In Brief: Evaluating the Impact of UVA-PLE on New Mexico Schools

The University of Virginia's Partnership for Leaders in Education (UVA-PLE) provides school districts with a multi-year commitment that leverages research-based practice to strengthen systems leadership capacity. UVA-PLE believes that district and school leaders must jointly re-examine their system to enrich conditions for sustainable, scalable improvements. To do this, UVA-PLE provides various design courses, executive education, multiple site visits, tailored supports, and various other professional learning services across four developmental stages.

- *Engage:* Together, UVA-PLE and district leaders identify the district's most significant needs and highest-leverage opportunities;
- *Design:* UVA-PLE architects long- and short-term strategies to address district-specific challenges, outline sustainable change, and prepare a learning lab of partner schools;
- *Activate:* School and district leaders immerse themselves in a rigorous campaign to ignite school performance, leadership commitment, and collective purpose; and,
- *Adapt:* Leaders integrate the success of year one and adapt their approach toward new areas of sustainable improvement.

Collectively, these professional learning opportunities are designed to result in the development of two core competencies for systems leaders. First, UVA-PLE is designed to increase district and school teams' ability to identify and confront their highest-leverage priorities for school improvement. Second, UVA-PLE builds district, school, and individual capacity to lead transformations (i.e., to execute rapid, high-impact school improvement).

The collective development of district and school leaders appears to be imperative to drive systemic change. Considerable research evidence underscores the importance of district leadership on school principals, classroom instruction, and student achievement (e.g., Leithwood, 2010; Marzano & Waters, 2009; Murphy & Hallinger, 1988). An even larger body of research suggests that principal quality is the second most important within-school factor on student achievement (Andrews & Soder, 1987; Louis et al., 2010; Robinson et al., 2008). Yet, the limited research on leader preparation programs indicates that what is taught varies considerably (Hess & Kelly, 2007), as do the knowledge and skills of district and school leaders completing them (Grissom et al., 2019).

As such, the evaluation of education leadership in-service programs seems critically important. These evaluations focus on key areas including identifying programs with proven success, demonstrating that leadership development broadly matters, and considering the adoption or adaptation of strategies embedded in the program more widely. In the only existing evaluation of the University of Virginia's Partnership for Leaders in Education (UVA-PLE), statistically significant positive effects on student achievement in two large Ohio districts were found in both English language arts and mathematics. Those positive effects persisted in the two years of data

available subsequent to the educational leaders' completion of the program (Player & Katz, 2016). According to Herman et al.'s (2018) review of school leadership interventions, the UVA-PLE is one of only two comprehensive school improvement models to have impact evidence substantial enough to meet tier I, II, or III levels of evidence, as defined in the Every Student Succeeds Act.

In this second evaluation of UVA-PLE impact on student achievement, we evaluate a pooled sample of traditional public schools in New Mexico. To do so, we employ a comparative interrupted time series model—a strong quasi-experimental design (Shadish et al., 2002)—to analyze student proficiency scores over time. Specifically, we answer the following questions:

- What is the impact of UVA-PLE on the percentage of students scoring proficient or higher in English language arts? In mathematics?

We turn next to provide an overview of the data we analyzed. Then, we briefly explain the methods used to conduct the analysis. After that, we present the results before concluding with some brief reflections on those results.

Data Analyzed for This Study

From 2010 through 2015, UVA-PLE began partnerships with 11 New Mexico districts, four of which partnered (or enlisted new schools) multiple times. In total, 56 schools began working with UVA-PLE during the five year window. The number of schools by district ranged from only 1 to 17. Proficiency data for both subjects, as well as student demographic percentages, were retrieved from the Public Education Department (PED) of New Mexico.[1] Relevant years of data files for this study were from 2007 through 2018. From them, we retrieved and/or calculated overall percent proficient by subject, as well as percent proficient by student demographics, for each school for students in Grades 3-8 and 11. Organizational data (e.g., school urbanicity) were retrieved from the National Center for Education Statistics (NCES)[2] and matched to the year of assessment by school code used by both PED and NCES.

For each school, the three years of data preceding Year 0 (the year in which their partnership with UVA-PLE began) were used for trend trajectory and prior achievement before partnering with PLE. Three years of post-PLE partnership initiation data were also collected, with Year 1 and Year 2 data taken from spring assessments while schools were still in the partnership (i.e., Year 2 testing would have been conducted before a school's partnership with UVA-PLE was completed). Year 3, however, was nearly one full academic year removed from the partnership.

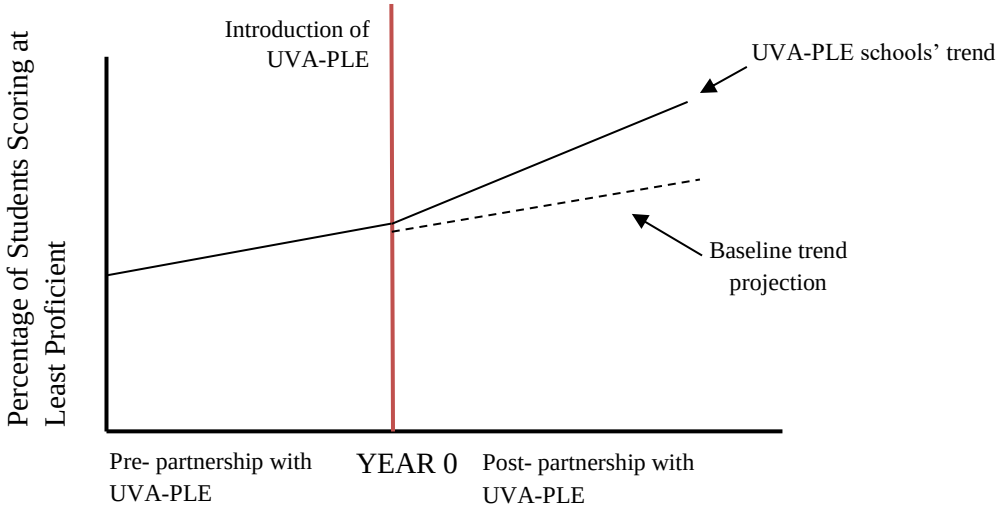
We conducted propensity score matching procedures to identify other New Mexico schools not partnering with UVA-PLE but statistically similar in prior achievement and student demographics to the study's sample to be the comparison group. The matching process was also conducted separately for ELA and mathematics. After conducting propensity score matching, we compared achievement trends of UVA-PLE schools with their matched comparison schools to verify that the trends were similar prior to UVA-PLE working with treatment districts and schools.

[1] https://webnew.ped.state.nm.us/bureaus/accountability/achievement-data/
[2] https://nces.ed.gov/ccd/elsi/

Methodology: Comparative Interrupted Times Series (CITS)

Comparative Interrupted Time Series (CITS) is one of the strongest quasi-experimental designs (Shadish et al., 2002). It is available to use when a comparison series (such as comparing groups of schools at different points in time) can be constructed. The CITS analysis allowed us to compare the percentage of students scoring proficient or higher for participating and comparison schools in the years preceding any partnerships with UVA-PLE with outcome trends in the years following those partnerships. Because schools began their partnerships with UVA-PLE in different years, we had to first establish a time series for each set of schools separately, where Year 0 for any treatment school (and its comparison school) would be the year that the treatment school began its partnership with UVA-PLE. Figure 1 illustrates of what we would expect to see if UVA-PLE had impact on the percentage of students scoring proficient or better. To the left of the red UVA-PLE time point where partnerships began, we would expect the trend scores for future UVA-PLE schools and their comparison schools to be statistically the same. Then once schools partner with UVA-PLE, we would expect those trends to diverge as UVA-PLE schools' achievement increases and/or accelerates faster than achievement in comparison schools over time.

Figure 1. Depiction of Comparative Interrupted Time Series Expected Pattern of Change in Student Test Scores



Note: The change in the post center implementation line does not represent an immediate change rather change overtime.

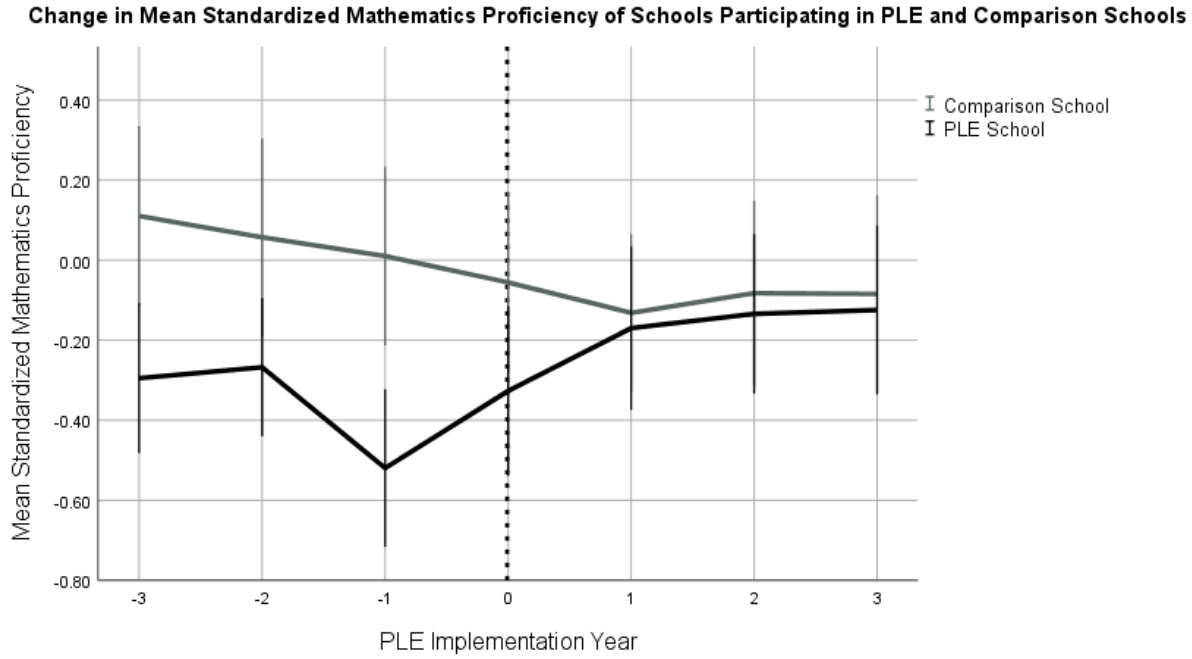
Results

Mathematics

When comparing the change in math proficiency trajectory, we observe a significant difference between UVA-PLE and comparison schools. Schools participating in UVA-PLE had a significant, positive increase in math proficiency compared to the comparison schools (see Figure 2). We added school demographic covariates to this model, and we still observe this same

trend even when controlling for school demographics. Additionally, while schools with higher percent ELL and Black student enrollment had significantly decreasing math achievement pre-UVA-PLE implementation, no school demographic covariates had a significant relationship to math proficiency trajectory following UVA-PLE implementation.

Figure 2.

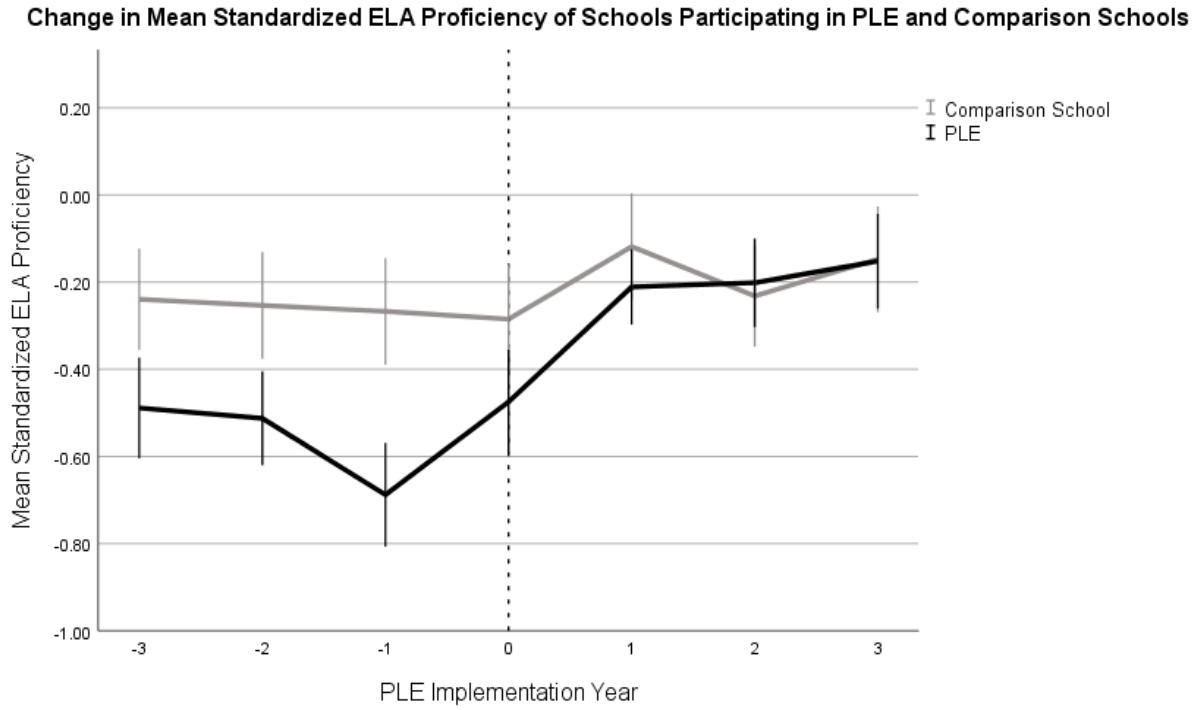


Note: Error bars represent 95% confidence intervals.

English/language arts

The slope of the change in the pre-UVA-PLE ELA achievement proficiency for all schools, PLE program and comparison, was not significant in the time only, UVA-PLE variable, or full models including school demographic covariates. This indicates that although UVA-PLE schools had more ELA improvement following participation in the PLE program compared to non-UVA-PLE schools, this difference in slopes was not statistically significant (see Figure 3). However, this difference in slopes following PLE implementation is nearly significant after controlling for school demographic characteristics ($\beta = .09$, $p < .10$). However, in our statistical model, with school demographics covariates added, we found that schools with *higher percent ELL student enrollment had significantly larger increases in ELA proficiency* ($\beta = 1.07$, $p < .01$ where β is in standard deviation units). This is a shift from a significantly decreasing trend for schools with higher ELL student enrollment prior to UVA-PLE implementation. This could indicate that UVA-PLE schools were better supporting ELL populations, and this was a driver of the near significant, positive slope difference between PLE and comparison schools.

Figure 3



Note: Error bars represent 95% confidence intervals.

Conclusions

This evaluation leverages CITS statistical approach, a strong quasi-experimental design that can be used to analyze publicly available data on aggregate school performance (Hallberg et al., 2018). Similar to the evaluation that Player and Katz (2016) conducted on Ohio data, we find that UVA-PLE had statistically significant impact on student mathematics achievement in New Mexico schools. Unlike the previous evaluation of UVA-PLE, we did not find statistically significant impacts on English/language arts, although the increases in proficiency were nearly statistically significant when controlling for demographics. In that vein, we want to underscore that ELA proficiency for English language learners was positive and statistically significant in ELA. These results are especially important because so few educational leadership programs have impact evidence on student achievement and this evaluation captured impact across eleven diverse school districts. UVA-PLE now has quasi-experimental impact evidence in two substantially different educational contexts.

References

Grissom, J. A., Mitani, H., & Woo, D. S. (2019). Principal preparation programs and principal outcomes. *Educational Administration Quarterly, 55*(1), 73-115. https://doi.org/10.1177/0013161X18785865

Hallberg, K., Williams, R., Swanlund, A., & Eno, J. (2018). Short comparative interrupted time series using aggregate school-level data in education research. *Educational Researcher, 47*(5), 295-306. https://doi.org/10.3102/0013189X18769302

Hess, F. M., & Kelly, A. P. (2007). Learning to lead: What gets taught in principal-preparation programs. *Teachers College Record, 109*(1), 1-28.

Leithwood, K. (2010). Characteristics of school districts that are exceptionally effective in closing the achievement gap. *Leadership and Policy in Schools, 9*(3), 245-291. https://doi.org/10.1080/15700761003731500

Louis, K. S., Leithwood, K., Wahlstrom, K. L., & Anderson, S. E. (2010). Investigating the links to improved student learning: Final report of research findings. Wallace Foundation.

Marzano, R. J., & Waters, T. (2009). *District leadership that works: Striking the right balance.* Solution Tree Press.

Murphy, J., & Hallinger, P. (1988). Characteristics of instructionally effective school districts. *The Journal of Educational Research, 81*(3), 175-181. https://doi.org/10.1080/00220671.1988.10885819

Richards, A. L., & Soder, R. (1987). Principal leadership and student achievement. *Educational Leadership, 44*(6), 9-11.

Robinson, V. M. J., Lloyd, C. A., & Rowe, K. J. (2008). The impact of leadership on student outcomes: An analysis of the differential effects of leadership types. *Educational Administration Quarterly, 44*(5), 635-674. https://doi.org/10.1177/0013161X08321509

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs of generalized causal inference.* Houghton Mifflin Company.

Contributing Authors

Coby Meyers is the Chief of Research of the Partnership for Leaders in Education and Associate Professor of Education in the University of Virginia School of Education and Human Development. Dr. Meyers' research focuses on understanding the role of school system leadership, especially in the context of improving low-performing schools.

Meredith Wronowski, Ph.D., is an Assistant Professor in the Educational Administration department at the University of Dayton. Her research interests are focused on issues of equity in schooling including the unintended effects of accountability policies on teachers and leaders, opportunity to learn, community-based school improvement, and resegregation of U.S. schools.

Laura LaMonica, MPP, is an independent research analyst conducting research and evaluation projects, specializing in education and public policy using both quantitative and qualitative research skills. Laura has worked in higher education for over five years at the University of Virginia, focusing on chronically underperforming schools with large proportions of students from economically disadvantaged backgrounds.