The Effects of Teacher Culturual Trainings on Student Achievement

All Authors Contributed Equally to this Paper

1st Desai, Sirsh
6yb@virginia.edu

2nd Landas, Stephanie
4rz@virginia.edu

3rd Thro, Noah
8eq@virginia.edu

3rd Wu, Michelle
8eq@virginia.edu

***Abstract*—Albemarle County Public School District implemented a cultural training program in 2015 for classroom teachers. Our question was to determine whether the cultural training program had an impact on students' growth and actual achievement on standardized tests. By using linear regression, gradient boosting regression, and Kruskal-Wallis tests, we were able to identify that the program did make an impact on students' growth and actual achievement; however it was not the most important factor. The students that had the highest increase in scores came from teachers that had the highest tenures and highest degrees.**

I. Introduction

Starting in X, the Albemarle County Public School District (ACPS) implemented a training program for teachers on cultural responsiveness. This training aimed to increase cultural awareness in the classroom, fostering a better relationship with students and their families. In 2020, this training was made mandatory with the goal of all teachers to be certified by 2026. In preparation for an upcoming restructuring of this program, we were tasked with determining whether or not this culturally responsive teacher training program had measurable effects on student testing outcomes, specifically, on reading and math scores.

Our main interests were to determine if the teacher certification had a statistically significant effect on student achievement and to determine what other aspects may be playing a stronger impact on these outcomes. This paper analyzes the differences in types of culturally responsive training certifications (full, micro, and none) and looks at other possible explanations for student growth. The results of this project will inform the ACPS leadership team on whether or not to implement culturally responsive teaching for all classroom teachers for this next upcoming school year of 2024-2025.

II. Related Work

A previous cohort of UVA data science master students worked with Albemarle County Public Schools to analyze the impact of the COVID-19 pandemic on the behavior, achievement, learning problems, and mental health of middle and high school students in ACPS. This was part of a larger initiative with the Virginia Center for Disease Control to create a longitudinal connecting education and medical records of children and adolescents statewide, enabling comprehensive research and intervention efforts.

A study by Kristina Laird-Arnold explored elementary teachers' perceptions of a cultural awareness curriculum. It investigated how teachers view the effectiveness of cultural awareness in their classrooms through interviews covering personal experiences, curriculum development, and professional training. Findings revealed teachers' beliefs in fostering cultural awareness and facilitating collaborative classroom environments. The study underscores the growing importance of cultural awareness in education, given the increasing diversity in the U.S. population, and emphasizes the significance of integrating cultural awareness into education for the holistic development of students, both socially and emotionally, and its role in building connections between schools and communities (Laird-Arnold, 2022).

III. Data Description

The data contains student demographic information (race, gender, English learner, disabilities, etc.), teacher demographic information (years experience, highest degree, birth year, gender, race, etc.), whether or not a teacher and a principle are certified as culturally responsive educators (and if so, what type of certification), and student results on Virginia Standards of Learning (SOL) such as growth, expected and actual achievement, and vertical scaled score (achievement scaled across grades and subjects). Our datasets contained all variables and outcomes on mathematics and reading SOLs from the 4th through 8th grade.

The main variables of interest were the teacher certification, growth, and actual achievement variables. The teacher certification variable reported on whether that student's teacher was fully certified, micro certified, or had no certification. Growth measured a student's growth between their expected achievement and their actual achievement. The actual achievement variable is the raw score of the student's standardized test score, on either reading or math. For the growth variable, there were 472 missing values in the reading dataset, 457 missing values in the math dataset, and 929 missing values in the combined dataset. These missing values are caused when there is no data on a student's expected achievement, usually caused when a student is new to ACPS.

IV. Methodology

To fit with the two goals of this analysis, we conducted different methodologies depending on whether we considered all predictors or only teacher certification.

A. Teacher Certification

We first wanted to identify whether the teacher certification alone had an effect on student growth. We first considered using an ANOVA table, but the data did not meet the normality assumption for any dataset given that we were testing on categorical data, and the reading dataset additionally did not meet the equal variances assumption when analyzing the actual achievement variable. Because of this, we used the Kruskal-Wallis test, which makes similar comparisons, but does not require a normal distribution or equal variances. To determine which, if any, type of certification had an effect, we used Dunn's test. Linear regression models were built for additional analysis.

B. All Possible Predictors

In order to build multivariate linear regression models, we looked at a stepwise regression for both growth and actual achievement. We also investigated random decision trees to determine the highest variable importance. To avoid high collinearity, any dependent variables that were not relevant to the model were dropped (eg. expected achievement, vertical scaled score, actual achievement percentile, and either growth or actual achievement, depending on the variable of interest).

To build our models, we looked at three different variable groupings: all predictors, student demographics alone, and only teacher and principal certification results. We used a python package called PyCaret, which evaluated the predictor variables with a number of different linear and nonlinear models, which then used cross validation to choose the best fitting model available. The best model for all of the variables while predicting for actual achievement was a gradient boosting model. The model trainings were done for both the reading and math datasets separately in order to evaluate whether the teacher certification status affected the actual achievements of the students' standardized test scores for either the reading or math tests.

V. Results

A. Teacher Certification in Isolation - Growth

1) Kruskal-Wallis Tests: The Kruskal-Wallis test found that teacher certification did have a statistically significant effect on growth, for all iterations of the data, though which certification pairs were significant did vary based on the data used. The reading dataset did not have a significant difference between the full certification and no certification, and the math dataset had no significant difference between the micro certification and no certification. The combined dataset showed statistically significant differences between all combinations of teacher certifications. Figure 1 shows the results of this test on the combined dataset.

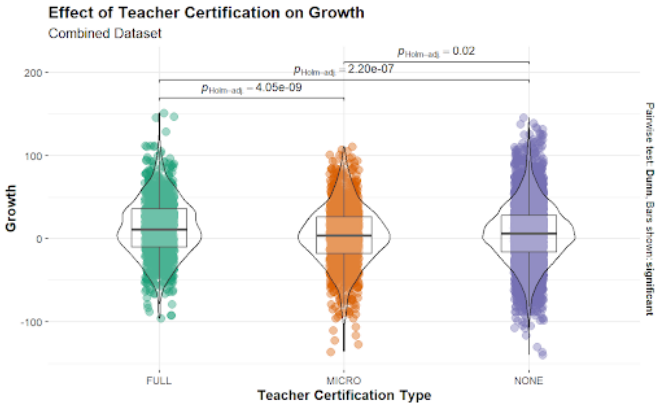

Fig. 1.

2) Linear Regressions: When looking at the linear regression between the teacher certification types and growth, we found that there was a statistically significant relationship, though the adjusted R2 values were extremely small. What was most surprising was that looking at the math data alone, having no certification improved growth more than having a full certification. Figure 2 showcases these results.

Linear Regression of Growth on Teacher Certification Type					
	Coefficients			Adjusted R^2	p-Value
	Intercept	Micro Certification	No Certification		
Reading Dataset	11.128	-7.394	-3.350	0.002731	9.936e-4
Math Dataset	5.3391	-1.5420	8.9282	0.009175	1.1e-09
Combined Dataset	12.6900	-8.9248	-6.1245	0.005041	1.382e-10

Fig. 2.

B. Teacher Certification in Isolation - Actual Achievement

1) Kruskal-Wallis Tests: The Kruskal-Wallis test found that teacher certification did have a statistically significant effect on actual achievement, but only for the math and combined dataset. The certification pairs that were significantly different varied between the math and combined datasets. The reading dataset had no statistically significant differences, and the math dataset showed statistically significant differences between all combinations of teacher certifications. The combined dataset had no significant difference between the micro certification and no certification. Figure 3 shows the results of this test on the combined dataset.

2) Linear Regressions: When looking at the linear regression between the teacher certification types and actual achievement, we found that there was no statistically significant relationship for the reading dataset, though the relationship was significant for the math and combined dataset. The math dataset again proves counterintuitive, where both the micro and no certification increase the actual achievement more than

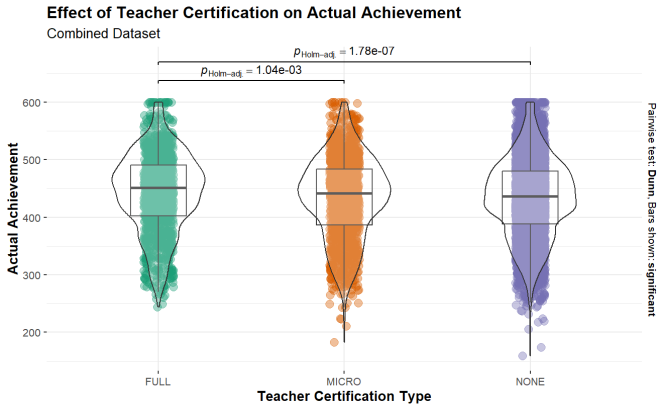

Fig. 3.

Linear Regression of Actual Achievement on Teacher Certification Type					
	Coefficients			Adjusted R^2	p-Value
	Intercept	Micro Certification	No Certification		
Reading Dataset	441.247	-8.252	-2.912	0.000759	0.05955
Math Dataset	428.250	6.770	16.751	0.007144	1.673e-8
Combined Dataset	443.088	-9.106	-9.783	0.002173	1.178e-5

Fig. 4.

the full certification. Figure 4 displays the results from this regression.

C. All Possible Predictors

1) Growth: When looking at a stepwise regression, the variables that were deemed the most important in predicting student growth were the type of teacher certification, the teacher's race, whether the student had a disability, the license type of the teacher, whether the student was an English learner, the student's gender, and the years of experience of the teacher. The final adjusted R2 was 0.03451.

The regression models predicting growth had low R2 values, meaning that the outcomes of these models may not be highly correlated together with predicting growth. For the regression model trained on all predictor variables and only teacher/principal demographics, the most important feature was the teacher birth year and teacher years experience, with the third most important being the year of teacher certification. For the regression model trained on only student demographics, the most important feature was grade 6. These results can be seen in Figure 5.

2) Actual Achievement: When looking at a stepwise regression, the variables that were deemed the most important in predicting student growth were whether the student was an English learner, whether the student had a disability, the student's race, the teacher's birth year, the teacher's highest degree, the type of principal certification, the type of teacher

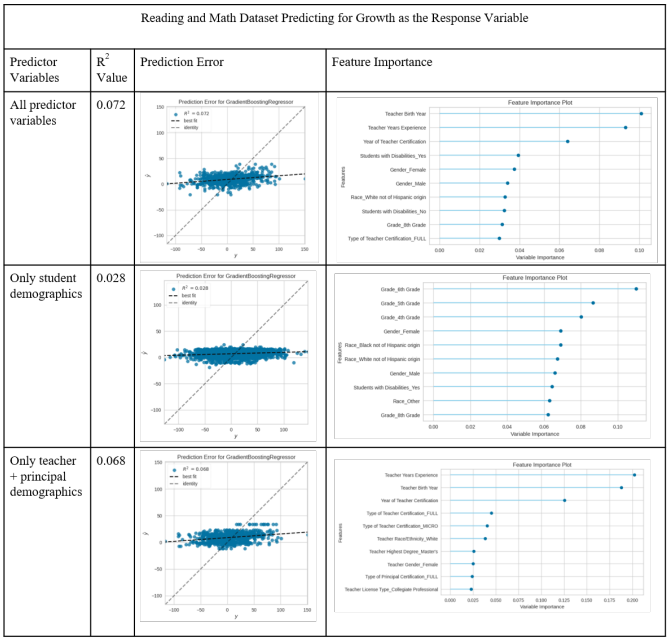

Fig. 5.

certification, the teacher's race, and the type of teacher license. The final adjusted R2 was 0.3417.

The regression models predicting for actual achievement that were trained with all predictors and with just student demographics revealed that the most important feature was not an English learner and a race of white, of not hispanic origin. The models trained with just teacher and principal demographics had the most important feature be the teacher birth year, followed by teacher's highest degree, meaning that it can be assumed that the experience of the teacher made a larger impact than the certification. From these models, the teacher certification did not make an impact on the students' performance on the standardized tests, and that there are other factors that would push the student's learning forward. These results can be seen in Figure 6.

VI. Discussion

Our analysis indicated that the teacher cultural responsiveness training did have a statistically significant effect on student growth. However, we found that when incorporating other possible predictor variables, such as student or teacher demographics, they played a stronger role in explaining the increase or decrease in a student's growth. The other most important variables were the years of experience of the teacher, whether the student was an English learner, and the highest degree of the teacher. Our analysis was affected by the small sample size, however, leading to less conclusive answers. There also was a fairly high degree of multicollinearity within our dataset and we find it worth mentioning that a large part of our analysis consisted of parsing through variables that displayed this collinearity. For example, one might expect teacher experience and teacher highest degree to be significantly related and much of our analysis consisted of

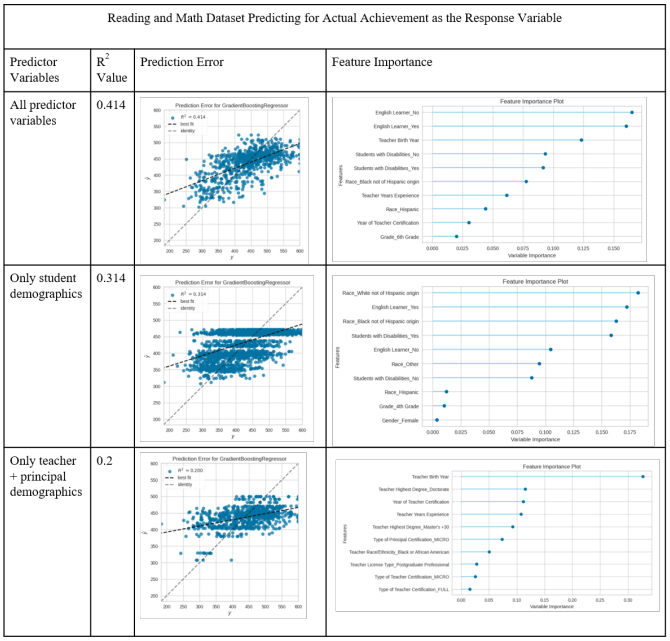

Fig. 6.

investigating which variables were the driving variables of influence.

Additionally, our analysis revealed that African American and Hispanic students tend to have both a lower actual achievement score as well as a lower average growth between years compared to their White peers. A contributing factor as to why this may be the case is directly tied to African American and Hispanic students having teachers with less experience than White students. Related to this is that teachers that were certified with the cultural training were disproportionately assigned to White students (with Black and Hispanic students receiving a smaller proportion of their groups being taught by such a teacher). In our discussions with our sponsor, this is likely due to different schools having different proportions of races but we encourage ACPS to investigate ways to make teacher qualifications between student race groups more equitable.

Recently, ACPS conducted a survey for teachers to review their experience with the cultural responsiveness training and its effects on their students. A vast majority of teachers indicated that the program improved learning outcomes for students and found that they were able to analyze their own biases that influenced their teaching. While the analysis does not indicate this training as the strongest predictor, the value teachers found should not be ignored.

VII. Conclusion / Future Work / Recommendations

Based on these results, we believe there is value in continuing the culturally responsive teaching training in the ACPS district. While this training did not have the strongest effect, it does have merit in a comprehensive training program. To encourage student growth, we recommend that ACPS focus on retaining teachers and hiring and retaining teachers with higher degrees. Teachers with greater experience, whether from teaching years or education, consistently appeared as strong, positive predictors in our analyses. Discussions with current ACPS should be conducted to determine what can be done to encourage longer and more consistent tenures.

In future analyses, we recommend conducting a longitudinal study to track student growth more closely and to allow for analysis based on different teacher experiences. This would allow for stronger assertions, with the assumption that a larger longitudinal study would include more cases than the one-year sample we received. It would also be beneficial to include more details about the school the student attended to track differences between the different environments.