WestEd

Evaluation of Mountaineer Mathematics Master Teachers (M3T)

Analysis of Student Math Achievement Data in West Virginia

Joshua Valcarcel, Kevin Huang, Bryan Matlen, and Rebecca Perry

June 2026

Suggested citation: Valcarcel, J., Huang, K., Matlen, B., & Perry, R. (2026). *Evaluation of Mountaineer Mathematics Master Teachers (M3T): Analysis of student math achievement data in West Virginia*. WestEd.

WestEd is a nonpartisan, nonprofit organization that aims to improve the lives of children and adults at all ages of learning and development. We do this by addressing challenges in education and human development, increasing opportunity, and helping build communities where all can thrive. WestEd staff conduct and apply research, provide technical assistance, and support professional learning. We work with early learning educators, classroom teachers, local and state leaders, and policymakers at all levels of government.

For more information, visit WestEd.org. For regular updates on research, free resources, solutions, and job postings from WestEd, subscribe to the *E-Bulletin*, our semimonthly e-newsletter, at WestEd.org/subscribe.

Contents

Analysis of Student Math Achievement in West Virginia

Highlights:

- When controlling for prescore and demographic information, M3T treatment exposure is associated with students' higher mathematics achievement, although statistical significance varied by dosage and teacher group.
- The largest effects were observed among students taught by Math Teacher Leaders (MTLs). A span of 2 years of exposure to MTLs was associated with statistically significantly higher student mathematics achievement outcomes.
- As with any nonrandomized study, we cannot definitively separate the effects of the M3T project from all other preexisting differences between teachers.

Introduction

The Mountaineer Mathematics Master Teacher (M3T) project is a group of researchers, secondary-level mathematics teachers, and instructional leaders from West Virginia who have been working together over a 5-year period in a networked improvement community (also known as a NIC) to support and grow a corps of MTLs from across the state. The project is using principles of improvement science to

- help teachers initiate and sustain improvements to mathematics instruction in their classrooms, schools, and districts and
- improve West Virginia's capacity to leverage local leaders in statewide initiatives by improving outcomes for teachers and students across the state.

M3T MTLs receive funding and ongoing professional learning support from project directors in order to build their leadership capacity. Through their codesign of the M3T project and leadership of local improvement teams (LITs), MTLs play a key role in their district's or county's ongoing improvement of mathematics instruction.

WestEd was invited to provide formative evaluation, summative evaluation, and feedback over the course of the project to assess the degree to which project goals are being met. Specifically, the evaluation has been organized to answer four evaluation questions (EQs):

EQ1. What are the key features of the M3T project? To what extent do those features change over time and shape how the project is defined?

EQ2. What are the perceptions of the MTLs regarding the project as it relates to the effectiveness of their own teaching, their ability to serve as teacher leaders, and their overall job satisfaction?

EQ3. What are the perceptions of participating teachers, administrators, and other interest holders (e.g., West Virginia Department of Education) regarding the M3T project as it relates to improving teaching and learning in their respective locales?

EQ4. Is participation in the M3T project associated with improved instruction and student math achievement in the classrooms of MTLs and in the classrooms of teachers they support?

This report describes findings from WestEd's analysis of the student data used to investigate EQ4. WestEd researchers obtained and analyzed data from the West Virginia Department of Education (WVDE) to investigate the extent to which participation in the M3T project is associated with student math achievement as measured by West Virginia's General Summative Assessment (GSA) Math scores. Although data were provided for grades 6–11, only the data for grades 6–8 included GSA Math scores and could be used for analysis.

Findings

Analysis indicated that, when prescore and demographic information is controlled for, M3T treatment exposure is associated with students' higher mathematics achievement. Positive treatment coefficients were observed across all models examined, although statistical significance varied by dosage and teacher group.

Students who received 2 years of instruction from M3T-participating teachers scored significantly higher than did the matched comparison students (adjusted treatment effect $\beta = 0.22$, $g = 0.24$, $p < .01$). The regression models explained approximately 62 percent of the variance in student mathematics outcomes. The largest effects were observed among students taught by MLTs. Exposure of 2 years was associated with statistically significantly higher mathematics achievement outcomes ($\beta = 0.48$, $g = 0.54$, $p < .01$).

Across the overall sample, the estimated treatment effect for 1 year of exposure was small and not statistically significant ($\beta = 0.02$, $g = 0.02$, $p = .56$). Similarly, 1-year analyses conducted separately for students taught by MTLs ($\beta = 0.04$, $g = 0.04$, $p = .52$) and LIT members ($\beta = 0.01$, $g = 0.01$, $p = .72$) did not yield statistically significant effects. Students taught by LIT members also demonstrated positive treatment effects after 2 years of exposure ($\beta = 0.15$, $g = 0.16$), although this effect did not reach conventional levels of statistical significance ($p = .10$).

Across models, R^2 values ranged from approximately .59 to .62, indicating that the regression models explained a substantial proportion of the variance in student mathematics achievement outcomes.

Overall, findings suggest that sustained exposure to teachers participating in the M3T project may be associated with improved student mathematics achievement outcomes, particularly among students taught by MTLs and among students receiving 2 years of treatment exposure. In contrast, 1 year of exposure was generally associated with minimal differences in achievement outcomes relative to outcomes for matched comparison students, suggesting that the effects of the project may emerge more strongly with long-term exposure. Given the limitations of West Virginia assessment data beyond grade 8, it may not be possible to model the impact of long-term exposure.

In addition, with the available data it is not possible to fully attribute the positive statistical findings to the M3T project; outcomes may also reflect "good" mathematics teaching. We cannot completely rule out the possibility that treated teachers differed from untreated teachers before the intervention. We did not have achievement data from the period before the M3T intervention, and teachers were not randomly assigned to treatment. To address this concern, we controlled for students' prior math achievement, which helps account for preexisting differences in the students served by each teacher. The treatment effect remained after adjusting for baseline achievement, suggesting that the findings are not solely attributable to treated teachers having stronger students. However, as with any nonrandomized study, we cannot definitively separate the effects of the intervention from all preexisting differences between teachers.

Data Provided

The data set provided included 613,647 observations across 140,040 unique students in grades 6–11. This included 15,756 students who were taught at least once by a teacher in the M3T project, either an MTL or a LIT member, (i.e., treatment students) and 124,284 students who were never taught by a teacher in the project (i.e., control students). Appendix A provides greater detail on the data provided.

The WVDE research team conducted the matching procedure using coarsened exact matching to create a comparison group of students who had similar key baseline characteristics as the treatment students. Students were matched within strata (i.e., distinct, homogeneous subgroups derived from a larger data set), and students without a match were excluded from the analytic sample. Matching was not based on school, grade level, or district, which allowed matched students within a stratum to differ across those characteristics. The WVDE also provided design weights to account for over- or underrepresentation of comparison students within matching strata. To investigate EQ4, standardized (i.e., *z*-scores) GSA Math scores were

provided as an outcome measure. Students' average elementary (grades 3–5) GSA Math *z*-score was included as a prescore.

Data Cleaning

WestEd restricted its analysis to students with outcome data, which was defined as valid GSA Math *z*-scores. Because the GSA Math is only administered in grades 3–8 in the state of West Virginia, only students with data in grades 6–8 were included in analysis. GSA Math *z*-scores for grade 11 were also included in the data set provided to WestEd. However, this was confirmed to reflect SAT School Day assessment scores rather than GSA scores. As such, grade 11 records were excluded because SAT School Day scores were not considered directly comparable to GSA Math scores. This reduced the number of observations from 613,647 across 140,040 unique students to 390,308 observations across 95,281 unique students.

The analytic sample was further restricted to students who received treatment during grades 6–8 and their matched comparison students within the corresponding matching strata. Because matching was conducted within strata rather than at the student-year level, records for comparison students were retained only for the years corresponding to the observations of treatment students within the same stratum. This filtering further reduced the data set used for analysis. Note that although MTLs included high school teachers, the data limitations did not enable WestEd to examine the impact of the project on the students of these teachers.

As shown in Appendix B, analyses examined the association between M3T treatment exposure dosage and students' mathematics achievement outcomes. Dosage was defined as the number of years in which a student was taught at least one mathematics course by a teacher who participated in the M3T project, with separate analyses conducted for students who received 1 year or 2 years of treatment exposure. It was not feasible to examine 3 years of treatment exposure due to the low number of students (88 total) who received 3 years of treatment. Note that comparison student data could vary across different analyses (e.g., overall 1-year vs. overall 2-year). This is because data were not available for all comparison students in the same years as their matches in the treatment group, so some comparison students were included in one analysis but not another.

As such, the final data sets used for analysis were 69,138 observations across 54,754 unique students for the 1-year analysis and 8,389 observations across 8,115 unique students for the 2-year analysis. Table C1 in Appendix C summarizes the steps in the data cleaning process and the number of observations and unique students left after each step.

Data analysis also indicated that recorded gender was not always consistent across years. To address inconsistencies in longitudinal demographic records, students with multiple gender values across years were recoded to the first gender value observed in the data set. A total of two students within the data set used in analysis had inconsistencies in recorded gender.

(Thirty-eight additional students in the data set provided also had inconsistencies in recorded gender, but these students were removed during the data cleaning process.)

MTLs worked alongside M3T project directors to codesign guidelines for how to conduct mathematics-focused inquiry cycles. They also worked with LITs of two to six teachers at their school sites. Separate 1-year and 2-year dosage analyses were conducted for both MTLs and LIT members. To ensure a clearer interpretation of treatment exposure, analyses were limited to students who were consistently taught by either MTLs or LIT members across the relevant treatment period.

Analysis

WestEd estimated weighted linear regression models to examine the association between M3T treatment exposure and students' mathematics achievement outcomes. Mathematics achievement, defined as GSA Math *z*-score, served as the outcome variable, and treatment status was included as the primary predictor. In addition, models controlled for grade level, student gender, socioeconomic status (SES) (measured in West Virginia by direct certification status), and prior mathematics achievement (defined as average elementary GSA Math *z*-score) in order to account for baseline differences between treatment and comparison students. Analyses incorporated the design weights provided by the WVDE and used cluster-robust standard errors at the school level to account for the nonindependence of observations among students who attended the same school.

Separate analyses were conducted for students who received 1 year or 2 years of M3T treatment exposure (Table 1). Analyses were first conducted using the full treatment sample and corresponding matched comparison students. Additional subgroup analyses were then conducted separately for students taught by MTLs and students taught by LIT members. (Students taught by both MTLs and LIT members were included in the overall 1-year and 2-year analyses but not in subgroup analyses.) One-year and 2-year treatment exposure models were estimated separately for both groups to examine whether the association between treatment exposure and students' mathematics achievement differed across the two implementation pathways. Effect sizes are reported using Hedges' *g*.

Because matching was done by the WVDE research team, WestEd was not able to confirm the quality of matching. However, to the extent possible, WestEd examined the baseline difference between treatment group and the matched control group (Appendix D). In addition, because matching was not performed separately for each study sample, reported impacts beyond those for the Overall Year 1 data set (i.e., Year 2 findings, and facilitator- and participant-specific findings) should be considered exploratory.

Table 1. 1-Year and 2-Year Outcomes for Overall Data, MTL-Specific Data, and LIT Member–Specific Data Analysis

Treatment group	Duration	Adjusted treatment effect (β)	Effect size (*g*)	SE	*p*	Observations	Unique students	R^2
Overall	1 year	0.02	0.02	0.04	.56	69,138	54,754	0.61
Overall	2 years	0.22	0.24	0.06	< .01*	8,389	8,115	0.62
MTLs	1 year	0.04	0.04	0.06	.52	34,464	29,395	0.61
MTLs	2 years	0.48	0.54	0.13	< .01*	3,910	3,886	0.59
LIT members	1 year	0.01	0.01	0.04	.72	46,891	40,217	0.61
LIT members	2 years	0.15	0.16	0.09	.10	4,901	4,780	0.60

Note. Models were estimated using weighted ordinary least squares (or OLS) regression with cluster-robust standard errors clustered at the school level. All models controlled for grade level, gender, SES, and prior achievement. Observations represent student-year records. Comparison students may contribute multiple years of data when matched to treatment students within strata across multiple treatment-entry years. The asterisk represents statistical significance.

Conclusions

Analysis of West Virginia GSA Math scores data indicated that students taught by teachers who participated in the M3T project generally demonstrated higher mathematics achievement than did matched comparison students who were not taught by participating teachers. These differences were most pronounced among students who received 2 years of instruction from participating teachers, particularly those taught by MTLs. In contrast, differences between students who received 1 year of instruction from teachers who were LIT members and matched comparison students were small and not statistically significant. Similarly, students taught by teachers who were LIT members demonstrated slightly higher mathematics achievement outcomes than did comparison students, but these differences were not statistically significant at either the 1-year or 2-year exposure level.

Overall, findings suggest that sustained exposure to mathematics teachers participating in the M3T project, particularly MTLs, may be associated with improved student mathematics achievement outcomes. Although results should be interpreted within the limitations of an observational matched comparison design, the pattern of findings provides preliminary evidence that long-term participation in the M3T project may positively impact student mathematics performance in West Virginia schools.

Appendix A. Sample Descriptives

Appendix A summarizes the descriptive statistics for all predictor variables for treatment and control students used in analysis. This includes both unweighted and weighted descriptive statistics. Tables A1–A12 are presented for Year 1 and Year 2 data sets for overall analysis, MTL analysis, and LIT member analysis.

Table A1. Unweighted Descriptives for Year 1 Overall Data Set

Variable	Control	Treatment
Unique students	44,827	9,927
Female	22,258	4,974
Male	22,569	4,953
Low SES	15,440	2,815
Not low SES	34,222	7,112
Grade 6	20,624	2,815
Grade 7	19,915	4,009
Grade 8	18,601	3,103
Elementary math *z* mean	0.28	0.20
Elementary math *z* SD	0.80	0.83
GSA Math *z* mean	0.25	0.20
GSA Math *z* SD	0.93	0.94

Note. Students with inconsistent gender values across years were recoded to their first observed gender value. Grade totals may exceed unique student counts because multiple years of data were retained for some comparison students.

Table A2. Weighted Descriptives for Year 1 Overall Data Set

Variable	Control	Treatment
Weighted student count	49,805.03	9,927
% Female	50%	50%
% Male	50%	50%
% Low SES	69%	72%
% Not low SES	31%	28%
% Grade 6	33%	28%
% Grade 7	33%	40%
% Grade 8	34%	31%
Elementary math *z* mean	0.27	0.20
Elementary math *z* SD	0.79	0.83
GSA Math z mean	0.24	0.20
GSA Math *z* SD	0.92	0.94

Note. Students with inconsistent gender values across years were recoded to their first observed gender value. Percentages and means in weighted tables reflect application of WVDE-provided design weights. Treatment student weights were uniformly equal to 1.

Table A3. Unweighted Descriptives for Year 2 Overall Data Set

Variable	Control	Treatment
Unique students	6,782	1,333
Female	3,325	685
Male	3,457	648
Low SES	1,246	118
Not low SES	5,604	1,215
Grade 6	2,449	8
Grade 7	2,376	461
Grade 8	2,231	864
Elementary math *z* mean	0.53	0.36
Elementary math *z* SD	0.76	0.83
GSA Math *z* mean	0.47	0.53
GSA Math *z* SD	0.90	0.90

Note. Students with inconsistent gender values across years were recoded to their first observed gender value. Grade totals may exceed unique student counts because multiple years of data were retained for some comparison students.

Table A4. Weighted Descriptives for Year 2 Overall Data Set

Variable	Control	Treatment
Weighted student count	5,790.198	1,333
% Female	50%	51%
% Male	50%	49%
% Low SES	81%	91%
% Not low SES	19%	9%
% Grade 6	33%	1%
% Grade 7	33%	35%
% Grade 8	34%	65%
Elementary math *z* mean	0.54	0.36
Elementary math *z* SD	0.75	0.83
GSA Math *z* mean	0.47	0.53
GSA Math *z* SD	0.90	0.90

Note. Students with inconsistent gender values across years were recoded to their first observed gender value. Percentages and means in weighted tables reflect application of WVDE-provided design weights. Treatment student weights were uniformly equal to 1.

Table A5. Unweighted Descriptives for Year 1 MTL Data Set

Variable	Control	Treatment
Unique Students	26,059	3,336
Female	12,780	1,708
Male	13,279	1,628
Low SES	9,649	1,152
Not low SES	17,936	2,184
Grade 6	10,822	638
Grade 7	10,496	1,615
Grade 8	9,780	1,083
Elementary math *z* mean	0.39	0.30
Elementary math *z* SD	0.79	0.81
GSA Math *z* mean	0.36	0.31
GSA Math *z* SD	0.93	0.94

Note. Students with inconsistent gender values across years were recoded to their first observed gender value. Grade totals may exceed unique student counts because multiple years of data were retained for some comparison students.

Table A6. Weighted Descriptives for Year 1 MTL Data Set

Variable	Control	Treatment
Weighted student count	26,625.68	3,336
% Female	50%	51%
% Male	50%	49%
% Low SES	64%	65%
% Not low SES	36%	35%
% Grade 6	33%	19%
% Grade 7	33%	48%
% Grade 8	34%	32%
Elementary math *z* mean	0.38	0.30
Elementary math *z* SD	0.62	0.55
GSA Math *z* mean	0.34	0.31
GSA Math *z* SD	0.93	0.94

Note. Students with inconsistent gender values across years were recoded to their first observed gender value. Percentages and means in weighted tables reflect application of WVDE-provided design weights. Treatment student weights were uniformly equal to 1.

Table A7. Unweighted Descriptives for Year 2 MTL Data Set

Variable	Control	Treatment
Unique Students	3,184	702
Female	1,495	356
Male	1,689	346
Low SES	286	46
Not low SES	2,910	656
Grade 6	1,197	8
Grade 7	1,063	259
Grade 8	948	435
Elementary math *z* mean	0.58	0.47
Elementary math *z* SD	0.73	0.82
GSA Math *z* mean	0.51	0.72
GSA Math *z* SD	0.89	0.85

Note. Students with inconsistent gender values across years were recoded to their first observed gender value. Grade totals may exceed unique student counts because multiple years of data were retained for some comparison students.

Table A8. Weighted Descriptives for Year 2 MTL Data Set

Variable	Control	Treatment
Weighted student count	2,621.498	702
% Female	49%	51%
% Male	51%	49%
% Low SES	90%	93%
% Not low SES	10%	7%
% Grade 6	37%	1%
% Grade 7	32%	37%
% Grade 8	30%	62%
Elementary math *z* mean	0.58	0.47
Elementary math *z* SD	0.74	0.82
GSA Math *z* mean	0.50	0.72
GSA Math *z* SD	0.90	0.85

Note. Students with inconsistent gender values across years were recoded to their first observed gender value. Percentages and means in weighted tables reflect application of WVDE-provided design weights. Treatment student weights were uniformly equal to 1.

Table A9. Unweighted Descriptives for Year 1 Participant Data Set

Variable	Control	Treatment
Unique Students	33,626	6,591
Female	16,724	3,266
Male	16,902	3,325
Low SES	9,754	1,663
Not low SES	26,531	4,928
Grade 6	14,123	2,177
Grade 7	13,578	2,394
Grade 8	12,566	2,020
Elementary math *z* mean	0.25	0.16
Elementary math *z* SD	0.80	0.83
GSA Math *z* mean	0.23	0.14
GSA Math *z* SD	0.93	0.93

Note. Students with inconsistent gender values across years were recoded to their first observed gender value. Grade totals may exceed unique student counts because multiple years of data were retained for some comparison students.

Table A10. Weighted Descriptives for Year 1 Participant Data Set

Variable	Control	Treatment
Weighted student count	32,949.27	6,591
% Female	50%	50%
% Male	50%	50%
% Low SES	74%	75%
% Not low SES	26%	25%
% Grade 6	34%	33%
% Grade 7	33%	36%
% Grade 8	33%	31%
Elementary math *z* mean	0.24	0.16
Elementary math *z* SD	0.49	0.40
GSA Math *z* mean	0.21	0.14
GSA Math *z* SD	0.92	0.93

Note. Students with inconsistent gender values across years were recoded to their first observed gender value. Percentages and means in weighted tables reflect application of WVDE-provided design weights. Treatment student weights were uniformly equal to 1.

Table A11. Unweighted Descriptives for Year 2 Participant Data Set

Variable	Control	Treatment
Unique Students	4,149	631
Female	2,074	329
Male	2,075	302
Low SES	982	72
Not low SES	3,193	559
Grade 6	1,427	0
Grade 7	1,438	202
Grade 8	1,405	429
Elementary math *z* mean	0.55	0.24
Elementary math *z* SD	0.79	0.84
GSA Math *z* mean	0.50	0.32
GSA Math *z* SD	0.92	0.91

Note. Students with inconsistent gender values across years were recoded to their first observed gender value. Grade totals may exceed unique student counts because multiple years of data were retained for some comparison students.

Table A12. Weighted Descriptives for Year 2 Participant Data Set

Variable	Control	Treatment
Weighted student count	3,481.437	631
% Female	49%	52%
% Male	51%	48%
% Low SES	75%	89%
% Not low SES	25%	11%
% Grade 6	30%	0%
% Grade 7	33%	32%
% Grade 8	37%	68%
Elementary math *z* mean	0.55	0.24
Elementary math *z* SD	0.77	0.84
GSA Math *z* mean	0.50	0.32
GSA Math *z* SD	0.91	0.91

Note. Students with inconsistent gender values across years were recoded to their first observed gender value. Percentages and means in weighted tables reflect application of WVDE-provided design weights. Treatment student weights were uniformly equal to 1.

Appendix B. Analysis Regression Models

Appendix B includes estimated coefficients, standard errors, and *p*-values for all predictors included in the regression models used for analysis.

Overall Analysis Regression Models

Table B1. Full Regression Results for Year 1 Overall Analysis

Independent variables	Estimate	SE	*p*
Treatment status	0.02	0.04	.56
Grade 7	–0.02	0.02	.32
Grade 8	–0.04	0.02	.09
Gender	–0.03	0.01	.00
SES	–0.10	0.02	.00
Prior math *z*-score	0.90	0.01	.00

Table B2. Full Regression Results for Year 2 Overall Analysis

Independent variables	Estimate	SE	*p*
Treatment status	0.22	0.06	< .01
Grade 7	0.00	0.03	.89
Grade 8	–0.02	0.03	.52
Gender	–0.02	0.02	.18
SES	–0.08	0.03	.01
Prior math *z*-score	0.92	0.02	.00

MTL Analysis Regression Models

Table B3. Full Regression Results for Year 1 MTL Analysis

Independent variables	Estimate	SE	*p*
Treatment status	0.04	0.06	.52
Grade 7	–0.02	0.02	.39
Grade 8	–0.07	0.03	.01
Gender	–0.03	0.01	.00
SES	–0.11	0.02	.00
Prior math *z*-score	0.91	0.01	.00

Table B4. Full Regression Results for Year 2 MTL Analysis

Independent variables	Estimate	SE	*p*
Treatment status	0.48	0.13	.00
Grade 7	–0.01	0.06	.89
Grade 8	–0.05	0.06	.38
Gender	–0.01	0.03	.80
SES	–0.10	0.05	.05
Prior math *z*-score	0.93	0.05	.00

Participant Analysis Regression Models

Table B5. Full Regression Results for Year 1 Participant Analysis

Independent variables	Estimate	SE	*p*
Treatment status	0.01	0.04	.72
Grade 7	–0.02	0.02	.43
Grade 8	–0.02	0.02	.34
Gender	–0.03	0.01	.00
SES	–0.08	0.02	.00
Prior math *z*-score	0.90	0.01	.00

Table B6. Full Regression Results for Year 2 Participant Analysis

Independent variables	Estimate	SE	*p*
Treatment status	0.15	0.09	.10
Grade 7	0.03	0.05	.55
Grade 8	0.01	0.05	.79
Gender	–0.01	0.05	.83
SES	0.00	0.05	.93
Prior math *z*-score	0.89	0.03	.00

Appendix C. Analytic Sample Sizes and Data Set Construction

Table C1. Number Of Observations and Unique Students in Data Sets Provided and Used in Analysis

Data set	Number of observations	Number of students
Original data set provided	613,647	140,040
Data set with valid outcome data	390,308	95,281
Data set for Year 1 overall analysis	69,138	54,754
Data set for Year 2 overall analysis	8,389	8,115
Data set for Year 1 MTL analysis	34,464	29,395
Data set for Year 2 MTL analysis	3,910	3,886
Data set for Year 1 LIT member analysis	46,891	40,217
Data set for Year 2 LIT member analysis	4,901	4,780

Appendix D. Baseline Equivalence Analysis

Appendix D presents baseline equivalence analyses that examine differences in prior mathematics achievement between treatment and comparison students prior to treatment exposure. Overall, baseline differences were small, although some treatment samples, particularly the Year 2 overall and Year 1 LIT member samples, demonstrated modestly lower prior mathematics achievement relative to matched comparison students.

Table D1. Baseline Differences in Prior Mathematics Achievement Across Analytic Samples

Treatment group	Duration	Adjusted treatment effect (β)	Effect size (*g*)	SE	*p*	Observations	Unique students	R^2
Overall	1 year	–0.07	–0.08	0.04	.14	69,138	54,754	0.00
Overall	2 years	–0.18	–0.23	0.09	.05	8,389	8,115	0.01
MTLs	1 year	–0.08	–0.14	0.08	.27	34,464	29,395	0.00
MTLs	2 years	–0.04	–0.06	0.25	.86	3,910	3,886	0.00
LIT members	1 year	–0.10	–0.21	0.05	.03	46,891	40,217	0.00
LIT members	2 years	–0.08	–0.10	0.10	.40	4,901	4,780	0.00


WestEd