TEXAS IMPACT 2020 Update: Evaluation Report For Teach For America

October 2020

Annie Wright, Ph.D.
Dylan Farmer
Yusuf Kara, Ph.D.
Yetunde Zannou, Ph.D.
Anne Ware, Ph.D.

SMU | Annette Caldwell Simmons School of Education & Human Development
Center on Research & Evaluation

EXECUTIVE SUMMARY

The Center on Research & Evaluation (CORE) at Southern Methodist University conducted two impact evaluations of Teach For America (TFA) corps members and alumni teachers in five Texas regions - Austin, Dallas-Fort Worth (DFW), Houston, the Rio Grande Valley (RGV), and San Antonio. These evaluations compare academic outcomes for students of TFA corps members and alumni compared to matched teachers with no TFA affiliation and with commensurate classroom experience. The evaluations replicate parts of existing TFA evaluations and expand both the scope and the rigor of existing evidence. This evaluation span ten grade levels (3rd grade to 12th grade) and nine content areas (Reading, Math, Science, Social Studies, Algebra 1, English 1, English 2, Biology and U.S. History). They both utilize the State of Texas Assessments of Academic Readiness (STAAR) standardized assessment.

The first study focused on six academic years: 2011-12 to 2016-17[1]. The second study focused on the next two years: 2017-18 and 2018-19.

Together, these two complementary evaluations provide a rigorous look at TFA impacts over multiple years and across a broad geographic area, the largest of all TFA regions. They have strong implications for public and private agencies seeking to continue or expand support for the TFA model and contains important feedback for TFA as they continue to support implementation of the model in a diverse array of schools, districts and communities.

Both studies provide a similar conclusion: Overall, students of TFA-affiliated teachers were *as likely as* or *more likely to* pass the STAAR assessment than students of non-TFA-affiliated teachers. TFA alumni are the most effective group of teachers compared to corps members, underscoring a need to retain TFA alumni as classroom teachers over time. These results vary across content areas, years, regions, student demographic subgroups, and grade levels. Table One describes the average relative advantage for students of TFA Corps Members and Alumni across content areas.

Table One. *Average difference in probability of passing STAAR by content area and teacher group*

Summary of Higher or Lower Probability of Passing STAAR Subjects for TFA Corps Members & Alumni				
	Novice Teachers (i.e., TFA Corps)		Veteran Teachers (i.e., TFA alum)	
	Study 1	**Study 2**	**Study 1**	**Study 2**
HSAlgebra 1	-0.53%	-3.85%	8.80%	2.50%
HS Biology	1.98%	4.73%	11.90%	2.68%
HS English 1	0.37%	1.62%	3.30%	3.26%
HS English 2	1.68%	2.55%	10.20%	7.54%
HS U.S. History	9.53%	-3.53%	5.20%	2.18%
3rd - 8th Math	-0.93%	-2.70%	7.70%	-3.45%
3rd - 8th Reading	-3.73%	-1.15%	5.50%	1.86%
5th & 8th Science	1.93%	-4.51%	8.20%	-2.07%
8th Social Studies	0.24%	6.58%	4.50%	3.06%
All Subjects	0.08%	-0.74%	7.20%	0.89%

The first study found that, on average, students of TFA-affiliated teachers are 3% more likely to meet state standards compared to students of comparable non-TFA-affiliated teachers. The second study, focused on the 2017-18 and 2018-19 school years, found that, on average, students of TFA-affiliated teachers were equally likely to meet state standards compared to students of non-TFA teachers. Specifically, the first study found that students of TFA alumni are 7.2% more likely, on average, to meet state standards than students of comparable veteran non-TFA-affiliated teachers. This is followed by brand new corps members in their first year with a very slight advantage (0.34% advantage). Students of TFA corps members are *as likely* to meet state standards as a

[1] https://www.smu.edu/-/media/Site/Simmons/Research/CORE/PDFs/TFA_TXImpact_Final-Report_February2019.pdf?la=en

student of a novice non-TFA-affiliated teacher (0.08% difference). The second study found that students of TFA alumni are 0.9% more likely, on average, to meet state standards than students of comparable veteran non-TFA-affiliated teachers. This advantage was overall lower than what was observed in the first evaluation. Similarly, the second study found no relative advantage of TFA or non-TFA conditions for all corps members (-0.74% disadvantage) or brand new corps members in their first year (-0.42% disadvantage).

When examining content areas across both studies, **students of TFA-affiliated teachers receive the overall greatest benefit in the high school tested subjects.**

There is variability within these overall trends across student demographics and school type. Relative benefits for having a TFA corps member or alumni teacher were seen for key groups of Texas students who best represent TFA's overall mission of education equity for all. On average, **TFA-affiliated teachers are more effective than non-TFA-affiliated teachers for key populations.**

- The first study found the advantage of having a TFA-affiliated teacher is the same for economically disadvantaged and non-economically disadvantaged students. The second study confirmed this finding for economically disadvantaged students, particularly students of TFA alumni.
- The first study found that there is an advantage of TFA for students of all races; the effect is strongest specifically for Black and Hispanic students who have a TFA alumni as their teacher. The second study confirmed that there is an advantage for Black and Hispanic students who have TFA alumni teachers.
- The first study found an advantage of TFA in both traditional public ISDs and charters and that the effect is strongest in ISDs. The second study confirmed similar small advantages of TFA in both ISDs and charter schools.
- The first study found an advantage of TFA for both Limited English Proficient (LEP) and non-LEP students; the effect was strongest for LEP students. The second study confirmed the TFA advantage for LEP students but found no relative advantage of TFA for non-LEP students.

Evaluation Methodology

These two studies were designed to allow for rigorous comparison of TFA and non-TFA conditions as well as to explore more deeply the various conditions that contribute to relative effectiveness. Similar to previous studies, these evaluations consider the role of factors related to the district, school, teacher and student in explaining variability. Unlike other studies, the data set used for analysis includes matched student and teacher-level data that allows for a more rigorous comparison group match and richer analysis of factors that contribute to differences in TFA effects on student outcomes.

The central evaluation question for both studies was whether there was a differential impact on academic outcomes attributable to having a TFA corps member or alumna as the classroom teacher of record for one academic year. CORE carefully described the variables that contribute to this impact such that deeper knowledge about the mechanisms of TFA impacts in different regions, and for different content areas, grade levels, school type and student demographics is illuminated.

Two aims for the methodology were paramount: (1) ensuring comparison group equivalence such that the effects of TFA could be appropriately isolated and (2) succinctly and accurately communicating findings from a large number of underlying analyses. A rigorous propensity score weighting process helped support confidence in claims of impact by controlling for covarying conditions including characteristics of students, teachers, schools and districts. Logistic regressions then assigned likelihood of passing STAAR to the TFA or non-TFA conditions. In tandem, these analytic strategies allowed for isolation of the effects of having a TFA corps member or alumna in the classroom compared to a non-TFA-affiliated teacher. Due to the large number of underlying analyses associated with this expansive dataset, CORE adopted a meta-analysis strategy and averaged the likelihood of passing STAAR in a given content area with a TFA-affiliated teacher (either corps member or alumni) or a non-TFA-affiliated teacher.

Combined, the two studies synthesize findings across 1,193 unique underlying analyses. The first study synthesized 699 underlying analyses across six academic years, and the second study synthesized 494 analyses across the next two academic years.

Table of Contents

Introduction

About CORE

A nationally ranked private university with seven degree-granting schools, Southern Methodist University (SMU) is a distinguished center for teaching and research. Housed within SMU's Simmons School of Education and Human Development, the Center on Research and Evaluation (CORE) provides a range of research, evaluation, and consultation services. CORE's overall aim is to use evaluation science to improve educational outcomes for youth. Many of CORE's projects center on early childhood, out-of-school learning opportunities, and educator preparation and professional development initiatives. CORE's work emphasizes the community contexts in which these educational initiatives are implemented as a key consideration for understanding effectiveness. CORE staff are interdisciplinary, representing educational research and evaluation, quantitative methodology, psychology and social work and many of CORE's evaluators have classroom teaching experience.

About Teach For America

Teach For America (TFA) is a national organization that recruits leaders early in their careers to teach for two years in one of 51 urban and rural regions across the U.S. According to the TFA Foundations document,[2] their mission is to find, develop, and support a diverse network of leaders who expand opportunity for children from classrooms, schools, and every sector and field that shapes the broader systems in which schools operate. In partnership with schools, families, local universities, other organizations, and businesses in the community, TFA provides initial training, ongoing professional development, and access to a resource and support network for corps members and alumni. In 2020, 632 corps members worked in Dallas, Houston, San Antonio, and the Rio Grande Valley compared to 638 California, where TFA's corps presence is currently the largest of any state.

About the Evaluations

The impact analyses in these evaluations replicate and expand on existing evidence of TFA impact by allowing for rigorous comparison of TFA and non-TFA conditions. A growing body of empirical evidence suggests that students of TFA corps member or alumni teachers outperform or perform as well as students of non-TFA-affiliated teachers based on subject matter, grade level, and student and teacher characteristics. Evaluation and research studies of the impact of TFA on student outcomes in Texas support these findings, but have limitations related to sample size, content area, student characteristics, and outdated state assessments to measure student achievement (Houston Independent School District [HISD], n.d., 2018; Mickelson & McEnturff, 2015; Raymond, Fletcher, & Luque, 2001; Turner, Goodman, Adachi, Brite, & Decker, 2012; Ware et al., 2011).

This evaluation addresses these limitations; it replicates existing findings and expands both the scope and the rigor of existing evidence. Unlike other studies, the data set used for analysis includes de-identified, matched student and teacher-level data that allows for a more rigorous comparison group match and richer analysis of factors that contribute to differences in student outcomes.

In addition to establishing robust understanding of impact at scale, this evaluation explores more deeply the various conditions that contribute to effectiveness. This evaluation includes a thorough description of variables that contribute to impacts such that deeper knowledge about the mechanisms of TFA impacts in different regions, and for different content areas, grade levels, school type, and student body composition is illuminated.

[2] R. Carreon, personal communication, January 7, 2018

Background

Prior Studies

General findings. A review of the literature points to strong evidence of TFA impact, where students of TFA-affiliated teachers outperform students of non-TFA-affiliated teachers in relatively well-designed studies. Some studies employed a rigorous random assignment design (Clark, Isenberg, Liu, Makowsky, & Zukiewicz, 2017; Clark et al., 2013; Glazerman, Mayer, & Decker, 2006), but most others employed a quasi-experimental design that created statistically matched groups of students or used statistical analyses to control student and/or teacher characteristics (Hansen, Backes, Brady, & Xu, 2015; Henry et al., 2015; Turner et al., 2012; Ware et al., 2011). Overall, the results of these prior studies suggest that TFA Mathematics teachers are more effective than their non-TFA counterparts, but that TFA and non-TFA-affiliated Reading/English Language Arts teachers perform about the same. In a rigorous review of seven studies investigating the effects of TFA-affiliated teachers on student outcomes, the What Works Clearinghouse (WWC) concluded that TFA-affiliated teachers have "positive effects on Mathematics achievement, potentially positive effects on Science achievement, and no discernible effects on Social Studies achievement and English Language Arts achievement" (United States Department of Education [USDOE], 2016, p. 1). Additional studies not included in the WWC review confirm the conclusions with at least five reporting a positive effect of TFA-affiliated teachers on student Math achievement (Decker, Mayer, & Glazerman, 2004; Hansen et al., 2015; Mickelson & McEnturff, 2015), but only Mickelson and McEnturff (2015) identified a positive effect for TFA-affiliated teachers in Reading/Language Arts.

TFA in Texas. Of particular importance for the current evaluation are the studies that have examined the effects of TFA-affiliated teachers in Texas and the gaps in evidence and understanding that have not yet been addressed. All of the Texas-based studies were quasi-experimental, as is the current studies described here; four were district level evaluations (Raymond et al., 2001; Mickelson & McEnturff, 2015; HISD, n.d., 2018) and two focused on multiple school districts and/or regions in Texas (Turner et al., 2012; Ware, et al., 2011). In general, these studies supported other findings that TFA-affiliated Math teachers outperformed their non-TFA-affiliated colleagues (at varied grade levels and teacher experience levels), and also provided some evidence of a positive impact for Reading/English Language Arts teachers. These Texas studies, however, are somewhat limited in their scope, as all studies necessarily are. These limitations highlight gaps that CORE sought to address in the current evaluations.

The Current Studies

The current studies expand on existing literature. Most significantly, previous studies estimated the effects of TFA by aggregating student achievement outcomes and attributing them to the school where TFA-affiliated teachers worked, in lieu of linking specific students with teachers. The current studies overcome that challenge by using linked student and teacher-level data. Additionally, the current studies are, together, the first comprehensive studies of TFA to use STAAR[3]. These studies address key limitations in prior studies of the TFA impact on student achievement in Texas by: (1) using STAAR exams as an indicator of student achievement, (2) considering school and district characteristics in estimates of the effect of TFA corps members and alumni teachers on student outcomes, and (3) linking de-identifiable student-level data with de-identified teachers, by grade level and course.

[3] The STAAR assessment represents increased rigor compared to TAKS which was phased out and replaced by STAAR from 2012 to 2014. The content and skills assessed by STAAR require a higher level of complexity and more authentic application of content and skills (e.g., less multiple choice items). Also, the TAKS assesses general knowledge that would accumulate across multiple school years, while STAAR goes into more depth on grade-specific content and skills. Results from analyses indicate that STAAR is more difficult than TAKS and that students are likely to answer fewer questions correctly on STAAR than on TAKS. Additionally, the STAAR assessment has more items and a shorter time limit. https://tea.texas.gov/student.assessment/staar

Methods

The geographic focus of the two evaluations is five identified Texas regions—Austin, Dallas-Fort Worth (DFW), Houston, Rio Grande Valley (RGV), and San Antonio, covering 8 counties—and the regional Independent School Districts (ISDs) and charter school systems in the counties in these regions. The grade levels and subjects tested in STAAR form the key outcomes of interest. These are: Reading and Math (grades 3-8), Science (grades 5 and 8), Social Studies (grade 8), and high school end of course exams for Algebra I, English I and II, Biology and U.S. History. STAAR Writing was excluded from these analyses due to the nature of that specific exam. The first study focused on six academic years: 2011-12 to 2016-17. The second study focused on the next two years: 2017-18 and 2018-19.

Data Management

Data Selection and Request

CORE received all data for the planned analyses from the TEA via a Public Information Requests (PIRs) as allowed by The Texas Public Information Act. The analyses for these studies did not require CORE to receive identifiable student level data; thus, student-level data was de-identified. Specifically, CORE requested STAAR test results for all students enrolled in all school districts across eight Texas counties in the aforementioned regions during the 2011-12 through 2018-19 school years. The counties included Travis, Dallas, Tarrant, Harris, Cameron, Hidalgo, Starr, and Bexar. Individual district indicators were not requested. In addition to de-identified student-level STAAR results, each student's test record also included information about the teacher of record for that course in that school year and the student's school. Each teacher of record was flagged as either a TFA corps member in the school year of record, a TFA alumni, or a non-TFA-affiliated teacher. All data were organized at the student level, with assessment, teacher and school level indicators matched at the student level.

Data Cleaning/Preparing for Analysis

CORE received two types of datasets from TEA. The first dataset was demographic data, including students, teachers, and campus-level indicators. The second data set contained student STAAR achievement data. Demographic datasets were cleaned and merged to form a single student-level final dataset for the statistical analyses. Thus, all analyses were conducted by using the student-level observations, and all campus and district level data such as school demographics and campus accountability rating were disaggregated to the student level. Finally, students' STAAR test data were merged to the final demographic dataset by using the encrypted student ID numbers.

The study sample includes:

- Teachers that have a unique TFA/non-TFA affiliation indicator within an academic year
- TFA corps members that had maximum two years of experience
- Only TFA and non-TFA-affiliated teachers of records (excludes teaching assistants and aides)
- Districts that have at least one TFA-affiliated teacher in any of the six academic years
- Students that did not change their schools within an academic year
- Students taught by a single teacher of a specific subject area (excludes co-teaching cases)
- Students that only took the standard version of the STAAR test
- For multiple-year analyses, students that did not repeat the same grade

Sample for Analysis

After preparing data for analyses and omitting student records as needed, a total of 7,298,318 unique students' data were used in the analyses in the first study, and a total of 1,786,112 unique students' data were used in the second study. These observations are distributed across ten grade levels, eight Texas counties (representing the five regions under study), and eight school years. CORE's analyses focused on single-year observations of student performance, not longitudinal analyses. Therefore, an individual student's performance within a school year is considered a single observation in the sample. That same student's performance the following school

year represents a separate observation. As described in detail in the Analysis Procedure section, evaluation questions were answered by conducting hundreds of individual analyses for sub-samples of the entire sample.

For the first study focused on the 2011-12 through 2016-17 school years, the final sample included 2,477 active TFA corps members (during any year used in the analyses), 1,479 TFA alumni, and 109,649 non-TFA-affiliated comparison teachers. For the second study focused on the 2017-18 and 2018-19 school years, the final sample included 738 active TFA corps members (during any year used in the analyses), 1,130 TFA alumni, and 89,216 non-TFA-affiliated comparison teachers. Tables describing student and teacher demographics across all years are provided in Appendix A.

Analysis Procedure

In order to determine TFA's impact on student achievement as well as to identify and succinctly communicate contributing conditions, CORE conducted several types of analyses:

- logistic regressions in order to assign the likelihood of passing STAAR to either a TFA or non-TFA condition
- propensity score weighting in order to control for covariate effects by matching two groups of students
- meta-analysis strategy to assign overall likelihood of passing into categories; identifying the percent of all analyses that showed a statistically significant likelihood in favor of TFA
- meta-analysis strategy using analysis of variance (ANOVA) to determine what broad factors (e.g., region, content area, or grade level) were associated with differentiated impact of TFA
- descriptive and comparative analyses of likelihoods in order to synthesize trends

All conditions (i.e., content area, geographic area, school year, and grade level) were considered as varying conditions. However, not all possible combinations of the various factors were used for final analyses due to incidences of low sample size. Using the various combinations of factors, the first study includes a total of 699 unique comparison analyses, and the second study includes a total of 494 unique comparison analyses (with the exclusion of analyses with low-sample sizes and perfect probability issues).

Logistic Regression

These analyses were designed to determine the likelihood that a student would meet grade level standards if they have a TFA corps member or alumni. The outcome measure for academic achievement was the dichotomized pass/fail indicator for a specific STAAR tested subject. CORE employed the logistic regression (LR) modeling technique to predict a student's pass/fail condition by using the dichotomous TFA indicator. The model predicted a student's passing probability for a specific subject, depending on whether they taught by a TFA-affiliated teacher or not.

Propensity Score Weighting

In order to control the potential covariate effects, such as being economically disadvantaged, CORE adopted the propensity score weighting (PSW) approach. Different from the propensity score matching (PSM) procedure, PSW uses the predicted propensity scores to calculate sampling weights for each of the data observations (Leite, 2017). The PSW method then derives the so-called "weight" scores for all comparison group students depending on their covariate measures and the treatment indicator—in this study, being taught or not being taught by a TFA-affiliated teacher. Different treatment effect evaluations require different weight score calculations by PSW. CORE preferred the "*average treatment effect on the treated (ATT)*" (Harder, Stuart, & Anthony, 2010), where all students taught by a TFA-affiliated teacher have a weight of one. All students taught by a non-TFA-affiliated teacher have weights lower than one depending on their measures of the multiple covariates used to estimate the propensity scores. Higher weight scores for those students indicate a better degree of matching with the students taught by a TFA-affiliated teacher in terms of the covariates. Using the ATT warrants inclusion of the entire group of students taught by a TFA-affiliated teacher (since their weights are all one) whose sample size are considerably lower than the students taught by a non-TFA-affiliated teacher.

CORE identified different possible analysis conditions by using multiple factors such as, but not limited to, region, academic year, subject of interest, and grade level. All condition-specific analyses were conducted separately for three treatment conditions: (1) brand-new TFA corps members in their first year, (2) all TFA corps members in

their first and second year of teaching, and (3) TFA alumni with a minimum two years of experience. Teacher experience was used as a balancing covariate during the PSW procedure for the alumni analyses only. CORE and TFA collaboratively identified a set of potentially important covariates to include in the PSW procedure. The contributing conditions or factors that were used for comparison group matching were student level (gender, race, economic disadvantage, LEP, special education status, bilingual status, ESL status, being "at risk" according to the TEA[4]), teacher level (years of experience), school level (percentage of the economically disadvantaged students in school, campus accountability rating, charter status) and district level (district accountability rating).

PSW was performed for each analysis condition prior to the main LR model that predicted the TFA impact. Thus, the covariates included during each of the PSW analyses differed depending on the selected conditions, such as region, grade level, academic year. This approach is thought to improve the covariate balance for each of the selected sub-samples, before the main LR model analyses. It is also important to note that CORE conducted some disaggregation analyses using student demographic characteristics. For these specific analyses, we excluded the demographic variable of interest from the PSW procedure. Sample sizes for each analysis were reported in terms of general, students taught by a TFA and non-TFA-affiliated teacher.

After fitting the main LR model along with the estimated weight scores, CORE obtained the degree of the difference between the groups of students taught by a TFA and non-TFA-affiliated teacher in terms of the predicted passing probabilities. The statistical significance of this difference, as well as the magnitude of the TFA effect in terms of odds ratio, were reported for each of the analyses. It is also important to note that some analysis results for specific conditions were not reported due to non-existing sub-samples and perfect probability levels (0/1) for either groups of students taught by a TFA or non-TFA-affiliated teacher. During the second study focusing on the 2017-18 and 2018-19 school years, CORE identified that many of those problematic analyses were associated with extremely low ATT weights for the non-TFA taught students. Following a method recommended by Leite (2017), the observations with extremely low values were gradually discarded from the final LR analyses until the perfect probabilities were changed. Moreover, CORE also excluded the analyses results where the sample sizes of students taught by a TFA-affiliated teacher were lower than 100. After these exclusions, a total of 494 valid analyses results (72.5% of the attempted) were reported for the second study.

[4] TEA At Risk Indicator Code indicates whether a student is currently identified as at-risk of dropping out of school using state-defined criteria only (TEC §29.081, Compensatory and Accelerated Instruction).

Results

Impact of TFA

Key Takeaways

- Both studies provide a similar conclusion: Overall, students of TFA-affiliated teachers were as likely as or more likely to pass the STAAR assessment than students of non-TFA-affiliated teachers.
- The first study found that, on average, students of TFA-affiliated teachers are 3% more likely to meet state standards compared to students of comparable non-TFA-affiliated teachers. The second study, focused on the 2017-18 and 2018-19 school years, found that, on average, students of TFA-affiliated teachers were equally likely to meet state standards compared to students of non-TFA teachers.
- Students of TFA-affiliated teachers receive the overall greatest benefit in the high school tested subjects.
- TFA-affiliated teachers are more effective than non-TFA-affiliated teachers for key populations:
    - The first study found the advantage of having a TFA-affiliated teacher is the same for economically disadvantaged and non-economically disadvantaged students. The second study confirmed this finding for economically disadvantaged students, particularly students of TFA alumni.
    - The first study found that there is an advantage of TFA for students of all races; the effect is strongest specifically for Black and Hispanic students who have a TFA alumni as their teacher. The second study confirmed that there is an advantage for Black and Hispanic students who have TFA alumni teachers.
    - The first study found an advantage of TFA in both traditional public ISDs and charters and that the effect is strongest in ISDs. The second study confirmed similar small advantages of TFA in both ISDs and charter schools.
    - The first study found an advantage of TFA for both Limited English Proficient (LEP) and non-LEP students; the effect was strongest for LEP students. The second study confirmed the TFA advantage for LEP students but found no relative advantage of TFA for non-LEP students.

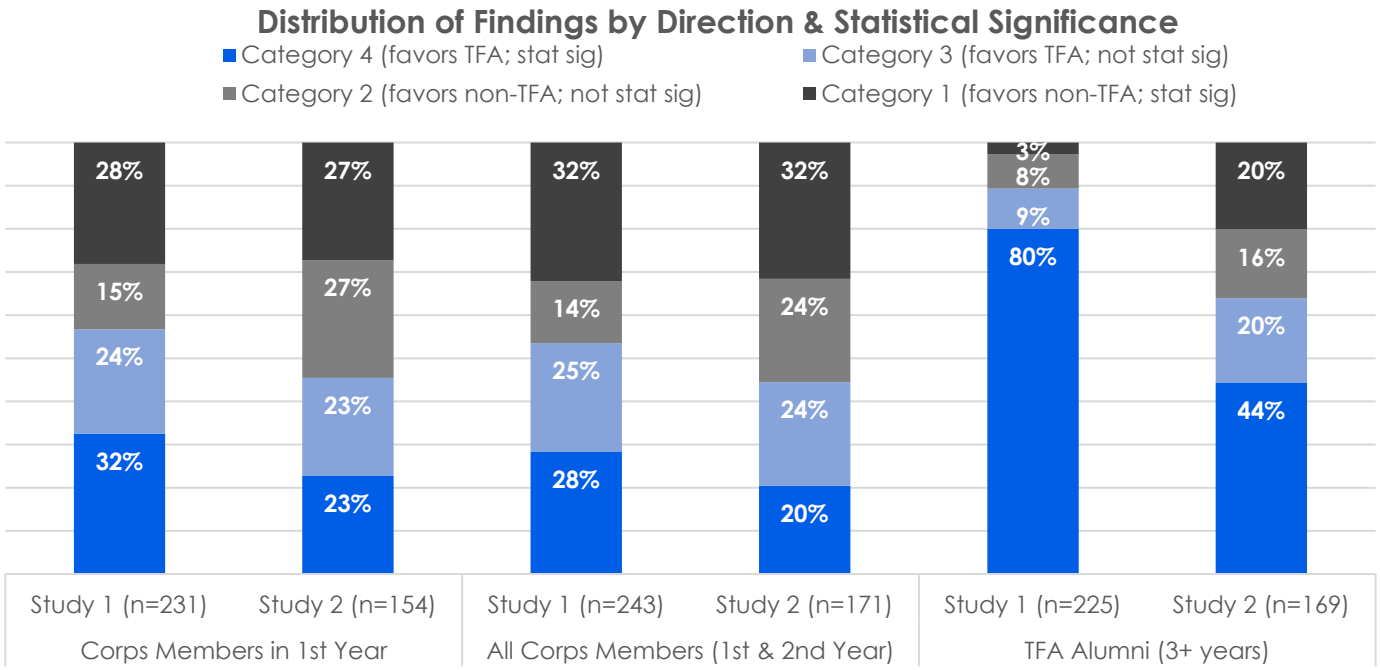
Categorizing All Findings

The result of each of the 1,193 unique analyses across the two studies categorized into one of four groups indicating both (1) which group (TFA or comparison) had a higher probability of passing, and (2) the statistical significance of the finding (see Table Two for legend). Figure One presents an aggregate summary of all analyses categorized by "direction" and significance of findings by study. These findings are summarized for specific content areas and regions in the figures on the next pages.

Table Two. *Categories of direction and significance of TFA/non-TFA analyses*

Finding category	Students of TFA-affiliated teachers are more likely to pass STAAR	Result is statistically significant
4	yes	yes
3	yes	no
2	no	no
1	no	yes
Did not converge	n/a	n/a

Figure One. *Aggregate synthesis of all analyses by direction & significance findings categories*

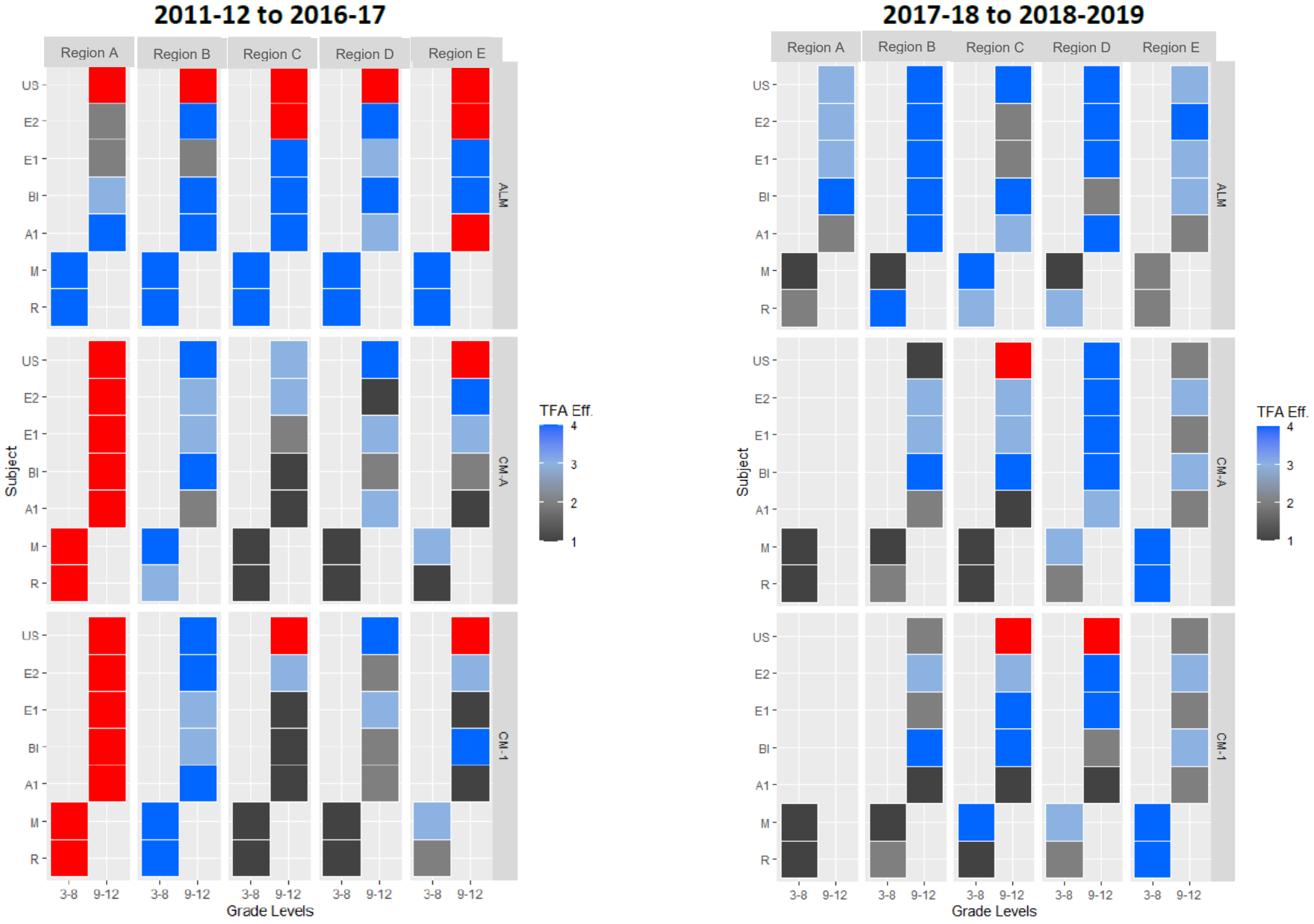


Note: The N for each group represents the number of analyses conducted for that group; the bar represents how the findings from that group of analyses are distributed within the four categories of possible findings.

Figures Two through Five on the next pages provide a detailed summary of findings across the two studies.

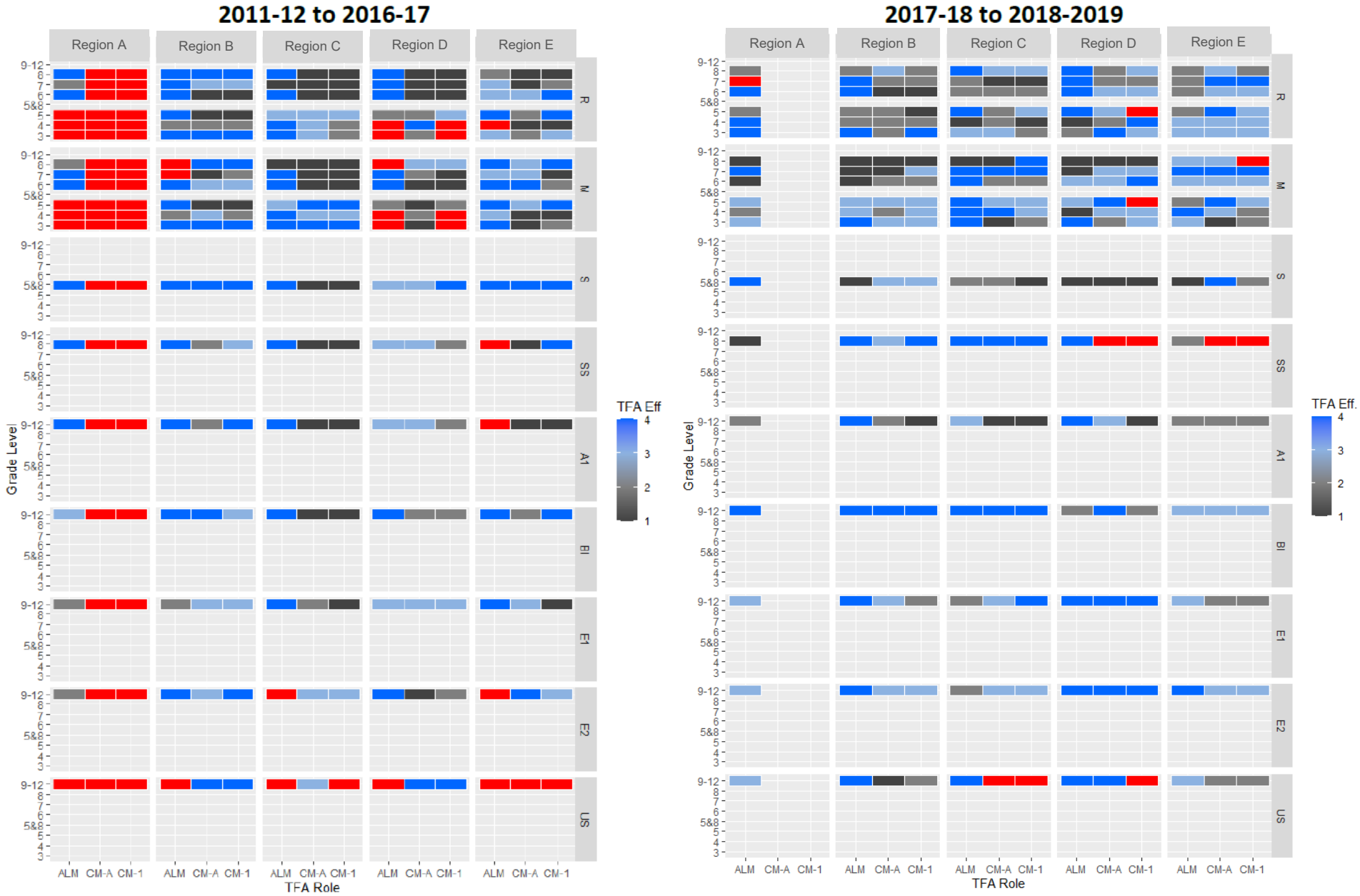
- **Figure Two** provides a heat map of the disaggregated comparison analyses by subject, geographic region, and TFA affiliation (alumni, corps member, etc.). Using the same "category-of-finding" logic described previously, this heat map provides a comprehensive snapshot of which TFA-affiliated teachers are most effective for specific content areas within five distinct Texas regions.
- **Figure Three** provides a heat map similar to Figure Two, but describes a more granular illustration of the findings organized by grade level. Using the same "category-of-finding" logic described previously, this heat map provides a comprehensive snapshot of which TFA-affiliated teachers are most effective for specific content areas and grade levels within five distinct Texas regions.

Figure Two. *Comparison analyses heat map disaggregated by TFA affiliation, geographic region, and content area*[5]



[5] Heatmap legend: CM=Corps Member, US=US History, R=Reading, M=Math, E2=English 2, E1=English 1, BI=Biology, A1=Algebra 1. ALM: Alumni, CM-A:All Corps Members, CM-1: New (1st year) Corps Members. Red cells represent non-converged analyses.

Figure Three. *Comparison analyses heat map disaggregated by TFA affiliation, geographic region, grade level and content area*[6]



[6] Heatmap legend: CM=Corps Member, US=US History, R=Reading, M=Math, E2=English 2, E1=English 1, BI=Biology, A1=Algebra 1. ALM: Alumni, CM-A:All Corps Members, CM-1: New (1st year) Corps Members. 5&8: 5th and 8th grades combined, 9-12: 9th through 12th grades combined. Red cells represent non-converged analyses.

Student and School Subgroups

Table Three describes the differentiated effect of TFA corps member and alumni teachers on specific demographic subgroups of the population, such as student race and economically disadvantaged, and on specific groups of campuses such as charter schools compared to traditional ISD campuses.

***Table Three**. Summary of overall trends observed for student subgroups*

Charter schools & ISD schools	**The first study found TFA-affiliated teachers are more effective than non-TFA-affiliated teachers, on average, in both traditional ISD campuses and charter schools; this advantage over non-TFA-affiliated teachers is greater in ISD campuses than charters. The second study confirmed similar small advantages of TFA in both ISDs and charter schools.**	
	Study One: In traditional ISDs, students of TFA alumni were 7% more likely to pass STAAR than students of non-TFA veteran teachers, compared to a 2% advantage in charter schools. Differences are less notable for TFA corps members; students of TFA corps members are slightly less likely to pass STAAR in charter schools (-0.8%), and slightly more likely to pass in traditional ISDs (0.2%).	**Study Two:** In traditional ISDs, students of TFA alumni were 1% more likely to pass STAAR than students of non-TFA veteran teachers. Like study one, this advantage for TFA alumni was 2% in charter schools. Also like study one, students of TFA corps members in both traditional ISDs and charters were as likely to pass STAAR, with differences in likelihoods near zero.
Student race	**The first study found TFA-affiliated teachers are more effective than non-TFA-affiliated teachers, on average, for all ethnicities of students; this advantage over non-TFA-affiliated teachers is slightly greater and more consistent for Black and Hispanic students of TFA alumni. The second study confirmed that there is an advantage for Black and Hispanic students who have TFA alumni teachers.**	
	Study One: Black students of TFA alumni were 7.6% more likely to pass STAAR than black students of non-TFA veteran teachers, while Hispanic students of TFA alumni were 6.6% more likely to pass STAAR. This advantage of TFA alumni teachers was smaller for White students at 3.2%. However, for TFA corps members, there is not a notable advantage for Black and Hispanic students.	**Study Two:** Black students of TFA alumni were 1.5% more likely to pass STAAR than black students of non-TFA veteran teachers, while Hispanic students of TFA alumni were 3.6% more likely to pass STAAR. There was not advantage of TFA for White students. Similar to study two, there is not a notable advantage specifically for Black and Hispanic students taught by TFA corps members.
Economically disadvantaged	**The first study found TFA-affiliated teachers are equally effective, on average, for both economically disadvantaged (EcoDis) and non-economically disadvantaged students. The second study confirmed this finding for economically disadvantaged students, particularly students of TFA alumni.**	
	Study One: EcoDis students of TFA-affiliated teachers are 2.3% more likely to pass STAAR than peers taught by non-TFA-affiliated teachers. This advantage is similar for non-EcoDis students with a 2.6% advantage for non-EcoDis students of TFA-affiliated teachers. This advantage is greater for students of TFA alumni; EcoDis students of TFA alumni are 6.7% more likely to pass STAAR, on average, and non-EcoDis students are 7.6% more likely to pass. Differences are less notable for TFA corps members.	**Study Two:** Similar to the first study, the second study found EcoDis students of TFA-affiliated teachers are slightly more likely to pass STAAR (1%) than peers taught by non-TFA-affiliated teachers. EcoDis students of TFA alumni are 3% more likely to pass STAAR than peers taught by non-TFA affiliated veteran teachers. EcoDis students of TFA corps members were equally likely to pass compared to peers taught by non-TFA novice teachers. Unlike the first study, the second study found no TFA advantage for non-EcoDis students.
Limited English proficient (LEP)	**The first study found TFA-affiliated teachers are more effective than non-TFA-affiliated teachers, on average, for both LEP and non-LEP students; this advantage over non-TFA-affiliated teachers is greater for LEP students. The second study confirmed the TFA advantage for LEP students but found no relative advantage of TFA for non-LEP students.**	
	Study One: LEP students of TFA-affiliated teachers are 5.3% more likely to pass STAAR than LEP peers taught by non-TFA-affiliated teachers. This advantage is 1.8% for non-LEP students. This advantage is greater for LEP students of TFA	**Study Two:** LEP students of TFA teachers are 5.4% more likely to pass STAAR than LEP peers taught by non-TFA affiliated teachers. There is no overall TFA advantage for non-TFA students. This advantage is greater for students of TFA alumni. On average,

	alumni; LEP students of TFA alumni are 9.3% more likely to pass STAAR, on average, compared to a 6.3% advantage for non-LEP students of TFA alumni. There is no relative advantage of TFA corps members for non-LEP students (-0.2% advantage), while LEP students of TFA corps members are 3.3% more likely to pass, on average.	LEP students of TFA alumni are 8.6% more likely to pass STAAR than LEP peers taught by non-TFA veteran teachers with similar experience. Also similar to study one, LEP students of TFA corps members are 3.4% more likely to pass STAAR compared to peers taught by novice non-TFA teachers.

Limitations

A notable limitation of the current studies—though additional analyses to address this are planned—is that previous achievement levels of the students were not controlled during the PSW procedure. Because the focus of this current study was Texas impact broadly, prior performance was not controlled for due to data-loss concerns. For instance, all 2011-12 school year and all 3rd grade data for six years would be excluded from that cohort view, as the records of prior achievement will not be available. Nevertheless, understanding the relative impacts of having a TFA corps member or alumni teacher in conjunction with all of the other covariates addressed in the PSW procedure provide actionable and meaningful outcome data. The current analyses predicted STAAR pass likelihood and controlled for a great deal of contributing variability but planned analyses will further strengthen evidence related to TFA impacts by also controlling for prior academic performance. This will require identification of a "good track" cohort of students who have sufficient data across multiple years. This cohort will be slightly smaller than the existing sample, although sample sizes should remain sufficient.

These analyses are not able to account for initiatives designed to improve student achievement. Statewide, since the 2011-12 school year, student achievement on STAAR has improved overall. For the State of Texas, 12% more students passed STAAR in 2018-19 than in 2011-12 (1,378,470 additional students meeting standards). These improvements are even larger in Dallas ISD, where 18% more students are passing STAAR over the same time period[7]. While this study does not account for how TFA is associated with this overall statewide improvement, it is reasonable to hypothesize that the effect of TFA relative to non-TFA teachers could have declined over time as all students are, on average, generally improving on statewide assessments of achievement.

For these evaluations we made an analytic decision to exclude: students who changed schools during the school year, and teachers who came in part way through a school year or who were not the teacher of record (e.g., teacher's aide or assistant). Additional exclusions—that were warranted given the central evaluation questions for this report—included deciding to only analyze data for teachers who were the *only* teacher of record for a given school year. This necessarily excluded teachers who may have co-taught with other teachers. Finally, we excluded special education students that took the modified version of the STAAR exam. Because these students take a modified test, their test scores are on a different scale and passing thresholds are modified. Including them with the majority group would have likely introduced skewness to the outcome data. These excluded samples represent critically important variation in student and teacher experiences that should be taken into account in future planned analyses.

While the current analyses identify differences in student outcomes that can be attributed to having a TFA-affiliated teacher or not, they do not tell us anything in particular about the non-TFA-affiliated group of teachers. Other than knowing the number of years the non-TFA-affiliated sample had been teaching, we do not have available data about their terminal degrees, whether they attended a traditional or alternative teacher certification program, or even whether they had teaching experience in another state. The nature of the data necessitated that we treat the non-TFA-affiliated group quite homogeneously while they certainly are not. Identifying sub-groups of non-TFA-affiliated teachers would have important implications for understanding TFA's impact relative to other alternative teacher certifications programs.

[7] Commit Partnership. STAAR Comparison Tool. https://commitpartnership.org/dashboard/visualizations/staar-comparison-tool

References

Clark, M. A., Chiang, H. S., Silva, T., McConnell, S., Sonnenfeld, K., Erbe, A., & Puma, M. (2013). The effectiveness of secondary Math teachers from Teach For America and the Teaching Fellows programs (NCEE 2013-4015). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Retrieved from https://eric.ed.gov/?q=Teach+for+America+&pg=2&id=ED544171

Clark, M. A., Isenberg, E., Liu, A. Y., Makowsky, L., & Zukiewicz, M. (2017). Impacts of the Teach for America Investing in Innovation scale-up. Princeton, NJ: Mathematica Policy Research. Retrieved from http://www. Mathematica-mpr.com/our-publications-and-findings/publications/impacts-of-the-teach-for-america-investing-in-innovation-scaleup

Cohen, J. E. (1988). Statistical power analysis for the behavioral sciences. Hillsdale, NJ: Lawrence Erlbaum Associates.

Decker, P. T., Mayer, D. P., and Glazerman, S. (2004). The effects of Teach For America on students: Findings from a national evaluation. Mathematica Policy Research. Retrieved from www.Mathematica-mpr.com/our-publications-and-findings

Glazerman, S., Mayer, D., & Decker, P. (2006). Alternative routes to teaching: The impacts of Teach for America on student achievement and other outcomes. Journal of Policy Analysis and Management, 25(1), 75-96.

Haddock, C., Rindskopf, D., & Shadish, W. (1998). Using odds ratios as effect sizes for meta-analysis of dichotomous data: A primer on methods and issues. Psychological Methods, 3(3), 339-353.

Hansen, M., Backes, B., Brady, V., & Xu, Z. (2015). Examining spillover effects from Teach for America corps members in Miami–Dade County Public Schools (CALDER Working Paper 113). Washington, DC: National Center for Analysis of Longitudinal Data in Education Research. Retrieved from https://caldercenter.org/sites/default/files/WP%20113_0.pdf

Harder, V. S., Stuart, E. A., & Anthony, J. C. (2010). Propensity score techniques and the assessment of measured covariate balance to test causal associations in psychological research. Psychological Methods, 15(3), 234-249.

Henry, G. T., Purtell, K. M., Bastian, K. C., Fortner, C. K., Thompson, C. L., Campbell, S. L., & Patterson, K. M. (2014). The effects of teacher entry portals on student achievement. Journal of Teacher Education, 65(1), 7-23.

Houston Independent School District. (n.d.). Teach for America (TFA) 2010−2011. Houston, TX: Houston Independent School District, Department of Research and Accountability. Retrieved from https://www.houstonisd.org/cms/lib2/TX01001591/Centricity/Domain/8269/TFA%202011full%20report.pdf

Houston Independent School District. (2018, January). Teach for America evaluation, 2017. Houston, TX: Houston Independent School District, Department of Research and Accountability. Retrieved from https://www.houstonisd.org/cms/lib2/TX01001591/Centricity/domain/8269/pe_cirriculum/TFA%20Program%20Eval%20Report_12%2018%2017-dy.pdf

Leite, W. L. (2016). Practical propensity score methods using R. Thousand Oaks, CA: Sage.

Lipsey, M. W., & Wilson, D. B. (2001). Practical meta-analysis. Thousand Oaks, CA: Sage.

Mickelson, N. R., & McEnturff, A. (2015, January). Evaluation of Teach for America: 2014-2015 (EA 15-536-2). Dallas, TX: Dallas Independent School District, Department of Evaluation and Assessment. Retrieved from https://www.dallasisd.org/cms/lib/TX01001475/Centricity/domain/98/evaluation/14-15/finalrpts/EA15-536-2%20Teach%20For%20America.pdf

Penner, E. K. (2016). Teaching for all? Teach For America's effects on the distribution of student achievement. Journal of Research on Educational Effectiveness, 9(3), 259-282. https://doi.org/10.1080/19345747.2016.1164779

Raymond, M., Fletcher, S. H., and Luque, J. (2001). *Teach For America: An evaluation of teacher differences and student outcomes in Houston, Texas*. Stanford, CA: Hoover Institution, CREDO. Retrieved from http://credo.stanford.edu/downloads/tfa.pdf

Selya, A. S., Rose, J. S., Dierker, L. C., Hedeker, D., & Mermelstein, R. J. (2012). A practical guide to calculating Cohen's f2, a measure of local effect size, from PROC MIXED. Frontiers in Psychology, 3, 111. https://doi.org/10.3389/fpsyg.2012.00111

Setoguchi, S., Schneeweiss, S., Brookhart, M. A., Glynn, R.J., & Cook, E. F. (2008). Evaluating uses of data mining techniques in propensity score estimation: A simulation study. Pharmacoepidemiology and Drug Safety, 17(6), 546-555. https://doi.org/10.1002/pds.1555

Turner, H. M., Goodman, D., Adachi, E., Brite, J., & Decker, L. E. (2012). Evaluation of Teach For America in Texas schools. San Antonio, TX: Edvance Research.

U.S. Department of Education. (2016, August). Teach for America. What Works Clearinghouse Intervention Report. Retrieved from https://ies.ed.gov/ncee/wwc/Docs/InterventionReports/wwc_tfa_083116.pdf

Ware, A., LaTurner, R. J., Parsons, J., Okulicz-Kozaryn, A., Garland, M., & Klopfenstein, K. (2011). Teacher preparation programs and Teach for America research study. Dallas, TX: Education Research Center, The University of Texas at Dallas.

Xu, Z., Hannaway, J., & Taylor, C. (2007). Making a difference? The effects of Teach For America in high school (CALDER Working Paper 17). Washington, DC: National Center for Analysis of Longitudinal Data in Education Research. Retrieved from: https://eric.ed.gov/?id=ED509654

Appendix. Student and Teacher Samples

Table A1. *Student sample demographics by school year*

		AY2011-12		AY2012-13		AY2013-14		AY2014-15		AY2015-16		AY2016-17		AY2017-18		AY2018-19	
		n	%	n	%	n	%	n	%	n	%	n	%	n	%	n	%
Total		1,192,774	100	1,207,608	100	1,212,302	100	1,214,361	100	1,197,446	100	1,273,827	100	1,469,792	100	1,510,979	100
Gender																	
	Male	602,345	50.5	608,818	50.4	611,703	50.5	610,403	50.3	601,982	50.3	646,727	50.8	738,054	50	762,796	51
	Female	590,429	49.5	598,790	49.6	600,599	49.5	603,958	49.7	595,464	49.7	627,100	49.2	731,738	50	748,183	50
Race																	
	Black	182,319	15.3	181,730	15	178,975	14.8	179,898	14.8	176,851	14.8	187,996	14.8	241,031	16	245,881	16
	Hispanic	709,146	59.5	728,799	60.4	738,961	61	742,746	61.2	735,629	61.4	791,652	62.1	895,096	61	926,786	61
	White	235,114	19.7	230,304	19.1	225,007	18.6	219,828	18.1	211,873	17.7	214,794	16.9	239,233	16	239,631	16
	Other	66,195	5.5	66,775	5.5	69,359	5.7	71,889	5.9	73,093	6.1	79,385	6.2	94,432	6.4	98,681	6.5
LEP*		156,675	13.1	156,680	13	168,630	13.9	180,663	14.9	183,745	15.3	226,541	17.8	279,170	19	324,926	22
Economically Disadvantaged		767,837	64.4	783,601	64.9	791,842	65.3	777,727	64	771,028	64.4	833,538	65.4	999,405	68	1,023,242	68
Grade Level																	
	ES (3rd-5th)	193,543	16.2	185,966	15.4	176,295	14.5	167,749	13.8	148,259	12.4	146,136	11.5	474,815	32	474,921	31
	MS (6th-8th)	445,434	37.3	454,688	37.7	464,763	38.3	463,596	38.2	464,713	38.8	503,920	39.6	433,323	30	461,807	31
	HS (9th-12th)	553,797	46.4	566,954	46.9	571,244	47.1	583,016	48	584,474	48.8	623,771	49	561,655	38	574,252	38

**LEP=limited English proficiency status*

The next table describes the final sample of teachers retained for the analyses. For the first study focused on the 2011-12 through 2016-17 school years, the final sample included 2,477 active TFA corps members (during any year used in the analyses), 1,479 TFA alumni, and 109,649 non-TFA-affiliated comparison teachers. For the second study focused on the 2017-18 and 2018-19 school years, the final sample included 738 active TFA corps members (during any year used in the analyses), 1,130 TFA alumni, and 89,216 non-TFA-affiliated comparison teachers.

Table A2. *Teacher sample by TFA affiliation status*

		Study One: 2011-12 through 2016-17						Study Two: 2016-17 through 2018-19					
		Corps Members		Alumni		Non-TFA		Corps Members		Alumni		Non-TFA	
		n	%	n	%	n	%	n	%	n	%	n	%
Race													
	Black	410	16.6	247	16.7	16,943	15.5	144	19.5	209	18.5	15,785	17.7
	Hispanic	485	19.6	262	17.7	31,208	28.5	210	28.5	332	29.4	27,399	30.7
	White	1,308	52.8	820	55.4	56,332	51.4	283	38.3	482	42.7	41,770	46.8
	Other	274	11.1	150	10.1	5,166	4.7	101	13.7	107	9.5	4,262	4.8
Grade Level													
	Elementary (3rd-5th)	361	14.4	197	12.6	25,872	22.6	206	27.9	337	29.8	34,305	38.4
	Middle (6th-8th)	1,157	46	740	47.3	44,322	38.7	310	42	413	36.5	26,207	29.4
	High School (9th - 12th)	997	39.6	628	40.1	44,443	38.8	222	30.1	380	33.6	28,710	32.2
Area													
	Austin	-	0	128	8.6	8,291	7.6	23	3.1	140	12.4	6,199	6.9
	DFW	900	36.4	349	23.4	34,767	31.7	329	44.6	431	38.1	29,921	33.5
	Houston	747	30.2	619	41.5	35,793	32.7	193	26.2	277	24.5	31,612	35.4
	Rio Grande Valley	391	15.8	216	14.5	14,736	13.4	97	13.1	132	11.7	9,804	11
	San Antonio	434	17.6	178	11.9	16,031	14.6	96	13	150	13.3	11,686	13.1

Note: Total count of teachers within each group fluctuate between race, grade level and area due to missing data. These numbers are based on the final cleaned dataset and do not reflect the actual teacher distribution within each region.