

# Relationships between Schoolwide Instructional Observation Scores and Student Academic Achievement and Growth in Low-Performing Schools in Massachusetts

The Massachusetts Department of Elementary and Secondary Education, like other state education agencies and school districts, recognizes that the quality of instruction is a key lever to turning around low-performing schools. As part of annual monitoring of state-designated low-performing schools, the department's external monitors observe instruction in low-performing schools using Teachstone's Classroom Assessment Scoring System. The external monitors rate low-performing schools on three instructional domains: emotional support, classroom organization, and instructional support.

This study examines the relationships between schoolwide instructional observation scores in these domains and schoolwide student academic achievement (measured by the percentage of students who met or exceeded expectations on state assessments) and growth in low-performing schools while taking into account what might be attributed to the schools' percentage of economically disadvantaged students and to school grade span. It found a statistically significant positive relationship between schoolwide instructional observation scores in the classroom organization domain and schoolwide student achievement in English language arts. There was no significant relationship between scores in any other domain and achievement in English language arts or between scores in any domain and achievement in math. The relationship between instructional observation scores and student achievement may be weak because achievement may be influenced by other factors, including students' prior academic achievement and the economic and social challenges their families face. The study also found statistically significant positive relationships between schoolwide instructional observation scores in each domain and schoolwide student growth in both English language arts and math. On a 7 point scale, a 1 point increase in schoolwide instructional observation score was associated with an increase in schoolwide student academic growth of 4.4 percentile points in English language arts and 5.1 percentile points in math.

## Why this study?

The Massachusetts Department of Elementary and Secondary Education (DESE), like other state education agencies and school districts, recognizes that the quality of instruction is a key lever to turning around low-performing schools. As part of annual monitoring of low-performing schools, DESE employs Teachstone's Classroom Assessment Scoring System (CLASS) to collect data on the quality of interactions between teachers and students during instruction. The CLASS observation tool rates the quality of interactions in three domains: emotional support, classroom organization, and instructional support.

Prior research suggests a relationship in individual classrooms between consistently high CLASS scores and improved academic outcomes for students in those classrooms. However, DESE uses schoolwide averages of single observations of multiple classrooms within a school and provides this feedback to the school rather than to individual classrooms. Thus, researchers at DESE and the Regional Educational Laboratory Northeast & Islands

partnered to examine whether there is a relationship in low-performing schools between schoolwide average domain scores and schoolwide student academic achievement and growth. The study draws on data from classroom monitoring visits in 2016/17 and 2017/18 as well as publicly available data on schoolwide student academic achievement and growth.<sup>11</sup>

**What was studied and how?**

The study examined two research questions focused on understanding the relationships between schoolwide instructional observation scores and schoolwide student academic achievement and growth.

1. What are the characteristics of low-performing schools in Massachusetts in terms of student demographics, schoolwide student academic achievement and growth, and the overall and within-school variation in instructional observation scores?
2. Are the instructional observation scores of low-performing schools associated with concurrent schoolwide student academic achievement or growth in English language arts and math while taking into account what might be attributed to the schools’ percentage of economically disadvantaged students and to school grade span?

To examine the relationship between schoolwide instructional observation scores and student outcomes, the study used data collected by certified CLASS observers in a sample of classrooms during the 2016/17 and 2017/18 annual monitoring of low-performing schools. The domain scores for emotional support, classroom organization, and instructional support were calculated as an average of a set of dimensions for each domain. The scores ranged from a low of 1 to a high of 7 (table 1).

Publicly available schoolwide student demographic and academic outcome data for 2016/17 and 2017/18 were used.<sup>2</sup> Outcome data included schoolwide student academic achievement in English language arts and math—the percentage of students who met or exceeded proficiency on the state assessment—and student growth in English language arts and math, calculated by DESE using the schoolwide median student growth percentile.

The 88 low-performing schools in Massachusetts 2016/17 or 2017/18 were categorized as serving students in the elementary grade span (grades 4–5), the secondary grade span (grades 6–12), or both to align with the CLASS observation tool designed specifically for those grade spans. The sample had 100 elementary and secondary

**Table 1. Classroom Assessment Scoring System domain score classifications**

Score range	Classification	Definition
1.00–2.99	Low range	The interactions observed within this domain are of minimal effectiveness. Effective interactions happen rarely, if ever, and when they do, they are isolated, brief, or low quality.
3.00–5.99	Mid-range	Effective interactions within this domain are observed sometimes or to some degree but are inconsistent or limited.
6.00–7.00	High range	Effective interactions are observed consistently—they are frequent, sustained, and high quality.

Source: Croasdale, N. (2015). *Cracking the code: How to interpret the CLASS score report as a coach*. Teachstone. Retrieved January 27, 2019, from <http://info.teachstone.com/blog/cracking-code-interpret-class-score-report-coach>.

1. The lead author supported the design of the overall monitoring process but did not have a role in designing the instructional observation process or the data collection tool used in this study. The observation protocol used in monitoring low-performing schools in Massachusetts is based on one used in other states and districts, and the instructional observation tool and processes were developed by Teachstone.  
 2. See [http://profiles.doe.mass.edu/state\\_report/](http://profiles.doe.mass.edu/state_report/) for DESE schoolwide demographic and student outcome data.

grade spans because 12 schools served students in both elementary and secondary grade spans and were counted in both categories.

For research question 1, descriptive analyses focused on the demographic makeup of schools and student academic achievement and growth at the school level. In addition, average instructional observation scores were calculated at the domain level. For research question 2, regression models analyzed how domain scores were related to schoolwide student academic achievement and schoolwide student academic growth while taking into account what might be attributed to the schools' percentage of economically disadvantaged students and to school grade span (see appendixes B and C in the main report for details).

### Findings

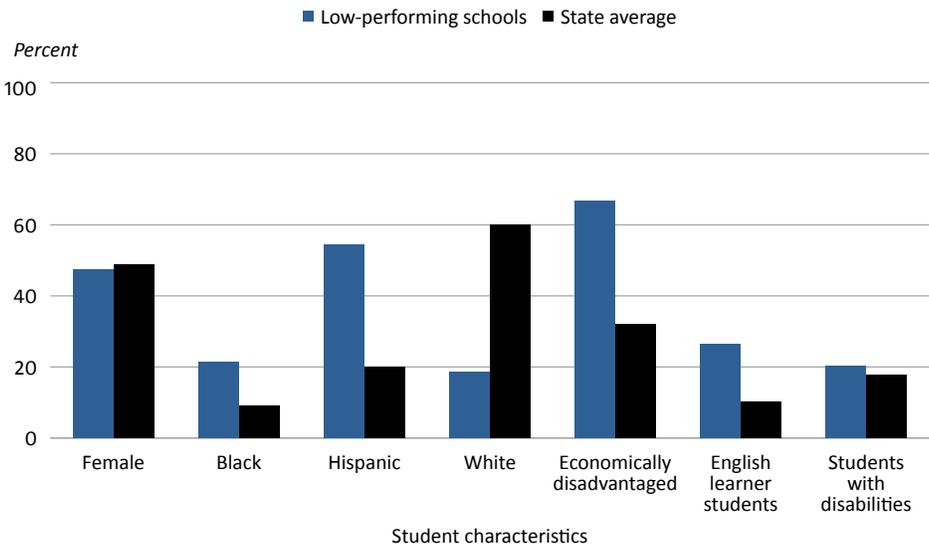
#### *Low-performing schools served higher percentages of Black, Hispanic, economically disadvantaged, and English learner students than the state average*

On average, 67 percent of students in low-performing schools were economically disadvantaged compared with 32 percent statewide, and 26 percent of students in low-performing schools were English learner students compared with 10 percent statewide (figure 1).

#### *Scores in the classroom organization domain had a statistically significant positive relationship with schoolwide student achievement in English language arts in concurrent years, but scores in other domains had no significant relationship with achievement in English language arts or math*

No significant relationship was found between scores in the emotional support or instructional support domain and schoolwide student achievement on the state English language arts assessment, and no significant relationship was found between scores in any domain and achievement on the state math assessment while taking into account what might be attributed to the schools' percentage of economically disadvantaged students and to school grade span (table 2). The relationship between schoolwide instructional observation scores in other

**Figure 1. Low-performing schools in Massachusetts served higher percentages of Black, Hispanic, economically disadvantaged, and English learner students compared with the state average, 2017/18**



Note: See table B2 in appendix B for more information.

Source: Authors' analysis based on data from the Massachusetts Department of Elementary and Secondary Education.

**Table 2. Relationships between instructional observation domain scores and schoolwide student academic achievement and growth in low-performing schools in Massachusetts, by domain and subject, 2016/17 or 2017/18**

Domain	Schoolwide student academic achievement		Schoolwide student academic growth	
	English language arts	Math	English language arts	Math
Emotional support	1.8	1.9	3.3**	3.6**
Classroom organization	5.1*	4.3	4.2*	7.6**
Instructional support	1.3	2.3	2.7**	2.6*
Three domains combined	2.9	3.4	4.4**	5.1**

\* Significant at  $p < .05$ ; \*\* significant at  $p < .01$ .

Note: The sample size was 100 grade spans in 88 low-performing schools that received a monitoring visit in 2016/17 or 2017/18 (46 schools with elementary grade spans, 54 schools with secondary grade spans, and 12 schools with both grade spans). Schoolwide student academic achievement is the percentage of students who met or exceeded expectations on annual state assessments. Schoolwide student academic growth was calculated by the Massachusetts Department of Elementary and Secondary Education as the school's median student growth percentile. Coefficients in this table show the increase in these outcomes (percentage points for schoolwide student achievement and median student growth percentile for schoolwide student academic growth) associated with a 1 point increase in the 7 point schoolwide observation score.

Source: Authors' analysis based on data from the Massachusetts Department of Elementary and Secondary Education.

domains and schoolwide student academic achievement may be weak because student achievement is influenced by factors that precede current instructional quality.

***Schoolwide instructional observation scores in all three domains had a statistically significant positive relationship with schoolwide student growth in English language arts and math***

Low-performing schools with higher instructional observation scores in any one domain tended to have higher student growth in English language arts and math while taking into account what might be attributed to the schools' percentage of economically disadvantaged students and to school grade span (see table 2). A statistically significant positive relationship was also found between observation scores for the three domains combined and student academic growth.

**Implications**

Schoolwide instructional observation scores had a positive relationship with schoolwide student academic growth in this exploration, which supports the way DESE uses the current tool to offer instructional feedback to low-performing schools through the monitoring process. As a mechanism for feedback on the quality of instruction to low-performing schools, a schoolwide instructional score or measure may be informative.

The results are, however, not causal and may not be generalized to schools outside Massachusetts. In addition, the small sample of schools and the restriction of the sample to low-performing schools are limitations. Finally, the study examined only academic outcomes. Future work could examine relationships between instructional observation scores and other outcomes, such as chronic absenteeism, student behavior, teacher turnover, and high school graduation rate, and between instructional observation scores and student outcomes in other than low-performing schools if classroom data are also collected for other schools.

This brief was prepared for the Institute of Education Sciences (IES) under Contract ED-IES-17-C-0008 by the Regional Educational Laboratory Northeast & Islands administered by Education Development Center, Inc. The content of the publication does not necessarily reflect the views or policies of IES or the U.S. Department of Education, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government. The full report is available on the Regional Educational Laboratory website at <http://ies.ed.gov/ncee/edlabs>.

Therriault, S. B., Walston, J., Li, Y., & Pan, J. (2020). *Relationships between schoolwide instructional observation scores and student academic achievement and growth in low-performing schools in Massachusetts* (REL 2020–026). U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Northeast & Islands.