

# SUPPLEMENTAL MATERIAL FOR GAO-24-106360:

## Technical Materials for Regression Analyses on Student, Teacher, and School Characteristics Associated with English Learners' Academic Performance

This product is a supplement to *K-12 Education: Student, Teacher, and School Characteristics Associated with English Learners' Academic Performance* ([GAO-24-106360](#)).

### Background

This electronic supplement serves as a companion to [GAO-24-106360](#), *K-12 Education: Student, Teacher, and School Characteristics Associated with English Learners' Academic Performance*. This supplement presents technical information about our regression analyses of three Department of Education datasets:

- The National Assessment of Educational Progress (NAEP) 2019 data for fourth and eighth grade reading and math.
- The Early Childhood Longitudinal Study (ECLS), Kindergarten Class of 2010-11, which followed a sample of children from kindergarten through the fifth grade.
- ED*Facts*, which centralizes performance data supplied by state educational agencies, including performance on state assessments and progress toward English proficiency. We analyzed data for the 2018-19, 2019-20, and 2021-22 school years.

To explore the relationship between several academic performance outcomes and a variety of student, teacher, and school characteristics, such as school attendance, we conducted regression analyses on these data sets. Such models allowed us to test the association between student performance outcomes and any one specific student, teacher, and school characteristic, while holding other characteristics (such as school demographics) constant. Technical details of our analyses are linked below.

The performance outcomes we included were

- English learners' scores on national reading and math assessments,
- growth in English learners' reading scores from the longitudinal study, and
- state reports of their students' progress toward English proficiency and rates of English proficiency.<sup>1</sup>

---

<sup>1</sup> In general, English learners are K-12 students whose native language is not English, and whose difficulties in speaking, reading, writing, or understanding English may affect their academic success in classrooms where instruction is in English. For federal reporting purposes, students are considered English learners only while they are eligible to receive language instruction services. See 20 U.S.C. § 7801(20) for the definition of an English learner under the Elementary and Secondary Education Act of 1965, as amended.

We conducted the work upon which this supplement is based in accordance with generally accepted government auditing standards from November 2022 to July 2024. Those standards require that we plan and perform the audit to obtain sufficient, appropriate evidence to provide a reasonable basis for our findings and conclusions based on our audit objectives. We believe that the evidence obtained provides a reasonable basis for our findings and conclusions based on our audit objectives.

## **Contents**

[Technical Description of Regression Analyses](#)

[Regression Analysis Using National Assessment of Educational Progress Data](#)

[Regression Analysis of English Learner Progress in Early Childhood Longitudinal Study, Kindergarten Class of 2010-2011](#)

[Regression Analysis of English Learner Progress toward Proficiency in ED\*Facts\* Data](#)

## **Contacts**

For more information, contact Jacqueline M. Nowicki at (202) 512-7215 or [NowickiJ@gao.gov](mailto:NowickiJ@gao.gov).

Office of Public Affairs; Sarah Kaczmarek, Acting Managing Director (202) 512-4800 or [Kaczmarkek@gao.gov](mailto:Kaczmarkek@gao.gov).

## **Copyright**

This is a work of the U.S. government and is not subject to copyright protection in the United States. The published product may be reproduced and distributed in its entirety without further permission from GAO. However, because this work may contain copyrighted images or other material, permission from the copyright holder may be necessary if you wish to reproduce this material separately.

## **GAO Contact**

Jacqueline M. Nowicki

Director, Education, Workforce, and Income Security Issues

[NowickiJ@gao.gov](mailto:NowickiJ@gao.gov)

(202) 512-7215

## Technical Description of Regression Analyses

We conducted multivariate regressions using student-level performance data from the 2019 NAEP, the ECLS, Kindergarten Class of 2010-2011, and ED*Facts* from 2018, 2019, and 2021 to explore the relationship between several academic performance outcomes and a variety of student, teacher, and school characteristics. We performed regressions to estimate the average differences in test scores associated with a given characteristic, while controlling for other characteristics. For example, we estimated the expected average difference in math test scores for English learners in rural schools versus urban schools, while holding other characteristics, such as student and school demographics, constant. The unit of analysis for these regressions is at the student-level.

All regression models are subject to limitations. For these models, the limitations included:

- Some variables that may be associated with academic performance of English learners are not available in the data. For example, it could be that a student's household income adjusted for family size could be related to an English learner student's academic performance. Omission of such factors could potentially introduce bias into the analysis. To account for this, we used a robust set of control factors at the school, student, and teacher levels.
- Missing data such as unit and item nonresponse cause a loss in sample size, which limits the power of the analysis to detect significant associations. To the extent such missing data is highly associated with characteristics we could not control for in our analysis, there is the potential for bias in our regression estimates and their standard errors.
- The results of our analyses are associational and do not imply a causal relationship. The data we used are observational in nature and were not gathered by a randomized controlled trial, where schools and students within schools would be randomized by English learner status.
- Our models controlled for factors of interest using fixed effects. To the extent that there is variability in academic performance of English learners in our sample by geography or school, the standard errors for model estimates may be too small and thus have a higher chance of detecting a significant effect for a factor that is not actually significant.
- Additionally, both NAEP and ECLS data are subject to both sampling and nonsampling error. While the analysis accounted for sampling error, survey data are also affected by nonsampling error, which cannot be accounted for. Nonsampling error could occur for many reasons, including a failure to sample a segment of the population, inability to obtain information for all respondents in the sample, inability or unwillingness of respondents to provide accurate information, mistakes by respondents, and errors made in the collection or processing of data (such as imputation or data quality checks).

We assessed the reliability of the data by reviewing technical documentation and Education reports on the data, and testing for missing values, calculation errors, and outliers. We determined the data were sufficiently reliable to use in our regressions.

## Regression Analysis Using National Assessment of Educational Progress Data

We used 2019 NAEP data for the fourth and eighth grade math and reading assessments to perform a regression analysis. We conducted this analysis to account for student-, teacher-, and school-level characteristics and policies that may be associated with the academic performance of English learners on national reading and math assessments. This type of analysis allows us to more accurately speak to characteristics that are significantly associated with differences in test scores.

For this analysis we limited our population to current or former English learners and excluded students attending private schools. Because the NAEP survey is not designed to generalize to the larger population of English learners, the results of our analysis are associational in nature. Some observations were discarded from the analysis due to having missing reading or math test score data. For the fourth grade analysis, 922 out of 19,539 observations were missing math test score data, and 1,166 out of 19,692 were missing data on reading scores. For the eighth grade analysis, 826 out of 12,699 observations were missing math test scores, and 981 out of 12,352 were missing reading scores. Some of the values for characteristics we controlled for in our model were missing or omitted due to item nonresponse or survey skip patterns. We treated item nonresponse as a valid categorical level in our regression model. We did this to leverage as much data as possible, while controlling for potential bias introduced into the analysis.

To better understand which characteristics are associated with academic achievement of English learners, we reviewed studies that were relevant to our research questions in addition to prior GAO work. We created composite variables and recoded some variables from those we identified. This was done for several reasons. First, variables were recoded so that the reference level in the regression is consistent across all variables. By doing this, the regression model results are easier to interpret. In other cases, we created composite variables to better measure certain concepts of interest. For example, we used proxy variables such as Title I funding and free or reduced-price lunch as a measure of a student's economic disadvantage in cases where that information was missing.<sup>2</sup> In other cases, categorical variables had levels collapsed together. This was done to increase the model degrees of freedom and to make the model more interpretable by having consistent variable levels. For example, combining categorical percent ranges into quartile ranges such as "1–25%," "26–50%," etc.

To select our final set of variables for use in the regression models, we checked for collinearity by examining polychoric correlations, since many of our identified variables were categorical and ordinal in nature. Additionally, we examined the distribution of identified variables for sparsity in cells. We then used backward and stepwise regression selection methods alongside manual iterative regressions to refine the variable set, starting with the fourth-grade data. Ideally, we would include random effects for teacher, school, and state. However, due to having hundreds of levels for school and thousands of levels for teacher, this was not computationally possible given available resources. In addition, English learners are a small population that is further reduced due to missing control and outcome variables. Therefore, we instead used a fixed effects regression model. However, we could not include fixed effects for teacher, school, and state for similar reasons—our model had too few degrees of freedom. As such, state,

---

<sup>2</sup> Title I, Part A (Title I) of the Elementary and Secondary Education Act of 1965, as amended, provides financial assistance to districts and schools with high numbers or high percentages of children from low-income families to help ensure that all children meet challenging state academic standards.

school, and classroom were not considered in our variable selection. See table 1 for our final variable selection.

The outcome of interest is student performance on national reading and math assessments. For NAEP reading and math tests, NCES computed twenty ‘plausible’ test scores to capture the distribution of possible test scores for each student. Because of this, the variance added through this process should be accounted for when calculating estimates using the NAEP survey data. Failing to account for these errors results in estimates which are less reliable than they appear to be. For the purposes of our analysis, we instead used the average of the twenty ‘plausible’ scores as our outcome of interest. This was done due to software limitations of incorporating the 20 individual plausible test scores into our analysis. NAEP Technical Documentation states that using averages of the twenty plausible values attached to a student’s file is inadequate to calculate group summary statistics such as proportions above a certain level or to determine whether group means differ from one another.<sup>3</sup> However, for the purposes of our analysis, we determined this approach to be sufficient for several reasons. First, estimates of statistics such as proportions and averages are not relevant to our analysis. Instead, we are interested in associations between test score performance and select control variables. Second, when using the average of the twenty plausible values as our outcome, the associated variance estimates are likely to be underestimated. This is because it ignores the additional error introduced by imputing test scores generated from a distribution of likely scores, rather than directly observing a test score for each student. While some research suggests the two variance estimates would be close, the variances estimated when using the average of the plausible values would not be the same. We also used an alternate variance estimation method – Taylor series linearization – which accounted for both the sample design and any additional variance added from having missing data. Because the variance estimated by averaging the plausible values are likely underestimated, using this approach would further minimize the bias in the variance.<sup>4</sup>

**Table 1: Outcome and Control Variables Used in the Regression Analysis Using Department of Education’s 2019 National Assessment of Education Progress for 4<sup>th</sup> and 8<sup>th</sup> Grades**

Outcomes	Control variables
<p>Math test score</p> <p>Reading test score</p>	<p><b>School characteristics:</b> Percent of students by race and ethnicity, percent of student receiving targeted Title I services,<sup>a</sup> type of school, school size, locale, parent volunteers available to assist English learners, reading specialists/literacy coaches available for English learners, certified English learner/bilingual education teachers, paraprofessionals/teacher aids trained to work with English learners, percent in grade held back and repeating</p>

<sup>3</sup>National Center for Education Statistics, “NAEP Technical Documentation: Plausible Values versus Individual Scores.” NAEP Analysis and Scaling – Plausible Values Versus Individual Scores, accessed June 27, 2024, last modified June 2, 2016, [https://nces.ed.gov/nationsreportcard/dw/analysis/est\\_pv\\_individual.aspx](https://nces.ed.gov/nationsreportcard/dw/analysis/est_pv_individual.aspx).

<sup>4</sup>Gregory J. Marchant, “How plausible is using averaged NAEP values to examine student achievement?,” *Comprehensive Psychology* 4, 1 (2015). <https://doi.org/10.2466/03.CP.4>

Outcomes	Control variables
	<p><b>Student characteristics:</b> Gender, race/ethnicity, type of English learner accommodation, limited English proficiency, lives with father/step father, lives with parental guardian/foster parent, lives with mother/step mother, days absent from school last month, reported sense of belonging and resiliency, economic disadvantage, length of time student has received instruction in English, English learner’s English proficiency (listen, speak, write, read), primary language, type of identified disability, enjoyment of complex problems</p>
	<p><b>Teacher characteristics:</b> Percent of teachers absent on an average day, number of years worked as elementary/secondary teacher, holds a valid regular/standard teaching certificate, satisfaction being a teacher at the school, gender, awarded tenure by school/district, student and teacher race/ethnicity match, joint teaching in the same class, reported severity of classroom overcrowding, class size, weekly time spent on subject instruction, teaching too many hours, enter teaching through alternative route/certification program, highest academic degree</p>

Source: GAO analysis of the Department of Education’s 2019 National Assessment of Educational Progress data. | GAO-24-107485

<sup>a</sup>Title I, Part A (Title I) of the Elementary and Secondary Education Act of 1965, as amended, provides financial assistance to districts and schools with high numbers or high percentages of children from low-income families to help ensure that all children meet challenging state academic standards. A targeted assistance program uses Title I funds to help improve the performance of students who are failing, or most at risk of failing, to meet the state’s academic achievement standards.

We used a multivariate linear regression model with full-sample weights. Rather than use the replicate student weights for variance estimation using the jackknife repeated replication method, we leveraged the sample design and weights to estimate variance using Taylor series linearization. We used this approach since including missing values without additional adjustment to variances implicitly assumes the missing values are missing completely at random. By reflecting the features of the sample design appropriately, we obtain approximately unbiased estimates of sampling variance. We assessed model fit by evaluating the R-square and adjusted R-square statistics, which is a measure of the variation in test scores the model explains. We also examined the root mean square error, which is a measure of how well the model fits the data (smaller is better). All models demonstrated reasonable fit by having an R-square of 0.5 or higher, and a root mean square error less than 25.

Results from our models are shown below in tables 2 and 3. Due to the limitations of our analysis, we report the relative margins of error instead of the 95% confidence intervals for each estimate. These are calculated at the 0.05 significance level. Variables denoted as “increase” are associated with an estimated increase in test scores for the subject matter at a 0.01 significance level; “decrease” means a particular variable was associated with an estimated decrease in test score at a 0.01 significance level. Variables which are associated but may not be precisely estimated due to having statistical significance between a 0.01 and 0.05 significance level are denoted by an asterisk “\*”. Factors which were not statistically significant at a 0.05 significance level are denoted by a “—”. Since the factors we controlled for are categorical in nature, we provided the comparison characteristic in the “Comparison group” column. For example, when looking at the results for the eighth grade NAEP analysis, we see that school where more than 10 percent of the teachers are absent on an average day are associated with decreases in reading score, relative to schools with less than five percent of teachers absent, holding other factors constant.

**Table 2: Results of Regression Analysis Using Department of Education’s 2019 National Assessment of Education Progress for 8<sup>th</sup> Grade**

Variable label	Effect (level of variable vs. comparison level)	8 <sup>th</sup> grade math		8 <sup>th</sup> grade reading	
		Direction of association	Absolute Relative Margin of Error	Direction of association	Absolute Relative Margin of Error
(Student) classified as current English learner	Not English learner vs. English learner	Increase	0.3	—	
(Student) English learner accommodations other than using non-English version of test	Yes vs. no	—		—	
(Student) English learner accommodations (non-English version of test used)	Yes vs. no	Increase*	1.0	Decrease*	0.8
(Student) lives with father or stepfather	Missing vs. no	—		Decrease*	0.9
	Yes vs. no	—		Decrease	0.8
(Student) lives with Mother or Stepmother	Missing vs. no	—		Increase	0.5
	Yes vs. no	—		Increase*	1.0
(Student) lives with parental guardian or foster parent	Missing vs. no	Decrease	0.6	Decrease	0.3
	Yes vs. no	Decrease	0.4	Decrease	0.3
(Student) gender	Female vs. male	Increase	0.3	Decrease	0.4
(Student) race and ethnicity, compared to White non-Hispanic or unknown ethnicity	American Indian or Alaska Native non-Hispanic or unknown ethnicity	—		—	
	Asian non-Hispanic or unknown ethnicity	Increase	0.4	Increase	0.4
	Black non-Hispanic or unknown ethnicity	—		Decrease	0.4
	Asian, American Indian, Alaska Native, Native Hawaiian, Pacific Islander, two or more races, or unknown race	—		—	
	Hispanic				
	Black-Hispanic	Decrease	0.7	Decrease*	0.8
	White-Hispanic	—		—	
	Native Hawaiian or Pacific Islander non-Hispanic or unknown ethnicity	—		—	

Variable label	Effect (level of variable vs. comparison level)	8 <sup>th</sup> grade math		8 <sup>th</sup> grade reading	
		Direction of association	Absolute Relative Margin of Error	Direction of association	Absolute Relative Margin of Error
	Two or more races, unknown race or two or more races unknown ethnicity non-Hispanic	—		—	
	Missing race and ethnicity information	—		—	
Student and teacher are of the same race ethnicity	Yes vs. no	—		Increase*	0.8
(Student) days absent from school last month	1–4 days vs. none	Decrease	0.6	Decrease	0.3
	5–10 days vs. none	Decrease	0.5	Decrease	0.3
	More than 10 days vs. none	Decrease	0.3	Decrease	0.2
	Missing vs. none	—		—	
(Student) I try hard after making mistakes	Somewhat/quite a bit like me vs. not at all/a little bit like me	Increase	0.8	Increase	0.5
	Very much like me vs. not at all/a little bit like me	Increase	0.6	Increase*	0.9
	Missing vs. not at all/a little bit like me	—		—	
(Student) I like complex problems more than easy ones	Somewhat/ quite a bit like me vs. not at all/a little bit like me	Increase	0.3	Increase	0.2
	Very much like me vs. not at all/a little bit like me	Increase	0.5	Increase	0.2
	Missing vs. not at all/a little bit like me	Increase*	0.8	—	
(Student) I felt awkward and out of place at school	Less than half the time vs. never or hardly ever	—		Increase	0.5
	About half the time vs. never or hardly ever	Decrease	0.4	—	
	More than half the time/all the time vs. never or hardly ever	Decrease*	0.9	—	
	Missing vs. never or hardly ever	—		Decrease*	0.9
(Student) I felt happy at school	Less than half the time vs. never or hardly ever	Increase	0.6	—	



Variable label	Effect (level of variable vs. comparison level)	8 <sup>th</sup> grade math		8 <sup>th</sup> grade reading	
		Direction of association	Absolute Relative Margin of Error	Direction of association	Absolute Relative Margin of Error
	About half the time vs. never or hardly ever	Increase	0.4	Increase*	0.8
	More than half the time/all the time vs. never or hardly ever	Increase	0.3	Increase	0.6
	Missing vs. never or hardly ever	—		Increase	0.7
(Student) economically disadvantaged	Missing vs. no	Decrease	0.5	—	
	Yes vs. no	Decrease	0.5	Decrease	0.3
(Student) how long has English learner received instruction in English	1–3 years vs. none or less than one year	—		Decrease	0.8
	Over 3 years vs. none or less than one year	—		Decrease	0.6
	Omitted/don't know vs. none or less than one year	—		—	
	Missing vs. none or less than one year	—		—	
(Student) proficiency in listening in English	Intermediate vs. no proficiency	—		—	
	English learner Advanced vs. no proficiency	Increase	0.6	Increase	0.4
	Omitted/don't know vs. no proficiency	—		—	
	Missing vs. no proficiency	Increase	0.2	Increase	0.4
(Student) proficiency in speaking in English	Intermediate vs. no proficiency	Increase	0.2	Increase	0.3
	Advanced vs. no proficiency	Increase	0.2	Increase	0.2
	Omitted/don't know vs. no proficiency	—		—	
	Missing vs. no proficiency	—		—	
(Student) proficiency in reading in English	Intermediate vs. no proficiency	Increase	0.5	—	
	Advanced vs. no proficiency	Increase	0.4	—	
	Omitted/don't know vs. no proficiency	—		—	
	Missing vs. no proficiency	—		—	
(Student) proficiency in writing in English	Intermediate vs. no proficiency	Increase	0.5	Increase	0.7

Variable label	Effect (level of variable vs. comparison level)	8 <sup>th</sup> grade math		8 <sup>th</sup> grade reading	
		Direction of association	Absolute Relative Margin of Error	Direction of association	Absolute Relative Margin of Error
	Advanced vs. no proficiency	Increase*	1	Increase	0.5
	Omitted/don't know vs. no proficiency	—		—	
	Missing vs. no proficiency	—		—	
(Student) English learner's primary language	Spanish vs. other or omitted	—		—	
	Missing vs. other or omitted	—		—	
(Student) identified disability	Specific learning disability	Decrease	0.2	Decrease	0.1
	Hearing impairment	Decrease*	0.8	—	
	Visual impairment	Decrease	0.2	—	
	Speech impairment	—		—	
	Emotional disturbance	—		Decrease	0.3
	Brain injury	—		—	
	Autism	Decrease	0.6	Decrease	0.8
	Developmental delay	Decrease	0.7	Decrease*	0.9
	Other health disability	Decrease	0.4	Decrease	0.5
	Intellectual disability	Decrease	0.2	Decrease	0.3
(Teacher) class size	16–20 students vs. 15 or fewer students	Increase*	0.9	—	
	21–30 students vs. 15 or fewer students	Increase	0.4	—	
	More than 30 students vs. 15 or fewer students	Increase	0.6	Increase*	1.0
	Missing vs. 15 or fewer students	—		—	
(Teacher) highest academic degree	Associates degree or vocational certificate vs. high school diploma	—		—	
	Bachelors degree vs. high school diploma	Decrease*	1	Increase*	0.9
	Education specialist degree vs. high school diploma	—		—	
	Masters, doctorate, or professional degree vs. high school diploma	—		Increase*	0.9

Variable label	Effect (level of variable vs. comparison level)	8 <sup>th</sup> grade math		8 <sup>th</sup> grade reading	
		Direction of association	Absolute Relative Margin of Error	Direction of association	Absolute Relative Margin of Error
	Missing vs. high school diploma	Decrease	0.7	—	
(Teacher) years worked as elementary/secondary teacher	1–2 years vs. less than 1 year	Decrease	0.8	Increase*	1.0
	3–5 years vs. less than 1 year	—		—	
	6–10 years vs. less than 1 year	—		—	
	11–20 years vs. less than 1 year	—		—	
	21 or more years vs. less than 1 year	—		—	
	Missing vs. less than 1 year	—		—	
(Teacher) hold valid regular/standard teaching certification	No, not planning vs. yes, permanent	—		—	
	No, working toward vs. yes, permanent	—		—	
	Yes, temporary vs. yes, permanent	—		—	
	Missing vs. yes, permanent	—		—	
(Teacher) entered teaching through alternative route/certificate program	Missing vs. no	—		Decrease*	0.9
	Yes vs. no	—		—	
(Teacher) teach jointly as a team in the same class	Once/twice a year vs. never	—		Decrease	0.6
	A few times a month vs. never	—		Decrease	0.6
	About every day vs. never	—		Decrease	0.6
	Missing vs. never	—		—	
(Teacher) how severe is classroom overcrowding	Minor problem vs. not a problem	—		—	
	Moderate problem vs. not a problem	Decrease*	0.9	Decrease	0.6
	Severe problem vs. not a problem	Decrease*	0.8	Decrease	0.7
	Missing vs. not a problem	—		—	
(Teacher) how severe of a problem is too many teaching hours	Minor problem vs. not a problem	—		—	
	Moderate problem vs. not a problem	—		—	

Variable label	Effect (level of variable vs. comparison level)	8 <sup>th</sup> grade math		8 <sup>th</sup> grade reading	
		Direction of association	Absolute Relative Margin of Error	Direction of association	Absolute Relative Margin of Error
	Severe problem vs. not a problem	—		—	
	Missing vs. not a problem	—		Decrease*	0.9
(Teacher) I'm satisfied being a teacher at this school	Somewhat/quite a bit like me vs. a little bit/not at all like me	—		—	
	Very much like me vs. a little bit/not at all like me	—		—	
	Missing vs. a little bit/not at all like me	—		Increase	0.6
(Teacher) gender	Female vs. male	—		—	
	Missing vs. male	—		—	
(Teacher) have you been awarded tenure by school or district?	No school or district tenure available vs. no	Increase*	0.9	—	
	Yes vs. no	—		—	
	Missing vs. no	Increase*	0.9	—	
(Teacher) time spent on math instruction in average week, in hours	5 to less than 7 vs. less than 5		b	Increase	0.7
	7 or more vs. less than 5		b	Increase*	0.9
	Missing vs. less than 5		b	—	
(Teacher) time spent on reading instruction in average week, in hours	5 to less than 10 vs. less than 5		—	b	
	10 or more vs. less than 5		—	b	
	Missing vs. less than 3		—	b	
(Teacher/School) percentage of teachers absent on an average day	6–10% vs. 0–5%		—	—	
	More than 10% vs. 0–5%	Decrease	0.5	—	
	Missing vs. 0–5%		—	—	
(School) percentage of students who are two or more races	1–25% vs. 0%	Increase*	0.9	—	
	26–50% vs. 0%		—	—	
	51% or more vs. 0%		—	—	
(School) percentage of Asian students	1–25% vs. 0%		—	—	
	26–50% vs. 0%		—	Increase	0.7
	51% or more vs. 0%		—	—	
(School) percentage of Black students	1–25% vs. 0%		—	—	
	26–50% vs. 0%		—	—	

Variable label	Effect (level of variable vs. comparison level)	8 <sup>th</sup> grade math		8 <sup>th</sup> grade reading	
		Direction of association	Absolute Relative Margin of Error	Direction of association	Absolute Relative Margin of Error
	51% or more vs. 0%	Increase	0.8	Decrease*	1.0
(School) percentage of Native Hawaiian or Pacific Islander students	1–25% vs. 0%	—		Decrease*	0.8
	26–50% vs. 0%	—		Decrease	0.4
	51% or more vs. 0%	—		Decrease	0.5
(School) percentage of Hispanic/Latino students	1–25% vs. 0%	—		—	
	26–50% vs. 0%	—		—	
	51% or more vs. 0%	—		—	
(School) percentage of American Indian or Alaska Native students	1–25% vs. 0%	—		Decrease*	0.9
	26–50% vs. 0%	—		—	
	51% or more vs. 0%	—		—	
(School) percentage of White students	1–25% vs. 0%	—		—	
	26–50% vs. 0%	—		—	
	51% or more vs. 0%	—		—	
(School) percentage of students receiving targeted Title I services <sup>a</sup>	1–25% vs. 0%	—		—	
	26–50% vs. 0%	—		—	
	51–75% vs. 0%	—		—	
	76% or above vs. 0%	—		—	
	Omitted vs. 0%	—		—	
(School) percentage of (School) 8th graders held back or repeating	1–2% vs. 0%	Decrease	0.7	—	
	3–5% vs. 0%	—		—	
	6% or more vs. 0%	—		—	
	Missing vs. 0%	Decrease	0.7	Increase	0.6
(School) percentage of students identified as limited English proficiency	1–25% vs. 0%	—		—	
	26–50% vs. 0%	—		—	
	51–75% vs. 0%	—		—	
	76% or more vs. 0%	—		—	
	Missing vs. 0%	—		—	
(School) type: Regular with magnet program	No vs. yes	—		—	
(School) type: Magnet or special emphasis	No vs. yes	Decrease*	1	—	
(School) type: Special education	No vs. yes	—		—	
(School) type: Alternative	No vs. yes	—		—	
(School) type: Independent Charter	No vs. yes	Decrease	0.6	Decrease	0.5

Variable label	Effect (level of variable vs. comparison level)	8 <sup>th</sup> grade math		8 <sup>th</sup> grade reading	
		Direction of association	Absolute Relative Margin of Error	Direction of association	Absolute Relative Margin of Error
(School) type: Charter administered by local district	No vs. yes	Decrease*	0.9	—	
(School) Teachers certified to teach English learners or bilingual ed teachers	Yes vs. no	—		b	
	Missing vs. no	—		b	
(School) Reading specialists/literacy coaches available to assist English learners:	Yes vs. no	—		b	
	Missing vs. no	—		b	
(School) Paraprofessionals or teacher aids trained in working with English learners	Yes vs. no	—		b	
	Missing vs. no	—		b	
(School) Parent volunteers available to assist English learners	Yes vs. no	—		b	
	Missing vs. no	—		b	
(School) size	Medium (577–781) vs. small (less than 576)	Increase*	0.9	Increase*	1.0
	Large (782–994) vs. small (less than 576)	Increase	0.7	Increase*	0.9
	Extra large (more than 994) vs. small (less than 576)	Increase	0.6	Increase	0.6
	Missing vs. small (less than 576)	Increase	0.7	—	
(School) locale	Rural vs. city	—		—	
	Suburb vs. city	—		—	
	Town vs. city	—		—	

Source: GAO analysis of the Department of Education’s 2019 National Assessment of Educational Progress data. | GAO-24-107485

Note: If the direction of association is listed as “—” then the association was not statistically different from zero.

\*While the variable is statistically significant between a 0.01 and 0.05 confidence level, it may be less precisely estimated.

<sup>a</sup>Title I, Part A (Title I) of the Elementary and Secondary Education Act of 1965, as amended, provides financial assistance to districts and schools with high numbers or high percentages of children from low-income families to help ensure that all children meet challenging state academic standards. A targeted assistance program uses Title I funds to help improve the performance of students who are failing, or most at risk of failing, to meet the state’s academic achievement standards.

<sup>b</sup>Variable not in subject-specific National Assessment of Educational Progress dataset.

**Table 3: Results of Regression Analysis Using Department of Education’s 2019 National Assessment of Education Progress for 4<sup>th</sup> Grade**

Variable label	Effect (level of variable vs. comparison level)	4 <sup>th</sup> grade reading		4 <sup>th</sup> grade math	
		Direction of association	Absolute relative margin of error	Direction of association	Absolute relative margin of error
(Student) classified as current English learner	Not English learner vs. English learner	—	—	—	—
(Student) English learner accommodations other than using non-English version of test	Yes vs. missing (no)	—	—	—	—
(Student) English learner accommodations (non-English version of test used)	Yes vs. missing (no)	—	—	—	—
(Student) lives with father or stepfather	Missing vs. no		*	Decrease	0.6
	Yes vs. no	Decrease	0.6	*	
(Student) lives with mother or stepmother	Missing vs. no	Increase	0.7	Increase	0.7
	Yes vs. no	—	—	—	—
(Student) lives with parental guardian or foster parent	Missing vs. no	Decrease	0.4	Decrease	0.3
	Yes vs. no	Decrease	0.2	Decrease	0.3
(Student) gender	Female vs. male	Increase	0.4	Decrease	0.2
(Student) race/ethnicity compared with White (non-Hispanic or unknown ethnicity)	American Indian or Alaska Native non-Hispanic or unknown ethnicity	—	—	—	—
	Asian non-Hispanic or unknown ethnicity	Increase	0.4	Increase	0.3
	Black non-Hispanic or unknown ethnicity	—	—	*	—
	Asian, American Indian Alaska Native, Native Hawaiian, Pacific Islander, two or more races, or unknown race	—	—	—	—
	Hispanic	—	—	—	—
	Black-Hispanic	Decrease	0.4	—	—
	White-Hispanic	—	—	—	—
	Missing race/ethnicity information	Increase	0.8	Increase	0.5
	Native Hawaiian or Pacific Islander (non-Hispanic or unknown ethnicity)	—	—	Decrease	0.4
	Two or more races or unknown race non-Hispanic or two or more races unknown ethnicity	—	—	Increase <sup>a</sup>	1.0

Variable label	Effect (level of variable vs. comparison level)	4 <sup>th</sup> grade reading		4 <sup>th</sup> grade math	
		Direction of association	Absolute relative margin of error	Direction of association	Absolute relative margin of error
Student and teacher are of the same race/ethnicity	Yes vs. no	Increase <sup>a</sup>	0.9	—	—
(Student) days absent from school last month	1–4 days vs. none	Decrease	0.3	Decrease	0.2
	5–10 days vs. none	Decrease	0.3	Decrease	0.2
	Missing vs. none	Decrease	0.7	Decrease	0.4
	More than 10 days vs. none	Decrease	0.2	Decrease	0.1
(Student) I try hard after making mistakes	Missing vs. not at all/a little bit like me	—	—	—	—
	Somewhat/quite a bit like me vs. not at all/a little bit like me	Increase	0.1	Increase	0.2
	Very much like me vs. not at all/a little bit like me	Increase	0.1	Increase	0.2
(Student) I like complex problems more than easy problems	A little bit like me vs. not at all like me	Increase <sup>a</sup>	0.8	Increase	0.4
	Quite a bit like me vs. not at all like me	Increase	0.3	Increase	0.2
	Somewhat like me vs. not at all like me	Increase	0.4	Increase	0.2
	Very much like me vs. not at all like me	—	—	Increase	0.2
	Missing vs. not at all like me	—	—	—	—
(Student) I felt awkward and out of place at school	Less than half the time vs. never or hardly ever	—	—	—	—
	About half the time vs. never or hardly ever	Decrease	0.5	—	—
	More than half the time vs. never or hardly ever	Decrease <sup>a</sup>	0.8	Decrease	0.6
	All or most of the time vs. never or hardly ever	Decrease	0.2	Decrease	0.3
	Missing vs. never or hardly ever	—	—	Increase <sup>a</sup>	0.8
(Student) I felt happy at school	Less than half the time vs. never or hardly ever	—	—	—	—
	About half the time vs. never or hardly ever	Increase	0.5	Increase	0.5
	More than half the time/all the time vs. never or hardly ever	Increase	0.4	Increase	0.4



Variable label	4 <sup>th</sup> grade reading			4 <sup>th</sup> grade math	
	Effect (level of variable vs. comparison level)	Direction of association	Absolute relative margin of error	Direction of association	Absolute relative margin of error
	Missing vs. never or hardly ever		—		—
(Student) economically disadvantaged	Missing vs. no	Decrease	0.7		—
	Yes vs. no	Decrease	0.2	Decrease	0.2
(Student) how often student gets help/tutoring with math after school	Every day or almost vs. never		n/a	Decrease	0.3
	Missing vs. never		n/a	Decrease	0.7
	Once/twice a month vs. never		n/a	Decrease	0.7
	Once/twice a week vs. never		n/a	Decrease	0.7
	Once/twice a year vs. never		n/a	Decrease	0.5
(Student) how long has student been receiving instruction in English	1–3 years vs. none or less than 1 year		—		—
	3+ years vs. none or less than 1 year	Increase	0.6		—
	Missing vs. none or less than 1 year	Increase	0.3	Increase	0.7
	Omitted/don't know vs. none or less than 1 year		—		—
(Student) English proficiency: listening comprehension	Advanced vs. none or beginner	Increase	0.4	Increase	0.2
	Intermediate vs. none or beginner	Increase	0.6	Increase	0.4
	Missing vs. none or beginner		—		—
	Omitted/don't know vs. none or beginner		—		—
(Student) English proficiency: speaking	Advanced vs. none or beginner	Increase	0.1	Increase	0.1
	Intermediate vs. none or beginner	Increase	0.2	Increase	0.1
	Missing vs. none or beginner		—		—
	Omitted/don't know vs. none or beginner	Increase	0.5		—
(Student) English proficiency: reading	Advanced vs. none or beginner	Increase	0.7		—
	Intermediate vs. none or beginner	Increase	0.7	Increase <sup>a</sup>	0.8
	Missing vs. none or beginner		—		—
	Omitted/don't know vs. none or beginner		—		—

Variable label	4 <sup>th</sup> grade reading			4 <sup>th</sup> grade math	
	Effect (level of variable vs. comparison level)	Direction of association	Absolute relative margin of error	Direction of association	Absolute relative margin of error
(Student) English proficiency: writing	Advanced vs. none or beginner	Increase	0.3	Increase	0.2
	Intermediate vs. none or beginner	Increase	0.3	Increase	0.4
	Missing vs. none or beginner	—	—	—	—
	Omitted/don't know vs. none or beginner	—	—	—	—
(Student) English learner's primary language	Missing vs. other/omitted	—	—	—	—
	Spanish vs. other/omitted	—	—	Decrease	0.5
(Student) identified disability	Specific learning	Decrease	0.1	Decrease	0.1
	Hearing impairment	—	—	—	—
	Visual impairment	Decrease	0.6	Decrease	0.2
	Speech impairment	Decrease	0.4	Decrease	0.4
	Emotional disturbance	Decrease <sup>a</sup>	0.9	—	—
	Brain injury	Decrease	0.5	Decrease	0.3
	Autism	Decrease	0.5	Decrease	0.6
	Developmental delay	Decrease	0.7	Decrease	0.5
	Other health	Decrease	0.5	Decrease	0.3
	Intellectual disability	Decrease	0.5	Decrease	0.2
(Teacher) highest academic degree	Bachelors degree vs. high school diploma	—	—	—	—
	Educational specialist vs. high school diploma	—	—	—	—
	Master's, doctoral, or professional degree vs. high school diploma	—	—	—	—
	Missing vs. high school diploma	—	*	—	*
	Some college/certificate vs. high school diploma	—	—	—	—
(Teacher) excluding student teaching, number of years worked as elementary/secondary teacher	1–2 years vs. less than 1 year	—	—	—	—
	11–20 years vs. less than 1 year	—	—	—	—
	21 or more years vs. less than 1 year	—	—	—	—
	3–5 years vs. less than 1 year	—	—	—	—

Variable label	4 <sup>th</sup> grade reading			4 <sup>th</sup> grade math	
	Effect (level of variable vs. comparison level)	Direction of association	Absolute relative margin of error	Direction of association	Absolute relative margin of error
	6–10 years vs. less than 1 year	—		—	
	Missing vs. less than 1 year	—		—	
(Teacher) hold valid regular/standard teaching certificate	Missing vs. yes, permanent	—		—	
	No, not planning vs. yes, permanent	—		Decrease <sup>a</sup>	0.9
	No, working toward vs. yes, permanent	—		—	
	Yes, temporary vs. yes, permanent	—		—	
(Teacher) entered teaching through alternative route	No vs. yes	—		—	
	Missing vs. yes	—		—	
(Teacher) teach jointly as a team in the same class	A few times a month vs. never	—		—	
	About every day vs. never	Decrease <sup>a</sup>	1.0	—	
	Missing vs. never	—		—	
	Once/twice a year vs. never	—		—	
(Teacher) how many students in class	16–20 vs. 15 or fewer	—		Increase <sup>a</sup>	0.9
	21–30 vs. 15 or fewer	Increase	0.7	—	
	More than 30 vs. 15 or fewer	Increase	0.5	Increase	0.8
	Missing vs. 15 or fewer	—		—	
(Teacher) how severe is classroom overcrowding	Minor problem vs. not a problem	Decrease <sup>a</sup>	0.9	—	
	Missing vs. not a problem	—		—	
	Moderate problem vs. not a problem	Decrease <sup>a</sup>	0.8	—	
	Severe problem vs. not a problem	Decrease	0.5	Decrease <sup>a</sup>	0.8
(Teacher) how severe a problem is too many teaching hours	Minor problem vs. not a problem	—		—	
	Moderate problem vs. not a problem	—		—	
	Serious problem vs. not a problem	—		—	
	Missing vs. not a problem	—		—	*
(Teacher) I'm satisfied being a teacher at this	Missing vs. not at all/a little bit like me	—		—	

Variable label	4 <sup>th</sup> grade reading			4 <sup>th</sup> grade math	
	Effect (level of variable vs. comparison level)	Direction of association	Absolute relative margin of error	Direction of association	Absolute relative margin of error
school	Somewhat/quite a bit like me vs. not at all/a little bit like me	Increase <sup>a</sup>	0.8	—	—
	Very much like me vs. not at all/a little bit like me	Increase <sup>a</sup>	0.9	—	—
(Teacher) gender	Female vs. male	—	—	—	—
	Missing vs. male	—	—	—	—
(Teacher) have you been awarded tenure by school/district?	Missing vs. no	Increase <sup>a</sup>	0.9	Increase <sup>a</sup>	0.9
	No school/district tenure vs. no	—	—	—	—
	Yes vs. no	Increase	0.3	—	—
(Teacher) time spent on English/Language Arts instruction in typical week	10 to less than 15 hours vs. 0 to less than 5 hours	—	—	—	n/a
	5 to less than 10 hours vs. 0 to less than 5 hours	—	—	—	n/a
	Missing vs. 0 to less than 5 hours	Increase <sup>a</sup>	1.0	n/a	—
(Teachers/school) percent of teachers absent on average day	3–5% vs. 0–2%	—	—	—	*
	6–10% vs. 0–2%	—	—	Decrease <sup>a</sup>	0.9
	Missing vs. 0–2%	—	—	*	—
	More than 10% vs. 0–2%	Decrease	0.6	—	—
(School) percentage of students of two or more races	1–25% vs. 0%	—	—	—	—
	26–50% vs. 0%	Increase	1.0	—	*
(School) percentage of Asian students	1–25% vs. 0%	—	—	—	—
	26–50% vs. 0%	—	—	—	—
	51% or more vs. 0%	—	—	Increase <sup>a</sup>	0.8
(School) percentage of Black students	1–25% vs. 0%	—	—	—	—
	26–50% vs. 0%	—	—	—	—
	51% or more vs. 0%	—	—	—	—
(School) percentage of Native Hawaiian or Pacific Islander students	1–25% vs. 0%	—	—	—	—
	26–50% vs. 0%	—	—	Decrease <sup>a</sup>	0.8
	51% or more vs. 0%	Decrease	0.7	—	—
(School) percentage of Hispanic students	1–25% vs. 0%	Increase	0.7	—	—
	26–50% vs. 0%	—	—	—	—
	51% or more vs. 0%	Increase <sup>a</sup>	0.8	—	—
(School) percentage of American Indian or Alaska Native students	1–25% vs. 0%	Decrease	0.7	—	—
	26–50% vs. 0%	—	—	—	—
	51% or more vs. 0%	—	—	—	—
(School) percentage of	1–25% vs. 0%	*	—	Increase <sup>a</sup>	0.8

Variable label	4 <sup>th</sup> grade reading			4 <sup>th</sup> grade math	
	Effect (level of variable vs. comparison level)	Direction of association	Absolute relative margin of error	Direction of association	Absolute relative margin of error
White students	26–50% vs. 0%	Increase	0.7	Increase	0.5
	51% or more vs. 0%	*		Increase	
(School) percent of enrollment identified as limited English proficient	1–25% vs. 0%	—		—	
	26–50% vs. 0%	—		—	
	51–75% vs. 0%	—		—	
	76% or above vs. 0%	—		—	
	Missing vs. 0%	—		—	
(School) percent of students at school receiving targeted Title I services <sup>b</sup>	1–25% vs. 0%	—		*	
	26–50% vs. 0%	—		Decrease	0.6
	51–75% vs. 0%	—		—	
	76% or above vs. 0%	—		*	
	Missing vs. 0%	—		Decrease	0.7
(School) percent of 4th graders held back and repeating	3–5% vs. 0–2%	—		—	
	6–10% vs. 0–2%	—		Increase	0.6
	Missing vs. 0–2%	—		—	
	More than 10% vs. 0–2%	—		—	
(School) type: Regular with magnet program	Missing (no) vs. yes	—		—	
(School) type: Magnet or special emphasis	Missing (no) vs. yes	—		Increase <sup>a</sup>	1.0
(School) type: Special education	Missing (no) vs. yes	—		—	
(School) type: Alternative	Missing (no) vs. yes	—		—	
(School) type: Independent charter	Missing (no) vs. yes	—		—	
(School) type: Charter administered by local district	Missing (no) vs. yes	—		—	
	Missing vs. yes	—		c	
(School) teachers certified to work with English learners or bilingual education teachers	Missing vs. yes	—		c	
	No vs. yes	*		c	
(School) Reading specialists or literacy coaches to work with English learners	Missing vs. yes	Decrease <sup>a</sup>	0.9	c	
	No vs. yes	—		c	
(School) paraprofessional teacher aides trained to work with English learners	Missing vs. yes	Increase	0.7	c	
	No vs. yes	—		c	
(School) parent volunteers to assist English learners	Missing vs. yes	Decrease	0.5	c	
	No vs. yes	*		c	

Variable label	Effect (level of variable vs. comparison level)	4 <sup>th</sup> grade reading		4 <sup>th</sup> grade math	
		Direction of association	Absolute relative margin of error	Direction of association	Absolute relative margin of error
(School) size (number of students)	Extra large (more than 994) vs. small (less than 576)	—		Increase <sup>a</sup>	0.8
	Large (782–994) vs. small (less than 576)	—		—	
	Medium (577–781) vs. small (less than 576)	—		—	
	Missing vs. small (less than 576)	—		—	
	Zero vs. small (less than 576)	*		n/a	
(School) location	Rural vs. city	—		—	
	Suburb vs. city	—		—	
	Town vs. city	—		—	

Source: GAO analysis of the Department of Education’s 2019 National Assessment of Educational Progress data. | GAO-24-107485

Note: If the direction of association is listed as “—” then the association was not statistically different from zero. If the direction of association is listed as “\*” the p-value for the variable category would indicate a significant association with the reading or math score, but we have chosen to disregard this association because of the presence of multi-collinearity that affects the significance of this variable.

<sup>a</sup> While the variable is statistically significant between a 0.01 and 0.05 confidence level, it may be less precisely estimated.

<sup>b</sup> Title I, Part A (Title I) of the Elementary and Secondary Education Act of 1965, as amended, provides financial assistance to districts and schools with high numbers or high percentages of children from low-income families to help ensure that all children meet challenging state academic standards. A targeted assistance program uses Title I funds to help improve the performance of students who are failing, or most at risk of failing, to meet the state’s academic achievement standards.

<sup>c</sup> Variable not in subject-specific National Assessment of Educational Progress dataset.

## Regression Analysis of English Learner Progress in Early Childhood Longitudinal Study, Kindergarten Class of 2010-2011

We used the ECLS Kindergarten Class of 2010-2011 data to perform a regression analysis analyzing which school, student-, and teacher-level characteristics and policies may be associated with growth in reading scores from second grade to fifth grade.<sup>5</sup>

<sup>5</sup> In the ECLS-K:2011 changes in reading score indicate changes in reading skills and reading abilities. For example, a score raising a score from 50 to 90 involves changes in skills and abilities that are different than those demonstrated for a change between 90 and 120.

The ECLS Kindergarten Class of 2010-2011 is a longitudinal survey consisting of nine rounds of data collection that follows students from kindergarten through fifth grade. Data collection included the fall and spring of kindergarten, first, and second grade. For third through fifth grade, data were collected in the spring only. Our analysis uses second grade and fifth grade reading scores to measure growth in reading ability because second grade was the first year the entire standard ECLS reading assessment was administered to all English learners. In our analysis we identified students as potential English learners using kindergarten and 1st grade data. We limited our analysis to students who did not switch schools during 6 school years of data collection.<sup>6</sup> However, we controlled for the difference between students who move schools and do not move schools by including variables in our model that we found to be significant when predicting if a student switched schools over the nine rounds of data collection. We also did not include students who did not have all six spring rounds of data collected. Additionally, we limited our analysis to students who scored below average on their second grade reading assessment, to avoid clouding our regression results with second graders who did not grow substantially from second grade to fifth grade in reading score due to high reading ability in second grade. The subset of respondents included in our model consisted of 621 students.

Some characteristics used in our analysis were missing or omitted for some observations. As the number of observations available was quite low, we wanted to leverage as much data as possible while controlling for such item nonresponse. As a result, we treated these missing responses as a valid categorical level in our regression model. Additionally, we treated our variance computation as not completely missing at random by analyzing the nonmissing values as a domain (subpopulation) when computing variance. The specific characteristics we controlled for in our model are shown in table 4.

**Table 4: Variables Included in Regression Model Using the Department of Education’s Early Childhood Longitudinal Study, Kindergarten Class of 2010-2011**

**Independent variables**

Growth in reading scores

**Dependent variables**

***Student characteristics:***

Born in the United States, classmates make me feel happy, disabled at some point, likes to come to school, language spoken at home, type of parents in household, gender, absences, race and ethnicity

***Teacher/classroom characteristics:***

Average number of years of experience, disagrees that the inclusion of English learners has gone well, agrees that they can teach English learners in class

***School characteristics:***

School type: Charter school

School location: Urbanicity

Staffing: Issues with teachers absent, percent of full-time or part-time teachers who teach English as a second language, bilingual education, language immersion, or English learner instruction, full-time/part-time teachers who teach English as a second language, bilingual education, language immersion, or English learner instruction,

<sup>6</sup>A student’s school could not be allowed to change because several variables are coded using school characteristics across multiple school years.

student-teacher-ratio, full-time student-teacher-ratio

Student body: Proportion of students with free or reduced-price lunch, percent receiving dual language instruction in pullout, average percent receiving dual language instruction in regular class, average percent receiving English language instruction, average percent receiving English as a second language pullout instruction, percent White, percent American Indian or Alaska Native, percent Asian, percent Black, percent Native Hawaiian or Pacific Islander, percent Hispanic

Other: Issues with overcrowding, translators available for parents, special meetings with English learner families

Source: GAO analysis of the Department of Education’s Early Childhood Longitudinal Study, Kindergarten Class of 2010-2011. | GAO-24-107485

Using the ECLS codebook, we selected a list of essential variables to include in our model unless there were concerns of multi-collinearity. The selection of variables was informed by our literature review of studies related to English learners. Through a series of iterative regressions, we then considered several nonessential education and demographic variables for our model that were also found to be present across the studies reviewed. Of these nonessential variables, we only kept the variables in the model that were found to be significant. We checked for collinearity among the variables considered for the model using polychoric and numeric correlations. When we found multi-collinearity, we removed those variables to reduce collinearity in our analysis.

We created composite or recoded variables to better measure certain concepts of interest. In some cases these variables were from multiple periods of data collection. (See table 5.)

**Table 5: Description of Variables in Regression Model Using the Department of Education’s Early Childhood Longitudinal Study, Kindergarten Class of 2010-2011**

Variable	Grade collected and additional description (when needed)
(Student) growth in reading scores	Difference between 5 <sup>th</sup> grade reading score and 2 <sup>nd</sup> grade reading score
(Student) born in the United States	Responded born in the United States at least once any point between kindergarten (K) and 5 <sup>th</sup> grade
(Student) sex	Grade collected not applicable
(Student) classmates make me feel happy	5 <sup>th</sup> grade
(Student) disabled at some point	Reported having a disability sometime between kindergarten and 5 <sup>th</sup> grade
(Student) likes to come to school	5 <sup>th</sup> grade
(Student) language spoken at home	In 5 <sup>th</sup> grade, English only, Spanish, or other non-English language spoken at home
(Student) type of parents in household	5 <sup>th</sup> grade
(Student) absences	From K-5 <sup>th</sup> grades, the number of years students had 5 or fewer absences or more than 5 absences
(Student) race/ethnicity	Grade collected not applicable. Race and ethnicity, including unknown race and unknown Hispanic status
(School) charter school	Kindergarten
(School) urbanicity of location	5 <sup>th</sup> grade
(School) issues with teachers absent	5 <sup>th</sup> grade
(School) percentage of full-time or part-time teachers who teach English as a second language, bilingual education, language immersion, or English learner instruction	1 <sup>st</sup> grade
(School) full-time or part-time teacher who teach	1 <sup>st</sup> grade: number of students in school divided by number of



<b>Variable</b>	<b>Grade collected and additional description (when needed)</b>
English as a second language, bilingual education, language immersion, or English learner instruction student-teacher-ratio	full-time and part-time teachers who teach English as a second language, bilingual education, language immersion, or English learner instruction
(School) full-time student-teacher-ratio	1 <sup>st</sup> grade: number of students in school divided by number of full-time teachers
(School) proportion of students with free or reduced-price lunch	1 <sup>st</sup> grade
(School) percent of students receiving dual language instruction in pullout	Kindergarten
(School) average percent of students receiving dual language instruction in regular class	Averaging kindergarten and 1 <sup>st</sup> grade
(School) average percent of students receiving English as a second language instruction in regular classroom	Averaging kindergarten and 1 <sup>st</sup> grade
(School) average percent receiving English as a second language instruction in a pullout model	Averaging kindergarten and 1 <sup>st</sup> grade
(School) percent White	5 <sup>th</sup> grade
(School) percent American Indian or Alaska Native	5 <sup>th</sup> grade
(School) percent Asian	5 <sup>th</sup> grade
(School) percent Black	5 <sup>th</sup> grade
(School) percent Native Hawaiian or Pacific Islander	5 <sup>th</sup> grade
(School) percent Hispanic	5 <sup>th</sup> grade
(School) issues with overcrowding	5 <sup>th</sup> grade
(School) translators available to parents	5 <sup>th</sup> grade
(School) special meetings with English learner families	Special meetings reported at least once, kindergarten through 3 <sup>rd</sup> grade
(Teacher) average number of years of teaching experience	Averaging kindergarten through 5 <sup>th</sup> grade
(Teacher) disagrees that the inclusion of English learners has gone well	Teacher disagrees/strongly disagrees that inclusion of English learners has gone well at least once, kindergarten through 2 <sup>nd</sup> grade
(Teacher) agrees that they can teach English learners in class	Teacher agrees/strongly agrees that they are adequately trained to teach English learners at least once, kindergarten through 2 <sup>nd</sup> grade

Source: GAO analysis of the Department of Education’s Early Childhood Longitudinal Study, Kindergarten Class of 2010-2011. | GAO-24-107485

The ECLS has several possible weights that analysts can choose from. Users of the data may choose from a matrix of fifth grade weights to control for the rounds of the variables used and whether the variables in the model are from the child, parent, and teacher questionnaires. We chose our weight considering the rounds of the data we used, the questionnaires of variables in our model, and the impact that the weights had on our number of observations available for regression analysis. Since our analysis uses a specific subset of ECLS respondents, the impact that our chosen weight had on the number of observations in our model was regarded as most important. We compared several of the weights and determined the W9C19P\_2T290 weight best fit our requirements. This weight is described in the manual as:

Child base weight adjusted for nonresponse associated with child assessment/child questionnaire data from both kindergarten rounds and spring fifth grade, as well as parent data from fall kindergarten or spring kindergarten, and either teacher-/classroom- or child-level teacher data from spring kindergarten (from a core or supplemental teacher questionnaire), spring first grade (from a first-grade or a kindergarten teacher questionnaire), spring second grade, and spring third grade, and either reading teacher-

/classroom- or child-level reading teacher data from spring fourth grade and spring fifth grade.<sup>7</sup>

We used a multivariate linear regression model with full-sample weights. The regression allows us to identify factors that were associated with students’ growth/decline in reading scores. We leveraged the sample design and weights to estimate variance using Taylor series linearization. By reflecting the features of the sample design appropriately, we obtain approximately unbiased estimates of sampling variance. Our approach may not reflect all students who are English learners. For example, the selection of potential English learners used in our analysis was based on kindergarten responses to questions and this approach does not consider English learners who may enter public schooling after kindergarten. Additionally, our approach does not address the issue of large changes between kindergarten and fifth grade respondents who speak a non-English language at home found in the ECLS nonresponse bias analysis. Furthermore, the variables in our regression analysis consisted of mostly variables found to be essential for our analysis of English learners. There were numerous other nonessential variables we identified that could have been considered for a model analyzing growth of English learners’ reading scores. However, due to the limited number of observations (621) available for our analysis, we limited the number of variables in our model.

Results from our model are shown in table 6. Variables denoted as “increase” are statistically significantly associated with an estimated increase in the outcome at the 0.1 significance level; “decrease” means a particular variable was statistically significantly associated with an estimated decrease in the outcome at the 0.1 significance level. The estimated change in the outcome is within the range of the lower and upper bound of the 90% confidence interval for the ECLS analysis. Factors that were not statistically significant are denoted by “—”. Since the factors we controlled for are categorical in nature, we provided the comparison characteristic in the “Comparison group” column.

**Table 6: Associations of Regression Model Variables with Reading Score Growth from Early Childhood Longitudinal Study, Kindergarten Class of 2010-2011**

Variable label	Effect (level of variable vs. comparison level)	Direction of association	90% confidence interval	
			Lower bound	Upper bound
(Student) sex	Female vs. male	Decrease	-3.76	-0.36
(Student) race/ethnicity compared to White non-Hispanic or unknown Hispanic status	American Indian or Alaska Native non-Hispanic or unknown Hispanic status	*		
	Asian non-Hispanic or unknown Hispanic status	—		
	Black non-Hispanic or unknown Hispanic status	—		
	Native Hawaiian or Pacific Islander non-Hispanic or unknown Hispanic status	*		
	Black-Hispanic	—		
	White-Hispanic	—		

<sup>7</sup>Department of Education, Institute of Education Sciences, National Center for Education Statistics, *User’s Manual for the ECLS-K:2011 Kindergarten–Fifth Grade, Data File and Electronic Codebook*, NCES 2019-051 (Washington, D.C.: July 2019).

Variable label	Effect (level of variable vs. comparison level)	Direction of association	90% confidence interval	
			Lower bound	Upper bound
	Asian-Hispanic, American Indian or Alaska Native-Hispanic, Native Hawaiian or Pacific Islander-Hispanic, two or more races-Hispanic, unknown race-Hispanic	—		
	Two or more races non-Hispanic, unknown race non-Hispanic, two or more races unknown Hispanic status	—		
	Missing	—		
(Student) disabled at some point	Yes vs. no	—		
	Missing vs. no	—		
(Student) type of parents in household	One biological/adoptive parent and one other parent/partner vs. two biological/adoptive parents	—		
	One biological/adoptive parent only vs. two biological/adoptive parents	Decrease	-7.37	-1.44
	Other guardian(s) vs. two biological/adoptive parents	Decrease	-6.28	-0.05
	Missing vs. two biological/adoptive parents	—		
(Student) born in the United States	No vs. yes	—		
	Missing vs. yes	—		
(Student) language spoken at home	Non-English language used at home including Spanish vs. only English used at home	—		
	Non-English language used at home not including Spanish vs. only English used at home	Decrease	-9.57	-0.63
	Missing vs. only English used at home	—		
(Student) classmates make me feel happy	Often/very often vs. never/rarely	—		
	Sometimes vs. never/rarely	—		
	Missing vs. never/rarely	—		
(Student) likes to come to school	Certainly applies (often displays this behavior) vs. doesn't apply (seldom displays this behavior)	—		
	Sometimes applies (occasionally displays this behavior) vs. doesn't apply (seldom displays this behavior)	—		
	Missing vs. doesn't apply (seldom displays this behavior)	—		
(Student) absences	Had more than five absences in most years from Kindergarten through 5 <sup>th</sup> grade (K-5) vs. had 5 or fewer absences in most years	—		
(Teacher) average number of years of teaching experience	not applicable	—		
(Teacher) disagrees that the inclusion of English learners has gone well	Yes vs. no	—		
	Missing vs. no	Increase	3.89	15.41

Variable label	Effect (level of variable vs. comparison level)	Direction of association	90% confidence interval	
			Lower bound	Upper bound
(Teacher) agrees that they can teach English learners in class	No vs. yes	—		
	Missing vs. yes	—		
(School) charter school	Yes vs. no	—		
	Missing vs. no	—		
(School) urbanicity of location	City vs. rural	Decrease	-13.46	-0.11
	Suburb vs. rural	Decrease	-15.4	-2.14
	Town vs. rural	Decrease	-17.5	-3.61
	Missing vs. rural	—		
(School) percent American Indian or Alaska Native	not applicable	—		
(School) percent Asian	not applicable	—		
(School) percent Black	not applicable	—		
(School) percent Native Hawaiian or Pacific Islander	not applicable	*		
(School) percent White	not applicable	Decrease	-0.37	0
(School) percent Hispanic	not applicable	Decrease	-0.33	-0.01
(School) proportion of students with free or reduced-price lunch	not applicable	—		
(School) full-time student-teacher ratio	not applicable	—		
(School) full-time English language teacher or part-time English language teacher student-teacher ratio	not applicable	—		
(School) percentage of full-time or part-time English language teachers	not applicable	Decrease	-12.2	-2.08
(School) issues with teachers absent	Serious problem vs. not a problem	Decrease	-13.85	-5.77
	Moderate problem vs. not a problem	—		
	Minor problem vs. not a problem	Increase	2.82	6.59
	Missing vs. not a problem	—		
School) issues with overcrowding	Serious problem vs. not a problem	Decrease	-22.12	-15.24
	Moderate problem vs. not a problem	Decrease	-7.08	-0.32
	Minor problem vs. not a problem	—		
	Missing vs. not a problem	—		
(School) translators available to parents	No vs. yes	—		
	Missing vs. yes	—		
(School) special meetings with English learners' families	No vs. yes	Decrease	-4.71	-0.74
	Missing vs. yes	—		

Variable label	Effect (level of variable vs. comparison level)	Direction of association	90% confidence interval	
			Lower bound	Upper bound
(School) average percent of students receiving English as a Second Language instruction in regular classroom	not applicable	—		
(School) average percent of students receiving English as a Second Language pullout instruction	not applicable	—		
(School) average percent of students receiving dual language instruction in regular class	not applicable	Increase	0.03	0.13
(School) percent of students receiving dual language instruction in pullout	not applicable	Decrease	-0.37	-0.04

Source: GAO analysis of the Department of Education’s Early Childhood Longitudinal Study, kindergarten class of 2010-2011. | GAO-24-107485

Note: If the direction of association is listed as “—” then the association was not statistically different from zero.

\*Results are not presented because operational problems prevented the ECLS from collecting data in some areas of the country where Native Hawaiian, Pacific Islander, American Indian, or Alaska Native sampled students resided.

**Table 7: Variables Considered for the Regression Using the Early Childhood Longitudinal Study, Kindergarten Class of 2010-2011, but Not Included Due to Nonsignificance**

Issues with teacher turnover in school

Did school receive Title I funding at some point<sup>a</sup>

Did any of the students’ teachers in kindergarten have training in teaching English learners

Average hours aide trained to work with English learners works with children

Source: GAO analysis of the Department of Education’s Early Childhood Longitudinal Study, Kindergarten Class of 2010-2011. | GAO-24-107485

Notes: These variables were not included in our model because they were considered nonessential for our analysis and were found to be not significant.

<sup>a</sup>Title I, Part A (Title I) of the Elementary and Secondary Education Act of 1965, as amended, provides financial assistance to districts and schools with high numbers or high percentages of children from low-income families to help ensure that all children meet challenging state academic standards.

## Regression Analysis of ED Facts Data on English Learner Progress toward Proficiency

To analyze the relationship between student outcomes and a variety of school and district characteristics, we conducted regression analysis of data from Department of Education’s ED Facts for school years 2017-2018, 2018-2019, and 2020-2021.<sup>8</sup> The ED Facts data set covers all K-12 public schools and includes data on English learners’ performance on the state

<sup>8</sup>We excluded the 2019-2020 school year because of issues with data collection during COVID-19.

English-language proficiency (ELP) assessment. This data includes the number of students who tested as English proficient as well as the number making progress toward English proficiency based on the ELP assessment. We assessed the reliability of the data by reviewing technical documentation and Education reports on the survey, and testing for check for missing values, calculation errors, and outliers. We determined the data were sufficiently reliable to use in our regression.

There are many characteristics that may be related to a school’s proportion of proficient English learners. We based our independent variable selections on existing literature and variables with a logical relationship. In addition to our preferred model specification, we created other specifications to understand whether the results we report were robust to different specifications. We generally found that direction and statistical significance of our reported results were consistent across specifications. This includes specifications where our model controlled for different states to limit the effect of different testing standards across states on our results.

To determine whether a school was in an urban or rural location, we combined the *EDFacts* data with data from the Common Core of Data and U.S. Department of Agriculture data on urban-rural locations by zip code. We define urban locations as zip codes with the “metropolitan” designation, suburban locations as zip codes with a “micropolitan” designation, towns as zip codes with a “small town” designation, and rural locations as zip codes with a “rural areas” designation.

To determine whether the school was located in a state that bordered Canada, Mexico, or a coast, we combined the *EDFacts* school state data with data from a 2006 Congressional Research Service report on the United States’ international borders.

Lastly, we used data from the WIDA consortium website on which states participated in the WIDA consortium.<sup>9</sup> From this we created a variable that equaled one if the school was located in a state that participated in the consortium.

The fractional logistic regression results from our preferred models are shown in table 8. Variables denoted as “increase” are statistically significantly associated with an estimated increase in the outcome at the 0.05 significance level; “decrease” means a particular variable was statistically significantly associated with an estimated decrease in the outcome at the 0.05 level. The estimated change in the outcome is within the range of the lower and upper bound of the 95% confidence level. Factors that were not statistically significant are denoted by “—”. Since the factors we controlled for are categorical in nature, we provided the comparison characteristic in the “Comparison group” column.

**Table 8: Associations of Regression Model Variables with Reading Scores and Progress toward English Proficiency in the Department of Education’s *EDFacts*, School Years 2017-18, 2018-19, and 2020-21**

Variable label	Effect (level of variable vs. comparison level)	Proportion of English learners proficient based on the English-language proficiency assessment		Proportion of English learners making progress on the English language proficiency assessment			
		Direction of association	95% confidence interval		Direction of association	95% confidence interval	
			Lower bound	Upper bound		Lower bound	Upper bound

<sup>9</sup>The WIDA Consortium is made up of 42 states, territories and federal agencies and designs the standards and assessments for English learners among other instructional support.

Variable label	Effect (level of variable vs. comparison level)	Proportion of English learners proficient based on the English-language proficiency assessment			Proportion of English learners making progress on the English language proficiency assessment		
		Direction of association	95% confidence interval		Direction of association	95% confidence interval	
			Lower bound	Upper bound		Lower bound	Upper bound
School level	Elementary vs. high school	Positive*	-0.0008	0.0644	Positive	0.0382	0.11637
	Middle vs. high school	—			Negative	-0.0672	-0.0079
Charter school	Yes vs. no	—			—		
School locale	Urban vs. town	—			—		
	Suburban vs. town	—			Negative	-0.0483	-0.0143
	Rural vs. town	—			Positive	0.0032	0.0542
School's demographic composition compared to percent of White students	Percentage of non-Hispanic Asian students	Positive	0.0048	0.0085	Positive	0.0015	0.0050
	Percentage of non-Hispanic Black students	—			Positive*	-0.0001	0.0022
	Percentage of Hispanic students	Positive	0.0011	0.0042	—		
	Percentage of non-Hispanic American Indian /Alaska Native students	—			—		
	Percentage of non-Hispanic Native Hawaiian or Pacific Islander students	Negative	-0.0105	-0.0045	Negative	-0.0037	-0.0005
	Percentage of students with multiple races, ethnicities, or cultures	—			Negative	-0.0154	-0.00004
	Gender	Percentage female vs. percentage male	—			Positive	0.0013
Percentage of English learners	Percentage English vs. non-English learners	Positive	0.0003	0.0057	—		
Percentage of economically disadvantaged students	Percentage economically disadvantaged vs. not	—			—		
Percentage of students with disability	Percentage of students with disability vs. no disability	Negative	-0.0061	-0.0010	—		

Variable label	Effect (level of variable vs. comparison level)	Proportion of English learners proficient based on the English-language proficiency assessment			Proportion of English learners making progress on the English language proficiency assessment		
		Direction of association	95% confidence interval		Direction of association	95% confidence interval	
			Lower bound	Upper bound		Lower bound	Upper bound
Quartile for total number English learner students	Top quartile vs. bottom-middle	Negative	-0.2706	-0.1156	—		
	Second quartile vs. bottom-middle	Negative	-0.1122	-0.0562	Negative	-0.0375	-0.0041
	Bottom quartile vs. bottom-middle	Positive	0.0810	0.1916	Positive	0.0910	0.1509
Quartile of the number of total students	Top quartile vs. bottom-middle	Positive*	-0.0024	0.0818	—		
	Second quartile vs. bottom-middle	—			—		
	Bottom quartile vs. bottom-middle	Positive*	-0.0419	0.0027	—		
Year of assessment	2019 vs. 2018	—			—		
	2021 vs. 2018	Negative*	-0.2770	0.0184	—		
Border state	Coastal state	Positive	0.0003	0.1320	—		
	Borders Mexico	—			Negative	-0.2575	-0.0386
State participation in WIDA Consortium <sup>a</sup>	State participates	Negative	-0.1415	-0.0210	Negative	-0.1901	-0.0133

Source: GAO analysis of the Department of Education's ED Facts 2017-2018, 2018-2019, and 2020-2021 data. | GAO-24-107485

Note: Directions of association with "\*" are statistically different from zero at the 90 percent confidence interval. If the direction of association is listed as "—" then the association was not statistically different from zero.

<sup>a</sup>The WIDA Consortium includes 42 states, territories, and federal agencies and designs the standards and assessments for English learners among other instructional support.