average CLASS scores less than 4.50) while the darkest blue represents the top fifth (20%) of network-level CLASS scores (scores greater than 5.04) from that year. Maps for each successive year also use the 2015-16 score groupings to illustrate growth over time (as the shading transitions to darker hues of blue). These maps show large differences, both across networks and regions of the state, in CLASS scores beginning in 2015-16. For instance, average CLASS scores in several networks across north and central Louisiana fell below 4.50; this indicates that sites in these networks were, on average, on the margin between an “Approaching Proficient” and “Proficient” rating. Conversely, average CLASS scores were often higher in southwestern Louisiana, as four networks in the region demonstrated scores at or above 5.04—well over proficient—in 2015-16.

These maps also highlight significant improvements over time. Though 20% of all networks earned average scores below 4.50 in 2015-16, none fell into this category by 2018-19. Further, while in 2015-16 13 networks had scores that placed them in the top category (scores greater than 5.04), by 2018-19, 45 networks had scores above that level. As an example, Rapides Parish was among the bottom quintile of networks in CLASS performance during 2015-16 with an average score of 4.41. Scores in Rapides grew rapidly such that its 2018-19 average of 5.24 would have placed it in the top 10% of 2015-16 network-level scores. Not all networks experienced similar growth, however, as illustrated by Grant and Avoyelles Parishes (just north and southeast of Rapides). In each year, these networks’ average CLASS scores remained among the lowest in the state and improved very little between 2015-16 and 2018-19. Finally, across the four-year period, scores remained low in parts of both northeast and southeast Louisiana, relative to other regions.

Figure 5. Percentage growth in sites rated “Proficient” or above, 2015-16 to 2018-19

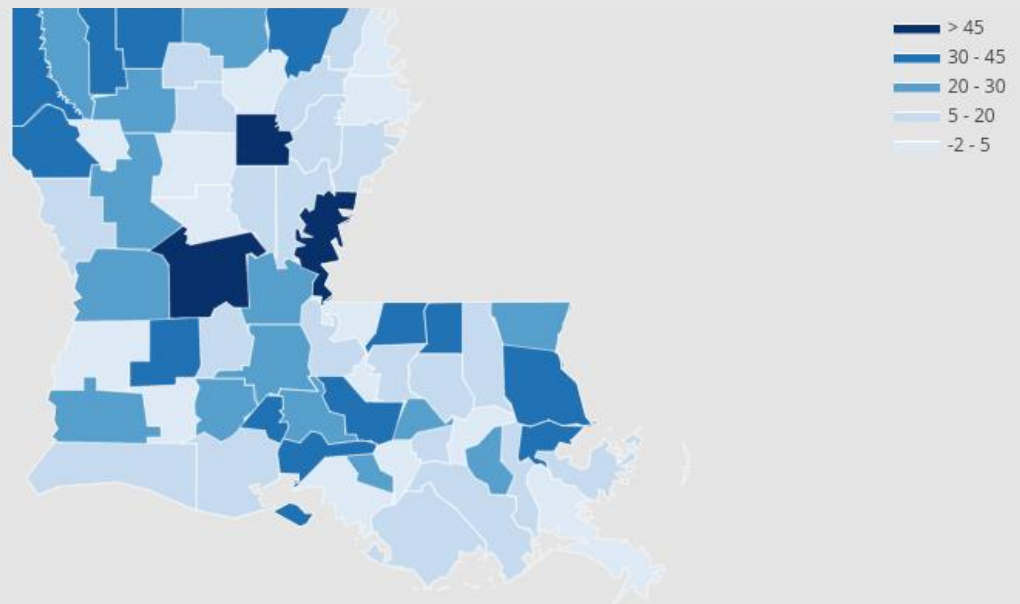


Figure 5 illustrates network-level *growth* in the percentage of sites rated “Proficient” or above between 2015-16 and 2018-19; darker shading indicates more growth in the share of “Proficient” sites. Between 2015-16 and 2018-19, nearly 50% of networks increased the share of proficient sites by over 20 percentage points. However, some networks demonstrated more modest gains and even slight declines, as demonstrated by the lightest shading in Figure 5. Specifically, networks across northwest and south-central Louisiana achieved some of the largest improvements while several networks across the northeastern and southeastern regions of the state improved marginally, if at all.

Differences Across Site Types

Both average CLASS scores and growth in those scores over time varied considerably by site types. Figure 6 presents average CLASS scores between 2015-16 and 2018-19 for each site type. In 2015-16, the average overall CLASS score for child care sites (4.30) was significantly lower than that for Head Start and school sites (4.64 and 5.16, respectively). Scores in child care sites increased by 0.57 points to 4.87 by 2018-19. Increases were smaller in Head Start (0.37) and school sites (0.27). Although average CLASS scores remained lower in child care sites than in the other site types throughout each of the four years, the size of this difference dropped. For instance, in 2015-16 school sites, on average,

scored nearly an entire point (0.86) higher than did child care sites. By 2018-19, this difference fell to closer to half a point (0.56). Appendix Table A1 provides overall CLASS score and proficiency trends statewide and by site type.

Figure 6.

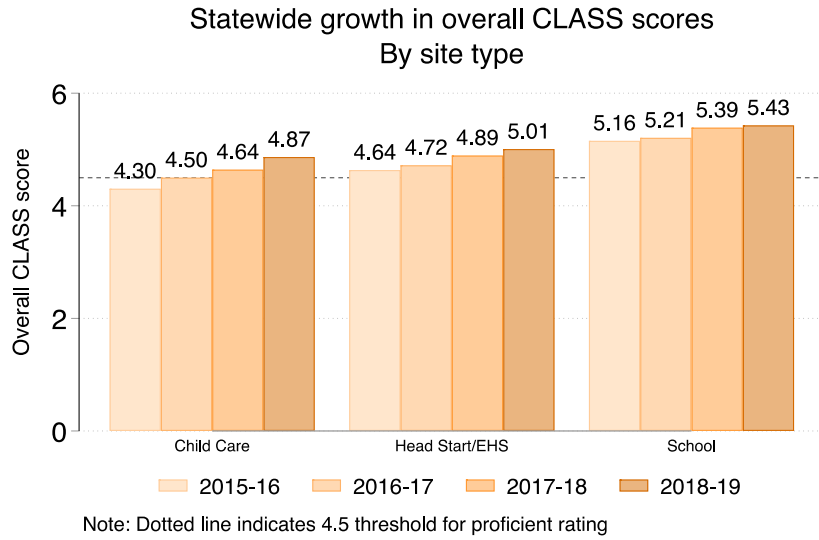


Figure 7.

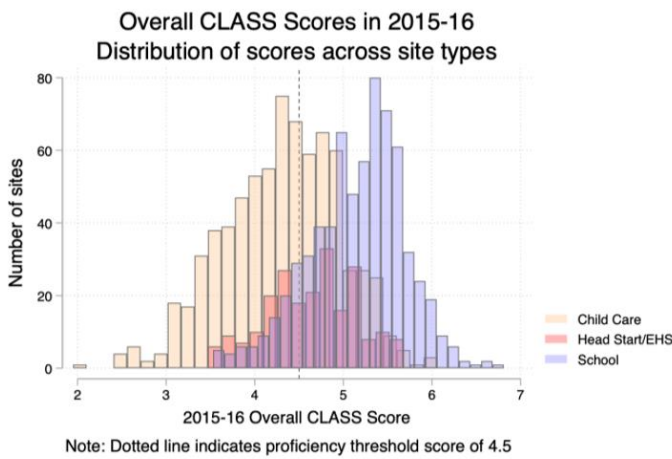
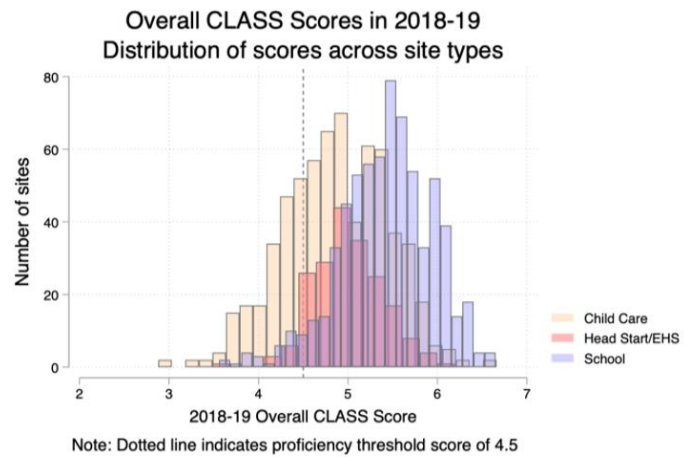


Figure 8.

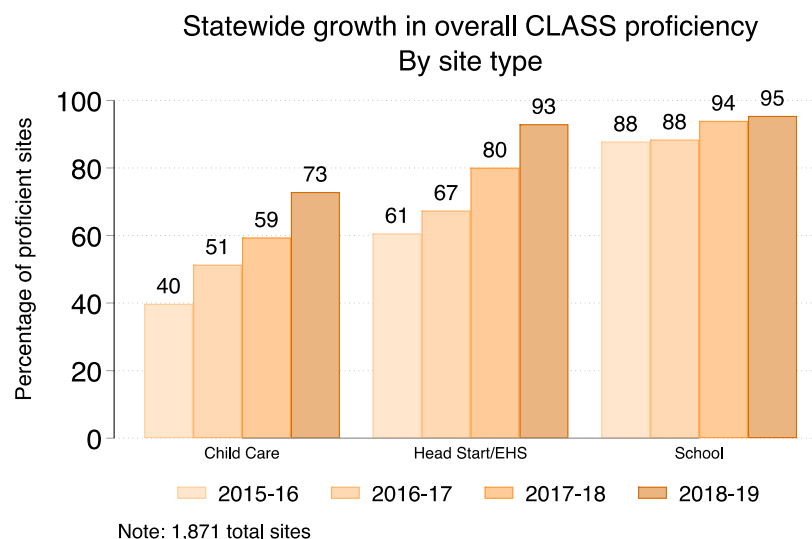


While average CLASS scores were lower for child care sites than for Head Start and school sites, scores varied widely within site types. Figure 7 presents the distributions of 2015-2016 scores for child care, Head Start and school sites along with a vertical line indicating the score threshold for a “Proficient” rating (4.50). The distributions reveal that despite lower average scores than school sites, some child care sites did outperform school and Head Start sites; and, though school and Head Start sites generally outperformed child care sites, a substantial proportion of school and Head Start sites had scores that put them below the average child care setting.

Figure 8 presents these distributions for the 2018-19 year and highlights noticeable improvements across each of the site types. Each distribution in this graph shifted to the right and a large share of each fell beyond the 4.5-point threshold necessary for a “Proficient” rating.

Finally, Figure 9 presents trends in rates of proficiency across site types. In 2015-16, 40% of child care sites were rated “Proficient” or above, as compared to 61% and 88% of Head Start and school sites, respectively. By 2018-19, the percentage of sites rated proficient or above in child care sites rose to 73%. Proficiency in Head Start and school sites reached 93% and 95%, respectively. While child care sites were still the least likely to achieve “Proficient” ratings by 2018-19, they achieved the greatest growth in the share proficient.

Figure 9.



The findings again suggest a narrowing in the quality gap between site types. For instance, in 2015-16, Head Start sites were 27 percentage points less likely than schools to be rated “Proficient” or above. By 2018-19, this difference fell to only 2 percentage points. Child care sites in 2015-16 were 48 percentage points less likely than schools to be rated “Proficient” or above. While a large gap remained in 2018-19, it fell to only 22 percentage points.

Changes across toddler and pre-K classrooms

A primary aim of Act 3 was to improve the quality of publicly-funded ECE for low-income children across age groups. From 2015-16 through 2018-19, all publicly-funded classrooms enrolling toddlers (one- and two-year-olds) and preschool-aged children (three- to five-year-olds) were observed using versions of the CLASS designed specifically for those settings. Figures 10 and 11 highlight trends in CLASS domain scores for the Toddler CLASS and Pre-K CLASS scores, respectively.

Figure 10.

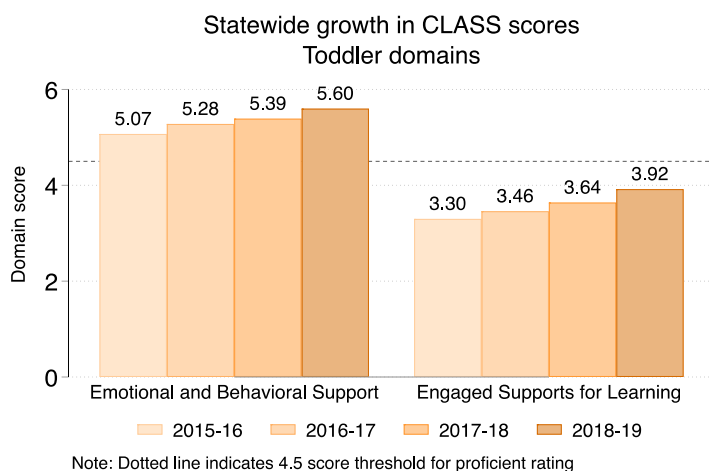


Figure 11.

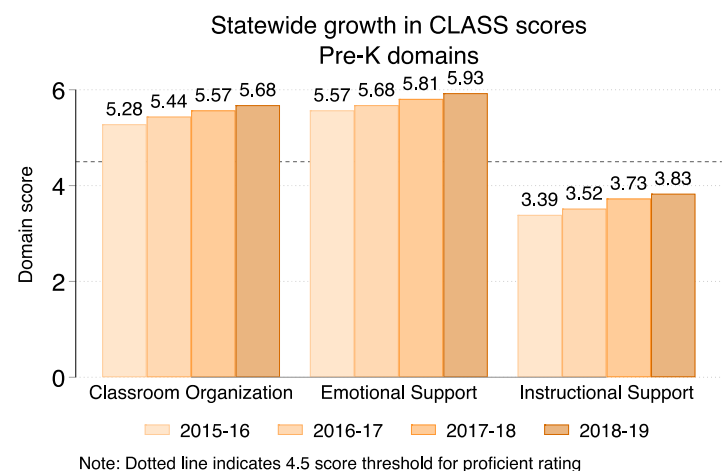


Figure 10 highlights that scores in each of the Toddler CLASS domains increased over the four-year period, but growth was particularly notable for Engaged Supports for Learning. Specifically, scores for Emotional and Behavioral Support increased by 0.53 points to 5.60 in 2018-19 and Engaged Support for Learning scores rose from 3.30 to 3.92 (+0.62) in 2018-19. Similarly, Figure 11 highlights growth in each of the pre-K domains, but most

notably in Instructional Support. While this domain demonstrated the lowest score in each of the four years, scores for Instructional Support, which research has linked more strongly to children’s learning outcomes than the other two domains, rose nearly half of a point (+0.44) from 2015-16 to 2018-19.

Implications for Louisiana’s Youngest Learners

This report highlights significant growth in both CLASS scores and proficiency ratings since 2015. This means that today, when children in Louisiana enroll in a publicly-funded early childhood setting—whether it is in a school, a Head Start, or a child care site—they are more likely than they were just five years ago to be in a classroom with the type of teacher-child interactions that have been shown to support children’s learning. They are more likely to be in a classroom where teachers are responsive to children’s needs, providing reassurance and encouragement, and helping children solve problems peacefully. They are more likely to be in a classroom with clear routines, consistent schedules, and well-organized learning centers. And they are more likely to be in a classroom where teachers introduce new material in a way that feels relevant to children’s lives and ask questions that extend children’s thinking and reasoning. Indeed, by 2018-19 Louisiana’s classrooms serving 3- to 5-year-olds had average quality ratings that were above empirically defined thresholds for “high-quality” (e.g. Weisenfeld, Frede, & Barnett, 2018) across all site types, and nearly all CLASS domains.

All of these changes are expected to lead to greater learning opportunities for young children. However, an important open question is to what extent these improvements have actually led to improvements in children’s early learning in Louisiana. Due largely to a lack of data, this is a difficult question to answer. Currently, there is not good longitudinal data available to assess whether multi-year improvement in CLASS scores is linked to increases in children’s early literacy and math skills, their social-emotional skills, and other key outcomes of interest. Systematically collecting data statewide about children’s learning, and using the same measures consistently over time, would help better answer that question. In the meantime, we can only speculate on the benefits of these quality improvements for children, based on findings from existing studies.

A large body of research does show that children in classrooms with higher CLASS scores tend to do better than those in classrooms with lower scores (e.g. Hamre, 2014). Hundreds of studies have explored the relationship between CLASS and child outcomes

and have generally found a positive, though modest, link. These studies have also shown that patterns differ depending on what domain of CLASS is considered (e.g. Instructional Support versus Emotional Support), the specific outcomes considered (e.g. literacy versus math), or the context for the study (e.g. state pre-K sites versus child care sites outside of the United States). One small study explored this question specifically in Louisiana, and therefore may be particularly relevant. It showed that local ratings from 85 classrooms serving four-year-olds in publicly-funded settings were positively linked to direct measures of children’s language, math, and executive function. Based on those findings, the sizable increase in the average overall CLASS scores in Louisiana over the four years this report examined—0.43 points on a 7-point scale—may be large enough to meaningfully impact children’s skills at school entry.⁴

Conclusion

The detailed data Louisiana has collected as part of its accountability system provide an unprecedented opportunity to examine large-scale early childhood quality improvement. Since 2015-16, LDOE has compiled multiple observations per year from every classroom in every publicly-funded school-based pre-K, Head Start, and child care site in the state. The goal of this report was to describe changes in the quality of publicly-funded early childhood settings in Louisiana between 2015-16 and 2018-19. We find evidence that, on average, the quality of teacher-child interactions in ECE settings in Louisiana has increased meaningfully over the past four years. Not only has the average CLASS score statewide increased by 0.43 points, but the number of sites deemed “Proficient” or above under Louisiana’s system has risen from 62% to 85%.

Encouragingly, this growth has been systemwide, with evidence of growth in nearly all communities, across all site types, and in classrooms serving both preschoolers and toddlers. Notably, child care and Head Start settings, which had substantially lower CLASS scores than did schools in 2015-16, showed greater growth. This means that today,

⁴ Vitiello and colleagues (2018) find that a 0.80-point increase in CLASS scores was associated with a 0.10, 0.14, and 0.09 of a standard deviation increase in language, math, and executive function skills, respectively. Since 2015-16, average overall CLASS scores in Louisiana have increased by 0.43 points. In “back-of-the-envelope” calculations, which assume that the relationship between CLASS and children’s development has remained stable over this time period and that the patterns observed in classrooms serving four-year-olds apply for other age groups, the magnitude of CLASS growth would be associated with an increase of about 0.05, 0.08, and 0.05 of a standard deviation in children’s language, math and executive function skills, respectively, over this time period. While small, these are considered meaningful gains.

the “quality gap” across site types is significantly less pronounced than it was five years ago.

The Toddler CLASS domain with the largest growth was Engaged Supports for Learning; the Pre-K CLASS domain with the largest growth was Instructional Support. While on average, both Engaged Supports for Learning and Instructional Support remain the lowest-scoring domain in the Toddler and Pre-K CLASS, respectively, growth in each domain was even larger than observed growth in overall CLASS scores. This is notable because these domains are both the lowest scoring domains in other nationwide data (e.g. Head Start, Barnett & Friedman-Krauss, 2018) and relatively more difficult for teachers to improve.

The findings in this report are consistent with the notion that Louisiana’s QRIS and related supports for publicly-funded sites have been effective in providing Louisiana children with higher quality ECE settings. Louisiana introduced several initiatives over the four year period examined, including providing grant funding to support the purchase and use of approved curricula; the introduction of a new early childhood teacher certification designed to better train teachers to engage in high-quality teacher-child interactions; the provision of particular supports for low-performing networks; and many other strategies. Understanding which of these supports has been most effective in generating change is a clear next step from both a research and policy perspective. Another goal is to use these data to identify commonalities among sites struggling to improve and to provide them with effective supports.

Finally, although the quality improvements documented in this report are expected to benefit young children, it is difficult to know exactly how much improvements in CLASS have translated to improvements in children’s outcomes. LDOE has not required the use of a single, comparable school-readiness assessment over the time period considered, making it difficult to assess whether children are now entering kindergarten more ready to learn. Further, the oldest cohort of children from the earliest year of data examined (2015-16) are only now in third grade. It is too early to assess whether Louisiana’s efforts to improve early childhood quality systemwide have led to improvements in those outcomes. However, tracking the benefits of systemwide quality improvements on child outcomes is a priority, and investments in the data collection and data systems needed to do so are essential.

Louisiana's Act 3 reform was an ambitious initiative that restructured early care and education in the state with the goal of improving the quality of teacher-child interactions. Data from this report indicate that ECE site quality has improved statewide and across various contexts, particularly those demonstrating the lowest initial levels of quality.

Works Cited

- Burchinal, M. (2017). Measuring Early Care and Education Quality. *Child Development Perspectives*. <https://doi.org/10.1111/cdep.12260>
- Campbell, F. A., Ramey, C. T., Pungello, E., Sparling, J., & Miller-Johnson, S. (2002). Early Childhood Education: Young Adult Outcomes From the Abecedarian Project. *Applied Developmental Science*, 6(1), 42–57. https://doi.org/10.1207/S1532480XADS0601_05
- Cannon et al., 2017
- Deming, D. (2009). Early Childhood Intervention and Life-Cycle Skill Development: Evidence from Head Start. *American Economic Journal: Applied Economics*, 1(3), 111–134.
- Early, D. M., Maxwell, K. L., Burchinal, M., Alva, S., Bender, R. H., Bryant, D., ... Zill, N. (2007). Teachers' Education, Classroom Quality, and Young Children's Academic Skills: Results From Seven Studies of Preschool Programs. *Child Development*, 78(2), 558–580. <https://doi.org/10.1111/j.1467-8624.2007.01014.x>
- Barnett, W. S., & Friedman-Krauss, A. H. (2016). *State(s) of Head Start*. Rutgers Graduate School of Education: National Institute for Early Education Research.
- Hamre, B. K. (2014). Teachers' Daily Interactions With Children: An Essential Ingredient in Effective Early Childhood Programs. *Child Development Perspectives*, 8(4), 223–230. <https://doi.org/10.1111/cdep.12090>
- Harms, Clifford, & Cryer, 2014
- Helburn, S. W. (1995). *Cost, Quality and Child Outcomes in Child Care Centers. Technical Report, Public Report, and Executive Summary*. Cost, Quality, and Child Outcomes Study, Economics Department, Campus Box 159, P. Retrieved from <https://eric.ed.gov/?id=ed386297>
- Mashburn, A. J., Pianta, R. C., Hamre, B., Downer, J. T., Barbarin, O. A., Bryant, D., & Howes, C. (2008). Measures of classroom quality in prekindergarten and children's development of academic, language, and social skills. *Child Development*, 79(3), 732–749.
- Phillips, D., Lipsey, M., Dodge, K., Haskins, R., Bassok, D., Burchinal, M., ... Weiland, C. (2017). *Puzzling it out: The current state of scientific knowledge of pre-kindergarten effects. A consensus statement*. Washington DC: Brookings Institution.
- Pianta, R. C., Paro, K. M. L., & Hamre, B. K. (2008). *Classroom Assessment Scoring System: Manual Pre-K*. <https://doi.org/10.14507/er.v0.1049>
- Vitiello, V. E., Bassok, D., Hamre, B. K., Player, D., & Williford, A. P. (2018). Measuring the quality of teacher–child interactions at scale: Comparing research-based and state observation approaches. *Early Childhood Research Quarterly*, 44, 161–169.

Weisenfeld, G.G., Frede, E., & Barnett, W.S. (2018). Implementing 15 essential elements for high-quality pre-k: An updated scan of state policies. Rutgers Graduate School of Education: National Institute for Early Education Research. Retrieved from: <http://nieer.org/wp-content/uploads/2018/10/Essential-Elements-FINAL-9.14.18.pdf>

Appendix
Table A1. Overall CLASS Scores, 2015-16 to 2018-19

	Scores					Proficiency (%)				
	<i>15-16</i>	<i>16-17</i>	<i>17-18</i>	<i>18-19</i>	<i>Change</i>	<i>15-16</i>	<i>16-17</i>	<i>17-18</i>	<i>18-19</i>	<i>Change</i>
Overall	4.70 (0.71)	4.84 (0.68)	5.00 (0.68)	5.13 (0.59)	+0.43	62	70	77	85	+23
Site Types										
Child Care	4.30 (0.65)	4.50 (0.63)	4.64 (0.65)	4.87 (0.58)	+0.57	40	51	59	73	+33
Head Start	4.64 (0.50)	4.72 (0.52)	4.89 (0.46)	5.01 (0.39)	+0.37	61	67	80	93	+32
School	5.16 (0.54)	5.21 (0.57)	5.39 (0.53)	5.43 (0.51)	+0.27	88	88	94	95	+7

Note: 1,871 total sites observed across years. Standard deviations are presented in parentheses.

Table A2. Instructional Support Scores, 2015-16 to 2018-19

	Scores					Proficiency (%)				
	<i>15-16</i>	<i>16-17</i>	<i>17-18</i>	<i>18-19</i>	<i>Change</i>	<i>15-16</i>	<i>16-17</i>	<i>17-18</i>	<i>18-19</i>	<i>Change</i>
Overall	3.39 (0.95)	3.52 (0.95)	3.73 (0.95)	3.83 (0.88)	+0.44	13	17	22	23	+10
Site Types										
Child Care	2.84 (0.84)	3.03 (0.89)	3.18 (0.91)	3.40 (0.89)	+0.56	3	7	9	12	+9
Head Start	3.39 (0.72)	3.47 (0.69)	3.64 (0.63)	3.68 (0.56)	+0.29	8	6	8	10	+2
School	3.91 (0.80)	3.96 (0.84)	4.22 (0.77)	4.21 (0.76)	+0.30	23	29	36	36	+13

Note: 1,748 total sites observed across years. Standard deviations are presented in parentheses.

Table A3. Emotional Support Scores, 2015-16 to 2018-19

	Scores					Proficiency (%)				
	<i>15-16</i>	<i>16-17</i>	<i>17-18</i>	<i>18-19</i>	<i>Change</i>	<i>15-16</i>	<i>16-17</i>	<i>17-18</i>	<i>18-19</i>	<i>Change</i>
Overall	5.57 (0.69)	5.68 (0.62)	5.81 (0.61)	5.93 (0.54)	+0.36	93	96	97	98	+5
Site Types										
Child Care	5.32 (0.80)	5.51 (0.67)	5.62 (0.68)	5.77 (0.59)	+0.45	85	92	94	97	+12
Head Start	5.39 (0.50)	5.48 (0.53)	5.60 (0.49)	5.73 (0.40)	+0.34	96	97	99	99	+3
School	5.85 (0.50)	5.88 (0.52)	6.03 (0.49)	6.10 (0.47)	+0.25	99	98	99	100	+1

Note: 1,748 total sites observed across years. Standard deviations are presented in parentheses.

Table A4. Classroom Organization Scores, 2015-16 to 2018-19

	Scores					Proficiency (%)				
	<i>15-16</i>	<i>16-17</i>	<i>17-18</i>	<i>18-19</i>	<i>Change</i>	<i>15-16</i>	<i>16-17</i>	<i>17-18</i>	<i>18-19</i>	<i>Change</i>
Overall	5.28 (0.78)	5.44 (0.71)	5.57 (0.71)	5.68 (0.64)	+0.40	85	90	93	95	+10
Site Types										
Child Care	4.88 (0.83)	5.12 (0.74)	5.24 (0.76)	5.39 (0.70)	+0.51	71	81	86	89	+18
Head Start	5.08 (0.60)	5.21 (0.61)	5.31 (0.53)	5.44 (0.44)	+0.36	81	89	92	98	+17
School	5.72 (0.52)	5.78 (0.53)	5.92 (0.51)	5.97 (0.49)	+0.25	98	98	99	99	+1

Note: 1,748 total sites observed across years. Standard deviations are presented in parentheses.

Table A5. Emotional and Behavioral Support Scores, 2015-16 to 2018-19

	Scores					Proficiency (%)				
	<i>15-16</i>	<i>16-17</i>	<i>17-18</i>	<i>18-19</i>	<i>Change</i>	<i>15-16</i>	<i>16-17</i>	<i>17-18</i>	<i>18-19</i>	<i>Change</i>
Overall	5.07 (0.73)	5.28 (0.66)	5.39 (0.69)	5.60 (0.60)	+0.53	80	88	89	96	+16
Site Types										
Child Care	5.05 (0.75)	5.26 (0.67)	5.36 (0.70)	5.58 (0.62)	+0.53	78	87	88	95	+17
Head Start	5.32 (0.53)	5.44 (0.55)	5.69 (0.45)	5.78 (0.36)	+0.45	97	96	98	100	+3
School	5.28 (0.63)	5.56 (0.59)	5.66 (0.53)	5.70 (0.45)	+0.42	100	89	100	100	0

Note: 938 total sites observed across years. Standard deviations are presented in parentheses.

Table A6. Engaged Support for Learning Scores, 2015-16 to 2018-19

	Scores					Proficiency (%)				
	<i>15-16</i>	<i>16-17</i>	<i>17-18</i>	<i>18-19</i>	<i>Change</i>	<i>15-16</i>	<i>16-17</i>	<i>17-18</i>	<i>18-19</i>	<i>Change</i>
Overall	3.30 (0.84)	3.46 (0.85)	3.64 (0.86)	3.92 (0.78)	+0.62	9	12	18	25	+16
Site Types										
Child Care	3.25 (0.83)	3.42 (0.84)	3.59 (0.86)	3.88 (0.79)	+0.53	7	11	16	23	+16
Head Start	3.91 (0.61)	3.91 (0.75)	4.21 (0.65)	4.24 (0.60)	+0.33	21	16	33	42	+21
School	3.67 (1.06)	3.89 (1.13)	4.07 (0.89)	4.24 (0.61)	+0.57	20	33	56	30	+10

Note: 938 total sites observed across years. Standard deviations are presented in parentheses.