IES Institute of Education Sciences

APPENDIXES
December 2022

Regional Educational Laboratory Southwest
At American Institutes for Research

Biliteracy Seals in a Large Urban District in New Mexico: Who Earns Them and How Do They Impact College Outcomes?

See https://ies.ed.gov/ncee/rel/Products/Publication/100913 for the full report.

Appendix A. About biliteracy seals

Students who attend schools in the large urban district that participated in this study can earn four types of biliteracy seals:

- State Seal of Bilingualism-Biliteracy (state seal).
- District Spanish Bilingual Seal (district seal).
- District Spanish Bilingual Seal of Distinction (district seal of distinction).
- Global Seal of Biliteracy (global seal).

Currently, only five dual language high schools in the district offer the district seal or the district seal of distinction. Students who attend any public high school can earn a state seal or a global seal. The requirements for each seal are in table A1.

The New Mexico Public Education Department has approved four pathways for earning a state seal. The pathways available to students may vary by district based on resources available to support implementation or student interest. The four pathways are:

- Tribal language proficiency certification: Receive certification of proficiency in a tribal language from the tribe.
- Units of credit and assessment (assessment pathway): Earn a C or higher for four credits in the same non-English language (including any combination of language courses, language arts courses, and content courses) and demonstrate proficiency on an assessment in a non-English language.[1]
- Units of credit and alternative process portfolio (portfolio pathway): Earn a C or higher for four credits in the same non-English language (including any combination of language courses, language arts courses, and

[1] A student can demonstrate proficiency by attaining a score of 3 or higher on an Advanced Placement exam for a non-English language, attaining a score of 4 or higher on an International Baccalaureate exam for a higher-level non-English language, passing a national assessment of language proficiency in a non-English language, or passing a New Mexico assessment in a non-English language.

content courses) and create a portfolio (including an oral presentation to demonstrate language proficiency, a written sample of work, and an interview with a three-member panel of educators or community members proficient in the target language).

- Assessment and alternative process portfolio: Demonstrate proficiency on an assessment in a non-English language and create a portfolio (including an oral presentation to demonstrate language proficiency, a written sample of work, and an interview with a three-member panel of educators or community members proficient in the target language).

There is no explicit English language proficiency requirement as part of the state seal; students who meet the state graduation requirements are assumed to be English proficient.

The large urban district in this study grants the state seal through the assessment pathway and the portfolio pathway. High schools select one of these two pathways for their students to earn the state seal.

Table A1. Requirements for biliteracy seals for high school students in the large urban school district, 2019/20

Requirement	State seal	District seal	District seal of distinction	Global seal
School where seal is offered	Any public high school	Five dual language high schools	Five dual language high schools	Any public high school
Languages	Multiple[a]	Spanish only	Spanish only	Multiple[b]
Grade point average	A grade of C or higher in four credits of a non-English language	Minimum 2.0 grade point average in required English and Spanish courses	Minimum 3.5 grade point average in required English and Spanish courses	No requirement
Assessment of proficiency	District option 1: Demonstrate proficiency on an assessment in a non-English language[c]	No requirement	One Advanced Placement course taught in Spanish	Standards-based Measurement of Proficiency assessment
Portfolio[d]	District option 2: Portfolio consisting of an oral presentation to demonstrate language proficiency, a written sample of student work, and an interview with a panel	Portfolio consisting of an oral presentation to demonstrate language proficiency, a written sample of student work, and an interview with a panel	Portfolio consisting of an oral presentation to demonstrate language proficiency, a written sample of student work, and an interview with a panel	No requirement
Required credits	Four credits in the same non-English language, including any combination of language courses, language arts courses, and content courses	• Four credits in Spanish language arts[e] • Four credits in English language arts or English as a second language • Four credits in core content courses taught in Spanish • Four credits in core content courses taught in English	• Four credits in Spanish language arts[e] • Four credits in English language arts or English as a second language • Four credits in core content courses taught in Spanish • Four credits in core content courses taught in English	Three credits in English language arts or English language development in grades 9-11
Teacher recommendation[d]	No requirement	One letter of recommendation from any educator	One letter of recommendation from any educator	No requirement

a. Between 2016 and 2018 the state seal was awarded in Chinese, French, German, Japanese, Keres, Navajo, Spanish, Tewa, Tiwa, and Zuni.
b. The Global Seal of Biliteracy is offered in more than 100 languages (see https://theglobalseal.com/languages).
c. Students can demonstrate proficiency by attaining a score of 3 or higher on an Advanced Placement exam for a non-English language, a score of 4 or higher on an International Baccalaureate exam for a higher-level non-English language, a proficient score on a national assessment of proficiency in a non-English language, or a proficient score on a New Mexico assessment in a non-English language.
d. The study team did not have access to data on portfolios or teacher recommendations.
e. Introduced for students in the 2019/20 graduating cohort. Previously, students only had to complete Spanish IV or higher.
Source: Authors' compilation based on New Mexico Public Education Department (2020).

Reference

New Mexico Public Education Department. (2020). *State of New Mexico: Diploma of Excellence State Seal of Bilingualism-Biliteracy guidance handbook.* https://webnew.ped.state.nm.us/wp-content/uploads/2020/12/SSBB-Guidance-Handbook-2020-2021.pdf.

Appendix B. Methods

This study used student-level data provided by the large urban school district from an administrative database that is overseen by the New Mexico Public Education Department. The district also provided college enrollment data from the National Student Clearinghouse and a separate data file that tracked biliteracy seals earned, which was collected by the district's bilingual education program. The study team accessed Integrated Postsecondary Education Data System (IPEDS) data to obtain the characteristics of the colleges attended by students in the district. The district provided anonymized student identifiers, school names, and school years to link data across the provided datasets.

The study team obtained data for students who graduated from high school in 2017/18–2019/20. The first cohort was selected because this was the school year when the district began tracking which students earned biliteracy seals. The last cohort was selected because it was the most recent year of data available. To fully account for background characteristics, the data for these cohorts included information from graduates' grade 8 year (2013/14 for 2017/18 graduates, 2014/15 for 2018/19 graduates, and 2015/16 for 2019/20 graduates).

Data elements

The study team used several sets of variables for the analysis: seal information, student characteristics, course information (course enrollment and grades), assessment data, school characteristics, and college information (college attended and college characteristics). Data on seals awarded to students were obtained from the district's bilingual education program. Student characteristics, course information, assessment data, and school characteristics were obtained from the Student Teacher Accountability Reporting System (STARS) data management system used by the New Mexico Public Education Department. College attended and some college characteristics were obtained by the district from National Student Clearinghouse records; other college characteristics were obtained from IPEDS (National Center for Education Statistics, n.d.). Data provided by the district were anonymized with unique student identifiers to allow for merging across the datasets. IPEDS data were joined with data provided by the district using an institution identifier crosswalk for IPEDS and National Student Clearinghouse data, publicly available from the National Student Clearinghouse.

Data provided by the district ranged from student data for each school and year to student course-level data for each school and year. School-level characteristics were calculated using student-level and course-level files. Students without demographic characteristics in STARS were not included in this analysis. Students with missing college enrollment data in the National Student Clearinghouse data were assumed to have not attended college because it was not possible to distinguish students who did not attend college from students who attended college but did not have reported data. Similarly, students without a graduation status were assumed to have not graduated because they could not be distinguished from students who graduated but did not have reported data. Table B1 defines each data element and the source of the provided data.

Table B1. Summary of variables used in analyses

Variable	Data source	Description	Research question
Seal information			
Type of seal awarded	District	State Seal of Bilingualism-Biliteracy, District Spanish Bilingual Seal, District Spanish Seal of Distinction, Global Seal of Biliteracy	1, 2, 3, 4, 5, 6, 7
Date seal awarded	District	Date the seal was awarded	1, 2, 3, 4, 5, 6, 7

Variable	Data source	Description	Research question
Student characteristics			
Current grade level	STARS	Grade level in a given school year	1, 3, 4
Gender	STARS	Male or female	1, 3, 4, 6, 7
Race/ethnicity	STARS	Caucasian, Black or African American, Asian, American Indian/Alaska Native, Native Hawaiian or Other Pacific Islander[a]	1, 3, 4, 6, 7
Hispanic indicator	STARS	Indicator of whether the student is of Hispanic ethnicity	1, 3, 4, 6, 7
Eligibility for the National School Lunch Program	STARS	Student eligibility for the National School Lunch Program, an indicator of economic disadvantage	1, 3, 4, 6, 7
Special education indicator	STARS	Indicator of whether a student received special education services in a given school year	1, 3, 4, 6, 7
Direct certification code	STARS	Economic disadvantaged status code; coded as 1 if a student or family members are Supplemental Nutrition Assistance Program direct certified or are identified as eligible for other direct certification (such as homeless, Food Distribution Program on Indian Reservations, foster, migrant, and Head Start)	1, 3, 4, 6, 7
English learner indicator	STARS	Indicator of whether a student is a current or former English learner student	1, 3, 4, 6, 7
Home language	STARS	Non-English language spoken at home	1, 3, 4, 6, 7
Graduation indicator	STARS	Indicator of whether a student graduated from high school	1, 3, 4, 6, 7
Gifted education indicator	STARS	Indicator of whether a student participated in a gifted education program	1, 3, 4, 6, 7
Course information			
Seal course indicator	District	Indicator of whether a course meets the course requirements for state or district seal	5
Course credit earned	STARS	Number of credits earned by the student for a course in a given semester year	5
Course grade	STARS	Letter grade earned by the student for a course in a given semester year	5
Assessment data			
Grade-specific math and English language arts assessment score	District	Scale score on standardized exam taken by students in grades 4–8 and at the end of courses in math and English language arts, normalized within grade and school year	1, 3, 4, 6, 7
PSAT score	District	PSAT score, normalized	1, 3, 4, 6, 7
Non-English language AP exam score	District	Non-English language AP exam score	5
STAMP language proficiency exam score	District	STAMP reading, listening, writing, and speaking scores	5
School characteristics			
Charter school indicator	District	Indicator of whether a school is a charter school	1, 2, 3, 4, 5, 6, 7
Dual language school indicator	District	Indicator of whether a school is a dual language school	1, 2, 3, 4, 5, 6, 7

Variable	Data source	Description	Research question
College information			
College enrollment indicator	NSC	Indicator of whether a student enrolled in college	6, 7
In-state college indicator	NSC	Indicator of whether a college is located in New Mexico	7
Four-year college indicator	NSC	Indicator of whether a college is a four-year college	7
Full-time enrollment indicator	NSC	Indicator of whether a student is attending college full time	7

AP is Advanced Placement. NSC is National Student Clearinghouse. PSAT is Preliminary Scholastic Achievement Test. STAMP is Standards-based Measurement of Proficiency. STARS is Student Teacher Accountability Reporting System.
a. Asian and Native Hawaiian or Other Pacific Islander were collapsed into a single category, referred to as Asian.
Source: Authors' compilation.

Sample

The district provided data for students who graduated in 2017/18–2019/20 (table B2). For research question 1 the study team used three years of data to compare the characteristics of the five dual language high schools in the district (which award district seals and district seals of distinction) with those of the eight high schools in the district that are not dual language high schools. For research questions 2, 3, and 4 the study team used three cohorts of data (2017/18–2019/20) on 12,184 high school graduates to examine the demographic characteristics and college readiness of students who earned different types of seals and students who did not (table B3). For research question 5 the study team examined five years of data for each cohort of graduates, beginning with the 2017/18 graduates' grade 8 year. These data were used to evaluate course information for the entire period during which a student could complete courses that met the seal course requirements. The sample for research question 5 was limited to 2,595 unique students who graduated in 2017/18–2019/20 and who completed at least two credits in a non-English language but did not earn any type of biliteracy seal.

Table B2. Years of data requested to address research questions 1–5

Graduating cohort	2014/15	2015/16	2016/17	2017/18	2018/19	2019/20
2017/18 graduates	Grade 9	Grade 10	Grade 11	Grade 12	College enrollment	
2018/19 graduates		Grade 9	Grade 10	Grade 11	Grade 12	College enrollment
2019/20 graduates			Grade 9	Grade 10	Grade 11	Grade 12

Source: Authors' compilation.

Table B3. Sample size for research questions 1–5

Research question	Sample size
School-level question	
1. Number of schools	13
Student-level questions	
2. Number of graduates between 2017/18 and 2019/20	12,184
3. Number of graduates between 2017/18 and 2019/20	12,184
4. Number of graduates between 2017/18 and 2019/20	12,184
5. Number of graduates who partially or fully met credit requirements, grade requirements, and assessment requirement but did not earn a state seal	2,595

Source: Authors' analysis of data provided by a large urban school district in New Mexico.

Finally, for research questions 6 and 7 the study team used data on graduates from 2017/18 to 2019/20 to examine whether they enrolled in college and, among those who enrolled, whether they enrolled in a New Mexico college, enrolled in a four-year college, and enrolled full time (table B4). This sample excluded groups of students in which fewer than 12 students had earned a state seal across the three years of the study (including students in special education, students in gifted education, American Indian students, Asian students, and Black students) and students who did not take the algebra I assessment. The study team then compared outcomes between graduates who earned any type of biliteracy seal and similar graduates who did not earn a seal and between graduates who earned a district seal and similar graduates who earned only a state seal.

Table B4. Number of high school students who graduated between 2017/18 and 2018/19 and are included in the analysis for research questions 6 and 7

	Outcome	
Comparison	College enrollment	Enrollment in a New Mexico college, enrollment in a four-year college, or enrollment in college full time
Any type of biliteracy seal versus no seal	6,801	4,861
District seal versus state seal only	351	271

Source: Authors' analysis of data provided by a large urban school district in New Mexico.

Data preparation

For all research questions, student demographic characteristics that were categorical were recoded into indicators. Standardized assessment scores were normalized to have a mean of 0 and a standard deviation of 1 within each grade, subject, and school year. For research question 1, student-level measures were aggregated to the school level and then summarized.

For research question 5, student courses were coded for whether they qualified for meeting seal requirements. Information about the total number of credits and the average grade a student received in each course was aggregated for each student. For state seal requirements, courses were averaged for the number of credits completed in seal courses and the number of credits completed in seal courses for which the student received a letter grade of C or higher.

Analytic methods and models

This study used descriptive statistics (research questions 1–5) and regression analysis with propensity score weighting (research question 6 and 7) to address the research questions.

Methods for addressing research questions 1. The study team compared mean differences in school-level characteristics for dual language high schools (which offer the district seal and the district seal of distinction) and high schools that do not offer the district seals. The school-level characteristics examined included student demographic composition (gender, percentage of students by race/ethnicity, eligibility for the National School Lunch Program, special education and gifted status, English learner status, home language, and student enrollment) and school-level average achievement in English language arts and math and on the PSAT. For binary characteristics differences that are statistically significant and 5 percentage points or greater based on Pearson's chi-squared tests between groups are considered meaningful and are highlighted in the narrative. For continuous characteristics differences that are statistically significant and 0.1 standard deviation or greater based on *t* tests between groups are considered meaningful and are highlighted in the narrative.

Methods for addressing research question 2. The study team used descriptive methods to examine the number of students who earned the four different types of biliteracy seals offered in the district over the three-year study

period. The study team summarized the number of graduates in school years 2017/18–2019/20 who received any type of biliteracy seal and each type of seal earned.

Methods for addressing research question 3 and 4. The study team used descriptive methods to compare the mean differences in the characteristics and college readiness of graduates who earned any type of biliteracy seal in 2017/18–2019/20 with those of graduates who did not earn a seal. Then, the study team compared the mean differences in the characteristics of graduates who earned a state seal with those of graduates of dual language high schools who earned a district seal or a district seal of distinction. Finally, the study team compared the mean differences in characteristics and college readiness of 2019/20 graduates who earned a global seal with those of graduates who did not earn a global seal (the findings from this supplemental analysis are in appendix C). For binary characteristics statistically significant differences of 5 percentage points or greater based on Pearson's chi-squared tests between groups are considered meaningful and are highlighted in the narrative. For continuous characteristics statistically significant differences of 0.1 standard deviation or greater based on *t* tests between groups are considered meaningful and are highlighted in the narrative.

Methods for addressing research questions 5. The study team used descriptive methods to assess the number of graduates who partially or fully met coursework and grade requirements for the state seal but did not earn one. The study team also used data from two eligible assessments[2]—a Spanish Advanced Placement exam and the Standards-based Measurement of Proficiency assessment (an online proficiency assessment specifically for the Seal of Biliteracy; Avant, n.d.)—to describe the number of students who met the assessment requirement for a state seal (in addition to the course and grade requirements) but did not earn one.

The study team also summarized the number of graduates who met the credit and grade requirements by whether graduates attended a high school that offered the assessment pathway or the portfolio pathway and by whether graduates had ever been an English learner student. The study team assessed graduates' courses taken, grades, and assessment scores and compared them with awarded seal data to determine whether students met course, grade, and assessment requirements but did not earn a seal.

Methods for addressing research questions 6 and 7. The study team used doubly robust methods that combine regression analysis with propensity score weighting to evaluate the extent to which earning any type of biliteracy seal compared with not earning a seal, and earning a district seal compared with earning a state seal, affects enrollment in college, enrollment in a New Mexico college, enrollment in a four-year college, and enrollment in college full time.

The study team conducted this analysis for two samples of students. First, the study team compared graduates who earned any type of seal with similar graduates who did not earn a seal. Second, the study team compared graduates who earned only a district seal with similar graduates who earned only a state seal. Separating the analyses into these steps allowed the study team to examine both the impact of earning any seal and the impact on student outcomes of including English language requirements along with the non-English language requirements (because the district seal includes both, and the state seal includes only the non-English language requirement).

The study team used a propensity score weighting method to control for observable differences between the intervention and comparison groups for each analysis. The rationale behind the weighting approach is that, for

[2] Eligible assessments fall into three main categories: Advancement Placement exams in language and literature, International Baccalaureate exams in language and literature, and national language tests such as the Standards-based Measurement of Proficiency assessment or the College-Level Examination Program tests created by the College Board.

the purpose of understanding any one of the two comparisons, the study team can construct a comparison group that is similar in observable characteristics to the intervention group.

To conduct this analysis, the study team defined the following control variables: indicator for having ever been an English learner student, indicator for speaks Spanish at home, indicator for Hispanic, indicator for eligible for the National School Lunch Program, indicator for whether the student took algebra I in grade 8, indicator for school year, indicator for school attended (school fixed effect), and grade 8 grade point average.[3] The study team conducted complete case analysis and did not impute missing values. The study team was interested in the average treatment effect on the treated, or the effect of earning any type of biliteracy seal relative to not earning a seal. This is represented by equation B1, with the underlying aspsumption represented by equation B2. In equations B1 and B2 intervention T_i is defined as students who earned a state seal, $Y_i(1)$ denotes the potential outcome of student *i* had student *i* earned a state seal ($T_i = 1$), and $Y_i(0)$ denotes the potential outcome of student *i* had student *i* not earned a seal ($T_i = 0$). The average treatment effect on the treated (*ATE*) is defined as:

$$ATE = E[Y_i(1) - Y_i(0)]. \qquad \text{(B1)}$$

For an individual student both potential outcomes cannot be observed, so the study team constructed an appropriate counterfactual. The analysis assumes conditional independence, namely that conditional on observable characteristics, the potential outcome under no intervention, $Y_i(0)$, is mean independent of intervention T_i:

$$Y_i(0) \perp T_i \mid X_i. \qquad \text{(B2)}$$

The conditional independence assumption enters the analysis in two ways. First, the study team used X_i to estimate the propensity of earning any type of seal (relative to not earning a seal) and weight the observable demographic characteristics and college readiness of graduates who did not earn a seal toward the observable demographic characteristics and college readiness of graduates who earned any type of seal. Second, the study team estimated propensity weights using the entropy balancing method (Hainmueller, 2012). This method constructs a set of matching weights that, by design, forces certain balance metrics to hold. Entropy balancing is a generalization of the traditionally used inverse probability weighting (Hainmueller, 2012; Zhao, 2019; Zhao & Percival, 2017).

In the first step the study team estimated the matching weights that require the intervention and comparison groups to be balanced on the first three moments of each covariate listed earlier.[4]

In the second step (equation B3), the study team estimated intervention effects using a logistic regression model with the logit link function:

$$\eta_{ij} = log\left(\frac{p_{ij}}{1-p_{ij}}\right) \qquad \text{(B3)}$$

where η_{ij} is the log odds of attaining the outcome (for example, enrolling in college) for student *i* in high school *j* and p_{ij} is the probability of attaining the outcome for student *i* in high school *j*. The general equation (equation B4) for the outcome variable is:

$$\eta_{ij} = \beta_0 + \beta_1 T_i + \boldsymbol{X_i}\boldsymbol{\beta}_2 + \lambda_j + \epsilon_i \qquad \text{(B4)}$$

[3] The study team also explored whether including students' high school demographic composition (at the time of graduation) as opposed to the school fixed effect differentially affected the quality of the match (for example, smaller differences in observable characteristics between intervention and comparison group students). No major differences were found.

[4] Entropy balancing is implemented using the Covariate Balancing & Weighting Web App (CobWeb), an rShiny web application that implements multiple propensity weighting options. For more information, see https://andreasmarkoulidakis.shinyapps.io/cobweb/.

where T_i is an indicator for earning a biliteracy seal, $\boldsymbol{X_i}$ is the same vector of covariates used in the propensity score model, λ_j are school fixed effects, and ϵ_i is an idiosyncratic error. The model includes the same set of covariates used to generate the weights in the first stage, in order to control for any remaining differences between the intervention and comparison groups and to increase precision. Because the study team included school fixed effects in the covariates, it is thus comparing seal recipients with similar students who attended the same high school. Standard errors are clustered at the high school level.

For each comparison (any type of seal versus no seal and district seal versus state seal), the study team examined two samples of students using this model. First, the study team explored enrollment in any college, using all high school graduates from a given year. Next, the study team restricted the sample to students who enrolled in any college, examining the following outcomes: enrollment in a New Mexico college, enrollment in a four-year college, and enrollment in college full time.

By including the weights when estimating the regression, the coefficient β_1 is a doubly-robust two-step estimator, termed a regression-adjusted weighted estimator. This type of estimator is considered better in practice than regression or weighting on its own (Abadie & Imbens, 2011; Imbens & Wooldridge, 2009; Wooldridge, 2010). The final regression adjustment reduces bias from small differences in observables left over after the weighting process; this estimator is also robust with regard to misspecification of the regression function in the second step (Abadie & Imbens, 2011). However, although the two-step estimator created two groups of students with similar observable characteristics who differ only in intervention status, it cannot remove unobserved sources of bias (for example, student motivation).

Interpretation of regression results. The interpretation of the coefficients from the regressions using the logit link function are slightly different from those for a linear regression. In a linear regression a one-unit change in the covariate results in a change in the dependent variable equal to the coefficient, while all the other predictors are held constant. In the results reported in this study, where the outcome is transformed using a logit link function, a one-unit change in the covariate changes the log odds of the dependent variable. To help readers interpret the regression estimates, the study team used the regression results to calculate the adjusted probability of attaining an outcome (for binary outcomes), while holding other covariates in the model at their overall sample means (not by group). The group differences in the adjusted probabilities can be interpreted as the estimated effects for graduates earning any type of biliteracy seal compared with those for similar graduates who did not earn a seal (or the effects for graduates who earned a district seal compared with those for similar graduates who earned only a state seal). The group differences are provided in the main report.

Nonresponse bias analysis. Research questions 1–3 used information on standardized assessment scores. The study team obtained scores for at least 85 percent of students in the sample on all assessments, except the algebra I assessment and the PSAT. Of the 12,184 graduates in the sample, 8,949 had algebra I assessment scores (a response rate of 79.2 percent), and 9,529 had PSAT scores (a response rate of 84.4 percent). To determine whether there was nonresponse bias, the study team compared respondents with assessment scores and the full study sample on covariates that were available for both groups and that are strongly related to algebra I assessment scores and PSAT scores. All differences between respondents with assessment scores and the full study sample are less than 0.05 standard deviation, indicating no clear evidence of bias (tables B5 and B6).

Table B5. Nonresponse bias analysis for algebra I assessment scores

Covariate	Students with algebra I data (n = 8,949)	Original study sample (n = 12,184)	Pooled standard deviation	Standardized mean difference (Hedges's g)	Correlation with algebra I score
Eligible for the National School Lunch Program	0.60	0.60	0.49	0.00	-0.33
In special education	0.17	0.18	0.37	0.01	-0.34
In gifted education	0.04	0.04	0.20	-0.00	0.20
Current English learner student	0.13	0.14	0.33	0.00	-0.26
Ever an English learner student	0.30	0.30	0.45	0.00	-0.19
Hispanic	0.67	0.67	0.47	0.00	-0.17

Source: Authors' analysis of data provided by a large urban school district in New Mexico.

Table B6. Nonresponse bias analysis for scores on the Preliminary Scholastic Achievement Test

Covariate	Students with PSAT data (n = 9,529)	Original study sample (n = 12,184)	Pooled standard deviation	Standardized mean difference (Hedges's g)	Correlation with PSAT
Eligible for the National School Lunch Program	0.59	0.60	0.49	0.02	-0.43
In special education	0.16	0.18	0.18	0.04	-0.35
In gifted education	0.04	0.04	0.04	-0.01	0.25
Current English learner student	0.13	0.14	0.14	0.03	-0.31
Ever an English learner student	0.29	0.30	0.30	0.01	-0.26
Hispanic	0.67	0.67	0.67	0.00	-0.31

PSAT is Preliminary Scholastic Achievement Test.
Source: Authors' analysis of data provided by a large urban school district in New Mexico.

References

Abadie, A., & Imbens, G. W. (2011). Bias-corrected matching estimators for average treatment effects. *Journal of Business & Economic Statistics, 29*(1), 1-11.

Avant. (n.d.). *Test overview*. Retrieved May 30, 2022, from https://avantassessment.com/test-overview.

Hainmueller, J. (2012). Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies. *Political Analysis, 20*(1), 25-46.

Imbens, G. W., & Wooldridge, J. M. (2009). Recent developments in the econometrics of program evaluation. *Journal of Economic Literature*, *47*(1), 5-86.

National Center for Education Statistics. (n.d.). *Integrated Postsecondary Education Data System*. Retrieved May 30, 2022, from https://nces.ed.gov/ipeds/use-the-data.

Wooldridge, J. (2010). *Econometric analysis of cross section and panel data*. MIT Press.

Zhao, Q. (2019). Covariate balancing propensity score by tailored loss functions. *Annals of Statistics, 47*(2), 965-993.

Zhao, Q., & Percival, D. (2017). Entropy balancing is doubly robust. *Journal of Causal Inference, 5*(1), 1-19.

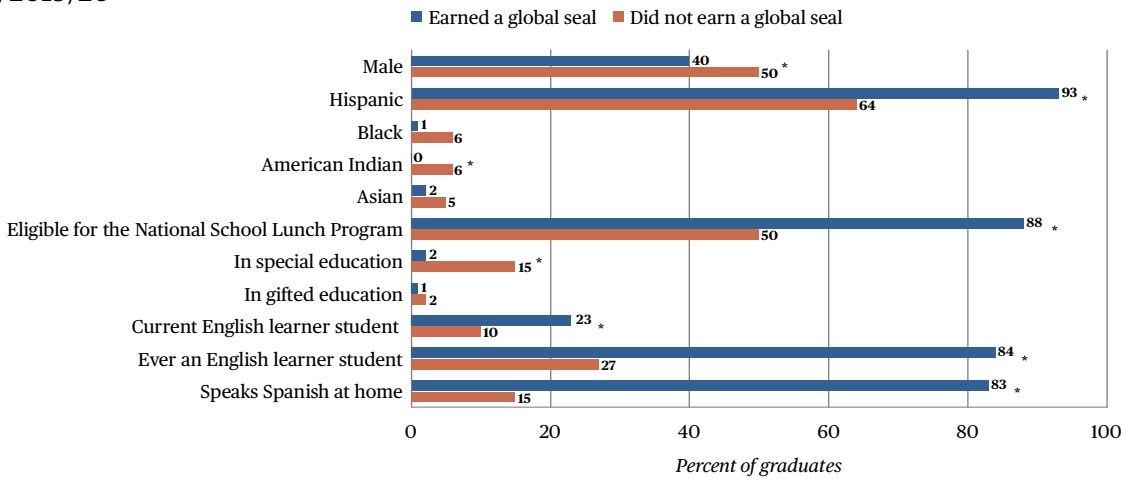
Appendix C. Supporting analyses

This appendix provides the results for the global seal and the full results for research questions 5 and 6.

Graduates who earned a global biliteracy seal were more likely than graduates who did not earn that seal to be Hispanic, to be eligible for the National School Lunch Program, to be a current English learner student, to have ever been an English learner student, and to speak Spanish at home

In 2019/20, the year in which the global seal was introduced, 354 graduates earned a global seal. Like graduates who earned any type of seal, graduates who earned a global seal were more likely than graduates who did not earn that seal to be Hispanic (93 percent versus 64 percent), to be eligible for the National School Lunch Program (88 percent versus 50 percent), to be a current English learner student (23 percent versus 10 percent), to have ever been an English learner student (84 percent versus 27 percent), and to speak Spanish at home (83 percent versus 15 percent; figure C1). Graduates who earned a global seal were less likely than graduates who did not earn that seal to be male (40 percent versus 50 percent), American Indian (0 percent versus 6 percent), or in special education (2 percent versus 15 percent).

Figure C1. In the New Mexico district, graduates who earned a global biliteracy seal were more likely than graduates who did not earn that seal to be Hispanic, to be eligible for the National School Lunch Program, to be a current English learner student, to have ever been an English learner student, and to speak Spanish at home, 2019/20



* Difference is considered meaningful (statistically significant at $p < .05$ and 5 percentage points or greater).

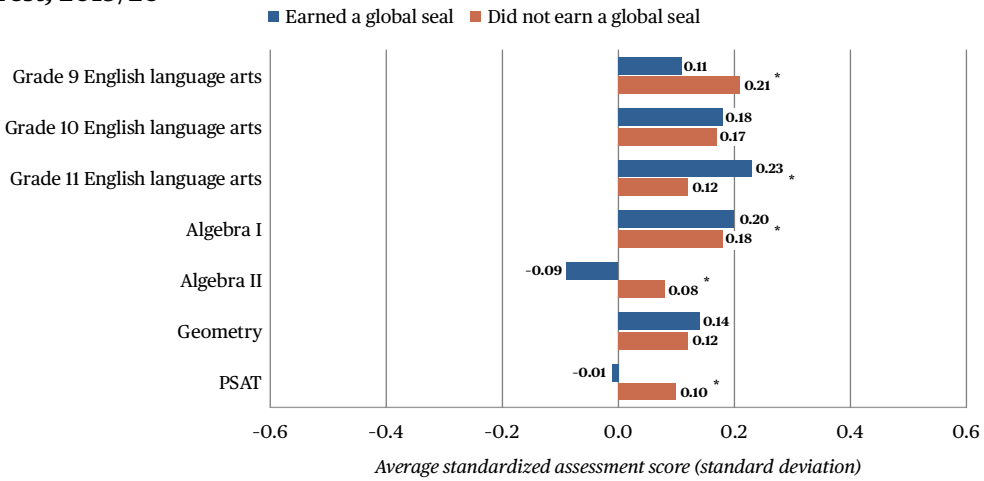
Note: The sample included 354 graduates who earned a global seal and 3,772 graduates who did not earn a global seal in 2019/20. Statistical significance was calculated using Pearson's chi-squared test.

Source: Authors' analysis of data provided by a large urban school district in New Mexico.

Graduates who earned a global seal performed better on some standardized assessments but worse on others than graduates who did not earn that seal

There were no clear patterns in the academic performance of graduates who earned a global seal compared with graduates who did not earn that seal (figure C2). Graduates who earned a global seal performed better than graduates who did not earn that seal on the grade 11 English language arts assessment but worse on the grade 9 English language arts assessment, the algebra II assessment, and the PSAT. There was no statistically significant difference in their performance on the grade 10 English language arts assessment, algebra I assessment, or geometry assessment.

Figure C2. In the New Mexico district, graduates who earned a global biliteracy seal performed better on the grade 11 English language arts assessment than graduates who did not earn that seal but worse on the grade 9 English language arts assessment, the algebra II assessment, and the Preliminary Scholastic Achievement Test, 2019/20



* Difference is considered meaningful (statistically significant at $p < .05$ and 0.1 standard deviation or greater).

PSAT is Preliminary Scholastic Achievement Test.

Note: The samples for each assessment are as follows. Grade 9 English language arts: 155 students who earned a global seal and 3,244 students who did not earn a global seal. Grade 10 English language arts: 157 students who earned a global seal and 3,377 students who did not earn a global seal. Grade 11 English language arts: 163 students who earned a global seal and 3,462 students who did not earn a global seal. Algebra I: 155 students who earned a global seal and 3,172 students who did not earn a global seal. Algebra II: 164 students who earned a global seal and 3,392 students who did not earn a global seal. Geometry: 159 students who earned a global seal and 3,318 students who did not earn a global seal. PSAT: 158 students who earned a global seal and 3,222 students who did not earn a global seal. Students took the Partnership for Assessment of College and Career Readiness assessment in 2017/18 and the New Mexico Standards-Based Transition Assessment of Mathematics and English Language Arts in 2018/19. Assessment scores were not available for 2019/20. Assessment scores were standardized within school year and grade in the district at the student level; therefore, a 0 represents the average score for all students in the district. Statistical significance was calculated using *t* tests.

Source: Authors' analysis of data provided by a large urban school district in New Mexico.

Baseline equivalence of sample groups for the analysis of college enrollment outcomes

Tables C1-C4 demonstrate the baseline equivalence of the intervention and comparison group samples for college enrollment, enrollment in a New Mexico college, enrollment in a four-year college, and enrollment in college full time for each of the two analyses: graduates who earned any type of seal compared with graduates who did not earn a seal and graduates who earned a district seal compared with graduates who earned a state seal. In each analysis the samples are equivalent at baseline.

Table C1. Characteristics of graduates who earned any type of biliteracy seal and of graduates who did not earn a seal in the New Mexico district, before and after entropy balancing algorithm for college enrollment, 2017/18–2019/20

		Graduates included in matching			Matched graduates included in analysis		
Covariate	All graduates (*n* = 6,801)	Earned any type of biliteracy seal (*n* = 522)	Did not earn a seal (*n* = 6,279)	Standardized difference	Earned any type of biliteracy seal (*n* = 522)	Did not earn a seal (*n* = 6,279)	Standardized difference
Male	0.47	0.30	0.48	0.18**	0.47	0.47	0.00
Ever an English learner student	0.28	0.79	0.24	-0.55**	0.28	0.28	0.00
Speaks Spanish at home	0.27	0.85	0.22	-0.63**	0.27	0.27	0.00
Eligible for the National School Lunch Program	0.30	0.40	0.29	-0.11**	0.30	0.30	0.00
Hispanic	0.74	0.95	0.72	-0.23**	0.74	0.74	0.00
Direct certification indicator	0.25	0.39	0.24	-0.15**	0.25	0.25	0.00
Grade 8 GPA	3.17	3.32	3.20	-0.11*	3.17	3.17	0.00
Took algebra in grade 8	0.11	0.18	0.10	-0.07**	0.11	0.11	0.00
2018/19 cohort	0.33	0.29	0.34	0.05*	0.33	0.33	0.00
2019/20 cohort	0.34	0.42	0.33	-0.09**	0.34	0.34	0.00

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

GPA is grade point average.

Note: Sample excludes student groups in which fewer than 12 students earned a state seal over the three years in the district studied (students in special education, students in gifted education, American Indian students, Asian students, and Black students). The correlation between grade 8 GPA and enrollment in college for this sample is 0.30.

Source: Authors' analysis of data provided by a large urban school district in New Mexico.

Table C2. Characteristics of graduates from dual language high schools who earned either type of district seal and of those who earned a state seal in the New Mexico district, before and after entropy balancing algorithm for college enrollment, 2017/18–2019/20

Covariate	All graduates of dual language high schools (*n* = 351)	Graduates included in matching			Matched graduates included in analysis		
		Earned a district seal (*n* = 181)	Earned a state seal (*n* = 170)	Standardized difference	Earned a district seal (*n* = 181)	Earned a state seal (*n* = 170)	Standardized difference
Male	0.31	0.31	0.31	-0.04	0.31	0.30	0.01
Ever an English learner student	0.72	0.86	0.58	-0.29**	0.72	0.72	0.00
Speaks Spanish at home	0.79	0.92	0.65	-0.27**	0.79	0.79	0.00
Eligible for the National School Lunch Program	0.39	0.45	0.33	-0.12*	0.39	0.39	0.00
Hispanic	0.91	0.99	0.83	-0.16**	0.91	0.91	0.00
Direct certification indicator	0.37	0.44	0.3	-0.14**	0.37	0.37	0.00
Grade 8 GPA	3.16	3.00	3.30	0.32**	3.16	3.17	-0.01
Took algebra in grade 8	0.15	0.07	0.23	0.16**	0.15	0.15	0.00
2018/19 cohort	0.33	0.27	0.39	0.12*	0.33	0.33	0.00
2019/20 cohort	0.39	0.36	0.42	0.05	0.39	0.39	0.00

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

GPA is grade point average.

Note: Sample excludes student groups in which fewer than 12 students earned a state seal over the three years in the district studied (students in special education, students in gifted education, American Indian students, Asian students, and Black students). The correlation between grade 8 GPA and enrollment in college for this sample is 0.14.

Source: Authors' analysis of data provided by a large urban school district in New Mexico.

Table C3. Characteristics of graduates who earned any type of biliteracy seal and of those who did not earn a seal in the New Mexico district, before and after entropy balancing algorithm for college outcomes (enrollment in a New Mexico college, enrollment in a four-year college, and enrollment in college full time), conditional on enrollment, 2017/18–2019/20

		Graduates included in matching			Matched graduates included in analysis		
Covariate	All graduates of dual language high schools who enrolled in college (*n* = 4,861)	Earned any type of biliteracy seal (*n* = 398)	Did not earn a seal (*n* = 4,463)	Standardized difference	Earned any type of biliteracy seal (*n* = 398)	Did not earn a seal (*n* = 4,463)	Standardized difference
Male	0.42	0.27	0.44	0.17**	0.42	0.42	0.00
Ever an English learner student	0.26	0.77	0.21	-0.56**	0.26	0.26	0.00
Speaks Spanish at home	0.26	0.84	0.20	-0.63**	0.26	0.26	0.00
Eligible for the National School Lunch Program	0.27	0.38	0.26	-0.13**	0.27	0.27	-0.01
Hispanic	0.72	0.95	0.70	-0.26**	0.72	0.72	0.00
Direct certification indicator	0.21	0.35	0.20	-0.16**	0.21	0.21	0.00
Grade 8 GPA	3.29	3.30	3.30	-0.03	3.29	3.29	0.00
Took algebra in grade 8	0.12	0.18	0.12	-0.06**	0.12	0.12	0.00
2018/19 cohort	0.34	0.32	0.34	0.02	0.34	0.34	0.00
2019/20 cohort	0.30	0.36	0.29	-0.02	0.30	0.30	0.00

** Significant at $p < .01$.

GPA is grade point average.

Note: Sample excludes student groups in which fewer than 12 students earned a state seal over the three years in the district studied (students in special education, students in gifted education, American Indian students, Asian students, and Black students). The correlation between grade 8 GPA and enrollment in a New Mexico college is -0.193, the correlation between grade 8 GPA and enrollment in a four-year college is 0.42, and the correlation between grade 8 GPA and enrollment in college full time is 0.25 for this sample.

Source: Authors' analysis of data provided by a large urban school district in New Mexico.

Table C4. Characteristics of graduates from dual language high schools who earned either type of district seal and of those who earned a state seal in the New Mexico district, before and after entropy balancing algorithm for college outcomes (enrollment in a New Mexico college, enrollment in a four-year college, and enrollment in college full time), conditional on enrollment, 2017/18–2019/20

		Graduates included in matching			Matched graduates included in analysis		
Covariate	All graduates who enrolled in college (*n* = 271)	Earned a district seal (*n* = 135)	Earned a state seal (*n* = 136)	Standardized difference	Earned a district seal (*n* = 135)	Earned a state seal (*n* = 136)	Standardized difference
Male	0.27	0.26	0.27	0.01	0.27	0.27	0.00
Ever an English learner student	0.70	0.85	0.54	-0.31**	0.69	0.69	0.00
Speaks Spanish at home	0.78	0.93	0.63	-0.30**	0.78	0.78	0.00
Eligible for the National School Lunch Program	0.38	0.41	0.36	-0.05	0.37	0.38	0.01
Hispanic	0.91	0.99	0.82	-0.17**	0.90	0.90	0.00
Direct certification indicator	0.34	0.39	0.29	-0.10*	0.32	0.33	0.01
Grade 8 GPA	3.22	3.00	3.40	0.35**	3.22	3.22	0.00
Took algebra in grade 8	0.14	0.05	0.24	0.18**	0.14	0.14	0.00
2018/19 cohort	0.36	0.28	0.43	0.15**	0.37	0.36	0.01
2019/20 cohort	0.32	0.30	0.35	0.06	0.31	0.32	0.01

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

GPA is grade point average.

Note: Sample excludes student groups in which fewer than 12 students earned a state seal over the three years in the district studied (students in special education, students in gifted education, American Indian students, Asian students, and Black students). The correlation between grade 8 GPA and enrollment in a New Mexico college is -0.29, the correlation between grade 8 GPA and enrollment in a four-year college is 0.37, and the correlation between grade 8 GPA and enrollment in college full time is 0.09 for this sample.

Source: Authors' analysis of data provided by a large urban school district in New Mexico.

The regression coefficients from the models predicting the four outcomes are in tables C5-C8.

Table C5. Regression results for enrolling in any college, 2017/18-2019/20

	Any type of seal versus no seal (n = 6,801)		District seal versus state seal (n = 351)	
Covariate or intervention group	Coefficient	Standard error	Coefficient	Standard error
Any seal	1.199*	0.545	na	na
District seal	na	na	0.864	0.634
Male	-0.468**	0.133	-0.075	0.487
Ever an English learner student	-0.115	0.231	-0.685	0.624
Speaks Spanish at home	-0.097	0.301	0.650**	0.226
Eligible for the National School Lunch Program	-0.209	0.220	0.040	0.134
Hispanic	-0.117	0.299	1.353	0.951
Direct certification indicator	-0.287	0.351	-0.787*	0.323
Grade 8 GPA	0.424	0.231	0.676	0.351
Took algebra in grade 8	0.872	0.558	-0.383	0.319
2018/19 cohort	-0.366	0.282	-0.394	0.425
2019/20 cohort	-1.007**	0.255	-1.458**	0.304
Constant	0.636	1.123	0.926	1.745
	Predicted probability (adjusted)	Standard error	Predicted probability (adjusted)	Standard error
Earned a seal	0.849	0.105	0.806	0.150
Did not earn a seal	0.716	0.185	0.754	0.222

* Significant at $p < .05$; ** significant at $p < .01$.
GPA is grade point average. na is not applicable.
Note: All models include school fixed effect indicators.
Source: Authors' analysis of data provided by a large urban school district in New Mexico.

Table C6. Regression results for enrolling in a New Mexico college, 2017/18–2019/20

	State seal (n = 4,861)		District seal (n = 255)	
Covariate or intervention group	Coefficient	Standard error	Coefficient	Standard error
Any seal	-0.441	0.564	na	na
District seal	na	na	-1.021	0.637
Male	-0.372	0.456	-1.457*	0.593
Ever an English learner student	0.967	0.535	1.964	1.435
Speaks Spanish at home	0.981	0.633	0.646	1.072
Eligible for the National School Lunch Program	0.174	0.335	-0.536	0.942
Hispanic	1.158**	0.234	1.710*	0.744
Direct certification indicator	-0.428	0.469	-0.799	1.351
Grade 8 GPA	-0.391	0.413	0.303	1.038
Took algebra in grade 8	-0.264	0.273	-0.0995	1.304
2018/19 cohort	1.337*	0.564	0.120	0.794
2019/20 cohort	-0.122	0.441	-1.998**	0.652
Constant	-0.372	0.456	-1.021	0.637
	Predicted probability (adjusted)	Standard error	Predicted probability (adjusted)	Standard error
Earned a seal	0.938	0.122	0.989	0.048
Did not earn a seal	0.841	0.178	0.819	0.258

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

GPA is grade point average. na is not applicable.

Note: All models include school fixed effect indicators.

Source: Authors' analysis of data provided by a large urban school district in New Mexico.

Table C7. Regression results for enrolling in a four-year college, 2017/18–2019/20

	State seal (n = 4,861)		District seal (n = 271)	
Covariate or intervention group	Coefficient	Standard error	Coefficient	Standard error
Any seal	1.029**	0.386	na	na
District seal	na	na	-0.272	0.576
Male	-0.072	0.332	0.273	0.424
Ever an English learner student	-0.651	0.338	-0.368	0.803
Speaks Spanish at home	0.061	0.209	0.499	1.033
Eligible for the National School Lunch Program	0.181	0.434	-0.643	0.567
Hispanic	-0.359	0.214	0.928	1.528
Direct certification indicator	-0.286	0.538	-0.896**	0.296
Grade 8 GPA	1.330**	0.329	0.422	0.260
Took algebra in grade 8	1.404**	0.305	0.308	0.437
2018/19 cohort	-0.246	0.342	-0.925	0.712
2019/20 cohort	0.058	0.254	0.469	0.511
Constant	-0.072	0.332	-0.272	0.576
	Predicted probability (adjusted)	Standard error	Predicted probability (adjusted)	Standard error
Earned a seal	0.606	0.244	0.240	0.177
Did not earn a seal	0.537	0.272	0.546	0.292

** Significant at $p < .01$.

GPA is grade point average. na is not applicable.

Note: All models include school fixed effect indicators.

Source: Authors' analysis of data provided by a large urban school district in New Mexico.

Table C8. Regression results for enrolling in college full-time, 2017/18–2019/20

	State seal (n = 4,861)		District seal (n = 241)	
Covariate or intervention group	Coefficient	Standard error	Coefficient	Standard error
Any seal	1.492**	0.410	na	na
District seal	na	na	-0.168	0.519
Male	-0.234	0.186	0.820**	0.264
Ever an English learner student	0.101	0.263	-0.732*	0.329
Speaks Spanish at home	-0.438	0.270	-0.727*	0.315
Eligible for the National School Lunch Program	0.242	0.238	0.685	0.593
Hispanic	-0.133	0.148	1.435	1.057
Direct certification indicator	-0.023	0.268	-1.044	1.250
Grade 8 GPA	0.741**	0.206	1.087**	0.298
Took algebra in grade 8	0.654*	0.329	-1.767**	0.362
2018/19 cohort	0.474**	0.125	0.485	0.602
2019/20 cohort	1.081**	0.314	1.881**	0.673
Constant	-0.234	0.186	-3.254	1.982
	Predicted probability (adjusted)	Standard error	Predicted probability (adjusted)	Standard error
Earned a seal	0.843	0.127	0.568	0.232
Did not earn a seal	0.683	0.189	0.692	0.204

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

GPA is grade point average. na is not applicable.

Note: All models include school fixed effect indicators.

Source: Authors' analysis of data provided by a large urban school district in New Mexico.