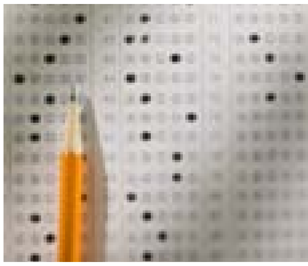
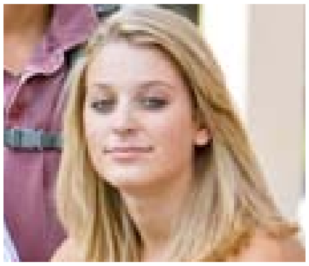
ISSUES & ANSWERS

REL 2011–No. 102

REL Northeast & Islands
Regional Educational Laboratory At Education Development Center, Inc.

How student and school characteristics are associated with performance on the Maine High School Assessment





ies NATIONAL CENTER FOR EDUCATION EVALUATION AND REGIONAL ASSISTANCE
Institute of Education Sciences

ISSUES & ANSWERS

REL 2011–No. 102

REL NORTHEAST & ISLANDS
Regional Educational Laboratory
At Education Development Center, Inc.

How student and school characteristics are associated with performance on the Maine High School Assessment

February 2011

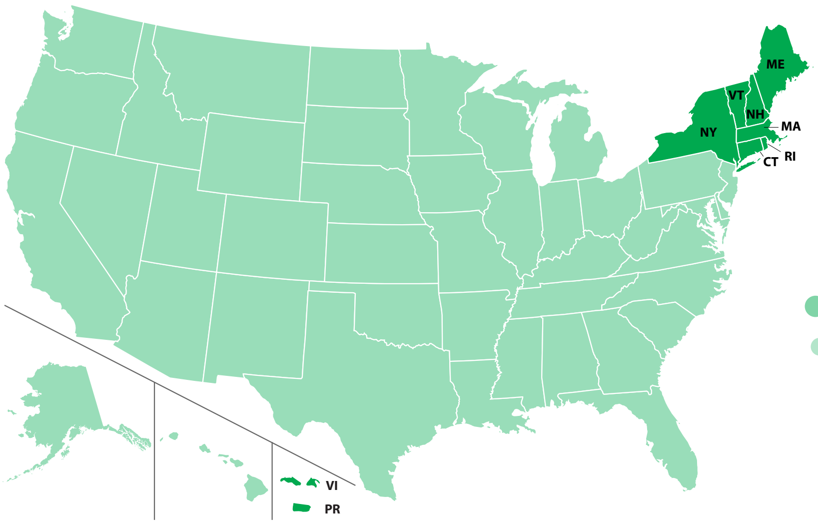
Prepared by

Craig D. Hoyle
Education Development Center, Inc.

Laura M. O'Dwyer
Boston College

Quincy Chang
Education Development Center, Inc.

ies NATIONAL CENTER FOR EDUCATION EVALUATION AND REGIONAL ASSISTANCE
Institute of Education Sciences



REL
NORTHEAST
& ISLANDS
Regional Educational Laboratory
At Education Development
Center, Inc.

Issues & Answers is an ongoing series of reports from short-term Fast Response Projects conducted by the regional educational laboratories on current education issues of importance at local, state, and regional levels. Fast Response Project topics change to reflect new issues, as identified through lab outreach and requests for assistance from policymakers and educators at state and local levels and from communities, businesses, parents, families, and youth. All Issues & Answers reports meet Institute of Education Sciences standards for scientifically valid research.

February 2011

This report was prepared for the Institute of Education Sciences (IES) under Contract ED-06-CO-0025 by Regional Educational Laboratory Northeast and 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.

Hoyle, C.D., O'Dwyer, L.M., and Chang, Q. 2011. *How student and school characteristics are associated with performance on the Maine High School Assessment.* (Issues & Answers Report, REL 2011–No. 102). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Northeast and Islands. Retrieved from http://ies.ed.gov/ncee/edlabs.

This report is available on the regional educational laboratory web site at http://ies.ed.gov/ncee/edlabs.

How student and school characteristics are associated with performance on the Maine High School Assessment

Using multilevel regression models to examine how student characteristics, student prior achievement measures, and school characteristics are associated with performance on the Maine High School Assessment, the study finds statistically significant relationships between several of these variables and assessment scores in reading, writing, math, and science.

The Maine Department of Education wanted to use longitudinal data from its data system to better understand whether and how student and school characteristics are associated with student performance on the state-mandated Maine High School Assessment (MHSA). It was particularly interested in understanding the factors associated with changes in test scores between the beginning of grade 10 and the end of grade 11. The MHSA, which comprises the College Board's SAT tests in critical reading, writing, and math and a science assessment, is administered in the spring of grade 11 to determine whether Maine high schools have made adequate yearly progress.

The following research question guided this study:

- How are student characteristics, student prior achievement measures, and school characteristics associated with students' grade 11 Maine High School Assessment scores in reading, writing, math, and science?

The Maine Department of Education provided data on MHSA scores, characteristics, and prior achievement measures for all grade 11 students in Maine in 2007/08. The prior achievement measures examined were grade 8 Maine Educational Assessment (MEA) scores in reading, writing, math, and science and grade 10 Preliminary SAT/National Merit Scholarship Qualifying Test (PSAT) verbal, math, and writing scores.

Data on school characteristics were either provided directly by the Maine Department of Education or gathered from publicly available data. The data consisted of percentage of racial/ethnic minority students, economically disadvantaged students, students in special education, English language learner students, and students who drop out; student–teacher ratio; mean years of teaching; cohort size; school location; whether the school made adequate yearly progress in reading and math in 2007; and whether the school was classified as Title I in 2007.

Multilevel regression modeling, which allows individual and group characteristics to be included in models of individual outcomes, was used to analyze the associations.

In the primary model used to address the research question, statistically significant relationships were found between some student and school characteristics and grade 8 MEA and grade 10 PSAT scores and grade 11 MHSA scores in reading, writing, math, and science, holding all other variables constant.

The major findings related to student characteristics as predictors of MHSA domain scores are:

- Gender was a significant predictor in all four MHSA domains. Male students were predicted to have significantly higher MHSA scores in reading (–0.052 standard deviation), math (–0.088 standard deviation), and science (–0.183 standard deviation) than female students. Female students were predicted to have significantly higher MHSA writing scores (0.060 standard deviation) than male students.

- Economically disadvantaged students were predicted to have significantly lower MHSA scores in reading (–0.035 standard deviation), writing (–0.061 standard deviation), and math (–0.039 standard deviation) than non–economically disadvantaged students.

- Students in special education were predicted to have significantly lower MHSA scores in reading (–0.065 standard deviation), writing (–0.117 standard deviation), and math (–0.105 standard deviation) than general education students.

- English language learner students were predicted to have significantly lower MHSA scores in reading (–0.061 standard deviation), writing (–0.102 standard deviation), and science (–0.100 standard deviation) than non–English language learner students.

The major findings related to prior achievement on the MEA and PSAT as predictors of MHSA domain scores are:

- With three exceptions, grade 8 MEA and grade 10 PSAT scores were significantly related to MHSA scores. Within each MHSA domain, the most recent measure of prior achievement with the closest match in content was the strongest predictor:

  - The strongest predictors of MHSA reading scores were grade 10 PSAT verbal and writing scores.

  - The strongest predictor of MHSA writing scores was PSAT writing scores.

  - The strongest predictor of MHSA math scores was grade 10 PSAT math scores.

  - The strongest predictor of MHSA science scores was grade 8 MEA science scores.

The major findings related to school characteristics as predictors of MHSA domain scores are:

- Students in schools with higher percentages of grade 11 economically disadvantaged students were predicted to have significantly lower MHSA reading scores.

- Students in schools with higher percentages of grade 11 students in special education were predicted to have significantly lower MHSA science scores.

- Students in schools with higher percentages of students in cohort who drop out were predicted to have significantly higher MHSA reading scores.

- Students in schools with higher student–teacher ratios were predicted to have significantly lower MHSA scores in reading, writing, and math.

- Students in schools with larger grade 11 cohorts were predicted to have significantly higher MHSA math scores.

- Compared with students in rural schools, students in suburban schools were predicted to have significantly higher MHSA writing scores.

- Students in schools that made adequate yearly progress in reading in 2007 were predicted to have significantly higher MHSA writing and math scores.

Because the PSAT is a preliminary assessment for the SAT, which is used for the reading, writing, and mathematics domains of the MHSA and was administered in the beginning of grade 10, the strong relationship between PSAT scores and MHSA scores could influence both the magnitude of the multilevel regression coefficients associated with the other variables and the overall interpretation of the model. To the extent that student characteristics, grade 8 MEA scores, and school characteristics relate to grade 10 PSAT scores as well as grade 11 MHSA scores, the regression coefficients for the remaining predictor variables may be attenuated relative to a model that does not include PSAT scores. Indeed, when PSAT scores are included in the model, the coefficients for the non-PSAT variables can, to some extent, be interpreted as the relationship between those variables and the change in achievement between grades 10 and 11. This is particularly the case with respect to the MHSA reading, writing, and math scores.

The findings from this study will inform further efforts by the Maine Department of Education to understand the association between student and school factors and high school achievement. To the extent that MHSA scores, student characteristics, PSAT scores, school characteristics, and unmeasured factors remain the same, results of the current study can be generalized to other cohorts of Maine students who complete all four sections of the MHSA in grade 11.[1] Results from this study could also be used to identify students and schools likely to benefit most from additional assistance. Because this is a descriptive rather than an experimental study, the findings are not sufficient to support causal inferences.

February 2011

Note

1. In 2009, the grade 8 assessment in Maine changed from the MEA, taken by the 2007/08 cohort of students, to the New England Common Assessment Program (NECAP). Despite the potential differences between the MEA and the NECAP, grade 8 MEA scores represent measures of achievement prior to students' entry to their respective high schools and are therefore still of interest to the Maine Department of Education as it examines how information in its evolving longitudinal data system can be used to improve the education of students in Maine.

TABLE OF CONTENTS

Why this study? **1**

Findings **5**
Association of student characteristics, student prior achievement measures, and school characteristics with grade 11 Maine High School Assessment scores (model 2) 5
Predictors of Maine High School Assessment domain scores 5
Predicted differences across Maine High School Assessment domains by measure of prior achievement 10
Implications of including Preliminary SAT scores in model 2 11
Variance explained by the models 12

Conclusions **12**
Study limitations 14
Topics for further research 15

Appendix A Data and assessments used in the report **16**

Appendix B Literature review **22**

Appendix C Methodology **24**

Appendix D Confidence intervals for testing differences **34**

Appendix E Multilevel regression model results **44**

Appendix F Predicted differences across Maine High School Assessment domains for student and school characteristics **52**

Appendix G Predictors of Maine High School Assessment domain scores when Preliminary SAT scores are excluded from the model (model 3) **55**

Notes **57**

References **59**

Boxes

1 Data sources and variable definitions 2

2 Methodology 4

Figure

A1 Overlap between the available Maine High School Assessment datasets and the Maine Educational Assessment and Preliminary SAT datasets, 2007/08 18

Tables

1 Maine High School Assessment domain scores regressed on student characteristics, student prior achievement measures, and school characteristics (model 2) 6

2 Summary of percentage of variance explained, by Maine High School Assessment domain 12

A1 Matching stages for linking Maine and College Board student identification numbers 18

A2 Summary of student characteristics, student prior achievement measures, and outcome variables for observed and imputed datasets, 2007/08 20

A3 Summary of school characteristics for the analysis dataset for 115 eligible schools, 2007/08 21

C1 Summary of the intercepts-and-slopes-as-outcomes models explored 26

C2 Description of variables and coding schemes for the multilevel regression models 31

D1 .95 confidence interval around regression coefficients within Maine High School Assessment content domains (based on model 2 results) 34

D2 .95 confidence interval around regression coefficients for Maine Educational Assessment and Preliminary SAT scores across Maine High School Assessment domains (based on model 2 results) 37

D3 .95 confidence interval around regression coefficients for student and school characteristics across Maine High School Assessment content domains (based on Model 2 results) 40

E1 Predictors of Maine High School Assessment reading scores for grade 11 students in the 2007/08 cohort 44

E2 Predictors of Maine High School Assessment writing scores for grade 11 students in the 2007/08 cohort 46

E3 Predictors of Maine High School Assessment math scores for grade 11 students in the 2007/08 cohort 48

E4 Predictors of Maine High School Assessment science scores for grade 11 students in the 2007/08 cohort 50

G1 Maine High School Assessment domain scores regressed on student characteristics, grade 8 Maine Educational Assessment scores, and school characteristics (model 3) 55

Using multilevel regression models to examine how student characteristics, student prior achievement measures, and school characteristics are associated with performance on the Maine High School Assessment, the study finds statistically significant relationships between several of these variables and assessment scores in reading, writing, math, and science.

WHY THIS STUDY?

The Maine Department of Education wanted to use longitudinal data from its data system to better understand whether and how student and school characteristics are associated with student performance on the state-mandated Maine High School Assessment (MHSA). Since 2008 the MHSA, which comprises the College Board's SAT tests in critical reading, writing, and math as well as a science assessment, has been administered in the spring of grade 11 to determine whether Maine high schools have made adequate yearly progress.

The Maine Department of Education is particularly interested in understanding the factors associated with changes in test scores between the beginning of grade 10 and the end of grade 11. This study uses MHSA scores, Preliminary SAT/National Merit Scholarship Qualifying Test (PSAT) data, grade 8 Maine Educational Assessment (MEA) data, and other student and school characteristics to examine the association (see box 1 and appendix A for more on the data used in this report).

The literature examining student and school effects has explored an array of factors potentially associated with student and school outcomes, including and beyond those examined in the current study. Information collected by the Maine Department Education that has been shown to relate to student achievement in the literature includes student characteristics (such as gender, economically disadvantaged status, and special education status), student prior achievement measures, and school characteristics (such as the student–teacher ratio and school location). Appendix B briefly reviews the literature examining the correlates of student performance.

The research question addressed in this report is:

- How are student characteristics, student prior achievement measures, and school characteristics associated with students' grade 11

BOX 1

Data sources and variable definitions

The student and school data for the study were provided by the Maine Department of Education. Longitudinal test score data and demographic information were provided for the cohort of grade 11 students in Maine during the 2007/08 school year who had been enrolled for a full academic year in a publicly funded high school (which includes some private schools; see appendix A). Appendix A describes the dataset matching procedures, cleaning procedures, and handling of missing data. Additional sources of school data include aggregate student information from the student cohort files, school staffing and composition files, and publicly available Maine High School Assessment (MHSA) School Summary Reports for 2007/08 (http://www.maine.gov/ education/ mhsa/08schoolreports/index.html).

This report uses the following student characteristics, student prior achievement measures, and school characteristics as the independent variables and the following outcome variables as the dependent variables. Tables A2 and A3 in appendix A summarize the student and school characteristics. Table C2 in appendix C summarizes how the variables were coded for the multilevel analysis.

Student characteristics

- *Gender.*
- *Racial/ethnic minority*: whether a student is a racial/ethnic minority. Based on a student's No Child Left Behind (NCLB) subgroup, students who were American Indian, Asian/Pacific Islander, Black, or Hispanic were classified as racial/ethnic minority students.
- *Economically disadvantaged*: students eligible for free or reduced-price meals at the time of testing.
- *Special education*: students identified as having a disability under the Individuals with Disabilities Education Act and educated in accordance with an Individual Education Program at the time of testing.
- *English language learner*: students categorized as English language learners in accordance with the NCLB at the time of testing.

Student prior achievement measures

- Grade 8 2004/05 Maine Educational Assessment (MEA) scale scores and standardized scores in reading, writing, math, and science for grade 11 students who took the MHSA in the spring of 2008: in 2004/05, the MEA scale score points ranged from 500 to 580 and were standardized at a mean of 0 and a standard deviation of 1.
- Grade 10 Preliminary SAT (PSAT) verbal, math, and writing scale scores and standardized scores for students who took the MHSA in the fall of 2007: PSAT scale scores range from 20 to 80 and were standardized at a mean of 0 and a standard deviation of 1.

School characteristics

- *Racial/ethnic minority students:* percentage of grade 11 students who are American Indian, Asian/Pacific Islander, Black, or Hispanic.
- *Economically disadvantaged students*: percentage of grade 11 students in the school during the testing period who were categorized as economically disadvantaged.
- *Special education students*: percentage of grade 11 students in the school during the testing period who were categorized as having a disability and who received special education services.
- *English language learner students*: percentage of grade 11 students in the school during the testing period who were categorized as English language learners.
- *Students in cohort who drop out*: percentage of students who would have taken the grade 11 MHSA in 2007/08 had they not left school between grades 9 and 11 prior to graduation for reasons other than transferring to another school without re-enrolling before the following October 1.
- *Student–teacher ratio:* number of students in a school divided by number of teachers (including

BOX 1 (CONTINUED)

Data sources and variable definitions

classroom and special education teachers and literacy specialists) in the school.

- *Mean years of teaching experience:* the average number of years that all teachers in the school (including classroom and special education teachers and literacy specialists) have been teaching in public and private schools.
- *Cohort size:* the number of grade 11 students in a school who were registered during the testing period (does not include dropouts or other students who transferred earlier in the year).
- *School location:* location of the school based on the urban-centric locale coding system: rural (fringe, distant, and remote), city (small), suburban (midsize and small), or town (fringe, distant, and remote) locales. No Maine schools are in large or midsize city locales.
- *Adequate yearly progress status indicators*: whether the school met adequate yearly progress requirements for reading or math in 2007 (former English language learner students are included in the subgroup up to two years after exiting English language learner status) and whether a school was classified as a Title I school (schoolwide or targeted assistance) in 2007.

Outcome variables

- *MHSA scale scores:* the MHSA domain scale scores in reading, writing, math, and science, which range from 1100 to 1180.
- *MHSA standardized scores:* the MHSA domain scale scores in reading, writing, math, and science standardized at a mean of 0 and a standard deviation of 1.

Maine High School Assessment scores in reading, writing, math, and science?

A series of multilevel regression models were applied to student assessment data from a single cohort of students who took the MHSA tests in 2007/08 (see box 2 and appendix C for more on the methodology of this report). Student characteristics, student prior achievement measures, and school characteristics were chosen by the Maine Department of Education based on interest in and availability of data.

Both grade 8 MEA scores and grade 10 PSAT scores measure student achievement prior to taking the MHSA in the spring of grade 11. The Maine Department of Education requested the inclusion of PSAT scores as predictors of MHSA scores partly because PSAT data have never been linked with longitudinal data for Maine. In addition, while the grade 8 MEA is administered in middle school, most students take the PSAT in the school where they take the MHSA. Since the PSAT was designed by the College Board to align with the SAT, a major component of the MHSA, it allows the change from the fall of grade 10 to the spring of grade 11 to be assessed.[1]

The findings from the foundational work in this study will inform further efforts by the Maine Department of Education to understand the association between student and school factors and eventual high school achievement. To the extent that MHSA scores, student characteristics, MEA and PSAT scores, school characteristics, and unmeasured factors remain the same, results of the current study can be generalized to other cohorts of Maine students who complete all four sections of the MHSA in grade 11.[2] Results from the current study could also be used to identify students and schools likely to benefit most from additional assistance. The observational nature of the data means that the findings are not sufficient to support causal inferences, but the data describe the associations between achievement on the MHSA and the student and school characteristics for the most recent cohort of students for which data are available.

BOX 2
Methodology

Multilevel regression modeling was used to examine how student characteristics, student prior achievement on the Maine Educational Assessment (MEA) and the Preliminary SAT (PSAT), and school characteristics are associated with scores on the Maine High School Assessment (MHSA). Separate analyses were conducted for each MHSA domain. See appendix C for a detailed description of the analyses.

To examine how the inclusion of variables could affect the statistical association of other variables in the model, the multilevel regression models were built in three stages:

- Model 1 includes only the MEA and PSAT scores as predictors. Including PSAT scores allows the model to reflect learning that takes place within the high schools being assessed.
- Model 2 adds student and school characteristics to the MEA and PSAT scores as predictors.
- Model 3 includes all the variables in model 2 except PSAT scores.

This report focuses on model 2 (see appendix E for the full results of all three models and appendix G for additional results for model 3). To the extent that grade 8 MEA scores and the student and school characteristics relate to grade 10 PSAT scores as well as MHSA scores, the regression coefficients for the remaining predictor variables may be attenuated relative to a model that excludes PSAT scores. Indeed, when PSAT scores are included (model 2), the coefficients for the non-PSAT variables can to some extent be interpreted as the relationship between those variables and the change in achievement between grades 10 and 11. This is particularly the case with respect to MHSA reading, writing, and math scores. To provide a contrast for the impact of including student and school characteristics, model 1, which includes only student prior achievement and excludes student and school characteristics, is briefly introduced. Model 3, which excludes PSAT scores, provides a contrast for assessing the impact of including PSAT scores.

The results for model 2 include the standardized intercept and regression coefficients for the model and regression coefficients converted to scale score points. Each regression coefficient assumes that all other variables are being held constant. The intercept is an estimate of the standardized MHSA score of a student for whom all variables in the model equal 0. The regression coefficients are the predicted standard deviation difference in MHSA scores for every one unit change in the predictor variable. For discrete variables, the coefficients are the predicted standard deviation increase in MHSA scores associated with going from 0 to 1 on the predictor variable. For the standardized MEA and PSAT scores, the coefficients are the predicted standard deviation difference in MHSA scores for every 1 standard deviation increase in the predictor variable above the mean of 0. For grand mean–centered variables the coefficients are the predicted standard deviation difference in MHSA scores for every one unit increase in the predictor variable above the average of that variable across all students and schools in the analytic sample (table C2 in appendix C shows what a one-unit difference represents for each variable).

Regression coefficients are also presented in scale score points—that is, converted back to their point values on the original scale and rescaled to represent predicted differences in MHSA scores for 10 scale score point changes in the continuous predictor variables (see appendix C for guidelines on interpreting the two forms of regression coefficients along with explanations of the statistical tests).

The multilevel regression models also generated estimates of the total percentage of variance in MHSA scores that is explained by the student and school characteristics and by the MEA and PSAT domain scores. This percentage is analogous to R^2 in a traditional ordinary least squares model in which higher percentages of explained variance are associated with stronger prediction models (see appendix C for more on how the percentages are calculated and table 2 and tables E1–E4 in appendix E for the values).

FINDINGS

This section presents the results of model 2 on the association of student characteristics, student prior achievement measures, and school characteristics with grade 11 MHSA scores; predictors of MHSA domain scores; predicted differences across MHSA domains for measures of prior achievement; implications of including PSAT scores in model 2; and the variance explained by the models.

Association of student characteristics, student prior achievement measures, and school characteristics with grade 11 Maine High School Assessment scores (model 2)

Model 2 regressed MHSA scores on all student characteristics, student prior achievement measures, and school characteristics (table 1). The intercept row in the right panel of table 1 shows the predicted MHSA scale score when all predictors are equal to the grand mean or 0 (depending on how the variables were coded or centered). The predicted MHSA scale scores at the intercept are 1,141.62 for reading, 1,139.45 for writing, 1,142.32 for math, and 1,141.64 for science.

Predictors of Maine High School Assessment domain scores

This section describes the multilevel model regression results presented in table 1 (model 2) for each MHSA domain (reading, writing, math, and science) and focuses on the absolute and relative predictive strength of each measure of prior achievement (MEA domain scores and PSAT domain scores) holding all other variables in the model constant. The association between MHSA domain scores and student characteristics, student prior achievement measures, and school characteristics is also examined. All prior achievement scores were standardized before being entered into the models, allowing the coefficients within models to be compared.

Reading

Student characteristics. Other factors being equal, being male was associated with statistically significant higher MHSA reading scores. But the increase reflected less than one scale score point difference between male and female students' MHSA reading scores (0.76). Racial/ethnic minority status was not a significant predictor of MHSA reading scores.

Economically disadvantaged, special education, and English language learner status were each associated with a decrease in MHSA reading scores

Economically disadvantaged, special education, and English language learner status were each associated with a decrease in MHSA reading scores. But this decrease was less than one scale score point when student prior achievement and school characteristics were included in the model. The standardized coefficients (left panel of table 1) show that special education status was associated with the largest standard deviation decrease in MHSA reading scores (–0.065), followed by English language learner status (–0.061), gender (–0.052), and economically disadvantaged status (–0.035).

Student prior achievement measures. The strongest predictor of MHSA reading scores was PSAT verbal scores, followed by PSAT writing scores. Both PSAT verbal and writing scores were significantly and positively related to MHSA reading scores, with a 1 standard deviation increase in PSAT verbal scores associated with a 0.41 standard deviation increase in MHSA reading scores and a 1 standard deviation increase in PSAT writing scores associated with a 0.22 standard deviation increase in MHSA reading scores. PSAT math scores were not significant predictors of MHSA reading scores. MEA math, reading, and science scores were all significantly and positively associated with MHSA reading scores. Notably, MEA writing scores were significantly and negatively associated with MHSA reading scores; a 1 standard deviation increase in MEA writing scores was associated with a 0.022 standard deviation decrease in MHSA reading scores. The regression coefficient associated with PSAT verbal scores was significantly larger than the regression coefficients

TABLE 1

Maine High School Assessment domain scores regressed on student characteristics, student prior achievement measures, and school characteristics (model 2)

	Standard deviation units				Scale score points			
Variable	Reading	Writing	Math	Science	Reading	Writing	Math	Science
Intercept	0.064	–0.036	0.100	0.095	1,141.62	1,139.45	1,142.32	1,141.64
Student characteristics								
Gender (0 = male, 1 = female)	–0.052**	0.060**	–0.088**	–0.183**	–0.76**	0.83**	–0.97**	–1.69**
Racial/ethnic minority status (0 = no, 1 = yes)	–0.020	–0.036	0.006	0.003	–0.29	–0.50	0.07	0.03
Economically disadvantaged status (0 = no, 1 = yes)	–0.035**	–0.061**	–0.039**	–0.015	–0.51**	–0.84**	–0.43**	–0.14
Special education status (0 = no, 1 = yes)	–0.065**	–0.117**	–0.105**	–0.032	–0.95**	–1.62**	–1.16**	–0.30
English language learner status (0 = no, 1 = yes)	–0.061**	–0.102**	–0.024	–0.100**	–0.89**	–1.41**	–0.26	–0.92**
Student prior achievement measures					Prior achievement scores in units of 10 scale points			
Grade 8 Maine Educational Assessment scores								
Reading	0.138a**	0.120**	–0.024**	0.022	1.67a**	1.38**	–0.22**	0.17
Writing	–0.022a**	0.063**	0.003	–0.050**	–0.36a**	0.98**	0.04	–0.52**
Math	0.058**	0.105**	0.313a**	0.127**	0.57**	0.99**	2.35a**	0.80**
Science	0.147**	0.053**	0.051**	0.314a**	1.80**	0.62**	0.47**	2.45a**
Grade 10 Preliminary SAT scores								
Verbal	0.410**	0.212**	0.056**	0.229a**	5.45**	2.68**	0.56**	1.93a**
Math	0.016	0.051**	0.429a**	0.171a**	0.21	0.64**	4.29a**	1.43a**
Writing	0.222**	0.370**	0.067**	0.071**	2.88**	4.56**	0.66**	0.58**
School characteristics								
Percentage of grade 11 racial/ethnic minority students[b]	–0.016	–0.038	–0.025	0.017	–0.23	–0.53	–0.28	0.16
Percentage of grade 11 students who are economically disadvantaged[b]	–0.013**	–0.006	0.003	0.003	–0.19**	–0.08	0.03	0.03
Percentage of grade 11 students in special education[b]	–0.015	0.002	–0.037	–0.054**	–0.22	0.03	–0.41	–0.50**
Percentage of grade 11 students who are English language learners[b]	0.043	0.058	0.009	–0.061	0.63	0.80	0.10	–0.56
Percentage of students in cohort who drop out[b]	0.037**	0.008	–0.005	–0.008	0.54**	0.11	–0.06	–0.07
Student–teacher ratio	–0.019**	–0.017**	–0.015**	–0.001	–0.28c**	–0.23c**	–0.17c**	–0.01c
Mean years of teaching experience	0.003	0.005	0.001	0.004	0.16	0.25	0.04	0.13
Cohort size[b]	0.002	0.002	0.003**	0.001	0.03	0.03	0.03**	0.01
School location								
City versus rural (0 = rural, 1 = city)	–0.042	–0.003	0.004	0.027	–0.61	–0.04	0.04	0.25
Suburb versus rural (0 = rural, 1 = suburb)	0.025	0.111**	0.029	0.058	0.36	1.53**	0.32	0.54
Town versus rural (0 = rural, 1 = town)	–0.016	–0.019	0.001	–0.019	–0.23	–0.26	0.01	–0.18

(CONTINUED)

TABLE 1 (CONTINUED)

Maine High School Assessment domain scores regressed on student characteristics, student prior achievement measures, and school characteristics (model 2)

Variable	Standard deviation units				Scale score points			
	Reading	Writing	Math	Science	Reading	Writing	Math	Science
Adequate yearly progress indicators								
Made adequate yearly progress in reading in 2007 (0 = no, 1 = yes)[d]	0.023	0.042**	0.064**	0.014	0.33	0.58**	0.71**	0.13
Made adequate yearly progress in math in 2007 (0 = no, 1 = yes)[d]	–0.007	0.027	0.027	0.006	–0.10	–0.37	0.30	0.06
Classified as Title I school in 2007 (0 = no, 1 = yes)	0.021	0.043	0.007	0.014	0.31	0.59	0.08	0.13

** significant at $p < 0.05$.

a. The slopes associated with these prior achievement measures were allowed to vary randomly across schools. See appendix C for statistical justification.

b. Rescaled by 10.

c. Represents the scale score point change for a one-student change in the student–teacher ratio. The discussion of scale point changes in the student–teacher ratio in the text is based on a five-student change.

d. Entered into the model as 0 = yes and 1 = no to make the model intercept interpretable. The coefficients are reverse coded to aid presentation.

Note: The regression coefficients reflect the variability introduced by the imputation procedures because they were derived by combining the results across the 10 imputed datasets.

Source: Authors' calculations based on student data from data files provided by the Maine Department of Education and school data from the Maine Department of Education's Maine High School Assessment Summary Reports (http://www.maine.gov/education/mhsa/08schoolreports/index.html).

associated with PSAT scores in writing and math and was significantly larger than the regression coefficients for MEA reading, writing, math, and science (see table D1 in appendix D).[3]

A 10 scale score point increase in PSAT verbal scores was associated with a 5.45 scale score point increase in MHSA reading scores. A 10 scale score point increase in PSAT writing scores was associated with a 2.88 scale score point increase in MHSA reading scores. The decrease in MHSA reading scores associated with a 10 scale score point increase in MEA writing scores noted above represents a 0.36 scale score point decrease.

School characteristics. Only three school characteristics were statistically significant predictors of students' MHSA reading scores: the percentage of grade 11 students who are economically disadvantaged, the percentage of students in cohort who drop out, and the student–teacher ratio. A 10 point increase in the percentage of students who are economically disadvantaged was associated with a 0.19 scale score point (0.013 standard deviation) decrease in MHSA reading scores. A 10 point increase in the percentage of dropouts in a student's cohort above the mean for schools in the sample was associated with a 0.54 scale score point (0.037 standard deviation) increase in MHSA reading scale scores. An increase of five students per teacher above the grand mean for all schools in the sample (11.32) was associated with a 1.38 scale score point decrease (–0.28*5) in MHSA reading scores.

Writing

Student characteristics. Female students had significantly higher MHSA writing scores than did male students, which translated to approximately 0.83 scale score point (0.060 standard deviation) on the MHSA writing test. Three student characteristics were associated with lower MHSA writing scores: economically disadvantaged status (0.84 scale score point or 0.061 standard deviation drop), special education status (1.62 scale score point or 0.117 standard deviation drop), and English language learner status (1.41 scale score point or 0.102 standard deviation drop).

The strongest significant predictor of MHSA writing scores was PSAT writing scores, followed by PSAT verbal scores

Student prior achievement measures. All student prior achievement variables examined were significantly and positively related to MHSA writing scores. The strongest significant predictor was PSAT writing scores, followed by PSAT verbal scores. A 10 scale score point increase in PSAT writing scores was associated with a 4.56 scale score point (0.370 standard deviation) increase in MHSA writing scores. A 10 scale score point increase in PSAT verbal scores was associated with a 2.68 scale score point (0.212 standard deviation) increase in MHSA writing scores. MEA reading and MEA math scores were significantly stronger predictors of MHSA writing scores than were MEA writing, MEA science, and PSAT math scores. But the regression coefficients associated with MEA reading and math scores were not statistically different from one another, nor were the coefficients associated with MEA math and MEA science scores (see table D1 in appendix D).

School characteristics. The only school characteristics significantly associated with MHSA writing scores were student–teacher ratio, school location, and whether the school made adequate yearly progress in reading in 2007. Students in schools with a student–teacher ratio higher than the grand mean for all schools in the sample were predicted to do worse than their counterparts in schools with lower student–teacher ratios. For example, the results suggest that, other things equal, students attending schools with a student–teacher ratio of 15 would be expected to score an average of 1.17 points higher on the MHSA writing test than students in a school with a student–teacher ratio of 20.

Being in a suburban compared with a rural school was significantly and positively associated with MHSA writing scores (1.53 scale score point or 0.111 standard deviation increase). Being in a school that made adequate yearly progress in reading in 2007 was associated with a 0.58 scale score point or 0.04 standard deviation increase in MHSA writing scores.

Math

Student characteristics. Male students were predicted to outperform female students on the MHSA math test by 0.97 scale score points (0.088 standard deviation). Economically disadvantaged status and special education status were each significantly and negatively associated with MHSA math scores. The decrease associated with being economically disadvantaged was 0.43 scale score point. Although special education students represented only a small portion (12.2 percent) of the cohort examined, being in special education was associated with a 1.16 scale score point (0.105 standard deviation) drop in MHSA math scores.

Student prior achievement measures. MEA math and science scores and PSAT verbal, math, and writing scores were significantly and positively related to MHSA math scores. MEA writing scores were not statistically associated with MHSA math scores. MEA reading scores were significant and negatively associated with MHSA math scores: a 10 scale score point increase in MEA reading scores above the mean was associated with a 0.22 scale score point (0.024 standard deviation) decrease in MHSA math scores. The strongest significant and positive predictor of MHSA math scores among the prior achievement variables was PSAT math scores, followed by MEA math scores (see table D1 in appendix D). A 10 point increase in PSAT math scores was associated with a 4.29 scale score point (0.429 standard deviation) increase in MHSA math scores. A 10 scale score point increase in MEA math scores was associated with a 2.35 scale score point (0.313 standard deviation) increase in MHSA math scores.

A 10 scale score point increase in MEA science, PSAT writing, or PSAT verbal scores was associated with less than a 1 scale score point increase in MHSA math scores. The regression coefficient associated with PSAT math scores was significantly larger than the regression coefficient associated with MEA math scores, which was larger than the regression coefficient associated with MEA science, PSAT writing, and PSAT verbal scores;

the coefficients for MEA science, PSAT writing, and PSAT verbal scores were not statistically significant from one another (see table D1 in appendix D).

School characteristics. School characteristics significantly associated with MHSA math performance were student–teacher ratio, cohort size, and whether the school made adequate yearly progress in reading in 2007. Increasing the number of students per teacher by five was associated with a 0.83 (–0.17*5) scale score point decrease in MHSA math scores. A 10-student increase in cohort size above the grand mean for all schools in the sample was associated with a statistically significant 0.03 scale score point increase in MHSA math scores.

Making adequate yearly progress in math in 2007 was not a significant predictor of MHSA math scores, but making adequate yearly progress in reading in 2007 was statistically and positively associated with MHSA math scores. A student in a school that made adequate yearly progress in reading in 2007 was predicted to score about 0.71 scale score point higher on the MHSA math assessment than a student in a school that did not.

Science

Student characteristics. Being female was associated with a 1.69 scale score point (0.183 standard deviation) decrease in MHSA science scores. Being economically disadvantaged, a member of a racial/ethnic minority, or in special education was not significantly related to MHSA science scores. Being an English language learner student was significantly and negatively related to MHSA science scores and associated with a 0.92 scale score point (0.100 standard deviation) decrease in science scores.

Student prior achievement measures. The strongest student prior achievement measure associated with MHSA science scores was MEA science scores: a 1 standard deviation increase in MEA science scores was associated with a 0.314 standard deviation increase in MHSA science scores, and a 10 scale score point increase in MEA science scores was associated with a 2.45 scale score point increase in MHSA science scores. PSAT verbal scores were the next strongest predictor: a 1 standard deviation increase in PSAT verbal scores was associated with a 0.229 standard deviation increase in MHSA science scores, and a 10 scale score point increase in PSAT verbal scores translated to a 1.93 scale score point increase in MHSA science scores. PSAT math scores (0.171 standard deviation) were the third strongest predictor, followed by MEA math scores (0.127 standard deviation). A 10 scale score point increase in PSAT math scores was associated with a 1.43 scale score point increase in MHSA science scores, and a 10 scale score point increase in MEA math scores was associated with a 0.80 scale score point increase.

PSAT writing scores were also significantly and positively associated with MHSA science scores. A 10 scale score point increase in PSAT writing scores was associated with a 0.58 scale score point (0.071 standard deviation) increase in MHSA science scores. MEA writing scores, on the other hand, were a significant but negative predictor of MHSA science scores. A 10 scale score point increase in MEA writing scores was associated with a 0.52 scale score point (0.05 standard deviation) decrease in MHSA science scores. All regression coefficients for the prior achievement variables were statistically different from one another (see table D1 in appendix D).

School characteristics. Percentage of grade 11 students in special education was the only school characteristic examined that was significantly associated with MHSA science scores. A 10 point increase in the percentage of grade 11 students in special education above the grand mean for all schools was associated with a 0.50 scale score point (0.054 standard deviation) decrease in MHSA science scores.

School characteristics significantly associated with MHSA math performance were student–teacher ratio, cohort size, and whether the school made adequate yearly progress in reading in 2007

Predicted differences across Maine High School Assessment domains by measure of prior achievement

This section compares the predicted differences across MHSA domains (reading, writing, math, and science) within each MEA domain (reading, writing, math, and science) and PSAT domains (verbal, math, and writing) (see table 1). Appendix F presents the predicted differences across MHSA domains for student and school characteristics.

The discussion focuses on the predicted scale score point differences and the .95 confidence intervals for comparing the strength of the standardized relationships across MHSA domains for a single measure of prior achievement, holding all other variables constant (see table D2 in appendix D for the .95 confidence intervals). This comparison allows drawing conclusions about whether a measure of prior achievement is a "significantly stronger" predictor in one MHSA domain than in another. Although 10 scale score point differences are not equivalent within a single model across the four MEA and three PSAT domains (because the scores in each domain have different standard deviations), it is possible to compare the predicted difference in MHSA scores associated with a 10 scale score point change within a MEA domain or PSAT domain across MHSA domains. So, for example, it is possible to compare how a 10 scale score point difference in MEA reading scores predicts differences in MHSA reading, writing, math, and science scores.

MEA reading scores were significant predictors of MHSA reading, writing, and math scores; the largest predicted difference associated with a 10 scale score point increase in MEA reading scores was in MHSA reading scores (1.67 scale score points increase)

Maine Educational Assessment reading scores

MEA reading scores were significant predictors of MHSA reading, writing, and math scores but not of MHSA science scores. The largest predicted difference associated with a 10 scale score point increase in MEA reading scores was in MHSA reading scores (1.67 scale score points increase) and MHSA writing scores (1.38 scale score points increase). A 10 scale score point difference in MEA reading scores was associated with a decrease in MHSA math scores of 0.22 scale score point. The relationship between MEA and MHSA reading scores is significantly stronger than the relationship between MEA reading scores and MHSA math and science scores. There was no difference in the strength of the relationship between MEA and MHSA reading scores and the relationship between MEA reading scores and MHSA writing scores.

Maine Educational Assessment writing scores

MEA writing scores were significant predictors of MHSA reading, writing, and science scores but not of MHSA math scores. The largest predicted difference associated with a 10 scale score point difference in MEA writing scores was in MHSA writing scores (0.98 scale score point increase). The predicted scale score point changes in MHSA reading and science scores associated with a 10 scale score point increase in MEA writing scores were both negative. In this case, a 10 scale score point increase in MEA writing scores was associated with a 0.36 scale score point decrease in MHSA reading scores and a 0.52 scale score point decrease in MHSA science scores. The relationship between MEA and MHSA writing scores was significantly stronger than the relationships between MEA writing scores and MHSA reading, math, and science scores.

Maine Educational Assessment math scores

MEA math scores were significant and positive predictors of MHSA reading, writing, math, and science scores. When comparing the predicted MHSA scale score point changes associated with a 10 scale score point increase in MEA math scores, the largest change was in MHSA math scores. For every 10 scale score point increase in MEA math scores, MHSA math scores were predicted to increase 2.35 scale score points. The predicted magnitude of increase associated with a 10 scale score point increase in MEA math scores was lower for MHSA reading (0.57 scale score point), writing

(0.99 scale score point), and science scores (0.80 scale score point). The relationship between MEA and MHSA math scores is significantly stronger than the relationships between MEA math scores and MHSA writing, reading, and science scores.

Maine Educational Assessment science scores

MEA science scores were significant and positive predictors of MHSA reading, writing, math, and science scores, and the largest predicted difference associated with a 10 scale score point increase in MEA science scores was in MHSA science scores (2.45 scale score point increase), followed by MHSA reading scores (1.80 scale score point increase). The relationship between MEA and MHSA science scores was significantly stronger than the relationships between MEA science scores and MHSA reading, writing, and math scores.

Preliminary SAT verbal scores

A 10 scale score point increase in PSAT verbal scores was associated with a significant and positive increase in MHSA reading, writing, math, and science scores. When comparing the predicted MHSA scale score point differences associated with a 10 scale score point increase in PSAT verbal scores, the largest increase occurred for MHSA reading scores (5.45 scale score points). The predicted magnitude of increase associated with a 10 scale score point increase in PSAT verbal scores was lower for MHSA writing (2.68 scale score points), math (0.56 scale score point), and science scores (1.93 scale score points). The relationship between PSAT verbal and MHSA reading scores was significantly stronger than the relationships between PSAT verbal scores and MHSA writing, math, and science scores.

Preliminary SAT math scores

PSAT math scores were significant and positive predictors of MHSA writing, math, and science scores but not MHSA reading scores. The largest predicted increase associated with a 10 scale score point increase in PSAT math scores was in MHSA math scores (4.29 scale score points), followed by MHSA science scores (1.43 scale score points) and MHSA writing scores (0.64 scale score point). The relationship between PSAT and MHSA math scores was significantly stronger than the relationships between PSAT science scores and MHSA reading, writing, and science scores.

MEA science scores were significant and positive predictors of MHSA reading, writing, math, and science scores, and the largest predicted difference associated with a 10 scale score point increase in MEA science scores was in MHSA science scores (2.45 scale score point increase)

Preliminary SAT writing scores

A 10 scale score point increase in PSAT writing scores was associated with a significant and positive increase in MHSA reading, writing, math, and science scores. The largest predicted increase was in MHSA writing scores (4.56 scale score points), followed by MHSA reading scores (2.88 scale score points). The predicted differences in MHSA math and science scores associated with a 10 scale score point increase in PSAT writing scores were similar, at 0.66 and 0.58 scale score point, respectively. The relationship between PSAT and MHSA writing scores was significantly stronger than the relationships between PSAT writing scores and MHSA reading, math, and science scores.

Implications of including Preliminary SAT scores in model 2

Contrasting the results of model 2 with those of model 3, which excludes PSAT scores as predictors, highlights the influence of PSAT scores on the associations of other variables in the model to MHSA scores. Except for MEA writing scores predicting MHSA reading scores, the coefficients associated with the grade 8 MEA scores were larger[4] in model 3 than in model 2 (see tables E1–E4 in appendix E and table G1 in appendix G). The coefficient associated with MEA reading scores for predicting MHSA math scores reversed direction when PSAT scores were removed from the model, changing from –0.024 to 0.034. MEA reading scores predicting MHSA science scores was the only student prior achievement measure predictor to go from nonsignificant (0.022) to significant (0.118). Similarly, the magnitudes of the regression coefficients for student and school characteristics also changed. In many cases the coefficients increased in model 3, and in several cases the predictors went from nonsignificant to significant (for example, the coefficient representing the difference between urban and rural schools in predicting MHSA reading, math, and science scores). Including or excluding key variables in the model can thus affect the interpretation of individual model regression coefficients even when all variables are held constant.

Variance explained by the models

The percentage of variance explained by a model provides a way to assess how well each model accurately reproduces the observed student scores. The greater the percentage of variance accounted for by the model, the better the model describes the data (a more complete description is in appendix C).

Comparing the variance explained by model 1 with that explained by models 2 and 3, when all the variables are considered together, confirms a strong association between student prior achievement measures and MHSA scores (table 2; full results are in tables E1–E4 in appendix E). Model 1, containing only student prior achievement measures, predicted 78.9 percent of the differences in MHSA reading scores, 77.8 percent of the differences in MHSA writing scores, 76.8 percent of the differences in MHSA math scores, and 67.9 percent of the differences in MHSA science scores. Comparing the explained differences with the percentage of the variance accounted for by model 2, containing all examined student and school characteristics as well as student prior achievement measures, shows that the difference in variance explained by each model is less than 1 percentage point. When PSAT scores are excluded (model 3), the percentage of variance explained in MHSA scores is lower.

TABLE 2

Summary of percentage of variance explained, by Maine High School Assessment domain

	Reading	Writing	Math	Science
Model 1	78.9	77.8	76.8	67.9
Model 2	79.1	78.4	77.3	68.7
Model 3	62.7	63.0	66.6	61.5
Scale score means	1,140.69	1,139.95	1,141.22	1,140.76

Note: Complete models are shown in tables E1–E4 in appendix E.

Source: Authors' calculations based on student data from data files provided by the Maine Department of Education and school data from the Maine Department of Education's Maine High School Assessment Summary Reports (http://www.maine.gov/education/mhsa/08schoolreports/index.html).

CONCLUSIONS

This study presents the findings from multilevel regression models that were applied to examine the association between grade 11 MHSA reading, writing, math, and science scores and student characteristics, student prior achievement measures, and school characteristics. The regression coefficients in the multilevel regression models represent the association between a student or school variable and the examined MHSA score when holding all other student and school variables in the model constant. Grade 8 MEA scores in reading, writing, math, and science served as measures of achievement prior to entry into high school. PSAT verbal, writing, and math scores served as measures of achievement in high school. When PSAT scores are included in the model to predict MHSA scores, the coefficients for the non-PSAT variables may be interpreted as the relationship between those variables and the change in achievement between grade 10 and grade 11. To provide a contrast for assessing the impact of including PSAT scores in the model, models that excluded grade 10 PSAT scores were also formulated.

The largest standardized regression coefficient when PSAT scores were included was the PSAT score most closely aligned with the MHSA domain being predicted. The exception was the model used to predict MHSA science scores. Because no PSAT assessment was directly aligned with the content assessed in the MHSA science domain, the largest standardized regression coefficient predicting MHSA science scores in model 2 was associated with grade 8 MEA science scores.

When the regression coefficients associated with student prior achievement measures in model 2 were compared with those in model 3, which excluded PSAT scores, the observed regression coefficients for MEA scores were larger in model 3. The observed MHSA scale score point increase associated with a given MEA score was also larger when PSAT scores were not included in the model. Understanding how the differences between models 2 and 3 relate to the utility of each of these models for informing decisions rests again on the interpretation of individual regression coefficients when other variables are held constant in the model.

The results from model 2—the model that includes scores on the grade 10 PSAT, which is designed to align with the SAT, which makes up three of the four MHSA domains—show that the relationship of other student and school characteristics to MHSA outcomes can be interpreted as their unique association with MHSA scores over and above that described by PSAT scores. Model 2 is thus potentially useful for examining how student and school factors are associated with change between the time when the grade 10 PSAT is taken in the fall to when the grade 11 MHSA is taken in spring of the following year. Model 3, which excludes PSAT scores, describes the association between student and school factors and eventual performance on the grade 11 MHSA over a longer time frame.

Even when prior achievement and other variables are accounted for in the model, being male was associated with higher MHSA scores in all but the writing domain—a finding that held regardless of whether PSAT scores were included. Assessment results that do not control for prior achievement and other factors, such as those reported by National Assessment of Educational Progress, typically report that female students outscore male students in reading and writing, while male students outscore female students in math and science. Because the findings for gender in this study were unexpected, the models were reformulated with only gender as a level-one predictor. The results from these models confirmed the patterns previously reported for other assessments: female students outscored male students in the MHSA reading and writing domains, and male students outscored female students in the math and science domains.[5]

The cohort dropout rate was a significant predictor of MHSA reading scores but not a significant predictor of MHSA writing, math, or science scores

Holding all other variables constant, economically disadvantaged students, special education students, and English language learner students were predicted to have MHSA reading and writing scores that were significantly lower than non–economically disadvantaged, general education, and non–English language learner students. These results are typical of those reported in the literature when other factors are not controlled for. Students who are economically disadvantaged and students in special education were each predicted to have MHSA science scores that were similar to those of non–economically disadvantaged and general education students. However, the regression coefficients indicated that there was no significant difference between English language learner and non–English language learner students' MHSA math scores.

The cohort dropout rate was a significant predictor of MHSA reading scores but not a significant predictor of MHSA writing, math, or science scores. While reasons for this result are not apparent from the data, other research such as Rumberger and Palardy's (2005) study of how school characteristics influence test scores, dropout rates, and other indicators of student performance found that high schools that are effective in promoting student learning are not necessarily more effective in reducing dropout rates. In model 2 the student–teacher ratio was a significant predictor of MHSA outcomes in reading, writing, and math, with higher student–teacher ratios associated with lower MHSA scores.

The student and school characteristics examined in this analysis are by no means an exhaustive list of potential factors that could influence performance on the MHSA. Further, the addition of other variables could change the observed relationship between currently examined student and school characteristics and MHSA scores.

Study limitations

This study has six main limitations that should be taken into accounting when interpreting the findings.

First, although this study finds statistically significant relationships between student prior achievement and MHSA scores after controlling for several student and school characteristics, these relationships represent correlations, which do not imply causation. The multilevel regression analysis describes statistical associations rather than causal relationships between student and school characteristics, including student prior achievement, and MHSA scores. Thus, this report does not support any conclusions about causality. The grade 11 students examined took the MHSA after completing the grade 8 MEA and grade 10 PSAT, but this report does not provide direct evidence that higher prior achievement causes or leads to higher MHSA scores. Unmeasured factors (such as student motivation or access to high-quality teachers) may have raised both prior achievement scores and MHSA scores.

Although this study finds statistically significant relationships between student prior achievement and MHSA scores after controlling for several student and school characteristics, these relationships represent correlations, which do not imply causation

Second, the student and school characteristics included in the models were collected by the Maine Department of Education and the College Board and do not represent all possible variables related to student performance on the MHSA. For example, individual or family factors such as student motivation, parental involvement, or parental education expectations may be related to MHSA scores. If measures of these characteristics were available, including them in the analyses might have strengthened the models and accounted for a greater proportion of the variance in performance on the MHSA.

Third, small predicted differences in MHSA domain scores were statistically significant in the multilevel regression models. When considering any policy relevance associated with the findings, the practical significance of the small point differences identified in this study must be taken into account.

Fourth, the grade 8 student prior achievement data consisted of available grade 8 MEA scores in reading, writing, math, and science. Beginning in 2009, the Maine Department of Education replaced the MEA in grades 3–8 with the New England Common Assessment Program (NECAP), an assessment used in New Hampshire, Rhode Island, and Vermont. While the current study shows that grade 8 student prior achievement is a significant predictor of grade 11 MHSA domain scores, possible differences between the MEA and NECAP in coverage, focus, and the like make findings from the current association between grade 8 achievement and grade 11 MHSA performance potentially less relevant. Further work needs to be done to explore whether similar relationships exist between middle school NECAP scores and MHSA performance.

Fifth, there are limitations regarding the generalizability of the study's findings. The results of the current study can be generalized only to Maine students who completed all four domains of the MHSA and not to the larger population of test takers. As such, the findings from this study may not be generalized to all states. In addition, the samples were dominated by non-Hispanic White students who were not special education or English language learner students and who were living in rural areas. The percentages of racial/ethnic minority students, students who are economically disadvantaged, and English language learner students for this study were different from regional and national averages.

Sixth, while every effort was made to account for errors associated with using imputed scores instead of observed scores, the possibility remains that imputing missing data led to additional noise in the data and in the subsequent models. Further, any error due to imperfect matching of PSAT data with the Maine Department of Education files could adversely impact imputation results. While the sample used for the analyses presented here is very similar to the population of students in the 115 eligible schools, the possibility of unintended differences between the analytic sample and the population remains.

Topics for further research

The findings of this study combined with its limitations suggest several areas for further research. While the results show an association between student prior achievement and performance on the MHSA, the extent to which school factors influence this relationship remains largely unexplained.

School characteristics included in the analysis did little to explain the differences in MHSA scores between schools when included with prior achievement. Specifically, school characteristics explained less than 1 percent of the variance in MHSA scores over the variance explained by MEA and PSAT scores. An area for additional research would involve identifying and measuring additional and more fine-grained school information to improve overall understanding of how school characteristics are associated with student achievement. Additional school variables that could be examined include school climate, resources, programs, leadership styles, teaching strategies, and teacher satisfaction. Looking at other earlier measures of student prior achievement could also prove useful for future research. Similarly, further research could focus specifically on examining factors that may explain improvement in performance at both the student and school levels.

While the results of this study show an association between student prior achievement and performance on the MHSA, the extent to which school factors influence this relationship remains largely unexplained

Student–teacher ratio was the most consistent school-level predictor of MHSA performance in the model. Specifically, a higher student–teacher ratio was a significant and negative predictor of MHSA reading, writing, and math scores. This relationship, while only correlational, is consistent with the research findings on class size and student–teacher ratio.

Additional research to examine the predictors of success for special populations such as racial/ethnic minority, special education, and English language learner students is warranted. Targeted research directly examining the factors relating to specific subgroups could potentially provide valuable insight into their specific needs. For example, a previous Regional Educational Laboratory Northeast and Islands Issues & Answers report, *New Measures of English Language Proficiency and Their Relationship to Performance on Large-Scale Content Assessments,* (Parker, Louie, and O'Dwyer 2009) examined how English language proficiency measures were related to NECAP outcomes in New Hampshire, Rhode Island, and Vermont. Another related Regional Educational Laboratory Northeast and Islands Issues & Answers report examined the performance patterns of Hispanic high school students on the Massachusetts state assessment (Sánchez et al. 2009).

APPENDIX A
DATA AND ASSESSMENTS USED IN THE REPORT

Data from several sources were merged to create the dataset examined in this report. This appendix presents information on the assessments used to create the dataset, the procedures for merging datasets, and the methods for imputing data.

Data sources

The dataset examined in this report was created by merging student domain scores on the Maine Educational Assessment (MEA), Preliminary SAT (PSAT), and the Maine High School Assessment (MHSA) and demographic data from the Maine Department of Education and combining those data with school data from the MHSA School Summary Reports 2007–2008 (http://www.maine.gov/education/mhsa/08schoolreports/index.html).

The Maine Department of Education uses the MEA and MHSA in its accountability system under the No Child Left Behind Act of 2001 (NCLB) and requires that both be aligned to state content standards. Although used for NCLB reporting, the assessments do not affect student grades. Before the SAT could be approved as a component of the state's accountability system, (the SAT is used for the reading, writing, and math portions of the MHSA), the Maine Department of Education had to ensure that the SAT was sufficiently aligned to the state's content standards. The PSAT is not used as part of the state's accountability system, but the state commissioned an alignment study of both the SAT and PSAT with its content standards (College Board 2005). The SAT and PSAT were found to be sufficiently aligned to Maine content standards, but factors other than student learning might influence performance from one test to the next.

The 2007/08 cohort of grade 11 students completed the MEA tests in reading, writing, math, and science in grade 8 (during 2004/05). The MEA is used to measure student progress toward achieving the state standards known as Learning Results (adopted in 1997). The MEA is administered in grades 3–8 and is used to meet state assessment and NCLB requirements. The reading portion includes questions related to literary and informational reading passages. The writing portion presents a narrative writing prompt. In grade 8 the MEA covers the following content areas in math: numbers and number sense; computation; data analysis and statistics; probability; geometry; measurement; patterns, relations, and functions; algebra concepts; and mathematical communication. In science, the MEA covers life sciences (classifying life forms, ecology, and cells), physical sciences (structure of matter, energy, and motion), earth and space sciences (continuity and change, the Earth, the universe), and the nature and implications of science (inquiry and problem solving, scientific reasoning, communication, implications of science and technology).

The 2007/08 cohort of grade 11 students had also completed the PSAT in their second and third years of high school. According to the College Board (http://collegeboard.com), the PSAT measures the reading, mathematical reasoning, and writing skills essential for success in college; the knowledge and skills acquired through school and outside the classroom; and the ability to think critically by reasoning with facts and concepts.

For the 2007/08 cohort of grade 11 students, the MHSA comprised the SAT in reading, writing, and math and a science and technology assessment developed from Maine's previous high school accountability assessment. The math portion of the SAT was augmented with a 25-item test to ensure that math results were properly aligned with the state's content standards. As part of Maine's state and federal accountability systems, all third-year high school students in Maine are required to complete all components of the MHSA (http://www.maine.gov/education/mhsa/index.htm).

MHSA School Summary Reports for 2007/08 display student achievement results for each school that receives at least 60 percent of funds from the public sector. The MHSA results for reading, writing, math, and science are disaggregated by student and school characteristics and reported according to Maine's achievement standards (exceeds, meets, partially meets, and does not meet). In the current analyses, the School Summary Reports provide an official school record that includes the percentages for the following student groups: gender, race/ethnicity, economically disadvantaged status, special education status, and English language learner status.

The final dataset covers only publicly funded schools in Maine, which includes public and some private schools. Private schools receiving public funds are entities that receive 60 percent or more of their funds from the public sector and are included in accountability determinations at the school administration unit and state levels. Of 159 schools, 116 schools were publicly funded. Of those 116 schools, 1 did not have school data available and was dropped from the sample. The remaining 115 schools with complete school data make up the eligible school sample and are referred to as public schools in this report for convenience, even though they include some private schools that receive public funding. During the 2007/08 school year, 14,065 students in grade 11 were enrolled in the 115 schools. Of those, 1,057 were dropped from the sample because they did not have any MHSA, MEA or PSAT scores (343 students), because they had all three PSAT scores but no MHSA or MEA scores (85 students), because they had all four MEA scores but no MHSA or PSAT scores (350 students), or because they had only PSAT and MEA scores but no MHSA scores (279 students). The resulting dataset included 13,008 students who had all four MHSA domain scores.

Procedures for merging datasets

Student MEA scores were linked to their MHSA domain scores by state-assigned unique student identification numbers. Student PSAT scores, provided to the Maine Department of Education by the College Board, did not contain the unique student identification number. After consultation with the Maine Department of Education and the College Board, it was determined that student names and dates of birth along with other demographic variables available for both datasets would provide the information necessary to build a link file to match PSAT data with Maine Department of Education data for the current analysis. The resulting link file would ultimately contain the unique student identification number used by the Maine Department of Education to track students with matched PSAT identification numbers kept by the College Board.

Because of individually identifiable information was used to build the link file, procedures for handling secure data were outlined in the data agreement between the Maine Department of Education and Regional Educational Laboratory Northeast and Islands. Student demographic information used to link files was stored on an encrypted drive kept in a locked cabinet for the duration of the study. While student names were not kept in the analysis files, the inclusion of necessary demographic information in the analysis files required these security procedures.

The common variables used to link Maine Department of Education student identification numbers to College Board student identification numbers were first name, last name, date of birth, gender, and school identification number. Because both datasets had incomplete data for the five matching variables, the matching was done in seven stages. Each matching stage was performed on the remaining unmatched data from subsequent stages. Only unique matches were retained and included in the subsequent link file. Table A1 shows the number of student PSAT identification numbers that were uniquely matched to Maine Department of Education student identification numbers in each stage.

All names were capitalized in both files to standardize format. No allowances were made for possible nicknames, alternative spellings, or changes in last names when conducting the initial matches. The underlying rationale for the matching strategy was that a smaller set of more accurate matches

TABLE A1

Matching stages for linking Maine and College Board student identification numbers

	Records matched		Matching variables				
Stage	Stage total	Total	First name	Last name	Date of birth	Gender	School identification number
1	8,740	8,740	✓	✓	✓	✓	✓
2	217	8,957	✓	✓	✓	✓	
3	41	8,998	✓	✓	✓		✓
4	640	9,638	✓	✓		✓	✓
5	622	10,260		✓	✓	✓	
6	426	10,686		✓		✓	✓
7	39	10,725		✓	✓		✓

Source: Authors' calculations based on student data from data files provided by the Maine Department of Education and school data from the Maine Department of Education's Maine High School Assessment Summary Reports (http://www.maine.gov/education/mhsa/08schoolreports/index.html).

would be less likely to introduce error into subsequent imputations than would a larger set with lower accuracy.

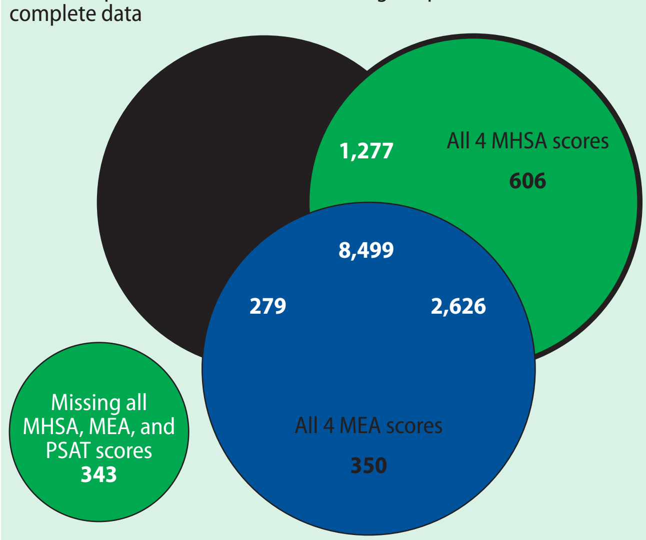
The initial PSAT file contained more than 25,000 student records. Because the data had not been collected with the intent of linking it with Maine Department of Education files, the data included all students who had taken the PSAT at the same time as the analytic sample. The first, most conservative stage identified 8,740 unique matches based on all five matching variables. Six additional stages were sequentially performed using different combinations of matching variables. After stage 7, 10,725 student identification numbers from the PSAT file were matched to student identification numbers in the files provided by the Maine Department of Education.

Considering only the PSAT scores that were matched to the 13,008 students with all four MHSA scores in the 115 schools for which data were available, 9,810 PSAT verbal scores, 9,805 PSAT math scores, and 9,778 PSAT writing scores were successfully matched. Ultimately, 9,776 cases with all three PSAT scores were matched to the 13,008 cases with all four MHSA scores that were selected as the final student sample.

FIGURE A1

Overlap between the available Maine High School Assessment datasets and the Maine Educational Assessment and Preliminary SAT datasets, 2007/08

Initial sample: 14,065 students in 115 eligible public schools with complete data



Note: The number of students within the boldface border (13,008) is the final analytic sample (analysis dataset with imputed cases).

Source: Authors' calculations based on student data from data files provided by the Maine Department of Education and school data from the Maine Department of Education's Maine High School Assessment Summary Reports (http://www.maine.gov/education/mhsa/08schoolreports/index.html).

Figure A1 shows the overlap of student cases with all four MEA scores and all three PSAT scores and student cases with all four MHSA scores. Of these 13,008 student cases, after matching, 8,499 had MEA, PSAT, and MHSA scores. For the remaining cases, missing MEA or PSAT scores were imputed.

Not all of the 13,008 students who had all four MHSA scores had all grade 8 MEA scores and grade 10 PSAT scores available. Of the 13,008 students with all four MHSA scores, 1,847 (14.20 percent) were missing MEA math scores, 1,850 (14.22 percent) were missing MEA reading scores, 1,867 (14.35 percent) were missing MEA writing scores, and 1,837 (14.12 percent) were missing MEA science scores. The 13,008 students were also missing 3,198 (24.58 percent) PSAT verbal scores, 3,203 (24.62 percent) PSAT math scores, and 3,230 (24.83 percent) PSAT writing scores.

Imputation methods

For the 13,008 students who had all of the MHSA scores, missing MEA and PSAT scores were imputed using multiple imputation procedures (Little and Rubin 1987; Rubin 1987). Multiple imputed datasets were created to represent a random sample of the missing values, thereby accounting for the uncertainty associated with using imputed values. According to Little and Rubin (1987), imputing 5–10 plausible values is sufficient for estimating the variability in the imputed scores, so 10 plausible values were imputed for each missing MEA and PSAT case. As recommended by Little and Rubin (1987) and Allison (2002), to avoid bias in the imputed scores, all available variables were used to impute the missing cases. Specifically, a Markov chain Monte Carlo procedure using multiple chains within the SAS software multiple imputation procedure was applied to create 10 imputed datasets. Student characteristics, school characteristics, and other available MEA and PSAT scores as well as the outcome MHSA domain scores were used to impute missing MEA and PSAT scores. No data were missing from the student or school characteristics, so none of these variables were imputed.

The imputation model assumes multivariate normality and data missing at random. While this assumption was not directly tested, the distribution of MEA and PSAT scores was examined first by looking at the distribution of each statistic and then by regressing each variable on all others. Residuals were approximately normally distributed. Possible clustering within schools was not accounted for in the imputations themselves. A subsequent examination that included school dummy variables in the model showed a negligible effect of not accounting for clustering in the original imputation models.

Once the 10 multiple-imputed datasets were created, standard analysis procedures were employed for each dataset. AM Statistical Software (http://am.air.org) was used to produce the descriptive results so that the means and standard errors could be calculated simultaneously using the 10 multiple-imputed datasets and so that the clustering of students within schools was accounted for in the calculations. Using HLM software (version 6.08), models were simultaneously formulated with the 10 multiple-imputed datasets, producing correct means and standard errors for the parameter estimates. The final sample (analysis dataset with imputed cases) comprised 13,008 students in 115 publicly funded schools in Maine. Table A2 compares the original observed sample of students in the 115 eligible schools (original dataset for 115 eligible schools) and the analysis dataset with imputed cases. Table A3 summarizes the school characteristics for the 115 eligible schools (no school characteristics were missing, so none required imputing).[6]

The student characteristics of the analysis dataset with imputed cases were very similar to those of the original dataset for 115 eligible schools. In the analysis dataset with imputed cases, 95.5 percent of students were not of a racial/ethnic minority White students (compared with 95.0 percent in the original dataset for 115 eligible schools), 75.0 percent were not economically disadvantaged (compared with 73.3 percent), 87.8 percent were general education students (compared with 85.8 percent), and 97.0 percent were not English language learner students (compared with 96.1 percent).

TABLE A2

Summary of student characteristics, student prior achievement measures, and outcome variables for observed and imputed datasets, 2007/08

Item	Original dataset for 115 eligible schools					Analysis dataset with imputed cases[a]				
Student characteristics	Total *N*	No *n*	No Percent	Yes *n*	Yes Percent	Total *N*	No *n*	No Percent	Yes *n*	Yes Percent
Male		7,125	50.7	6,940	49.3		6,553	50.4	6,455	49.6
Racial/ethnic minority		13,357	95.0	708	5.0		12,420	95.5	588	4.5
Economically disadvantaged	14,065	10,316	73.3	3,749	26.7	13,008	9,758	75.0	3,250	25.0
In special education		12,066	85.8	1,999	14.2		11,422	87.8	1,586	12.2
English language learner		13,520	96.1	545	3.9		12,619	97.0	389	3.0
Student prior achievement measures	Total *N*		Mean	Standard error	Standard deviation	Total *N*	Mean	Standard error	Standard deviation	Number imputed
Grade 8 Maine Educational Assessment scores										
Reading	11,798		538.17	0.36	12.21		538.51	0.33	11.99	1,850
Writing	11,781		537.95	0.29	8.97	13,008	538.18	0.27	8.85	1,867
Math	11,798		532.44	0.45	14.72		532.78	0.42	14.68	1,847
Science	11,815		530.39	0.35	11.93		530.75	0.33	11.85	1,837
Grade 10 Preliminary scores[b]										
Verbal	10,178		44.49	0.37	10.94		43.57	0.32	10.94	3,198
Math	10,174		44.63	0.37	11.13	13,008	43.71	0.33	11.02	3,203
Writing	10,142		43.21	0.39	11.19		42.16	0.34	11.22	3,230
Outcome variables										
Grade 11 Maine High School Assessment scores										
Reading	13,174		1,140.60	0.43	14.58		1,140.69	0.43	14.55	na
Writing	13,176		1,139.86	0.45	13.83	13,008	1,139.95	0.46	13.82	na
Math	13,433		1,140.87	0.34	11.08		1,141.22	0.33	11.02	na
Science	13,338		1,140.57	0.27	9.29		1,140.76	0.26	9.23	na
Listwise	8,499					13,008				

na is not applicable.

a. The means and standard errors were calculated using the 10 multiple-imputed datasets, and the clustering of students within schools was accounted for in the calculations. AM Statistical Software was used to conduct these analyses (http://am.air.org).

b. The number of imputed PSAT scores is larger than the difference between the number in the original dataset for 115 eligible schools and the 13,008 cases in the final sample because 279 students had PSAT scores but did not have all four MHSA scores.

Source: Authors' calculations based on data files provided by the Maine Department of Education.

TABLE A3

Summary of school characteristics for the analysis dataset for 115 eligible schools, 2007/08

Characteristic	*n*	Mean	Standard error	Standard deviation
Percentage of grade 11 racial/ethnic minority students	115	4.16	0.55	5.88
Percentage of grade 11 students who are economically disadvantaged	115	30.63	1.55	16.63
Percentage of grade 11 students in special education	115	15.42	0.61	6.52
Percentage of grade 11 students who are English language learners	115	3.02	0.43	4.59
Percentage of students in cohort who drop out	115	6.39	0.44	4.76
Student–teacher ratio	115	11.32	0.20	2.13
Mean years of teaching experience	115	16.11	0.26	2.78
Cohort size	115	12.23	0.76	8.14
School location	***n***	**Percent**		
Rural: fringe, distant, remote	73	64.0		
City: small	8	7.0		
Suburb: midsize, small	12	10.0		
Town: fringe, distant, remote	22	19.0		
Adequate yearly progress indicators	**No**		**Yes**	
	n	**Percent**	***n***	**Percent**
Made adequate yearly progress for reading in 2007	77	67.0	38	33.0
Made adequate yearly progress for math in 2007	26	22.6	89	77.4
Classified as Title I school in 2007	90	78.3	25	21.7

Source: Authors' calculations based on Maine Department of Education’s Maine High School Assessment Summary Reports (http://www.maine.gov/education/mhsa/08schoolreports/index.html).

APPENDIX B
LITERATURE REVIEW

Widely regarded as the beginning of a serious interest in the study of factors associated with effective schools in the United States, Coleman's landmark 1966 *Equality of Educational Opportunity,* more commonly known as the "Coleman report," was one of the first large-scale studies to simultaneously examine how student and school factors were related to student-level achievement (Reynolds et al. 2000). The report found that 5–9 percent of the variance in student achievement was associated with school characteristics. Student household socioeconomic status was among the strongest predictors of individual achievement. Although the report was widely criticized for considering only a limited number of school resources, the statistical association of household background characteristics on achievement, both nationally and internationally, continues to hold (Martin et al. 2001).

In a reanalysis of Coleman's 1966 data, Jencks et al. (1972) took Coleman's study a step further by assessing the impact of home and school in terms of effect sizes, rather than the percentage of variance accounted for in student achievement (Teddlie, Reynolds, and Sammons 2000). Jencks and his colleagues concluded that effect sizes for school, controlled for student intake and prior achievement, would be "no larger than 0.17 for White students and 0.20 for Black students" (Teddlie, Reynolds, and Sammons 2000, p. 97). Walberg's (1980) "productivity model" argued that achievement in school can be described as a function of seven factors: student ability, motivational factors, quality of instruction, quantity of instruction, classroom variables, home environment, and age or mental development (Marzano 2000). Although Walberg reported average effect sizes for a variety of variables in each category, he mixed different types of effect sizes (that is, correlations and standardized mean differences) without specifying which metric was being used, making it difficult to ascertain the relative impact of each factor (Marzano 2000).

Various researchers have also indicated that measures of student prior achievement are at least as important as household background in determining background variables (Teddlie, Reynolds, and Sammons 2000). Teddlie, Reynolds, and Sammons also note that among researchers, it is generally accepted that student (household) background and prior achievement are two types of variables that should be included in school effectiveness research models (Fitz-Gibbon 1996; Sammons, Mortimore, and Thomas 1996). Smith and Tomlinson's (1989) study of multiracial comprehensiveness produced some evidence of differential effectiveness for students with different levels of prior attainment. The largest differential effect found in their study of 18 comprehensive schools in the United Kingdom was for students with different patterns of prior achievement (Teddlie, Reynolds, and Sammons 2000).

Since the Coleman report, the factors examined in relation to student and school outcomes have included characteristics such as school type (public versus private), school size, and student body demographics (Parcel and Dufur 2001; Carbonaro 2005; Rumberger and Palardy 2005). A review of school effectiveness literature by Reynolds et al. (2000) describing additional school factors associated with achievement includes the work of Murnane (1975), which shows an association of school assignment (and classroom assignment) to student achievement over and above that explained by student background and prior achievement. In a more recent review of characteristics commonly associated with high-performing schools, Shannon and Bylsma (2007) identified nine factors: effective leadership; a clear and shared focus; high expectations for all students; curriculum, instruction, and assessments aligned with state standards; and high levels of family and community involvement. The authors concluded that no one factor on its own was associated with schools performing higher than their demographic characteristics would predict (Shannon and Bylsma 2007).

Student-level factors of particular interest to Maine—race/ethnicity, economic disadvantage status, special education status, and English language learner status—are well represented in the literature. Researchers in previous studies have identified parental social class as a considerable influence on a child's educational outcome; structural arguments fault or credit differences in achievement and attainment on parental socioeconomic status (Sewell and Shah 1968; Sewell, Haller, and Portes 1969; Murnane, Maynard, and Ohls 1980; Baker and Stevenson 1986; Astone and McLanahan 1991; Bankston and Caldas 1998). More recent studies echo these previous findings, suggesting an association between low academic achievement and school poverty (Lee 2000; Ma and Wilkins 2002; Myers, Kim, and Mandala 2004). Sirin (2005) found that greater proportions of students eligible for free or reduced-price lunch in a school were associated with lower test scores for the school. Kalambouka et al.'s (2007) review of studies examining students with disabilities suggests that placing students with special needs in mainstream schools is unlikely to have a negative impact on achievement outcomes for students without special needs. In addition, various researchers have found that English language learner students have lower achievement in reading and math than non–English language learner students do (Eamon 2005; Reardon and Galindo 2007; Terwilliger and Magnuson 2005). Reardon and Galindo (2007) further contend that gaps in student achievement can depend on the proportion of first, second, and third-plus generation English language learner students in a school.

School factors related to Maine's interest in the current study, such as school size, student–teacher ratio, and teacher experience, are also well represented in the literature. In previous studies of class size, results suggested that smaller class sizes tend to produce higher achievement (Glass and Smith 1979; Hedges and Stock 1983). More recent studies have supported this finding. Several researchers have also found that smaller school size leads to positive academic outcomes for students (Johnson, Crosnoe, and Elder 2001; McMillen 2004; Rumberger and Palardy 2005). Further, Hanushek and Krueger's (2000) two meta-analyses comparing Asian and Black students with White students also included class-size studies. They found that small class size can make a difference in the achievement of students, especially racial/ethnic minority students and those with low socioeconomic status (as stated in Rice 2002).

While the school effectiveness literature is replete with factors, studies, and even meta-analyses of studies linking factors to student and school outcomes, criticisms of the quality and comprehensiveness of school effectiveness research remain a part of the evolving dialogue. While Coe and Fitz-Gibbon (1998) acknowledge that school effectiveness research has developed significantly in recent years, they also acknowledge the importance of maintaining consistency within its methodology and inferences.

APPENDIX C METHODOLOGY

Hierarchical linear regression models were used to examine how student characteristics, student prior achievement on the Maine Educational Assessment (MEA) and the Preliminary SAT (PSAT), and school characteristics are associated with scores on the Maine High School Assessment (MHSA). Separate analyses were conducted for each MHSA domain. Multilevel modeling procedures were used to model the dependence among individuals who are clustered in groups (such as schools), thereby producing unbiased estimates of the standard errors associated with the regression coefficients. These types of models allow individual and group characteristics to be included in models of individual outcomes and, where the data permit, allow the relationship between individual-level measures and outcome variables to vary across groups (Raudenbush and Bryk 2002). For all the multilevel models formulated, the regression coefficients and their standard errors, as well as the random components, reflect the variability introduced by the imputation procedures. Specifically, the estimates were derived by combining the results across the 10 imputed datasets using HLM software (version 6.08), which allows up to 10 multiple-imputed datasets to be modeled simultaneously.

Subsequent to running an unconditional (or null) model that included only a random school effect, three two-level regression models were formulated:

- Model 1 includes only the MEA and PSAT scores as predictors.
- Model 2 includes the MEA and PSAT scores as well as student and school characteristics. Including PSAT scores allows the model to reflect learning that takes place within the high schools being assessed.
- Model 3 includes the MEA scores and student and school characteristics but excludes PSAT scores.

A two-level model was used because students were nested in schools, and no classroom or grouping information was available. The results of all three modeling efforts are presented in appendix E.

Building the models

Each model used in the analysis was a two-level model in which grade 11 students were nested within schools. The general two-level model assumes a random sample of i students within j schools, such that Y_{ij} is the MHSA domain score for student i in school j (Raudenbush and Bryk 2002). The general level-one or student model was:

$$Y_{ij} = \beta_{0j} + \beta_{1j}X_{1ij} + \ldots + \beta_{kj}X_{kij} + r_{ij}.$$

MHSA domain scores, Y_{ij}, were modeled as a function of an intercept and a linear combination of student characteristics, X_{kij}. These X_{kij}s were grade 8 MEA scores and grade 10 PSAT scores in model 1; grade 8 MEA scores, grade 10 PSAT scores, and student and school characteristics in model 2; and grade 8 MEA scores and student and school characteristics in model 3 (grade 10 PSAT scores were excluded from model 3). The predicted outcome was composed of a unique intercept, β_{0j}, and slope for each predictor variable, β_{kj}, as well as a random student effect, r_{ij}.

Through empirical examination of the variability in the relationships (that is, level-one slopes) between prior achievement (MEA and PSAT scores) and the outcome variables (MHSA domain scores) across schools, the research team found significant variation in several of the relationships:

- MEA reading scores and MHSA reading scores.
- MEA writing scores and MHSA reading scores.
- MEA math scores and MHSA math scores.
- PSAT math scores and MHSA math scores.
- MEA science scores and MHSA science scores.

- PSAT math scores and MHSA science scores.

- PSAT verbal scores and MHSA science scores.

To produce models 2 and 3 for each MHSA domain, the level-one intercept and the slopes associated with this subset of student prior achievement measures were allowed to vary randomly across schools. In these models, the variation in the level-one intercept across schools was modeled at the second level as a function of an intercept, γ_{00}, and a linear combination of school characteristics, W_{pj}. Each school had a unique random effect, u_{0j}. No school characteristics were included in the models to predict the variability in the level-one slopes (the statistical justification for this is described below). The level-one slopes that did not vary significantly across schools were fixed. There was no significant variation in the relationships between prior achievement and MHSA writing scores across schools. Similarly, there was no significant variation between any of the MHSA domain scores and student characteristics across schools. As such, each of the level-one slopes in the model for predicting MHSA writing scores was fixed, and for the remaining MHSA domains, the level-one slopes associated with student characteristics were fixed (that is, the X_{kij} representing student characteristics were constrained to have the same fixed value for each school).

The research team explored school characteristics as predictors of the variation in the random level-one slopes (described above) across schools. Using model-building strategies outlined by Raudenbush and Bryk (2002), Bickel (2007), Snijders and Bosker (1999), and Gelman and Hill (2007), the research team systematically examined school characteristics as predictors of the level-one random slopes associated with the (varying) subset of prior achievement measures. First, after finalizing model 2, where school characteristics were included in the random intercept model and where a (varying) subset of the level-one slopes associated with student prior achievement were allowed to vary randomly across schools, school characteristics were added in an effort to predict a single randomly varying level-one slope. Next, the research team evaluated the slopes-as-outcomes-model (Raudenbush and Bryk 2002) based on the reliability of the slope, the statistical significance of the coefficient (that is, from 0) associated with the school characteristic for predicting the level-one slope, the percentage of variability in the random slope that was explained with the addition of the school characteristics and the significance of the residual variance (that is, the variance that remained in the slope after the addition of the school characteristics), and the fit of the model.

Examining the reliabilities showed that overall, the slopes were estimated with reasonable levels of confidence. None of the randomly varying slopes had reliability estimates lower than 0.15. In examining the statistical significance of the school characteristics for predicting the random level-one slopes, only a handful of significant coefficients were found; most of the school characteristics were not significantly related to the variability in the relationship between student prior achievement and MHSA domain scores across schools. The exceptions to this were:

- The percentage of grade 11 English language learner students in a school was a significant predictor of the variability in the relationship between MHSA reading scores and grade 8 MEA reading scores.

- The percentage of grade 11 racial/ethnic minority students and the percentage of grade 11 economically disadvantaged students in a school were significant predictors of the variability in the relationship between MHSA math scores and grade 8 MEA math scores.

- Grade 11 cohort size and school location were significant predictors of the variability in the relationship between MHSA math scores and grade 10 PSAT math scores.

- The percentage of grade 11 special education students in a school was a significant predictor of the variability in the relationship between MHSA science scores and grade 8 MEA science scores.

Table C1 summarizes both the percent of variability in each level-one slope that was explained with the addition of the subset of school characteristics and the significance of the residual variance. The percent of variance explained in the level-one slopes was low, ranging from 6 percent to 14 percent. In each case, the residual variance was significant, suggesting that other variables not available for inclusion in the models were associated with the variability in the level-one slopes.

For the final determination of which variables to retain for predicting the variation in the level-one slopes associated with student prior achievement, the fit of the models was examined by evaluating the change in information criteria compared with the random coefficients model and the slopes-as-outcomes model. To do this, the deviance statistics (–2 times the value of the log-likelihood function estimated at the maximum) was examined, and the significance of the deviance difference between the more complex (the slopes-as-outcomes model) and the simpler model (the random coefficients model) was calculated. When the deviance difference was significant, the model with the smallest

TABLE C1

Summary of the intercepts-and-slopes-as-outcomes models explored

	Random coefficients model		Exploratory intercepts-and-slopes-as-outcomes model			
Model and slope	Unconditional variance in level-one slope	Significance	Predictors for each slope	Residual variance[a]	Significance	Percentage explained
Model for predicting Maine High School Assessment reading scores						
Slope for Maine Educational Assessment reading scores	0.00106	0.009	Percentage of students who are English language learners	0.00100	0.011	6
Slope for Maine Educational Assessment writing scores	0.00099	0.011	No variables added	0.00099	0.011	0
Model for predicting Maine High School Assessment math scores						
Slope for Maine Educational Assessment math scores	0.00395	< .001	Percentage of students who are economically disadvantaged; percentage of racial/ethnic minority students	0.00347	< .001	12
Slope for Preliminary SAT math scores	0.01036	< .001	School location; cohort size	0.00914	< .001	12
Model for predicting Maine High School Assessment science scores						
Slope for Maine Educational Assessment science scores	0.00393	0.001	Percentage of students in special education	0.00337	0.002	14
Slope for Preliminary SAT verbal scores	0.00197	0.033	No variables added	0.00197	0.033	0
Slope for Preliminary SAT math scores	0.00229	0.006	No variables added	0.00229	0.006	0

a. The variance estimates reflect the variability introduced by the imputation procedures because they were derived by combining the results across the 10 imputed datasets.

Note: There was no significant variation in the relationships between prior achievement and MHSA writing scores across schools.

Source: Authors' calculations based on student data from data files provided by the Maine Department of Education and school data from the Maine Department of Education's Maine High School Assessment Summary Reports (http://www.maine.gov/education/mhsa/08schoolreports/index.html).

deviance statistic was chosen. When the deviance difference was not significant, the most parsimonious model (that is, the model with the smallest number of estimated parameters) was chosen.

Based on this criterion, including school characteristics in the models for predicting the randomly varying level-one slopes produced poorly fitting models (that is, larger deviance statistics) when compared with the simpler random coefficient model, where the slopes for some level-one measures of student prior achievement were allowed to vary randomly, but no school characteristics were included to predict the variation. Based on the low percentages of variability in the level-one slopes explained with the addition of school characteristics, the significance of the residual variation in the slopes, and the poorer fit statistics for these intercepts-and-slopes-as-outcomes models, the research team opted to allow the slopes to vary randomly for the subset of student prior achievement measures described previously for the final models, but did not include school characteristics as predictors of that variation.

The research team also explored whether school characteristics were associated with the nonrandomly varying slopes (that is, fixed level-one slopes). Similar to the previous approach, the statistical significance of the coefficient (that is, from 0) associated with the school characteristic for predicting the level-one slope and the fit of the model was examined. In examining the statistical significance of school characteristics for predicting the fixed level-one slopes, only a handful of significant coefficients were found; most of the school characteristics were not significantly related to the relationship between prior achievement and MHSA domain scores across schools. The exceptions to this all involved cohort size as a significant predictor of the relationship between:

- MEA math scores and MHSA writing scores ($b = 0.002(.001)$, $t = 2.901$, $p = .004$).
- PSAT verbal scores and MHSA writing scores ($b = 0.002(.001)$, $t = 2.626$, $p = .009$).
- PSAT math scores and MHSA writing scores ($b = 0.002(.001)$, $t = 3.193$, $p = .002$).
- PSAT verbal scores and MHSA math scores ($b = 0.002(.001)$, $t = 2.434$, $p = .015$).
- PSAT writing scores and MHSA math scores ($b = 0.001(.001)$, $t = 2.200$, $p = .028$).

Comparing the fit for these models with the models where no school characteristics were included to predict the slopes showed no significant decrease in the deviance statistics (–2 times the value of the log-likelihood function estimated at the maximum). Therefore, the most parsimonious model (that is, the model with the smallest number of estimated parameters) was chosen, and school characteristics were not chosen to predict the fixed level-one slopes.

The consequence of these decisions was that the models did not support the inclusion of cross-level interactions, so this study cannot discuss whether the relationship between grade 8 MEA scores or grade 10 PSAT scores and grade 11 MHSA scores were moderated by school characteristics. The final general level-two (or school) models were:

$\beta_{0j} = \gamma_{00} + \gamma_{01}W_{1j} + \ldots + \gamma_{0P}W_{Pj} + u_{0j}$

$\beta_{kj} = \gamma_{k0}$ for $k = 1, 2, \ldots, k$ when there is no significant variation in the slopes across schools

and

$\beta_{kj} = \gamma_{k0} + u_{k0}$ for $k = 1, 2, \ldots, k$ when there is significant variation in the slopes across schools.

Model 2, the final intercept-only model that allowed the research team to examine the relationship between MHSA scores (Y_{ij}) and MEA and PSAT scores and student and school characteristics was as follows:

Level one

$Y_{ij} = \beta_{0j} + \beta_{1j}$ (student economic disadvantage status)$_{ij}$ + β_{2j} (student English language learner status)$_{ij}$ + β_{3j} (student special education status)$_{ij}$ + β_{4j} (student gender)$_{ij}$ + β_{5j} (student racial/ethnic minority status)$_{ij}$ + β_{6j} (MEA math score)$_{ij}$ + β_{7j} (MEA reading score)$_{ij}$ + β_{8j} (MEA science score)$_{ij}$ + β_{9j} (MEA writing score)$_{ij}$ + β_{10j} (PSAT verbal score)$_{ij}$ + β_{11j} (PSAT math score)$_{ij}$ + β_{12j} (PSAT writing score)$_{ij}$ + r_{ij}.

Level two

β_{0j} = γ00 + γ01 (city dummy variable)$_j$ + γ_{02} (suburb dummy variable)$_j$ + γ_{03} (town dummy variable)$_j$ + γ_{04} (mean years of teaching experience)$_j$ + γ_{05} (cohort size)$_j$ + γ_{06} (student–teacher ratio)$_j$ + γ_{07} (percent economic disadvantage)$_j$ + γ_{08} (percent English language learners)$_j$ + γ_{09} (percent special education)$_j$ + γ_{10} (percent racial/ethnic minority) $_j$ + γ_{11} (adequate yearly progress in reading)$_j$ + γ_{12} (adequate yearly progress in math)$_j$ + γ_{13} (school title I status)$_j$ + γ_{14} (percent dropout)$_j$ + u_{0j}.

$\beta_{(1\to12)j} = \gamma_{(1\to12)0}$ for each level-one slopes k = 1–12 when there is no significant variation in the slopes across schools

and

$\beta_{(1\to12)j} = \gamma_{(1\to12)0} + u_{(1\to12)0}$ for each level-one slopes k = 1–12 when there is significant variation in the slopes across schools.

Applying model 2

For the purposes of the multilevel regression modeling, the MHSA score variables (Y_{ij}) were each standardized to have a mean of 0 and a standard deviation of 1. The dichotomous level-one student characteristic variables (gender, racial/ethnic minority status, economic disadvantage status, special education status, and English language learner status) were included in the models uncentered. The MEA and PSAT scores were standardized to have a mean of 0 and a standard deviation of 1 across all schools and entered into the model uncentered. Through standardizing the MHSA scores and the MEA and PSAT scores, the regression coefficients are standardized (similar to the β in an ordinary least squares model) and can be compared within a model. The same students were included in the models for predicting MHSA scores—and because they had the same background characteristics (gender, race/ethnicity, and the like) in each model, the regression coefficients for the student characteristics are comparable across MHSA domains. Table C2 at the end of this appendix summarizes how the student characteristics variables were coded prior to inclusion in the multilevel regression models.

At level two, the continuous school measures (percentage of racial/ethnic minority students, percentage of economically disadvantaged students, percentage of special education students, percentage of English language learner students, student–teacher ratio, mean years of teaching experience, cohort size, and percentage of students in cohort who drop out) were entered into the model grand mean–centered. The school percentage variables were also rescaled by a factor of 10 points. The dichotomous school covariates (city, suburb, and town dummy variables, adequate yearly progress in reading and math, and school Title I status) were entered into the level-two models uncentered.

Interpreting the results of model 2

The standardized intercept and regression coefficients for model 2 can be interpreted as follows.

Intercept. A student with 0 for all the student characteristics variables would be male, not a racial/ethnic minority, not economically disadvantaged, not in special education, and not an English language learner and would score at the grand mean in all MEA and PSAT domains. A school with 0 for all the school characteristics variables would be a rural school that has made adequate yearly progress in reading and math, is not classified as Title 1, and is at the average of all schools in the sample on percentage of racial/ethnic minority students, percentage of economically disadvantaged students, percentage of special education students, percentage of English language learner students, student–teacher ratio, mean years of teaching experience, cohort size, and percentage of students in cohort who drop out. The intercept indicates how many standard deviations (or proportion of a standard deviation) above or below the average MHSA domain score a student with 0 on all the student characteristics variables who attends a school with 0 for all the school characteristics variables is predicted to score. For example, for the MHSA writing assessment, the mean score was 1,139.86, and the standard deviation was 13.83 (see table A2 in appendix A). The intercept for model 2 for MHSA writing indicates that the student described here would, on average, score –0.036 standard deviations below the mean (1,139.86). Because the standard deviation was 13.83, this corresponds to approximately 0.50 (0.036*13.83) scale score points below 1,139.86.

Regression coefficients for discrete predictor variables. For each of the discrete student and school characteristics variables, the coefficient can be interpreted as the change in standard deviation units on the scale score of the outcome variable (MHSA domain scores) by going from 0 to 1 in each of the categories (see table C2 for descriptions of 0 and 1 for each variable). Following the previous example, if all variables for the student were the same except that the student is in special education, this would represent a change in the special education variable from 0 to 1, with a corresponding coefficient of –0.117. Multiplying this by the standard deviation of the MHSA writing scores (13.83) yields an expected change of 1.62 scale score points below a general education student with the same background characteristics and in the same school.

Regression coefficients for continuous predictor variables. The continuous student prior achievement measures (MEA and PSAT scores) were standardized to have a mean of 0 and standard deviation of 1, and the remaining continuous student and school characteristics variables (such as percentage of students who are English language learners) were centered on the grand mean across all students and schools in the analytic sample. As such, the coefficients can be interpreted as the predicted standard deviation difference in MHSA domain scores for every 1 standard deviation increase (or one unit increase if the variable is grand mean–centered) in the predictor variable above the mean of 0 (or, if the variable is grand mean–centered, the average of that variable across all students and schools in the analytic sample). See table C2 to determine what one unit of difference represents for each variable.

The regression coefficient for any single variable assumes that all other variables are being held constant. Therefore, although there may be an association between a particular variable and student performance found in the literature (such as school locale), these variables may not be significantly related after accounting for other variables included in the model (such as percentage of economically disadvantaged students, if it was highly correlated with the locale of a school). So that the regression coefficients can be interpreted in relation to the scale scores, the regression coefficient estimates were reported two ways: in standard deviation units and in scale score points.

The left panel in table 1 in the body of the report as well as table G1 in appendix G and tables E1–E4 in appendix E present the standardized regression coefficients. The regression coefficients for the MEA and PSAT scores and continuous school characteristics variables represent standard deviation changes in MHSA domain scores for every 1 standard deviation difference in the predictor variables, holding all other variables constant. Regression coefficients for the dichotomous student and school background variables represent standard deviation differences in MHSA domain scores for every unit change in the predictor value (defined in table C2).

To aid in the interpretation of the predicted MHSA domain scores, the regression coefficients were converted back to their point values on the original scale. These are presented in addition to the standardized estimates in the right panel of table 1 in the body of the report and for model 3 in table G1 in appendix G. In converting back to the original scale scores, the regression coefficients were rescaled to represent predicted differences in MHSA scores for 10 scale score point differences in the continuous predictor variables. Readers are cautioned not to compare the predicted differences in MHSA domain scores associated with 10 scale score point shifts in MEA or PSAT domain scores within a single model. This is because 10 scale score point shifts are not equivalent across the MEA or PSAT domains because the scores from each domain have different standard deviations. However, it is possible to compare the predicted differences in MHSA domain scores associated with a 10 scale score point shift within a single MEA or PSAT domain or for a particular student or school characteristic across MHSA domains. For example, it is possible to compare the predicted differences in MHSA writing, reading, math, or science scores associated with a 10 scale score point shift in MEA reading scores, keeping in mind that identical point changes in reading, writing, math, or science do not necessarily have equivalent meaning. Predicted differences in MHSA domain scale scores were calculated for unit changes in the student and school characteristics as defined in tables B2 and B3 in appendix B and table C2.

Three sets of statistical tests were conducted and reported. First, when the probability (p-value or significance) of observing a regression coefficient this size or larger by chance (if there was no relationship) was less than .05, the coefficient was reported as statistically significant (that is, statistically significantly different from 0). With this statistical test, for example, MEA reading scores were found to be a statistically significant predictor of MHSA reading scores (0.138). In this case, the probability of observing this regression coefficient by chance if there was no relationship was less than .05.

Second, when .95 confidence intervals constructed around the difference between two standardized regression coefficients within the same regression model did not include 0, the larger coefficient was reported as significantly stronger than the other coefficient. This type of statistical test allows the conclusion that one variable was a stronger predictor than another for predicting an MHSA domain score. In this report stronger predictors are defined as those with regression coefficients that are larger than those of other predictors in the report's regression models.[7] With this statistical test, for example, MEA reading scores were found to be a stronger predictor than MEA writing scores for predicting MHSA reading scores. In this case, the confidence interval around the difference between the two standardized regression coefficients did not include 0. Table D1 in appendix D presents the .95 confidence intervals around the regression coefficients for MHSA content domains (within models).

Third, when .95 confidence intervals constructed around the difference between two standardized regression coefficients for the same predictor variable across MHSA domain models did not include 0, the larger coefficient was reported as significantly stronger than the other coefficient. This type of statistical test allows the conclusion that a variable is a stronger predictor in one MHSA domain than in another MHSA domain. Again, stronger predictors are defined as those with regression coefficients that are larger than those of other predictors in the report's regression models. For example, MEA reading scores were a stronger predictor of MHSA reading scores than they were of MHSA math scores. Tables D2 and D3 in appendix D present the .95 confidence intervals around the regression coefficients across MHSA content domains (within models).

Calculating percentage of variance explained by the models

In addition to the regression coefficients and their associated significance levels, the regression models allowed the research team to estimate the total percentage of variance in the MHSA domain scores that was explained by the student and school characteristics and by the MEA and PSAT domain scores. This percentage was calculated for models 1, 2, and 3 by comparing the residual variance to the available variance in the unconditional model. Specifically, the percentage of variance explained (analogous to R^2) was estimated for each model using the following equation:

$$1 - \frac{\text{Total residual variance under the conditional model}}{\text{Total unconditional variance}}$$

Table 2 in the body of the report and tables E1–E4 in appendix E present the variance explained by the models.

The unconditional percent of total variation in the MHSA domain scores between schools (the intraclass correlation coefficient) was nonzero (and statistically significant) in all domains. When no predictors were included in the model, the unconditional intraclass correlation coefficients for the MHSA scores were 0.079 for reading, 0.093 for writing, 0.075 for math, and 0.068 for science. Dependence among scores within groups (that is, a nonzero intraclass correlation coefficient) has an impact on the effective sample; larger intraclass correlation coefficients result in larger design effects and smaller effective sample sizes (holding total sample size and the average number of individuals within groups constant).[8]

The size of these unconditional intraclass correlation coefficients is consistent with those found in prior research indicating that the proportion of variability in achievement between schools tends to be lower in rural schools than in the general population of schools (Hedges and Hedberg 2007). For example, for math achievement, Hedges and Hedberg found that across all grades, the average intraclass correlation in rural schools was 0.149, compared with 0.220 for all schools. They observed a similar pattern for reading outcomes: 64 percent of the schools in the Maine sample were in rural locales (see table A3 in appendix A), compared with about 31 percent of schools nationwide.

TABLE C2

Description of variables and coding schemes for the multilevel regression models

Variable	Values	Notes
Outcome variables		
MHSA writing, reading math, and science scores	Scale scores range from 1100–1180, standardized (mean = 0, standard deviation = 1).	Scores were converted to standard deviation units. See table A2 in appendix A for summary statistics for the original variable.
Student characteristics		
Gender	0 = male, 1 = female	
Racial/ethnic minority status	0 = White, 1 = racial/ethnic minority	A student's No Child Left Behind subgroup (defined as a student's racial/ethnic classification first assigned during enrollment in a Maine public school) was used to identify racial/ethnic minority students. Students who were White were classified as non–racial/ethnic minority students, and students who were Alaskan/Native American, Asian/Pacific Islander, Black, or Hispanic were classified as racial/ethnic minority students. See table A2 in appendix A for summary statistics.
Economically disadvantaged status	0 = no, 1 = yes	An economically disadvantaged student is a student who is eligible for free or reduced-price meals. See table A2 in appendix A for summary statistics.
Special education status	0 = no, 1 = yes	A student in special education is a student who has been identified under the Individuals with Disabilities Education Act and educated in accordance with an Individualized Education Plan. See table A2 in appendix A for summary statistics.
English language learner status	0 = no, 1 = yes	An English language learner student is a student who is identified in accordance with No Child Left Behind as a student with limited English proficiency (for adequate yearly progress determinations, former English language learner students are included in the subgroup for up to two years after exiting English language learner status). See table A2 for summary statistics.

(CONTINUED)

TABLE C2 (CONTINUED)

Description of variables and coding schemes for the multilevel regression models

Variable	Values	Notes
Student prior achievement		
MEA reading, writing, math, and science scores	Scale scores designed to range from 500 to 580, standardized (mean = 0, standard deviation = 1).	Scores were standardized to have a mean of 0 and a standard deviation of 1. See table A2 in appendix A for summary statistics for the original variable.
PSAT verbal, math, and writing scores	Scale scores designed to range from 20 to 80, standardized (mean = 0, standard deviation = 1).	Scores were standardized to have a mean of 0 and a standard deviation of 1. See table A2 in appendix A for summary statistics for the original variable.
School characteristics		
Percentage of racial/ethnic minority students	0–100 percent, rescaled by a factor of 10, grand mean–centered	The percentage of grade 11 racial/ethnic minority students in the school during the testing period. The variable was rescaled to units of 10 percentage points and centered on the grand mean. See table A3 in appendix A for summary statistics for the original variable.
Percent of students who are economically disadvantaged	0–100 percent, rescaled by a factor of 10, grand mean–centered	Defined as the percentage of grade 11 economically disadvantaged students in the school during the testing period. The variable was rescaled to units of 10 percentage points and centered on the grand mean. See table A3 in appendix A for summary statistics for the original variable.
Percentage of students in special education	0–100 percent, rescaled, grand mean–centered	Defined as the percentage of grade 11 students who are categorized as having a disability and who receive special education services in the school during testing period. The variable was rescaled to units of 10 percentage points and centered on the grand mean. See table A3 in appendix A for summary statistics for the original variable.
Percentage of students who are English language learners	0–100 percent, rescaled by a factor of 10, grand mean–centered	Defined as the percentage of grade 11 English language learner students in the school during testing period. The variable was rescaled to units of 10 percentage points and centered on the grand mean. See table A3 in appendix A for summary statistics for the original variable.
Percentage of students in cohort who drop out	0–100 percent, rescaled by a factor of 10, grand mean–centered	• *Cohort size:* the number of grade 11 students in a school who were registered during the testing period. This does not include dropouts or other students who might have transferred earlier in the year. • *School location:* location of the school based on the urban-centric locale coding system: rural (fringe, distant, and remote), city (small), suburban (midsize and small), or town (fringe, distant, and remote) locales. No Maine schools are in large or midsize city locales. • *Adequate yearly progress status indicators:* whether the school met adequate yearly progress requirements for reading or math in 2007 (former English language learner students are included in the subgroup up to two years after exiting English language learner status) and whether a school was classified as a Title I school (schoolwide or targeted assistance) in 2007. Defined as the percentage of students from the analysis cohort who left school in grades 9–12 prior to graduation for reasons other than transferring to another school and who did not re-enroll before the following October 1. See table A3 in appendix A for summary statistics for the original variable.

(CONTINUED)

TABLE C2 (CONTINUED)

Description of variables and coding schemes for the multilevel regression models

Variable	Values	Notes
Student–teacher ratio	Grand mean–centered only	Defined as the total number of students in the school divided by the total number of teachers (including classroom and special education teachers and literacy specialists) in the school. This variable was centered on the grand mean. See table A3 in appendix A for summary statistics.
Mean years of teaching experience (not rescaled)	Grand mean–centered only	Defined as the average number of years that all teachers in the school (including classroom and special education teachers and literacy specialists) have been teaching in public and private schools. This variable was centered on the grand mean. See table A3 in appendix A for summary statistics for the original variable.
Cohort size	Rescaled by a factor of 10, grand mean–centered	Defined as the number of grade 11 students in a school who were registered during the testing period. This does not include dropouts or other students who might have transferred earlier in the year. This variable was centered on the grand mean. See table A3 in appendix A for summary statistics.
Rural versus city	Rural = 0, city = 1	Based on the urban-centric locale coding system. Rural (fringe, distant, and remote) compared with city (small) locales. See table A3 in appendix A for summary statistics.
Rural versus suburb	Rural = 0, suburb = 1	Based on the urban-centric locale coding system. Rural (fringe, distant, and remote) compared with suburban (midsize and small) locales. See table A3 in appendix A for summary statistics.
Rural versus town	Rural = 0, town = 1	Based on the urban-centric locale coding system. Rural (fringe, distant, and remote) compared with town (fringe, distant, and remote) locales. See table A3 in appendix A for summary statistics.
Made adequate yearly progress in reading	0 = yes, 1 = no	Accountability status of a school based on the federally mandated measures of performance outlined by the No Child Left Behind Act. A school makes adequate yearly progress if the students and subpopulations of students in the tested grade or grades and all required subgroups meet the participation targets of 95 percent, meet or exceed the performance targets established for reading and math, and meet attendance goals (K–8) or graduation-rate targets (high schools) (Maine Department of Education 2008). See table A3 in appendix A for summary statistics.
Made adequate yearly progress in math		
School Title I status	0 = no, 1 = yes	Those schools operating under Title I either schoolwide or as targeted assistance programs. See table A3 in appendix A for summary statistics.

Note: MHSA is Maine High School Assessment. MEA is Maine Educational Assessment. PSAT is Preliminary SAT.

Source: Authors.

APPENDIX D CONFIDENCE INTERVALS FOR TESTING DIFFERENCES

TABLE D1

.95 confidence interval around regression coefficients within Maine High School Assessment content domains (based on model 2 results)

Prior achievement measure	β[a]	Standard error	Prior achievement measure	β[a]	Standard error[a]	Difference between βs	Standard error of difference between βs	.95 confidence interval[b] Lower boundary	.95 confidence interval[b] Upper boundary
Model for predicting MHSA reading scores									
MEA math scores	0.058	0.011	MEA reading	0.138	0.010	–0.080	0.016	–0.111	–0.049**
			MEA writing	–0.022	0.007	0.080	0.013	0.054	0.106**
			MEA science	0.147	0.010	–0.089	0.016	–0.120	–0.058**
			PSAT verbal	0.410	0.009	–0.352	0.014	–0.379	–0.325**
			PSAT math	0.016	0.011	0.042	0.017	0.008	0.076**
			PSAT writing	0.222	0.010	–0.164	0.015	–0.193	–0.135**
MEA reading scores	0.138	0.010	MEA writing	–0.022	0.007	0.160	0.013	0.134	0.186**
			MEA science	0.147	0.010	–0.009	0.015	–0.039	0.021
			PSAT verbal	0.410	0.009	–0.272	0.014	–0.300	–0.244**
			PSAT math	0.016	0.011	0.122	0.015	0.093	0.151**
			PSAT writing	0.222	0.010	–0.084	0.014	–0.112	–0.056**
MEA science scores	0.147	0.010	MEA writing	–0.022	0.007	0.169	0.012	0.145	0.193**
			PSAT verbal	0.410	0.009	–0.263	0.014	–0.291	–0.235**
			PSAT math	0.016	0.011	0.131	0.015	0.102	0.160**
			PSAT writing	0.222	0.010	–0.075	0.014	–0.103	–0.047**
MEA writing scores	–0.022	0.007	PSAT verbal	0.410	0.009	–0.432	0.011	–0.454	–0.410**
			PSAT math	0.016	0.011	–0.038	0.013	–0.064	–0.012**
			PSAT writing	0.222	0.010	–0.244	0.012	–0.268	–0.220**
PSAT verbal scores	0.410	0.009	PSAT math	0.016	0.011	0.394	0.015	0.365	0.423**
			PSAT writing	0.222	0.010	0.188	0.015	0.158	0.218**
PSAT math scores	0.016	0.011	PSAT writing	0.222	0.010	–0.206	0.016	–0.236	–0.176**
Model for predicting MHSA writing scores									
MEA math scores	0.105	0.010	MEA reading	0.120	0.010	–0.015	0.015	–0.044	0.014
			MEA writing	0.063	0.006	0.042	0.012	0.019	0.065**
			MEA science	0.053	0.010	0.052	0.015	0.022	0.082**
			PSAT verbal	0.212	0.010	–0.107	0.014	–0.134	–0.080**
			PSAT math	0.051	0.010	0.054	0.016	0.022	0.086**
			PSAT writing	0.370	0.009	–0.265	0.014	–0.292	–0.238**
MEA reading scores	0.120	0.010	MEA writing	0.063	0.006	0.057	0.013	0.032	0.082**
			MEA science	0.053	0.010	0.067	0.015	0.037	0.097**
			PSAT verbal	0.212	0.010	–0.092	0.015	–0.121	–0.063**
			PSAT math	0.051	0.010	0.069	0.014	0.042	0.096**
			PSAT writing	0.370	0.009	–0.250	0.014	–0.277	–0.223**

(CONTINUED)

TABLE D1 (CONTINUED)

.95 confidence interval around regression coefficients within Maine High School Assessment content domains (based on model 2 results)

Prior achievement measure	β[a]	Standard error	Prior achievement measure	β[a]	Standard error[a]	Difference between βs	Standard error of difference between βs	.95 confidence interval[b] Lower boundary	Upper boundary
MEA science scores	0.053	0.010	MEA writing	0.063	0.006	–0.010	0.012	–0.033	0.013
			PSAT verbal	0.212	0.010	–0.159	0.015	–0.188	–0.130**
			PSAT math	0.051	0.010	0.002	0.014	–0.026	0.030
			PSAT writing	0.370	0.009	–0.317	0.014	–0.344	–0.290**
MEA writing scores	0.063	0.006	PSAT verbal	0.212	0.010	–0.149	0.012	–0.172	–0.126**
			PSAT math	0.051	0.010	0.012	0.012	–0.011	0.035
			PSAT writing	0.370	0.009	–0.307	0.011	–0.329	–0.285**
PSAT verbal scores	0.212	0.010	PSAT math	0.051	0.010	0.161	0.015	0.132	0.190**
			PSAT writing	0.370	0.009	–0.158	0.015	–0.188	–0.128**
PSAT math scores	0.051	0.010	PSAT writing	0.370	0.009	–0.319	0.014	–0.347	–0.291**
Model for predicting MHSA math scores									
MEA math scores	0.313	0.012	MEA reading	–0.024	0.011	0.337	0.017	0.304	0.370**
			MEA writing	0.003	0.008	0.310	0.015	0.281	0.339**
			MEA science	0.051	0.009	0.262	0.016	0.230	0.294**
			PSAT verbal	0.056	0.010	0.257	0.015	0.227	0.287**
			PSAT math	0.429	0.014	–0.116	0.022	–0.159	–0.073**
			PSAT writing	0.067	0.009	0.246	0.015	0.216	0.276**
MEA reading scores	–0.024	0.011	MEA writing	0.003	0.008	–0.027	0.014	–0.055	0.001
			MEA science	0.051	0.009	–0.075	0.015	–0.105	–0.045**
			PSAT verbal	0.056	0.010	–0.080	0.016	–0.111	–0.049**
			PSAT math	0.429	0.014	–0.453	0.018	–0.488	–0.418**
			PSAT writing	0.067	0.009	–0.091	0.015	–0.120	–0.062**
MEA science scores	0.051	0.009	MEA writing	0.003	0.008	0.048	0.012	0.024	0.072**
			PSAT verbal	0.056	0.010	–0.005	0.014	–0.033	0.023
			PSAT math	0.429	0.014	–0.378	0.017	–0.411	–0.345**
			PSAT writing	0.067	0.009	–0.016	0.013	–0.041	0.009
MEA writing scores	0.003	0.008	PSAT verbal	0.056	0.010	–0.053	0.013	–0.078	–0.028**
			PSAT math	0.429	0.014	–0.426	0.016	–0.458	–0.394**
			PSAT writing	0.067	0.009	–0.064	0.012	–0.088	–0.040**
PSAT verbal scores	0.056	0.010	PSAT math	0.429	0.014	–0.373	0.018	–0.408	–0.338**
			PSAT writing	0.067	0.009	–0.011	0.016	–0.042	0.020
PSAT math scores	0.429	0.014	PSAT writing	0.067	0.009	0.362	0.017	0.328	0.396**

(CONTINUED)

TABLE D1 (CONTINUED)

.95 confidence interval around regression coefficients within Maine High School Assessment content domains (based on model 2 results)

Prior achievement measure	β[a]	Standard error	Prior achievement measure	β[a]	Standard error[a]	Difference between βs	Standard error of difference between βs	.95 confidence interval[b] Lower boundary	.95 confidence interval[b] Upper boundary
Model for predicting MHSA science scores									
MEA math scores	0.127	0.014	MEA reading	0.022	0.012	0.105	0.019	0.067	0.143**
			MEA writing	–0.050	0.008	0.177	0.016	0.145	0.209**
			MEA science	0.314	0.012	–0.187	0.020	–0.226	–0.148**
			PSAT verbal	0.229	0.012	–0.102	0.018	–0.137	–0.067**
			PSAT math	0.171	0.013	–0.044	0.021	–0.086	–0.002**
			PSAT writing	0.071	0.011	0.056	0.018	0.021	0.091**
MEA reading scores	0.022	0.012	MEA writing	–0.050	0.008	0.072	0.015	0.042	0.102**
			MEA science	0.314	0.012	–0.292	0.018	–0.328	–0.256**
			PSAT verbal	0.229	0.012	–0.207	0.018	–0.242	–0.172**
			PSAT math	0.171	0.013	–0.149	0.017	–0.183	–0.115**
			PSAT writing	0.071	0.011	–0.049	0.017	–0.082	–0.016**
MEA science scores	0.314	0.012	MEA writing	–0.050	0.008	0.364	0.014	0.336	0.392**
			PSAT verbal	0.229	0.012	0.085	0.018	0.049	0.121**
			PSAT math	0.171	0.013	0.143	0.018	0.107	0.179**
			PSAT writing	0.071	0.011	0.243	0.016	0.211	0.275**
MEA writing scores	–0.05	0.008	PSAT verbal	0.229	0.012	–0.279	0.014	–0.307	–0.251**
			PSAT math	0.171	0.013	–0.221	0.015	–0.251	–0.191**
			PSAT writing	0.071	0.011	–0.121	0.014	–0.148	–0.094**
PSAT verbal scores	0.229	0.012	PSAT math	0.171	0.013	0.058	0.018	0.023	0.093**
			PSAT writing	0.071	0.011	0.158	0.019	0.121	0.195**
PSAT math scores	0.171	0.013	PSAT writing	0.071	0.011	0.100	0.018	0.065	0.135**

MHSA is Maine High School Assessment. MEA is Maine Educational Assessment. PSAT is Preliminary SAT.

** .95 confidence interval does not contain 0.

a. The regression coefficients and their standard errors reflect the variability introduced by the imputation procedures because they were derived by combining the results across the 10 imputed datasets.

b. The significance of the difference between the standardized regression coefficients was calculated by constructing a .95 confidence interval around the difference between the standardized regression coefficients. The interval was calculated as $(\beta_1 - \beta_2) \pm 1.96(SE_{\beta_1-\beta_2})$, where $SE_{\beta_1-\beta_2} = \sqrt{(SE_1)^2 + (SE_2)^2 - 2Cov_{\beta_1\beta_2}}$. Tests of the confidence intervals around the difference between model coefficients are not adjusted for multiple comparisons.

Source: Authors' calculations based on student data from data files provided by the Maine Department of Education and school data from the Maine Department of Education's Maine High School Assessment Summary Reports (http://www.maine.gov/education/mhsa/08schoolreports/index.html).

TABLE D2

.95 confidence interval around regression coefficients for Maine Educational Assessment and Preliminary SAT scores across Maine High School Assessment domains (based on model 2 results)

Predictor 1	β[a]	Standard error	Predictor 2	β[a]	Standard error[a]	Difference between βs	Standard error of difference between βs	.95 confidence interval[b] Lower boundary	.95 confidence interval[b] Upper boundary
MEA reading									
Predicting MHSA reading	0.138	0.010	Predicting MHSA writing	0.120	0.010	0.018	0.011	−0.004	0.040
			Predicting MHSA math	−0.024	0.011	0.162	0.012	0.139	0.185**
			Predicting MHSA science	0.022	0.012	0.116	0.013	0.090	0.142**
Predicting MHSA writing	0.120	0.010	Predicting MHSA math	−0.024	0.011	0.144	0.012	0.121	0.167**
			Predicting MHSA science	0.022	0.012	0.098	0.014	0.071	0.125**
Predicting MHSA math	−0.024	0.011	Predicting MHSA science	0.022	0.012	−0.046	0.014	−0.073	−0.019**
MEA writing									
Predicting MHSA reading	−0.022	0.007	Predicting MHSA writing	0.063	0.006	−0.085	0.007	−0.098	−0.072**
			Predicting MHSA math	0.003	0.008	−0.025	0.009	−0.042	−0.008**
			Predicting MHSA science	−0.050	0.008	0.028	0.009	0.010	0.046**
Predicting MHSA writing	0.063	0.006	Predicting MHSA math	0.003	0.008	0.060	0.008	0.044	0.076**
			Predicting MHSA science	−0.050	0.008	0.113	0.009	0.096	0.130**
Predicting MHSA math	0.003	0.008	Predicting MHSA science	−0.050	0.008	0.053	0.010	0.034	0.072**
MEA math									
Predicting MHSA reading	0.058	0.011	Predicting MHSA writing	0.105	0.010	−0.047	0.011	−0.069	−0.025**
			Predicting MHSA math	0.313	0.012	−0.255	0.013	−0.281	−0.229**
			Predicting MHSA science	0.127	0.014	−0.069	0.016	−0.100	−0.038**
Predicting MHSA writing	0.105	0.010	Predicting MHSA math	0.313	0.012	−0.208	0.012	−0.231	−0.185**
			Predicting MHSA science	0.127	0.014	−0.022	0.015	−0.051	0.007
Predicting MHSA math	0.313	0.012	Predicting MHSA science	0.127	0.014	0.186	0.015	0.157	0.215**

(CONTINUED)

TABLE D2 (CONTINUED)

.95 confidence interval around regression coefficients for Maine Educational Assessment and Preliminary SAT scores across Maine High School Assessment domains (based on model 2 results)

Predictor 1	β[a]	Standard error	Predictor 2	β[a]	Standard error[a]	Difference between βs	Standard error of difference between βs	.95 confidence interval[b] Lower boundary	.95 confidence interval[b] Upper boundary
MEA science									
Predicting MHSA reading	0.147	0.010	Predicting MHSA writing	0.053	0.010	0.094	0.013	0.069	0.119**
			Predicting MHSA math	0.051	0.009	0.096	0.011	0.074	0.118**
			Predicting MHSA science	0.314	0.012	−0.167	0.013	−0.192	−0.142**
Predicting MHSA writing	0.053	0.010	Predicting MHSA math	0.051	0.009	0.002	0.012	−0.022	0.026
			Predicting MHSA science	0.314	0.012	−0.261	0.015	−0.291	−0.231**
Predicting MHSA math	0.051	0.009	Predicting MHSA science	0.314	0.012	−0.263	0.013	−0.288	−0.238**
PSAT verbal									
Predicting MHSA reading	0.410	0.009	Predicting MHSA writing	0.212	0.010	0.198	0.011	0.176	0.220**
			Predicting MHSA math	0.056	0.010	0.354	0.013	0.329	0.379**
			Predicting MHSA science	0.229	0.012	0.181	0.013	0.155	0.207**
Predicting MHSA writing	0.212	0.010	Predicting MHSA math	0.056	0.010	0.156	0.013	0.131	0.181**
			Predicting MHSA science	0.229	0.012	−0.017	0.015	−0.046	0.012
Predicting MHSA math	0.056	0.010	Predicting MHSA science	0.229	0.012	−0.173	0.014	−0.200	−0.146**
PSAT math									
Predicting MHSA reading	0.016	0.011	Predicting MHSA writing	0.051	0.010	−0.035	0.011	−0.056	−0.014**
			Predicting MHSA math	0.429	0.014	−0.413	0.014	−0.441	−0.385**
			Predicting MHSA science	0.171	0.013	−0.155	0.014	−0.182	−0.128**
Predicting MHSA writing	0.051	0.010	Predicting MHSA math	0.429	0.014	−0.378	0.013	−0.403	−0.353**
			Predicting MHSA science	0.171	0.013	−0.120	0.014	−0.147	−0.093**
Predicting MHSA math	0.429	0.014	Predicting MHSA science	0.171	0.013	0.258	0.015	0.228	0.288**

(CONTINUED)

TABLE D2 (CONTINUED)

.95 confidence interval around regression coefficients for Maine Educational Assessment and Preliminary SAT scores across Maine High School Assessment domains (based on model 2 results)

Predictor 1	β[a]	Standard error	Predictor 2	β[a]	Standard error[a]	Difference between βs	Standard error of difference between βs	.95 confidence interval[b] Lower boundary	.95 confidence interval[b] Upper boundary
PSAT writing									
Predicting MHSA reading	0.222	0.010	Predicting MHSA writing	0.370	0.009	−0.148	0.011	−0.169	−0.127**
			Predicting MHSA math	0.067	0.009	0.155	0.012	0.131	0.179**
			Predicting MHSA science	0.071	0.011	0.151	0.012	0.128	0.174**
Predicting MHSA writing	0.370	0.009	Predicting MHSA math	0.067	0.009	0.303	0.011	0.281	0.325**
			Predicting MHSA science	0.071	0.011	0.299	0.013	0.274	0.324**
Predicting MHSA math	0.067	0.009	Predicting MHSA science	0.071	0.011	−0.004	0.013	−0.029	0.021

MHSA is Maine High School Assessment. MEA is Maine Educational Assessment. PSAT is Preliminary SAT.

** .95 confidence interval does not contain 0.

a. The regression coefficients and their standard errors reflect the variability introduced by the imputation procedures because they were derived by combining the results across the 10 imputed datasets.

b. The significance of the difference between the standardized regression coefficients was calculated by constructing a .95 confidence interval around the difference between the standardized regression coefficients. The interval was calculated as $(\beta_1 - \beta_2) \pm 1.96(SE_{\beta_1-\beta_2})$, where $SE_{\beta_1-\beta_2} = \sqrt{(SE_1)^2 + (SE_2)^2 - 2Cov_{\beta_1\beta_2}}$. Tests of the confidence intervals around the difference between model coefficients are not adjusted for multiple comparisons.

Source: Authors' calculations based on student data from data files provided by the Maine Department of Education and school data from the Maine Department of Education's Maine High School Assessment Summary Reports (http://www.maine.gov/education/mhsa/08schoolreports/index.html).

TABLE D3

.95 confidence interval around regression coefficients for student and school characteristics across Maine High School Assessment content domains (based on Model 2 results)

Predictor 1	β[a]	Standard error	Predictor 2	β[a]	Standard error[a]	Difference between βs	Standard error of difference between βs	.95 confidence interval[b] Lower boundary	Upper boundary
Gender									
Predicting MHSA reading	−0.052	0.010	Predicting MHSA writing	0.060	0.010	−0.112	0.012	−0.136	−0.088**
			Predicting MHSA math	−0.088	0.009	0.036	0.013	0.011	0.061**
			Predicting MHSA science	−0.183	0.011	0.131	0.014	0.104	0.158**
Predicting MHSA writing	0.060	0.010	Predicting MHSA math	−0.088	0.009	0.148	0.012	0.124	0.172**
			Predicting MHSA science	−0.183	0.011	0.243	0.013	0.217	0.269**
Predicting MHSA math	−0.088	0.009	Predicting MHSA science	−0.183	0.011	0.095	0.012	0.071	0.119**
Economic disadvantage									
Predicting MHSA reading	−0.035	0.011	Predicting MHSA writing	−0.061	0.011	0.026	0.013	0.000	0.052**
			Predicting MHSA math	−0.039	0.011	0.004	0.015	−0.026	0.034
			Predicting MHSA science	−0.015	0.012	−0.020	0.015	−0.049	0.009
Predicting MHSA writing	−0.061	0.011	Predicting MHSA math	−0.039	0.011	−0.022	0.014	−0.050	0.006
			Predicting MHSA science	−0.015	0.012	−0.046	0.016	−0.078	−0.014**
Predicting MHSA math	−0.039	0.011	Predicting MHSA Science	−0.015	0.012	−0.024	0.015	−0.053	0.005
Special education status									
Predicting MHSA reading	−0.065	0.019	Predicting MHSA writing	−0.117	0.021	0.052	0.019	0.015	0.089**
			Predicting MHSA math	−0.105	0.019	0.040	0.019	0.002	0.078**
			Predicting MHSA science	−0.032	0.018	−0.033	0.022	−0.077	0.011
Predicting MHSA writing	−0.117	0.021	Predicting MHSA math	−0.105	0.019	−0.012	0.020	−0.051	0.027
			Predicting MHSA science	−0.032	0.018	−0.085	0.021	−0.127	−0.043**
Predicting MHSA math	−0.105	0.019	Predicting MHSA science	−0.032	0.018	−0.073	0.021	−0.113	−0.033**

(CONTINUED)

TABLE D3 (CONTINUED)

.95 confidence interval around regression coefficients for student and school characteristics across Maine High School Assessment content domains (based on Model 2 results)

Predictor 1	β[a]	Standard error	Predictor 2	β[a]	Standard error[a]	Difference between βs	Standard error of difference between βs	.95 confidence interval[b] Lower boundary	.95 confidence interval[b] Upper boundary
English language learner status									
Predicting MHSA reading	−0.061	0.028	Predicting MHSA writing	−0.102	0.032	0.041	0.034	−0.027	0.109
			Predicting MHSA math	−0.024	0.032	−0.037	0.038	−0.112	0.038
			Predicting MHSA science	−0.100	0.030	0.039	0.031	−0.021	0.099
Predicting MHSA writing	−0.102	0.032	Predicting MHSA math	−0.024	0.032	−0.078	0.042	−0.161	0.005
			Predicting MHSA science	−0.100	0.030	−0.002	0.042	−0.085	0.081
Predicting MHSA math	−0.024	0.032	Predicting MHSA science	−0.100	0.030	0.076	0.040	−0.002	0.154
Percentage of students who are economically disadvantaged									
Predicting MHSA reading	−0.013	0.006	Predicting MHSA writing	−0.006	0.006	−0.007	0.006	−0.020	0.006
			Predicting MHSA math	0.003	0.007	−0.016	0.007	−0.029	−0.003
			Predicting MHSA science	0.003	0.010	−0.016	0.009	−0.034	0.002
Predicting MHSA writing	−0.006	0.006	Predicting MHSA math	0.003	0.007	−0.009	0.008	−0.025	0.007
			Predicting MHSA science	0.003	0.010	−0.009	0.011	−0.031	0.013
Predicting MHSA math	0.003	0.007	Predicting MHSA science	0.003	0.010	0.000	0.011	−0.022	0.022
Percentage of students who are in special education									
Predicting MHSA reading	−0.015	0.013	Predicting MHSA writing	0.002	0.018	−0.017	0.018	−0.052	0.018
			Predicting MHSA math	−0.037	0.014	0.022	0.016	−0.009	0.053
			Predicting MHSA science	−0.054	0.021	0.039	0.022	−0.004	0.082
Predicting MHSA writing	0.002	0.018	Predicting MHSA math	−0.037	0.014	0.039	0.020	−0.001	0.079
			Predicting MHSA science	−0.054	0.021	0.056	0.027	0.002	0.110**
Predicting MHSA Math	−0.037	0.014	Predicting MHSA science	−0.054	0.021	0.017	0.024	−0.030	0.064

(CONTINUED)

TABLE D3 (CONTINUED)

.95 confidence interval around regression coefficients for student and school characteristics across Maine High School Assessment content domains (based on Model 2 results)

Predictor 1	β[a]	Standard error	Predictor 2	β[a]	Standard error[a]	Difference between βs	Standard error of difference between βs	.95 confidence interval[b] Lower boundary	Upper boundary
Percentage of students in cohort who drop out									
Predicting MHSA reading	0.037	0.015	Predicting MHSA writing	0.008	0.019	0.029	0.014	0.001	0.057**
			Predicting MHSA math	−0.005	0.016	0.029	0.018	−0.007	0.065
			Predicting MHSA science	−0.008	0.025	0.029	0.022	−0.014	0.072
Predicting MHSA writing	0.008	0.019	Predicting MHSA math	−0.005	0.016	0.013	0.020	−0.025	0.051
			Predicting MHSA science	−0.008	0.025	0.008	0.019	−0.030	0.046
Predicting MHSA math	−0.005	0.016	Predicting MHSA science	−0.008	0.025	0.003	0.016	−0.028	0.034
Student–teacher ratio									
Predicting MHSA reading	−0.019	0.005	Predicting MHSA writing	−0.017	0.007	−0.002	0.007	−0.015	0.011
			Predicting MHSA math	−0.015	0.005	−0.004	0.006	−0.016	0.008
			Predicting MHSA science	−0.001	0.008	−0.018	0.009	−0.035	−0.001
Predicting MHSA writing	−0.017	0.007	Predicting MHSA math	−0.015	0.005	−0.002	0.008	−0.017	0.013
			Predicting MHSA science	−0.001	0.008	−0.016	0.011	−0.038	0.006
Predicting MHSA math	−0.015	0.005	Predicting MHSA science	−0.001	0.008	−0.014	0.009	−0.031	0.003
Cohort size									
Predicting MHSA reading	0.002	0.001	Predicting MHSA writing	0.002	0.002	0.000	0.002	−0.004	0.004
			Predicting MHSA math	0.003	0.001	−0.001	0.001	−0.003	0.001
			Predicting MHSA science	0.001	0.002	0.001	0.002	−0.003	0.005
Predicting MHSA writing	0.002	0.002	Predicting MHSA math	0.003	0.001	−0.001	0.002	−0.005	0.003
			Predicting MHSA science	0.001	0.002	0.001	0.003	−0.005	0.007
Predicting MHSA math	0.003	0.001	Predicting MHSA science	0.001	0.002	0.002	0.002	−0.002	0.006

(CONTINUED)

TABLE D3 (CONTINUED)

.95 confidence interval around regression coefficients for student and school characteristics across Maine High School Assessment content domains (based on Model 2 results)

Predictor 1	β[a]	Standard error	Predictor 2	β[a]	Standard error[a]	Difference between βs	Standard error of difference between βs	.95 confidence interval[b] Lower boundary	Upper boundary
School location: suburban versus rural									
Predicting MHSA reading	0.025	0.023	Predicting MHSA writing	0.111	0.033	−0.086	0.016	−0.116	−0.056**
			Predicting MHSA math	0.029	0.026	−0.086	0.027	−0.139	−0.033**
			Predicting MHSA science	0.058	0.039	−0.086	0.047	−0.177	0.005
Predicting MHSA writing	0.111	0.033	Predicting MHSA math	0.029	0.026	0.082	0.025	0.034	0.130**
			Predicting MHSA science	0.058	0.039	0.111	0.047	0.018	0.204**
Predicting MHSA math	0.029	0.026	Predicting MHSA science	0.058	0.039	−0.029	0.030	−0.089	0.031
Made adequate yearly progress in reading									
Predicting MHSA reading	−0.023	0.015	Predicting MHSA writing	−0.042	0.020	0.019	0.017	−0.014	0.052
			Predicting MHSA math	−0.064	0.019	0.041	0.019	0.004	0.078
			Predicting MHSA science	−0.014	0.031	−0.009	0.034	−0.075	0.057
Predicting MHSA writing	−0.042	0.020	Predicting MHSA math	−0.064	0.019	0.022	0.024	−0.025	0.069
			Predicting MHSA science	−0.014	0.031	−0.028	0.038	−0.103	0.047
Predicting MHSA math	−0.064	0.019	Predicting MHSA science	−0.014	0.031	−0.050	0.041	−0.131	0.031

MHSA is Maine High School Assessment. MEA is Maine Educational Assessment. PSAT is Preliminary SAT.

** .95 confidence interval does not contain 0.

a. The regression coefficients and their standard errors reflect the variability introduced by the imputation procedures because they were derived by combining the results across the 10 imputed datasets.

b. The significance of the difference between the standardized regression coefficients was calculated by constructing a .95 confidence interval around the difference between the standardized regression coefficients. The interval was calculated as $(\beta_1 - \beta_2) \pm 1.96(SE_{\beta_1-\beta_2})$, where $SE_{\beta_1-\beta_2} = \sqrt{(SE_1)^2 + (SE_2)^2 - 2Cov_{\beta_1\beta_2}}$. Tests of the confidence intervals around the difference between model coefficients are not adjusted for multiple comparisons.

Note: The .95 confidence intervals are presented only for school variables that are statistically significant predictors in more than one MHSA domain. Racial/ethnic minority status was not significant for any MHSA domains.

Source: Authors' calculations based on student data from data files provided by the Maine Department of Education and school data from the Maine Department of Education's Maine High School Assessment Summary Reports (http://www.maine.gov/education/mhsa/08schoolreports/index.html).

APPENDIX E
MULTILEVEL REGRESSION MODEL RESULTS

TABLE E1

Predictors of Maine High School Assessment reading scores for grade 11 students in the 2007/08 cohort

	Unconditional model[a]					
	Fixed components			Random components		
Variable	Coefficient	Standard error	Significance	Within school	Between school	Intraclass correlation (percent)
Outcome: MHSA reading scores (standardized, mean = 0, standard deviation = 1)	–0.046	0.028	0.101	0.924	0.078	**7.78**

	Model 1[a]			Model 2[a]			Model 3[a]		
	Coefficient	Standard error	Significance	Coefficient	Standard error	Significance	Coefficient	Standard error	Significance
Intercept	0.002	0.007	0.755	**0.064**	**0.015**	**< 0.001**	**0.076**	**0.030**	**0.014**
Student characteristics									
Gender (0 = male, 1 = female)	na	na	na	**–0.052**	**0.010**	**< 0.001**	–0.005	0.012	0.640
Racial/ethnic minority status (0 = no, 1 = yes)	na	na	na	–0.020	0.025	0.406	–0.064	0.036	0.077
Economically disadvantaged status (0 = no, 1 = yes)	na	na	na	**–0.035**	**0.011**	**0.002**	**–0.065**	**0.014**	**< 0.001**
Special education status (0 = no, 1 = yes)	na	na	na	**–0.065**	**0.019**	**0.001**	**–0.115**	**0.025**	**< 0.001**
English language learner status (0 = no, 1 = yes)	na	na	na	**–0.061**	**0.028**	**0.028**	**–0.158**	**0.033**	**< 0.001**
Student prior achievement measures									
Grade 8 MEA reading[b]	**0.138**	**0.009**	**< .001**	**0.138**	**0.010**	**< 0.001**	**0.324**	**0.011**	**< 0.001**
Grade 8 MEA writing[b]	**–0.027**	**0.007**	**< .001**	**–0.022**	**0.007**	**0.004**	–0.017	0.009	0.063
Grade 8 MEA math	**0.057**	**0.011**	**< .001**	**0.058**	**0.011**	**< 0.001**	**0.184**	**0.011**	**< 0.001**
Grade 8 MEA science	**0.156**	**0.010**	**< .001**	**0.147**	**0.010**	**< 0.001**	**0.332**	**0.011**	**< 0.001**
Grade 10 PSAT verbal	**0.411**	**0.009**	**< .001**	**0.410**	**0.009**	**< 0.001**	na	na	na
Grade 10 PSAT math	**0.030**	**0.010**	**0.004**	0.016	0.011	0.134	na	na	na
Grade 10 PSAT writing	**0.218**	**0.010**	**< .001**	**0.222**	**0.010**	**< 0.001**	na	na	na
School characteristics									
Percentage of racial/ethnic minority students[c]	na	na	na	–0.016	0.017	0.330	–0.059	0.031	0.060
Percentage of students who are economically disadvantaged[c]	na	na	na	**–0.013**	**0.006**	**0.035**	–0.021	0.011	0.068
Percentage of students in special education[c]	na	na	na	–0.015	0.013	0.239	–0.012	0.025	0.617
Percentage of English language learner students[c]	na	na	na	0.043	0.022	0.050	**0.120**	**0.040**	**0.004**
Percentage of students in cohort who drop out[c]	na	na	na	**0.037**	**0.015**	**0.018**	0.046	0.034	0.176

(CONTINUED)

TABLE E1 (CONTINUED)

Predictors of Maine High School Assessment reading scores for grade 11 students in the 2007/08 cohort

	Model 1[a]			Model 2[a]			Model 3[a]		
	Coefficient	Standard error	Significance	Coefficient	Standard error	Significance	Coefficient	Standard error	Significance
Student–teacher ratio	na	na	na	**–0.019**	**0.005**	**< 0.001**	**–0.034**	**0.010**	**0.001**
Mean years of teaching experience	na	na	na	0.003	0.003	0.235	0.002	0.006	0.676
Cohort size[c]	na	na	na	0.002	0.001	0.129	0.003	0.002	0.183
City versus rural location (0 = rural, 1 = city)	na	na	na	–0.042	0.025	0.090	0.053	0.048	0.268
Suburb versus rural (0 = rural, 1 = suburb)	na	na	na	0.025	0.023	0.285	**0.130**	**0.040**	**0.002**
Town versus rural (0 = rural, 1 = town)	na	na	na	–0.016	0.016	0.322	0.006	0.033	0.845
Made adequate yearly progress in reading (0 = no, 1 = yes)[d]	na	na	na	–0.023	0.015	0.138	**0.101**	**0.032**	**0.003**
Made adequate yearly progress in math (0 = no, 1 = yes)[d]	na	na	na	0.007	0.017	0.696	–0.011	0.031	0.719
Title 1 school (0 = no, 1 = yes)	na	na	na	0.021	0.020	0.304	0.013	0.035	0.706
Random components	Residual variance			Residual variance			Residual variance		
Within schools	0.208	$p < .001$		0.207	$p < .001$		0.345	$p < .001$	
Between schools	0.003	$p < .001$		0.002	$p < .001$		0.017	$p < .001$	
Total residual	0.212	$p < .001$		0.209			0.363		
Total variance explained (percent)	78.87			79.13			62.72		

na is not applicable because the variable is not included in the model.

MHSA is Maine High School Assessment. MEA is Maine Educational Assessment. PSAT is Preliminary SAT.

a. The fixed and random components in the model reflect the variability introduced by the imputation procedures because they were derived by combining the results across the 10 imputed datasets.

b. The slopes associated with these prior achievement measures were allowed to vary across schools

c. Rescaled by 10.

d. Entered into the model as 0 = yes and 1 = no to make the model intercept interpretable. The coefficients are reverse coded in the table to aid presentation.

Note: Bolded values are significant at $p < .05$.

Source: Authors' calculations based on student data from data files provided by the Maine Department of Education and school data from the Maine Department of Education's Maine High School Assessment Summary Reports (http://www.maine.gov/education/mhsa/08schoolreports/index.html).

TABLE E2

Predictors of Maine High School Assessment writing scores for grade 11 students in the 2007/08 cohort

Variable	Unconditional model[a]					
	Fixed components			Random components		
	Coefficient	Standard error	Significance	Within school	Between school	Intraclass correlation (percent)
Outcome: MHSA reading scores (standardized, mean = 0, standard deviation = 1)	–0.046	0.030	0.131	0.908	0.093	**9.25**

Variable	Model 1[a]			Model 2[a]			Model 3[a]		
	Coefficient	Standard error	Significance	Coefficient	Standard error	Significance	Coefficient	Standard error	Significance
Intercept	0.005	0.010	0.627	–0.036	0.023	0.125	0.048	0.032	0.142
Student characteristics									
Gender (0 = male, 1 = female)	na	na	na	**0.060**	**0.010**	**< .001**	**0.121**	**0.012**	**< 0.001**
Racial/ethnic minority status (0 = no, 1 = yes)	na	na	na	–0.036	0.023	0.113	**–0.084**	**0.035**	**0.016**
Economically disadvantaged status (0 = no, 1 = yes)	na	na	na	**–0.061**	**0.011**	**< .001**	**–0.094**	**0.014**	**< 0.001**
Special education status (0 = no, 1 = yes)	na	na	na	**–0.117**	**0.021**	**< .001**	**–0.168**	**0.027**	**< 0.001**
English language learner status (0 = no, 1 = yes)	na	na	na	**–0.102**	**0.032**	**0.002**	**–0.183**	**0.040**	**< 0.001**
Student prior achievement measures									
Grade 8 MEA reading	**0.145**	**0.010**	**< .001**	**0.120**	**0.010**	**< .001**	**0.288**	**0.011**	**< 0.001**
Grade 8 MEA writing	**0.078**	**0.006**	**< .001**	**0.063**	**0.006**	**< .001**	**0.082**	**0.082**	**< 0.001**
Grade 8 MEA math	**0.101**	**0.010**	**< .001**	**0.105**	**0.010**	**< .001**	**0.257**	**0.011**	**< 0.001**
Grade 8 MEA science	**0.043**	**0.010**	**< .001**	**0.053**	**0.010**	**< .001**	**0.209**	**0.209**	**< 0.001**
Grade 10 PSAT verbal	**0.214**	**0.010**	**< .001**	**0.212**	**0.010**	**< .001**	na	na	na
Grade 10 PSAT math	**0.053**	**0.010**	**< .001**	**0.051**	**0.010**	**< .001**	na	na	na
Grade 10 PSAT writing	**0.378**	**0.009**	**< .001**	**0.370**	**0.009**	**< .001**	na	na	na
School characteristics									
Percentage of racial/ethnic minority students[b]	na	na	na	–0.038	0.022	0.081	–0.061	0.035	0.082
Percentage of students who are economically disadvantaged[b]	na	na	na	–0.006	0.006	0.316	–0.015	0.012	0.217
Percentage of students in special education[b]	na	na	na	0.002	0.018	0.897	–0.007	0.030	0.817
Percentage of English language learner students[b]	na	na	na	0.058	0.037	0.121	**0.110**	**0.055**	**0.049**
Percentage of students in cohort who drop out[b]	na	na	na	0.008	0.019	0.688	0.005	0.037	0.886
Student–teacher ratio	na	na	na	**–0.017**	**0.007**	**0.013**	**–0.034**	**0.011**	**0.004**
Mean years of teaching experience	na	na	na	0.005	0.004	0.213	0.006	0.007	0.420

(CONTINUED)

TABLE E2 (CONTINUED)

Predictors of Maine High School Assessment writing scores for grade 11 students in the 2007/08 cohort

	Model 1[a]			Model 2[a]			Model 3[a]		
	Coefficient	Standard error	Significance	Coefficient	Standard error	Significance	Coefficient	Standard error	Significance
Cohort size[b]	na	na	na	0.002	0.002	0.291	0.005	0.003	0.104
City versus rural location (0 = rural, 1 = city)	na	na	na	–0.003	0.045	0.949	0.059	0.067	0.387
Suburb versus rural (0 = rural, 1 = suburb)	na	na	na	**0.111**	**0.033**	**0.002**	**0.223**	**0.053**	**< 0.001**
Town versus rural (0 = rural, 1 = town)	na	na	na	–0.019	0.025	0.445	0.021	0.042	0.612
Made adequate yearly progress in reading (0 = no, 1 = yes)[c]	na	na	na	**–0.042**	**0.020**	**0.035**	**0.125**	**0.037**	**0.001**
Made adequate yearly progress in math (0 = no, 1 = yes)[c]	na	na	na	–0.027	0.022	0.220	0.026	0.038	0.491
Title 1 school (0 = no, 1 = yes)	na	na	na	0.043	0.028	0.124	0.030	0.046	0.512
Random components	Residual variance			Residual variance			Residual variance		
Within schools	0.214	$p < .001$		0.211	$p < .001$		0.336	$p < .001$	
Between schools	0.009	$p < .001$		0.005	$p < .001$		0.019	$p < .001$	
Total residual	0.223			0.216			0.355		
Total variance explained (percent)		77.75			78.37			63.00	

na is not applicable because the variable is not included in the model.

MHSA is Maine High School Assessment. MEA is Maine Educational Assessment. PSAT is Preliminary SAT.

a. The fixed and random components in the model reflect the variability introduced by the imputation procedures because they were derived by combining the results across the 10 imputed datasets.

b. Rescaled by 10.

c. Entered into the model as 0 = yes and 1 = no to make the model intercept interpretable. The coefficients are reverse coded in the table to aid presentation.

Note: Bolded values are significant at $p < .05$.

Source: Authors' calculations based on student data from data files provided by the Maine Department of Education and school data from the Maine Department of Education's Maine High School Assessment Summary Reports (http://www.maine.gov/education/mhsa/08schoolreports/index.html).

TABLE E3

Predictors of Maine High School Assessment math scores for grade 11 students in the 2007/08 cohort

	Unconditional model[a]					
	Fixed components			Random components		
Variable	Coefficient	Standard error	Significance	Within school	Between school	Intraclass correlation (percent)
Outcome: MHSA reading scores (standardized, mean = 0, standard deviation = 1)	–0.052	0.028	0.063	0.913	0.074	**7.53**

	Model 1[a]			Model 2[a]			Model 3[a]		
	Coefficient	Standard error	Significance	Coefficient	Standard error	Significance	Coefficient	Standard error	Significance
Intercept	–0.016	0.008	0.055	**0.100**	**0.018**	**< 0.001**	**0.180**	**0.024**	**< .001**
Student characteristics									
Gender (0 = male, 1 = female)	na	na	na	**–0.088**	**0.009**	**< 0.001**	**–0.143**	**0.011**	**< 0.001**
Racial/ethnic minority status (0 = no, 1 = yes)	na	na	na	0.006	0.027	0.828	–0.014	0.034	0.684
Economically disadvantaged status (0 = no, 1 = yes)	na	na	na	**–0.039**	**0.011**	**0.001**	**–0.081**	**0.013**	**< 0.001**
Special education status (0 = no, 1 = yes)	na	na	na	**–0.105**	**0.019**	**< 0.001**	**–0.213**	**0.024**	**< 0.001**
English language learner status (0 = no, 1 = yes)	na	na	na	–0.024	0.032	0.456	–0.065	0.039	0.096
Student prior achievement measures									
Grade 8 MEA reading	**–0.025**	**0.011**	**0.018**	**–0.024**	**0.011**	**0.027**	**0.034**	**0.012**	**0.004**
Grade 8 MEA writing	–0.003	0.007	0.542	0.003	0.008	0.701	0.008	0.009	0.349
Grade 8 MEA math[b]	**0.311**	**0.012**	**< .001**	**0.313**	**0.012**	**< 0.001**	**0.602**	**0.012**	**< 0.001**
Grade 8 MEA science	**0.066**	**0.009**	**< .001**	**0.051**	**0.009**	**< 0.001**	**0.131**	**0.011**	**< 0.001**
Grade 10 PSAT verbal	**0.056**	**0.010**	**< .001**	**0.056**	**0.010**	**< 0.001**	na	na	na
Grade 10 PSAT math[b]	**0.451**	**0.014**	**< .001**	**0.429**	**0.014**	**< 0.001**	na	na	na
Grade 10 PSAT writing	**0.059**	**0.009**	**< .001**	**0.067**	**0.009**	**< 0.001**	na	na	na
School characteristics									
Percentage of racial/ethnic minority students[c]	na	na	na	–0.025	0.019	0.202	–0.059	0.023	0.013
Percentage of students who are economically disadvantaged[c]	na	na	na	0.003	0.007	0.646	0.004	0.010	0.715
Percentage of students in special education[c]	na	na	na	–0.037	0.014	0.008	–0.034	0.020	0.088
Percentage of English language learner students[c]	na	na	na	0.009	0.024	0.698	0.049	0.031	0.115
Percentage of students in cohort who drop out[c]	na	na	na	–0.005	0.016	0.779	0.005	0.027	0.851
Student–teacher ratio	na	na	na	**–0.015**	**0.005**	**0.004**	**–0.022**	**0.007**	**0.005**
Mean years of teaching experience	na	na	na	0.001	0.003	0.962	–0.001	0.004	0.736

(CONTINUED)

TABLE E3 (CONTINUED)

Predictors of Maine High School Assessment math scores for grade 11 students in the 2007/08 cohort

	Model 1[a]			Model 2[a]			Model 3[a]		
	Coefficient	Standard error	Significance	Coefficient	Standard error	Significance	Coefficient	Standard error	Significance
Cohort size[c]	na	na	na	**0.003**	**0.001**	**0.032**	**0.006**	**0.002**	**0.008**
City versus rural location (0 = rural, 1 = city)	na	na	na	0.004	0.023	0.853	0.001	0.032	0.976
Suburb versus rural (0 = rural, 1 = suburb)	na	na	na	0.029	0.026	0.272	**0.089**	**0.039**	**0.025**
Town versus rural (0 = rural, 1 = town)	na	na	na	0.001	0.017	0.976	0.014	0.024	0.570
Made adequate yearly progress in reading (0 = no, 1 = yes)[d]	na	na	na	**–0.064**	**0.019**	**0.001**	**0.128**	**0.027**	**< 0.000**
Made adequate yearly progress in math (0 = no, 1 = yes)[d]	na	na	na	–0.027	0.016	0.092	0.032	0.024	0.186
Title 1 school (0 = no, 1 = yes)	na	na	na	0.007	0.018	0.713	0.006	0.027	0.834
Random components	Residual variance			Residual variance			Residual variance		
Within schools	0.224		$p < .001$	0.222		$p < .001$	0.305		$p < .001$
Between schools	0.005		$p < .001$	0.002		$p < .001$	0.018		$p < .001$
Total residual	0.229			0.224			0.323		
Total variance explained (percent)		76.82			77.29			66.59	

na is not applicable because the variable is not included in the model.

MHSA is Maine High School Assessment. MEA is Maine Educational Assessment. PSAT is Preliminary SAT.

a. The fixed and random components in the model reflect the variability introduced by the imputation procedures because they were derived by combining the results across the 10 imputed datasets.

b. The slopes associated with these prior achievement measures were allowed to vary across schools.

c. Rescaled by 10.

d. Entered into the model as 0 = yes and 1 = no to make the model intercept interpretable. The coefficients are reverse coded in the table to aid presentation.

Note: Bolded values are significant at $p < 0.05$.

Source: Authors' calculations based on student data from data files provided by the Maine Department of Education and school data from the Maine Department of Education's Maine High School Assessment Summary Reports (http://www.maine.gov/education/mhsa/08schoolreports/index.html).

TABLE E4

Predictors of Maine High School Assessment science scores for grade 11 students in the 2007/08 cohort

	Unconditional model[a]								
	Fixed components			Random components					
Variable	Coefficient	Standard error	Significance	Within school	Between school	Intraclass correlation (percent)			
Outcome: MHSA reading scores (standardized, mean = 0, standard deviation = 1)	–0.049	0.026	0.068	0.930	0.068	**6.77**			
	Model 1[a]			Model 2[a]			Model 3[a]		
	Coefficient	Standard error	Significance	Coefficient	Standard error	Significance	Coefficient	Standard error	Significance
Intercept	–0.009	0.013	0.491	**0.095**	**0.033**	**0.005**	**0.110**	**0.036**	**0.003**
Student characteristics									
Gender (0 = male, 1 = female)	na	na	na	**–0.183**	**0.011**	**<0.001**	**–0.187**	**0.011**	**<0.001**
Racial/ethnic minority status (0 = no, 1 = yes)	na	na	na	0.003	0.028	0.920	–0.018	0.033	0.581
Economically disadvantaged status (0 = no, 1 = yes)	na	na	na	–0.015	0.012	0.200	**–0.040**	**0.013**	**0.003**
Special education status (0 = no, 1 = yes)	na	na	na	–0.032	0.018	0.085	**–0.097**	**0.022**	**<0.001**
English language learner status (0 = no, 1 = yes)	na	na	na	**–0.100**	**0.030**	**0.001**	**–0.156**	**0.032**	**<0.001**
Student prior achievement measures									
Grade 8 MEA reading	–0.004	0.012	0.712	0.022	0.012	0.067	**0.118**	**0.012**	**<0.001**
Grade 8 MEA writing	**–0.075**	**0.008**	**<.001**	**–0.050**	**0.008**	**<0.001**	**–0.044**	**0.009**	**<0.001**
Grade 8 MEA math	**0.128**	**0.014**	**<.001**	**0.127**	**0.014**	**<0.001**	**0.290**	**0.011**	**<0.001**
Grade 8 MEA science[b]	**0.348**	**0.012**	**<.001**	**0.314**	**0.012**	**<0.001**	**0.419**	**0.014**	**<0.001**
Grade 10 PSAT verbal[b]	**0.227**	**0.012**	**<.001**	**0.229**	**0.012**	**<0.001**	na	na	na
Grade 10 PSAT math[b]	**0.200**	**0.013**	**<.001**	**0.171**	**0.013**	**<0.001**	na	na	na
Grade 10 PSAT writing	**0.052**	**0.011**	**<.001**	**0.071**	**0.011**	**<0.001**	na	na	na
School characteristics									
Percentage of racial/ethnic minority students[c]	na	na	na	0.017	0.027	0.525	–0.011	0.026	0.679
Percentage of students who are economically disadvantaged[c]	na	na	na	0.003	0.010	0.775	0.001	0.012	0.959
Percentage of students in special education[c]	na	na	na	**–0.054**	**0.021**	**0.013**	**–0.048**	**0.024**	**0.049**
Percentage of English language learner students[c]	na	na	na	–0.061	0.032	0.057	–0.027	0.032	0.401
Percentage of students in cohort who drop out[c]	na	na	na	–0.008	0.025	0.758	–0.018	0.030	0.549
Student–teacher ratio	na	na	na	–0.001	0.008	0.898	–0.011	0.010	0.279
Mean years of teaching experience	na	na	na	0.004	0.004	0.300	0.003	0.005	0.551

(CONTINUED)

TABLE E4 (CONTINUED)

Predictors of Maine High School Assessment science scores for grade 11 students in the 2007/08 cohort

	Model 1[a]			Model 2[a]			Model 3[a]		
	Coefficient	Standard error	Significance	Coefficient	Standard error	Significance	Coefficient	Standard error	Significance
Cohort size[c]	na	na	na	0.001	0.002	0.923	0.001	0.002	0.683
City versus rural location (0 = rural, 1 = city)	na	na	na	0.027	0.047	0.560	0.082	0.053	0.123
Suburb versus rural (0 = rural, 1 = suburb)	na	na	na	0.058	0.039	0.146	**0.140**	**0.045**	**0.003**
Town versus rural (0 = rural, 1 = town)	na	na	na	–0.019	0.029	0.508	0.006	0.035	0.871
Made adequate yearly progress in math (0 = no, 1 = yes)[d]	na	na	na	–0.014	0.031	0.641	0.046	0.035	0.194
Title 1 school (0 = no, 1 = yes)	na	na	na	–0.006	0.026	0.823	0.000	0.030	0.991
City versus rural location (0 = rural, 1 = city)	na	na	na	0.014	0.028	0.628	–0.001	0.030	0.977
Random components	Residual variance			Residual variance			Residual variance		
Within schools	0.305		$p < .001$	0.298		$p < .001$	0.358		$p < .001$
Between schools	0.016		$p < .001$	0.014		$p < .001$	0.026		$p < .001$
Total residual	0.321			0.312			0.384		
Total variance explained (percent)		67.87			68.70			61.53	

na is not applicable because the variable is not included in the model.

MHSA is Maine High School Assessment. MEA is Maine Educational Assessment. PSAT is Preliminary SAT.

a. The fixed and random components in the model reflect the variability introduced by the imputation procedures because they were derived by combining the results across the 10 imputed datasets.

b. The slopes associated with these prior achievement measures were allowed to vary across schools.

c. Rescaled by 10.

d. Entered into the model as 0 = yes and 1 = no to make the model intercept interpretable. The coefficients are reverse coded in the table to aid presentation.

Note: Bolded values are significant at $p < .05$.

Source: Authors' calculations based on student data from data files provided by the Maine Department of Education and school data from the Maine Department of Education's Maine High School Assessment Summary Reports (http://www.maine.gov/education/mhsa/08schoolreports/index.html).

APPENDIX F
PREDICTED DIFFERENCES ACROSS MAINE HIGH SCHOOL ASSESSMENT DOMAINS FOR STUDENT AND SCHOOL CHARACTERISTICS

In examining the results for model 2, the primary model used to answer this study's research question, it is also possible to compare the predicted differences across MHSA domains (reading, writing, math, and science) for each student and school characteristic, holding all other variables in the model constant. For example, it is possible to compare the predicted difference in MHSA reading scores associated with being a racial/ethnic minority student to the predicted change in MHSA writing scores associated with being a racial/ethnic minority student. The discussion below focuses on the predicted scale score point differences only, and the .95 confidence intervals for comparing the strength of the standardized relationships across MHSA domains for student or school characteristics allow conclusions to be made about whether a student or school characteristic was a significantly stronger predictor for one MHSA domain than for another. The .95 confidence intervals are reported in table D3 in appendix D.

Gender

Across all four MHSA domains, student gender was a significant predictor. For MHSA reading, math, and science domains, male students were predicted to have significantly higher scores than female students. Among these three domains, the largest scale score point difference was observed for MHSA science (1.69 scale score points or 0.183 standard deviation). The relationship between gender and MHSA science scores was significantly stronger than the relationships between gender and MHSA scores in reading and math. For MHSA writing, the predicted scores for female students were 0.83 scale score point (0.060 standard deviation) higher than those for male students.

Racial/ethnic minority status

Student racial/ethnic minority status was not associated with significant changes in predicted MHSA scale scores.

Economically disadvantaged status

Economically disadvantaged students were predicted to have significantly lower MHSA reading, writing, and math scores than non–economically disadvantaged students. MHSA science scores were not significantly different for economically disadvantaged students. Across MHSA domains, being economically disadvantaged was associated with a decrease of 0.51 scale score point in reading scores, 0.84 scale score point in writing scores, and 0.43 scale score point in math scores. Based on the confidence intervals associated with the coefficients, writing and reading were significantly different, and writing and science were statistically different.

Special education status

Special education students were predicted to have significantly lower MHSA reading, writing, and math scores than general education students. There was no significant difference between special education and general education students' predicted MHSA science scores. The magnitude of the predicted difference between special education and general education students' scores was largest in the MHSA writing (–1.62 scale points) and math domains (–1.16 scale score points). The .95 confidence interval shows that the relationships observed for the MHSA writing and math domains were not significantly different from each other. The relationship observed for MHSA reading scores was significantly different from both MHSA writing and math scores.

English language learner status

English language learner students were predicted to have significantly lower MHSA reading, writing, and science scores than non–English language learner students. MHSA math scores were not significantly different for English language learner students. The largest decrease in MHSA scores associated with English language learner status was observed for MHSA writing scores; English language learner students were predicted to score 1.41 scale score points (0.102 standard deviation) lower than non–English language learner students. The predicted decrease in MHSA reading and science scores associated with being an English language learner student were similar in magnitude (0.89 and 0.92 scale score points, respectively). There were no significant differences observed between the strength of the relationships across MHSA domains.

Percentage of economically disadvantaged students

Students in schools with higher percentages of grade 11 economically disadvantaged students were predicted to have significantly lower MHSA reading scores. Specifically, a student in a school where the percentage of grade 11 economically disadvantaged students is 10 points higher than the mean for all schools in the sample is predicted to have MHSA reading scores that are 0.19 scale score point (0.013 standard deviation) lower than a similar student in a school with the sample average percentage of grade 11 economically disadvantaged students. There were no significant differences observed between the strength of the coefficients for the percentage of grade 11 economically disadvantaged students in a school across MHSA domains.

Percentage of special education students

Significantly lower MHSA science scores were associated with schools with higher percentages of grade 11 special education students. For every 10 point increase in the percentage of grade 11 special students over the mean for the sample, MHSA science scores were predicted to decrease by 0.50 scale score point (0.054 standard deviation). The relationship between the percentage of grade 11 special education students in a school and MHSA writing scores was significantly weaker than the relationship between the percentage of grade 11 special education students in a school and MHSA science scores.

Percentage of students in cohort who drop out

Students in schools with higher cohort dropout rates were predicted to have significantly higher MHSA reading scores. A student in a school where the percentage of grade 11 dropouts is 10 points higher than the mean for all schools in the sample is predicted to have an MHSA reading score that is 0.54 scale score point higher than a similar student in a similar school with the sample average cohort dropout rate. While the reason for this is unclear from the results, this finding is based on holding multiple variables constant and may not appear the same in a simpler model where other student and school characteristics are not considered. The relationship between the percentage of students in the cohort who drop out and MHSA reading scores was significantly stronger than the relationship between the percentage of students in the cohort who drop out and MHSA writing scores.

Student–teacher ratio

Students in schools with student–teacher ratios larger than the average for the whole sample were predicted to have significantly lower MHSA reading, writing, and math scores. The predicted decreases in MHSA score for an increase of five students per teacher were 1.38 scale score points for reading, 1.17 scale score points for writing, and 0.83 scale score point for math. There were no significant differences observed between the strength of the coefficient for student–teacher ratio across MHSA domains.

Cohort size

For MHSA math scores a 10 student increase in the size of the grade 11 cohort above the grand mean for all schools in the sample was associated with an increase of 0.03 scale score point (0.003 standard deviation). There were no significant differences observed between the strength of the coefficients for cohort size across MHSA domains.

School location

Across MHSA domains, there were no significant differences between students' scores in rural and city locales or between students' scores in rural and town locales. However, when compared with similar students in rural schools, students in suburban schools were predicted to have MHSA writing scale scores that were 1.53 scale score points (0.111 standard deviation) higher. The difference in MHSA writing scores between suburban and rural schools was significantly larger than the difference in MHSA reading, math, and science scores between suburban and rural schools. The difference in MHSA math scores between suburban and rural schools was significantly larger than the difference in MHSA reading scores between rural and suburban schools.

Made adequate yearly progress in reading

Whether a school made adequate yearly progress in reading was significantly associated with MHSA writing and math scores. In this case, the direction of the relationship between making adequate yearly progress in reading and MHSA writing scores is the opposite of the direction of the relationship between making adequate yearly progress in reading and MHSA math scores. A student in a school that failed to make adequate yearly progress in for reading was predicted to have an MHSA writing score that was 0.58 scale score point (0.042 standard deviation) lower than a similar student in a similar school that did make adequate yearly progress in reading. For MHSA math, failing to make adequate yearly progress in reading was associated with a decrease of 0.71 scale score point (0.064 standard deviation). There were no significant differences observed between the strength of the coefficient associated with having made adequate yearly progress in reading across MHSA domains.

Other school characteristics

Percentage of racial/ethnic minority students, percentage of economically disadvantaged students, mean years of teaching experience, made adequate yearly progress in math, and Title I status were not significantly associated with differences in MHSA scale scores for any domain.

APPENDIX G
PREDICTORS OF MAINE HIGH SCHOOL ASSESSMENT DOMAIN SCORES WHEN PRELIMINARY SAT SCORES ARE EXCLUDED FROM THE MODEL (MODEL 3)

TABLE G1

Maine High School Assessment domain scores regressed on student characteristics, grade 8 Maine Educational Assessment scores, and school characteristics (model 3)

	Standard deviation units[a]				Scale score points[a]			
Variable	Reading	Writing	Math	Science	Reading	Writing	Math	Science
Mean	0.00	0.00	0.00	0.00	1,140.69	1,139.95	1,141.22	1,140.76
Standard deviation	1.00	1.00	1.00	1.00	14.55	13.82	11.02	9.23
	Model results (coefficient)[a]							
Intercept	**0.076**	0.048	**0.180**	**0.110**	1,141.8	1,140.68	1,143.2	1,149.99
Student characteristics								
Gender (0 = male, 1 = female)	−0.005	**0.121**	**−0.143**	**−0.187**	−0.07	**1.67**	**−1.58**	**−1.72**
Racial/ethnic minority status (0 = no, 1 = yes)	−0.064	**−0.084**	−0.014	−0.018	−0.93	**−1.16**	−0.15	−0.17
Economic disadvantage status (0 = no, 1 = yes)	**−0.065**	**−0.094**	**−0.081**	**−0.040**	**−0.95**	**−1.3**	**−0.89**	**−0.37**
English language learner status (0 = no, 1 = yes)	**−0.115**	**−0.168**	**−0.213**	**−0.097**	**−1.67**	**−2.32**	**−2.35**	**−0.89**
Gender (0 = male, 1 = female)	**−0.158**	**−0.183**	−0.065	**−0.156**	**−2.30**	**−2.53**	−0.72	**−1.44**
Student prior achievement measures					Prior achievement scores in units of 10 scale points			
Grade 8 MEA scores								
Reading	**0.324**[b]	**0.288**	**0.034**	**0.118**	**3.93**[b]	**3.31**	**0.32**	**0.91**
Writing	−0.017[b]	**0.082**	0.008	**−0.044**	−0.28[b]	**1.28**	0.100	**−0.46**
Math	**0.184**	**0.257**	**0.602**[b]	**0.290**	**1.82**	**2.42**	**4.52**[b]	**1.82**
Science	**0.332**	**0.209**	**0.131**	**0.419**[b]	**4.08**	**2.44**	**1.22**	**3.75**[b]
School characteristics								
Percentage of grade 11 racial/ethnic minority students[c]	−0.059	−0.061	**−0.059**	−0.011	−0.86	−0.84	**−0.65**	−0.10
Percentage of grade 11 students who are economically disadvantaged[c]	−0.021	−0.015	0.004	0.001	−0.31	−0.21	0.04	0.01
Percentage of grade 11 students in special education[c]	−0.012	−0.007	−0.034	**−0.048**	−0.18	−0.10	−0.037	**−0.44**
Percentage of grade 11 students who are English language learners[c]	**0.120**	**0.110**	0.049	−0.027	**1.75**	**1.52**	0.54	−0.25
Percentage of students in cohort who drop out[c]	0.046	0.005	0.005	−0.018	0.67	0.07	0.06	−0.17
Student–teacher ratio	**−0.034**	**−0.034**	**−0.022**	−0.011	**−0.50**	**−0.47**	**−0.24**	−0.10

(CONTINUED)

	Standard deviation units[a]				Scale score points[a]			
Variable	Reading	Writing	Math	Science	Reading	Writing	Math	Science
Mean years of teaching experience	0.002	0.006	–0.001	0.003	0.11	0.30	–0.04	0.10
Cohort size[c]	0.003	0.005	**0.006**	0.001	0.04	0.07	**0.07**	0.01
City versus rural (0 = rural, 1 = city)	0.053	0.059	0.001	0.082	0.77	0.81	0.01	0.76
Suburb versus rural (0 = rural, 1 = suburb)	**0.130**	**0.223**	**0.089**	**0.140**	**1.89**	**3.08**	**0.98**	**1.29**
Town versus rural (0 = rural, 1 = town)	0.006	0.021	0.014	0.006	0.09	0.29	0.15	0.06
City versus rural (0 = rural, 1 = city)	0.053	0.059	0.001	0.082	0.77	0.81	0.01	0.76
Made adequate yearly progress for reading in 2007 (0 = no, 1 = yes)[d]	**0.101**	**0.125**	**0.128**	0.046	**1.47**	**1.73**	**1.41**	0.43
Made adequate yearly progress for math in 2007 (0 = no, 1 = yes)[d]	–0.011	0.026	0.032	0.000	–0.16	0.37	0.35	–0.00
Classified as Title I school in 2007 (0 = no, 1 = yes)	0.013	0.030	0.006	–0.001	0.19	0.42	0.07	–0.01

a. The regression coefficients reflect the variability introduced by the imputation procedures because they were derived by combining the results across the 10 imputed datasets.

b. The slopes associated with these prior achievement measures were allowed to vary across schools. See appendix C for statistical justification for allowing these slopes to vary randomly.

c. Rescaled by 10.

d. Entered into the model as 0 = yes, and 1 = no to make the model intercept interpretable. The coefficients are reverse coded to aid presentation.

Note: Values in bold are significant at $p < .05$.

Source: Authors' calculations based on student data from data files provided by the Maine Department of Education and school data from the Maine Department of Education's Maine High School Assessment Summary Reports (http://www.maine.gov/education/mhsa/08schoolreports/index.html).

NOTES

1. Because the PSAT is a preliminary assessment for the SAT (which constitutes the reading, writing, and math sections of the MHSA) and is administered at the beginning of grade 10, the strong relationship between the PSAT and MHSA could influence both the magnitude of the multilevel regression coefficients associated with the other variables and the overall interpretation of the model.

2. In 2009, the grade 8 assessment in Maine changed from the MEA, taken by the 2007/08 cohort of students, to the New England Common Assessment Program (NECAP). Adopting the NECAP allows the Maine Department of Education to compare grade 3–8 student performance in Maine with that in the other NECAP states (New Hampshire, Rhode Island, and Vermont). Despite the potential differences between the MEA and the NECAP now being adopted in the state, available grade 8 MEA scores represent measures of achievement prior to students' entry to their respective high schools and are therefore still of interest to the Maine Department of Education as it examines how information in its evolving longitudinal data system can be used to improve the education of students in Maine.

3. Tests of the confidence intervals around the difference between model coefficients are not adjusted for multiple comparisons.

4. While coefficients are described as increasing or decreasing, the significance of the difference between coefficients for the same predictor with and without PSAT scores was not tested. Therefore, the relative strength of the coefficients across the two models should be interpreted with caution.

5. With no other variables in the model, the results showed that female students had significantly higher MHSA reading and writing scores than did male students ($b = 0.110$, $t = 6.499$, df = 114, 13,006, $p < .001$ and $b = 0.286$, $t = 15.768$, df = 114, 13,006, $p < .001$, respectively). Conversely, male students had significantly higher MHSA math and science scores than did female students ($b = 0.126$, $t = -7.042$, df = 114, 13,006, $p < .001$ and $b = -0.120$, $t = -11.444$, df = 114, 13,006, $p < .001$, respectively). These findings for gender are consistent with National Assessment of Educational Progress results.

6. The descriptive statistics in table A2 in appendix A are based on the sample of 13,008 individual students who were eligible to be included in the final analysis dataset and were enrolled in the 115 publicly funded schools. Conversely, the summary statistics in table A3 in appendix A for the 115 schools are based on the data provided by the school reports prepared by the Maine Department of Education (http://www.maine.gov/education/mhsa/08schoolreports/index.html) and represent the average percentages of grade 11 students enrolled during the testing period who have a particular characteristic (such as being a minority, economically disadvantaged, in special education, or an English language learner student) across all schools.

7. The significance of the difference between the standardized regression coefficients was calculated by constructing a .95 confidence interval around the difference between the standardized regression coefficients. The interval was calculated as follows:

   $$\beta_1 - \beta_2 \pm 1.96\ (SE_{\beta_1-\beta_2})$$

   where

   $$SE_{\beta_1-\beta_2} = \sqrt{(SE_1)^2 + (SE_2)^2 - 2Cov_{\beta_1-\beta_2}}.$$

   When 0 was not in the interval around the difference between the regression coefficients, it was concluded that there was a statistically significant difference between the coefficients. This means that over an infinite number of random samples there is 95 percent confident that the interval around the difference between the regression coefficients does not include the population mean difference between the regression coefficients.

8. Using *Design effect* = 1 + (*n* – 1) * *ICC,* the design effects were 9.61 for reading, 11.14 for writing, 9.18 for math, and 8.41 for science.

REFERENCES

Allison, P. D. (2002). *Missing data*. Thousand Oaks, CA: Sage Publications.

Astone, N. M., and McLanahan, S. (1991). Family structure, parental practices, and high school completion. *American Sociology Review, 56,* 309–320.

Baker, D. P., and Stevenson, D. L. (1986). Mothers' strategies for children's school achievement: managing the transition to high school. *Sociology and Education, 59*, 156–166.

Bankston, C. L., III, and Caldas, S. J. (1998). Family structure, schoolmates, and racial inequalities in school achievement. *Journal of Marriage and Family, 60,* 715–723.

Bickel, R. (2007). *Multilevel analysis for applied research: it's just regression!* New York: The Guilford Press.

Carbonaro, W. (2005). Tracking, students' effort, and academic achievement. *Sociology of Education, 78*, 27–49.

Coe, R., and Fitz-Gibbon, C. T. (1998). School effectiveness research: criticisms and recommendations. *Oxford Review of Education*, *24*(4), 421–438.

Coleman, J. S. (1966). *Equality of educational opportunity*. U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.

College Board. (2005). *Report for the state of Maine on the alignment of the SAT and PSAT/NMSQT to the Maine Learning Results*. Internal report provided to the Maine Department of Education. Augusta, ME: College Board.

Eamon, M. K. (2005). Social-demographic, school, neighborhood, and parenting influences on the academic achievement of Latino young adolescents. *Journal of Youth and Adolescence, 34*(2), 163–174.

Fitz-Gibbon, C. (1996). *Monitoring education: indicators, quality, and effectiveness.* London: Cassell Press.

Gelman, A., and Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models.* Cambridge, UK: Cambridge University Press.

Glass, G. V., and Smith, M. E. (1979). Meta-analysis of research on class size and achievement. *Educational Evaluation and Policy Analysis, 1,* 2–16.

Hanushek, E., and Krueger, A. (2000). The class size policy debate. (Working Paper 121). Washington, DC: Economic Policy Institute.

Hedges, L. V., and Hedberg, E. (2007). Intraclass correlations for planning group randomized experiments in rural education. *Journal of Research in Rural Education, 22*(10), 1–15.

Hedges, L. V., and Stock, W. (1983). The effects of class size: an examination of rival hypotheses. *American Educational Research Journal, 20*, 63–85.

Jencks, C., Smith, M., Acland, H., Bane, M. J., Cohen, D., Gintis, H., Heyns, B., and Michaelson, S. (1972). *Inequality: a reassessment of the effect of family and schooling in America*. New York: Basic Books.

Johnson, M. K., Crosnoe, R., and Elder, G., Jr. (2001). Students' attachment and academic engagement: the role of race and ethnicity. *Sociology of Education, 74,* 318–340.

Kalambouka, A., Farrell, P., Dyson, A., and Kaplan, I. (2007). The impact of placing pupils with special education needs in mainstream schools on the achievement of their peers. *Educational Research, 49*(4), 365–382.

Lee, V. E. (2000). Using hierarchical linear modeling to study social contexts: the case of school effects. *Educational Psychologist, 35*(2), 125–141.

Little, R. J. A., and Rubin, D. B. (1987). *Statistical analysis with missing data.* New York: John Wiley & Sons, Inc.

Ma, X., and Wilkins, J. L. M. (2002). The development of science achievement in middle and high school:

individual differences and school effects. *Evaluation Review, 26*(4), 23.

Maine Department of Education. (2008). *State accountability manual, SY 2007–2008*. Augusta, ME: Maine Department of Education.

Martin, M. O., Mullis, I. V. S., Gregory, K. D., Hoyle, C., and Shen, C. (2001). *Effective schools in science and math: IEA's Third International Math and Science Study* (*TIMSS*). Chestnut Hill, MA: Boston College, Lynch School of Education, TIMSS International Study Center.

Marzano, R. J. (2000). *A new era of school reform: going where the research takes us*. Aurora, CO: Mid-Continent Research for Education and Learning.

McMillen, B. J. (2004). School size, achievement, and achievement gaps. *Education Policy Analysis Archives, 12*(58), 1–26.

Murnane, R. J. (1975). *The impact of school resources on the learning of inner city children*. Cambridge, MA: Ballinger Publishing Company.

Murnane, R. J., Maynard, R. A., and Ohls, J. C. (1980). Home resources and children's achievement. *Review of Economics Statistics, 63*(3), 369–376.

Myers, S. L., Jr., Kim, H., and Mandala, C. (2004). The effect of school poverty on racial gaps in test scores: the case of the Minnesota Basic Standards Tests. *Journal of Negro Education, 73*(1), 81–98.

Parcel, T. L., and Dufur, M. J. (2001). Capital at home and at school: effects on student achievement. *Social Forces, 79*(3), 881–992.

Parker, C. E., Louie, J., and O'Dwyer, L. M. (2009). *New measures of English language proficiency and their relationship to performance on large-scale content assessments* (Issues & Answers Report, REL 2009–066). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Northeast and Islands. Retrieved March 18, 2010, from http://ies.ed.gov/ncee/edlabs.

Raudenbush, S. W., and Bryk, A. S. (2002). *Hierarchical linear models: applications and data analysis methods* (*advanced quantitative techniques in the social sciences*). Thousand Oaks, CA: Sage Publications.

Reardon, S. F., and Galindo, C. (2007). Patterns of Hispanic students' math skill proficiency in the early elementary grades. *Journal of Latinos and Education, 6*(3), 229–251.

Reynolds, D., Teddlie, C., Creemers, B., Scheerens, J., and Townsend, T. (2000). An introduction to school effectiveness research. In D. Reynolds (Ed.), *The international handbook of school effectiveness research*. London: Falmer Press.

Rice, J. K. (2002). Making the evidence matter: implications of the class size research debate for policymakers. In L. Mishel and R. Rothstein (Eds.), *The class size debate*. Washington, DC: Economic Policy Institute.

Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. New York: John Wiley and Sons.

Rumberger, R. W., and Palardy, G. J. (2005). Test scores, dropout rates, and transfer rates as alternative indicators of high school performance. *American Educational Research Journal, 42*, 3–42.

Sammons, P., Mortimore, P., and Thomas, S. (1993, September). *Do schools perform consistently across outcomes and areas?* Paper presented at the annual conference of the British Educational Research Association, Oxford, UK.

Sammons, P., Mortimore, P., and Thomas, S. (1996). Do schools perform consistently across outcomes and areas? In J. Gray, D. Reynolds, C. Fitz-Gibbon, and D. Jesson (Eds.), *Merging traditions: the future of research on school effectiveness and school improvement*. London: Cassell Press.

Sánchez, M. T., Ehrlich, S., Midouhas, E., and O'Dwyer, L. M. (2009). *Analyzing performance by grade 10 Hispanic students on the Massachusetts state assessment* (Issues & Answers Report, REL 2009–071). Washington, DC:

U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Northeast and Islands. Retrieved March 18, 2010, from http://ies.ed.gov/ncee/edlabs

Sewell, W. H., Haller, A. O., and Portes, A. (1969). The educational and early occupational attainment process. *American Sociological Review, 34*, 82–92.

Sewell, W. H., and Shah, V. P. (1968). Social class, parental encouragement, and educational aspirations. *American Journal of Sociology, 73*, 559–572.

Shannon, G. S., and Bylsma, P. (2007). *Nine characteristics of high-performing schools: a research-based resource for schools and districts to assist with improving student learning.* (2nd. ed.). Olympia, WA: Washington Office of the State Superintendent of Public Instruction.

Sirin, S. R. (2005). Socioeconomic status and academic achievement: a meta-analytic review of research. *Review of Educational Research, 75*(3), 417–453.

Smith, D., and Tomlinson, S. (1989). *The school effect: a study of multiracial comprehensives.* London: Policy Studies Institute.

Snijders, T., and Bosker, R. (1999). *Multilevel analysis: an introduction to basic and advanced multilevel modeling.* Thousand Oaks, CA: Sage Publications.

Teddlie, C., Reynolds, D., and Sammons, P. (2000). The methodology and scientific properties of school effectiveness research. In D. Reynolds (Ed.), *The international handbook of school effectiveness research*. London: Falmer Press.

Terwilliger, J. S., and Magnuson, P. (2005). *Limited English proficiency, race/ethnicity and socio-economic status as influences on scores in large-scale assessments*. Roseville, MN: Minnesota Department of Education.

Walberg, H. J. (1980). *A psychological theory of educational productivity.* Washington, DC: National Institute of Education. (ERIC Document Reproduction Service No. ED 206 042).