MEMORANDUM July 18, 2019

TO: Staci Taylor Fullmighter
Officer, Student Assessment

FROM: Carla Stevens
Assistant Superintendent, Research & Accountability

SUBJECT: **A Quasi-Experimental Study On The Utility And Efficacy Of Lead4ward In HISD Toward Improving Instructional Planning And Student Achievement, 2018–2019**

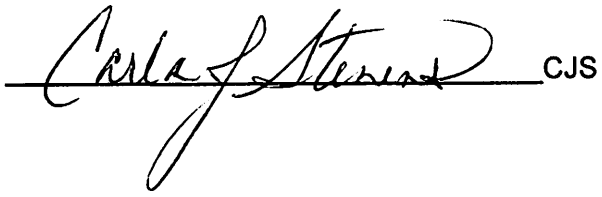
CONTACT: Carla Stevens, 713-556-6700

Attached is a copy of the program evaluation for the Student Assessment Department, measuring the utility and efficacy of lead4ward toward facilitating meaningful instructional planning, and improving the STAAR performance of students whose teachers participated in lead4ward professional development during the 2018–2019 academic year. A web-based survey was distributed to educators who attended lead4ward sessions to gather their perceptions. Moreover, paired samples t-tests and matching techniques were used to create comparable treatment and control groups to assess students' test performance.

Key findings include:

- Survey responses of 340 teachers revealed that lead4ward was closely aligned to TEKS (87.1), with 69.8% noting that the Field Guides were helpful in planning and implementing lessons well.
- Respondents perceived that lead4ward resources improved proficiency among students "more than a little" in mathematics, English language arts/reading, science, social studies, teacher induction, and intervention/RTI.
- Learning Videos, Field Guides, State Accountability Quicklooks, and Test Accessibility and Special Education Resources were rated as "fairly useful."
- Statistically significant increases were found in the STAAR 3–8 mean math scale scores of student groups as they progressed from 3rd to 4th through 7th to 8th grades. Small to medium effects were found among 3rd to 5th and 6th to 8th-grade students in reading.
- Propensity-score matching yielded no statistically significant differences in the 2019 English I and Algebra I EOC mean scale scores of treatment and control-group students, after controlling for background characteristics and 8th-grade reading and math scores. However, gains in favor of treatment group students were observed on the Algebra I EOC.
- Implications of the research included positive benefits of lead4ward based on students' test scores and educators' perceptions.

Further distribution of this report is at your discretion. Should you have any further questions, please contact me at 713-556-6700.



CJS

Attachment

cc: Margarita Gardea
Michael Dorsey
Courtney Busby

HOUSTON INDEPENDENT SCHOOL DISTRICT

RESEARCH

Educational Program Report

A QUASI-EXPERIMENTAL STUDY ON THE UTILITY AND EFFICACY OF LEAD4WARD IN HISD TOWARD IMPROVING INSTRUCTIONAL PLANNING AND STUDENT ACHIEVEMENT, 2018-2019

HISD | Research and Accountability
ANALYZING DATA, MEASURING PERFORMANCE.

2019 BOARD OF EDUCATION

Diana Dávila
President

Holly Maria Flynn Vilaseca
First Vice President

Elizabeth Santos
Second Vice President

Sergio Lira
Secretary

Susan Deigaard
Assistant Secretary

Wanda Adams
Jolanda Jones
Rhonda Skillern-Jones
Anne Sung

Grenita Lathan, Ph.D.
Interim Superintendent of Schools

Carla Stevens
Assistant Superintendent
Department of Research and Accountability

Venita Holmes, Dr. P.H.
Research Manager

Tamara Elijah
Applications Developer

Houston Independent School District
Hattie Mae White Educational Support Center
4400 West 18th StreetHouston, Texas 77092-8501

www.HoustonISD.org

It is the policy of the Houston Independent School District not to discriminate on the basis of age, color, handicap or disability, ancestry, national origin, marital status, race, religion, sex, veteran status, political affiliation, sexual orientation, gender identity and/or gender expression in its educational or employment programs and activities.

HISD | Research and Accountability
ANALYZING DATA, MEASURING PERFORMANCE.

EVALUATION REPORT
BUREAU OF PROGRAM EVALUATION

A Quasi-experimental Study on the Utility and Efficacy of lead4ward in HISD toward Improving Instructional Planning and Student Achievement, 2018–2019

Prepared by Venita R. Holmes, Dr.P.H.

Abstract

This program evaluation assessed the utility and efficacy of lead4ward toward facilitating meaningful instructional planning, and improving STAAR reading, English language arts, and math performance of students whose teachers participated in lead4ward professional development. A web-based survey on HISD's HUB yielded responses from 340 educators who had direct exposure to lead4ward. The highest percentages of respondents found that lead4ward was closely aligned to TEKS (87.1). In addition, 79.0% of respondents indicated that they used lead4ward in grade level, developmental planning, and PLCs meetings. Nearly 70.0% of respondents noted that the Field Guides were helpful in planning and implementing lessons well (69.8%). Respondents found that lead4ward resources improved proficiency among students "more than a little" in mathematics, English language arts/reading, science, social studies, teacher induction, and intervention/RTI. Instructional Tools were found to be the "most useful" among tools. Performance Standards, Data Tools, Academic Vocabulary Resources, lead4ward Reports in OnTrack, and lead4ward App were considered "more than fairly useful" by respondents. Learning Videos, Field Guides, State Accountability Quicklooks, and Test Accessibility and Special Education Resources were rated as "fairly useful". Statistically significant increases were found in the STAAR 3–8 mean math scale scores of student groups as they progressed from 3rd to 4th through 7th to 8th grades. Small to medium effects were found among 3rd to 5th and 6th to 8th-grade students in reading. Propensity-score matching yielded no statistically significant differences in the 2019 English I and Algebra I EOC mean scale scores of treatment and control-group students, after controlling for background characteristics and 8th-grade reading and math scores. However, gains in favor of treatment group students were observed on the Algebra I EOC. Implications included positive benefits of lead4ward based on students' test scores and educators' perceptions.

Background

For decades, school improvement has been at the forefront of education reform efforts to advance student learning and to facilitate the successful transition of students into the workforce. National and local educator organizations, including the National Commission on Teaching and America's Future (NCTAF), continued to emphasize the role of quality educator professional development toward meeting the demands of school systems with competing pressure to exist in this changing environment (Darling-Hammond & McLaughlin, 2016; Fuller, Hollingworth, & Pendola, 2017; NCTAF, 2016; Young, Winn, & Reedy, 2017).

The U.S. Department of Education's Every School Succeeds Act (ESSA, 2016) defined high-quality professional learning (**Figure 1**) as sustained, intensive, collaborative, job-embedded, data-driven, and classroom-focused (ESSA, 2016). ESSA expanded the scope of traditional professional development activities to all educators who work directly and indirectly with students, including principals and school support staff (Davis, 2017).

Figure 1: lead4ward data dive to identify students who need intervention using assessment data, 2018–2019

There was a call for programs and interventions—including those aimed at improving teacher practice—to be "evidence based," reflecting standards and practices that ranged from promising to strong. Educators must become the initiators of professional development and view professional development as an ongoing process to effectively provide coherent, improved instruction and foster student learning (Guskey & Sparks, 1996; Harfitt & Tavares, 2004; Stephens & Hartmann, 2004). A common professional development system for teachers and instructional support staff is vital to meet the current federal demand of a quality education for all students (Cohen, Spillane, & Peurach, 2018).

To that end, this program evaluation assessed the utility and efficacy of lead4ward professional development and resources toward improving student performance and helping HISD educators plan meaningful and quality instruction. Utility was assessed through inquiry about the usefulness of lead4ward and efficacy was assessed through an inquiry about the effectiveness of lead4ward. Both inquiries were survey-based. The evaluation also measured the impact of lead4ward professional development on students' mathematics and reading/English language arts. This led to the following research questions.

Research Questions:

1.What extent was lead4ward professional development and resources implemented in HISD?
2. How often did HISD educators utilize lead4ward resources during the 2018–2019 academic year?
3. How useful were lead4ward resources for planning meaningful instruction?
4. How helpful were Field Guides toward connecting state standards, building content knowledge, and for instructional planning?
5. How effective were lead4ward resources toward improving students' proficiency in content areas, for special populations, teacher induction, and intervention/response to intervention (RTI)?
6. What were HISD educators' overall perceptions of lead4ward?
7. What was the impact of lead4ward on the mathematics and reading/English language arts performance of students whose teachers participated in lead4ward professional development during the 2018–2019 academic year?

Foremost, the alignment of lead4ward to Darling-Hammond, Hyler, and Gardner (2017) criteria for effective professional development was considered. The researchers noted that effective professional development consists of "structured professional learning that results in changes in teacher practices and improvements in student learning outcomes" (p. v). The alignment of lead4ward with Darling-Hammond et al.'s work was captured from in-depth interviews with the Department of Student Assessment administrators and presented below.

Alignment of the HISD lead4ward Professional Development Model to the Effective Teacher Model[1]

Content Focused

Teachers were provided sessions related to their content areas, including reading, math, science, and social studies. Special education teachers were exposed to training materials to help them connect standards to instruction for students receiving special education services.

1 Darling-Hammond, Hyler, & Gardner. (2017). Effective Teacher Professional Development. Palo Alto, CA: Learning Policy Institute

Figure 2: Curriculum and Intervention staff engage in learning to support Achieve 180 campuses, 2018–2019

Collaboration

The HISD Student Assessment Department partnered with the department of Curriculum and Development to provide a cohesive professional development strategy. Curriculum staff provided guidance in the core content areas, Leadership Development shared the system with principals who shared the information with their staff. Student Assessment supported understanding of the data and data reporting. Future plans are to incorporate overarching sessions with campus leaders, then, build follow-up sessions with Tier 2 campus leaders, including assistant principals, instructional specialists, and deans, to strengthen collaboration among teachers at campuses.

Models of Effective Practice

HISD campus administrators selected champions for each academic content area to function as the site-based leadership team. The sessions embedded instructional strategies, through the Instructional Strategies Playlist within learning sessions. The training-of-trainers model was used to bring data back to campuses. Capacity-building among district leaders was emphasized to support the continuation of the work. For sustainability of this approach in the district, HISD Student Assessment staff plans to conduct one-on-one refresher courses (**Figure 2)**.

Active Learning

lead4ward professional development sessions were designed to be interactive and engaging. HISD worked with lead4ward staff to provide working sessions. Participants had time to plan how to integrate the new learning at their campuses. For example, during "Rockin Reviews," teachers developed plans on how to apply learning relative to the state assessment.

Coaching and Expert Support

Teacher Development Specialists (TDS) were considered experts in core content areas, and the Data-Driven Instruction Specialists (DDIS) were the expert specialists for data reporting. DDIS helped Achieve 180 campuses, who serve under performing students, use data in professional learning communities (PLCs). HISD Student Assessment helped campuses learn how to use lead4forward resources to prioritize instruction.

Feedback and Reflection

Questioning was built into sessions to allow educators to reflect on their current practices. Collaboration among colleagues included time to reflect on best practices and build plans to take back to campuses.

Sustained Duration

Teachers had an opportunity to develop deeper skills sets during the lead4ward professional development. Achieve 180 campuses had DDIS and TDS on site to help support their new learning. The importance of using the same vocabulary and using consistent concepts helped with transferability of information and sustainability across the district, while building more capacity with Tier 2 leaders.

Limitations

There were several limitations to the study. First, lead4ward professional development was implemented using a hybrid model, with classroom sessions facilitated by a lead4ward instructor, and additional support, primarily, provided by HISD Curriculum and Development and Student Assessment staff. Educators could also access materials online. Given the model, teachers' direct exposure to lead4ward professional development may not have been fully captured. The study was also limited to the number of HISD educators who chose to participate in the lead4ward Educator Survey and their willingness to provide accurate responses to survey items. In order to mitigate the limitations, the study design used strategies to expand the reach of prospective study participants, including the HISD employee portal and social media. Moreover, multiple measures were used to assess the impact of lead4ward in HISD.

Review of the Literature

Research has shown that creating an educational infrastructure with coordinated roles, structures, and resources to support and coordinate instruction, maintain instructional quality, and enable instructional improvement is pivotal toward school improvements (Cohen, Spillane, & Peurach, 2018; Woulfin, 2015). The U.S. Department of Education (n.d.) recognizes that teachers and other educators in schools are valued experts in building this infrastructure, prompting the need to engage in quality professional development opportunities that are informative and that drive systematic progress in student learning. Research emphasizes that "major changes required to reform schools cannot be accomplished without professional development nor can it be achieved with outdated models of professional development" (Chukwu, 2009, p. 112).

Professional development has the potential "to empower educators and communities of educators to make complex decisions; to identify and solve problems, to connect theory, practice, and student outcomes" (Katamei & Omwono, 2015, p. 112). Gusky (1997) highlighted that learning opportunities for teachers occur as they teach lessons, administer assessments, and review curriculum or instructional material. Desimone (2009) acknowledged that some of the most powerful teacher professional development models can occur in school communities, the teacher's own classroom, and through individual activities, such as engagement in online venues. On-the-job, teachers need professional development, which serves as capacity-building, to stay abreast of new teaching technologies (Katamei & Omwono, 2015).

Several studies have shown that effective professional development programs required from 50 to 80 hours of instruction, practice, and coaching before teachers arrived at mastery (French, 1997; Banilower, 2002; Yoon et al., 2007). Yoon et al. (2007) conducted meta-analyses to determine hours of professional development needed to improve student achievement. The study found that teachers who received well-designed professional development, for an average of 49 hours over 6 to 12 months, increased student achievement by as much as 21 percentile points. Documentation of more than 14 hours of professional development showed a positive and significant effect on student achievement, while workshops lasting 14 hours or less showed no statistically significant effect on student learning (Darling-Hammond, et al., 2009). The three studies that involved the least amount of professional development (5–14 hours) showed no statistically significant effects on student achievement. Considering these findings, it is imperative that school districts provide sufficient time for ongoing professional development to achieve changes in students' learning and behavior (Darling-Hammond, et al., 2009; Mortimore & Sammons, 1987).

At the same time, the rapid growth in hybrid forms of professional development that incorporate face-to-face, online resources, and media have the potential to extend learning opportunities for teachers and instructional support staff while transforming the classroom to one that is more engaging and effective (Walker, et al., 2012; Lawless & Pellegrino, 2007). In hybrid models, visualization tools are more accessible and manipulation of the data can be readily shared and adapted by communities of learners within school districts (Walker, et al., p. 422). However, more research is needed that links hybrid teacher professional development models to student achievement (Desimone, 2009). This study seeks to expand the body of research in this area, while assessing the effectiveness and utility of such a model.

Methods

Study Population and Sample

The study population consisted of all teachers, school administrators, and instructional support staff in HISD who had the opportunity to participate in lead4ward professional development. Two samples were generated from the study population based on (1) participation in the lead4ward Educator Survey, (2) participation in OneSource lead4ward training, and (3) having a direct link to students with specific background characteristics and STAAR reading and math data. Details regarding sample selection follows.

Study Sample 1 consisted of HISD staff who completed the web-based lead4ward Educator Survey on the HISD HUB. Survey access and messaging were provided to 12,457 educators with support from the Instructional Technology (IT) Department, HISD Communications, and the Department of Student Assessment. A total of 399 HISD staff started the survey; however, 340 staff completed the survey. The 59 staff who exited the survey indicated that they lacked exposure to lead4ward professional development. Consequently, survey analyses were based on the responses of the 340 staff who completed the survey. **Table 1** (**Appendix A**, p. 11) presents background characteristics of Study Sample 1.

Table 2. Type of Power Analysis: A priori: Compute Required Sample Size for Dependent Samples

Input Parameters		Output Parameters	
Effect size d	0.5	Noncentrality parameter δ	3.674234
a err prob (Type I error rate)	0.05	Critical t	2.005746
Power (1 – β err prob)	0.95	Df	53
		Total Sample size	54

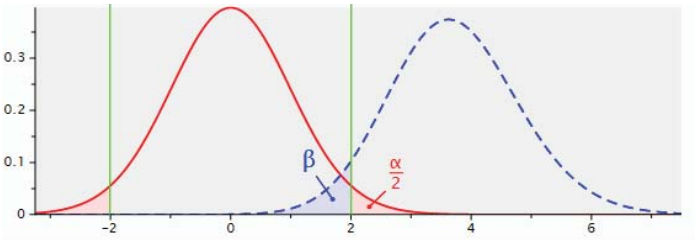


Figure 3: Critical t = 2.005746

Study Sample 2 was extracted from the HISD OneSource database that provided lead4ward course attendance information. These staff along with staff who completed the lead4ward Educator Survey were used to measure impact on student achievement. A total of 4,172 HISD staff met these criteria, with 3,464 staff extracted from OneSource training completion reports.

IBM Cognos was used to match teachers to students. G*Power software was used to estimate the sample size needed prior to the study to detect a statistically significant difference in students' test performance, considering that a difference truly existed (Faul, Erdfelder, Lang, & Buchner, 2007; Faul, Erdfelder, Buchner, & Lang, 2009; Pitner, Yu, & Brown, 2013). For a two-tailed dependent t-test, G*Power determined that a sample size of 54 was needed, with a critical t-value at 2.005, a significance factor of 0.05, beta value equal to 0.5, and degrees of freedom equal to 53. An effect size (d) of .5 is considered medium effect (Cohen, 1988). There was a 95% chance of correctly rejecting the null hypothesis that there was no difference in the academic performance of students whose teachers participated in lead4ward professional development. A graphical representation of the test parameters are shown in **Table 2** and **Figure 3**, with the sampling distribution as a dotted blue line, the population distribution represented by a solid red line, a red shaded area delineating the probability of a type 1 error, a blue area the type 2 error, and a pair of green lines demarcating the critical points t. A total of 3,135 teachers were found with 67,742 students to conduct the analyses.

Data Collection/Data Analyses

At the elementary and middle-school levels, paired t-test analyses were conducted to detect statistically significant differences in the English language version of students' 2018 State of Texas Assessments of Academic Readiness (STAAR) Grades 3–8 reading and mathematics scale scores (pretest) compared to their 2019 scale scores (posttest) on comparable assessments. The first administration of students' results were used in the analyses. The level of statistical significance was $p < .05$. Effect size analyses were also conducted using Cohen's d (Rosenthal, 1991). Interpretation of Cohen's d is: .2 = small effect; .5 = medium effect, and .8 = large effect (Cohen, 1988). According to the What Works Clearinghouse (n.d.), effect sizes of 0.25 standard deviations or larger are considered to be substantively important. Effect sizes at least this large are interpreted as a qualified positive (or negative) effect, even though they may not reach statistical significance.

The first test administration of students' spring 2019 STAAR English I and Algebra I End-of-Course (EOC) results were used as outcome measures for English language arts and math at the secondary level. First, propensity score matching was conducted using SPSS software to adjust for treatment effects. Propensity score matching is considered an alternative to the commonly used regression adjustment (Stuart, 2010). For this analysis, nearest-neighbor matching with replacement was used. The logic behind propensity score methods is that balance on observed covariates is achieved through careful matching on a single score – the estimated propensity of selecting the treatment (Stuart, 2010). The propensity score is defined as the probability of receiving treatment based on measured covariates. Covariates used for matching in this evaluation were gender, economic status, at risk, gifted/talented, and special education, along with students' eighth-grade 2018 STAAR reading and math scores. The quality of matches can be affected by the order in which subjects were selected for matching and the maximum permitted difference between matched subjects (the "caliper") (Lunt, 2014). The caliper for math was .25 standard deviations considering that Rosenbaum and Rubin (1985) used this caliper based on Cochran and Rubin's (1973) logistic regression model to predict exposure. A tighter caliper was used for reading to reduce bias, considering the large difference that existed between the mean reading scores of the treatment and control groups' scores before matching (Lunt, 2014). Matches were selected with replacement to improve imbalance between the groups, and in a randomized order.

Results

What extent was lead4ward professional development and resources implemented in HISD?

A variety of lead4ward resources were offered to HISD educators in professional development sessions (**Figure 4**). Educators learned how to use these resources to improve student achievement and to plan meaningful instruction for students. In general, Field Guides allowed professional learning communities (PLCs) to develop content knowledge around Texas Essential Knowledge and Skills (TEKS) standards. Instructional Tools offered strategies that educators could use to engage learners, provide practice, encourage interaction, and to better understand what students were thinking. These tools also allowed instructional leaders to become more knowledgeable of changes in content-area TEKS to ensure a smoother transition period in developing new skills.

Performance Reports depicted students' progress on the state assessment. Data Tools included the Leadership Report Card, Student Learning Reports, and Heat Maps to monitor students' progress. The Academic Vocabulary Resource exposed educators to terms used on the state assessment. Many of these reports were available through the OnTrack data system. The lead4ward App allowed educators convenient access to view and use these resources. lead4ward Videos provided an alternative means to receive information related to the state assessment standards

Figure 4: HISD lead4ward professional development resources, 2018-2019

and expectations. State Accountability Outlooks summarized the Texas Accountability System for campuses and the district. Tests Accessibility and Special Education Resources provided information related to the use of accessibility features and designated supports for state assessments.

The "Rockin Review" was among the more prevalent lead4ward professional development sessions (**Figure 5**). "Rocking Review" helped instructors build independent thinkers and problem solvers among students. Participants learned to create purposeful plans for review (daily review, reviewing for tests, etc.), transitioning from newly-introduced concepts to understanding, application, review, and transfer of information. The intent was to build confident students by providing strategies for transferring learning in multiple ways, with varying instructional strategies and classroom activities to make meaningful connections.

"lead4ward Instructional Support/Planning with Data" provided in-depth training of lead4ward resources, tools, and strategies to support schools in using data effectively and to plan instruction. "lead4ward Wave Goodbye, Say Hello New ELAR TEKS (K-5)" helped participants explore and compare the new and old TEKS to determine curriculum and instructional "keepers" and shifts. This course was also offered in various content areas.

Finally, "lead4ward Deep Dive" helped participants advance their use of lead4ward tools and processes, including the Quintile Report. Participants were able to identify evidence of student learning and design engaging experiences to address areas where both teachers and students have experienced difficulty. Campus leaders already experiencing success with lead4ward tools and processes learned how to expand implementation and refine their practice to advance student achievement.

During the 2018–2019 academic year, 3,464 HISD teachers, support staff, and administrators attended various lead4ward

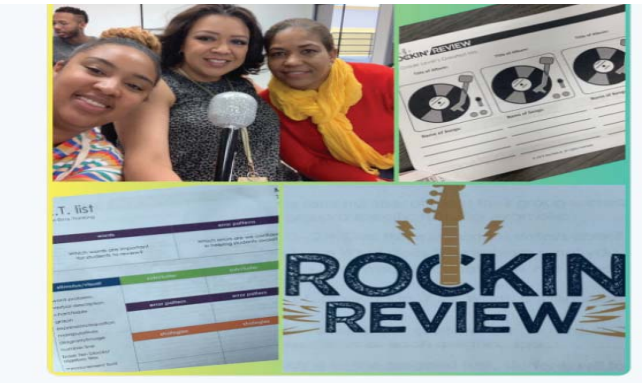


Figure 5: Elementary math staff learning how to identify a grade level's greatest hits through Rockin Review, 2018-2019

professional development courses. **Tables 3a** through **3f** in **Appendix B** (p. 12) provide the title and focus areas of the courses.

Descriptive statistics depicted in **Figure 6** show the number and percent of lead4ward professional development sessions by focus area. It is evident that 12 lead4ward courses offered during the 2018–2019 academic year focused on strategies and practices related to implementation and support (38%). A slightly higher number and percentage of content-area professional development courses addressed reading/ELA (English language arts) (n = 5, 16%) compared to other content areas.

Figure 7 presents the number and percent of HISD lead4ward professional development participants by focus area. The majority of HISD staff who participated in the professional development were exposed to strategies and practices related to implementation and support (n = 2,122 or 51%). Math and reading/ELA professional development sessions were the next highly-attended sessions (n = 874 or 21% and n = 579 or 14%, respectively). Science, social studies/U.S. history, and writing sessions were attended by less than 10% of the staff (6%, 5%, and 3%, respectively).

How often did HISD educators utilize lead4ward resources during the 2018–2019 academic year?

Data were gathered from HISD educators to estimate how often lead4ward resources were used during the 2018–2019 academic year. The results are presented in **Table 4** (**Appendix C**, p. 14). The highest percentage of survey participants that

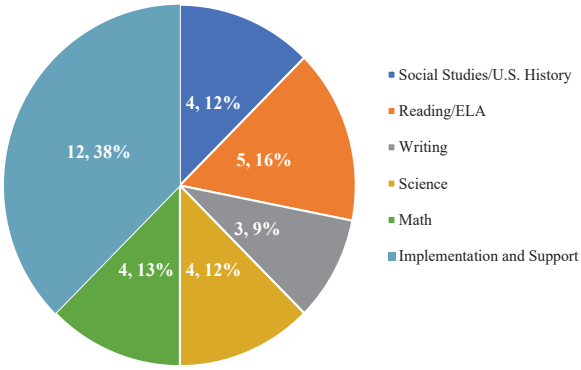


Figure 6: Number and percent of HISD lead4ward professional development sessions by focus area, 2018-2019

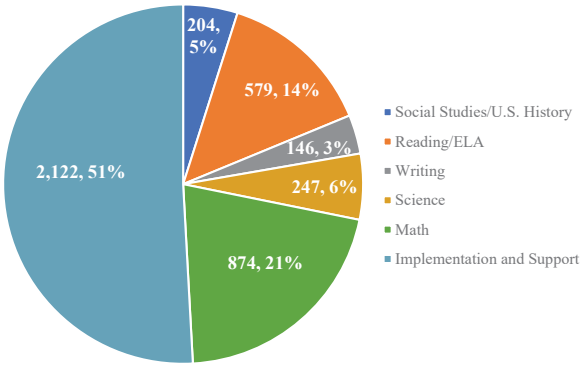


Figure 7: Number and percent of HISD lead4ward professional development participants by focus area, 2018–2019

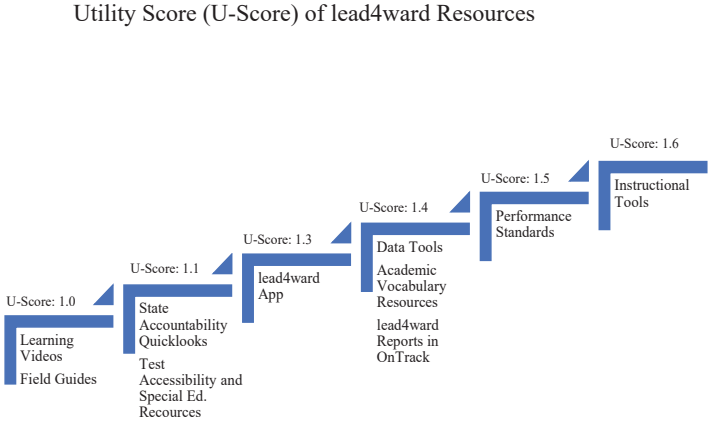


Figure 8: Usefulness of lead4ward resources, spring 2019
(Note: Scale ranged from -2 to +2)

indicated using lead4ward resources between "one and three times a month" included Field Guides (46.4%), Instructional Tools (42.0%), Performance Standards (47.4%), Data Tools (39.1%), Academic Vocabulary Resources (39.6%), lead4ward Reports in OnTrack (40.7%), and the lead4ward App (40.7%). lead4ward resources used "everyday" were more likely to be Instructional Tools (13.6%) and Academic Vocabulary Resources (11.1%). The majority of survey respondents indicated that they "never used" Learning Videos (56.7%), closely followed by Test Accessibility and Special Education Resources (49.1%) and State Accountability Quicklooks (46.2%).

How useful were lead4ward resources for planning meaningful instruction?

Table 5 (Appendix C, p. 14) shows survey participants' perceptions regarding how useful lead4ward resources were for planning meaningful instruction during the 2018–2019 academic year. The majority of respondents indicated that Instructional Tools (60.2%), Performance Standards (52.3%), and Field Guides (50.3%) were "extremely useful." Data Tools were perceived as "extremely useful" by 48.8% of respondents, while Academic Vocabulary Resources and lead4ward Reports in OnTrack were considered "extremely useful" by 45.7% and 45.8% of respondents, respectively. Between 32.2% and 40.5% of respondents revealed that they "never used" State Accountability Quicklooks, Test Accessibility and Special Education Resources, and Learning Videos.

A "Utility Score" (U-Score) was calculated based on response options ("extremely useful" = 2, "fairly useful" = 1, "not sure" = 0, "not very useful" = -1, and "not useful at all" = -2). "Never used this resource" was not included in the calculations. The Instructional Tools resource was considered the most useful among all tools rated in this evaluation (**Figure 8**, U-Score = 1.6 out of 2.0). Performance Standards, Data Tools, Academic Vocabulary Resources, lead4ward Reports in OnTrack, and lead4ward App were considered "more than fairly useful" by respondents. The U-Scores ranged from 1.3 to 1.5. The Learning Videos, Field Guides, State Accountability Quicklooks, and Test Accessibility and Special Education Resources were rated as "fairly useful" among respondents who used the resources (U-Score = 1.0 to 1.1 out of 2.0).

How helpful were Field Guides toward connecting state standards, building content knowledge, and for instructional planning?

Table 6 (Appendix C, p. 15) shows survey respondents' perceptions regarding the helpfulness of Field Guides toward connecting state standards, building students' content knowledge, and for instructional planning. The majority of respondents indicated that the Field Guides were "extremely helpful" in most areas addressed in the survey. The findings are presented in descending order, according to the percentage of respondents who indicated the Field Guides were "extremely helpful."

- Connecting TEA standards to the curriculum (60.6%);
- Building content knowledge with explanations, stimulus identification, and essential vocabulary (56.9%);
- Making connections to instructional implications (56.4%);
- Helping with instructional planning and PLCs (56.3%);
- Attaining insight into the type of mistakes students make (53.7%), and
- Providing relevant context that shows how each student expectation fits in the big picture (52.6%).

Although ratings of "not at all helpful" were low, the highest percentage of respondents rated the Field Guides as "not at all helpful" with instructional planning and PLCs (4.3%).

How effective were lead4ward resources toward improving students' proficiency in content areas, for special populations, teacher induction, and intervention/response to intervention (RTI)?

Table 7 (Appendix C, p. 15) shows survey participants' perceptions regarding the effectiveness of lead4ward resources toward improving student proficiency in content areas, for special populations, teacher induction, and intervention/RTI. The majority of respondents indicated that they were "not sure" whether lead4ward resources were effective toward improving student proficiency in social studies (57.1%) and science (50.9%). Nearly half of the respondents indicated that they were "not sure" whether the resources improved the proficiency of special

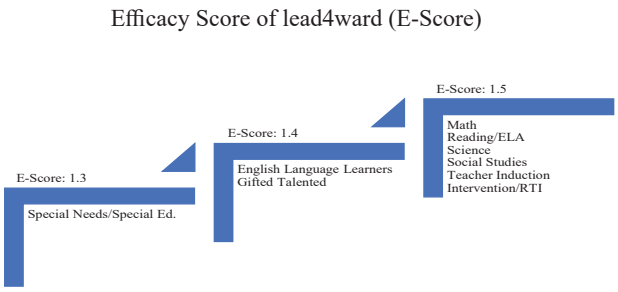


Figure 9: Effectiveness of lead4ward resources, spring 2019
Note: Scale ranged from -1 to +2)

needs and special education students (48.5%). However, 44.3% of survey respondents noted that lead4ward resources improved student proficiency "a lot" for intervention/response to intervention (RTI) services and teacher induction (40.7%).

Relative to content areas, Table 7 (p. 15) shows that 40.4% of respondents expressed that the lead4ward resources improved student proficiency in math "a lot," followed by reading/English language arts (36.0%), science (31.3%), and social studies (25.4%). There were perceptions that lead4ward resources were more successful toward improving proficiency "a lot" among English language learners (39.9%) compared to gifted/talented students (34.9%), and special needs/special education students (27.2%).

An "Efficacy Score" (E-Score) was calculated based on response options ("improves proficiency a lot" = 2, "improves proficiency a little" = 1, "makes no difference" = 0, "hinders proficiency" = -1). "Not sure" was not included in the calculations to reduce skewness of the results. Respondents mostly considered the lead4ward resources improved proficiency among students "more than a little" in mathematics, reading/English language arts, science, social studies, teacher induction, and intervention/RTI (**Figure 9**, E-Score = 1.5 out of 2.0). The E-Score for English language learners and gifted/talented students was 1.4 out of 2.0, and 1.3 out of 2.0 for special needs/special education students, which indicated that lead4ward resources also improved proficiency among these groups "more than a little."

What were HISD educators' overall perceptions of lead4ward?

HISD educators' overall perceptions of lead4ward are shown in **Table 8** (Appendix C, p. 16). A summary, which reflects the percentage of respondents in agreement (strongly agree or agree), is depicted in **Figure 10**.

Overall, 81.3% of respondents noted that lead4ward was effective in facilitating instructional planning. The highest percentages of respondents found that lead4ward was closely aligned to TEKS (87.1); was user-friendly (80.1%), and was used in grade level, developmental planning, and PLC meetings (79.0%). Nearly 70.0% of respondents indicated that the Field Guides were helpful in planning and implementing lessons well (69.8%). In addition, 74.0% of respondents noted that lead4ward helped them to prioritize, plan, and implement effective teaching strategies aligned to STAAR.

What was the impact of lead4ward on the mathematics and reading/English language arts performance of students whose teachers participated in lead4ward professional development during the 2018–2019 academic year?

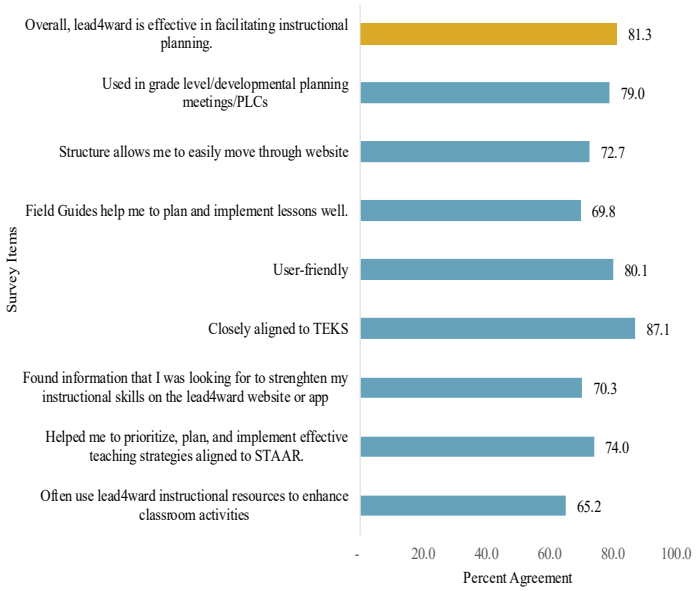


Figure 10: Survey respondents' overall percent agreement ratings of lead-4ward, spring 2019

Multiple methods were applied to determine the impact of lead4ward on students' mathematics and reading/English language arts performance. A paired t-test was conducted using spring 2018 (pretest) and spring 2019 (posttest) STAAR 3–8 reading and mathematics scale scores of students whose teachers participated in lead4ward professional development during the 2018–2019 academic year. Data for first-time test takers who were administered the English version of the tests in both years were included in the analyses. Reading and math results can be found in **Figure 11** and **Figure 12**, respectively, and in **Appendix D** (**Tables 9a** and **9b**, respectively, p. 17). Data depicted in graphs were rounded to the nearest whole number.

Figure 11 shows the paired STAAR English reading t-test results of the student sample. There was a statistically significant increase in the mean reading scale scores as students progressed from 3rd to 4th, 4th to 5th, 6th to 7th, and 7th to 8th grades. The largest increase was found for the 3rd to 4th-grade student group (M = 1424 vs. M = 1514, p = .000; Mean diff. = 90), while the low-

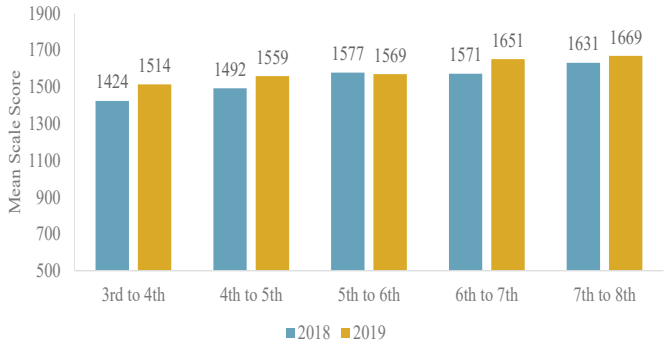


Figure 11: Paired t-tests, English reading STAAR 3–8, 2018 vs. 2019

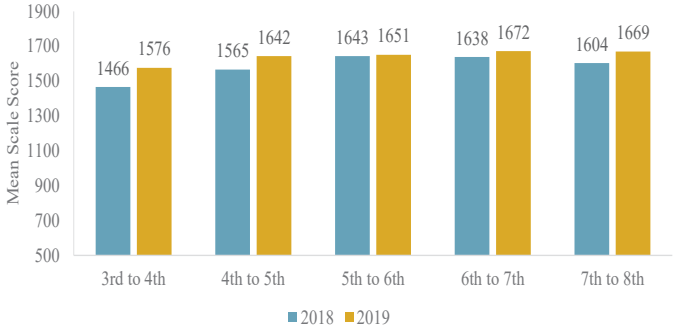


Figure 12: Paired t-tests, English math STAAR 3–8, 2018 vs. 2019

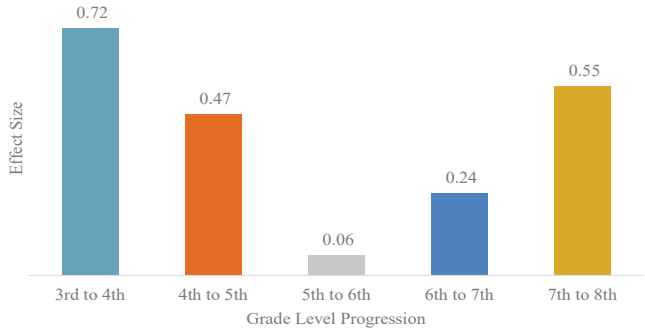


Figure 14: Cohen's d effect sizes, English math STAAR 3–8, spring 2018

est increase was found for the 7th to 8th-grade student group (M = 1631 vs. M = 1669, p = .000; Mean diff. = 38). At the same time, there was a statistically significant decrease in the mean reading scale scores of students as they progressed from 5th to 6th grades (M = 1577 vs. M = 1569, p = .000; Mean diff. = -8).

Statistically significant increases were found in students' STAAR math assessment performance as they progressed from grades three through eight. The largest increase was reflected in the performance of the 3rd to 4th-grade student group (M = 1466 vs. M = 1576, p = .000; Mean diff. = 110) (Figure 12), while the lowest increase was noted for the 5th to 6th-grade student group (M = 1643 vs. M = 1651, p = .000; Mean diff. = 8).

Cohen's d statistics were calculated to determine the effect size of lead4ward professional development on students' 2018 compared to 2019 reading and math English STAAR 3–8 performance. The findings indicate a medium effect in students' reading scores as they progressed from 3rd to 4th (d = .63), 4th to 5th (d = 0.46), and 6th to 7th grades (d = 0.57). A small effect was detected as students progressed from 7th to 8th grades (d = 0.26), and a negligible effect as students progressed from 5th to 6th grades (d = -0.05) (**Figure 13**).

In math, the effect size analyses revealed a medium effect of lead4ward professional development on students' math STAAR performance as they progressed from 3rd to 4th (d = .72), 4th to 5th (d = 0.47), and 7th to 8th grades (d = 0.55). A small effect was found in math as students progressed from 6th to 7th grades (d = 0.24), and a negligible effect was detected as they progressed from 5th to 6th grades (d = 0.06) (**Figure 14**).

At the secondary level, propensity score, nearest-neighbor matching, with replacement, was used to create a treatment and a control group to further assess the impact of lead4ward professional development on students' academic performance. This matching procedure helped to ensure that groups were comparable without randomization. The treatment group consisted of students whose teachers attended lead4ward professional development and the control group was gathered from a larger pool of HISD secondary students. Students used in the matching procedure were required to have specific covariates, including gender, economic status, at risk, gifted/talented, special education, and eighth-grade 2018 STAAR reading and math scores. Students' performance on the spring 2019 Algebra I EOC exam, first test administration, was used to measure treatment impact in math. Students' 2019 English I EOC, first test administration results were used to measure treatment impact in English language arts.

Before propensity matching, the sample sizes for the treatment and control groups using the Algebra I EOC were 3,230 and 4,037, respectively (**Appendix E**, **Table 10a**, p. 18). After matching, the sample sizes for the groups were 3,230 and 3,204, respectively. Descriptive statistics on matched groups for the Algebra I EOC are depicted in **Table 10b** (p. 18). A summary is depicted in **Figure 15**.

The mean 2019 Algebra I EOC scale score was lower for the treatment group relative to the control group before matching (M = 3958, SD = 503.2 vs. M=3982, SD = 517.5) (Figure 15). However, propensity score matching yielded a slightly higher 2019 mean Algebra 1 scale score for the treatment group compared to the control group (M=3958, SD = 503.22 vs. M = 3949, SD = 506.6). While the results estimated a positive effect

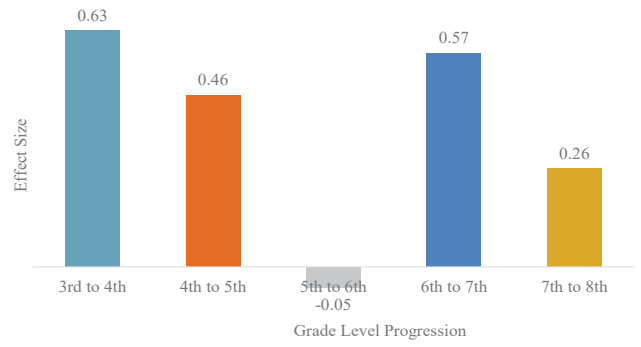


Figure 13: Cohen's d effect sizes, English reading STAAR 3–8, spring 2018 and 2019

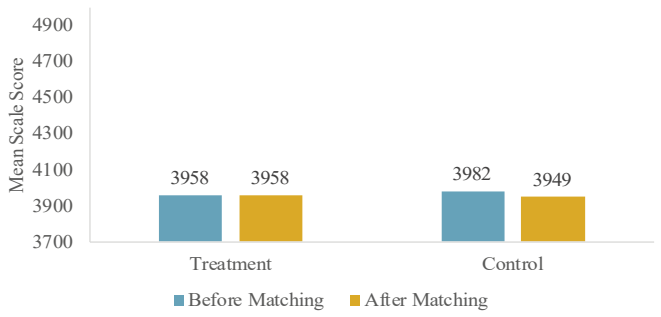


Figure 15: Mathematics propensity score matching results based on spring 2019 Algebra I EOC, first test administration

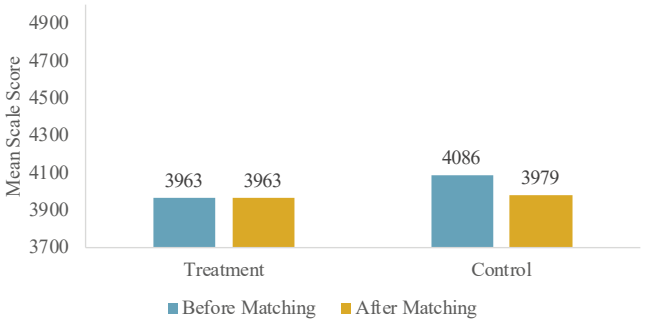


Figure 16: English language arts propensity score matching results based on spring 2019 English I EOC, first test administration

of lead4ward professional on students' Algebra I performance, an independent t-test found the difference between the groups was not statistically significant at the p <. 05 level (p = .733).

The sample sizes used to detect the impact of lead4ward on students' STAAR English I EOC was 4,045 for the treatment group and 5,971 for the control group before matching, and 4,045 for the treatment group and 4,020 for the control group after matching (Appendix E, Table 10a, p. 18). Descriptive statistics on matched groups for the English I EOC are depicted in **Table 10c** (p. 18).

The mean 2019 English I EOC scale score was lower for the treatment group compared to the control group before matching (M = 3963, SD = 543.5 vs. M = 4086, SD = 596.1) (**Figure 16**). Propensity score matching yielded a higher 2019 mean English 1 EOC scale score for the control group compared to the treatment group (M = 3979, SD = 550.7 vs. M = 3963, SD = 543.5). However, the difference between the groups was not statistically significant at the p <. 05 level (p = .193).

Further analyses of matched student data was conducted and reflected in **Figure 17**. The mean math (Algebra I EOC) scale score difference between the treatment and control groups was 24 points in favor of the control group before matching, but changed to 9 points in favor of the treatment group after matching. Figure 17 also reveals that mean scale score difference in students' English language arts (English I EOC) performance was 123 points in favor of the control group before matching, but decreased to 16 points after matching.

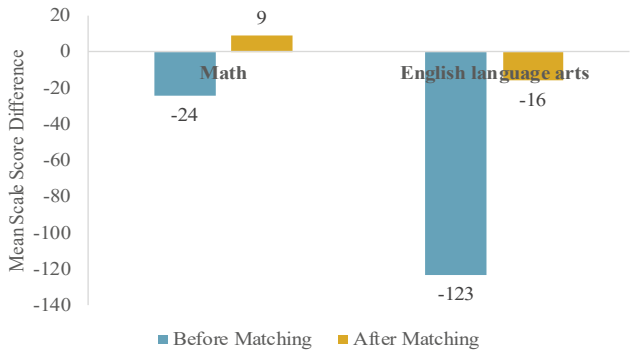


Figure 17: Mean scale score differences before and after propensity score matching in math (Algebra I EOC) and English language arts (English I EOC), first test administration, spring 2019

Discussion

The implementation of effective educator professional development is critical toward fostering academic learning and achievement among students. Various forms of professional development have been implemented in school districts across the United States, including hybrid models that incorporate access to materials and resources that are aligned to state standards, using technology and face-to-face collaboration.

During the 2018–2019 academic year, HISD expanded the use of lead4ward to facilitate meaningful instructional planning for educators districtwide. lead4ward resources provided strategies to address students' academic needs in content areas, including math, science, reading/English language arts, and social studies. Educators were equipped with information to support teacher induction, intervention/RTI, and to improve the academic performance of all HISD students, including English language learners, students with special needs, and gifted/talented students.

This study found that an, overwhelming, majority of the 2018–2019 educator sample who had direct exposure to lead4ward professional development perceived benefits from implementation. The instructional resources, including the Field Guides, were often used to enhance classroom activities, and for grade level, developmental planning, and PLCs meetings. The Instructional Tools were found to be the "most useful" among all tools rated in this evaluation.

Moreover, survey respondents noted that lead4ward resources improved proficiency among students in the major content areas, including mathematics, English language arts/reading, science, social studies "more than a little." Resources were also found to be effective for teacher induction and intervention/RTI. Survey respondents rated the Learning Videos, State Accountability Quicklooks, and Test Accessibility and Special Education Resources as "fairly useful." More information is needed to determine which features of these resources were of particular concern for educators.

The measurement of the academic benefits of lead4ward on student achievement was accomplished through multiple methods. Paired samples data of students whose teachers participated in lead4ward professional development detected statistically significant increases in their STAAR 3–8 mean math scale scores as they progressed from 3rd to 4th through 7th to 8th grades, with effect sizes ranging from negligible to medium effects. Small to medium effects were found among 3rd to 5th and 6th to 8th grade students in reading. Propensity-score matching yielded no statistically significant differences in the 2019 English I and Algebra I EOC mean scale scores of treatment and control-group students, after controlling for background characteristics and 8th grade reading and math STAAR scores. However, positive changes in test performance were noted among students in Algebra I whose teachers participated in lead4ward after matching. Considering these findings, school administrators should ensure that all teachers have access to lead4ward professional development and resources, particularly for special populations and interventions.

The research has shown that effective professional development programs require from 50 to 80 hours of instruction, practice, and coaching before teachers arrive at mastery. Plans to expand capacity for additional professional learning in HISD

may help educators build on their current knowledge. Additional opportunities to offer exposure to lead4ward should be sought at Special Education conferences and through planning and intervention sessions for special populations. This will help to ensure the utility and efficacy of lead4ward tools and resources among all student groups in the district.

Future research may examine the perceptions and the use of lead4ward, independently, at the elementary and secondary levels. In addition, the research could examine expectations of leadership regarding lead4ward use at their campuses as an effective tool to facilitate planning and instruction.

References

Banilower, E. (2002). Results of the 2001-2002 study of the impact of the local systemic change initiative on student achievement in science. Arlington, VA: National Science Foundation.

Chukwu, J.O. (2009). Quality Education and the Teacher in the Nigerian Society. Nigerian Journal of Sociology of Education, 111(3), 190 – 196.

Cochran, W. G., & Rubin, D. B. (1973). Controlling bias in observational studies: A review. The Indian Journal of Statistics, Series A. 35, 417–466.

Cohen, J. (1988). Statistical power analysis for the behavioral sciences. Hillsdale,NewJersey:LawrenceErlbaumAssociates.

Cohen, D. K., Spillane, J. P., & Peurach, D. J. (2018). The Dilemmas of Educational Reform. Educational Researcher, 47(3), 204–212. Retrieved from https://doi-org.contentproxy.phoenix.edu/10.3102/0013189X17743488

Darling-Hammond, L., Hyler, M., & Gardner, M. (2017). Effective Teacher Professional Development. Palo Alto, CA: Learning Policy Institute.

Darling-Hammond, L. & McLaughlin, M. (2016). Policies that support professional development in an era of reform. Retrieved from https://www.kappanonline.org/policies-that-support-professional-development-in-an-era-of-reform/

Darling-Hammond, L., Wei, R., Andree, A, Richardson, N., & Orphanos, S. (2009). Professional Learning in the Learning Profession: A Status Report on Teacher Development in the United States and Abroad. Retrieved from https://learningforward.org/docs/default-source/pdf/nsdcstudy2009.pdf

Davis, M. (2017). School Districts Update Professional Development. Education Week. Vol. 36, Issue 29, Pages 3-4. Retrieved from https://www.edweek.org/ew/articles/2017/04/26/school-districts-update-professional-development.html

Desimone, L. M. (2009). Improving impact studies of teachers' professional development: Toward better conceptualizations and measures. Educational Researcher, 38(3), 181-199. Retrieved from https://search-proquest-com.contentproxy.phoenix.edu/docview/216902889?accountid=134061

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behavior Research Methods, 39, 175-191.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. Behavior Research Methods, 41, 1149-1160.

French, V. W. (1997). Teachers Must Be Learners, Too: Professional Development and National Teaching Standards. NASSP Bulletin, 81(585), 38–44. Retrieved from https://doi.org/10.1177019263659708158507

Fuller, E. J., Hollingworth, L., & Pendola, A. (2017). The Every Student Succeeds Act, State Efforts to Improve Access to Effective Educators, and the Importance of School Leadership. Educational Administration Quarterly, 53(5), 727–756. Retrieved from https://doi.org/10.1177/0013161X17711481

Guskey, T.R. (1997). Research needs to link professional development and student learning. Journal of Staff Development, 18(2), 36-40.

Guskey, T. R., & Sparks, D. (1996). Exploring the Relationship between Staff Development and Improvements in Student Learning. Journal of Staff Development, 17, 34-38.

Harfitt, G. J., & Tavares, N. J. (2004). Obstacles as opportunities in the promotion of teachers' learning. International Journal of Educational Research, 41(4-5), 353-366. Retrieved from https://doi.org/10.1016/j.ijer.2005.08.006

Katamei, M., & Omwono, G. (2015). Intervention Strategies to Improve Students'Academic Performance in Public Secondary Jacob Schools in Arid and Semi-Arid Lands in Kenya. International Journal of Social Science Studies. 3(4); ISSN 2324-8033 E-ISSN 2324-8041. Retrieved from http://dx.doi.org/10.11114/ijsss.v3i4.796

Lawless, K. A., & Pellegrino, J. W. (2007). Professional Development in Integrating Technology Into Teaching and Learning: Knowns, Unknowns, and Ways to Pursue Better Questions and Answers. Review of Educational Research, 77(4), 575–614. Retrieved from https://doi.org/10.3102/0034654307309921

Lunt M. (2014). Selecting an appropriate caliper can be essential for achieving good balance with propensity score matching. American journal of epidemiology, 179(2), 226–235. doi:10.1093/aje/kwt212

Mortimore, P., & Sammons, P. (1987). New evidence on effective elementary schools. Educational Leadership, 45: 4–8.

National Commission on Teaching and America's Future. (2016). What Happens Now: A New Compact for Teaching and Learning- The Evidence Base. Retrieved from https://files.eric.ed.gov/fulltext/ED572506.pdf

Pitner, R., Yu, M., & Brown, E. (2013). Which factor has more impact? An examination of the effects of income level, perceived neighborhood disorder, and crime on community care and vigilance among low-income African American residents. Race and Social Problems, 5, 57-64.

Rosenthal, R. (1991). Meta-analytic procedures for social research. Newbury Park, CA: Sage.

Rosenbaum, P. R., & Rubin, D. B. (1985). Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. The American Statistician, 39(1): 33-38.

Stephens, A. C., & Hartmann, C. E. (2004). A Successful Professional Development Project's Failure to Promote Online Discussion about Teaching Mathematics with Technology. Journal of Technology and Teacher Education. 12(1), 57. Retrieved from https://www.questia.com/library/journal/1G1-116578436/a-successful-professional-development-project-s-failure

Stuart E. A. (2010). Matching methods for causal inference: A review and a look forward. Statistical science: a review journal of the Institute of Mathematical Statistics, 25(1), 1–21. doi:10.1214/09-STS313

U.S. Department of Education. (n.d.). Transforming Teaching

and Leading. Retrieved from https://www.ed.gov/teaching?src=rn

U.S. Department of Education. (2016). Non-regulatory guidance for Title II, Part A: Building systems of support for excellent teaching and leading. Retrieved from https://www2.ed.gov/policy/elsec/leg/essa/essatitleiipartaguidance.pdf

Walker, A., Recker, M., Ye, L., Robertshaw, B., Sellers, L., & Leary, H. (2012). Comparing Technology-related Teacher Professional Development Designs: A Multilevel Study of Teacher and Student Impacts. The Instructional Architect Research Group. Paper 6. Retrieved from https://digitalcommons.usu.edu/iagroup/6

What Work's Clearinghouse. (n.d.). Procedures and Standards Handbook (Version 3.0). Retrieved from, http://ies.ed.gov/ncee/wwc/pdf/reference_resources/wwc_procedures_v3_0_draft_standards_handbook.pdf

Woulfin, S. L. (2015). Highway to reform: The coupling of district reading policy and instructional practice. Journal of Educational Change, 16(4), 535–557.

Yoon, K. S., Duncan, T., Lee, S. W.-Y., Scarloss, B., & Shapley, K. (2007). Reviewing the evidence on how teacher professional development affects student achievement (Issues & Answers Report, REL 2007–No. 033). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Southwest. Retrieved from http://ies.ed.gov/ncee/edlabs

Young, M. D., Winn, K. M., & Reedy, M. A. (2017). The Every Student Succeeds Act: Strengthening the Focus on Educational Leadership. Educational Administration Quarterly, 53(5), 705–726. Retrieved from https://doi.org/10.1177/0013161X17735871

Appendix A

Table 1: Background Characteristics of lead4ward Educator Survey Respondents, 2018–2019 (N = 340)		
	N	**%**
How many years have you been an educator, including the 2018-2019 academic year?		
Less than 1	7	2.1
1 to 3	26	7.6
4 to 6	55	16.2
7 to 10	36	10.6
11 or more	216	63.5
Are you a classroom teacher?		
Yes, I am a full-time classroom teacher.	256	75.3
Yes, I am a part-time classroom teacher.	1	.3
No, I am a school administrator (principal, assistant principal, dean, etc.)	59	17.4
No, I am a district administrator (teacher development specialist, curriculum staff, data driven instructional specialist, etc.)	23	6.8
Not answered	1	.3
If you are a classroom teacher, school administrator, or district administrator, which grade levels do you currently teach or support? Check all that apply.		
Kindergarten	55	16.2
First through Fifth	166	48.8
Sixth through Eighth	66	19.4
Ninth through Twelfth	104	30.6
None of the Above	15	4.4
Which content areas do you currently teach or support? Check all that apply.		
English/Language Arts/Reading	177	52.1
Math	169	49.7
Science	150	44.1
Social Studies/History	130	38.2
Ancillary (CTE, Health and Physical Education, Fine Arts, Language Other Than English. etc.)	48	14.1
How many hours of training did you receive on lead4ward during the 2018-2019 academic year?		
0	37	10.9
1 to 4 hours	140	41.2
5 to 8 hours	71	20.9
9 or more hours	92	27.1
lead4ward training that you received occurred at which level or in which setting?		
Campus level	174	51.2
District level	193	56.8
Online Webinars	38	11.2
What is the highest level of education that you completed?		
High School Diploma		
Bachelor's Degree	152	44.7
Master's Degree	169	49.7
Doctorate Degree	19	5.6

Appendix B

Table 3a: HISD lead4ward Professional Development, 2018–2019
Implementation and Support
TE_lead4ward Instructional Planning
LD_lead4ward CSI Special Education
LD_lead4ward Deep Dive
LD_lead4ward Next Level Supporting Implementation Campus
LD_lead4ward Intermediate Overview: (Lead4ward) (All School Levels)
TE_lead4Ward Instructional Planning K-5
TE_lead4ward: Instructional Support/Planning with Data
LD_lead4ward Principal as Process Champion

Table 3b: HISD lead4ward Professional Development, 2018–2019
Reading/English Language Arts
LD_lead4ward Comprehension Strategies in Action Reading (K-5)
LD_lead4ward Wave Goodbye, Say Hello New ELAR TEKS (6-EOC)
LD_lead4ward Reading Rockin' Review Volume II
LD_lead4ward Wave Goodbye, Say Hello New ELAR TEKS (K-5)
LD_lead4ward Comprehension Strategies in Action Reading (6-EOC)

Table 3c: HISD lead4ward Professional Development, 2018–2019
Math
LD_lead4ward Problem Solving in the Math Classroom (K-12)
LD_lead4ward Number Sense Elem 1-5
LD_lead4ward Parts that Make Up the Whole of Fractions (Grades 2-5)
LD_ lead4ward Math Rockin' Review Volume II

Table 3d: HISD lead4ward Professional Development, 2018–2019
Science
LD_ lead4ward Science Rockin' Review Volume II
LD_lead4ward Physical Science: Force, Motion, and Energy (6-8)
LD_lead4ward The Power of Process in Science (6-EOC)
LD_lead4ward The Power of Process in Science (K-5)

Table 3e: HISD lead4ward Professional Development, 2018–2019
Social Studies\U.S. History
LD_lead4ward Social Studies Rockin' Review: Volume II
LD_lead4ward Power of Process Social Studies (4-EOC)
LD_lead4ward Early Republic Social Studies (8th Grade)
LD_ lead4ward Contemporary America (HS US History)

Table 3f: HISD lead4ward Professional Development, 2018–2019
Writing
LD_ lead4ward Souped-Up Super 8 Writing (6-EOC)
LD_ lead4ward Writing Rockin' Review
LD_lead4ward Souped-Up Super 8 Writing (3-5)

TE = Technology
LD = Leadership Development

Appendix C

Table 4: How Often lead4ward Resources Were Used During the 2018–2019 Academic Year

	Everyday		At least once a week		1 to 3 times a month		Never Used	
	n	%	n	%	n	%	n	%
Field Guides	19	6.0	93	29.2	148	46.4	59	18.5
Instructional Tools (e.g., TEKS Side-by-Sides, Planning Guides, Instructional Strategies, Quickchecks, Think It Up!, Thinking Stems)	44	13.6	102	31.5	136	42.0	42	13.0
Performance Standards (e.g., Raw Score Conversion Tables, STAAR Progress Measures)	17	5.2	86	26.3	155	47.4	69	21.1
Data Tools (e.g., Frequency Distribution Charts)	18	5.5	98	30.2	127	39.1	82	25.2
Academic Vocabulary Resources	36	11.1	82	25.4	128	39.6	77	23.8
lead4ward Reports in OnTrack	16	4.9	74	22.8	132	40.7	102	31.5
lead4ward App	25	7.8	86	26.7	131	40.7	80	24.8
Learning Videos	5	1.6	30	9.6	100	32.1	177	56.7
State Accountability Quicklooks	6	1.9	41	13.1	122	38.9	145	46.2
Test Accessibility and Special Education Resources	8	2.5	38	11.9	116	36.5	156	49.1

Table 5: Usefulness Ratings of lead4ward Resources for Planning Meaningful Instruction, 2018–2019

	Extremely useful		Fairly useful		Not sure		Not very useful		Not at all useful		Never used this resource		U-Score
	n	%	n	%	n	%	n	%	n	%	n	%	Rating
Field Guides	161	50.3	86	26.9	19	5.9	5	1.6	49	15.3	0	0.0	.95
Instructional Tools (e.g., TEKS Side-by-Sides, Planning Guides, Instructional Strategies, Quick-checks, Think It Up!, Thinking Stems)	198	60.2	91	27.7	9	2.7	6	1.8	1	.3	24	7.3	1.6
Performance Standards (e.g., Raw Score Conversion Tables, STAAR Progress Measures)	170	52.3	83	25.5	23	7.1	6	1.8	1	.3	42	12.9	1.5
Data Tools (e.g., Frequency Distribution Charts)	156	48.8	87	27.2	21	6.6	5	1.6	2	.6	49	15.3	1.4
Academic Vocabulary Resources	148	45.7	86	26.5	27	8.3	7	2.2	1	.3	55	17.0	1.4
lead4ward Reports in OnTrack	148	45.8	65	20.1	29	9.0	5	1.5	3	.9	73	22.6	1.4
lead4ward App	128	39.5	99	30.6	24	7.4	7	2.2	5	1.5	61	18.8	1.3
Learning Videos	73	23.1	59	18.7	45	14.2	7	2.2	4	1.3	128	40.5	1.0
State Accountability Quicklooks	87	27.2	79	24.7	43	13.4	5	1.6	3	.9	103	32.2	1.1
Test Accessibility and Special Education Resources	82	25.9	62	19.6	47	14.8	5	1.6	1	.3	120	37.9	1.1

Note: Values assigned to each category were as follows.
"Not at all useful" = -2 "Fairly useful" = 1
"Not very useful" = -1 "Extremely useful" = 2
"Not sure" = 0 "Never used this resource" were not used in calculations.

Appendix C (cont'd)

Table 6: How Helpful Field Guides Were Toward Connecting State Standards, Building Content Knowledge, and for Instructional Planning, 2018–2019

STAAR Field Guides' Sections	Extremely helpful		Fairly helpful		Not sure		Not very helpful		Not at all helpful	
	n	%	n	%	n	%	n	%	n	%
Connecting TEA standards to the curriculum.	198	60.6	80	24.5	37	11.3	1	.3	11	3.4
Providing relevant context that shows how each student expectation fits into the big picture.	170	52.6	84	26.0	54	16.7	6	1.9	9	2.8
Building content knowledge with explanations, stimulus identification, and essential vocabulary.	185	56.9	84	25.8	43	13.2	7	2.2	6	1.8
Making connections to instructional implications.	184	56.4	89	27.3	40	12.3	3	.9	10	3.1
Attaining insight into the type of mistakes students make.	173	53.7	86	26.7	49	15.2	5	1.6	9	2.8
Helping with instructional planning and PLCs.	182	56.3	86	26.6	36	11.1	5	1.5	14	4.3

Table 7. Effectiveness of lead4ward Resources Toward Improving Student Proficiency in Content Areas, Among Special Populations, and for Teacher Induction and Intervention, 2018–2019

	Hinders proficiency		Makes no difference		Improves proficiency a little		Improves proficiency a lot		Not sure	
	n	%	n	%	n	%	n	%	n	%
Math	12	4.1	10	3.4	31	10.6	118	40.4	121	41.4
Science	6	2.1	14	4.8	32	11.0	91	31.3	148	50.9
Reading/English Language Arts (ELA)	6	2.1	9	3.1	50	17.3	104	36.0	120	41.5
Social Studies	3	1.1	11	3.9	35	12.5	71	25.4	160	57.1
English Learners	7	2.5	15	5.3	55	19.6	112	39.9	92	32.7
Special Needs/Special Education	5	1.8	16	5.9	45	16.5	74	27.2	132	48.5
Teacher Induction	8	2.9	14	5.0	34	40.7	114	40.7	110	39.3
Gifted/Talented Students	7	2.5	15	5.3	39	13.7	99	34.9	124	43.7
Intervention/RTI	8	2.8	16	5.6	41	14.3	127	44.3	95	33.1

Note: Values assigned to each category were:
"Hinders proficiency" = -1
"Makes no difference" = 0
"Improves proficiency a little" = 1
"Improves proficiency a lot" = 2
"Not sure" was not used in effectiveness calculations

Appendix C (cont'd)

Table 8: HISD Educators' Overall Perceptions of lead4ward, 2018–2019

	Strongly Agree		Agree		Neutral		Disagree		Strongly Disagree	
	n	%	n	%	n	%	n	%	n	%
I feel that lead4ward material is user-friendly (easy to understand and apply) when planning individual lessons.	149	44.7	118	35.4	47	14.1	14	4.2	5	1.5
Information accessed through lead4ward has been used in grade level/developmental planning meetings/PLCs.	147	45.2	110	33.8	43	13.2	18	5.5	7	2.2
lead4ward is closely aligned to the Texas Essential Knowledge and Skills (TEKS).	193	59.4	90	27.7	38	11.7	3	.9	1	.3
The structure of lead4ward resources allows me to easily move through the website to access the Field Guides, ELAR Side-by-Sides, etc.	129	39.9	106	32.8	67	20.7	14	4.3	7	2.2
I often use lead4ward instructional resources to enhance classroom activities.	116	35.5	97	29.7	80	24.5	28	8.6	6	1.8
I found the information that I was looking for to strenghten my instructional skills on the lead4ward website or app.	124	37.9	106	32.4	78	23.9	13	4.0	6	1.8
lead4ward Field Guides help me to plan and implement lessons well.	118	36.3	109	33.5	81	24.9	9	2.8	8	2.5
lead4ward has helped me to prioritize, plan, and implement effective teaching strategies aligned to STAAR.	142	44.0	97	30.0	66	20.4	14	4.3	4	1.2
Overall, lead4ward is effective in facilitating instructional planning.	157	48.2	108	33.1	52	16.0	4	1.2	5	1.5

Appendix D

Table 9a: English STAAR 3–8 Reading Performance, Dependent Samples T-test, Spring 2018 vs. 2019

	3rd to 4th		4th to 5th		5th to 6th		6th to 7th		7th to 8th	
	Pre 2018	Post 2019	Pre 2018	Post 2019	Pre 2018	Post 2019	Pre 2018	Post 2019	Pre 2018	Post 2019
N	4701	4701	7264	7264	3942	3942	4081	4081	7021	7021
Mean Scale Score	1423.92	1514.11	1491.56	1559.20	1576.67	1569.24	1571.44	1650.95	1630.72	1668.66
Std	140.494	143.954	147.086	146.650	137.296	135.747	134.645	142.011	146.423	141.696
Mean Diff.	90.19		67.64		-7.43		79.51		37.94	
t	62.37		58.08		-5.15		52.22		35.48	
Sig.	.000		.000		.000		.000		.000	

*** p < .0001, ** p < .001, *p<.05
Data Source: 2018 and 2019 STAAR 3-8 test files, spring administration

Table 9b: English STAAR 3–8 Math Performance, Dependent Samples T-test, Spring 2018 vs. 2019

	3rd to 4th		4th to 5th		5th to 6th		6th to 7th		7th to 8th	
	Pre 2018	Post 2019	Pre 2018	Post 2019	Pre 2018	Post 2019	Pre 2018	Post 2019	Pre 2018	Post 2019
N	4790	4790	7328	7328	3968	3968	3936	3936	5315	5315
Mean Scale Score	1466.00	1575.85	1565.22	1642.38	1642.50	1650.89	1637.88	1672.09	1603.59	1668.64
Std	141.131	163.262	156.537	172.126	134.754	149.893	138.644	147.018	104.995	129.998
Mean Diff.	109.85		77.16		8.39		34.21		65.05	
t	70.80		62.88		5.13		24.25		49.20	
Sig.	.000		.000		.000		.000		.000	

*** p < .0001, ** p < .001, *p<.05
Data Source: 2018 and 2019 STAAR 3-8 test files, spring administration

Table 10a: Propensity Score Matching Sample Based on Reading and Math, Algebra I and English I End-of-Course Exam Results, First-time Testers, 2019										
	Before Matching					After Matching				
	Treatment		Control		Total	Treatment		Control		Total
	n	%	n	%	N	n	%	n	%	N
Algebra I	3,230	44.4	4,037	55.6	**7,267**	3,230	50.2	3,204	49.8	6,434
English I	4,045	40.4	5,971	59.6	**10,016**	4,045	50.2	4,020	49.8	8,065

Table 10b: Propensity Score Matching Results Based on Algebra I (Math) End-of-Course Exam, First-time Testers, 2019								
	Before Matching				After Matching			
	Treatment (n = 3,230)		Control (n = 4,037)		Treatment (n = 3,230)		Control (n = 3,204)	
Males	1587	49.1	2019	50.0	1587	49.1	1580	49.3
Females	1643	50.9	2018	50.0	1643	50.9	1624	50.7
Eco Disadv.	2741	85.1	3149	78.2	2741	85.1	2682	83.7
At Risk	2656	82.2	3132	77.6	2656	82.2	2624	81.9
LEP	818	25.5	988	24.6	818	25.5	781	24.4
Special Ed.	350	10.8	288	7.1	350	10.8	215	6.7
Gifted/Talented	191	5.9	332	8.2	191	5.9	180	5.6

Table 10c: Propensity Score Matching Results based on English I (English Language Arts) End-of-Course Exam, First-time Testers, 2019								
	Before Matching				After Matching			
	Treatment (n = 4,045)		Control (n = 5,971)		Treatment (n = 4,045)		Control (n = 4,020)	
Males	1993	49.3	2970	49.7	1993	49.3	1975	49.1
Females	2052	50.7	3001	50.3	2052	50.7	2045	50.9
Eco Disadv.	3303	81.8	4347	72.9	3303	81.8	3336	83.0
At Risk	2910	71.9	3732	62.5	2910	71.9	2897	72.1
LEP	893	22.2	1134	19.1	893	22.2	849	21.1
Special Ed.	370	9.1	321	5.4	370	9.1	300	7.5
Gifted/Talented	608	15.0	1435	24.0	608	15.0	596	14.8