TEXAS A&M UNIVERSITY

RESEARCH BRIEF

TEXAS A&M UNIVERSITY
Center for Research & Development in Dual Language & Literacy Acquisition

TEXAS A&M UNIVERSITY
Education Leadership Research Center

School Enhancement Programs

Beverly J. Irby, Ed.D., Rafael Lara-Alecio, Ph.D., Fuhui Tong, Ph.D., Betsy Lockhart, Ph.D., Elsa Villarreal, Ph.D., Matthew J. Etchells, Ph.D., Rudolfo Gaytan, MBA, Reza Norouzian, Ph.D., Kara Sutton Jones, Ph.D, & Hailey Harris, M.S.

Introduction

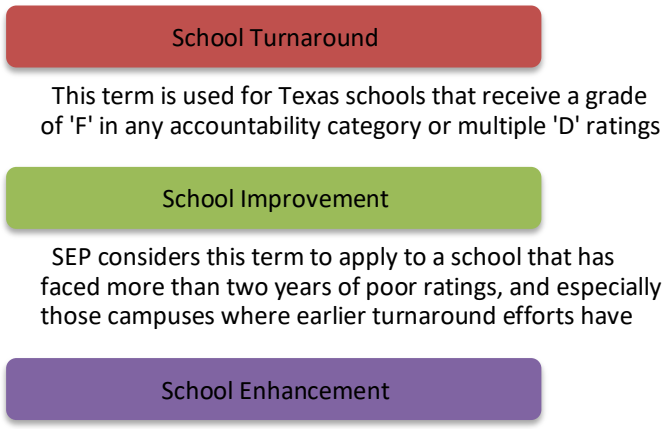
Project Accelerated Preparation of Leaders for Underserved Schools (A-PLUS; Irby, Lara-Alecio, Tong, & Torres, 2017) is building instructional capacity to impact diverse learners through a three-year, $13.7 million project funded by the Supporting Effective Educator Development (SEED) Grant Program. The project incorporates multiple innovative approaches to developing school leaders, including training practicing principals in building instructional capacity at the campus level in the education of English learners and students experiencing economic challenges.

The entire project is expected to serve 2,320 Texas principals and school leaders.

Consider the analogy of foundation repair. If a building is having issues with its foundation, it is necessary to assess the entire foundation, elevate the affected areas, provide support to problematic places, and attend to the seemingly unaffected, yet nearby, portions. To do a proper job that will last for years to come, many lifting points must exist.

This example, however, is rarely the view of school improvement, enhancement, or turnaround. As long as schools have been struggling with low campus ratings, leadership teams have spent countless hours and dollars in trying to determine the 'one thing' that they need to do that will help their schools improve. The A-PLUS School Enhancement Project (SEP) has been working to demonstrate how a panacea does not exist, but that schools can instead utilize a systematic approach to improve student outcomes through deliberate efforts from teachers and campus administrators, instead of a single, 'one-size-fits-all' approach.

Figure 1
Common Terms Defined by SEP



Theoretical Framework

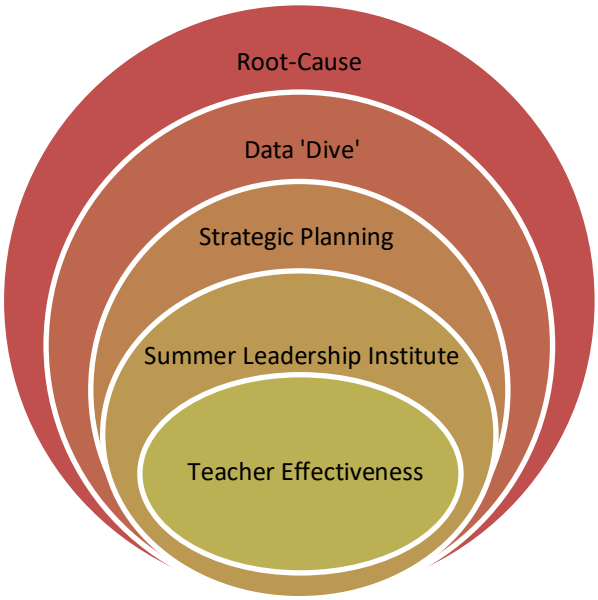
From a theoretical perspective, various requirements are involved in the school improvement process. Huberman et al. (2011) stressed the importance of developing a clear process for determining the needs of a school. Dufour and Mattos (2013) espoused collaboration among the instructional staff for improving the culture of teaching in the school. Day, Gu, and Sammons (2016) focused on transformational and instructional strategies for enhancing student outcomes. Ayscue and Siegel-Hawley (2019) discussed the importance of shared goals, not only between faculty, staff, and students, but with parents as well. Sun and Leithwood focused on direction setting practices as a means to influencing organizational outcomes, while Irby and Lunenburg (2014) included a communication section on a

survey regarding ethical principal practices. However, Leithwood and Riehl (2003) proposed a core set of leadership practices valuable in educational contexts: identifying and articulating a vision; creating shared meanings; creating high performance expectations; fostering the acceptance of group goals; monitoring organization performance; and communicating. As we shall see, A-PLUS's approach was premised on a combination of these considerations.

How do you improve a school?

The A-PLUS SEP has been working to answer this question for the past three school years. Simply framing this question is an extreme challenge, since the types of issues faced by schools that have received lower ratings by the State of Texas are as diverse as the student populations served by the faculty and staff of the schools involved in this component of the A-PLUS grant.

Figure 2
Operational Process of School Improvement



Schools involved in our SEP were first identified via the Texas school rating system. Schools that were not State Improvement Required, but that were within 2 points of becoming improvement required (IR) were recruited to participate in the SEP. Next, a clustered randomized control trial (RCT) design was employed with 8 eligible elementary schools from one participating school district in Texas. These schools were randomly assigned to Treatment or Control conditions.

As is common in educational interventions, attrition occurred. Two participating schools unexpectedly withdrew from the SEP around the conclusion of the 1st year of implementation. Thus, three Treatment and three Control schools remained for the full two-years of intervention.

During the first two years of the grant, the three Treatment schools were each involved in all "Turnaround" efforts, as each campus needed intensive improvement. Control schools participated in several interventions as well. Principals at the control campuses were simply not assigned leadership coaches as a part of their intervention. However, their participation in the Summer Leadership Institute proved highly effective as discussed later.

Despite the number of difficulties each of the cooperating schools faced, the SEP has seemingly 'cracked the code' for school improvement.

School Improvement Process

A reflection on the scholarly literature reveals that school improvement often involves focusing on instructional effectiveness, providing leadership coaches and tutors, and evaluating school improvement plans based on a campus assessment of Root-Cause (see Tschannen-Moran & Tschannen-Moran, 2011). As a result, A-PLUS attempted to make use of all such elements as a comprehensive way to approach the "school turnaround". Figure 2 summarizes these elements. At the writing of this brief, no other University-based project has brought together each of these elements, and the following list includes the most essential components.

A Randomized Controlled Trial (RCT) with a Cooperating District

The first year (October 2017 - June 2018) served as the launching point of this project. During this preliminary year, the emphasis for the eight randomly selected elementary schools with similar issues (student mobility, number of students in bilingual programs, etc.) was to establish the professional networks and relationships necessary to engage in the meaningful work of Instructional Change through leadership coaching. This first year placed two Leadership Coaches across the Treatment campuses, along with an Instructional Specialist, with the focused effort of assisting these struggling campuses.

By the time Year 3 (October 2019 – June 2020) began, there were four Leadership Coaches and five Instructional Specialists working to assist the Treatment schools to be their best selves:

- Building and Sustaining a School Vision
- Sharing Leadership
- Leading and Learning in Professional Learning Communities
- Cultivating Leadership in Other
- Using Data to Make Instructional Decisions
- Monitoring Curriculum and Instruction
- Critical Dialogues
- Understanding English Language Learners' Need and Dual Language Programs

Effects on Student Outcomes

The State of Texas Assessments of Academic Readiness (STAAR) are the chosen state standardized measures used in Texas public schools. STAAR Reading and Math data was collected from the six schools that were retained throughout the first two years of implementation (2017-2019). Data was unable to be collected from the 2019-2020 school year due to school emergency closure at the onset of the COVID-19 pandemic. A total of 683 students were followed for these three testing periods, including 468 students in the Treatment schools and 218 students in the Control schools. These students were followed for their third to fifth grade years. Preliminary analyses found no significant pretest differences ($\alpha = .05$) between missing (those who did not attend their respective schools for the entire 2017-2019 periods) and non-missing participants (those who constantly attended their respective school during the 2017-2019 periods) on all of the study's measures of interest. These findings suggested that data may be missing at random (Little & Rubin, 1987). Therefore, missing data imputation was deemed unnecessary.

Student demographic data is provided in Table 1.

Table 1
Student Demographics by Condition

	Students			
	Hispanic		Not Hispanic	
Condition	Male	Female	Male	Female
Control	88	85	26	19
Treatment	204	195	32	37
Total	292	280	58	55

Average scale scores on both exams were compared between the two years using a matched pairs t-test and the gains were indeed significant ($ps<.001$). Cohen's D effect size was calculated as 1.63 for reading and 1.76 for math. These are both large effects. Results are provided below in Table 2.

Table 2
Matched Pairs t-test Results

	Mean	Standard Deviation	t-value	Degrees of Freedom	p-value	Effect Size
	Reading					
Year 0	1372.60	155.09				
Year 2	1563.10	137.63				
Diff	*190.51*	*116.76*	*42.74*	*685*	*<.001***	*1.631*
	Math					
Year 0	1439.82	136.62				
Year 2	1626.97	138.79				
Diff	*187.15*	*106.12*	*46.19*	*685*	*<.001***	*1.76*

The overall effectiveness of the SEP on student outcomes is impressive. Students from all schools involved in the SEP gained an average of over 190 scale points on the STAAR Reading Exam, an increase of over 1.2 standard deviations from their average baseline scores. The STAAR Math Exam yielded similar results with students averaging a 187 scale point increase over the two years. This is an increase of 1.38 standard deviations from their average baseline scores. As documented by the Institute of Educational Sciences (IES) Expert Panel, turnaround schools are defined as showing substantial gains if they increase their standardized reading or math scores by an average of 10 percentage points or approximately .25 standard deviations within three years (Aladjem et al., 2010). Schools involved in our SEP showed increases that were far greater than this definition of "substantial gains" in only two years.

Scale score data for the three testing periods from both the math and reading exams was analyzed longitudinally using latent growth modeling with ethnicity and gender covariates. The subsequent models (Model 1: Reading and Model 2: Math) are given by:

$$Y_{ti} = \beta_{00} + \beta_{10}Time_{ti} + \beta_{01}Hispanic_i + \beta_{02}Female_i + \beta_{11}Hispanic_i * Time_{ti} + \beta_{12}Female_i * Time_{ti} + \mu_{0i} + \mu_{1i}Time_{ti} + e_{ti}$$

Results from the analysis are provided in Table 3.

Table 3
STAAR Reading and Math Latent Growth Models Results

	Reading			Math		
Fixed Effect	Coefficient	Standard Error	P-value	Coefficient	Standard Error	P-value
Intercept (β_{00})	1378.63	15.51	<.001**	1457.62	13.85	<.001**
Time (β_{10})	97.52	5.9	<.001**	93.85	5.36	<.001**
Hispanic (β_{01})	-24.65	15.72	.121	-11.97	14.00	.392
Female (β_{02})	22.84	11.7	.051^{+}	-14.57	10.41	.162
Hispanic*Time (β_{11})	-1.48	5.98	.804	-6.22	5.42	.251
Female*Time (β_{12})	-.910	4.45	.838	9.47	4.03	.019*

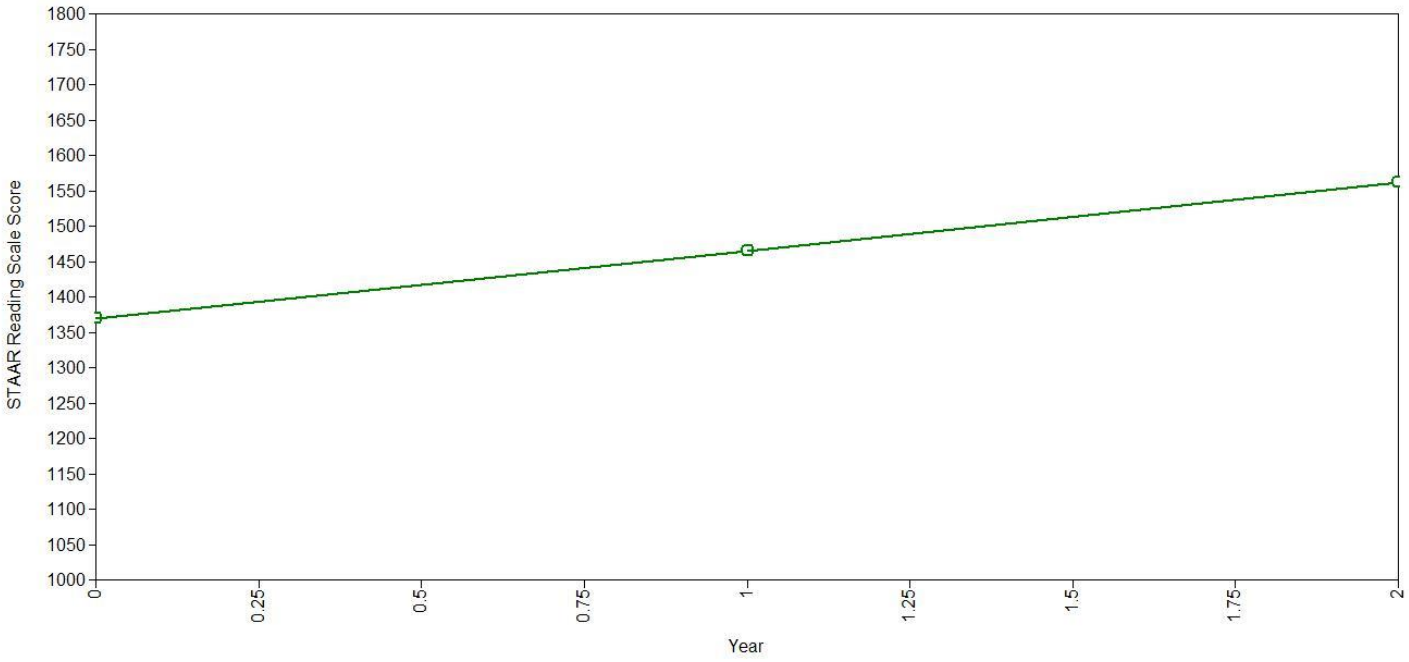
Note: **p<.001, *p<.05, $^{+}$p<.1

Both linear models fit the data adequately well noting that the Chi-Square test is subject to sample size and not always a good indicator of model fit (Reading: Chi-Square p-value<.001, RMSEA=.093, CFI=.962 and SRMR=.032; Math: Chi-Square p-value<.001, RMSEA=.104, CFI=.957 and SRMR=.039).

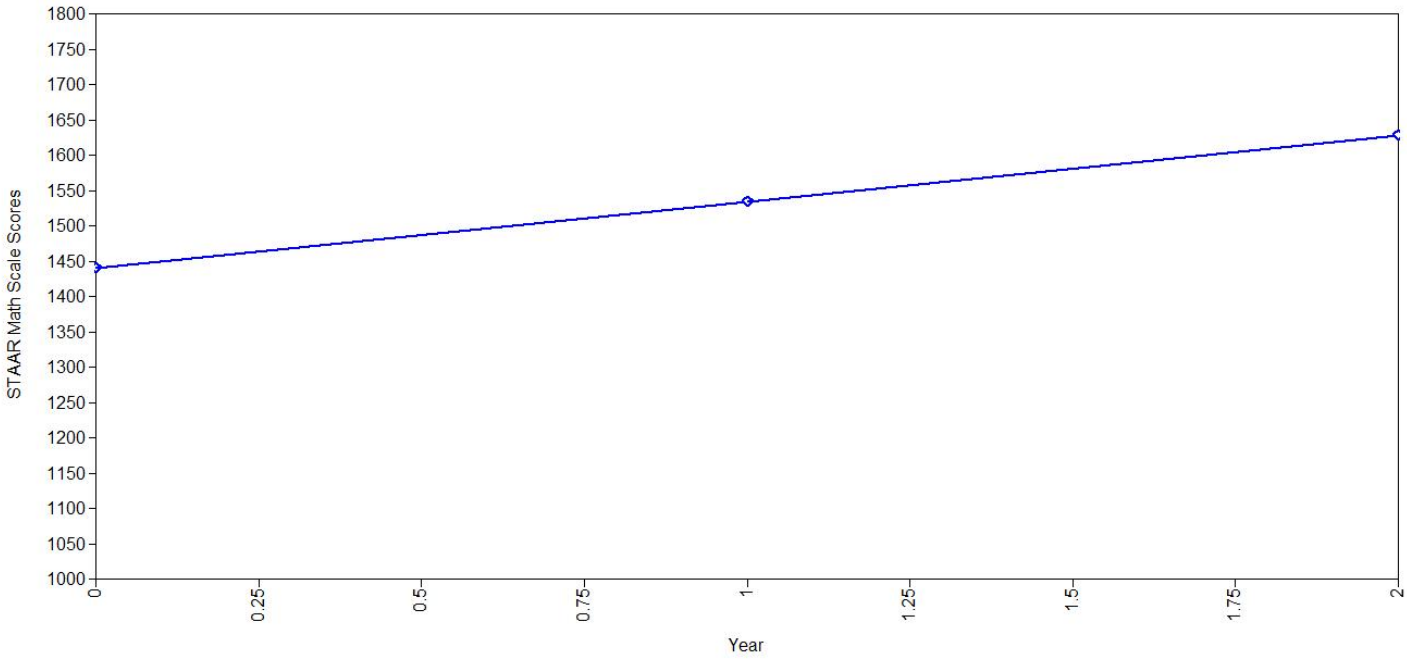
The overall effectiveness of the SEP is again substantiated with the significant intercepts and time (slope) variables. This shows that students on average scored 1378.63 in reading for Year 0 (2017) and that their scale score increased 97.52 points per year, on average ($p<.001$). Furthermore, students on average scored 1457.62 in math for Year 0 (2017) and their scale score increased 93.85 points per year, on average ($p<.001$). These results are illustrated in Graph 1 and Graph 2. Interestingly, females though initially scoring lower than males in math at Year 0, increased at a greater rate than males over the two years ($p=.019$). No differences in scale scores due to ethnicity were detected.

Graph 1

STAAR Reading Scale Scores Latent Growth Model



Graph 2
STAAR Math Scale Scores Latent Growth Model



In comparing the Treatment and Control schools, Treatment schools yielded higher gains on STAAR Math scale scores. Treatment school students were also more likely than Control school students to increase their STAAR Math proficiency rating by one-level or to attain a Masters proficiency rating after the two years. They were also more likely than Control school students to increase both their STAAR Reading and Math proficiency ratings by two-levels or to attain a Masters proficiency rating on these exams after two years.

A multilevel logistic model was used to determine if these increases in proficiency ratings were significant. A Hispanic, Female and baseline proficiency rating (Prof_Pre) were used as covariates in the model. Treatment was also added as a school-level variable. Model 3's outcome was rated as 1 if the student increased by at least one proficiency level or Mastered by the end of the second year, and 0 otherwise. Model 4 is identical to Model 3 except the outcome measure was rated as 1 if a student increased by at least two proficiency levels or received a proficiency rating of Masters by the end of the second year, and 0 otherwise. The full multilevel logistics models (Model 3 and 4) are given below:

$$Y_{ij}^{*} = \gamma_{00} + \gamma_{10}Prof\_pre_{ij} + \gamma_{20}Hispanic_{ij} + \gamma_{30}Female_{ij} + \gamma_{01}Treatment_{j} + e_{ij} + \mu_{0j},$$

$$where\ Y^{*} = \ln(odds) = \ln\left[\frac{P(Y_{ij}=1)}{1-P(Y_{ij}=1)}\right].$$

Results of Model 3 and Model 4 are provided in Tables 4 and 5 below.

Table 4
Model 3 Results: STAAR Reading and Math Proficiency Increase of at Least One

	Reading			Math		
Fixed Effect	Odds Ratio	Standard Error	P-value	Odds Ratio	Standard Error	P-value
Intercept (γ_{00})	1.36	.3438	.218	1.09	.2579	.712
Prof_pre (γ_{10})	1.77	.3575	.005**	3.741	.8905	<.001**
Hispanic (γ_{20})	.6718	.1493	.073^{+}	.6006	.1297	.018*
Female (γ_{30})	.9562	.1480	.772	1.34	.2111	.066^{+}
Treatment (γ_{01})	.8587	.1699	.441	1.06	.1806	.727

Note: **p<.001, *p<.05, $^{+}$p<.1

Table 5
Model 4 Results: STAAR Reading and Math Proficiency Increase of at Least Two

	Reading			Math		
Fixed Effect	Odds Ratio	Standard Error	P-value	Odds Ratio	Standard Error	P-value
Intercept (γ_{00})	.2446	.0689	<.001**	.4190	.1179	.002*
Prof_pre (γ_{10})	8.02	1.73	<.001**	13.00	3.23	<.001**
Hispanic (γ_{20})	.7049	.1758	.161	.4562	.1170	.002*
Female (γ_{30})	.9649	.1889	.855	1.23	.2345	.284
Treatment (γ_{01})	1.08	.2285	.721	1.04	.2131	.860

Note: **p<.001, *p<.05

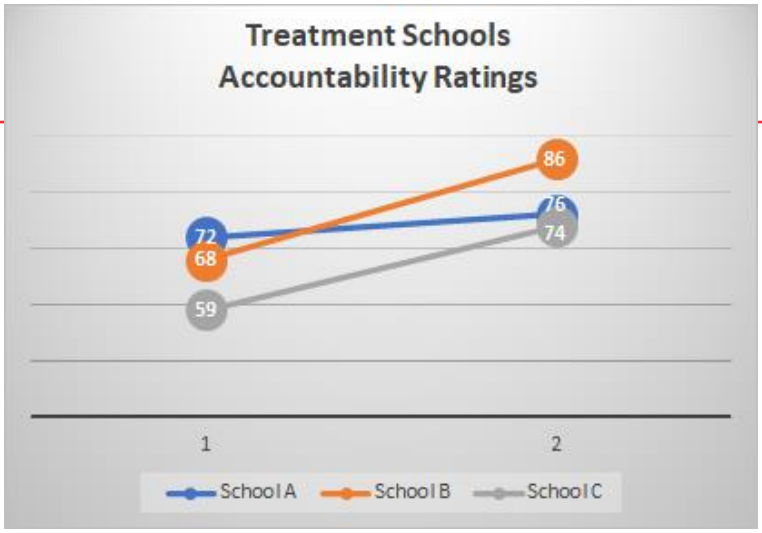
Findings again highlight Treatment school students being more likely than control students to increase proficiency ratings on three out of the four measures evaluated. Indeed, Treatment school students were 1.08 times more likely than control students to increase two proficiency levels in reading or to Master by the end of the second year. Though these findings were not statistically significant, they are noteworthy.

Effect on School Accountability Ratings

The state of Texas provides annual academic accountability ratings to all public school districts, charters and schools. The accountability ratings are based on more than just student performance on standardized tests (STAAR). These ratings are based on a combination of performance on state standardized tests, graduation rates, and college and career readiness outcomes. The accountability ratings examine student achievement, school progress, and if schools/districts are closing achievement gaps amongst various student demographic groups (Texas Education Agency, 2021).

Texas public schools implemented an A-F scale rating for the first time during the 2017-2018 academic year. Results for the 2017-2018 and 2018-2019 school years are provided below for Treatment (Graph 3) and Control schools (Graph 4).

Graph 3 and Graph 4
Treatment and Control Schools Accountability Ratings



As evidenced in the graph, all Treatment schools showed substantial gains in their accountability ratings over the two academic years. Likewise, one of the Control schools showed substantial gains in their accountability rating as well. However, two of the Control schools demonstrated a slight decline in their accountability ratings over the two academic years. Both schools scored a 76 in 2017-2018 and declined to a 75 in 2018-2019. Though this is not significant, it is important to note the differences in growth of the accountability ratings between Treatment and Control schools. This signifies the importance of the Treatment intervention, the mentor coaching, on overall school performance.

Conclusion

Through implementation of various diagnostic measures (e.g., root cause analysis) and interventions (e.g., SLI, coaching and mentoring), A-PLUS has shown that "school turnaround" is multi-faceted concept that requires a comprehensive approach.

To capture the true sense of "school turnaround", component 5 of A-PLUS employed a systematic approach consisting of (a) a Root Cause Analysis (RCA) conducted on all 8 schools—with a report for each school; (b) a SLI after the RCA was conducted (c) on-going training and implementation for the Treatment leadership team on Virtual Mentoring and Coaching real-time so that they can calibrate their observations and provide uniform feedback including, supervisory functions, instructional capacity, and instructional transformation; and (d) an on-site coach for 10 hours per week per Treatment campus.

The collective results indicated substantial improvement on student outcomes and school accountability ratings resulting from the appropriate diagnostic measures as well as interventions employed to address the factors that kept the participating schools from reaching their full potential. It is, therefore, concluded that A-PLUS's success is particularly due to its comprehensive approach to school turnaround; one that broadly considers the triangle of school leaders, teachers, and students.

References

Aldajem, D. K., Birman, B. F., Orland, M., Harr-Robins, J., Parrish, T. B. & Ruffini, S. J. (2010). *Achieving dramatic school improvement: An exploratory study*. Washington D. C.: US department of education.

Ayescue, J. B. & Siegel-Hawley, G. (2019). Magnets and school turnarounds: Revisiting policies for promoting equitable, diverse schools. *Education Policy Analysis Archives, 27*(72).

Brown, G., & Irby, B. J. (2001). *The principal portfolio (2nd ed.)*. Corwin Press.

Day, C., Gu, Q., Sammons, P. (2016). The impact of leadership on student outcomes: How successful school leaders use transformational instructional strategies to make a difference. *Educational Administration Quarterly, 52*(2), 221-258.

Dufour, R. & Mattos, M. (2013). How do principals really improve schools? *Educational Leadership, 70(7), 34-41*.

Huberman, M., Parrish, T., Hannan, S., Arellanes, M., & Shambaugh, L. (2011). Turnaround schools in California: Who are they and what strategies do they use? (ED592859). ERIC. https://eric.ed.gov/?id=ED592859

Irby, B. J., & Brown, G. (1999). A Gendered Dichotomy in Written Reflections in Professional Development Portfolios. Paper presented at the Annual Meeting of the American Educational Research Association, Montreal, Canada.

Irby, B. J., Brown, G., Duffy, J. A., & Trautman, D. (2002). The synergistic leadership theory. *Journal of Educational Administration, 40(4),* 304-322.

Irby, B., J., Lara-Alecio, R., Tong, F., and Torres, M. (2017). Accelerated Preparation of Leaders for Underserved Schools (A-PLUS): Building Instructional Capacity to Impact Diverse Learners. Project Sponsored by the Supporting Effective Educator Development Grant Program. (SEED), US Department of Education.

Irby, B. J. & Lunenburg, F. C. (2014). Promoting ethical behavior for school administrators. *Journal of Ethical Educational Leadership, 1*(4), 1-28.

Lara-Alecio, R., Tong, F., Irby, B. J., & Mathes, P. (2009). Teachers' pedagogical differences during ESL block among bilingual and English-immersion kindergarten classrooms in a randomized trial study. *Bilingual Research Journal, 32*(1), 77-100.

Leithwood, K. A., & Riehl, C. (2003). *What we know about successful school leadership*. Nottingham: National College for School Leadership. Philadelphia: Laboratory for Student Success, Temple University.

Little, R. J. A., & Rubin, D. B. (1987). *Statistical analysis with missing data.* New York: Wiley.

Sun, J. & Leithwood, K. (2014). Direction-setting school leadership practices: a meta-analytical review of evidence about their influence. *School Effectiveness and School Improvement, 26,* 499-523.

Texas Education Agency. (2021). *Academic Accountability*. https://tea.texas.gov/texas-schools/accountability/academic-accountability

Tschannen-Moran, B. & Tschannen-Moran, M. (2011). Evaluation and coaching can work at cross-purposes if schools blur the distinctions between them. *Educational Leadership, 69(2),* 10-16.