Reenergizing Leadership to Achieve Greater Student Success

EIR EARLY04 Grant Implementation & Impact Findings Report

May 15, 2023

Submitted to:
Intercultural Development Research Association (IDRA)
5815 Callaghan Road, Suite 101, San Antonio, TX 78228

Prepared by:
Brian Freeman, Supriya Tamang, Jesse Wood, & Liz M.B. Yadav

Submitted by:
Abt Associates
6130 Executive Boulevard, Rockville, MD 20852

About This Report

This report provides findings from the Abt Associates (Abt) evaluation of the Reenergizing Leadership to Achieve Greater Student Success (Reenergize) program. The evaluation was funded by the Intercultural Development Research Association (IDRA) Education Innovation and Research (EIR) Early Phase grant (EARLY04; Award No. U411C170153).

The views expressed in this report do not necessarily reflect the views or policies of the U.S. Department of Education.

Acknowledgements

We would like to extend our sincere thanks to the many individuals who contributed to the Reenergize evaluation.

Our work would not have been possible without the guidance and support of many individuals at IDRA, including Dr. Nilka Avilés and Hector Gonzalez, who led the implementation of the Reenergize program and provided information to the evaluation team. Their cooperation and willingness to participate in the evaluation have been invaluable. We also thank the IDRA consultants and local teachers and administrators who participated in interviews to help inform our understanding of the Reenergize program's implementation.

We also wish to acknowledge our colleagues at Abt who contributed to this evaluation and report: Jill Cusick, Raquel González, Jennifer Poulos, and former Abt staff members Rebecca Jackson, Karen Loya, Alyssa Rulf-Fountain, Melissa Velez, and Rian Watt contributed to the evaluation design, data collection and analysis; Daniel Litwok served as the Project Quality Reviewer; and Bry Pollack provided editing support.

Lastly, we are grateful for the EIR evaluation technical assistance provided by Geoffrey Borman (Arizona State University) and Jessaca Spybrook (Western Michigan University).

Contents

Exhibits

Executive Summary

The Intercultural Development Research Association (IDRA) was awarded an Education Innovation and Research (EIR) Early Phase grant from the U.S. Department of Education in 2017 to develop, implement, and test a promising program to improve student achievement and educational progress for high-need students: Reenergizing Leadership to Achieve Greater Student Success (Reenergize). This report provides findings from the EIR-funded Abt Associates implementation and impact evaluation of the Reenergize program.

Reenergize Overview

Reenergize is intended to improve the achievement and educational progress of high-need students by selecting, training, coaching, and empowering principals and school leadership teams to be effective advocates, collaborators, and risk-taking innovators. School leadership teams may include, but not limited to, principals, vice-principals, instructional coaches, and master teachers. Reenergize also aims to contribute to an environment in which principals and leadership teams are empowered to create a more open and equitable environment for their students. Reenergize is explicitly designed to focus on key issues that districts with high-need students require support in addressing: student attendance, reading and writing achievement, and college and career readiness.

The Reenergize program has six key activities:

1. Root cause analysis (using student achievement data to identify key problems)
2. Professional development for principals and leadership teams
3. Education for master teachers to become principals
4. Campus climate alignment efforts
5. Community partnerships
6. Professional learning communities

Implementation Overview

Reenergize was introduced beginning in school year 2019–20 in two public school districts serving high-need students around San Antonio, Texas. Within these two districts, Reenergize was introduced in ten schools: four elementary schools, four middle schools, and two high schools. All ten schools, serving a total of 8,464 students, were included in the study.

The students served by Reenergize schools are primarily Hispanic/Latino and economically disadvantaged (receive free or reduced-price lunch). Most of these schools also have a higher proportion of students with disabilities and English language learners compared with the national average. Nearly all Reenergize schools have student achievement scores below the state average and were identified in 2019 for either comprehensive or targeted support by the Texas Education Agency's Accountability Ratings.

As part of this study, IDRA intended to implement Reenergize in ten schools for three school years: 2019–20, 2020–21, and 2021–22. However, during Year 2 (2020–21), only seven of the initial ten intervention

schools participated in any Reenergize activities; during Year 3 (2021–22), only three intervention schools participated in any Reenergize activities.

Master teacher education was provided by the University of Texas at San Antonio (UTSA) as part of its principal preparation program. Over the three years of implementation, 13 master teachers from the ten intervention schools participated in the UTSA principal preparation program. Two IDRA consultants who are former local teachers with experience in reading and writing instruction provided program-specific teacher professional development and one-on-one coaching. The professional development and coaching activities included an equity audit, co-teaching, modeling, lesson planning, and observations. The leadership team and master teachers received professional coaching services both individually and as a group to support the implementation of the school's action plan. IDRA staff, who are predominately Hispanic/Latino, led or provided the other program activities (root cause analysis, campus alignment efforts, professional development for leadership teams, and establishing community partnerships and professional learning communities).

Over the three years of implementation, more than 60 principals and leadership team members received professional development services for a total of more than 250 hours of professional development over 13 different modules. The modules covered topics such as IDRA's Quality Schools Action Framework™, asset-based practices, teacher diversity, family leadership, bullying and harassment, early childhood development, and bilingual education in addition to topics around the three key district issues of student attendance, reading and writing achievement, and college and career readiness.

Implementation Fidelity Findings

The implementation study describes the extent to which the six key Reenergize program activities were implemented with fidelity in study schools during the evaluation period (school years 2019–20, 2020–21, and 2021–22).

Overall program fidelity for the full sample of intervention schools was not met for the three years of the Reenergize intervention implementation, as the intervention schools did not meet the fidelity thresholds set for the study for all six key activities each year.[1]

- **In Year 1**, Reenergize met the fidelity threshold for one key activity, professional learning communities, but did not meet the threshold for the other five activities. Nonetheless, more than half of the ten schools implemented these five activities with fidelity.

- **In Year 2**, Reenergize did not meet the fidelity thresholds for any of the key activities. However, at least half of the schools implemented root cause analysis, campus climate alignment efforts, and professional learning communities with fidelity.

- **In Year 3**, fewer than half of the schools implemented the key activities with fidelity.

[1] Indicator scores were defined for each key activity to reflect adequate implementation of the Reenergize model (See Appendix A). To meet the *fidelity threshold,* 86 percent or more of schools had to receive the highest indicator scores for each activity.

Relatively few intervention schools experienced the full Reenergize program as designed. The impact study includes all ten intervention schools regardless of which activities they implemented or to what degree, so its results represent the effect of a partial implementation of Reenergize. It is worth noting that the low fidelity scores for each year could be partially attributed to various factors, such as changes in school administration and school restructuring, which led to schools not participating in the intervention at all. Furthermore, the fidelity findings might not present a complete picture of the Reenergize implementation, particularly given the challenges posed by the COVID-19 pandemic, which began at the end of Year 1 (2019–20). Challenges such as teacher turnover, changes in school administration, food insecurity, and limited access to critical services have all contributed to the difficulties faced by schools in implementing the Reenergize intervention.

Impact Study Design

The impact study examines the effect of Reenergize on student academic success outcomes: reading/writing and math achievement, high school graduation, college and career readiness, postsecondary enrollment, and attendance. Student outcomes were targeted in the impact study rather than principal or teacher outcomes because student outcomes can be readily measured by instruments that meet What Works Clearinghouse outcome measure standards. Each outcome was estimated from publicly available school-level data from the Texas Education Agency (TEA) spanning school years 2014–15 through 2021–22.

To estimate the impact of the Reenergize program, we designed an impact study where we matched intervention schools that agreed to implement Reenergize to a set of comparison schools that did not initially participate in Reenergize. The impact study is quasi-experimental because we did not use a random process to decide which schools would participate in the Reenergize intervention and which schools would be in the comparison group. Comparison schools conducted business-as-usual instruction, systems, professional development, and leadership. Notwithstanding our design, however, three comparison schools in one of the intervention districts ended up receiving the Reenergize program in the second and third years of the evaluation.[2] These three schools remain in our comparison sample because our analysis follows an intent-to-treat approach; thus, our intervention group is limited to schools that agreed to implement Reenergize at the beginning of the study. All other comparison schools never had the opportunity to participate in Reenergize during the evaluation period.

Impact study research questions were designated as confirmatory or exploratory prior to analysis. Confirmatory outcomes are the main outcomes of interest for an evaluation for which favorable impacts signal the program is meeting its goals. The confirmatory outcomes in this impact study are the most precise measure in each outcome domain estimated after the full three years of Reenergize implementation. Several confirmatory research questions are not evaluated in this report because the needed public use data was not available prior to the analysis and release of this report. Exploratory outcomes are either less precise measures in an outcome domain (e.g., a binary threshold rather than a

[2] One school participated in the model with full fidelity during school year 2021–22, one school participated in the model with some fidelity, and in the third school, one master teacher participated in the education for master teachers.

continuous scale score), are measured at an interim time point (two years of Reenergize implementation) or are longer-term outcomes where impacts may not be observed during the evaluation period.

We use a comparative short interrupted time-series (C-SITS) model with baseline mean projection to estimate the impacts of the Reenergize intervention.[3] For the quasi-experimental impact study to estimate an unbiased impact of Reenergize, post-intervention outcome measures for comparison schools should be reasonable counterfactuals for the intervention schools (meaning the counterfactuals represent what would have happened in the intervention schools if they had not been selected to participate in the Reenergize program).[4] Assessing the similarity of the intervention and comparison schools on pre-intervention measures, such as student demographics, TEA school Accountability Ratings, and outcomes of interest is one way to assess whether outcomes for comparison schools are a reasonable counterfactual.

The baseline differences in pre-intervention outcomes between intervention and comparison schools are either below or within the adjustment range according to What Works Clearinghouse Standards (Versions 4.1 and 5.0) for all outcomes. In both the intervention and comparison schools, students are primarily Hispanic/Latino, economically disadvantaged (receive free or reduced-price lunch), and are more likely to be classified as English language learners than in average U.S. schools.

Impact Study Findings

- With the exception of attendance in Year 2 (2020–21), there were no statistically significant differences in educational outcomes based on a school's participation in the Reenergize program.
- The Reenergize program had no detectable impact on student achievement (reading/writing and math standardized test scores), high school graduation, college and career readiness, or postsecondary enrollment.

Given the small sample size of assigned units (schools), the evaluation's ability to detect small statistically significant effects is limited. This is particularly true for the high school and postsecondary outcomes, as only two high schools participated in the Reenergize intervention. The relatively low level of implementation fidelity and limited participation of many intervention schools in Reenergize suggest that the intervention contrast was relatively small, limiting this evaluation's ability to detect significant impacts. In addition, the standard errors of some impact estimates are large relative to the scales of the measures, which would preclude detection of effects of small to moderate magnitude.

---

[3] A C-SITS model with baseline mean projection estimates how much more (or less) outcomes changed in the intervention group than the comparison group between the years before the study and the years of the study. It is equivalent to a difference-in-differences model.

[4] Establishing baseline equivalence mitigates the concern for selection bias distorting the estimated impact of the program on student outcomes. Selection bias is the risk that intervention schools are not similar to comparison schools at the beginning of the study, so the comparison schools cannot be reasonable counterfactuals. Additional assumptions required for the estimate to be unbiased include: conditional on the factors within our model, the outcome is independent of a school's intervention status, the instrumentation of each outcome remains constant over the study period, and no other policies were introduced during this time that target the outcomes of interest.

Conclusion

Through the Reenergize program, IDRA aimed to enhance instructional leadership and management skills of school personnel, which it hypothesized would lead to increased student achievement, increased college and career readiness, and improved rates of attendance, high school graduation, and postsecondary enrollment.

The COVID-19 pandemic, combined with high rates of teacher and administration turnover, severely attenuated the implementation of the Reenergize program. Only seven of the initial ten intervention schools participated in any Reenergize activities during Year 2 (2020–21), and only three of the initial ten intervention schools participated in any Reenergize activities during Year 3 (2021–22). Consequently, the Reenergize program did not meet any of the program-level thresholds for adequate fidelity of implementation in those years, and none of the impact estimates for confirmatory student outcomes was statistically significant.

Despite the challenges IDRA faced implementing Reenergize, IDRA learned valuable lessons that can inform future efforts. Assessing stakeholder capacity and aligning the program to fit within this capacity, cultivating district and school champions to advocate for and support the program, and flexibly adapting to shifting stakeholder needs are crucial when implementing a program in this context. IDRA was able to develop and maintain effective partnerships in some Reenergize schools by building on existing relationships. The effective planning and collaboration from school administrators allowed the Reenergize coaches to integrate into these schools smoothly. As IDRA continues to support schools and districts around San Antonio, the findings from the Reenergize study and the lessons learned can inform future implementation and evaluation efforts.

1. Reenergize Overview

The need for effective and equitable instructional leadership is more evident than ever (Theoharis & Brooks, 2012) as schools are becoming increasingly diverse, and diverse schools that fail to meet state standards are on the rise. Nationally, the rate of school-age children who were White decreased by more than 10 percentage points between 2000 and 2017 (from 62 to 51 percent), while the percentage of Hispanic, Asian, and multiracial school-age children increased (de Brey et al., 2019).[5] Alongside a diversifying student population, there is a shortage of principals and teachers, particularly in urban and rural schools that serve predominately racial/ethnic minority student populations. In school year 2011–12, more than two-thirds of schools nationwide had at least one teaching vacancy, and high-poverty and high-minority schools continue to face persistent staffing challenges (Aragon, 2016). Principals and school leadership teams also face numerous challenges working with diverse student populations. Racial/ethnic minority students and low-income students are underrepresented in high school graduation and postsecondary enrollment (National Center for Education Statistics, 2019, Tables 302.20 and 302.44)[6] whereas minority students are overrepresented in special education and disciplinary actions (National Center for Education Statistics, 2019, Tables 204.40 and 233.28).[7]

One strategy to address these needs and challenges is to have culturally proficient administrators, aspiring administrators, and teachers in low-performing schools who are visionary leaders, student advocates, instructional leaders, collaborators, and risk-taking innovators.

Principals and school leadership teams are immensely important to student success. Research suggests that school leadership's impact on student academic success is second only to the classroom and teacher among school-related factors that influence student learning (Leithwood et al., 2008; Wallace Foundation, 2013). Roughly one-quarter of student academic achievement is due to the school leadership alone, compared to the classroom instruction (including the teacher), which accounts for one-third to one-half of the noticeable differences in academic achievement (Hallinger & Heck [1998] as cited in Leithwood et al., 2008; Marzano et al. [2005] as cited in Perilla, 2014). The school leadership's effect on student achievement is likely due to their influence on the school learning environment and classroom practices. Further, school leadership is key to the professional growth of school staff and for retaining teachers (Kraft et al., 2016; Learning Policy Institute, 2017; Wallace Foundation, 2013). For example, the principal is key to helping instructional coaches establish a climate for professional growth and for classroom success (Heineke & Polnick, 2013) and for creating an instructional climate that allows teachers and staff to improve their instruction and become leaders (Wallace Foundation, 2013). Evidence from New York City suggests

[5] This data combines race and ethnicity. Note that racial categories (e.g., White, Black, Asian, multiracial) in this data exclude children of Hispanic ethnicity.

[6] In 2018, around 65% of Black and Hispanic high school graduates enrolled in postsecondary education compared to 71% of white high school graduates (Table 302.20). 30 percent of low-income 9th graders (those in the lowest fifth of the income brackets in 2009) ever enrolled in college (Table 302.44).

[7] Around 16% of Black students and 18% of American Indian students receive special education services compared to 14% of White students and 13% of Hispanic students (Table 204.40). Black students are more than three times as likely to receive a suspension or have a school-related arrest and are twice as likely to be expelled compared to their white counterparts (Table 233.38).

that teacher turnover is twice as high when leadership is not seen as strong, supportive, and effective and that leadership and professional development are the most dominant school context factors that affect turnover (Kraft et al., 2016).

Reenergizing Leadership to Achieve Greater Student Success (Reenergize), developed by the Intercultural Development Research Association (IDRA) through an Education Innovation and Research (EIR) Early Phase grant (EARLY04) from the U.S. Department of Education in 2017, is a culturally proficient, specific transformation intervention for administrators, aspiring administrators, and teachers in low-performing schools. Reenergize is intended to improve student achievement and educational progress by selecting, training, coaching, and empowering principals and school leadership teams to be effective advocates, collaborators, and risk-taking innovators. Reenergize also aims to contribute to an environment in which principals and leadership teams are empowered to create a more open and equitable environment for their students. Reenergize is explicitly designed to focus on district needs centered around student attendance, reading and writing achievement, and college and career readiness.

1.1 Program Model

The Reenergize program model is rooted in theory and research on successful strategies that have been strategically positioned to have the greatest impact in schools that need improvement according to state standards. This section describes the theory of the six key activities that drive the program model as theorized by IDRA in their EIR grant application. Section 3.1 describes the Reenergize intervention as actually implemented in the evaluation period.

1. **Root cause analysis.** First, IDRA and participating school leadership teams conduct root cause analysis using student achievement data to identify key problems that have presented themselves among the three key district issues (student attendance, reading and writing achievement, and college and career readiness). The results of the issue analysis guide the content of the technical assistance provided by IDRA.

   The program design incorporates five practices recommended by the What Works Clearinghouse (WWC) for using achievement data to assess and improve student instruction (Hamilton et al., 2009). Those WWC recommendations are to (1) "make data part of an ongoing cycle of instructional improvement" including using a variety of data sources, (2) involve "students to examine their own data and set learning goals", (3) "establish a clear vision for schoolwide data use" including "collaboration across and within grade levels and subject areas,"(4) invest in "supports that foster a data-driven culture within the school," and (5) "maintain a districtwide data system". In particular, Reenergize uses multiple data sources as part of the root cause analysis to inform the school's vision and improvement process, and involves students and families so that they can set their own goals.

   The design of the Reenergize intervention also incorporates four practices the WWC suggests are linked with turnaround of low-performing schools (Herman et al., 2008): (1) "signal[ing] the need for dramatic change" including the urgency for change, (2) "maintain[ing] a consistent focus on improving instruction," (3) celebrate "visible improvements early in the…process (quick wins)" to inspire and gain the collaboration of staff, and (4) "build a committed staff" who are open to change.

2. **Professional development for principals and leadership teams.** Second, as a group who shares collective responsibility for systemic effectiveness and student learning, current principals, the school leadership team, and master teachers receive professional development. IDRA provides joint training to principals and their leadership teams; bespoke coaching for each leadership team and master teachers; experiences in the design, implementation, and evaluation of schoolwide projects to all participants; and myriad opportunities to build trust as a group and understand their role as a leadership team. School leadership teams may include, but not limited to, principals, vice-principals, instructional coaches, and master teachers.

   Providing professional development to current principals and the leadership team together is a powerful way to develop collective administrative efficacy. IDRA's professional development focuses on the underlying pedagogical and institutional changes required to implement practices around equity and equality to counteract systemic inequalities and oppressive behaviors (Richardson, 2010). Thus, each school leader can become an agent for social change.

   Historically, few effective leadership development models, have integrated the services of a leadership coach (Psencik, 2011). Leadership coaches utilize individualized and group coaching strategies that provide opportunities to demonstrate and implement effective instructional and leadership practices. Given the important role coaches can play, Reenergize includes coaching services aimed at improving systemic effectiveness and student instruction for the leadership team and master teachers individually and as a group. The coach also aims to enable and consolidate the integration of theory into practice by providing advice or alternative solutions for consideration. The combination of formal professional development and coaching allows school leaders to learn about, plan for, and implement effective practices in their schools.

3. **Education for master teachers to become principals.** Third, participating school districts select, prepare, and certify master teachers to become principals of diverse schools that serve minority and high-need students. Master teachers are ideal candidates to move into administrative roles where they can strengthen and develop the leadership competency skills needed to become an effective principal for struggling schools (Wallace Foundation, 2013). Jackson & McDermott (2012) are unequivocal in their view that almost any teacher can become a master teacher under the right instructional leader if the instructional leader works with and through other people to accomplish the vision and goals of the institution. The selection and training of new principals is consistent with the participating districts' plans for an enhanced leadership development pipeline.

4. **Campus climate alignment efforts.** Fourth, the program supports campus climate alignment efforts intended to align school climate with social justice and equity principles. Changing a school's culture is not an easy task as it often involves uprooting values, myths, and ways of operating deeply embedded in long-held traditions. IDRA coaches each school's leadership team through the process of building a school culture aligned to district-level priorities. Each school reimagines and takes steps to implement instructional strategies where (1) implied factual knowledge (content mastery) includes a range of ideas and perspectives and not just the mainstream ideology; (2) content knowledge (critical thinking) is consistently open to debate and critique from multiple points of view, which may spur change processes; (3) students can use intergroup dialogue, literature, and writing to propel action and social change; (4) personal

reflection by teachers and leaders to inform future practice becomes common practice; and (5) teachers and school leaders take into account awareness of multicultural group dynamics, not only on the content to be taught but on the demographics of the school and of the school community (Hackman, 2005).

5. **Community partnerships.** Fifth, the program establishes community partnerships with parents and community stakeholders to share the school's vision of success and collectively partner with the school to improve the achievement of all students. Principals examine ways to engage parents in school decision making processes and community collaborative partnerships that assist in improving and cultivating support efforts for improved outcomes in low-performing schools. Parent and community organizing can take charge in holding schools accountable for educational outcomes (Henderson & Mapp, 2002). The Reenergize program builds upon IDRA's Education CAFE™ framework[8] to emphasize the importance of empowering families through a community-based organization to improve the educational process in the three key outcomes through leadership, capacity-building, and accountability to support student achievement and close existing gaps. An Education CAFE is a parent group that is rooted in a community-based organization, rather than in a single school. To establish an Education CAFE, parents hold initial planning and organizational meetings, and then establish partnerships with their school community. As part of the CAFE, parents and school partners collaboratively work on activities to improve student success and education.

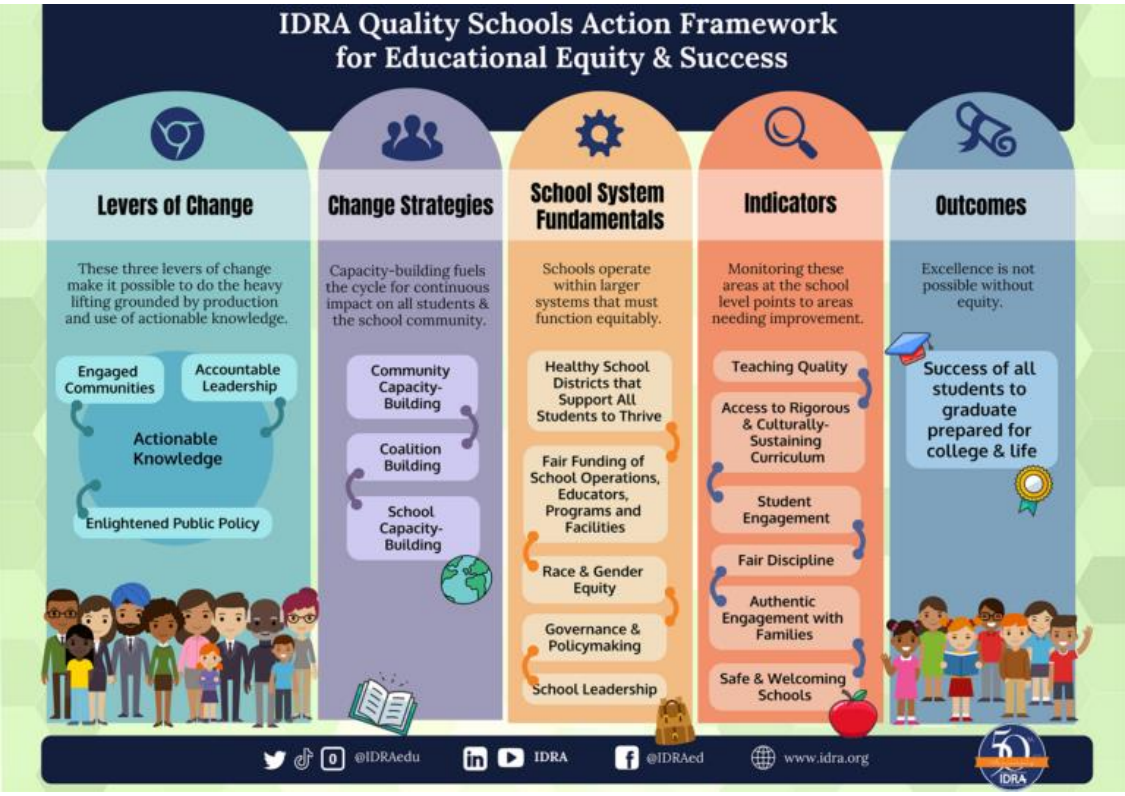
6. **Professional learning communities (PLCs).** Sixth, the program supports current principals to establish and facilitate a schoolwide professional learning community that will share a collective responsibility for curriculum and instructional effectiveness and student learning. IDRA assists school leadership teams in the establishment and maintenance of PLCs, a well-established peer coaching model, to facilitate peer-to-peer learning around the three key issues.

IDRA's Quality Schools Action Framework™, as shown in Exhibit 1, underpins all of IDRA's efforts, including the Reenergize program.

Further, all of IDRA's efforts are rooted in asset-based practices that assume all students, regardless of race/ethnicity and socio-economic background, are able to succeed and be prepared for college. IDRA has pioneered frameworks, systemic efforts and professional development around how instructional leaders, teachers and communities can create change based on asset-based, rather than deficit-based, practices. Asset-based institutional practices are primarily defined as educational practices in the classroom, school, and district that bank on student's funds of knowledge and capacities to be used in the educational process.

[8] This framework was developed under an i3 grant. See https://www.idra.org/families-and-communities/education-cafe/ for more information on the Education CAFE framework.

Exhibit 1. IDRA's Quality Schools Action Framework



Source: IDRA (n.d). https://www.idra.org/change-model/quality-schools-action-framework/.

1.2 Logic Model

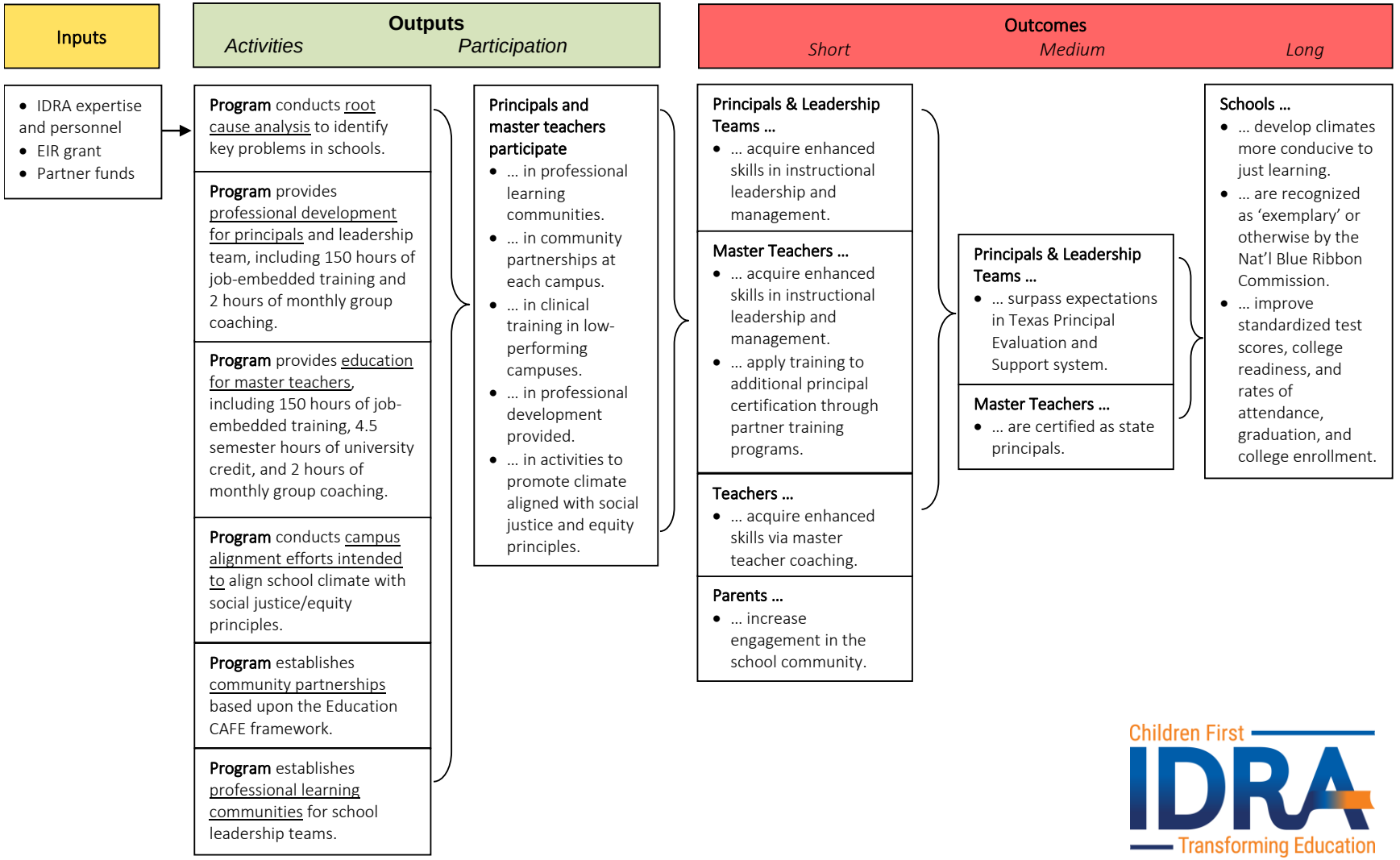
The Reenergize logic model (Exhibit 2 on next page) illustrates the inputs, outputs, and expected outcomes, moving from left to right. The inputs represent the elements that together support the Reenergize program model, including IDRA's expertise and personnel and financial supports (EIR grant and partner funds). The outputs are the six activities that embody the program strategies: root cause analysis, professional development for principals and leadership teams of participating schools and districts, education for master teachers, campus alignment efforts, establishing community partnerships, and establishing professional learning communities.

If the model is implemented with fidelity, the theorized short-term outcomes include enhanced instructional leadership and management skills of principals, leadership teams, and master teachers; additional principal certification for master teachers; and increased parental engagement in the school community. As a result of the short-term outcomes, it is theorized that principals and leadership teams would surpass expectations in the Texas Principal Evaluation and Support System, and that more master teachers are certified as state principals. The theorized longer-term outcomes include developing school climates more conducive to just (equitable) learning as well as increased student achievement, as captured by state standardized test scores; increased college and career readiness; and improved rates of attendance, graduation, and postsecondary enrollment.

For the impact study, we evaluate the long-term outcome of student academic success, as it can be readily measured by instruments that meet WWC outcome measure standards.

Exhibit 2. Reenergize Logic Model

Reenergizing Leadership to Achieve Greater Student Success (Reenergize)



2. Evaluation Overview

The evaluation of Reenergize seeks to learn whether participation in the Reenergize program improves student outcomes. To do this, the evaluation includes two elements: a fidelity of implementation study and a study of impacts on student outcomes.

The implementation study (Section 3) helps us understand whether the program model was carried out with fidelity. The implementation study analyzes fidelity of implementation by year for the full sample of schools that started participating in Reenergize in 2019–20. In addition, the study documents additional implementation findings based on focus groups with teachers and Reenergize coaches. Findings from the implementation study provide important context for the evaluation of the program's impact on student outcomes.

The impact study, discussed in Section 4, is designed to assess five long-term student outcomes: English language arts and math achievement, graduation rates, college and career readiness, attendance, and postsecondary enrollment. To estimate the effect of participating in Reenergize, this quasi-experimental impact study relies on a school-level comparative short interrupted time series (C-SITS) that compares student outcomes for the intervention schools that agreed to implement Reenergize to a matched set of comparison schools that did not initially participate in Reenergize.[9]

2.1 Independence of the Evaluation

IDRA hired Abt Associates (Abt) as the independent external evaluator for the Reenergize program. IDRA was responsible for developing and implementing the Reenergize program and for collecting fidelity of implementation data. Abt collected the publicly available outcome data used in the impact analyses as well as qualitative implementation information. Abt independently conducted the implementation and impact analyses.[10]

2.2 Pre-Registration of the Study Plan

The evaluation study plan was pre-registered (Study No. 1806.1) in the Registry of Efficacy and Effectiveness Studies (REES).[11] Revisions to the analysis plan due to extensions to the implementation timeline and updates to the procedures for selecting comparison schools were documented in REES prior to any outcome data collection or analysis.

---

[9] The impact study is quasi-experimental because we did not use a random process to decide which schools would participate in the Reenergize intervention and which schools would be in the comparison group.

[10] Abt held the EIR Evaluation Technical Assistance Support contract that provided evaluation support to the 2017 cohort of EIR grantees. To avoid the appearance of a conflict of interest, the EARLY04 Abt evaluation team received support from non-Abt EIR Technical Assistance Liaisons at Arizona State University and Western Michigan University. No one on the EARLY04 Abt evaluation team had or has a role on the EIR Technical Assistance contract and did not have any information about other EIR grants or forthcoming resources and guidance not yet provided to other EIR grantees or their evaluators.

[11] The REES study plan for the Reenergize evaluation can be downloaded here: https://sreereg.icpsr.umich.edu/sreereg/subEntry/17801/pdf?section=all&action=download

Several pre-registered confirmatory research questions are not evaluated in this report because the public use data was not available prior to the analysis and release of this report (see Section 2.3).

The evaluation plan did not pre-register the fidelity of implementation study. However, the fidelity matrix was designed and drafted prior to the start of implementation.

2.3 Research Questions

Impact evaluation research questions were designated as confirmatory or exploratory prior to analysis. Confirmatory outcomes are the main outcomes of interest for an evaluation for which favorable impacts signal the program is meeting its goals. The confirmatory outcomes in this impact study are the most precise measure in each outcome domain measured after the full three years of Reenergize implementation.[12] Exploratory outcomes are either less precise measures in an outcome domain (e.g., a binary threshold rather than a continuous scale score), are measured at an interim time point (two years of Reenergize implementation), or are longer-term outcomes where impacts may not be observed during the evaluation period. Due to data availability, not all research questions can be addressed in this evaluation report. Asterisks (*) indicate research questions for which data was not available prior to the publication deadline for this report.[13]

The Reenergize intervention was implemented for three school years (2019–20 through 2021–22). The impact of participation in Reenergize on school attendance, student achievement, high school graduation, and college and career readiness are assessed using both **confirmatory (three-year impacts)** and **exploratory (two-year impacts)** analyses:

Confirmatory Research Questions

- Do elementary and middle schools that participate in Reenergize have higher reading scores (i.e., improved reading achievement) at the end of three years of intervention compared to schools not experiencing the intervention?
- Do elementary and middle schools that participate in Reenergize have higher math scores (i.e., improved mathematics achievement) at the end of three years of intervention compared to schools not experiencing the intervention?
- Do high schools that participate in Reenergize have higher graduation rates (i.e., improved student progression) at the end of three years of intervention compared to schools not experiencing the intervention? *

[12] All college enrollment outcomes were designated as exploratory, as three-year outcomes were not expected to be released in advance of the evaluation report publication deadline.

[13] All evaluation outcome measures are published by the Texas Education Agency (TEA). The evaluation team expected data to be available for all confirmatory and exploratory research questions based on the timing of TEA data publication in prior school years. However, school attendance and high school graduation measures for Year 3 (the 2021–22 school year) had not been published by TEA by January 2023.

- Do high schools that participate in Reenergize score higher on measures of college readiness (i.e., improved general achievement) at the end of three years of intervention compared to schools not experiencing the intervention?
- Do elementary, middle, and high schools that participate in Reenergize have higher mean rates of daily attendance (i.e. improved attendance) at the end of three years of intervention compared to schools not experiencing the intervention? *

Exploratory Research Questions

- Do elementary and middle schools that participate in Reenergize have higher reading scores (i.e., improved reading achievement) at the end of two years of intervention compared to schools not experiencing the intervention?
- Do elementary and middle schools that participate in Reenergize have higher math scores (i.e., improved mathematics achievement) at the end of two years of intervention compared to schools not experiencing the intervention?
- Do high schools that participate in Reenergize have higher graduation rates (i.e., improved student progression) at the end of two years of intervention compared to schools not experiencing the intervention?
- Do high schools that participate in Reenergize score higher on measures of college readiness (i.e., improved general achievement) at the end of two years of intervention compared to schools not experiencing the intervention?
- Do elementary, middle, and high schools that participate in Reenergize have higher mean rates of daily attendance (i.e., improved attendance) at the end of two years of intervention compared to schools not experiencing the intervention?

All analyses conducted around **proportions of students meeting Texas state standards** are considered **exploratory** because these are additional analyses of student achievement that utilize a binary threshold for achievement rather than a continuous raw achievement score.

Exploratory Research Questions (Alternate Measures of Student Achievement)

- Do elementary and middle schools that participate in Reenergize have higher proportions of students meeting Texas state standards (i.e., achievement Level II or higher) in reading achievement at the end of three years of intervention compared to schools not experiencing the intervention?
- Do elementary and middle schools that participate in Reenergize have higher proportions of students meeting Texas state standards (i.e., achievement Level II or higher) in math achievement at the end of three years of intervention compared to schools not experiencing the intervention?
- Do elementary and middle schools that participate in Reenergize have higher proportions of students meeting Texas state standards (i.e., achievement Level II or higher) in reading achievement at the end of two years of intervention compared to schools not experiencing the intervention?

- Do elementary and middle schools that participate in Reenergize have higher proportions of students meeting Texas state standards (i.e., achievement Level II or higher) in math achievement at the end of two years of intervention compared to schools not experiencing the intervention?

All analyses conducted around **postsecondary (college) enrollment** are **exploratory**, as this is a longer-term outcome where impacts may not be observed when the evaluation was conducted. Further, prior to study pre-registration in REES, the evaluation team anticipated that data would not be available for the final intervention year by the report publication deadline.

Exploratory Research Questions (Postsecondary Enrollment)

- Do high schools that participate in Reenergize have higher rates of postsecondary enrollment (i.e., improved college access and enrollment) at the end of three years of intervention compared to schools not experiencing the intervention? *
- Do high schools that participate in Reenergize have higher rates of postsecondary enrollment (i.e., improved college access and enrollment) at the end of two years of intervention compared to schools not experiencing the intervention?

3. Implementation Study

This section describes the Reenergize intervention as implemented for the duration of the evaluation period. The implementation study assesses whether the six key activities of the Reenergize intervention were implemented with fidelity and provides implementation context from conversations with implementing staff and participating school leadership teams.

3.1 Implementation Overview

Reenergize was implemented beginning in school year 2019–20 in two public school districts serving high-need students around San Antonio, Texas. Within these two districts, Reenergize was implemented in ten schools: four elementary schools, four middle schools, and two high schools. As part of this study, Reenergize was to be implemented for three school years: 2019–20, 2020–21, and 2021–22.[14,15] However, of the initial ten intervention schools, seven schools participated in any Reenergize activities during Year 2 (2020–21) and only three schools participated in any Reenergize activities during Year 3 (2021–22).

The Reenergize program implemented six key activities in the participating schools:

1. **Root cause analysis.** Together, IDRA and district and school leadership engaged in root cause analysis of problems that have presented themselves within each of the three key issue domains (student attendance, reading and writing achievement, and college and career readiness). The results of that issue analysis guided the content of the technical assistance provided.

2. **Professional development for principals and leadership teams.** IDRA provided joint training to principals and their leadership teams; bespoke coaching for each leadership team; experiences in the design, implementation, and evaluation of schoolwide projects to all participants; and myriad opportunities to build trust as a group and understand their role as a leadership team. School leadership teams may include, but not limited to, principals, vice-principals, instructional coaches, and master teachers. Over the three years of implementation, more than 60 principals and leadership team members received professional development services for a total of more than 250 hours of professional development over 13 different modules. The modules covered topics such as IDRA's Quality Schools Action Framework™, asset-based practices, teacher diversity, family leadership, bullying and harassment, early childhood development, and bilingual education in addition to topics around the three key district issues (student attendance, reading and writing achievement, and college and career readiness).

[14] There was a two-year pilot period for Reenergize in school years 2017–18 and 2018–19. IDRA used the pilot period to design, test, and refine the Reenergize program. Only one intervention school district participated in the pilot period. Two of the intervention schools participated in both years of the pilot period, and two additional intervention schools participated in the second pilot year. For this report, Year 1 of the Reenergize intervention is the first year of full implementation (2019–20).

[15] The Reenergize EIR proposal specified a two-year implementation period, but IDRA received permission from the Department of Education to extend the implementation period by an additional year due to disruptions to instruction caused by the COVID-19 pandemic.

3. **Education for master teachers to become principals.** IDRA prepared and certified new principals from an existing pool of master teachers. New principal candidates were recruited from participating program schools; enrolled at University of Texas at San Antonio's (UTSA's) principal preparation program; participated in program-specific on-the-job professional development and group coaching; and completed clinical internship requirements. Master teacher training was provided by UTSA as part of its principal preparation program that takes place in the fall, spring, and summer semesters and can lead to a master's or doctoral degree in education. Over the three years of implementation, thirteen teachers from the two districts participated in the UTSA principal preparation program as well as professional development and group coaching led by IDRA.[16]

4. **Campus climate alignment efforts**. IDRA coached each school's leadership team through the process of building a school culture aligned to district-level priorities. Through professional development, targeted coaching, and mentoring, participants develop a passion for advocacy and equity that is expected to influence all school activity and result in schoolwide improvements.

5. **Community partnerships.** IDRA implemented its Education CAFE framework to emphasize the importance of empowering families through a community-based organization to improve the educational process in the three key district issues through leadership, capacity-building, and accountability to support student achievement and close existing gaps.

6. **Professional learning communities (PLCs).** IDRA assisted school leadership teams to establish and maintain PLCs to facilitate peer-to-peer learning around the three key district issues.

All IDRA professional development efforts are rooted in asset-based practices that assume all students, regardless of race/ethnicity and socio-economic background, can succeed and be prepared for college.

Two IDRA consultants (one man and one woman) who are former local teachers with more than 20 years of experience in reading and writing instruction provided the program-specific master teacher professional development and one-on-one coaching. The professional development and coaching activities included an equity audit, co-teaching, modeling, lesson planning, and observations. The leadership team and master teachers received professional coaching services both individually and as a group to support the implementation of the school's action plan.

IDRA staff, who are predominately Hispanic/Latino, led or provided the other program activities (root cause analysis, campus alignment efforts, professional development for leadership teams, and establishing community partnerships and professional learning communities).

Note that due to COVID-19, some activities that were intended to be offered in person were provided via video conferencing or Google Classrooms.

[16] Over the five-year EIR grant, IDRA supported 25 Master Teachers through the UTSA principal preparation program across the two school districts, including eight Master Teachers in the pilot period (2017–19) and five teachers in comparison schools in Year 3 (2021–22).

Intervention School Characteristics

Students served by the Reenergize schools are primarily Hispanic/Latino and economically disadvantaged (receive free or reduced-price lunch). Most participating schools also have a higher proportion of students with disabilities and English Language Learners compared with the national average–14.1 percent and 10.1 percent, respectively (National Center for Education Statistics, 2019, Tables 204.20 and 204.70). The majority of Reenergize schools have lower student achievement scores than the state average of 79 out of 100, and were identified for either comprehensive or targeted support by the Texas Education Agency's Accountability Ratings in 2019. Exhibit 3 describes the characteristics of the intervention schools.

Exhibit 3. Characteristics of Intervention Schools

School	Grades Served	Number of Students	Hispanic/ Latino	Economically Disadvantaged	Students with Disabilities	English Language Learners	Student Achievement Scaled Score [a]
School District 1							
School A	6-8	735	98%	94%	29%	12%	60
School B	PK-5	486	93%	96%	24%	8%	58
School C	9-12	1,141	98%	92%	10%	11%	74
School D	PK-5	455	96%	95%	19%	11%	54
School E	6-8	621	96%	95%	24%	12%	59
School District 2							
School F	PK-5	608	90%	89%	23%	16%	76
School G	6-8	870	93%	90%	15%	12%	73
School H	6-8	904	85%	73%	8%	13%	75
School I	PK-5	628	89%	78%	22%	12%	75
School J	9-12	2,016	89%	75%	6%	13%	84

Source: 2018–19 Texas Academic Performance Reports (TAPR) and 2019 Accountability Ratings

[a] The Student Achievement Scaled Score (out of 100) is a school accountability rating defined by the Texas Education Agency. The Student Achievement Scaled Score evaluates district and school performance based on student achievement in three areas: performance on STAAR assessments; College, Career, and Military Readiness (CCMR) indicators; and graduation rates. For more information on this score, see https://tea.texas.gov/sites/default/files/Adopted%202019%20Accountability%20Manual_final.pdf.

3.2 Implementation Measurement

Abt worked closely with IDRA to develop fidelity measures and specified thresholds for adequate implementation of the Reenergize intervention (see Appendix A). IDRA collected the implementation data for school years 2019–20, 2020–21, and 2021–22, corresponding to Years 1-3 of the study, and provided them to Abt at the end of the intervention period. Implementation data comes from multiple sources, including trainer logs, coach logs, and grant performance reports. To document implementation of the program and inform IDRA's future efforts, Abt also conducted a focus group with school leadership and master teachers in May 2021 and a group interview with two of the IDRA coaches in February 2023.

The IDRA Reenergize logic model (Exhibit 2) includes six key activities (see Section 3.1) essential to IDRA's professional development efforts, which intend to improve student achievement by selecting, training, coaching, and empowering principals and school leadership teams to be effective advocates, collaborators, and risk-taking innovators.

For each key activity, IDRA and Abt created indicators to reflect adequate implementation of the Reenergize program model. Specifically, Abt measured implementation fidelity by rolling up indicator scores at the unit level (e.g., school, master teacher) to the sample-level and comparing them to the threshold for adequate implementation (see Appendix A). Meeting overall program fidelity expectations would mean that each component (and indicator, where applicable) meets the established fidelity thresholds. The fidelity matrix required a high degree of fidelity of implementation across most schools (86 percent or more of schools receive the highest fidelity threshold) for the program element to be rated as implemented with fidelity.

Abt measured program fidelity for the full sample of intervention schools and master teachers across all three implementation years, regardless of the number of schools implementing Reenergize activities. Appendix B includes a supplemental analysis of fidelity limited to those schools that implement Reenergize in each year.

3.3 Implementation Fidelity Findings

Overall program fidelity for the full sample of ten intervention schools was not met for the three years of the Reenergize intervention implementation, as the intervention schools did not meet the fidelity thresholds for all key activities each year.

- In Year 1, Reenergize met the fidelity threshold for one key component, professional learning communities, but did not meet the fidelity threshold for the other five activities. Nonetheless, more than half of the schools implemented these activities with fidelity.
- In Year 2, Reenergize did not meet the fidelity thresholds for any of the key activities. However, at least half of the schools implemented root cause analysis, campus climate alignment efforts, and professional learning communities with fidelity.
- In Year 3, fewer than half of the schools implemented the key activities with fidelity.

Separately by year, Exhibit 4 specifies whether the fidelity threshold was met for each key activity used in the overall program fidelity measure. (For additional details on implementation fidelity in each year, see Appendix B.) Exhibit 4 also reports the number of schools implementing the intervention each year—ten in Year 1, seven in Year 2, and three in Year 3.

Exhibit 4. Overall Program Fidelity Results

Key Activity	Fidelity Threshold Met?		
	Year 1 2019–20	Year 2 2020–21	Year 3 2021–22
Root Cause Analysis	No	No	No
Professional Development for Principals and Leadership Teams	No	No	No
Education for Master Teachers to Become Principals	No	No	No
Campus Climate Alignment Efforts	No	No	No
Community Partnerships	No	No	No
Professional Learning Communities	Yes	No	No
# Schools Implementing Reenergize [a]	10	7	3

[a] Overall program fidelity was measured for the full sample of intervention schools (N=10) across all three implementation years, regardless of the number of schools implementing Reenergize activities.

Relatively few intervention schools experienced the full Reenergize program as designed. The impact study includes all intervention schools regardless of level of implementation, so its results represent the effect of a partial implementation of Reenergize. It is worth noting that the low fidelity scores for each year could be partially attributed to various factors, such as changes in school administration and school restructuring, which led to many schools not participating in the intervention at all.

To explore this further, Abt conducted a supplemental analysis of program fidelity using the subset of intervention schools that implemented Reenergize activities each year. This subset of schools, Reenergize met the fidelity thresholds for the training for master teachers component for all implementing years and the establishing a professional learning community component for Years 1 and 2. (For additional details on the supplementary fidelity analysis, see Appendix Table B.2).

The fidelity findings might not present a complete picture of Reenergize implementation, particularly given the challenges posed by the COVID-19 pandemic. The next section provides additional details on the impact of COVID-19 and other challenges on the implementation of the Reenergize intervention.

3.4 Implementation Context

The **COVID-19 pandemic**, which began at the end of Year 1 (2019–20), had a significant impact on schools and communities, which provides important context for the Reenergize intervention. In Texas, the closure of schools from March 2020 until the end of the 2019–20 school year resulted in limited access to meals, social services, and other critical support services for low-income communities, families, and students who depended on them, which had a profound impact on their well-being. An Urban Education Institute at UTSA (2020) survey found that food insecurity was a significant issue for many students and families in San Antonio during the COVID-19 pandemic school closures, including in the districts in this study. The survey also found that food-insecure high school students were less motivated during distance learning. Specifically, 25 percent of high school students who said they were never engaged by classroom lessons also experienced food insecurity.

The **shift to virtual learning** also affected professional development for teachers and school leaders. Observations, feedback, coaching, mentoring, and community building are all essential elements for their growth. Interviews with Reenergize coaches revealed that adapting these practices to a virtual setting was difficult:

> *I think the fact that we were doing it virtually really limited what we could hear and sometimes see what teachers were doing…. It just seemed like less helpful from my estimation to do a debrief with the teacher online, because although I could see the teacher or hear the teacher, I don't know that I made any kind of an impact. But versus being there in person and that personal face-to-face, person-to-person rather than online, that I thought was more helpful for teachers.*

Texas schools were also experiencing **high rates of teacher turnover**. Texas teacher surveys found that 77 percent of respondents were considering leaving the profession in 2022, a 9 percent increase from 2021 and a 19 percent increase from 2020 (Charles Butt Foundation, 2022). In the focus group conducted with

the Reenergize coaches, teacher turnover was identified as a significant challenge to the implementation of the intervention activities:

> *Particularly in the middle school, there was a problem with staffing with teachers. There was turnover with teachers, the special ed teacher became a seventh grade teacher, one of the teachers resigned, and there was a series of substitute teachers…. It was a challenge. [A] couple of times in the classroom I was in, the sub was not terribly qualified, and would let me do the teaching. But then oftentimes she'd leave the room, so it wasn't a situation where I was actually modeling for her. I did the teaching so she could run out and do something else.*

The turnover of teachers meant disrupting instruction for many vulnerable students and additional responsibilities for the remaining teachers while adapting to Reenergize's new teaching and learning modalities. It also meant that sustaining the intervention within a school over the long term was more challenging when the coaches had to work with new groups of teachers every school year.

Furthermore, as noted in Implementation Fidelity Findings (Section 3.3 and Appendix B), many schools also saw **changes in principals and school administration**; one of the intervention schools even had a principal assigned to lead two schools during the second year of the intervention. Previous research has shown positive and significant associations between principal influence and elements of organizational climate such as collegial leadership, professional teacher behavior, push to maintain high academic standards, and institutional vulnerability (Smith et al., 2020). These changes in school administration could have affected the success of the Reenergize intervention and the ability of IDRA to implement the program with fidelity.

Despite facing significant challenges with the implementation of Reenergize, **IDRA persevered in its efforts to reach out and engage with districts and schools**. It continued to offer coaching support to schools and adapted its services to meet the varying needs of different schools. For example, one school district requested an adaptation to the Reenergize model to focus more specifically on assistant principals and instructional coaches after a change in superintendent. IDRA worked collaboratively with the district to respond to requests and kept the district involved through extensive outreach and engagement. Additionally, IDRA shifted its services to a virtual format and expanded its reach to non-intervention schools based on the schools' specific needs and capabilities. It is crucial to consider the challenges and IDRA's response to these challenges when assessing the overall Reenergize implementation and fidelity.

4. Impact Study

The Reenergize impact study is designed to assess the program's effect on five long-term student outcomes: school attendance, reading/writing and math achievement, high school graduation rates, college and career readiness, and postsecondary enrollment. This section describes the impact study's design and methods, including the outcome measures, intervention and comparison conditions, study sample, and analytic approach. We also establish baseline equivalence between our intervention and comparison groups. Finally, we describe the results of each confirmatory and exploratory impact analysis.

4.1 Outcome Measures

The impact study examines the Reenergize intervention's effects on the average school-grade–level reading/writing and math performance on the State of Texas Assessments of Academic Readiness (STAAR) standardized tests for students in elementary and middle schools (grades 3-8); on the average school-level high school graduation rates, college and career readiness,[17] and postsecondary enrollment measures for students in high schools (grades 9-12); and on the average school-level attendance rates for students in all schools (grades 1-12).[18]

All confirmatory outcomes (attendance, reading/writing and math achievement, graduation rates, college and career readiness) are measured at the end of the full three years of implementation (2021–22). Additionally, the evaluation measures these same outcomes after two years of implementation (2020–21) as exploratory outcomes. Additional exploratory outcomes, including proportions of students meeting Texas state standards and postsecondary enrollment, were measured after two and three years of implementation.

At the time this report was prepared, school attendance and high school graduation data for school year 2021–22 and postsecondary enrollment data for Fall 2022 were not available. Thus, impact estimates for these contrasts have been omitted from this report.[19]

Data for each outcome was obtained from publicly available school-level data from the Texas Education Agency (TEA) and spans the school years of 2014–15 through 2021–22. Baseline data was downloaded by Spring 2021, when the final baseline year data (2018–19) became publicly available. Outcome data was

[17] College and career readiness is measured using the Texas Success Initiative (TSI) criteria for Reading Language Arts and Mathematics. The TSI benchmark rate is the percentage of annual graduates who meet predefined criteria for a Texas state high school exit-level test, the SAT, or the ACT or by successfully completing and earning credit for a college prep course. The TSI benchmark rate calculation and criteria changed over time but remained consistent across schools within a given year. To account for this, we standardize the TSI benchmark rate within each year and analyze the Reenergize program's impact in terms of the changes in standard deviations to a school's TSI benchmark rate from the previous year. See Appendix Table C.1 for more details.

[18] Due to the COVID-19 pandemic, many schools transitioned to remote learning during the 2019–20 school year. In response, TEA temporarily altered the definition for attendance rate from "the percentage of days students were present over the entire school year" to "the percentage of days students were present through the fourth six weeks" (i.e., roughly February 2020, or two-thirds of the school year). The TEA returned to the same calculation for the 2020–21 school year, according to IDRA, as attendance continued to be a concern into Fall 2020.

[19] These outcomes are not expected to be available until Summer 2023 and might be included in an update to the final evaluation report.

downloaded from the TEA website at the end of each school year as it became available. Data was not analyzed until Spring 2023, however, after the comparison matching school procedure and baseline equivalence was established and the study pre-registration in REES was finalized. Refer to Appendix C.1 for more information about each outcome measure.

4.2 Intervention and Comparison Conditions

The Reenergize program model includes six key activities implemented over three school years for participating schools: (1) root cause analysis, (2) professional development for principals and leadership teams, (3) education for master teachers to become principals, (4) campus climate alignment efforts, (5) establishing community partnerships to empower families, and (6) establishing professional learning communities. Relatively few intervention schools fully implemented the Reenergize program, so the contrast between the intervention and comparison conditions does not reflect the full impact of the Reenergize program. See Section 3 for more information on how these activities were implemented during the study period, as well as important implementation context.

To evaluate the impact of the Reenergize program, intervention schools were compared with a matched set of comparison schools. For details on the comparison school selection procedure, see Appendix C.2. Comparison schools participated in business-as-usual instruction, systems, professional development, and leadership. Given that the evaluation team selected the comparison schools from publicly available data, we did not have the ability to track non-Reenergize services received by the comparison schools.

Three comparison schools in one of the intervention districts ended up receiving the Reenergize program in the last two years of implementation. One school participated in the model with full fidelity during school year 2021–22, one school participated in the model with some fidelity, and in the third school, one master teacher participated in the education for master teachers. These three schools remain in our comparison sample because our analysis follows an intent-to-treat approach; thus, our intervention group is limited to schools that agreed to implement Reenergize at the beginning of the study.[20] All other comparison schools never had the opportunity to participate in Reenergize during the evaluation period.

4.3 Study Sample

The impact study included all ten schools (four elementary schools, four middle schools, and two high schools) implementing Reenergize at the start of the 2019–20 school year, serving a total of 8,464 students across two school districts.

As described in Appendix C.2, each of the ten Reenergize schools is matched to a group of 15 comparison schools of the same school type (elementary, middle or high). There are 160 unique schools in the impact study (10 intervention, 150 comparison). For each school, our analysis includes seven school years of data: pre-intervention data from school years 2014–15 through 2018–19, and post-intervention data from school

[20] The intent-to-treat approach measures the impact of introducing the Reenergize program on schools, rather than just looking at its direct effect on student outcomes. This approach helps to address concerns about schools in the comparison groups potentially starting to use the program and affecting the results.

years 2020–21 and 2021–22.[21] Where outcome data is missing, we used case-deletion to delete the missing observation from the analysis. Therefore, the analysis samples are defined as cases with non-missing pre-intervention outcome data.

Not all outcomes of interest are measured across all grade levels. To account for this, we define three analytic samples based on the grades a school serves and the corresponding outcomes of interest measured:

- **Elementary and Middle Schools (Grades 3-8)**: STAAR Reading Language Arts and Mathematics
- **High Schools (Grades 9-12)**: graduation rate, college and career readiness, and postsecondary enrollment
- **All Schools (Grades 1-12)**: school attendance[22]

Exhibit 5 describes the analytic sample size for each outcome measure. The unit of analysis for the attendance, graduation rate, college and career readiness, and postsecondary enrollment outcomes is the school. The unit of analysis for the STAAR Reading Language Arts (RLA) and Mathematics outcomes is the school-grade (i.e., grade within a school).[23]

Exhibit 5. Analytic Sample Size for Each Outcome

Sample	Reenergize Intervention Group			Comparison Group		
	Baseline (2018–19)	Year 2 (2020–21)	Year 3 (2021–22)	Baseline (2018–19)	Year 2 (2020–21)	Year 3 (2021–22)
Elementary and Middle Schools						
RLA (z-score)	24	24	22	345	342	330
Mathematics (z-score)	22	22	20	317	314	302
Proficient RLA (rate)	24	24	22	345	342	330
Proficient Mathematics (rate)	22	22	20	317	314	302
High Schools						
Graduation Rate	2	2	-	30	30	-
College and Career Readiness	2	2	2	30	30	30
Postsecondary Enrollment	2	2	-	30	30	-
All Schools						
School Attendance	10	10	-	150	145	-

RLA=Reading Language Arts. RLA and Mathematics test scores are standardized (z-score) within each grade (for grades 3-8) and school year. The numbers in this table represent the total number of schools within each condition, except for the elementary and middle school STAAR outcomes, which represent the number of grades within schools. Outcome data not yet available at the time of this report are represented by a dash. We do not include data from Year 1 (2019–20) because outcome data collection and reporting for this school year was disrupted by the COVID-19 pandemic.

[21] We do not include data from Year 1 (2019–20) in any analytic models. Outcome data collection and reporting for this school year was disrupted by the COVID-19 pandemic.

[22] While the intervention elementary schools serve Pre-K and Kindergarten students, TEA's attendance rate formula only includes students in grades 1 through 12. For more information, refer to https://rptsvr1.tea.texas.gov/perfreport/tapr/2020/glossary.pdf#page=8.

[23] Since the schools in our study sample do not include a consistent grade range, the number of school-grades is not a multiple of the number of elementary and middle schools in the study (8 intervention schools and 120 comparison schools).

4.4 Analytic Approach

We use a comparative short interrupted time-series (C-SITS) model with baseline mean projection to estimate impacts of the intervention on school attendance, reading/writing and math achievement, high school graduation rates, college and career readiness, and postsecondary enrollment.[24] For the STAAR RLA and Mathematics outcomes measured at the grade level, we pool the baseline and outcome data across grades to calculate a single, combined impact estimate for grades 3-8. All of the outcome data is from a single state, which simplifies the process of combining across grades. All other outcomes are measured at the school level.

The impact model includes fixed effects for matching blocks which indicate each intervention school and its corresponding matched comparison schools. All models also include random effects for schools (and grades, when data is reported at the grade level), the baseline measure of the outcome, and baseline school characteristics. See Appendix C.3 for the full impact model specification. Note that because the confirmatory outcomes are in different WWC outcome domains, the study team did not adjust any confirmatory analyses for multiple comparisons.[25] No adjustments were made to the exploratory analyses because the WWC does not require multiple comparisons adjustments for exploratory analyses.

The goal of these analyses is to estimate the impacts of Reenergize on intervention schools. The analyses use a matched comparison group to quasi-experimentally test whether intervention school outcomes differ from what would be expected had they not participated in the Reenergize program. If differences in outcomes are observed between the intervention and matched comparison groups, there is evidence that those differences in outcomes can be attributed to the effects of participating in Reenergize. However, because the evaluation uses a quasi-experimental design, we cannot completely eliminate other potential explanations for any observed differences in outcomes between the intervention and comparison groups.

Additionally, we conduct a supplemental analysis to evaluate the impact of Reenergize on the three schools that implemented the program for the intended three years (three-year exposure). We limit the analytical sample to include only these schools and their respective comparison group schools. As these three schools are either elementary or middle schools, we only analyze reading/writing and math achievement and attendance outcomes. Similar to the approach taken for the confirmatory and exploratory analyses, no adjustments were made for multiple comparisons.

4.5 Baseline Equivalence

For the quasi-experimental impact study to estimate an unbiased impact of Reenergize, post-intervention outcome measures for comparison schools should be reasonable counterfactuals for the intervention

[24] A C-SITS model with baseline mean projection estimates how much more (or less) outcomes changed in the intervention group than the comparison group between the years before the study and the years of the study. It is equivalent to a difference-in-differences model.

[25] See Appendix C.1 for the WWC outcome domain for each outcome measure.

schools (i.e., they represent what would have happened in the intervention schools if they had not been selected to participate in the Reenergize program).[26]

Assessing the similarity of the intervention and comparison schools on pre-intervention measures, such as student demographics, TEA school accountability ratings, and outcomes of interest is one way to assess whether outcomes for comparison schools are a reasonable counterfactual (see Appendix C.4 for baseline equivalence model specifications). According to guidelines outlined by the U.S. Department of Education's What Works Clearinghouse Version 4.1 and 5.0 Standards (What Works Clearinghouse, 2020; What Works Clearinghouse, 2022), baseline differences larger than 0.25 standard deviations in quasi-experimental design (QED) studies suggest that the comparison schools are different enough from the intervention schools that any observed outcome differences cannot be confidently attributed to the intervention.

Appendix D.1 compares the pre-intervention (2018–19) school-level demographic characteristics, Texas Education Agency's school accountability ratings, and outcomes of interest between Reenergize and comparison schools by analytic sample and outcome. The baseline differences in pre-intervention outcomes between intervention and comparison schools are either below or within the adjustment range according to What Works Clearinghouse Standards (Versions 4.1 and 5.0) for all outcomes. As such, analyses proceeded with the full analytic sample for each contrast.

This cluster-level assignment study uses school- and grade-level administrative data as outcomes without any systematic exclusions and therefore meets WWC cluster representativeness standards without the need to present the full population of individuals in clusters at follow-up.

4.6 Student Outcome Findings

In this section, we describe the impact of the Reenergize intervention on the key long-term student outcomes: reading/writing and math achievement, high school graduation rates, college and career readiness, postsecondary enrollment, and school attendance.

Outcomes for Elementary and Middle Schools (Grades 3-8): Reading Language Arts (RLA) and Mathematics Test Scores and Proficiency

As seen in Exhibits 6 through 9, the Reenergize program did not have a significant impact on any of the confirmatory or exploratory STAAR outcomes. Focusing on confirmatory outcomes (school year 2021–22) for each subject, our findings indicate that Reenergize schools, on average, scored lower than comparison schools by 0.04 standard deviations (SD) in reading/writing and 0.11 SD in math after adjusting for baseline differences. However, we are unable to statistically distinguish these findings from no impact of the program.

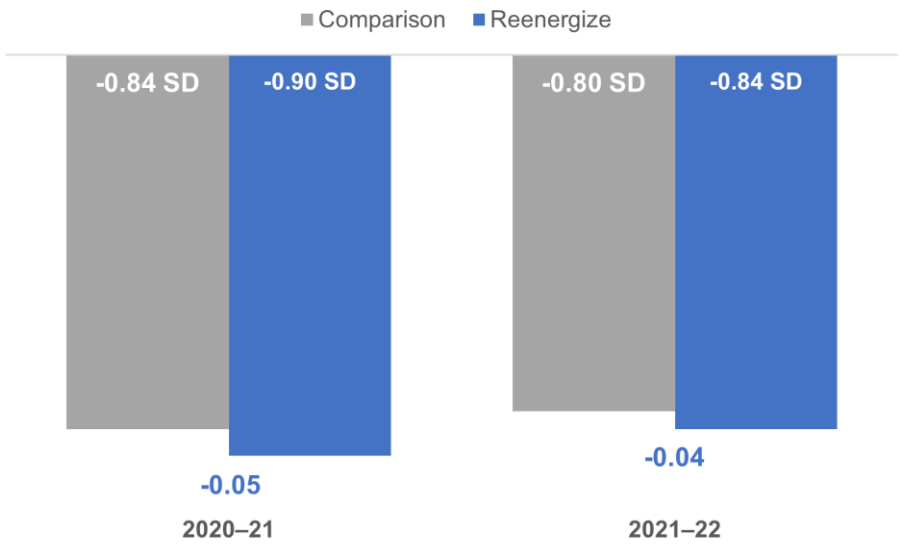
With the exception of standardized math test scores for the intervention schools, comparing the standardized test score means within Exhibits 6 and 7 to those found in the baseline equivalence table for

[26] Establishing baseline equivalence mitigates the concern for selection bias distorting the estimated impact of the program on student outcomes. Selection bias is the risk that intervention schools are not similar to comparison schools at the beginning of the study, so the comparison schools cannot be reasonable counterfactuals. Additional assumptions required for the estimate to be unbiased include: conditional on the factors within our model, the outcome is independent of a school's intervention status, the instrumentation of each outcome remains constant over the study period, and no other policies were introduced during this time that target the outcomes of interest.

school year 2021–22 outcomes (Appendix Table D.1b) indicates that test scores decreased by about 0.04 SD across both the intervention and comparison groups between the 2018–19 school year and the 2020–21 and 2021–22 school years.[27] For example, the average standardized reading/writing test score for intervention schools in 2018–19 was -0.80 SD, which compares to -0.84 SD for the 2021-22 school year. The impact that the COVID-19 pandemic had on schooling across the state of Texas could possibly explain this phenomenon.

For the three elementary and middle schools that participated in the intervention for the intended three years (three-years exposure sample), Reenergize did not significantly impact RLA or math test scores (Appendix Table D.2d).

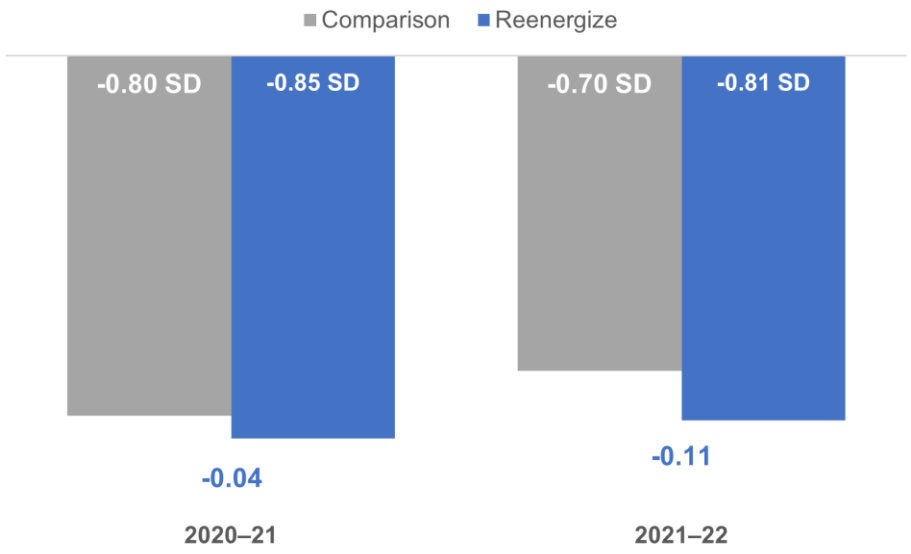
Exhibit 6. Impact of Reenergize on Standardized RLA Test Scores



Comparison bar indicates the unadjusted mean for 342 and 330 comparison school-grades during the 2020–21 and 2021–22 school years, respectively. Reenergize bar indicates the model-adjusted mean for 24 and 22 intervention school-grades during the 2020–21 and 2021–22 school years, respectively. The intervention-comparison difference may not be the difference between the displayed intervention and comparison means due to rounding. *Outcome:* State of Texas Assessments of Academic Readiness (STAAR) RLA for grades 3-8 in standard deviation units (z-scores) in a school year. *Data source:* Texas Education Agency's STAAR Aggregate Data (Campus-Level).

[27] The Reenergize means have been adjusted for differences in student demographic characteristics, aggregated to the school level, between the intervention and comparison groups for each outcome measure using data from school years 2014–15 through the baseline year (2018–19). For more details on the model, see Appendix C.3.

Exhibit 7. Impact of Reenergize on Standardized Mathematics Test Scores

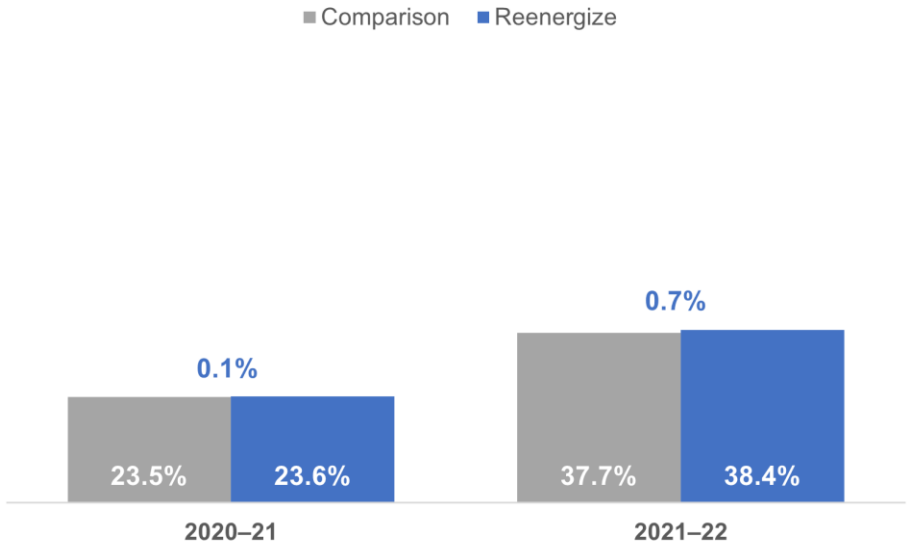


Comparison bar indicates the unadjusted mean for 314 and 302 comparison school-grades during the 2020–21 and 2021–22 school years, respectively. Reenergize bar indicates the model-adjusted mean for 22 and 20 intervention school-grades during the 2020–21 and 2021–22 school years, respectively. The intervention-comparison difference may not be the difference between the displayed intervention and comparison means due to rounding. *Outcome:* State of Texas Assessments of Academic Readiness (STAAR) Mathematics for grades 3-8 in standard deviation units (z-scores) in a school year. Data source: Texas Education Agency's STAAR Aggregate Data (Campus-Level).

In addition to test scores, we also explore the program's impact on STAAR RLA and Mathematics proficiency rates (Exhibits 8 and 9) and do not find any statistically significant differences between the comparison and intervention schools. Appendix Table D.2a provides the number of clusters and the number of students across clusters, as well as the statistical details for the estimated impact of Reenergize on all elementary and middle school outcomes.

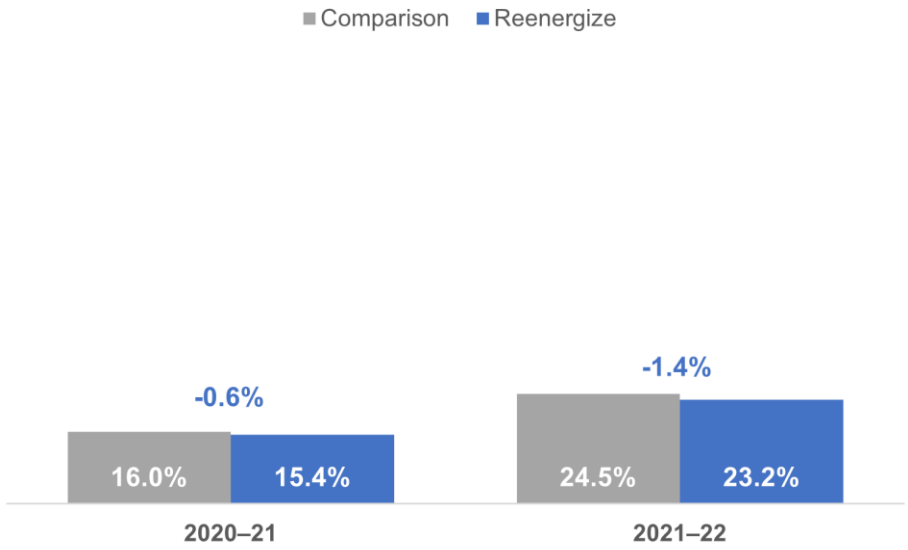
For the three elementary and middle schools that participated in Reenergize for the intended three years (three-years exposure sample), we did not find any statistically significantly impact on RLA or math proficiency (Appendix Table D.2d.).

Exhibit 8. Impact of Reenergize on RLA Proficiency



Comparison bar indicates the unadjusted mean for 342 and 330 comparison school-grades during the 2020–21 and 2021–22 school years, respectively. Reenergize bar indicates the model-adjusted mean for 24 and 22 intervention school-grades during the 2020–21 and 2021–22 school years, respectively. The intervention-comparison difference may not be the difference between the displayed intervention and comparison means due to rounding. *Outcome:* Percentage of students meeting Level II: Satisfactory Academic Performance or above in State of Texas Assessments of Academic Readiness (STAAR) RLA for grades 3-8 in a school year. *Data source:* Texas Education Agency's STAAR Aggregate Data (Campus-Level).

Exhibit 9. Impact of Reenergize on Mathematics Proficiency

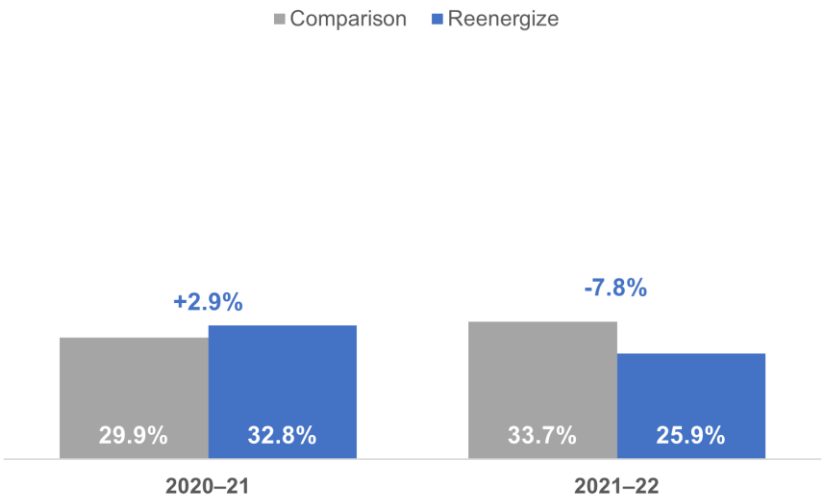


Comparison bar indicates the unadjusted mean for 314 and 302 comparison school-grades during the 2020–21 and 2021–22 school years, respectively. Reenergize bar indicates the model-adjusted mean for 22 and 20 intervention school-grades during the 2020–21 and 2021–22 school years, respectively. The intervention-comparison difference may not be the difference between the displayed intervention and comparison means due to rounding. *Outcome:* Percentage of students meeting Level II: Satisfactory Academic Performance or above in State of Texas Assessments of Academic Readiness (STAAR) Mathematics for grades 3-8 in a school year. *Data source:* Texas Education Agency's STAAR Aggregate Data (Campus-Level).

Outcomes for High Schools (Grades 9-12): High School Graduation, College and Career Readiness, and Postsecondary Enrollment

The Reenergize program did not have a statistically significant impact on the confirmatory (2021–22 school year) or exploratory (2020–21 school year) college and career readiness outcome (Exhibit 10).[28] Though the difference in the 2021–22 Texas Success Initiative (TSI) rates between intervention and comparison schools might seem substantial (-7.80 percentage points), the small sample of intervention schools for these outcomes (*n*=2) likely contributed to the imprecision of the impact estimate. This imprecision could also explain the change in sign in the impact finding from the 2020–21 school year.

Exhibit 10. Impact of Reenergize on College and Career Readiness



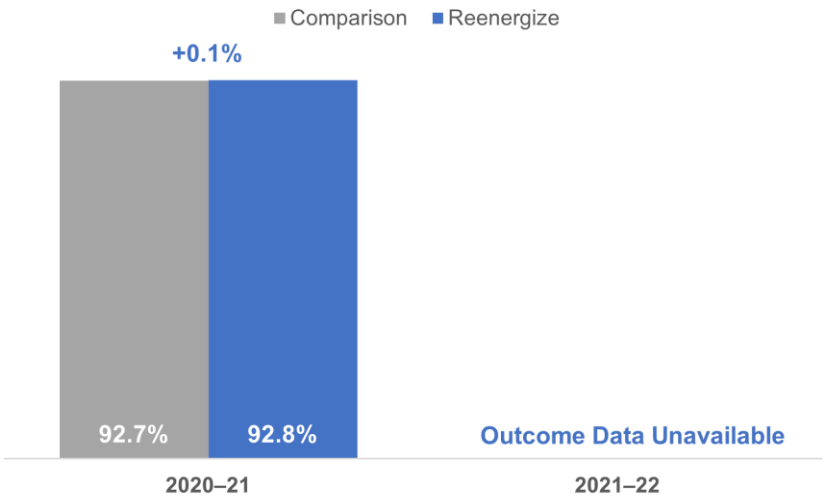
Comparison bar indicates the unadjusted mean for 30 comparison schools during the indicated year. Reenergize bar indicates the model-adjusted mean for two intervention schools during the indicated year. The intervention-comparison difference may not be the difference between the displayed intervention and comparison means due to rounding. *Outcome:* Percentage of graduates who met the Texas Success Initiative (TSI) criteria in both RLA and Mathematics in a school year. *Data source:* Texas Education Agency's Texas Accountability Rating System: Accountability Ratings.

Due to lack of data availability, we were unable to analyze the 2021–22 school year outcomes for high school graduation and postsecondary enrollment rate (confirmatory and exploratory research questions, respectively). However, as shown in Exhibits 11 and 12, the Reenergize program did not have a significant impact on graduation or postsecondary enrollment rates for the 2020–21 school year. Similar to the trend for STAAR test scores, these averages decreased for both groups from the baseline year, though less so for graduation rates. However, the average postsecondary enrollment rate decreased by about 7 to 8 percentage points between 2018–19 and 2020–21.

[28] The analysis uses within-year standardized TSI rates to account for changes in how TSI was calculated during the 2016–17 school year. For simplicity, the estimated impact of Reenergize has been converted from the number of standard deviations from the mean TSI rate to the corresponding TSI rate for the given outcome year. For example, the average number of standard deviations from the TSI rate mean for the comparison schools in the 2020–21 school year was -0.41 SD (see Appendix Table D.2b), which corresponds to a TSI benchmark rate of 29.9% for that school year.

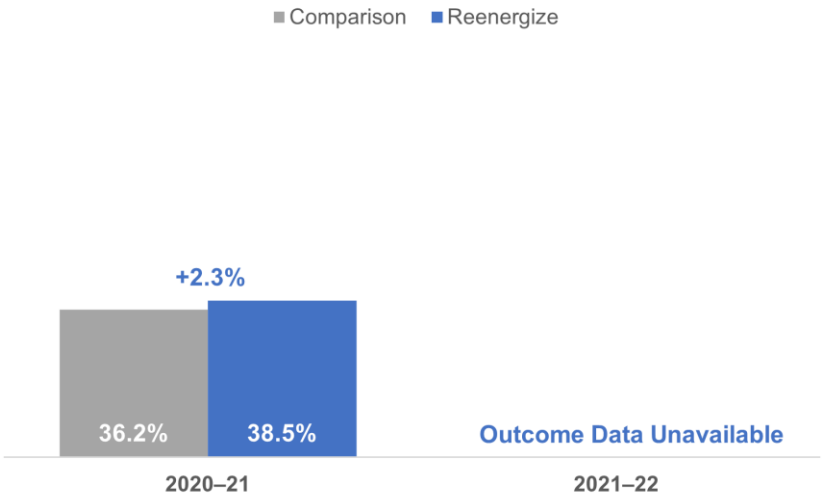
Appendix Table D.2b contains details on sample sizes, impact means, and standardized differences for each high school impact estimate.

Exhibit 11. Impact of Reenergize on Graduation Rate



Comparison bar indicates the unadjusted mean for 30 comparison schools during the indicated year. Reenergize bar indicates the model-adjusted mean for two intervention schools during the indicated year. The intervention-comparison difference may not be the difference between the displayed intervention and comparison means due to rounding. School year 20211–22 graduation data was unavailable. *Outcome*: Percentage of 12th grade students graduating (4-Year Longitudinal Rate) in a school year. *Data source:* Texas Education Agency's Completion, Graduation, and Dropout Data.

Exhibit 12. Impact of Reenergize on Postsecondary Enrollment Rate



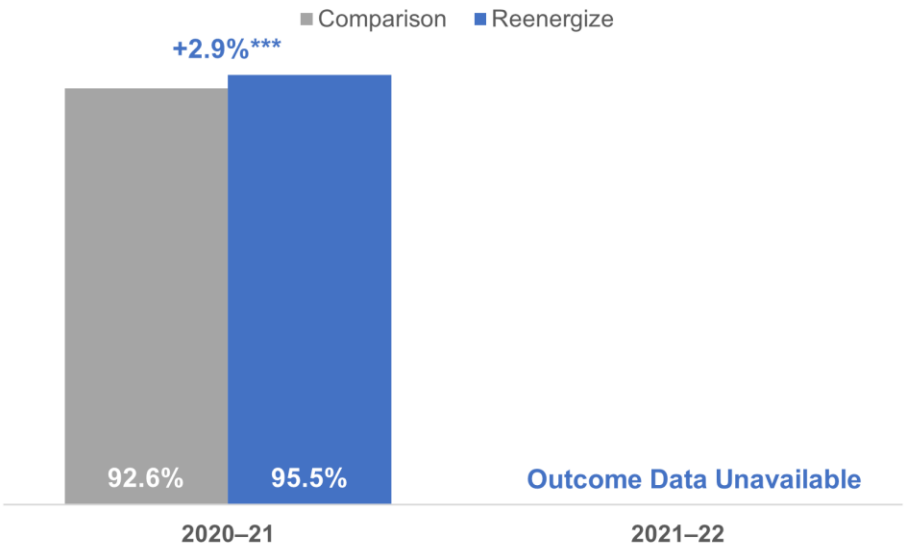
Comparison bar indicates the unadjusted mean for 30 comparison schools during the indicated year. Reenergize bar indicates the model-adjusted mean for two intervention schools during the indicated year. The intervention-comparison difference may not be the difference between the displayed intervention and comparison means due to rounding. School year 2021–22 postsecondary enrollment data was unavailable. *Outcome*: Percentage of students enrolling in a postsecondary institution in Texas in a school year. *Data source*: Texas Higher Education Coordinating Board: High School Graduates Enrolled in Higher Education the Following Fall by High School County, School District, High School.

Outcomes for All Schools (Grades 1-12): School Attendance

Due to lack of data availability (see Section 3.2), we were unable to analyze our confirmatory attendance rate research question after the full three years of implementation (after school year 2021–22). However, Reenergize increased student attendance rates in Year 2. Exhibit 13 shows a statistically significant impact of Reenergize on student attendance rates in Year 2 (school year 2020–21; p <. 001). Attendance rates in Reenergize schools were 2.9 percentage points higher than attendance rates in comparison schools after controlling for matching group and baseline measures (Exhibit 13). See Appendix Table D.2c for more details on the attendance rates impact estimates.

The Year 2 attendance rate finding is robust to restricting the sample to the three schools that participated in the Reenergize intervention across Year 1, Year 2, and Year 3. As shown in Appendix Table D.2d., for the three-years exposure sample, the mean attendance rate for the comparison group is 93 percent, and the model adjusted mean attendance rate for the intervention schools is 3 percentage points higher (*p<0.01*) at 96 percent.

Exhibit 13. Impact of Reenergize on Attendance Rate



Comparison bar indicates the unadjusted mean of for 145 comparison schools during the indicated year. Reenergize bar indicates the model-adjusted mean for 10 intervention schools during the indicated year. The intervention-comparison difference may not be the difference between the displayed intervention and comparison means due to rounding. School year 2021–22 attendance data was unavailable. *Outcome:* Percentage of days that students were present in a school year. *Data source:* Texas Education Agency's Texas Academic Performance Reports (TAPR). *** Difference is statistically significant at the .001 level (p < .001).

5. Conclusion

IDRA implemented Reenergize, a culturally proficient, specific transformation intervention for administrators, aspiring administrators, and teachers, in two public school districts serving high-need students around San Antonio, Texas between the 2019–20 and 2021–22 school years. IDRA aimed to enhance instructional leadership and management skills of school personnel, which they hypothesized would lead to increased student achievement, increased college and career readiness, and improved rates of attendance, high school graduation, and postsecondary enrollment.

The COVID-19 pandemic, combined with high rates of teacher and administration turnover, severely attenuated the implementation of the Reenergize program. Only seven of the initial ten intervention schools participated in any Reenergize activities during Year 2 (2020–21), and only three of the initial ten intervention schools participated in any Reenergize activities during Year 3 (2021–22). Consequently, the Reenergize program did not meet any of the program-level thresholds for adequate fidelity of implementation.

With the exception of attendance in Year 2, there were no statistically significant differences in educational outcomes based on a school's participation in the Reenergize program. The relatively low level of implementation fidelity and limited participation of many intervention schools in Reenergize suggests that the intervention contrast was relatively small; these factors attenuated the potential impact of the intervention. In addition, given the small sample size of assigned units (schools), the evaluation's ability to detect small statistically significant effects is limited. This is particularly true for the high school and postsecondary outcomes, as only two high schools participated in the Reenergize intervention. The standard errors of the impact estimates are large relative to the scales of the measures, precluding detection of effects of small to moderate magnitude.

Despite the challenges faced during the implementation of Reenergize, IDRA learned valuable lessons that can inform future efforts. Abt conducted focus groups with teachers and school leaders who participated in Reenergize as well as IDRA coaches and leaders to collect insights into factors crucial to the engagement and participation of Reenergize stakeholders.

One of the key lessons learned was the importance of assessing stakeholder capacity and aligning the Reenergize program to fit within this capacity. Aligning Reenergize with existing initiatives, resources and constraints of the participating schools and districts could lessen the stakeholders' responsibilities and increase their likelihood of participating in the program across multiple school years. Moreover, IDRA recognized the need for district and school champions who can advocate for the Reenergize program and support its implementation. These champions help increase buy-in and participation from other stakeholders in their district or school.

Another important lesson learned was the need to flexibly adapt to shifting stakeholder needs. Adaptivity and flexibility was especially important during the COVID-19 pandemic. As stakeholder needs changed rapidly, IDRA was able to pivot and offer alternative services to support Reenergize schools and districts. Furthermore, leveraging existing relationships allowed IDRA to develop and maintain partnerships with Reenergize districts and schools, and effective planning and collaboration from school administration allowed the Reenergize coaches to integrate into schools smoothly. As IDRA continues to support schools and districts around San Antonio, the findings from the Reenergize study and the lessons learned can inform future implementation and evaluation efforts.

Appendix

Appendix A: Fidelity of Implementation Matrix

This appendix includes information on the implementation fidelity of key activities including fidelity indicators and scoring. The implementation fidelity matrix appears in Appendix Table A.1 below.

Implementation fidelity was measured by rolling up indicator scores at the unit level (e.g., school, master teacher) to the sample-level and comparing them to the threshold for adequate implementation. The thresholds are defined for each key activity as follows:

1. ***Root Cause Analysis*** was measured through one indicator, collaboration among leadership, defined as the leadership team meeting at least twice in the first semester to conduct root cause analysis. At the unit level, schools received a score of 0 if the leadership team did not meet, a score of 1 if the team met once, and a score of 2 if the team met twice for root cause analysis. The sample-level threshold for adequate implementation was achieved if 86 percent or more of schools received a score of 2, indicating that the majority of schools worked frequently on root cause analysis.

2. ***Professional Development for Principals and Leadership Teams*** was measured through two indicators: 1) Training for the leadership team and 2) Coaching for the leadership team. Training for the leadership team was defined as receiving 75 hours of training each school year. Schools received a score of 0 if the leadership team received less than 40 hours of training, a score of 1 if the team received 40 to 74 hours of training, and a score of 2 if the team received 75 or more hours of training. Coaching for the leadership team was defined as the leadership team receiving a minimum of two hours of coaching a month. Schools received a score of 0 if the leadership team received less than 8 hours of individual coaching, a score of 1 if the team received 8 to 15 hours of individual coaching, and a score of 2 if the team received 16 or more hours of individual coaching. The sample-level threshold for adequate implementation was achieved if 86 percent or more of schools received a score of 4 across the two indicators.

3. ***Education for Master Teachers to become Principals*** was measured through four indicators: 1) Training for master teachers, 2) Semester hours of university credit for master teachers, 3) Individual coaching for master teachers, and 4) Clinical training for master teachers. Semester hours of university credit for master teachers was defined as each master teacher working towards completing 4.5 semester hours of university credit. A master teacher received a score of 0 if the teacher worked towards completing less than 2 semester hours of university credit, a score of 1 if the teacher worked towards completing 2 to 3.25 semester hours of university credit, and a score of 2 if the teacher worked towards completing 3.5 or more semester hours of university credit.[29] Clinical training for master teachers was defined as each master teacher participating in one clinical training in a low-performing school. A master teacher received a score of 0 if the teacher did not participate in the clinical training and a score of 1 if the teacher participated in the clinical training. The school-level threshold for adequate implementation was achieved if 86 percent or more of master teachers in a school received a score of 7 across the four indicators. The sample-level threshold for adequate implementation was achieved if 86 percent or more of schools met the school-level threshold for adequate implementation.

[29] Semester hours are measured in 0.25-hour increments.

4. ***Campus climate alignment efforts*** was measured through collaboration among the leadership team and master teachers. This indicator was defined as principals, leadership team, and master teachers, i.e., school stakeholders, meeting at least twice a semester to collaborate on ways to develop a school culture around social justice and equity. Schools received a score of 0 if the school stakeholders did not meet, a score of 1 if the school stakeholders met once per semester, and a score of 2 if the school stakeholders met at least twice in one semester and one or two times in the second semester to develop a school culture around social justice and equity. The sample-level threshold for adequate implementation was achieved if 86 percent or more of schools received a score of 2, indicating that the majority of schools had frequent collaborations on campus climate alignment efforts.

5. ***Community partnerships*** was measured through two indicators: 1) Schools use the IDRA Education CAFE Framework to build partnerships and 2) Collaborative team meets to discuss ways to involve and serve families in schools. The first indicator was defined as parents, school leaders, and other stakeholders building relationships between school, community, and families using the IDRA Education CAFE Framework. Schools received a score of 0 if meetings on partnership formation were not held at all, a score of 1 if at least one meeting on partnership formation was held but no partnerships were formed, and a score of 2 if after partnership formation, at least two community partnership meetings were held. The second indicator was defined as collaborative meetings occurring at least 2 times a year. Schools received a score of 0 if the collaborative team did not meet at all, a score of 1 if the collaborative team met once or twice, and a score of 2 if the collaborative team met three or more times. The sample-level threshold for adequate implementation was achieved if 86 percent or more of schools received a score of 4 across the two indicators.

6. ***Professional Learning Communities*** was measured through collaboration among teachers, defined as teachers in a school meeting at least twice a semester to collaborate on improving school culture to promote achievement and expectations for success. Schools received a score of 0 if teachers did not meet in professional learning community (PLC) during the year, a score of 1 if teachers met in PLC once per semester, and a score of 2 if teachers met in PLC at least 2 times in one semester and one or two times in the second semester. The sample-level threshold for adequate implementation was achieved if 86 percent or more schools received a score of 2, indicating that 86 percent or more of schools had frequent collaborations in PLCs.

Appendix Table A.1. Implementation Fidelity Matrix

Indicator	Unit of measurement	Indicator Scoring at Unit Level	Indicator Scoring at School Level	Indicator Scoring at Sample Level
Key Activity 1: Root Cause Analysis				
Collaboration among leadership team	School	0 = team did not meet to conduct root cause analysis at all during first semester 1 = team met to conduct root cause analysis once during the first semester 2 = team met to conduct root cause analysis at least twice during the first semester		
Key Activity 1: Root Cause Analysis		Sum of school-level indicator scores (range=0-2) **Adequate school score = 2**	N/A (School is unit for this Key Component)	0 = Less than half of schools receive score of 2 1 = 50-85% of schools receive score of 2 2 = 86%–100% of schools receive score of 2 **Adequate Sample score = 2 (86%–100% of schools work frequently on root cause analysis).**
Key Activity 2: Professional Development for Principals and Leadership Teams				
Training for leadership team	School	0 = team received less than 40 hours of training 1 = team received 40-74 hours of training 2 = team received 75 or more hours of training		

Indicator	Unit of measurement	Indicator Scoring at Unit Level	Indicator Scoring at School Level	Indicator Scoring at Sample Level
Coaching for leadership team	School	0 = team received less than 8 hours of individual coaching 1 = team received 8-15 hours of individual coaching 2 = team received 16 or more hours of individual coaching		
Key Activity 2: Professional Development for Principals and Leadership Teams		Sum of school-level indicator scores (range=0-4) **Adequate school score = 4**	N/A (School is unit for this Key Component)	0 = Less than half of schools receive score of 4 1 = 50-85% of schools receive score of 4 2 = 86%–100% of schools receive score of 4 **Adequate sample score = 2 (86%–100% of school leadership teams receive full training and coaching).**
Key Activity 3: Education for Master Teachers to become Principals				
Training for master teachers	Individual Master Teacher	0 = individual received less than 40 hours of training 1 = individual received 40-74 hours of training 2 = individual received 75 or more hours of training		
Semester Hours of University Credit for master teachers	Individual Master Teacher	0 = individual worked towards less than 2 semester hours of university credit 1 = individual worked towards 2-3.25 semester hours of university credit 2 = individual worked towards 3.5 or more semester hours of university credit		
Coaching for master teachers--Individual	Individual Master Teacher	0 = individual received less than 8 hours of coaching 1 = individual received 8-15 hours of coaching 2 = individual received 16 or more hours of coaching		

Indicator	Unit of measurement	Indicator Scoring at Unit Level	Indicator Scoring at School Level	Indicator Scoring at Sample Level
Clinical Training for master teachers	Individual Master Teacher	0 = individual did not participate in clinical training 1 = individual participated in clinical training		
Key Activity 3: Education for Master Teachers to become Principals		Sum of teacher-level indicator scores (range=0-7) **Adequate teacher score = 7**	0 = LT half of individuals in a school receive score of 7 1 = 50-85% of individuals in a school receive score of 7 2 = 86%–100% of individuals in a school receive score of 7 **Adequate school score = 2**	0 = Less than half of schools receive score of 2 1 = 50-85% of schools receive score of 2 2 = 86%–100% of schools receive score of 2 **Adequate sample score = 2 (in 86%–100% of schools, 86%–100% of master teachers receive full training, university credit, coaching, and clinical training).**
Key Activity 4: Campus climate alignment efforts				
Collaboration among leadership team and master teachers	School	0 = school stakeholders did not meet during year to develop a school culture around social justice and equity 1 = school stakeholders met once per semester to develop a school culture around social justice and equity 2 = school stakeholders met at least 2 times in one semester and one or two times in 2nd semester to develop a school culture around social justice and equity		

Indicator	Unit of measurement	Indicator Scoring at Unit Level	Indicator Scoring at School Level	Indicator Scoring at Sample Level
Key Activity 4: Campus climate alignment efforts		Sum of school-level indicator scores (range=0-2) **Adequate school score = 2**	N/A (School is unit for this Key Component)	0 = Less than half of schools receive score of 2 1 = 50-85% of schools receive score of 2 2 = 86%–100% of schools receive score of 2 **Adequate sample score = 2 (in 86%–100% of schools there is frequent collaboration).**
Key Activity 5: Community partnerships				
Schools use IDRA Education Café Framework to build partnerships	School	0 = no collaborative meetings on partnership formation 1 = at least one meeting but no partnerships formed 2=at least 2 community partners meetings after partnership formation		
Collaborative team meets to discuss ways to involve and serve families in schools	School	0 = no collaborative meetings held 1 = 1-2 collaborative meetings held 2= 3 or more collaborative meetings held		
Key Activity 5: Community partnerships		Sum of school-level indicator scores (range=0-4) **Adequate school score = 4**	N/A (School is unit for this Key Component)	0 = Less than half of schools receive score of 3 or higher 1 = 50-85% of schools receive score of 3 or higher 2 = 86%–100% of schools receive score of 3 or higher **Adequate sample score = 2 (in 86%–100% of schools there is high partnership formation and frequent collaborative meetings).**
Key Activity : 6: Professional Learning Communities (PLCs)				

Indicator	Unit of measurement	Indicator Scoring at Unit Level	Indicator Scoring at School Level	Indicator Scoring at Sample Level
Collaboration among teachers	School	0 = teachers did not meet in PLC during year 1 = teachers met in PLC once per semester 2 = teachers met in PLC at least 2 times in one semester and one or two times in 2nd semester		
Key Activity 6: Professional Learning Communities (PLCs)		Sum of school-level indicator scores (range=0-2) **Adequate school score = 2**	N/A (School is unit for this Key Component)	0 = Less than half of schools receive score of 2 1 = 50-85% of schools receive score of 2 2 = 86%–100% of schools receive score of 2 **Adequate sample score = 2 (86%–100% of schools have frequent collaborations in PLCs).**

Appendix B. Fidelity of Implementation Findings

Abt measured program fidelity for the full sample of intervention schools and master teachers across all three implementation years, regardless of the number of schools implementing Reenergize activities. In Year 1 (2019–20), ten intervention schools participated in the Reenergize program and were to implement the six key activities, including root cause analysis, professional development for principals, campus climate alignment efforts, and professional learning communities. Five master teachers across four intervention schools participated in education for master teachers in the first year. In Year 2 (2020–21) and Year 3 (2021–22), seven and three schools respectively participated in any Reenergize activities. Four master teachers across four intervention schools and three master teachers across three intervention schools participated in education for master teachers in Year 2 and Year 3, respectively.

Appendix Table B.1 presents information on each indicator used in the program fidelity measure, specifying whether program fidelity was met for each year of the intervention, as well as the observed fidelity score for each key component, and the sample size for implementation and fidelity measurement (the number of units in which the component was implemented, and the number of units in which fidelity was measured, respectively.) Appendix Table B.2 presents the same information as table B.1 for the subset of schools implementing any component of the Reenergize intervention in the specified school year.

Year 1 (2019–20)

In Year 1 of the intervention, Reenergize only met fidelity thresholds for professional learning communities component with 90% of schools implementing the component with fidelity. Reenergize received a sample-level fidelity score of 2 for this component, meeting the threshold for adequate implementation.

In addition, 80% of schools implemented root cause analysis and campus climate alignment efforts with fidelity, 70% implemented professional development for principals and leadership teams with fidelity, and 60% implemented community partnerships with fidelity. Reenergize received a sample-level fidelity score of 1 for these activities and failed to meet the threshold for adequate implementation.

Reenergize did not meet the threshold for adequate implementation for the education for master teachers to become principals component with a sample-level fidelity score of 0. Although the teacher-level adequate score of 7 and the school-level adequate score of 2 were met, the sample-level threshold for adequate implementation was not met because only 40% of schools implemented the component with fidelity.

Although Year 1 of intervention had the highest level of school participation, as reflected by the sample level fidelity scores, changes in school administration led one school to not participate in majority of the intervention activities. With a small sample size of 10 schools, the absence of even one school increased the chances of the sample not meeting the fidelity thresholds for the different activities.

Year 2 (2020–21)

In Year 2 of the intervention, the first school year beginning remotely under COVID-19 school closures, Reenergize did not meet fidelity thresholds for any of the key activities.

Specifically, 60% of schools implemented root cause analysis and professional learning communities with fidelity and 50% implemented campus climate alignment efforts with fidelity. Reenergize received a sample-level fidelity score of 1 for these activities and failed to meet the threshold for adequate implementation. Reenergize received a sample-level fidelity score of 0 for professional development for principals and leadership teams, community partnerships, and education for master teachers to become principals, and failed to meet the threshold for adequate implementation. However, all four master teachers met the teacher-level adequate score of 7, and all four schools met the school-level adequate score of 2. Nevertheless, Reenergize failed to meet the sample-level threshold for adequate implementation for education for master teachers because only 40% of schools implemented it with fidelity.

The decrease in the number of schools implementing key activities with fidelity can be attributed to three of the intervention schools not participating in the intervention due to changes in principals and school restructuring.

Year 3 (2021–22)

In Year 3 of the intervention, Reenergize did not meet fidelity thresholds for any of the key activities.

Reenergize received a sample-level fidelity score of 0 for root cause analysis, professional development for principals and leadership teams, campus climate alignment efforts, community partnerships, and professional learning communities, indicating that less than 50% of schools implemented the activities with fidelity. Only 20% of schools implemented root cause analysis and campus climate alignment efforts with fidelity. Similarly, only 10% of schools implemented professional development for principals and leadership teams, community partnerships, and professional learning communities with fidelity.

In Year 3, the participation of schools in the intervention decreased significantly, with only three of the original intervention schools continuing to participate. Three schools continued their non-participation from Year 2, and four additional schools did not participate in Year 3.

Appendix Table B.1. Implementation Fidelity Findings

Key Activities, Number of Indicators, Units, and Threshold				Year 1 (2019–20 School Year)			Year 2 (2020–21 School Year)			Year 3 (2021–22 School Year)		
Key Activity	Total # of Measurable Indicators	Unit of Implementation	Sample-Level Threshold	Number of Units in which Activity was Implemented [a]	Number of Units in which Fidelity of Activity was Measured [b]	Achieved Fidelity Score and whether Program met Sample-Level Threshold	Number of Units in which Activity was Implemented [a]	Number of Units in which Fidelity of Activity was Measured [b]	Achieved Fidelity Score and Whether Program Met Sample-Level Threshold	Number of Units in which Activity was Implemented [a]	Number of Units in which Fidelity of Activity was Measured [b]	Achieved Fidelity Score and whether Program Met Sample-Level Threshold
Root Cause Analysis	1	School	**Sample score of 2**	10 schools	10 schools	Score is 1 ***Program fidelity = No***	7 schools	10 schools	Score is 1 ***Program fidelity = No***	3 schools	10 schools	Score is 0 ***Program fidelity = No***
Professional Development for Principals and Leadership Teams	2	School	**Sample score of 2**	10 schools	10 schools	Score is 1 ***Program fidelity = No***	7 schools	10 schools	Score is 0 ***Program fidelity = No***	3 schools	10 schools	Score is 0 ***Program fidelity = No***
Education for Master Teachers to become Principals	4	Master Teacher	**Sample score of 2**	5 master teachers across 4 schools	5 master teachers across 10 schools	Score is 0 ***Program fidelity = No***	4 master teachers across 4 schools	4 master teachers across 10 schools	Score is 0 ***Program fidelity = No***	4 master teachers across 4 schools	4 master teachers across 10 schools	Score is 0 ***Program fidelity = No***

Key Activities, Number of Indicators, Units, and Threshold				Year 1 (2019–20 School Year)			Year 2 (2020–21 School Year)			Year 3 (2021–22 School Year)		
Key Activity	Total # of Measurable Indicators	Unit of Implementation	Sample-Level Threshold	Number of Units in which Activity was Implemented [a]	Number of Units in which Fidelity of Activity was Measured [b]	Achieved Fidelity Score and whether Program met Sample-Level Threshold	Number of Units in which Activity was Implemented [a]	Number of Units in which Fidelity of Activity was Measured [b]	Achieved Fidelity Score and Whether Program Met Sample-Level Threshold	Number of Units in which Activity was Implemented [a]	Number of Units in which Fidelity of Activity was Measured [b]	Achieved Fidelity Score and whether Program Met Sample-Level Threshold
Campus Climate Alignment Efforts	1	School	**Sample score of 2**	10 schools	10 schools	Score is 1 ***Program fidelity = No***	7 schools	10 schools	Score is 1 ***Program fidelity = No***	3 schools	10 schools	Score is 0 ***Program fidelity = No***
Community Partnerships	2	School	**Sample score of 2**	10 schools	10 schools	Score is 1 ***Program fidelity = No***	7 schools	10 schools	Score is 0 ***Program fidelity = No***	3 schools	10 schools	Score is 0 ***Program fidelity = No***
Professional Learning Communities	1	School	**Sample score of 2**	10 schools	10 schools	Score is 2 ***Program fidelity = Yes***	7 schools	10 schools	Score is 1 ***Program fidelity = No***	3 schools	10 schools	Score is 0 ***Program fidelity = No***

[a] Defined as the number of schools that implemented at least one Reenergize activity in the school year.

[b] Fidelity was measured across all intervention schools regardless of the number of schools implementing Reenergize activities in a given school year.

Appendix Table B.2. Supplemental Implementation Fidelity Findings for Implementing Schools

Key Activities, Number of Indicators, Units, and Threshold				Year 1 (2019-20 School Year)			Year 2 (2020-21 School Year)			Year 3 (2021-22 School Year)		
Key Activity	Total # of Measurable Indicators	Unit of Implementation	Sample-Level Threshold	Number of Units in which Activity was Implemented [a]	Number of Units in which Fidelity of Activity was Measured [b]	Achieved Fidelity Score and whether Program Met Sample-Level Threshold	Number of Units in which Activity was Implemented	Number of Units in which Fidelity of Activity was Measured	Achieved Fidelity Score and whether Program Met Sample-Level Threshold	Number of Units in which Activity was Implemented	Number of Units in which Fidelity of Activity was Measured	Achieved Fidelity Score and whether Program Met Sample-Level Threshold
Root Cause Analysis	1	School	**Sample score of 2**	10 schools	10 schools	Score is 1 ***Program fidelity = No***	7 schools	7 schools	Score is 1 ***Program fidelity = No***	3 schools	3 schools	Score is 1 ***Program fidelity = No***
Professional Development for Principals and Leadership Teams	2	School	**Sample score of 2**	10 schools	10 schools	Score is 1 ***Program fidelity = No***	7 schools	7 schools	Score is 0 ***Program fidelity = No***	3 schools	3 schools	Score is 0 ***Program fidelity = No***
Education for Master Teachers to Become Principals [c]	4	Master Teacher	**Sample score of 2**	5 master teachers across 4 schools	5 master teachers across 4 schools	Score is 2 ***Program fidelity = Yes***	4 master teachers across 4 schools	4 master teachers across 4 schools	Score is 2 ***Program fidelity = Yes***	4 master teachers across 4 schools	4 master teachers across 4 schools	Score is 2 ***Program fidelity = Yes***

Key Activities, Number of Indicators, Units, and Threshold				Year 1 (2019-20 School Year)			Year 2 (2020-21 School Year)			Year 3 (2021-22 School Year)		
Key Activity	Total # of Measurable Indicators	Unit of Implementation	Sample-Level Threshold	Number of Units in which Activity was Implemented [a]	Number of Units in which Fidelity of Activity was Measured [b]	Achieved Fidelity Score and whether Program Met Sample-Level Threshold	Number of Units in which Activity was Implemented	Number of Units in which Fidelity of Activity was Measured	Achieved Fidelity Score and whether Program Met Sample-Level Threshold	Number of Units in which Activity was Implemented	Number of Units in which Fidelity of Activity was Measured	Achieved Fidelity Score and whether Program Met Sample-Level Threshold
Campus Climate Alignment Efforts	1	School	**Sample score of 2**	10 schools	10 schools	Score is 1 ***Program fidelity = No***	7 schools	7 schools	Score is 1 ***Program fidelity = No***	3 schools	3 schools	Score is 1 ***Program fidelity = No***
Community Partnerships	2	School	**Sample score of 2**	10 schools	10 schools	Score is 1 ***Program fidelity = No***	7 schools	7 schools	Score is 0 ***Program fidelity = No***	3 school	3 school	Score is 0 ***Program fidelity = No***
Professional Learning Communities	1	School	**Sample score of 2**	10 schools	10 schools	Score is 2 ***Program fidelity = Yes***	7 schools	7 schools	Score is 2 ***Program fidelity = Yes***	3 schools	3 schools	Score is 0 ***Program fidelity = No***

[a] Defined as the number of schools that implemented at least one Reenergize component in the school year.

[b] Fidelity was measured only across the schools that implemented at least one Reenergize component in the given school year.

[c] For Education for master teacher component, Abt measured fidelity only across 4 schools that implemented the component.

Appendix C: Impact Study Design Supplemental

This appendix includes additional information on the outcome measures used in the impact analysis, the comparison school selection procedure, and the impact study model specifications.

Appendix C.1 Outcome Measures

Appendix Table C.1 provides information on the outcome measures used in the impact study.

Appendix Table C.1. Outcome Measures

Domain[a]	Outcome	Outcome Construction	Reliability/ Validity	Baseline Measure(s)	Data Source
School Attendance	Attendance	Average percentage of days that students were present in a school year [b]	State Standard educational measure	*Same as outcome:* Average percentage of days that students were present in a school year	Texas Education Agency's Texas Academic Performance Reports (TAPR)
High School Completion	Graduation	Percentage of 12th grade students graduating (4-Year Longitudinal Rate) [c]	State Standard educational measure	*Same as outcome:* Percentage of 12th grade students graduating (4-Year Longitudinal Rate)	Texas Education Agency's Completion, Graduation, and Dropout Data
College Readiness	College and Career Readiness: TSI Rate	Percentage of graduates who met TSI criteria in both ELA and math [d]	Standardized test	*Same as outcome:* Percentage of graduates who met TSI criteria in both ELA and math standardized at year-grade level [a]	Texas Education Agency's Texas Accountability Rating System: Accountability Ratings
College Enrollment	Postsecondary Enrollment	Share of students enrolling in a postsecondary institution in Texas	State Standard educational measure	*Same as outcome:* Share of students enrolling in a postsecondary institution in Texas	Texas Higher Education Coordinating Board: High School Graduates Enrolled in Higher Education the Following Fall by High School County, School District, High School
Proficiency in the English Language	STAAR RLA	STAAR RLA for grades 3-8 in standard deviation units (z-scores) [e]	Standardized test	*Same as outcome:* STAAR RLA for grades 3-8 in standard deviation units (z-scores)	Texas Education Agency's State of Texas Assessments of Academic Readiness (STAAR) Aggregate Data (Campus-Level)
Mathematics Achievement	STAAR Math	STAAR Math for grades 3-8 in standard deviation units (z-scores) [e]	Standardized test	*Same as outcome:* STAAR Math for grades 3-8 in standard deviation units (z-scores)	Texas Education Agency's State of Texas Assessments of Academic Readiness (STAAR) Aggregate Data (Campus-Level)

Proficiency in the English Language	STAAR RLA	Percentage of students meeting Level II: Satisfactory Academic Performance or above in STAAR RLA for grades 3-8	Standardized test	*Same as outcome:* Percentage of students meeting Level II: Satisfactory Academic Performance or above in STAAR RLA for grades 3-8	Texas Education Agency's STAAR Technical Digest Appendices: STAAR Score Distributions and Statistics by Content Area and Grade
Mathematics Achievement	STAAR Math	Percentage of students meeting Level II: Satisfactory Academic Performance or above in STAAR Math for grades 3-8	Standardized test	*Same as outcome:* Percentage of students meeting Level II: Satisfactory Academic Performance or above in STAAR Mathematics for grades 3-8	Texas Education Agency's STAAR Technical Digest Appendices: STAAR Score Distributions and Statistics by Content Area and Grade

ELA=English Language Arts. RLA = Reading Language Arts.

[a] Outcome domain definitions are aligned with the WWC Study Review Protocol (Version 5.0).

[b] While the intervention elementary schools in our sample serve Pre-K and Kindergarten students, TEA's attendance rate formula only includes students in grades 1 through 12. For more information, see https://rptsvr1.tea.texas.gov/perfreport/tapr/2020/glossary.pdf#page=8. Note that due to the COVID-19 pandemic, many schools transitioned to remote learning during the 2019–20 school year. In response, TEA temporarily altered the definition for attendance rate from "the percentage of days students were present over the entire school year" to "the percentage of days students were present through the fourth six weeks" (i.e., roughly February 2020 or two-thirds of the school year). For more information, see https://rptsvr1.tea.texas.gov/perfreport/tapr/2021/glossary.pdf. While the TEA returned to the same calculation for the 2020–21 school year, according to IDRA, attendance continued to be a concern into the Fall of 2020.

[c] The four-year longitudinal graduation rate is the percentage of students who graduate with a high school diploma within four years of beginning ninth grade. This does not include those who leave high school upon receipt of a Texas Certificate of High School Equivalency (TxCHSE). For more information, see https://tea.texas.gov/reports-and-data/school-performance/accountability-research/completion-graduation-and-dropout/four-year-graduation-and-dropout-data-class-of-2021

[d] Students must meet at least one of the Texas Success Initiative (TSI) criteria for reading/ELA and Math to enroll at a Texas public institution of higher education, specifically, high school graduates who met the college-ready criteria in both reading/ELA and math on the Texas Assessment of Knowledge and Skills (TAKS) exit-level test, the SAT, or the ACT, or by successfully completing and earning credit for a college prep course. Students must exceed a benchmark score on the standardized TSI Assessment (TSIA) to enroll in credit-bearing coursework. Students can be exempted from the TSIA if they exceed a minimum score on the ACT, SAT, or English III/Algebra II STAAR end-of-course exam. For more information, see Chapter 4 in https://tea.texas.gov/sites/default/files/Adopted%202019%20Accountability%20Manual_final.pdf . Note that the TSI criteria changed over time but remained consistent across schools within a given year. To account for this, we standardize the TSI rate within each year and analyze the Reenergize program's impact in terms of the changes in standard deviations to a school's TSI rate from the previous year.

[e] The STAAR (State of Texas Assessments of Academic Readiness) data are comprised of school-level mean scores for each grade level. To combine data across grades, we converted the school-by-grade scores into standardized z-scores using the student-level statewide score means and standard deviations for the corresponding grade and year. Following What Works Clearinghouse convention, we use the statewide student-level standard deviation (1.0, since the outcomes are z-scored) to calculate the standardized difference rather than the pooled standard deviation from the study sample.

Appendix C.2 Comparison School Selection Procedure

For each of the 10 Reenergize intervention schools, a set of 15 comparison schools were selected from a pool of potential comparison schools. For a given Reenergize school, the pool of potential comparison schools includes all, traditional public or charter schools in Texas not receiving Reenergize services in Year 1 (2019–20) that share the *school type* (i.e., elementary, middle, or high) of the intervention school.[30] We calculated the linear distance between each intervention school and all potential comparison schools on a vector of student and school characteristics (described below). The linear distance formula is given by:

$$Distance_s^c = \sqrt{\sum_{i=1}^{n}(t_i - x_i^c)^2},$$

where *Distance* represents the linear distance in characteristics between the intervention school (*s*) and each potential comparison school (*c*), *i* indexes the school and aggregated student characteristics used within the distance formula, *t* represents the value of the indexed characteristic for the intervention school, and *x* represents the value of the indexed characteristic for potential comparison school *c*.

Not all school types are represented in each analytic sample. For example, the graduation rate outcome only applies to high schools, so elementary and middle schools are excluded from this analysis. Due to these differences, some of the student and school matching characteristics used to select comparison schools differ across school types. We list the matching characteristics used for each school type below:

- Matching characteristics for all school types (elementary, middle, and high schools): attendance rate (all students), percent Hispanic students, percent White students, percent students with a disability, Overall rating, School Progress rating, Academic Growth rating, Relative Performance rating, Closing the Gaps rating[31]
- Additional matching characteristics for high schools: graduation rate (all students), graduation rate (Hispanic students), TSI proficiency rate
- Additional matching characteristics for elementary and middle schools: percent of students who are English Language Learners, percent of students who are economically disadvantaged, Student Achievement rating

For each analytic sample, each potential comparison school is ranked relative to each intervention school by the calculated linear distance. The 15 best matched comparison schools (those with the smallest linear distance) are selected to form the initial comparison group for that intervention school.

To ensure each intervention school had 15 unique comparison schools, we iteratively replaced duplicate comparison schools in each comparison group based on the match rank. For example, if potential comparison school A has a match rank of 1 (i.e., the best match) for intervention school B and a match rank of 2 (i.e., the second-best match) for intervention school C, potential comparison school A is removed from intervention school C's comparison group and the next-closest potential comparison school match for

[30] In consultation with IDRA, we decided to exclude all schools in the Uvalde Consolidated Independent School District from the pool of potential comparison schools due to tragic events within the timeframe of this study.

[31] All ratings are based on TEA's accountability rating system. For detailed information on each rating, see https://tea.texas.gov/sites/default/files/Adopted%202019%20Accountability%20Manual_final.pdf.

intervention school C is substituted. If the match rank of a potential comparison school is tied between multiple intervention schools, we preserved the match with the smallest linear distance.

We verify that at least one comparison school from one of the two intervention districts is included in each analytic sample to ensure that our analysis assesses the impact of the intervention at the school level rather than at the district level.

Appendix C.3 Impact Model Specification

The impact model is a three-level model with repeated observations over years (level-1), nested within grades (level-2), and multiple grades, nested in schools (level-3):

$$Z_{iGj} = \beta_0 + \beta_1(T_{ij}) + \beta_2(TrtYr_{ij}) + \beta_3(T_{ij} * TrtYr_{ij}) + X_{iGj}\Gamma + \sum_{m=1}^{M-1} \beta_{(3+m)}(MatchingBlock_m) + \mu_j^{Schs} + r_{Gj}^{Grades} + \varepsilon_{iGj}^{Years}$$

The subscripts i, G, and j represent the ith time point for the Gth grade, in the jth school, respectively. The outcome variable (Z_{iGj}) represents the z-score (for STAAR RLA, STAAR Mathematics, and college and career readiness) or percentage (for all other outcomes). Each outcome is described in more detail in Appendix C.1.

The model also includes matching block fixed effects, a vector of baseline school characteristics for each ith year (X_{iGj}), and indicator variables for intervention school, intervention years, and a school-by-intervention year interaction term. In each analytic model, we control for five years of pre-intervention outcome data (school years 2014–15 to 2018–19) and the following pre-intervention school characteristics: percent Hispanic students, percent of students identified as English language learners, percent of students with disabilities, and percent of students identified as economically disadvantaged.

The coefficient of this model interaction (denoted as β_3) is the expected difference in differences in the academic outcome between the Reenergize group and the comparison group, and can be interpreted as the impact of the Reenergize program under the intent-to-treat framework.

The random intercepts for schools (denoted as μ_j^{Schs}), random intercepts for grades (denoted as r_{Gj}^{Grades}) and residual error term (denoted as $\varepsilon_{iGj}^{Years}$) are assumed to be normally distributed and independent from one another.

Z_{iGj}	the mean standardized test score (z-score) or percentage from the ith time point in the Gth grade in the jth school.
β_0	the intercept, which is the comparison school mean score in pre-intervention years for schools in the omitted matching block.
β_1	the average difference between intervention and comparison schools during pre-intervention years.
T_j	1 if school j is an intervention school,

	and 0 if comparison school.
β_2	the average difference between pre-intervention years and post- intervention years for comparison schools.
$TrtYr_{ij}$	1 if year is after 2019, And 0 if year is on or before 2019
β_3	The intervention effect. This is the difference-in-difference estimator. It is the difference between intervention and comparison schools in their average differences between pre-intervention and intervention years.
Γ	A vector of effects of the baseline school characteristics X_{iGj}: percent Hispanic students, percent of students eligible for English Language Learner status, percent of students with disabilities, and percent of students identified as economically disadvantaged for each *jth* year.
X_{iGj}	A vector of baseline averaged student characteristics from the ith time point in the jth school: percent Hispanic students, percent of students eligible for English Language Learner status, percent of students with disabilities, and percent of students identified as economically disadvantaged for each *jth* year.
$\beta_{(3+m)}$	is the *m*th coefficient for the *m*th matching block dummy variable.
$MatchingBlock_m$	An indicator variable that takes the value 1 if school was in the *m*th of *M* matching blocks, and 0 otherwise. (Note: if there are M matching blocks, there will be M-1 indicator variables for matching blocks).
μ_j^{Schs}	is the deviation of school j's intercept from the mean intercept, conditional on T_j, distributed with mean 0 and variance σ^2_{Schs}
r_{Gj}^{Grades}	is the deviation of grade G's intercept from the mean intercept, conditional on T_j, distributed with mean 0 and variance σ^2_{Grades}
$\varepsilon_{iGj}^{Years}$	the random error effect representing the difference between mean score at year i in grade G for school j and the predicted mean score for school j. These residual effects are assumed normally distributed with mean 0 and variance σ^2_{Years}, and are assumed to have 1st order autoregressive correlation. They are assumed independent from μ_j^{Schs} and r_{Gj}^{Grades}.

Appendix C.4 Baseline Equivalence Model Specification

We assessed baseline equivalence for school-level outcomes (attendance, graduation, college and career readiness, and postsecondary enrollment) using one-level models that include fixed effects for matching blocks (indicators for each intervention school and its corresponding matched comparison schools).

$$X_{j,2019} = \beta_0 + \beta_1(T_{2019}) + \sum_{m=1}^{M-1} \beta_{(1+m)}(MatchingBlock_m) + \varepsilon_{j,2019}$$

We assessed baseline equivalence for school-by-grade-level outcomes (STAAR RLA and Mathematics) using two-level models that include fixed effects for matching blocks and random effects for grades (level-1) nested within schools (level-2).

$$X_{Gj,2019} = \beta_0 + \beta_1(T_{2019}) + \sum_{m=1}^{M-1} \beta_{(1+m)}(MatchingBlock_m) + r_{Gj}^{Grades} + \varepsilon_{j,2019}$$

The subscripts G and j represent the Gth grade, in the jth school, respectively. The outcome variable ($X_{j,2019}$ or $X_{Gj,2019}$) is the pre-intervention (2018–19) school-level demographic characteristics, Texas Education Agency's school accountability ratings, or outcome of interest between Reenergize and comparison schools by analytic sample. The coefficient of the intervention indicator (denoted as β_1) is the expected difference in the pre-intervention measure between the Reenergize group and comparison group during the year prior to the intervention. See Section 4.6 and Appendix Table D.1 for baseline equivalence testing results.

Appendix D: Impact Study Findings Tables

This appendix includes the full tables for establishing baseline equivalence and the impact study findings.

Appendix D.1 Baseline Equivalence Tables

Appendix Tables D.1a through D1.d compare the pre-intervention (2018–19) school-level demographic characteristics, Texas Education Agency's school accountability ratings, and outcomes of interest (shown in *italics*) between Reenergize and comparison schools by analytic sample.

Across intervention status, students are primarily Hispanic/Latino, economically disadvantaged (receive free or reduced-price lunch), and are more likely to be classified as English Language Learners. Since the Reenergize program specifically targeted schools with lower academic success, it is unsurprising to find that the intervention and comparison schools exhibited, on average, lower rates in most outcome measures as compared to the state average for school year 2018–19.[32] The exceptions are graduation and attendance rates, which are over 93% for the intervention group, comparison group, and state averages. The baseline differences in pre-intervention outcomes between intervention and comparison schools are either below or within the adjustment range according to What Works Clearinghouse Standards (Version 4.1) for all outcomes.[33]

We exclude attendance, graduation, and postsecondary enrollment rates from Table D.1b due to the lack of outcome data for school year 2021–22. The baseline equivalence for TSI in school year 2021–22 is established in Table D.1d because there was no sample attrition in this outcome variable.

[32] The state average for each outcome within the 2018–19 school year are as follows: proficiency in ELA (48%), proficiency in math (52%), graduation (95%), TSI achievement (38%), college enrollment (51%), and attendance (95%). See https://rptsvr1.tea.texas.gov/perfreport/snapshot/2019/state.html for more information.

[33] We assessed baseline equivalence for school-level outcomes (attendance, graduation, TSI, and postsecondary enrollment) using one-level models that include fixed effects for matching blocks (indicators for each intervention school and its corresponding matched comparison schools). We assessed baseline equivalence for school-by-grade-level outcomes (STAAR ELA and Mathematics) using two-level models that include fixed effects for matching blocks and random effects for grades (level-1) nested within schools (level-2). See Appendix C.4 for the full models.

Appendix Table D.1a. Baseline Equivalence for SY2020–21 Elementary and Middle Schools Sample

	Reenergize			Comparison			Mean Difference	Pooled SD	Hedges' g	Cox's Index
	N	Mean	SD	N	Mean	SD				
Elementary and Middle Schools										
Mathematics										
Test Scores (Z-Score) [a]	22	-0.80	0.82	314	-0.67	0.73	-0.14	0.74	-0.19	-
Proficient Rate	22	0.29	0.15	314	0.33	0.15	-0.04	0.15	-0.27	-0.12
% Hispanic	22	0.92	0.04	314	0.92	0.06	0.00	0.05	0.03	0.01
Student Achievement	22	66.83	9.03	314	67.39	9.76	-0.56	9.71	-0.06	-
RLA										
Test Scores (Z-Score) [a]	24	-0.81	0.49	342	-0.75	0.67	-0.06	0.66	-0.09	-
Proficient Rate	24	0.30	0.10	342	0.31	0.12	-0.01	0.12	-0.06	-0.02
% Hispanic	24	0.92	0.04	342	0.92	0.06	0.00	0.06	0.07	0.03
Student Achievement	24	66.22	8.87	342	66.95	9.54	-0.73	9.50	-0.08	-

Outcome measures appear in italics. See Appendix C.1 for more information about the outcome measures. N is the total number of schools within each condition, except for the STAAR (State of Texas Assessments of Academic Readiness) exam outcomes, which are represented at the school-grade level. The standardized difference is reported as Hedge's g for continuous outcomes (the baseline intervention-comparison difference divided by the pooled standard deviation) and Cox's Index for dichotomous outcomes.[a] The STAAR data are comprised of school-level mean scores for each grade level. To combine data across grades, we converted the school-by-grade scores into standardized z-scores using the student-level statewide score means and standard deviations for the corresponding grade and year. Following WWC convention, we use the statewide student-level standard deviation (1.0) to calculate the standardized difference rather than the pooled standard deviation from the study sample.

Appendix Table D.1b. Baseline Equivalence for SY2021–22 Elementary and Middle Schools Sample

	Reenergize			Comparison			Mean Difference	Pooled SD	Hedges' g	Cox's Index
	N	Mean	SD	N	Mean	SD				
Elementary and Middle Schools (School-Grade Level Analysis)										
Mathematics										
Test Scores (Z-Score)[b]	20	-0.85	0.84	302	-0.66	0.73	-0.18	0.74	-0.25	-
Proficient Rate	20	0.29	0.16	302	0.34	0.15	-0.05	0.15	-0.33	-0.14
% Hispanic	20	0.92	0.04	302	0.92	0.06	0.00	0.05	-0.01	0.00
Student Achievement	20	66.82	8.42	302	67.64	9.64	-0.82	9.58	-0.09	-
RLA										
Test Scores (Z-Score)[b]	22	-0.80	0.43	330	-0.75	0.66	-0.05	0.65	-0.08	-
Proficient Rate	22	0.31	0.10	330	0.31	0.12	-0.01	0.12	-0.06	-0.02
% Hispanic	22	0.92	0.04	330	0.92	0.06	0.00	0.06	0.04	0.02
Student Achievement	22	66.18	8.40	330	67.15	9.43	-0.98	9.37	-0.10	-

Outcome measures appear in italics. See Appendix C.1 for more information about the outcome measures. N is the total number of school-grades within each condition. The standardized difference is reported as Hedge's g for continuous outcomes (intervention-comparison difference divided by the pooled standard deviation) and Cox's Index for dichotomous outcomes.

[a] The STAAR data are comprised of school-level mean scores for each grade level. To combine data across grades, we converted the school-by-grade scores into standardized z-scores using the student-level statewide score means and standard deviations for the corresponding grade and year. Following WWC convention, we use the statewide student-level standard deviation (1.0) to calculate the standardized difference rather than the pooled standard deviation from the study sample.

[b] Student Achievement is a school accountability rating defined by the Texas Education Agency. For detailed information on each rating, see https://tea.texas.gov/sites/default/files/Adopted%202019%20Accountability%20Manual_final.pdf.

Appendix Table D.1c. Baseline Equivalence for High Schools Sample

	Reenergize			Comparison			Mean Difference	Pooled SD	Hedges' g	Cox's Index
	N	Mean	SD	N	Mean	SD				
High Schools										
Graduation (Rate)	2	0.95	0.02	30	0.93	0.04	0.02	0.04	0.52	0.25
TSI (Rate)	2	0.19	0.06	30	0.23	0.07	-0.04	0.07	-0.50	-0.13
Postsecondary enrollment (Rate)	2	0.46	0.03	30	0.44	0.09	0.02	0.08	0.18	0.04
% Hispanic	2	0.93	0.06	30	0.93	0.06	0.01	0.06	0.09	0.05

Outcome measures appear in italics. See Appendix C.1 for more information about the outcome measures. N is the total number of schools within each condition. The standardized difference is reported as Hedge's g for continuous outcomes (the baseline intervention-comparison difference divided by the pooled standard deviation) and Cox's Index for dichotomous outcomes.

Appendix Table D.1d. Baseline Equivalence for All Schools Sample

	Reenergize			Comparison			Mean Difference	Pooled SD	Hedges' g	Cox's Index
	N	Mean	SD	N	Mean	SD				
All Schools										
Attendance (Rate)	10	0.94	0.02	145	0.94	0.02	0.00	0.02	-0.16	-0.03
% Hispanic	10	0.93	0.04	145	0.92	0.06	0.00	0.06	0.07	0.03
% White	10	0.03	0.03	145	0.04	0.04	-0.01	0.04	-0.25	-0.16
% Black	10	0.03	0.01	145	0.03	0.04	0.00	0.04	-0.05	-0.04
% Asian	10	0.01	0.01	145	0.00	0.01	0.00	0.01	0.79	0.53
% Two or more Races	10	0.01	0.01	145	0.00	0.01	0.00	0.01	0.14	0.10
% American Indian	10	0.00	0.00	145	0.00	0.00	0.00	0.00	0.25	0.20
% Pacific Islander	10	0.00	0.00	145	0.00	0.00	0.00	0.00	0.20	0.28
% Economically Disadvantaged	10	0.88	0.09	145	0.87	0.08	0.01	0.09	0.11	0.05
% English Language Learner	10	0.18	0.08	145	0.21	0.11	-0.03	0.10	-0.25	-0.10
% Students with Disability	10	0.12	0.02	145	0.11	0.03	0.01	0.03	0.37	0.06
Overall Rating [a]	10	70.89	13.04	145	72.52	11.98	-1.64	12.04	-0.14	-
School Progress [a]	10	75.78	9.94	145	76.73	10.62	-0.95	10.58	-0.09	-
Academic Growth [a]	10	64.10	9.18	145	65.30	9.59	-1.20	9.56	-0.12	-
Relative Performance [a]	10	75.80	9.94	145	76.41	10.89	-0.62	10.83	-0.06	-
Closing the Gap [a]	10	62.33	18.93	145	62.79	16.93	-0.47	17.05	-0.03	-

Outcome measures appear in italics. See Appendix C.1 for more information about the outcome measures. N is the total number of schools within each condition. The standardized difference is reported as Hedge's g for continuous outcomes (the baseline intervention-comparison difference divided by the pooled standard deviation) and Cox's Index for dichotomous outcomes.

[a] Overall, School Progress, Academic Growth, Relative Performance, Closing the Gap, and Student Achievement are school accountability ratings defined by the Texas Education Agency. For detailed information on each rating, see https://tea.texas.gov/sites/default/files/Adopted%202019%20Accountability%20Manual_final.pdf.

Appendix D.2 Impact Findings Tables

Appendix Tables D.2a, D.2b, and D.2c provide impact estimate statistics for all confirmatory and exploratory outcomes. Appendix Table D.2d provides the impact estimate statistics for the supplemental analysis on the full three years of exposure of Reenergize.

Appendix Table D.2a. Impact of Reenergize on RLA and Mathematics (Elementary and Middle Schools)

Outcome Measure [a]	Comparison Group				Intervention Group				I/C Difference	Standard Error	Standardized Difference	p-value
	Sample Size[b]		Mean	Standard Deviation	Sample Size[b]		Model -Adj. Mean	Standard Deviation				
	# Clusters	# Students			# Clusters	# Students						
Mathematics Z-Score: Year 2 (2020–21) [Exploratory]	314	174,795	-0.80	0.69	22	13,266	-0.85	0.54	-0.04	0.11	-0.06	0.72
Mathematics Z-Score: Year 3 (2021–22) [Confirmatory]	302	163,382	-0.70	0.70	20	11,539	-0.81	0.60	-0.11	0.12	-0.15	0.36
Mathematics Proficiency: Year 2 (2020–21) [Exploratory]	314	174,795	0.16	0.14	22	13,266	0.15	0.11	-0.01	0.02	-0.05	0.80
Mathematics Proficiency: Year 3 (2021–22) [Exploratory]	302	163,382	0.25	0.15	20	11,539	0.23	0.12	-0.01	0.03	-0.09	0.59
RLA Z-Score: Year 2 (2020–21) [Exploratory]	342	192,911	-0.84	0.66	24	14,556	-0.90	0.48	-0.05	0.09	-0.08	0.54
RLA Z-Score: Year 3 (2021–22) [Confirmatory]	330	180,729	-0.80	0.72	22	12,636	-0.84	0.69	-0.04	0.09	-0.05	0.68
RLA Proficiency: Year 2 (2020–21) [Exploratory]	342	192,911	0.23	0.11	24	14,556	0.24	0.09	0.00	0.02	0.01	0.95
RLA Proficiency: Year 3 (2021–22) [Exploratory]	330	180,729	0.38	0.14	22	12,636	0.38	0.12	0.01	0.02	0.05	0.69

I/C difference = Intervention-Comparison Difference. The intervention-comparison difference may not be the difference between the intervention and comparison means due to rounding.

[a] Outcome is reported at the school-year-grade cluster level (# clusters = # of schools-grades).

[b] Analytic sample includes 128 unique elementary and middle schools (8 Reenergize and 120 comparison schools). Since the schools in our study sample do not include a consistent grade range, the number of school-grades is not a multiple of the number of elementary and middle schools in the sample.

Appendix Table D.2b. Impact of Reenergize on High School Graduation, College and Career Readiness, and Postsecondary Enrollment (High Schools)

Outcome Measure[a]	Comparison Group				Intervention Group				I/C Difference	Standard Error	Standardized Difference	p-value
	Sample Size[b]		Mean	Standard Deviation	Sample Size[b]		Model-Adj. Mean	Standard Deviation				
	# Clusters	# Students			# Clusters	# Students						
Graduation: Year 2 (2020–21) [Exploratory]	30	41,781	0.93	0.04	2	3,187	0.93	0.05	0.00	0.02	0.02	0.96
Graduation: Year 3 (2021–22) [Confirmatory][c]	-	-	-	-	-	-	-	-	-	-	-	-
College and Career Readiness: Year 2 (2020–21) [Exploratory]	30	41,781	-0.41	0.70	2	3,187	-0.23	0.88	0.19	0.47	0.26	0.69
College and Career Readiness: Year 3 (2021–22) [Confirmatory]	30	41,222	-0.27	0.89	2	3,147	-0.49	0.59	-0.22	0.48	-0.24	0.65
Postsecondary Enrollment: Year 2 (2020–21) [Exploratory]	30	41,781	0.36	0.10	2	3,187	0.38	0.12	0.02	0.04	0.23	0.58
Postsecondary Enrollment: Year 3 (2021–22) [Exploratory][c]	-	-	-	-	-	-	-	-	-	-	-	-

I/C difference = Intervention-Comparison Difference. The intervention-comparison difference may not be the difference between the intervention and comparison means due to rounding. Outcome data not yet available at the time of this report are represented by a dash.

[a] Outcome is reported at the school-year cluster level (# clusters = # of schools).

[b] Analytic sample includes 32 unique high schools (2 Reenergize and 30 comparison schools).

[c] Outcomes may be included in an update to the final evaluation report, as these are not expected to be available until Summer 2023.

Appendix Table D.2c. Impact of Reenergize on Attendance (All Schools)

Outcome Measure[a]	Comparison Group				Intervention Group				I/C Difference	Standard Error	Standardized Difference	p-value
	Sample Size[b]		Mean	Standard Deviation	Sample Size[b]		Model-Adj. Mean	Standard Deviation				
	# Clusters	# Students			# Clusters	# Students						
Attendance: Year 2 (2020–21) [Exploratory]	145	107,639	0.93	0.04	10	8,039	0.96	0.04	0.03	0.01	0.72	< 0.001
Attendance: Year 3 (2021–22) [Confirmatory][c]	-	-	-	-	-	-	-	-	-	-	-	-

I/C difference = Intervention-Comparison Difference. The intervention-comparison difference may not be the difference between the intervention and comparison means due to rounding. Outcome data not yet available at the time of this report are represented by a dash.

[a] Outcome is reported at the school-year cluster level (# clusters = # schools).

[b] Analytic sample includes all 160 unique study schools.

[c] Outcome may be included in an update to the final evaluation report, as these are not expected to be available until Summer 2023.

Appendix Table D.2d. Impact of Reenergize (Three-year exposure)

Outcome Measure[a]	Comparison Group				Intervention Group				I/C Difference	Standard Error	Standardized Difference	p-value
	Sample Size[b]		Mean	Standard Deviation	Sample Size[b]		Model-Adj. Mean	Standard Deviation				
	# Clusters	# Students			# Clusters	# Students						
Attendance: Year 2 (2020–21) [Exploratory]	43	26,843	0.93	0.04	3	1,992	0.96	0.06	0.03	0.01	0.67	0.01
Attendance: Year 3 (2021–22) [Confirmatory]	-	-	-	-	-	-	-	-	-	-	-	-
Mathematics Z-Score: Year 2 (2020–21) [Exploratory]	111	69,570	-0.69	0.72	8	5,383	-0.67	0.41	0.02	0.18	0.03	0.90
Mathematics Z-Score: Year 3 (2021–22) [Confirmatory]	106	65,624	-0.59	0.65	8	5,012	-0.72	0.44	-0.13	0.17	-0.21	0.43
Mathematics Proficiency: Year 2 (2020–21) [Exploratory]	111	69,570	0.17	0.16	8	5,383	0.18	0.08	0.01	0.04	0.08	0.75
Mathematics Proficiency: Year 3 (2021–22) [Exploratory]	106	65,624	0.25	0.16	8	5,012	0.22	0.09	-0.03	0.04	-0.17	0.47
RLA Z-Score: Year 2 (2020–21) [Exploratory]	126	78,566	-0.75	0.68	9	5,976	-0.76	0.45	-0.02	0.13	-0.02	0.90
RLA Z-Score: Year 3 (2021–22) [Confirmatory]	121	74,450	-0.69	0.71	9	5,517	-0.90	0.70	-0.20	0.14	-0.28	0.13
RLA Proficiency: Year 2 (2020–21) [Exploratory]	126	78,566	0.25	0.12	9	5,976	0.27	0.07	0.01	0.03	0.12	0.61
RLA Proficiency: Year 3 (2021–22) [Exploratory]	121	74,450	0.40	0.15	9	5,517	0.37	0.13	-0.03	0.03	-0.19	0.31

Three-year exposure schools = Intervention schools (n = 3) that implemented the Reenergize program across all three evaluation years as intended. I/C difference = Intervention-Comparison Difference. The intervention-comparison difference may not be the difference between the intervention and comparison means due to rounding. Outcome data not yet available at the time of this report are represented by a dash.

[a] Attendance is reported at the school-year cluster level (# clusters = # schools). Z-score and Proficiency outcomes are reported at the school-year-grade cluster level (# clusters = # of schools-grades).

[b] Analytic sample includes 48 unique schools (3 Reenergize and 45 comparison schools). Since the schools in our study sample do not include a consistent grade range, the number of school-grades is not a multiple of the number of elementary and middle schools in the sample.

References

Aragon, S. (2016). *Teacher shortages: What we know*. Education Commission of the States. https://eric.ed.gov/?id=ED565893

Charles Butt Foundation. (2022). *The 2022 Texas Teacher Poll: Persistent problems and a path forward.* https://charlesbuttfdn.org/wp-content/uploads/2022/09/2022-teacher-poll.pdf

de Brey, C., Musu, L., McFarland, J., Wilkinson-Flicker, S., Diliberti, M., Zhang, A., Branstetter, C., & Wang, X. (2019). *Status and trends in the education of racial and ethnic groups 2018* (NCES 2019-038). U.S. Department of Education, National Center for Education Statistics. https://nces.ed.gov/programs/raceindicators/

Dong, N., & Maynard, R. A. (2013). PowerUp!: A tool for calculating minimum detectable effect sizes and sample size requirements for experimental and quasi-experimental designs. *Journal of Research on Educational Effectiveness*, *6*(1), 24-67. https://doi.org/10.1080/19345747.2012.673143

Hackman, H. W. (2005). Five essential activities for social justice education. *Equity & Excellence in Education*, *38*(2), 103-109. https://doi.org/10.1080/10665680590935034

Hamilton, L., Halverson, R., Jackson, S., Mandinach, E., Supovitz, J., & Wayman, J. (2009). *Using student achievement data to support instructional decision making* (NCEE 2009-4067). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. https://ies.ed.gov/ncee/wwc/practiceguides

Heineke, S. F., & Polnick, B. (2013). Pave the way for coaches. *Journal of Staff Development, 34*(3), 48-53. https://learningforward.org/wp-content/uploads/2013/06/pave-the-way-for-coaches.pdf

Henderson, A. T., & Mapp, K. L. (2002). *A new wave of evidence: The impact of school, family, and community connections on student achievement.* National Center for Family and Community Connections with Schools, Southwest Educational Development Laboratory (SEDL). https://eric.ed.gov/?id=ED474521

Herman, R., Dawson, P., Dee, T., Greene, J., Maynard, R., Redding, S., & Darwin, M. (2008). *Turning around chronically low-performing schools: A practice guide* (NCEE #2008-4020). National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. https://ies.ed.gov/ncee/wwc/practiceguides

Jackson, Y., & McDermott, V. (2012). *Aim high, achieve more: How to transform urban schools through fearless leadership.* Association for Supervision and Curriculum Development (ASCD).

Intercultural Development Research Association. (n.d.). *Quality Schools Action Framework*. https://www.idra.org/change-model/quality-schools-action-framework/

Intercultural Development Research Association. (2020). *Education Cafe™*. https://www.idra.org/families-and-communities/education-cafe/

Learning Policy Institute. (2017). *The role of principals in addressing teacher shortages.* Learning Policy Institute. https://learningpolicyinstitute.org/product/role-principals-addressing-teacher-shortages-brief

Leithwood, K., Harris, A., & Hopkins, D. (2008). Seven strong claims about successful school leadership. *School Leadership and Management, 28*(1), 27-42. https://doi.org/10.1080/13632434.2019.1596077

National Center for Education Statistics, U.S. Department of Education. (2019). Table 204.20. English language learner (ELL) students enrolled in public elementary and secondary schools, by state: Selected years, fall 2000 through fall 2017. *2019 Digest of Education Statistics.* https://nces.ed.gov/programs/digest/2019menu_tables.asp

National Center for Education Statistics, U.S. Department of Education. (2019). Table 204.40. Children 3 to 21 years old served under Individuals with Disabilities Education Act (IDEA), Part B, by race/ethnicity and age group: 2000-01 through 2018-19. *2019 Digest of Education Statistics.* https://nces.ed.gov/programs/digest/2019menu_tables.asp

National Center for Education Statistics, U.S. Department of Education. (2019). Table 204.70. Number and percentage of children served under Individuals with Disabilities Education Act (IDEA), Part B, by age group and state or jurisdiction: Selected years, 1990-91 through 2018-19. *2019 Digest of Education Statistics.* https://nces.ed.gov/programs/digest/2019menu_tables.asp

National Center for Education Statistics, U.S. Department of Education. (2019). Table 233.28. Percentage of students receiving selected disciplinary actions in public elementary and secondary schools, by type of disciplinary action, disability status, sex, and race/ethnicity: 2013-14. *2019 Digest of Education Statistics.* https://nces.ed.gov/programs/digest/2019menu_tables.asp

National Center for Education Statistics, U.S. Department of Education. (2019). Table 302.20. Percentage of recent high school completers enrolled in college, by race/ethnicity: 1960 through 201. *2019 Digest of Education Statistics.* https://nces.ed.gov/programs/digest/2019menu_tables.asp

National Center for Education Statistics, U.S. Department of Education. (2019). Table 302.44. Percentage distribution of all 2009 9th-graders and those who ever attended a postsecondary institution, by socioeconomic status in 2009 and selected postsecondary outcomes: 2016. *2019 Digest of Education Statistics.* https://nces.ed.gov/programs/digest/2019menu_tables.asp

Perilla, N. (2014). Leading the future: Rethinking principal preparation and accountability frameworks. *Harvard Journal of Hispanic Policy, 26*(1), 59-69.

Psencik, K. (2011). *The coach's craft: Powerful practices to support school leaders.* Learning Forward.

Price, C. (2020). Example study design plan for a comparative short interrupted time series (C-SITS) design: An i3/EIR technical assistance guidance document. Abt Associates.

Richardson, J. (2010). Equitable education. *Principal Leadership, 10*(07), 68-71.

Smith, P. A., Escobedo, P., & Kearney, W. S. (2020). Principal influence: A catalyst for positive school climate. *International Journal of Education Policy & Leadership, 16*(5). http://journals.sfu.ca/ijepl/index.php/ijepl/article/view/961 doi: 10.22230/ijepl.2020v16n5a961

Texas Education Agency. (2022). *Four-Year Graduation and Dropout Data, Class of 2021.* https://tea.texas.gov/reports-and-data/school-performance/accountability-research/completion-graduation-and-dropout/four-year-graduation-and-dropout-data-class-of-2021

Texas Education Agency. (2021). *2020–21 Texas Academic Performance Report (TAPR) Glossary.* https://rptsvr1.tea.texas.gov/perfreport/tapr/2021/glossary.pdf

Texas Education Agency. (2020). *2019–20 Comprehensive Texas Academic Performance Report Glossary*. https://rptsvr1.tea.texas.gov/perfreport/tapr/2020/glossary.pdf#page=8

Texas Education Agency. (2019). *Snapshot 2019: State Totals.* https://rptsvr1.tea.texas.gov/perfreport/snapshot/2019/state.html

Texas Education Agency. (2019). *2019 Accountability manual for Texas public school districts and campuses.* https://tea.texas.gov/sites/default/files/Adopted%202019%20Accountability%20Manual_final.pdf

Theoharis, G., & Brooks, J. S. (Eds.). (2012). *What every principal needs to know to create equitable and excellent schools*. Teachers College Press.

Urban Education Institute at UTSA. (2020). *Teaching and learning in the time of COVID-19. Research Brief: Food insecurity, digital divide, work, & caregiving.* https://uei.utsa.edu/_files/pdfs/DistanceLearningBrief3-9-25-20.pdf

Wallace Foundation. (2013). *The school principal as leader: Guiding schools to better teaching and learning.* https://www.wallacefoundation.org/knowledge-center/pages/the-school-principal-as-leader-guiding-schools-to-better-teaching-and-learning.aspx

What Works Clearinghouse. (2020). *What Works Clearinghouse standards handbook, version 4.1*. U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance. https://ies.ed.gov/ncee/wwc/handbooks

What Works Clearinghouse. (2022). *What Works Clearinghouse procedures and standards handbook, version 5.0*. U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance (NCEE). https://ies.ed.gov/ncee/wwc/handbooks

What Works Clearinghouse. (2021). *What Works Clearinghouse study review protocol version 1.0*. U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance (NCEE). https://ies.ed.gov/ncee/wwc/handbooks