A Redesigned Training Program for New Teachers

Findings from a Study of Teach For America's Summer Institutes

Shelley Rappaport
Marie-Andrée Somers
Kelly Granito

March 2019

mdrc
BUILDING KNOWLEDGE TO IMPROVE SOCIAL POLICY

Funding for this report came from a Supporting Effective Educator Development (SEED) grant from the U.S. Department of Education. The goal of the SEED initiative is to increase the number of highly effective educators by supporting the implementation of evidence-based preparation, development, or enhancement opportunities for educators. Teach For America received a SEED grant to redesign its summer training for corps members, and MDRC conducted an independent evaluation of the implementation and impacts of that training.

Dissemination of MDRC publications is supported by the following funders that help finance MDRC's public policy outreach and expanding efforts to communicate the results and implications of our work to policymakers, practitioners, and others: The Annie E. Casey Foundation, Charles and Lynn Schusterman Family Foundation, The Edna McConnell Clark Foundation, Ford Foundation, The George Gund Foundation, Daniel and Corinne Goldman, The Harry and Jeanette Weinberg Foundation, Inc., The JPB Foundation, The Joyce Foundation, The Kresge Foundation, Arnold Ventures, Sandler Foundation, and The Starr Foundation.

In addition, earnings from the MDRC Endowment help sustain our dissemination efforts. Contributors to the MDRC Endowment include Alcoa Foundation, The Ambrose Monell Foundation, Anheuser-Busch Foundation, Bristol-Myers Squibb Foundation, Charles Stewart Mott Foundation, Ford Foundation, The George Gund Foundation, The Grable Foundation, The Lizabeth and Frank Newman Charitable Foundation, The New York Times Company Foundation, Jan Nicholson, Paul H. O'Neill Charitable Foundation, John S. Reed, Sandler Foundation, and The Stupski Family Fund, as well as other individual contributors.

The findings and conclusions in this report do not necessarily represent the official positions or policies of the funders.

For information about MDRC and copies of our publications, see our website: www.mdrc.org.

Overview

In summer 2015, roughly six years after the Great Recession of 2007-2009 ended, the American media began reporting on an emerging teacher shortage. Public schools were reinstating programs and classes that they had cut during the recession, and student enrollment was increasing. Meanwhile, enrollment in teacher training programs was decreasing. Teacher shortages are especially problematic for schools serving mainly students from low-income families and communities of color.

Teach For America (TFA), founded in 1989, has grown to become one of the largest providers of educators in the country for high-needs, underresourced schools. TFA recruits high-performing college graduates and professionals and prepares them for teaching during five to seven weeks of intensive training at its regional and national summer "institutes" before their first teaching job. In 2015, TFA was awarded a Supporting Effective Educator Development (SEED) grant from the U.S. Department of Education to redesign its summer training for new teachers, or "corps members" (CMs). In summer 2016, TFA piloted a redesigned training model, developed specifically for adult learners, that incorporated more rigorous academic standards and training methods and practices with an eye toward college and career readiness. The new model differed substantially from TFA's traditional training, most notably in the *method of delivering training* to CMs and the use of *content-based instructional activities*.

This report presents the findings from a study that examined how the redesigned training was implemented in a TFA training site (program group) and how it compared with the usual training at other TFA sites (comparison group). It also examined the promise of the redesigned training to improve the short-term outcomes of the first cohort of TFA teachers to receive it — that is, CMs' use of the new instructional strategies and teaching practices, their perceptions of the value of the training, their commitment to teaching and to equity, and their retention rates in the TFA program. Not surprisingly, given the complexity of the new training, TFA encountered a number of implementation challenges:

- It was more difficult to implement the redesigned training than anticipated, as it involved preparing TFA summer staff, in less than one year, to use radically different methods from those of TFA's traditional training to teach corps members.
- Lead instructors in the program group did not consistently or deliberatively model the new teaching strategies and practices, and as a result, no differences were found between CMs in the program group and the comparison group in their use of these new strategies and practices.
- CMs in the program group did not have more positive perceptions of the value of their summer training than did those in the comparison group.
- The redesigned training does not appear to have improved CMs' commitment to teaching and educational equity, or to have improved their retention rates in the TFA program.

Although the hoped-for outcomes did not materialize, the first cohort of CMs who received the redesigned training performed no better or worse than CMs who received TFA's usual training. The lessons from this study have helped TFA address the challenges that were encountered during implementation and to strengthen the redesigned training model. TFA has now revamped the model and in summer 2017 began to scale it up at their national institutes.

Contents

Appendix

List of Exhibits

Table

Table (continued)

Table (continued)

Figure

Figure (continued)

Box

Acknowledgments

This evaluation of Teach For America's (TFA's) redesigned training model and the resulting report reflect the efforts of a great many people. Our first debt of gratitude is to the corps members and TFA training staff who took time out of their busy schedules to participate in focus groups and interviews and to complete logs during the study. The assistance and cooperation of these individuals were vital for enabling the study to move forward and for providing the rich and detailed information on which this report is based.

At TFA, Hana Merkle, Jamila Singleton Munson, Rita Zota, Jennifer Salman, and Jessica Besser-Rosenberg spent countless hours giving us a strong understanding of the redesigned training model and making it possible for us to be observers at conferences in which TFA lead instructors and coaches were trained to work with corps members. They were also unfailing in their availability to answer questions and provide support. Grant Van Eaton, Hana Merkle, Chrissy Heyne, and Kristen Lambertz-Berndt at TFA provided useful critiques during the report drafting process. Throughout the project, Shane Traister, Amirah Patterson, and Semra Malik answered our many questions about the surveys and administrative data collected by TFA, and they provided several well-organized datasets that were used to describe the characteristics and outcomes of the corps members in this study.

Stephanie Levin from IMPAQ, as a technical assistance provider for the Supporting Effective Educator Development (SEED) grants, offered thoughtful and welcome advice about all aspects of the study design and the findings.

At MDRC, Elena Serna-Wallender led the operations work to recruit research participants and support data collection efforts; she also conducted some of the focus groups. Sonia Drohojowska was the resource manager during the proposal stage and during the first few months of working on the report. Andrea Shane, Kelly Quinn, Nicholas Commins, Deborah Van Kummer, Bulent Can, Usha Krishnan, Jaye Song, Zachary Pinto, and Laura Wang provided programming and analysis support. Larissa Saco, Christopher Boland, and Joseph Quinn contributed in various ways to the qualitative research, including the collection and coding of data and its analysis. Larissa and Nicholas also coordinated the production of the report, including making sure that deadlines were met, preparing exhibits, and verifying the accuracy of the report content. Fred Doolittle, Alice Tufel, Jean Grossman, Leigh Parise, Michelle Maier, and Janet Quint carefully reviewed earlier drafts of the report and offered helpful critiques throughout the writing process. Christopher Boland edited the report, and Carolyn Thomas prepared it for publication.

The Authors

Executive Summary

In summer 2015, roughly six years after the Great Recession of 2007-2009 ended, the American media began reporting on an emerging teacher shortage. Public schools were reinstating programs and classes that they had cut during the recession, and student enrollment was increasing after years of little or no growth. Meanwhile, enrollment in teacher training programs was decreasing, giving rise to concerns about an impending teacher shortage.[1] Teacher shortages are especially problematic for schools serving mainly students from low-income families and communities of color. These schools are generally underresourced and usually have the highest attrition rates and the most vacancies for teachers.[2]

Founded in 1989, Teach For America (TFA) has grown to become one of the largest providers of educators in the country for high-needs schools. Since its inception, TFA has trained over 50,000 teachers, also known as corps members (CMs), who are placed in high-needs schools to fill vacancies. CMs commit to teach in underresourced schools for at least two years. The majority of CMs have never taught and have no background in education; however, they receive intensive training during the summer before beginning to teach, at what is familiarly known as an "institute," that includes teaching summer school students. Several studies, including three large randomized controlled trials (RCTs), have found that TFA teachers are either more effective than or as effective as the non-TFA teachers with whom they were compared.[3]

Nonetheless, similar to many novice teachers, CMs face many challenges as they enter the classroom for the first time, and the institute is where TFA can have the biggest impact on how well CMs are prepared to teach effectively and to have positive effects on their students. One of TFA's core values is to learn continuously.[4] In this spirit, TFA applied for — and was awarded — a Supporting Effective Educator Development (SEED) grant from the U.S. Department of Education to redesign the way it trains CMs before their first year of teaching by enhancing the rigor and the relevance of the institute training. This new approach was piloted at one of TFA's institutes in summer 2016.

---

[1]Leib Sutcher, Linda Darling-Hammond, and Desiree Carver-Thomas, *A Coming Crisis in Teaching? Teacher Supply, Demand, and Shortages in the U.S.* (Palo Alto, CA: Learning Policy Institute, 2016).

[2]Frank Adamson and Linda Darling-Hammond, *Addressing the Inequitable Distribution of Teachers: What It Will Take to Get Qualified, Effective Teachers in All Communities* (Stanford, CA: Stanford Center for Opportunity Policy in Education, 2011).

[3]Paul T. Decker, Daniel P. Mayer, and Steven Glazerman, *The Effects of Teach For America on Students: Findings from a National Evaluation* (Princeton, NJ: Mathematica Policy Research, 2004); Melissa Clark, Hanley S. Chiang, Tim Silva, Sheena McConnell, Kathy Sonnenfeld, Anastasia Erbe, and Michael Puma, *The Effectiveness of Secondary Math Teachers from Teach For America and the Teaching Fellows Programs* (Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education, 2013); Melissa Clark, Eric Isenberg, Albert Liu, Libby Makowsky, and Marykate Zukiewicz, *Impacts of the Teach For America Investing in Innovation Scale-Up* (Princeton, NJ: Mathematica Policy Research, 2017).

[4]Teach For America, "Our Values," accessed October 16, 2018, website: www.teachforamerica.org.

This report presents the findings from an independent, quasi-experimental evaluation that examined TFA's efforts to implement its new training model and its effect on CMs' perceptions of the value of the training, their commitment to teaching and equity, their use of instructional strategies aligned with college- and career-ready standards,[5] their use of teaching practices that were part of the redesign, and their retention in the TFA program. MDRC — a nonprofit, nonpartisan education and social policy research organization — conducted the evaluation. The study found that the implementation of the redesigned summer training was a complex undertaking and more difficult to accomplish than anticipated, especially given that TFA had less than one year to prepare. (The pilot began 10 months after TFA was awarded the SEED grant.) It required preparing staff to train CMs in the use of methods that were radically different from TFA's traditional training, as well as restructuring the training process at the pilot institute. Some of the TFA staff who trained CMs were not sufficiently trained themselves for their role and reported that they felt inadequately prepared to support CMs during training at the institute. Moreover, they reported that the restructuring of the training process at the pilot institute did not take into account the additional time needed for staff to collaborate in order to align their curricula and goals. Not surprisingly, given these implementation challenges, the redesigned training does not appear to have had an effect on any of the outcomes that were examined (listed above). Although the hoped-for outcomes did not materialize, the outcomes of the first cohort of teachers to experience the redesigned training were not affected adversely. Moreover, lessons from the first year of implementing the redesigned training were invaluable for helping TFA strengthen the new model, which was scaled up to all TFA institutes beginning in summer 2017.

Teach For America's Redesigned Summer Training Model

Once they are selected, CMs are assigned to one of over 50 TFA regions throughout the United States.[6] CMs then look for teaching positions within their assigned regions.[7] In the summer before their first year of teaching, CMs receive five to seven weeks of intensive training from TFA, and as part of that training they have the opportunity to teach summer school students.[8] The training takes place at regional and national institutes and is typically led by TFA alumni, or "lead instructors," many of whom teach in schools during the academic year. The regional institutes host CMs from one region each, while each national institute — of which there were six in 2016 — hosts

[5]The college- and career-ready standards that TFA used were based on the Common Core Standards and the Next Generation Science Standards.

[6]In 2016, there were 53 TFA regions; in 2018, there were 51 TFA regions. Unlike in the past, CMs now have considerable say over the regions to which they are assigned. This gives CMs the choice to live and work in an area where they may decide to stay, increasing the likelihood that CMs will continue to teach beyond their two-year commitment and in a community where they can more easily advance educational equity and serve as leaders.

[7]TFA supports CMs in their endeavors to find teaching positions through its partnership with districts and by helping CMs create resumes and attend job fairs.

[8]CMs are responsible for teaching their own classes, but the teacher of record, who works for the school district (and not for TFA), is often present in the classroom and may help support CMs.

CMs from several regions. Where and when CMs are trained depends on the region to which they are assigned.

In summer 2016, a redesigned training model was piloted at the national institute in Tulsa, which trained CMs from eight TFA regions. The training was radically different from TFA's traditional training, as mentioned above and explained in more detail below. The goal of the new training was to better address the needs of students in schools that serve mainly low-income students by enhancing the rigor and the relevance of the summer training. TFA's traditional summer training was offered at the other five national institutes.[9]

The Learning Cycle

The redesigned training piloted in Tulsa included several specific features that differentiated it from the traditional training offered at the other national institutes, the most notable being the *method of delivering training* to CMs and the *content of the training*. With respect to its delivery, the training in Tulsa was conveyed via the Learning Cycle — a four-step process in which teachers observe, practice, teach (in this case, to summer school students), and then receive feedback on how they implement different types of content-based instructional activities, such as engaging students in a discussion about a book or counting out loud according to a pattern.

CMs in Tulsa received their training in cohorts where all of the members would be teaching in like subjects and grade levels. Lead instructors conducted the Learning Cycle sessions, and coaches provided support (for instance, giving feedback to CMs as they practiced the lessons they would teach, observing CMs teaching, and helping them to reflect on their practice) within the Learning Cycle sessions and in small groups outside of those sessions.

TFA had first tried using the Learning Cycle in 2014 at one of its regional institutes, in partnership with the Teacher Education by Design program at the University of Washington, where the Learning Cycle was developed. TFA used its adaptation of the Learning Cycle independently for the first time during the 2016 pilot for the redesigned training. It incorporated several "andragogies," or training methods, meant specifically for adult learners. For example, the CMs had many opportunities to practice what they would teach in front of their peers and they were given time to reflect on their teaching of students with their coaches and peers. The emphasis of the training in Tulsa was on *how* to teach with equity and according to their students' needs. CMs there were also introduced to rigorous *content-specific* lessons with predictable structures that were both appropriate for their summer school students and manageable for novice teachers to use while learning how to teach. For instance, while practicing the teaching of a lesson on counting according to a pattern, CMs were asked to consider strategies to get students' attention, to make sure that directions were clear, to include all students, to get students engaged in rigorous discussion with each other, and so forth. Taken together, the strategies that CMs were asked to consider make up the *core practices* they would need to incorporate into their teaching for it to be successful — how to create and maintain a productive learning environment, position students

[9]The other five national institutes were located in Philadelphia, Pennsylvania; Houston, Texas; Greater Delta (comprising Mississippi and Arkansas); Phoenix, Arizona; and Atlanta, Georgia.

so they can make sense of what is being taught, and teach toward an instructional goal and do this with equity.[10] In contrast, CMs attending the other traditional national institutes — the comparison sites — received training on discrete, content-neutral topics such as classroom management and lesson planning largely through a lecture format, with fewer systematic opportunities to practice teaching in front of peers or students.

Because the redesigned training differed from the traditional training, TFA had to develop new training for the lead instructors and coaches. Lead instructors and coaches at the other five national institutes (the comparison sites) also received training beforehand. However, their training was the training traditionally provided by TFA.

The underlying logic model for the new training posits that the training and coaching received by the CMs in Tulsa will enable them to effectively teach rigorous lessons grounded in core practices via the Learning Cycle. The training was expected to help the CMs start the school year feeling well prepared and confident in their teaching skills. The CMs' bolstered confidence and instructional abilities were, in turn, expected to strengthen their commitment to teaching and to TFA's mission, and to make them more likely to remain in the program for the duration of the two-year commitment.

Evaluation Design

The evaluation of the redesigned training had two objectives: (1) to examine how the redesigned summer training was implemented at the national institute in Tulsa, and how it compared with the traditional training that TFA offered at the other national institutes; and (2) to examine the effects of the new training on the short-term outcomes of the first cohort of CMs to participate in it.

Several types of data were collected to achieve these goals. In summer 2016, the study team visited the institute in Tulsa, as well as two of the five national institutes that offered the traditional training. During the visits to Tulsa, the study team observed the training and the CMs' summer school teaching. In Tulsa and in the two comparison sites, the team conducted focus groups with CMs, their lead instructors, and coaches. To learn about CMs' experience teaching in the classroom, open-ended logs were sent to CMs in the 2016 cohort every two weeks during their first year of teaching, to capture their use of the core practices taught at the institute. The study team also conducted follow-up phone interviews with a sample of CMs at the end of their first year of teaching. To measure outcomes for CMs, the study team leveraged TFA's administrative records; teacher surveys that TFA administered regularly; and biweekly, closed-ended teacher logs that the study team administered during CMs' first year of teaching.

A comparative interrupted time series (CITS) design was used to look at the effect of the redesigned training on most outcomes (perceptions of the training, commitment to teaching, and retention). The study compared CMs in two groups of TFA regions: One group trained at the national institute in Tulsa where the redesigned training was offered (the program group, which

[10]The core practices stem from a research-based collaboration among many teacher preparation programs that make up the Core Practices Consortium. See www.corepracticeconsortium/corepractice.

included CMs from 8 regions) and another group trained at the five other national institutes where the traditional summer training was offered (the comparison group, which included CMs from 13 regions). The study followed the outcomes of five cohorts of CMs in these regions: four "baseline" cohorts that were trained before the pilot at the national institute in Tulsa began (2012 to 2015 cohorts), and one follow-up cohort that was trained the year in which the pilot was implemented (2016 cohort). Using information from the baseline cohorts, the study team compared the estimated trends in the outcomes of CMs in each region with their expected outcomes had the regional baseline trends persisted ("deviations from baseline trend"). The effect of the redesigned trained was then estimated by looking at whether deviations from trend were larger on average (and more positive) for the program regions than the comparison regions.

Due to data limitations, the effect of the redesigned training on CMs' use of standards-aligned instructional strategies could not be examined using a CITS design. Instead, the study team used a cross-sectional comparison group design to compare the self-reported instructional strategies used by CMs in the program and comparison regions in the 2016 cohort, controlling for CMs' characteristics.

Because both the CITS and cross-sectional comparison group study designs are quasi-experimental, the findings presented in this report cannot definitively be interpreted as the *causal* effect of the redesigned summer training. The observed changes in CMs' outcomes could have been caused by other confounding factors (rather than the redesigned training). For this reason, the findings in this report are referred to as *estimates* of the effect of the redesigned training. However, it is more plausible that the estimated effects found using the CITS design can be attributed to the redesigned training because this design makes it possible to control for more of the factors that could be confounded with the effect of the redesigned training.

Key Findings

Implementation of the Redesigned Training Model

Overall, the study found that evaluating the redesigned summer training during its pilot year was premature, as strong implementation required more time for planning, practicing, and perfecting the model.

- **At the national institute in Tulsa, the lead instructors, who trained the CMs, reported that they felt prepared to teach via the Learning Cycle, while the coaches did not feel prepared for their role. The lead instructors and coaches in the comparison institutes said they felt more prepared for their training responsibilities than their counterparts in Tulsa.**

Lead instructors attended numerous training sessions led by TFA staff who were familiar with the new model, in which they were introduced to the Learning Cycle as "learners" (via role play) and given time to practice its teaching approach. That is, lead instructors played the part of CMs as TFA staff taught them the same content they would be expected to teach to CMs. They

then had opportunities to practice presenting the same content via the Learning Cycle with the coaches role-playing as CMs.

Coaches only experienced the Learning Cycle as learners. While they learned about the instructional activities when role-playing as CMs during the time that lead instructors practiced, they did not have the benefit of practicing teaching the content and the deeper learning that comes from that experience. In fact, they had a very limited understanding of what their responsibilities as coaches would be. Most coaches felt that they had not received adequate training.

- **Lead instructors at the national institute in Tulsa used some of the andragogies to train CMs via the Learning Cycle, but their incorporation of the core practices was minimal.**

CMs had opportunities to observe the instructional activities being modeled by the lead instructors and to discuss and practice some basic ways in which the core practices might be integrated into the instructional activities. However, because so many CMs were working with one lead instructor or coach at the same time and because the coaches were unclear about their role, CMs were often rehearsing for other CMs and they did not receive feedback from their lead instructors or coaches, who were much more experienced. While CMs had daily opportunities to teach summer school, they very rarely had the opportunity to get feedback from or reflect on their teaching with their coaches.

Lead instructors and coaches were not successful in helping CMs use the core practices to guide and shape the instructional activities. A possible explanation for this could be that lead instructors were directed to use their own judgment regarding which and how many core practices to feature in any Learning Cycle session. In addition, the lead instructors were not given specific instructions about how and when to incorporate the core practices in their sessions. Thus, the use or mention of the core practices varied considerably across Learning Cycle sessions, and, in general, CMs were not observed integrating the core practices into their summer school teaching. The teaching strategies that CMs deployed during the subsequent school year did not look very different regardless of whether they had received their training in Tulsa or at a comparison site.

Effect on CMs' Experience During the School Year

This study also examined the first-year teaching experience of the first cohort of CMs to participate in the redesigned training. The study team looked at different aspects of the school-year experience for both CMs trained at the national institute in Tulsa (program regions) and CMs trained at the other national institutes (comparison regions), to better understand how the experiences of CMs in these two groups differed.

- **CMs who participated in the redesigned training did not have more positive perceptions of the value of their summer training than CMs in the comparison national institutes.**

CMs' perceptions of the overall value of their summer training provides a broad measure of their satisfaction with it as preparation for their teaching placement. The study team collected

information about CMs' perceptions from surveys that TFA regularly administered at different points during CMs' first year of teaching (beginning, middle, and end) and in the fall of their second year.[11] Estimated effects were examined using a CITS design. Overall, the findings suggest that the redesigned training may have had a statistically significant negative effect on CMs' perceptions of the value of the summer training during their first year of teaching, but this effect ultimately dissipates in the fall of CMs' second year. By their second year, CMs may have figured out how to perform the tasks they initially found difficult (such as lesson planning and classroom management), which in turn may have improved their perceptions of the training.

- **Although CMs in the program regions used standards-aligned instructional strategies in the classroom, they did not use these strategies more often than CMs in the comparison regions.**

Information about CMs' use of instructional strategies aligned with college- and career-readiness standards comes from closed-ended logs that the study team administered every other week to CMs in the 2016 cohort who were teaching English language arts (ELA), mathematics, and general education during the first year of their placement. The items in the closed-ended logs asked CMs to report on their use of instructional practices that are commonly recognized as effective. The study team prioritized the following three instructional strategies that are most aligned with the redesigned training: CMs' use of culturally responsive texts (cultural responsiveness domain), the extent to which CMs asked students to cite text evidence (ELA domain), and a composite measure of CMs' use of standards-aligned math practices (math domain). The study team estimated the effect of the redesigned training by comparing the practices used by CMs in the program and comparison regions. Overall, it does not appear that the redesigned training increased CMs' use of standards-aligned instructional strategies. (See Table ES.1.) CMs in the program regions did use strategies aligned with college- and career-readiness standards; however, they did not use them more frequently than CMs in the comparison regions. Across all three measures, no statistically significant difference was observed between the program and comparison regions in the CMs' use of these strategies.

Interviews with CMs conducted in the spring of their first year of teaching support the quantitative findings. Most CMs felt that they learned the most about instruction by teaching during the school year. One among many possible explanations could be that teaching is hard and it is difficult for anybody — particularly for those teaching in underresourced schools that serve high-poverty populations — regardless of their training, to feel prepared for their first year.[12] Another explanation could be that the particular expectations of the schools in their assigned

[11]TFA also administers surveys in the middle and end of CMs' second year, but these were not examined because, at the time of the analysis, the 2016 cohort of CMs had not yet had a chance to complete their second year.

[12]See, for example, Amanda Bowsher, Dinah Sparks, and Kathleen Mulvaney Hoyer, "Preparation and Support for Teachers in Public Schools: Reflections on the First Year of Teaching," *Statistics in Brief*, NCES 2018143 (April 3, 2018); OECD, "Do New Teachers Feel Prepared for Teaching?," *Teaching in Focus,* No. 17 (2017).

Table ES.1

Estimated Effects on Self-Reported Instructional Strategies Aligned with College- and Career-Readiness Standards

Outcome	Program Regions	Comparison Regions	Estimated Difference	Effect Size	P-Value
English Language Arts					
Text used for the lesson featured an author, characters, and/or a community with similar background to focal student (% of classes)	38.33	49.25	-10.92	-0.21	0.118
Students were asked to cite text evidence in their writing (% of classes)[a]	75.82	81.17	-5.35	-0.12	0.447
Number of regions	8	12			
Number of corps members	87	138			
Number of logs	739	1143			
Mathematics					
Math instructional practice composite (z-score)[b]	-0.15	-0.04	-0.11	-0.12	0.415
Number of regions	8	13			
Number of corps members	80	151			
Number of logs	628	1197			

SOURCE: Closed-ended teacher logs administered by MDRC.

NOTES: Online teacher logs were administered to corps members (CMs) in the 2016 cohort in the program and comparison regions. CMs teaching English Language Arts (ELA), mathematics, or general education, in the first year of their teaching placement were eligible for the logs. Secondary school teachers, as well as elementary school teachers in departmentalized schools, were sent the log for their content area and level every two weeks (about twice per month). Elementary school teachers teaching both content areas (general education teachers) were sent the ELA log or the math log every two weeks, with the content area alternating between logs (about one log per content area per month). At the start of each log, teachers were given a randomly selected letter of the alphabet and asked to report on their instructional practices with a focal student whose name starts with that letter (or the closest letter) on a particular day that week.

The values in the "Program Regions" column are the observed mean outcome for teachers in the program regions. The values in the "Comparison Regions" column are the regression-adjusted mean outcome for teachers in the comparison regions, using the mean covariate values for the program region as the basis for the adjustment. The values in the "Estimated Difference" column are the difference between the program and comparison regions, adjusted for differences in the characteristics of the focal students and teachers. Values in the "Effect Size" column are the estimated effect divided by the standard deviation for the sample. A two-tailed t-test was applied to estimated differences. The statistical significance of estimated differences is indicated as follows: *** = 1 percent; ** = 5 percent; * = 10 percent. Rounding may cause slight discrepancies in calculating sums and differences.

Sample sizes in this table are the number of teachers in the log sample. A CM is included in the log sample if the CM completed at least one log during the school year.

[a]This item is only available for upper elementary and secondary teachers. The number of teachers included in this analysis is 60 in the program group and 93 in the comparison group.

[b]This composite combines the items in the math logs that map onto college- and career-readiness standards for math instruction. The scale of the items differs, so each item was first z-scored based on the mean and standard deviation of the comparison group logs; these z-scores were then averaged across items, and the average score was z-scored, again, based on the mean and standard deviation for the comparison logs. Because some of the log items differ across levels, the composite score was created separately for the lower elementary school and upper elementary or secondary school logs. The reliability (Cronbach's alpha) for the composite is 0.71 at the lower elementary school level and 0.75 at the upper elementary or secondary school level.

regions did not align perfectly with the expectations of the national institute. Or perhaps the CMs in Tulsa were not given specific instruction on how to incorporate the strategies leading to the core practices into their teaching. (For example, they were not sure how to engage their students in the content, sustain classroom discussions, or manage their classrooms.)

In addition, the CMs in Tulsa did not have a good grasp of how to plan their own lessons, as this was not part of their training. Although more of the CMs in the comparison regions felt that they could plan a lesson and were familiar with a particular classroom management system learned at their institutes, they neither knew the content better than CMs in Tulsa nor had a better sense of how to engage their students in the lesson or maintain classroom discussions.

However, CMs did report one aspect of the summer training to be quite useful: what they learned about diversity, equity, and inclusiveness. This was true for CMs who attended the national institute in Tulsa as well as for CMs who were trained at the comparison national institutes.

Commitment to Teaching and Retention

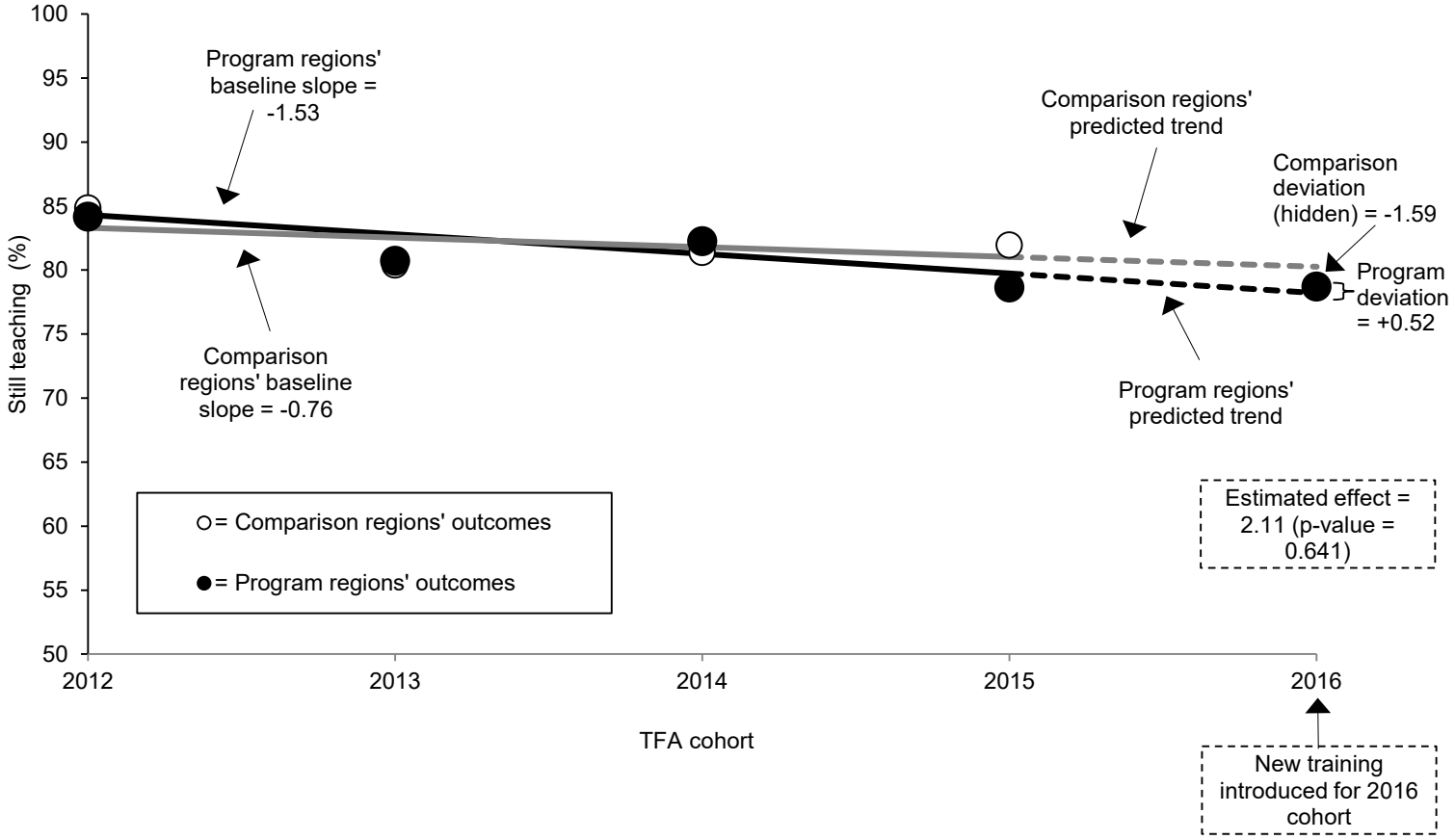
- **The redesigned training does not appear to have improved CMs' commitment to teaching and educational equity. Nor does it appear to have improved their retention rates in the TFA program.**

The study team assessed CMs' commitment to teaching and equity using items from teacher surveys that TFA administered regularly. The measure is a composite score created from eight items about CMs' beliefs in TFA's mission of ensuring that each child has an opportunity to receive an excellent education, and their beliefs in their own capacity to contribute effectively to this mission. The team examined CMs' commitment to teaching at three time points during their first year of teaching (beginning, middle, and end) and at the beginning of their second year. The study used a CITS design to estimate the effects of the redesigned training on this outcome. Overall, the findings suggest that the redesigned training does not appear to have had an effect on CMs' commitment to teaching and educational equity, neither during their first year of teaching nor in the fall of their second year.

The study also examined whether the redesigned summer training that was piloted in Tulsa had an effect on CMs' retention rates in the TFA program. Whether a CM was still a TFA teacher was measured at three points in time: the fall of CMs' first year of teaching, the end of their first year, and the fall of their second year. Overall, the findings from the CITS design suggest that the redesigned summer training did not have an effect on CMs' retention rates. Figure ES.1, for example, shows that before the redesigned training was launched (2012 to 2015 cohorts), second-year retention rates had been generally declining in the program and comparison regions. In both groups, the 2016 cohort's retention rates did not deviate from this trend visibly or statistically. This is true at all three time points when retention was measured.

Figure ES.1

Trends in Teacher Retention Rates in the TFA Program, Fall of Second Year



SOURCE: Teach For America administrative records.

Conclusion

Taken together, the findings from this study are consistent with what is known about the challenges of implementing new teacher training programs. The approach that was piloted at the national institute in Tulsa, which used the Learning Cycle and focused on content and core practices, was not only ambitious in scope but also a stark departure from TFA's traditional approach to training its CMs. The challenges that arose implementing it — and the associated lack of effects on teacher outcomes — are consistent with the complexity of the new training model and the fact that it was being tested for the first time in summer 2016. Nonetheless, the findings suggest that TFA was able to radically change its training approach without adversely affecting the outcomes of the first cohort of teachers to experience it.

Although the hoped-for outcomes did not materialize, TFA responded to the findings by adjusting and strengthening the model, which is now in its third year and is being scaled up to all TFA national institutes. As MDRC researchers observed in 2018, the training that lead instructors at the national institutes now undergo to prepare them to train CMs addresses almost all of the challenges encountered in summer 2016. TFA lead instructors and coaches are trained at a much deeper level about each of the steps of the Learning Cycle. Similarly, lead instructors and coaches are trained to make explicit to CMs how the core practices can be incorporated to guide and shape instruction, help them manage their classrooms, and create more equitable classrooms.

Chapter 1

Introduction

In summer 2015, roughly six years after the Great Recession, the American media began reporting on an emerging teacher shortage. Public schools were reinstating programs and classes that they had cut during the recession, and student enrollment in schools was increasing after years of little or no growth. Meanwhile, enrollment in teacher training programs was decreasing, giving rise to concerns about an impending teacher shortage.[1]

Teacher shortages are especially problematic for schools serving mainly students from low-income families and communities of color. These schools are generally underresourced and usually have the highest attrition rates and the most vacancies for teachers.[2] Attrition rates are highest among teachers with the least preparation, and these teachers are most often found in underresourced schools.[3] Students of these least-prepared teachers learn the least over the school year compared with their peers in better-resourced schools.[4] Given these realities, the training and placement of well-prepared teachers in high-needs schools has become an especially important policy goal.

Teach For America (TFA) has grown to become one of the largest providers of educators in the country for high-needs schools. Founded by Wendy Kopp in 1989, TFA describes itself as "a diverse network of leaders working to confront educational inequity through teaching and at every sector of society to create a country free from this injustice." Their vision is that "one day all children in this nation will have the opportunity to attain an excellent education."[5] To do this, TFA: (1) recruits outstanding and diverse leaders to be teachers (known as "corps members," or CMs) for a two-year commitment, (2) supports CMs in classrooms, (3) develops systems-change leaders, and (4) fosters the collective leadership of their CMs and alumni. TFA's long-term goal is to cultivate leaders who understand the issues facing the nation's schools to advance educational equity. Since its founding, TFA has grown from 489 CMs in its first cohort in 1990 to more than 50,000 alumni.[6] The majority of CMs enter TFA having never taught before and without a background in education.

[1]Sutcher, Darling-Hammond, and Carver-Thomas (2016).

[2]Adamson and Darling-Hammond (2011).

[3]Sutcher, Darling-Hammond, and Carver-Thomas (2016).

[4]Peske and Haycock (2006).

[5] See Teach For America website: https://www.teachforamerica.org/what-we-do/who-we-are

[6]All CMs who fulfill their two-year teaching commitment to TFA are considered alumni whether or not they stay in teaching.

Once selected, CMs are assigned to one of over 50 TFA regions throughout the United States.[7] CMs then look for teaching positions within their assigned regions.[8] In the summer before their first year of teaching, CMs receive five to seven weeks of intensive training from TFA that includes teaching summer school students.[9] The training, which is typically led by TFA alumni who work in schools during the school year, takes place at regional institutes and at multiregional national institutes. Where and when CMs are trained depends on the region to which they are assigned.

One of TFA's core values is to learn continuously.[10] In this spirit, TFA applied for — and was awarded — a Supporting Effective Educator Development (SEED) grant from the U.S. Department of Education in 2015 to redesign the way it trains CMs before their first year of teaching. The redesigned training is based on the University of Washington's Teacher Education by Design program.[11] In partnership with the university, TFA had originally piloted its new training model at the Chicago regional institute in summer 2014. The pilot relied heavily on expert staff members from both the University of Washington and TFA, and thus was neither sustainable nor scalable (that is, suitable for expansion to other TFA summer training institutes). Yet, TFA national staff thought that the redesigned training showed promise, and they sought to implement components of it in a scalable, sustainable way. During the 2015-2016 school year, full-time TFA staff members who had worked on the Chicago pilot worked alongside other full-time TFA staff to prepare TFA alumni to offer the redesigned training to new CMs the following summer. The redesigned training was implemented at the national institute in Tulsa, which CMs from several regions attended in summer 2016. TFA offered its standard training at the other national institutes.

The goal of the redesigned training was to enhance the rigor and the relevance of the summer training that CMs received by incorporating the newest college- and career-readiness standards for students into its curricula and using training methods and practices meant specifically for adult learners. This report presents the findings from a study conducted by MDRC of TFA's efforts to implement the redesigned training at the national institute in Tulsa in summer 2016, and it explores whether the redesigned training shows promise for improving CMs' perceptions of the training, their commitment to teaching and educational equity, their use of instructional strategies aligned with college- and career-readiness standards, and their retention rates

[7]Unlike in the past, CMs now have considerable say over the regions to which they are assigned. This gives CMs the choice to live and work in an area where they will choose to stay, increasing the likelihood that CMs will continue to teach beyond their two-year commitment and in a community where they can more easily advance educational equity and serve as leaders in the community. The number of regions has changed from 53 in 2016 to 51 in 2018.

[8]TFA supports CMs in their endeavors to find teaching positions through its partnership with districts and by helping CMs create resumes and attend job fairs.

[9]CMs are responsible for teaching their own classes, but a teacher of record, who works for the school district (and not for TFA), is often present in the classroom and may help support CMs.

[10]Teach For America (n.d.)

[11]Teacher Education by Design (TEDD), is a project of the University of Washington. TEDD.org hosts an online, open-source library of curricular and pedagogical tools for teacher educators, many of which were adopted by TFA for the redesigned training.

during and beyond their two-year commitment (that is, the extent to which they remained in teaching).[12] The study's findings are based on a comparison of the outcomes of CMs in program regions, defined as regions whose CMs received the redesigned training at the national institute in Tulsa, with the outcomes of CMs in a group of matched comparison regions who were trained at one of the other national institutes and received "business as usual" summer training.[13]

This effort exemplifies TFA's commitment to learn continuously and to take risks in order to effectively innovate.[14] TFA leaders regarded the SEED grant as an opportunity to catalyze their efforts to improve the program and bring additional rigor to their assessment of efficacy, with the aim of drawing from the findings to strengthen network-wide training efforts and to share with other members of the broader teacher preparation community (within and beyond TFA). To that end, MDRC provided formative feedback to TFA about the implementation of the redesigned training and CMs' and staffs' reactions to the redesign as the study progressed. The end of the report summarizes how TFA has refined the training based on the results presented here and observed throughout the study.

Background

Several studies, including three large randomized controlled trials conducted by Mathematica Policy Research, have found that TFA teachers were as effective as, and in some subjects more effective than, non-TFA teachers with whom they were compared.[15] In the first study, from 2004, elementary level students were randomly assigned to either a TFA teacher or a non-TFA teacher at their schools. This study found impacts on elementary math equivalent to one additional month of school for the average student assigned to a TFA teacher compared with students with new or certified non-TFA teachers.[16]

The second study, in 2013, randomly assigned students in grades 6 to 12 to either a TFA math teacher or a non-TFA math teacher. This study found that students in math classes with TFA teachers gained the equivalent of an additional 2.6 months of school compared with students in

[12]This study also includes a descriptive analysis of value-added scores for a smaller subset of CMs in the study.

[13]In 2016, there were 10 regional institutes and 6 national institutes. The national institutes were located in Tulsa, Oklahoma; Philadelphia, Pennsylvania; Houston, Texas; Greater Delta (comprising Mississippi and Arkansas); Phoenix, Arizona; and Atlanta, Georgia.

[14]See Teach For America website: www.teachforamerica.org/what-we-do/values.

[15]In a randomized controlled trial, study participants are randomly assigned to either a program group that is eligible to participate in the intervention being tested or a control group that cannot participate, to ensure that all members of the sample are similar at the start of the study. Differences between the program and control groups' outcomes reflect the program's impacts. Statistically significant differences indicate that the impacts can be attributed with a high degree of confidence to the intervention rather than to chance. The analyses in this MDRC report are nonexperimental — that is, the study is not a randomized controlled trial, in which the outcomes for two similar groups (one that receives the intervention and one that does not) are compared to determine the effect. Thus, the study identifies whether outcomes changed over time, but not whether implementation *caused* any change in outcomes.

[16]Decker, Mayer, and Glazerman (2004).

math classes with non-TFA teachers, regardless of whether their non-TFA teacher was a new or a certified teacher.[17]

In the third study, in 2017, Mathematica evaluated the effectiveness of TFA elementary school teachers by randomly assigning students in prekindergarten (pre-K) through fifth grade to TFA or non-TFA teachers. On average, new TFA teachers had a positive, statistically significant impact on the reading achievement of lower elementary school (pre-K to second-grade) students, equivalent to an additional 1.3 months of instruction, compared with new non-TFA teachers in the same schools. TFA teachers also had a positive impact on first- and second-grade math achievement.[18] Overall, TFA teachers were equally as effective as new or certified comparison teachers.

Rationale for the Redesigned Training

Critical to the redesigned training was the recognition that teaching is a complex undertaking. Successful teaching involves more than delivering a curriculum. Teachers need to consider what they teach and how to teach it. For example, some considerations include the relevance of particular topics to their students, how to proceed with lessons that are not going as planned, how to ask questions that engage students in reasoning, how to redirect students not on task, how to respond to student errors in a way that helps them to understand the problem without discouraging them, and so on. In making these decisions, teachers also need to consider their own identities and the power dynamic between themselves and their students. Teachers need to consider these decisions at the lesson planning stage, but they also need to be ready to make these decisions in the moment. The goal of the redesigned training was to support CMs in making the types of decisions that skilled teachers make to prepare their students for success.[19]

The hypothetical examples developed by MDRC researchers described below illustrate the differences between a class taught by a teacher who simply delivers a lesson without any context-specific considerations and a teacher who engages in complex decision making throughout the lesson delivery. This extreme example is used to clarify the focus of the redesigned training rather than to show what teachers would do without this specific training. The activity in each of these examples from lower elementary school math classrooms is called "Choral Counting."[20] Choral Counting is an activity in which teachers lead students in counting aloud together by a given number (for example, counting to 10 by twos). To begin, teachers must decide on a number by which the students should count, whether to count forward or backward, and the numbers at which to start and end the count. As students call out each number in unison, the teacher records

[17]Clark et al. (2013).

[18]Clark et al. (2017). The p-value on the math impact was 0.054, so it was almost significant at a conventional level (0.05).

[19]The work of Gloria Ladson-Billings, Geneva Gay, and Adrienne Dixson were foundational to the development of the redesigned institute curriculum. See the KIPP Teachers Resource Guide website for their writings on culturally relevant pedagogy: https://trg.kipp.org/culturally-relevant-pedagogy.

[20]"Choral Counting" is an instructional activity developed by Teacher Education by Design.

the count on the board, pausing the count at strategic moments. The goal of this activity is not to practice counting by rote learning, but to engage children in reasoning, predicting outcomes, and justifying their answers. To do so, teachers record the count so that patterns within the numbers are readily noticeable, and pause during the count to ask questions, such as "What do you think will come next?" and "How do you know?"[21]

To conduct this activity, teachers must be able to engage their students. The chart below compares the level of student participation in Choral Counting in two hypothetical teachers' classrooms.

Teacher 1: Delivering a Lesson	**Teacher 2: Embracing Complex Decision Making**
[Twenty second-grade students are rolling around on the classroom rug.]	[Twenty second-grade students are rolling around on the classroom rug.]
Teacher: Okay everybody, we are going to count backward from 50 to 0 by fives. Begin! [Six of the students sit up and begin counting with the teacher, and the rest continue rolling around.]	**Teacher:** One, two, three, all eyes on me. Excellent! I see Kayla is looking at me, Carlos is looking at me. Great job, it looks like everybody is looking at me. Today we are going to do Choral Counting, again, but this time instead of counting from 0 to 50 by fives, we are going to count from 50 to 0 by fives.

Teacher 2 begins the lesson by getting the students' attention. Although all students may not respond to "One, two, three, all eyes on me," the teacher follows up by noticing and naming students who are following directions, which gets the attention of the rest of the students. Students in this class are ready to hear about the activity they will do next. Teacher 1, on the other hand, goes on with the lesson even though most of the students are still rolling around and likely not aware of what they are being asked to do.

Before beginning the count, Teacher 2 further prepares students for the activity in the following manner:

Teacher 2: Embracing Complex Decision Making
Teacher 2: Remember, yesterday we counted from 0 to 50, and today we are counting from 50 to 0. When I say "go," turn to your partner, and in a Level 2 voice, talk to your partner about what you think is different about today's count. Go! [Students turn to their partners and start sharing their ideas.] **Teacher 2:** One, two, three, eyes on me. Okay, give me a thumbs up if you know what is different about today's count. [Almost all of the students have a thumb up.]

[21]Teacher Education by Design (n.d.(a)).

Teacher 2: Brenda, I see you have your thumb up. What do you think is different about today's count?

Brenda: It's backward.

Teacher 2: Put a finger on your nose if you agree with Brenda.

[All students put a finger on their nose.]

Teacher 2 makes sure that all of the students are clear on what they will be doing by giving them context ("Remember yesterday?"), while continuing to give student-centered directions (directions that make clear to all learners the "what and how" of engaging in a task) and offering them the chance to think about what they will be doing (counting backward). Finally, Teacher 2 provides an opportunity for all students to feel included in the lesson ("turn to your partner" and "touch your nose") and sets expectations for how to participate ("in a Level 2 voice"). The counting has not yet begun, but the students are all primed to begin. Teacher 2 continues as follows:

Teacher 2: Embracing Complex Decision Making

Teacher: Carlos, how did you figure out that we would be counting backward?

Carlos: Well, 50 is bigger than 0, so I know that the numbers will be getting smaller.

Teacher: Malia, did you figure it out the same way or a different way than Carlos?

Malia: I was imagining the number line and thinking that we would have to go backward instead of forward.

Teacher: Wonderful. We just heard about two different strategies that our friends used to figure out what is different about today's count. By a show of hands, can anyone imagine a situation in which you might find it useful to count backward by fives? I see Jasmine's hand is raised. Jasmine, tell us about a situation where you might have to count backward by fives.

Jasmine: Like spending your money.

Teacher: Like spending your money. Hmm, can you say a little bit more about what you mean by that?

Jasmine: So, like, if you have 50 cents and every time you put 5 cents in the gumball machine you have to take away 5 cents of your money.

Teacher: Right, counting backward by fives can help you figure out how much money you have left. Okay, I think we are just about ready to start. I want everyone to count together in a Level 3 voice. Ben, would you please count to five in a Level 3 voice to show us how we should sound?

Ben: 1, 2, 3, 4, 5.

Teacher: That was the perfect level, and it was also the right pace. Remember, this is not a race. We will count at a medium pace. Please don't say the next number until you see that I have written the number down that you already said. When I say "pause," it means you should not say the next number out loud. Ready? Pick up your whiteboards and markers.

The students in Teacher 2's class are now invested in the activity. They all know what they are being asked to do, how to do it, and why they should care about doing the activity. They also know that there may be different ways to solve a problem. Teacher 2 is ready to begin the count. What happens during the count in each of the two classrooms is shown below.

Teacher 1: Delivering a Lesson	Teacher 2: Embracing Complex Decision Making
Six students: 50, 45, 40, 35.	**All students:** 50, 45, 40, 35,
Teacher: Pause! Brenda, what number do you think will come next?	**Teacher:** Pause. On your mini-whiteboard, please write down the number that you think comes next.
Brenda: Huh?	[Most students write the number 30, but a few write 25 and one student writes 20.]
Teacher: If you had been paying attention, you would know the answer. You lose one point on the behavior chart. Jerry?	**Teacher:** Okay, I see that a lot of you wrote 30. Celia, you wrote 30. Would you tell us why you think the next number counting backward by fives is 30?
Jerry: 25.	**Celia:** I see a pattern, 0-5-0-5, so I know the next one must be a number that ends in 0.
Teacher: Anyone else?	**Teacher:** Jerry [who had written 25], do you agree with what Celia said about the 0-5-0-5 pattern?
Carlos: 30.	**Jerry:** Yes, after 35, the next number must end with a zero. I was thinking too fast.
Teacher: Excellent! How did you know that?	**Jasmine:** I've done that, too.
Carlos: Because you said to count backward by fives.	**Teacher:** All right, when I say 30, let's continue our countdown. 30….
Teacher: Let's keep going.	
Four students, including Carlos: 25, 20, 15, 10, 5, 0.	

By having the students write their answers on their whiteboards, Teacher 2 actively engages all the students in the activity. It also gives the teacher an opportunity to assess which students might need more help, based on the students' responses. The students in Teacher 2's class also have the opportunity to think more deeply about the problem by hearing from students who figured out the right answer how they thought through it, and they are given a chance to change their answer if their peers make a convincing argument. While Teacher 1 seems to have a classroom management system in place ("You lose one point on the behavior chart"), the directions the students receive do not adequately prepare them to act in the way the teacher thinks they should act. They are unsure of what they should be doing, how they should be doing it, or why they should be doing it. The emphasis seems to be on someone getting the right answer rather than on everybody understanding the problem.

Teacher 2, on the other hand, makes continuous and purposeful decisions based on students' academic, socio-emotional, and physical needs. Unlike Teacher 1's students, Teacher 2's students are engaged in the activity and they appear to be learning. Routines are in place to get the students' attention, and once students are paying attention, they are given clear examples of how to participate in the activity. Teacher 2 provides many opportunities for individual students to actively participate in the activity by using their voices, writing, or gesturing. By giving students a variety of ways to express themselves, Teacher 2 makes it possible to participate for students who might be shy, hesitant to reply for fear of getting the wrong answer, or new to speaking English. There is also evidence in Teacher 2's class that students understand that even if they get the wrong answer, it is okay. For example, in Teacher 2's class, Jerry ends up with the wrong answer (25) but is given a chance to see where he might have erred and to change his answer. In contrast, Teacher 1 summarily dismisses Jerry's answer with her response ("Anyone else?"). Also significant is the way Teacher 2 signals to students that what they have to say is important — for instance, by asking Jasmine to say more about what she means by "spending your money" and by inviting students to share different solutions to the same problem.

TFA has long sought to develop teachers who make the same types of thoughtful decisions as Teacher 2; the redesigned training was created to introduce new methods of teacher education that would accelerate the growth of new teachers to realize this type of skill while teaching rigorous and engaging culturally relevant content.[22] The next section provides a description of this training.

Overview of the Redesigned Training

Research indicates that learning accompanied by practice has the most impact on individuals' performance.[23] Accordingly, the redesigned training was developed to address the complexities of teaching by contextualizing the skills that CMs were learning and by stressing reflection and decision making. The redesigned training provided CMs the opportunity to try out decisions through public practice with their peers and trainers and to consider the potential impact of those decisions on students when made in the classroom. The supportive environment allowed them to weigh the benefits and drawbacks of their decisions in the classroom alongside other CMs. In this vein, the redesigned training aimed not simply to teach discrete skills (such as how to get students' attention) isolated from the complexities of teaching, but rather to build skills while strengthening CMs' judgment as teachers through content- and context-specific practice. Given the emphasis on practicing teaching, the redesigned training called for strong lesson plans to be provided to

[22]According to Ladson-Billings (1994), culturally relevant education (1) supports academic success by setting high expectations for students and providing ample opportunities for them to succeed, (2) embraces cultural competence, including a curriculum that builds on students' prior knowledge and cultural experience, and (3) promotes critical consciousness by providing students with the tools to critique and challenge institutions that perpetuate inequality. Preparing CMs to teach in culturally relevant ways is a major goal of the training that CMs receive from TFA. However, it is not the focus of this study.

[23]Ericsson, Krampe and Tesch-Romer (1993); Grossman et al. (2009).

CMs throughout most of their training so they could concentrate on the *how* and not lose themselves in the *what* of teaching. This provided an opportunity for CMs to internalize effective teaching techniques and strategies while learning rigorous content to teach their students.

Instructional Activities and Core Practices

Positioning CMs to develop their students' skills and knowledge and instill in them the competencies and dispositions needed for success, such as self-efficacy, persistence, and a sense of belonging and of equity, in college and the workforce is central to TFA's vision. With this goal in mind, the redesigned program identified two key areas of pedagogical development for CMs: *instructional activities* and *core practices*. Instructional activities were developed at the University of Washington. They are content-based learning routines that are simple enough for new teachers to implement while teaching to rigorous academic standards.[24]

Instructional activities have a predictable structure that specifies how teachers, their students, and the content interact to offer students opportunities to engage in complex learning. Choral Counting is an example of an instructional activity. The curricula that CMs receiving the redesigned training taught their summer school students were built around several content- and grade-specific instructional activities.

Core practices, developed by the Core Practices Consortium, are big ideas that successful teachers keep in mind as they teach in order to make the instructional judgments that will set up their students for success.[25] Teacher 2's decisions, described above, were grounded in some of the core practices. The four core practices adopted by TFA for the redesigned training were *creating and maintaining a productive learning environment, positioning students as competent sensemakers, teaching toward an instructional goal,* and *teaching with society in mind*.

The core practices are what make the instructional activities within the redesigned training more than just subject matter lessons.[26] Whereas an instructional activity provides rigorous content, the core practices consist of strategies and routines that teachers can use to set up the learning environment, engage students in the content, and help students make sense of the content. For example, in the case of Choral Counting, Teacher 2 integrated core practices such as giving student-centered directions and creating opportunities for students to share their voices.

The four core practices, developed by the Core Practices Consortium and prioritized by TFA for the redesigned training, are listed below along with some strategies and routines that teachers can use in support of those practices.

1. **Creating and maintaining a productive learning environment:** Teachers build relationships with students, give directions that are student centered and easy for students to understand, support disengaged learners, have positive expectations of

[24]Teacher Education by Design (n.d.(b)).

[25]The core practices stem from a research-based collaboration between many teacher preparation programs who make up the Core Practice Consortium. See www.corepracticeconsortium.com/core-practice.

[26]Core Practice Consortium (n.d.).

students, recognize and reinforce positive student behaviors, build excitement, and so on.

2. **Positioning students as competent sensemakers:** Teachers nurture a culture of respect and safety for academic risk-taking; create opportunities for students to share their voices and make personal connections to their schoolwork; elicit student thinking to inspire new ideas; express curiosity in all students; react in a way that promotes student inquiry, even when students' verbal contributions are off topic or miss the mark; and so on.

3. **Teaching toward an instructional goal:** Teachers establish efficient procedures in their classrooms; make their presence in front of classroom engaging by clearly identifying what they are teaching and why it matters; create opportunities for students to work in pairs, groups, as a whole class, and individually, and to express themselves in ways that are aligned with the instructional goal; expose students to content that is deeply engaging; and so on.

4. **Teaching with society in mind:** Teachers tailor classroom systems to meet students' needs in ways that push against larger social and systemic oppression; elevate considerations of students' communities, families, interests, and identities in their decision making around curricula and classroom activities; use inclusive language and challenge socially offensive language; promote equitable participation in the classroom; understand the effect that their own identity, power, and privileges may have on their students; and so on.

The purpose of the redesigned training, in part, was to expose CMs to this prioritized set of core practices and help them integrate these practices into their teaching. The instructional activities were intended to provide rigorous, content-specific routines that CMs could use while executing the core practices. Based on research indicating that pedagogical practices are learned most effectively when applied in the classroom, TFA included instructional activities in the curricula for both training CMs how to teach and teaching summer school students. The student curricula, used by CMs in their summer school classrooms and designed by full-time TFA staff, were aligned with rigorous content standards (for example, Common Core State Standards and Next Generation Science Standards) and grounded in culturally relevant pedagogy.[27]

The Learning Cycle

To ensure CMs absorbed the summer training as much as possible, the redesigned training incorporated several "andragogies," or training methods and practices meant specifically for adult learners. These andragogies are part of the Learning Cycle, a method for training new teachers such as CMs that was informed by research and developed at the University of Washington's

[27]TFA repackaged selected standards into something called the Core Components of Instruction.

Teacher Education by Design program.[28] The Learning Cycle consists of the following four activities, or quadrants: (1) experienced teachers model a content-specific lesson; (2) new teachers rehearse the lesson while their peers and trainers observe them; (3) new teachers teach the lesson to a group of students; and (4) the new teachers and their trainers reflect upon the lesson as a group. TFA's use of the whole cycle takes about two or three days, and then begins again with a different lesson being modeled.

Below is a detailed description of the quadrants in the Learning Cycle as used by TFA in the redesigned training.

- **Quadrant 1: Introduce.** New teachers observe strong teaching, either by playing the role of students while their lead instructor (the CMs' main trainer) assumes the role of a classroom teacher and conducts a lesson, or by watching a video in which an experienced teacher is teaching. New teachers observe the lesson twice: once all the way through (modeling), and then in chunks in which actions are paused at key points in order to discuss the motivation behind different teacher strategies and routines (modeling to decompose).

- **Quadrant 2: Prepare.** New teachers collaboratively rehearse teaching the lesson they have seen modeled, and their lead instructors, coaches (who support the CMs in small groups a few times a week and help the lead instructor during Learning Cycle sessions), and other CMs provide constructive feedback to help them strengthen their teaching skills. This could be done by rehearsing the whole lesson (rehearsal) or by rehearsing a very small segment of the lesson (microteaching) — for example, introducing the lesson.

- **Quadrant 3: Enact.** New teachers teach the lesson they have rehearsed to their summer school students. The lessons are recorded for use in Quadrant 4.

- **Quadrant 4: Analyze.** New teachers consider the impact of their actions or decisions and practice on student learning by watching video recordings of themselves teaching lessons or by reviewing their students' work along with constructive feedback from their trainers and peers. In this way, new teachers see the connection between their instruction and their students' results.

In the redesigned training, CMs were taught how to teach according to the core practices through the use of instructional activities, such as Choral Counting, via the Learning Cycle. The Learning Cycle allowed CMs to experience what it was like to teach, make mistakes, and learn from those mistakes through public practice. Beginning with Quadrant 1, CMs observed an instructional activity modeled by an experienced teacher, whether on video or by their lead instructor. The instructional activity was then viewed a second time, with the lead instructor periodically pausing to allow discussion about different decisions the experienced teacher might or could have made while teaching the instructional activity. For instance, if a video recording of Teacher 2

[28]See the University of Washington's Teacher Education by Design (TEDD) program website: tedd.org

from the earlier example were to have been shown, the lead instructor might have stopped the video when Teacher 2 got the students' attention for the first time. In this instance, the lead instructor might have explained that Teacher 2 used a strategy informed by the core practices when the teacher said "One, two, three, all eyes on me," as a signal to get some students' attention, and when the teacher followed up by noticing and naming students who had followed the directions to attract the attention of the rest of the students.

In Quadrant 2, CMs planned for and publicly rehearsed teaching an instructional activity, or microtaught part of an instructional activity.[29] CMs practiced using the attention-getting signals modeled and discussed in Quadrant 1, for instance, thus giving CMs the opportunity to integrate core practices into their teaching. Quadrant 2 happened either during a Learning Cycle session or in small groups with a coach. The lead instructor, coaches, and peers provided feedback.

In Quadrant 3, CMs taught an instructional activity using core practices to their summer school students. Later the same day or the following day, CMs returned to the Learning Cycle session for Quadrant 4. In it, the group collectively reflected on each CM's teaching experience in Quadrant 3, paying close attention to how well each one integrated the core practices in the lesson he or she had just given, either by watching and analyzing a video of the teaching or by reviewing students' work.

Figure 1.1 summarizes how the Learning Cycle was used to train CMs on how to implement the instructional activities and integrate the core practices into their teaching strategies and routines.

Figure 1.2 presents the redesigned training's logic model and its key components and corresponding expected outcomes.[30] The logic model posits that the training and coaching received by the CMs in Tulsa will enable them to effectively teach rigorous lessons grounded in the core practices. The training is expected to help the CMs start the school year feeling well prepared and confident in their teaching skills. The CMs' confidence and strong instructional abilities are expected to strengthen their commitment to teaching and to TFA's mission, and to make them more likely to remain in teaching for the duration of the two-year commitment.

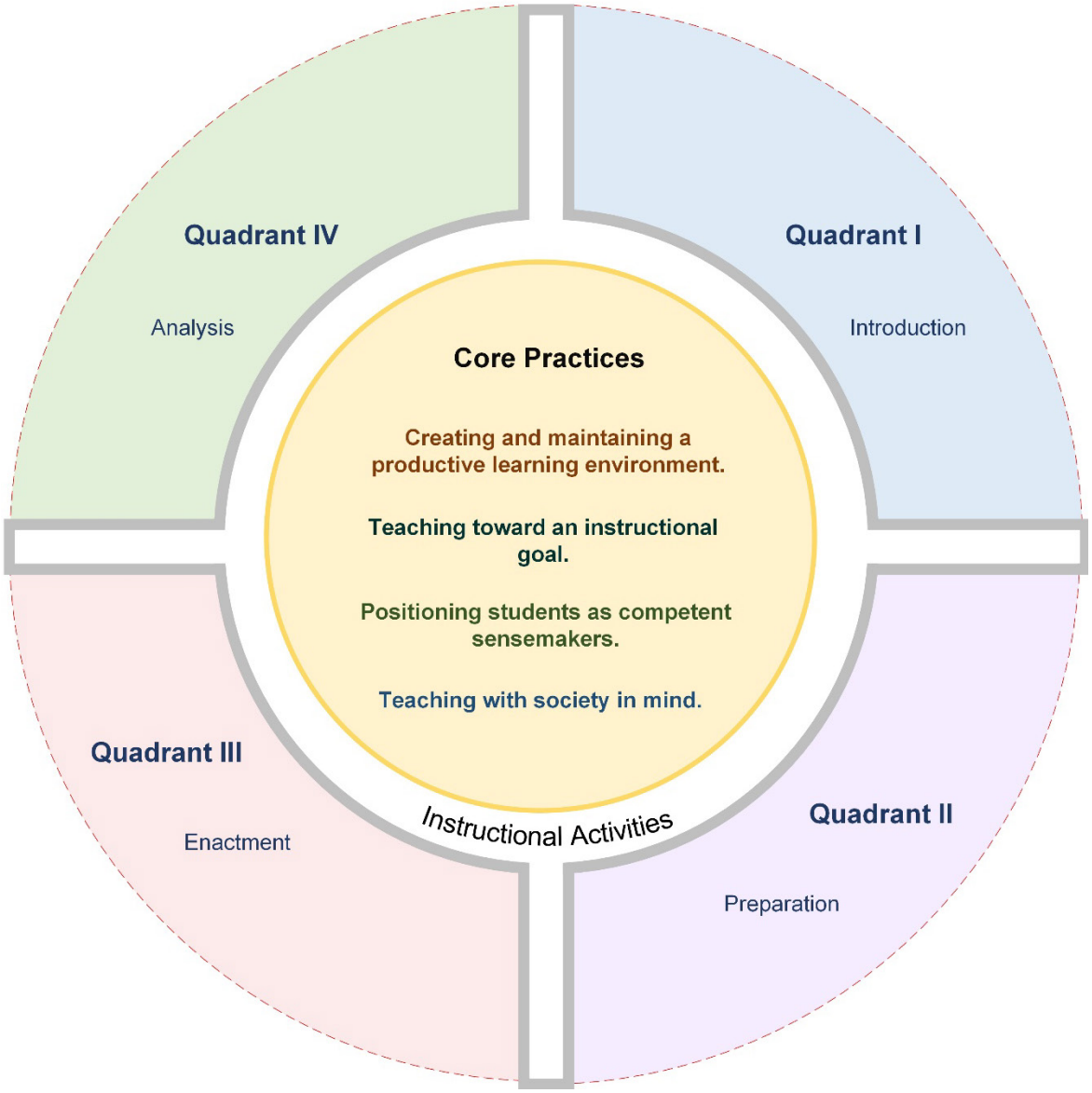
The redesigned training was implemented at the national institute in Tulsa for five weeks in summer 2016. During their first week in Tulsa, CMs attended Learning Cycle sessions,

[29] Microteaching is a technique used by teachers in which they practice a specified teaching behavior that they want to master. For instance, a teacher might practice (microteach) giving students clear directions on how to quietly get in line before recess.

[30] The study team created Figures 1.1 and 1.2 based on an understanding of the Learning Cycle and the logic behind the redesigned training.

Figure 1.1

The Learning Cycle and Core Practices

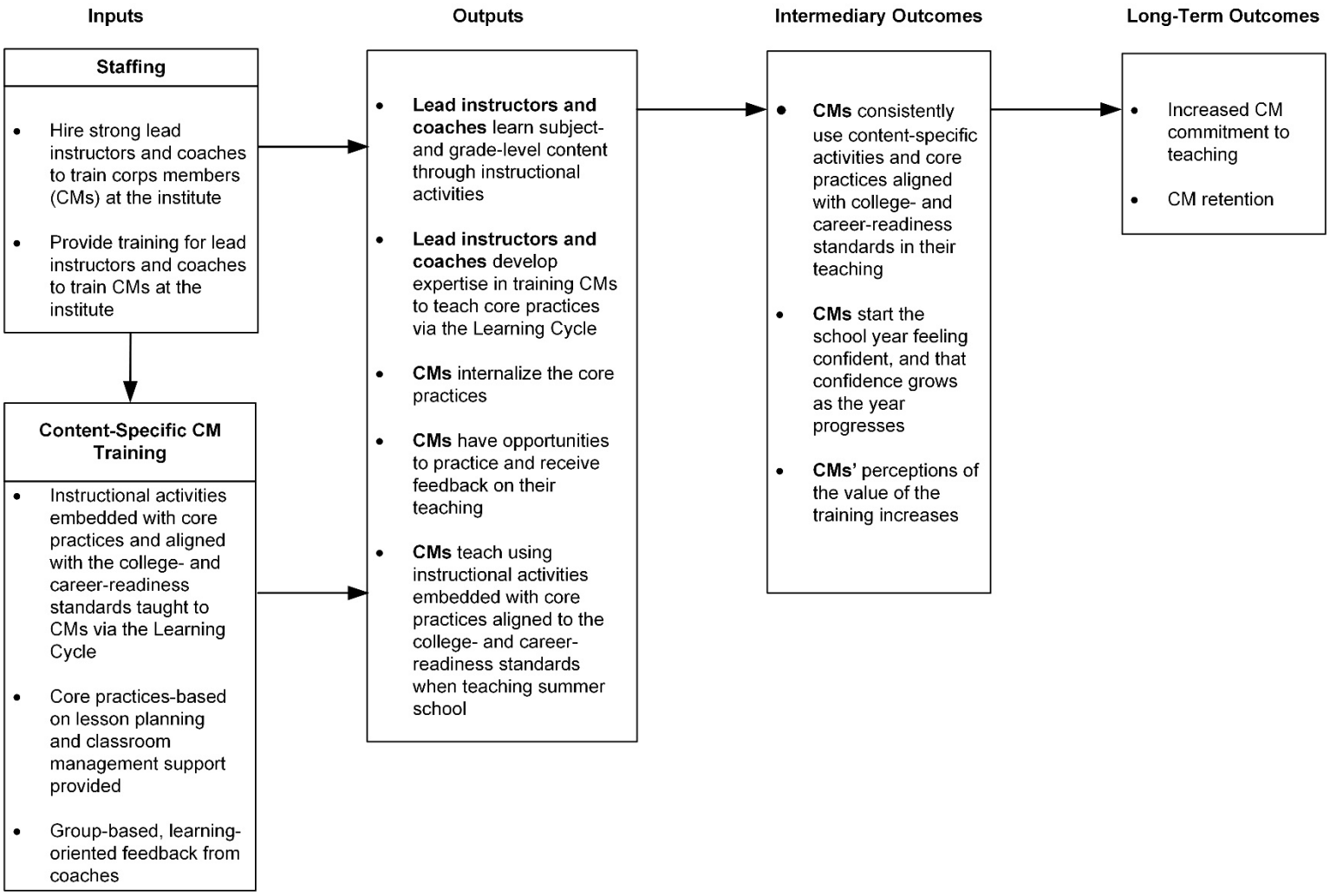


attended other professional development experiences, and met with their coaches.[31] Beginning in their second week and continuing throughout the rest of their training, CMs started each day at an assigned summer school teaching site. In the mornings, CMs taught summer school students in their assigned content area and grade level and attended either a social justice seminar

[31]CMs also attended other types of sessions, such as social justice seminars, in which they explored issues related to diversity, equity, and inclusion and their impact on student learning. However, the MDRC evaluation focused only on the training related to the instructional activities, core practices, and the Learning Cycle and thus these other aspects of the summer training are not a part of this report.

Figure 1.2

Logic Model for Teach For America's Redesigned National Summer Training Institute



or a meeting with their coach, depending on the day. When meeting with coaches, CMs worked collaboratively with a small group of other CMs teaching the same content area and grade level, rehearsed their summer school lessons, and received feedback from the coaches. Their afternoons were spent attending Learning Cycle sessions.

Differences Between the Redesigned Training and the Business-as-Usual Training

The redesigned training that was piloted at the national institute in Tulsa differed from the business-as-usual training at the comparison sites in several ways (shown in Table 1.1). Whereas the CMs in Tulsa experienced teaching in all its complexity by practicing how to make decisions that would lead to student success and to teach rigorous, standards-based content embedded with the core practices, CMs attending the other national institutes received training on discrete, content-neutral topics, such as classroom management and lesson planning, largely through lecture format. CMs in the comparison institutes therefore had fewer systematic opportunities to practice their teaching in front of their peers.

Table 1.1

TFA's Redesigned Teacher Training and the Traditional Training

Aspect of the Training	Redesigned Model (national institute in Tulsa)	Traditional Model (other national institutes)
Organization of the training topics	The interconnectivity of teaching skills and judgment is addressed using the Learning Cycle, in which instructional activities are the vehicle through which teachers can enact the core practices.	The elements of teaching are taught discretely.
Practice	Rehearsals and microteaching are performed publicly.	There is no expectation to practice publicly.
Content-area focus	CMs are trained through content.	Training is content neutral.
Grouping of corps members (CMs)	CMs are grouped by subject area and grade level for training.	CMs are sometimes grouped by subject area and sometimes not.
Staff responsibilities	CMs receive the redesigned training from lead instructors and the social justice seminars by the social justice seminar leaders.	CMs receive teacher training and diversity, equity, and inclusiveness training from lead instructors as they received pedagogical training.
Lesson plan creation	Strong lesson plans are provided to CMs throughout most of their training, so they can concentrate on the *how* of teaching and not get lost in the *what.*	CMs create their own lesson plans based on end-of-course assessments and objectives provided by TFA.
Coaching dosage and format	150 minutes per week of group coaching by subject area and grade.	30 to 120 minutes a week of one-on-one coaching.
Amount of summer school teaching	30 to 36 hours.	At least 20 hours.

There were other differences as well. Because the training in Tulsa emphasized *how* to teach, the CMs in Tulsa were provided with strong lesson plans so that they could learn *how* to deliver lessons instead of spending time developing those lessons. The CMs in Tulsa had to write their own lesson plans only in the last week of training at the institute. The CMs at the other national institutes, on the other hand, had to create all of their own daily lesson plans using the discrete skills they had learned and drawing on a unit plan and assessment provided by TFA. In addition, since the CMs in Tulsa were trained with set content, the CMs in Tulsa were grouped by content area and grade level. The CMs at the other institutes were not systematically grouped this way and the training at these institutes was content-neutral. As a result, and because CMs in the comparison institutes were not given an opportunity to practice teaching publicly, most of the coaching support they received was one-on-one.

Planning for and implementing the redesigned training within a 10-month period was a huge undertaking. The training itself looked nothing like the type of training that TFA had been offering to CMs at regional and national institutes for more than 20 years. In addition to providing the business-as-usual training to the staff members at the other national institutes during summer 2016, TFA had to provide newly developed training to the staff members responsible for delivering the redesigned training at the national institute in Tulsa. This required the heavy involvement of full-time TFA staff members who had attended the regional pilot of the redesigned training in 2014 in Chicago, along with other full-time staff. Typically, full-time staff members are not as involved in the national institute training.

Since the redesigned training was substantially different from the business-as-usual training, considerable time was needed to get everybody involved up to speed. TFA convened conferences to train the national institute staff who would be implementing the redesigned training. The redesigned training was built around new student curricula that included the instructional activities embedded with the core practices, and TFA had to design these new activities. Because the instructional activities differ by grade level and content, TFA had to develop several new curricula. In addition, at the national institute in Tulsa, both a lead instructor and a social justice seminar leader were expected to train CMs, whereas at the other national institutes CMs received their teacher training and social justice training from the same person. Consequently, compared with the business-as-usual model, the redesigned training called for much more preparation time, additional institute staff roles, and new full-time staff "headcount" and capabilities to support and provide training for these new institute roles, which had not been previously established in the organization.

The redesigned training also changed the way the institute was organized in significant ways and required operational staff members to rethink the logistics and find the resources to implement it. For example, in the business-as-usual model, CMs did all of their summer school teaching and got their training at their assigned school, whereas CMs in Tulsa taught at their assigned school site in the morning and received their teacher training at a TFA hub in the afternoon so that all CMs teaching the same content and grade level could undergo the training together. This meant that TFA had to plan for transporting the CMs in Tulsa to the hub every afternoon.

Contents of This Report

The remainder of this report is organized as follows:

- Chapter 2 describes the study design that was used to evaluate the implementation of the redesigned training and its potential effects on teacher outcomes, the data sources, and the samples of CMs used in the analyses.
- Chapter 3 focuses on the experiences of CMs in Tulsa during their training. It examines the implementation of the redesigned training based on observations of Learning Cycle sessions, observations of CMs teaching their summer school students, and focus groups with CMs and the lead instructors and coaches. The chapter also looks at how the implementation of the redesigned training differed from the business-as-usual training.
- Chapter 4 focuses on the experiences during the school year of CMs who were at the institute in Tulsa, compared with that of CMs who were trained at the other national institutes. It examines CMs' perceptions of the overall value of their summer training, their use of instructional strategies aligned with college- and career-readiness standards, their use of the core teaching practices, and their reflections on their first year of teaching.
- Chapter 5 considers whether the redesigned training shows promise for improving CMs' longer-term outcomes. The outcomes examined include CMs' commitment to teaching and educational equity and their retention rates in the TFA program.
- Chapter 6 concludes with the lessons that TFA has learned since it implemented the redesigned training in summer 2016.

Chapter 2

Study Design

This chapter describes the analytical approach used to evaluate the implementation of the redesigned summer training that Teach For America (TFA) piloted at its national institute in Tulsa, as well as the model's promise for improving the outcomes of the first cohort of corps members (CMs) to receive it. The first section presents the data sources for the study. The second section describes the two quasi-experimental approaches used to explore whether the redesigned training shows promise for improving teacher outcomes. Both study designs compare the outcomes of CMs in the TFA regions trained at the national institute in Tulsa (program regions) with the outcomes of CMs in a group of matched comparison regions who were trained at one of the other national institutes and therefore received TFA's traditional summer training (comparison regions). The third section describes the characteristics of the CMs in the program and comparison regions. The key takeaways from this chapter are the following:

- Several types of qualitative data were collected to assess the implementation of the redesigned training to better understand how it differs from the training offered in the other national institutes. The study team made site visits to the institute in Tulsa to observe the training and summer school teaching. The team also conducted focus groups with CMs, their trainers, and coaches at the institute in Tulsa and at the two comparison institutes. To learn about CMs' experience in the classroom during the school year, the study team sent open-ended logs to CMs in the program and comparison regions every two weeks to capture their use of the core practices. The study team also conducted follow-up phone interviews with CMs at the end of their first year of teaching.

- The key teacher outcomes examined in this study include CMs' perceptions of the summer training, their commitment to teaching and educational equity, their use of instructional strategies aligned with college- and career-readiness standards, and their retention in the TFA program. The study team measured these outcomes using TFA's administrative records and teacher surveys that are regularly administered by TFA. The study team also collected closed-ended teacher logs for the purposes of this study.

- Because the study designs used in this evaluation are quasi-experimental, the findings presented in this report cannot definitively be interpreted as the *causal* effect of the redesigned summer training, and for this reason they are referred to as *estimates* of the effect. However, for some CM outcomes, there is a relatively higher level of confidence that the findings might represent the causal effect of the new training. The findings that can most likely be attributed to the new training are those pertaining to CMs' perception of the training, their commitment to teaching, and their retention rates. In contrast, findings related to

CMs' use of instructional strategies aligned with college- and career-readiness standards should be interpreted with greater caution.

Measures and Data Sources

One of the goals of this study was to look at how the redesigned summer training was implemented at the Tulsa national institute and how it compared with the training offered by TFA in the other national summer institutes. In summer 2016, the study team visited the institute in Tulsa as well as two of the five summer institutes offering the traditional training. During these visits, the following types of data were collected.[1] (See Table 2.1 for a summary.)

- **Observations:** During the visits to the national institute in Tulsa, the study team observed sessions delivered by the lead instructors and coaches to assess the fidelity of implementation of the Learning Cycle and how core practices were embedded in the training. The team also observed CMs in Tulsa as they taught their summer school classes to identify their use of the core practices.

- **Focus groups:** During the visits to the institute in Tulsa as well as to the comparison institutes, the study team conducted focus groups with CMs to learn more about the key ideas, practices, or experiences that CMs gained from their training and how prepared they felt to teach. To learn about CMs' readiness to teach from the perspective of their trainers, as well as the trainers' own readiness to train CMs, the study team also conducted focus groups with English language arts (ELA), mathematics, and elementary school teacher educators (lead instructors and coaches).

Another main goal of the study was to learn more about CMs' first-year teaching experience and their use of the core practices, how these were shaped by their summer training, and to what extent they differed for CMs trained in Tulsa and CMs trained in the other institutes. The study team collected the following data during the school year for this purpose:

- **Open-ended teacher logs:** Every two weeks during their first year of teaching, the study team sent an open-ended teacher log to CMs in the 2016 cohort. The purpose of the open-ended logs was to capture CMs' use of the core practices, which were a focus of the redesigned training. Each open-ended log focused on a particular core practice, asking CMs to describe the instructional practices they used in the last one to two weeks in relation to that core practice. The team sent the logs to CMs in the program and comparison regions who were teaching ELA, mathematics, or general education, and who consented to

[1] As described in more detail in the next section, the study design used in the evaluation included 13 comparison regions. The study team visited the two national institutes that trained the greatest number of CMs from the comparison regions (the institutes in Houston and Philadelphia).

Table 2.1

Qualitative Data Collected During the Summer and School Year

Data Source	Timing	Domains	Sample	Institute/Regions
Observations of lead instructor and coaching sessions	During institute (beginning, middle)	Fidelity of implementation of the Learning Cycle; how core practices were embedded in the model	A sample of lead instructors and coaching sessions	Institute in Tulsa
Observations of summer school teaching	During the middle of institute	Use of core practices	Sample of summer classes taught by CMs	Institute in Tulsa
Focus groups with lead instructors, coaches, and CMs	Last week of summer institute 2016	Whether and how CMs were prepared for their first year of teaching	Sample of teacher educators and CMs teaching ELA/math/ general education who volunteered	Tulsa and two comparison institutes[a]
Follow-up interviews with CMs	Spring 2017	Whether and how CMs were prepared for first and second years of teaching	Sample of ELA/math/general education CMs who volunteered	Tulsa and two comparison institutes[a]
Open-ended teacher logs	Every two weeks during the 2016-17 school year	CMs' use of TFA core practices, which were a focus of the redesigned training	CMs teaching ELA/math/general education in the 2016 cohort who volunteered (elementary and secondary school)	Program and comparison regions

NOTES:
[a]The two comparison institutes are the two national institutes that trained the largest number of CMs from the comparison regions in summer 2016.

participate in this data collection.[2] Although each item in the open-ended log was designed based on TFA's Core Practices Evidence Guides, the log items did not explicitly mention the core practices, so that they could also be used with CMs in the comparison group.[3]

- **Phone interviews during the school year:** In the spring of CMs' first year of teaching, the study team conducted follow-up phone interviews with a subsample of CMs who participated in the focus groups in Tulsa and in the com-

[2]About 58 percent of the CMs who were eligible for the logs (that is, they taught ELA, mathematics, or general education) consented to participate in the logs. CMs who consented and those who did not consent had similar application scores and taught in schools with similar characteristics. However, CMs who consented to the logs were less likely than nonconsenting CMs to be a person of color (45 percent of consenting CMs, compared with 60 percent of nonconsenting CMs) and less likely to have received a Pell grant (45 percent, compared with 55 percent).

[3]TFA's Core Practices Evidence Guides lists the specific actions expected of the CMs when they teach with the core practices in mind.

parison institutes, to learn how prepared CMs felt when they first started teaching, what type of support they received from TFA during the school year, and how ready they felt going into their second year of teaching.[4]

The final key goal of the study was to look at the effect of the redesigned training on CMs' outcomes. The study examined outcomes in several domains: CMs' perceptions of the summer training, their commitment to teaching, their use of instructional strategies aligned with college- and career-readiness standards, and their retention in the TFA program. The study measured these outcomes at different points in time from the end of the summer training to the fall of CMs' second year of teaching. The following data sources were used for this purpose:[5]

- **Teacher surveys:** TFA regularly administers surveys to its CMs to learn more about their habits and mindsets, their experience as teachers and as TFA CMs, and their perceptions of the support provided by TFA, with the aim of improving the program. The surveys are administered online three times during each school year (eight weeks into the school year, midyear, and end of year). The study team used a subset of items from these surveys to examine CMs' perceptions of the value of their training and their commitment to teaching. Each of these items asked CMs to rate their response on a 7-point agreement scale.[6]

- **Closed-ended teacher logs:** In addition to the open-ended logs, the study team sent closed-ended logs to CMs in the 2016 cohort in the program and comparison regions, on alternate weeks (every two weeks). Closed-ended teacher logs have been effectively used in other studies to capture teachers' practices at more frequent intervals than would be feasible with classroom observations.[7] The closed-ended logs asked CMs to report on how often they used different instructional strategies aligned with college- and career-readiness standards.[8] When reporting, CMs were asked to reflect on their interactions with a quasi-randomly selected focal student.[9] The items in the closed-ended log differ across grade levels (lower elementary versus upper elementary or secondary) and content area (ELA versus mathematics). Each teacher was asked to complete 15 closed-ended logs during the school year. Unlike the open-ended logs,

[4]Appendix A provides further information on the data collected during the site visits and the phone interviews.

[5]The study team also collected data on CMs' value-added scores. However, these data were only available for CMs teaching in Florida. See Appendix G for a time series analysis of these data.

[6]The response categories were the following: 1 = strongly disagree, 2 = disagree, 3 = somewhat disagree, 4 = neutral, 5 = somewhat agree, 6 = agree, 7 = strongly agree. See Appendix B for further information about the surveys.

[7]Rowan, Camburn, and Correnti (2004); Rowan, Correnti, Miller, and Camburn (2009).

[8]See Appendix C for further information about the closed-ended teacher logs and how the items map onto different career- and college-readiness standards. Internal to TFA, these standards are called the "core components of instruction."

[9]At the start of each log, CMs were given a randomly selected letter of the alphabet and asked to report on their instructional strategies with a focal student in mind whose name starts with that letter (or the closest letter).

which were more closely aligned with the core practices, the items in the closed-ended logs reflected practices that are commonly recognized as effective instruction. The math logs were designed based on the Common Core State Standards for mathematics, and the ELA logs were based on TFA's core components of instruction, which are based on the Common Core State Standards.

- **TFA administrative records:** TFA maintains records of the CMs' characteristics, information about their teaching placements, and whether they are still teaching with TFA at the beginning and end of each placement year. The study team used these records to describe the characteristics of CMs and their placements, and to measure their retention in the TFA program.

These data sources were used to create measures of CM outcomes across several domains. The confirmatory outcomes that were used to gauge the effects of the redesigned training are listed in Table 2.2. The team chose these outcomes because of their alignment with the goals of the redesigned summer training.[10]

Finally, several publicly available data sets were used to describe the characteristics of the schools in which CMs were placed. These sets include the Common Core of Data (CCD) and the Office of Civil Rights Data Collection (OCRDC). The most recent available data sets were used (school year 2015-2016 for the CCD; school year 2013-2014 for the OCRDC).

Study Designs

One of the goals of this study was to examine the effect of the redesigned training on the short-term outcomes of the first cohort of CMs to participate in the new training. As explained in Chapter 1, TFA piloted the redesigned training at the national institute in Tulsa in summer 2016, while offering the traditional summer training at the other five national institutes.[11] The specific institute that CMs attended depended on their regional assignment. When CMs are recruited into TFA,

[10]When conducting an impact analysis and making judgments about statistical significance, it is important to recognize the potential problems associated with conducting multiple hypothesis tests across many different outcomes. Specifically, when hypothesis tests are conducted for estimated effects on several different outcomes, it increases the risk of concluding that a given estimated effect is statistically significant when, in fact, it is not. This is called a Type I error or false positive. For example, when a 10 percent statistical significance level is used, one would expect to see a false positive for every 10 hypothesis tests conducted. (See Box 4.1 for additional information on statistical significance.) To reduce this risk, the study team limited the number of outcomes and selected one "confirmatory" outcome per domain. The confirmatory outcomes are used to make conclusions about the redesigned training's effectiveness.

[11]The other five national institutes were located in Philadelphia, Pennsylvania; Houston, Texas; Greater Delta (comprising Mississippi and Arkansas); Phoenix, Arizona; and Atlanta, Georgia.

Table 2.2

Confirmatory Corps Member (CM) Outcomes, by Domain

Domain	Outcome	Scale	Data Source	Timing	Population
Perceptions of the summer training	Value of the summer training	7-point agreement scale (1 item)[a]	TFA CM surveys	End of institute; first year of teaching (fall, middle, end); fall of second year	All CMs in the 2012-2016 cohorts
Commitment to the teaching profession	Commitment to teaching and equity	7-point agreement scale (8 items)[b]	TFA CM surveys	End of institute; first year of teaching (fall, middle, end); fall of second year	All CMs in the 2012-2016 cohorts
Use of instructional strategies aligned with college- and career-readiness standards (cultural responsiveness)	Text used for the lesson featured an author, characters, or a community with a similar background to the focal student	Binary indicator (1/0)[c]	Teacher logs	Biweekly closed-ended logs (maximum of 15 per teacher)	CMs in the 2016 cohort who were teaching ELA/general education and who consented to the logs
Use of instructional strategies aligned with college and career readiness standards (English Language Arts)	Student was asked to cite text evidence in their writing	Binary indicator (1/0)[d]	Teacher logs	Biweekly closed-ended logs (maximum of 15 per teacher)	Same as above (but for CMs teaching at the upper elementary and secondary levels only)
Use of instructional strategies aligned with college and career readiness standards (math)	Composite of self-reported instructional practices (math)	z-score[e]	Teacher logs	Biweekly closed-ended logs (maximum of 15 per teacher)	CMs in the 2016 cohort who were teaching math/general education and who consented to the logs
Retention	Whether CM is still teaching	Binary indicator (1/0)	TFA records	Fall of first year, end of first year, and fall of second year	All CMs in the 2012-2016 cohorts

NOTES:

[a]This measure is based on the following question: “Overall, I believe the preparation I received from Teach For America prior to starting at my school was valuable in my efforts to become a successful teacher."

[b]This measure is a composite created from eight items related to CMs’ beliefs in TFA's mission of ensuring that each child has an opportunity to receive an excellent education, and CMs’ beliefs in their own capacity to effectively contribute to this mission.

[c]The log item asked CMs whether the lesson featured an author, characters, or a community with similar background to the focal student’s. This item was coded 1 if a CM reported that this practice was “a focus” of instruction, and 0 if they reported that they “touched on it briefly” or “not touched today.”

(continued)

Table 2.2 (continued)

[d]The log items asked teachers whether the focal student was asked to cite text evidence in their writing. This item was coded 1 if a teacher reported that this practice was "a focus" of instruction, and 0 if they reported that they "touched on it briefly" or "not touched today."

[e]This composite was created based on the items in the math logs that map onto the TFA core components. The scale of the items differs, so each item was first z-scored based on the mean and standard deviation of the comparison group logs; these z-scores were then averaged across items, and the average score was z-scored again based on the mean and standard deviation for the comparison logs. Because some of the log items differed across levels, the composite score was created separately for the lower elementary school and upper elementary or secondary school logs. The reliability (Cronbach's alpha) for the composite is 0.71 at the lower elementary school level (13 items) and 0.75 at the upper elementary or secondary school level (14 items).

they are assigned to a TFA region. In turn, CMs' regional assignment determines where they will be trained. Some regions run their own "regional" institute for their CMs, whereas other regions send their CMs to one of the national institutes. Each national institute hosts CMs from several TFA regions. In summer 2016, six national institutes trained CMs from 44 TFA regions. The national institute in Tulsa hosted eight regions in summer 2016.[12]

The study team used two types of quasi-experimental research design to explore whether the redesigned training shows promise for improving the outcomes of the 2016 cohort of CMs in the eight program regions. Because both study designs are quasi-experimental, the findings in this report should not be interpreted as the *causal* effect of the new training, because the observed changes in CMs' outcomes could have been caused by other confounding factors (rather than the redesigned training). For this reason, the findings in this report are referred to as *estimates* of the effect of the redesigned training.

However, for some CM outcomes, there is a relatively higher level of confidence that the findings *could* represent the causal effect of the new training, because more information was available to account for confounding factors. The findings that are most likely to represent the effect of the new training are those pertaining to CMs' perception of the summer training, their commitment to teaching, and their retention rates. For these outcomes, the findings can be interpreted as being strongly suggestive of the presence or absence of a causal effect. In contrast, findings related to CMs' use of instructional strategies aligned with college- and career-readiness standards should be interpreted with greater caution. Box 2.1 summarizes the levels of confidence related to the findings in this study. The sections below describe the two study designs in greater detail, beginning with the design whose findings are most likely to represent the effect of the redesigned TFA training.

[12]The regions were Detroit, Idaho, Jacksonville, Miami-Dade, Greater Tulsa, Orlando, Washington, and the Bay Area. However, not all of the Bay Area region's CMs were trained at the Tulsa national institute — some were trained at the regional institute. See Appendix D for a discussion of how the study design accounted for this discrepancy.

Box 2.1

Levels of Confidence with which the Estimated Effects in This Study Can Be Attributed to the Redesigned Training

The two study designs used in this evaluation are quasi-experimental, so the findings presented in this report cannot be interpreted as the causal effect of the redesigned summer training. However, one of the two study designs can account for many confounding factors, so there is a relatively higher level of confidence that the findings observed using this design *could* be attributable to the effect of the new training.

Greater (but not perfect) confidence: Comparative interrupted time series (CITS) design

Effects on the following outcomes:

- Perceived value of the training
- Commitment to teaching and educational equity
- Retention in the TFA program

Key potential confounder:

- A policy change or state reform that happened in 2016 (the same year as the redesigned institute) and that affected CMs' outcomes to a different degree in the program and comparison regions

Less confidence: Cross-sectional Program-Comparison Group Design

Effects on the following outcomes:

- Use of instructional strategies aligned with career- and college-readiness standards

Key potential confounders:

- Unobserved preexisting differences between the characteristics of CMs in the program and comparison regions
- Unobserved program-comparison differences in the characteristics of CMs' schools

Other issues:

- High attrition (non-consent and nonresponse rates)

Approach 1: Comparative Interrupted Time Series Design (CMs' Perceptions of the Training, Commitment, and Retention)

The study team evaluated whether the redesigned training improved CMs' perceptions of the training, their commitment to teaching, and their retention rates, using a comparative interrupted time series (CITS) design. This design has been used for decades to evaluate interventions in areas such as epidemiology, political participation, substance abuse, advertising,

and employment programs.[13] In education, CITS designs have been used to evaluate federal policies such as No Child Left Behind, as well as whole school reform models.[14] Studies have shown that, when well implemented, CITS designs can reproduce the results of a randomized experiment in some circumstances.[15]

In this study, a CITS design was used to look at the trends in the outcomes of consecutive cohorts of CMs, before and after the redesigned training was launched. The study followed CMs in two groups of regions: one group that was trained at the national institute in Tulsa where the redesigned training was offered (the eight program regions) and another group that was trained at one of the other national institutes where the traditional summer training was offered (the comparison regions). The study team estimated the effect of the redesigned summer training by looking at whether the program regions "deviated" from their baseline trends by a greater amount than the comparison regions. Figure 2.1 illustrates how the study team applied the CITS design in this study. The team took the following steps:

- **Mean outcomes by cohort:** First, the study team used CM data to create a time series that represents the average outcomes of consecutive cohorts of CMs in the program regions (black dots) and the comparison regions (white dots). A cohort is defined as a group of CMs who joined TFA and attended the summer institute the same year. The cohorts that received their summer training before the launch of the redesigned model are the baseline cohorts (2012-2015 cohorts), and the 2016 cohort is the follow-up cohort.

- **Baseline trends:** The next step was to estimate the baseline trend in CMs' outcomes, using data for the baseline cohorts. The study team estimated baseline trends separately for the program regions (solid black line) and the comparison regions (solid gray line). These trends were then projected into the follow-up period, thereby making it possible to predict the outcomes that the 2016 cohort of CMs would have experienced had their regions' baseline trends persisted (dotted lines).

- **Estimated deviations from the baseline trend:** The next step was to compare the *actual* mean outcomes of the 2016 cohort (the dots in the follow-up period) with what these CMs' mean outcomes would have been had their regions' baseline trends persisted (the dotted lines). The study team estimated these "deviations from baseline trend" for the program regions and the comparison regions.

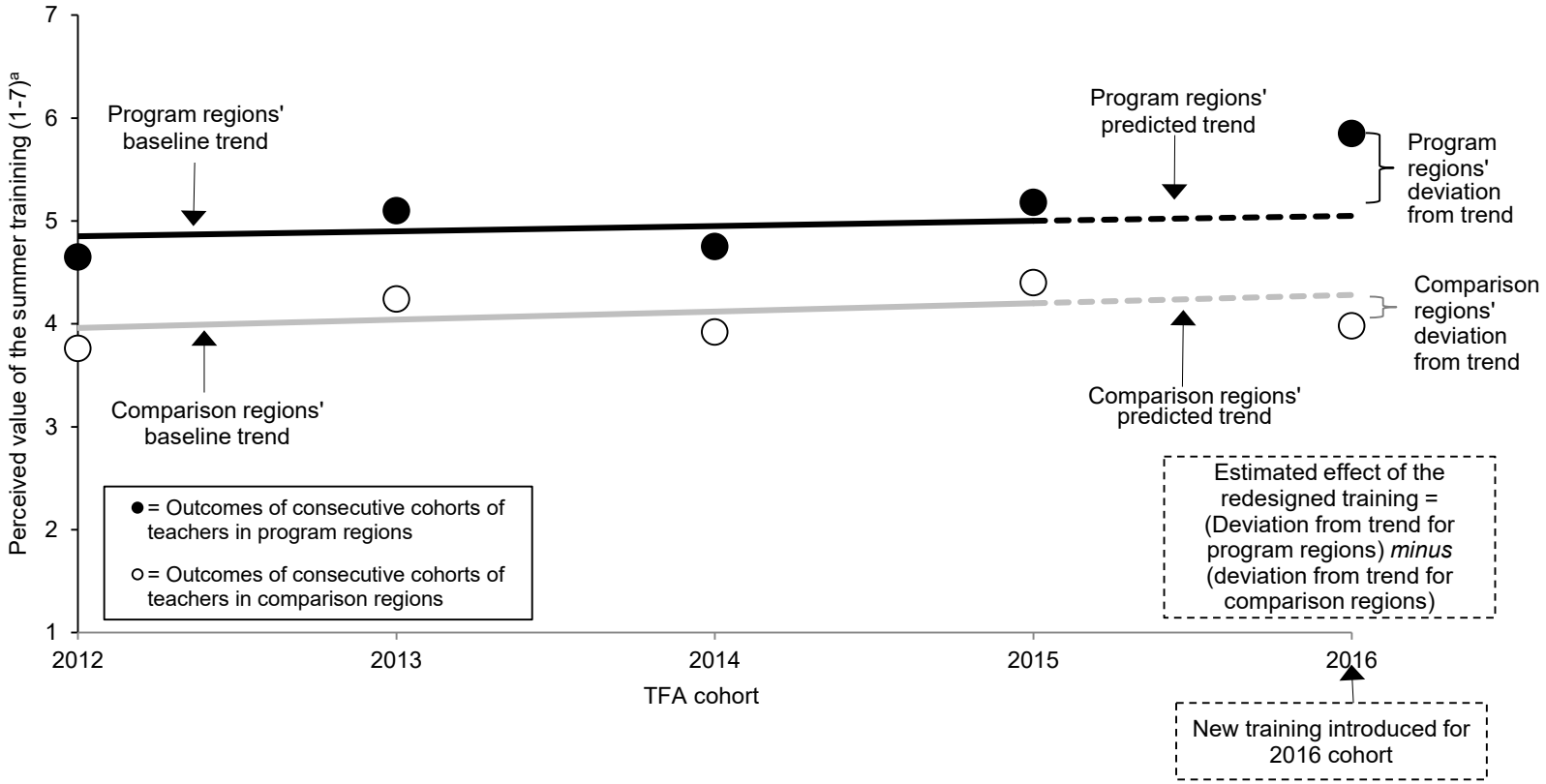
[13]Bloom and Riccio (2005); Ballart and Riba (1995); Campbell and Ross (1968); Mulford, Ledolter, and Fitzgerald (1992). For a discussion and history of CITS designs, see Shadish, Cook, and Campbell (2002). For a discussion of these designs in the context of education research, see Bloom (2003).

[14]Dee and Jacob (2011); Wong, Cook, and Steiner (2009); Kemple, Herlihy, and Smith (2005).

[15]St. Clair, Cook, and Hallberg (2014); Somers, Zhu, Jacob, and Bloom (2013); Fretheim et al. (2013).

Figure 2.1

Using a Comparative Interrupted Time Series Design to Estimate the Effect of the Redesigned Summer Training on the Perceived Value of the Summer Training: A Hypothetical Example



SOURCE: Teach For America corps member surveys.

NOTES: [a]The survey outcome measures are based on survey items with a 7-point agreement scale. The "value of the summer training" measure is based on the following question: "Overall, I believe the preparation I received from Teach For America prior to starting at my school was valuable in my efforts to become a successful teacher."

- **Estimated effect:** Finally, the study team estimated the effect of the redesigned training as the difference between the average deviation from baseline trend for the program regions and the average deviation from trend for the comparison regions. If the redesigned training is more effective than the traditional training, then the program regions should have experienced improvements relative to their baseline trend that exceed the improvements found in the comparison regions.

Thus, in a CITS design, the actual outcomes of the two groups are not compared directly (the black and white dots for the 2016 cohort). Instead, the analysis compares the amount by which the two groups *deviated* from their separate baseline trends.

The CITS design is more rigorous than most other quasi-experimental designs because it combines time series data with a matched comparison group. Together, these two design elements make it more plausible that estimated effects from a CITS design can be attributed to the effect of the redesigned training, because they can eliminate most alternative explanations for why deviations from trend would be different for the program and comparison regions.[16] First, by using several years of preintervention (baseline) data, the CITS design eliminates the possibility that estimated effects are due to differences in preexisting trends between the program regions and the comparison regions.[17] Second, by including a comparison group, the CITS design also reduces the possibility that estimated effects are confounded with the effect of a state or district reform launched at the same time as the redesigned training.[18] In the CITS design, the effect of other educational reforms or initiatives is captured by the comparison regions' deviations from their baseline trend, which is subtracted out when estimating the effect of the new training.

This illustrates an important assumption of the CITS design. For the design to accurately estimate the causal effect of the redesigned training, the comparison regions' deviations from their baseline trend must represent the deviations from trend that the program regions would have experienced *had they not attended the redesigned training* ("the counterfactual"). To maximize the likelihood that this assumption holds true, the study team matched the program regions to comparison regions whose CMs had similar outcomes and characteristics as CMs in the program regions during the baseline period. Matching based on the outcomes of earlier cohorts of CMs makes it more probable that CMs in the program and comparison regions were affected by the same contextual factors in the follow-up period.

In total, 21 TFA regions were included in the CITS analysis: 8 program regions and 13 comparison regions. CMs in the program regions were trained at the national institute in Tulsa and therefore received the redesigned training. The study team selected the comparison regions from among the regions whose CMs were trained at one of the other five national institutes in summer 2016. To increase the likelihood that the comparison regions' deviations from their baseline trend provide the right counterfactual, the team matched each of the program regions to the

[16]Corrin and Cook (1998).

[17]This is also referred to as maturation bias (Shadish, Cook, and Campbell, 2002).

[18]This is also referred to as historical bias (Shadish, Cook, and Campbell, 2002).

two comparison regions whose baseline cohorts of CMs had the most similar outcomes and characteristics. Because the pool of comparison regions from which to choose was small, the team conducted the matching "with replacement," which means that some comparison regions were chosen as the "best match" for more than one program region. Regions chosen more than once were given a greater weight in the CITS analysis.[19]

Approach 2: Cross-Sectional Program-Comparison Group Design (Instructional Strategies Aligned with College and Career Standards)

Data on CMs' use of standards-aligned instructional strategies — which were measured using closed-ended teacher logs — are available only for the 2016 cohort of CMs (the follow-up cohort). Therefore, a CITS design could not be used to examine whether the redesigned training shows promise for improving teachers' use of these strategies. Instead, the study team used a cross-sectional program-comparison group design to estimate effects on this outcome. This study design was used to compare the self-reported instructional strategies used by CMs in the program and comparison regions in the 2016 cohort, controlling for CMs' characteristics.[20]

A key limitation of this study design is that it cannot account for differences in prior outcomes between the two groups of CMs, nor can it fully control for differences in the local context. Therefore, differences in the use of standards-aligned instructional strategies between the program and comparison groups should be interpreted with caution because these differences could be due to factors other than the effect of the redesigned training.

Analysis Samples

This section discusses the samples of CMs included in each type of analysis and their characteristics, for each group of regions (program and comparison). As noted earlier, the program and comparison regions in the evaluation should be as similar as possible in terms of the characteristics and outcomes of the CMs and the schools that they serve. This increases the likelihood that CMs in the two groups were motivated and affected by the same factors and that they faced similar contextual challenges. This, in turn, makes the comparison regions a more credible reference point for the program regions.

When comparing the characteristics of CMs in the program and comparison regions, the size of the difference between the two groups is more important than the statistical significance of those differences. A useful rule of thumb is that differences should not exceed an effect size of

[19]See Appendix D for additional information on the selection of the study regions and Appendix E for information on the statistical model used.

[20]See Appendix E for further information on the statistical model used.

0.25.[21] An effect size is the difference in values for an outcome measure or characteristic expressed as a proportion of the standard deviation for that outcome measure or characteristic.[22]

Full Study Sample (CITS Design)

In total, 21 TFA regions were included in the CITS analysis (8 program regions and 13 comparison regions). From these 21 regions, the CITS analysis included five cohorts of CMs: the 2012 to 2015 cohorts (the baseline cohorts) and the 2016 cohort (of which some members received the redesigned training at the national institute in Tulsa). The full study sample includes 8,324 CMs across all regions and across all cohorts (an average of about 79 CMs per region per cohort). The study team used this sample to examine whether the redesigned training shows promise for improving teacher retention rates.[23]

Overall, CMs in the comparison regions in the full study sample appear to be a credible reference point for the CMs in the program regions. Table 2.3 presents the outcomes of the 2015 cohort of CMs in the program and comparison regions. These results show that, before the redesigned training was launched, CMs in both groups had similar retention rates, as well as similar perceptions of the summer training and commitment to teaching. For example, in the fall of the second year of their tenure in TFA, about 80 percent of CMs in both groups were still teaching. At the end of their first year, both groups "somewhat agreed" that the summer training was valuable (an average rating of 4.5 out of 7), and both groups "agreed" that they were committed to teaching and educational equity (an average rating of almost 6 out of 7). The magnitude (as an effect size) of all baseline differences in this table is less than 0.25.

Whereas Table 2.3 focuses on the outcomes of a baseline cohort of CMs pre-dating the redesigned summer training, Table 2.4 looks at the characteristics of the CMs in the 2016 cohort, some of whom received the redesigned training. This table shows that the CMs in the program and comparison regions had similar levels of education and preparation before TFA, as well as similar demographic characteristics. For example, about 16 percent of CMs in both groups majored in an education-related discipline, and 35 percent were the first in their family to attend college. About 52 percent of CMs in both groups were persons of color, and about three-fourths were women. The magnitude (as an effect size) of all baseline differences in this table is less than 0.25.

[21]This benchmark is based on studies showing that quasi-experimental designs are more likely to produce biased results when baseline differences exceed this threshold. See Ho, Imai, King, and Stuart (2007). It is also the criterion used by the What Works Clearinghouse, which is the clearinghouse for education research hosted by the Institute for Education Sciences in the U.S. Department of Education. See What Works Clearinghouse (2017).

[22]Effects sizes in this section are based on the standard deviation for the pooled sample (program and comparison regions combined).

[23]Retention records are available for all CMs in the study sample.

Table 2.3

Retention Rates and Teacher Perceptions, by Group, Last Baseline Cohort, 2015 Cohort

Outcome	Program Regions	Comparison Regions	Estimated Difference	Effect Size	P-Value
Retention					
Fall of first year (% of CMs still with TFA)	93.7	96.5	-2.7	-0.13	0.188
End of first year (% of CMs still with TFA)	83.8	87.5	-3.7	-0.10	0.192
Fall of second year (% of CMs still with TFA)	80.2	82.1	-1.9	-0.05	0.504
Number of regions	8	13			
Number of corps members	420	887			
Perceptions in the fall of first year					
Value of the summer training (1-7)	4.72	4.80	-0.08	-0.04	0.691
Commitment to teaching and equity (1-7)	5.83	5.85	-0.02	-0.03	0.687
Number of regions	8	13			
Number of corps members	360	781			
Perceptions in the middle of first year					
Value of the summer training (1-7)	4.62	4.65	-0.03	-0.01	0.905
Commitment to teaching and equity (1-7)	5.67	5.78	-0.11	-0.13	0.217
Number of regions	8	13			
Number of corps members	335	746			
Perceptions at the end of first year					
Value of the summer training (1-7)	4.49	4.53	-0.04	-0.02	0.861
Commitment to teaching and equity (1-7)	5.82	5.89	-0.07	-0.08	0.541
Number of regions	8	13			
Number of corps members	331	715			
Perceptions in the fall of second year					
Value of the summer training (1-7)	4.61	4.68	-0.07	-0.04	0.727
Commitment to teaching and equity (1-7)	5.91	5.93	-0.02	-0.03	0.764
Number of regions	8	13			
Number of corps members	328	660			

(continued)

Table 2.3 (continued)

SOURCES: Teach For America corps member survey and adminstrative records.

NOTES: The findings in this table are for the 2015 cohort of corps members (CMs), who joined TFA in the year before the launch of the redseigned summer institute in summer 2016.

The reported outcome measures are based on survey items with a 7-point agreement scale. The "value of the summer training" measure is based on the following question: "Overall, I believe the preparation I received from Teach For America prior to starting at my school was valuable in my efforts to become a successful teacher." The commitment to teaching and equity measure is a composite created from eight items related to CMs ' beliefs in TFA's mission of ensuring that each child has an opportunity to receive an excellent education, and CMs' beliefs in their own capacity to effectively contribute to this mission.

The values in the columns "Program Regions" and "Comparison Regions" are observed means. Values in the "Effect Size" column are the estimated difference divided by the standard deviation for the full sample. Rounding may cause slight discrepancies in calculating sums and differences.

A two-tailed t-test was applied to estimated deviations and estimated differences between program regions and comparison regions. The statistical significance of estimated differences is indicated as follows: *** = 1 percent; ** = 5 percent; * = 10 percent.

Table 2.5 presents the characteristics of the 2016 cohort's teaching placements. As shown in this table, the CMs in the program and comparison regions taught similar grade levels and subjects, and in many respects they also taught in schools characterized by similarly high levels of need. For example, about 96 percent of CMs in both groups taught in a Title I school, and more than 80 percent of students in these schools were eligible for free or reduced-price lunch. Overall, the schools where CMs were placed also had a high proportion of first-year teachers (about a third), and the teachers at these schools had high rates of absenteeism (about 29 percent of teachers). On the other hand, there are also some notable contextual differences between the two groups of CMs. For instance, the CMs in the program regions taught in schools where 14 percent of students were chronically absent, compared with 21 percent in the comparison CMs' schools.[24] Students in the program region schools were also somewhat less likely to be eligible for free or reduced-price lunch and to be suspended.

Yet despite these differences, *both* groups of CMs taught in schools with high levels of disadvantage. Therefore, even though some of the differences in school characteristics exceed an effect size of 0.25, the comparison regions' outcome trends may still potentially provide a credible counterfactual for the program regions' trends. As explained earlier, for a CITS design to produce an accurate estimate of the effect of the redesigned training, one must assume that the comparison regions' deviation from their baseline trend represents what would have happened in the program regions had they not received the redesigned training. For example, if there were a policy change or a labor market shock that affected teacher outcomes (perceptions and retention) in 2016, then one must assume that this policy change or shock would have similarly affected CMs in the program and comparison regions. This may still be a plausible assumption for two reasons. First,

[24]These differences do not reflect the effect of the CMs in the program regions on their students, because the school characteristics data are for the school years *preceding* CMs' placement (either the 2015-2016 or 2013-2014 school year).

Table 2.4

Characteristics of the 2016 Cohort, Full Study Sample

Characteristic	Program Regions	Comparison Regions	Estimated Difference	Effect Size	P-Value
Preparation and skills					
Composite application score (1-5)	3.62	3.63	-0.01	-0.02	0.813
Educational attainment (%)					
Bachelor	84.5	86.0	-1.5	-0.04	0.731
Masters	12.5	12.1	0.4	0.01	0.910
Doctorate	1.7	2.0	-0.4	-0.02	0.723
Other graduate degree	0.8	0.2	0.6	0.10	0.132
Major or minor in Education (%)	16.1	15.9	0.2	0.00	0.939
Prospect type (%)					
Undergraduate	56.0	60.5	-4.5	-0.08	0.469
Graduate	6.9	5.5	1.5	0.06	0.375
Professional	36.5	34.4	2.1	0.04	0.710
Demographic					
Age at entry into TFA	25.61	25.15	0.46	0.07	0.567
First in family to attend college (%)	35.1	35.3	-0.2	0.00	0.973
Received a Pell grant (%)	48.2	51.7	-3.5	-0.06	0.573
Race and ethnicity (%)					
Hispanic	10.9	17.5	-6.5	-0.18	0.401
Black	22.1	22.1	0.0	0.00	0.997
White	44.5	47.4	-2.9	-0.05	0.730
Asian	6.0	3.9	2.1	0.09	0.265
Other	14.5	9.1	5.4 **	0.16	0.015
Person of color (%)	52.1	51.1	1.0	0.02	0.905
Female (%)	72.8	75.8	-3.0	-0.06	0.460
Number of regions	8	13			
Number of corps members	361	815			

SOURCE: Teach For America administrative records for corps members.

NOTES: A cohort is defined as a group of TFA corps members (CMs) who joined TFA and attended the summer institute the same year. The findings in this table are for the 2016 cohort of CMs, who joined TFA when the redesigned summer institute was launched in summer 2016. The number of CMs represents the total number of CMs in this cohort. CMs' characteristics are measured during the application process to TFA. The composite application score is based on a 5-point scale measuring CMs' skills in several areas, using information from their application materials and interviews.

The values in the columns "Program Regions" and "Comparison Regions" are observed means. Values in the "Effect Size" column are the estimated difference divided by the standard deviation for the pooled sample. Rounding may cause slight discrepancies in calculating sums and differences.

A two-tailed t-test was applied to estimated deviations and estimated differences between program regions and comparison regions. The statistical significance of estimated differences is indicated as follows: *** = 1 percent; ** = 5 percent; * = 10 percent.

Table 2.5

Characteristics of First-Year Teaching Placements, 2016 Cohort, Full Study Sample

Placement Characteristic	Program Regions	Comparison Regions	Estimated Difference	Effect Size	P-Value
Grade and subject					
Grade level (%)					
Lower elementary school	18.5	17.1	1.4	0.03	0.729
Upper elementary school	29.7	22.7	7.0	0.15	0.153
Middle school	27.6	30.1	-2.5	-0.05	0.615
High school	24.0	30.2	-6.2	-0.13	0.237
Subject taught (%)					
General education	39.6	27.7	12.0	0.24	0.125
English Language Arts	21.3	23.1	-1.9	-0.04	0.690
Mathematics	17.6	20.1	-2.5	-0.06	0.380
Science	16.2	15.3	0.8	0.02	0.739
Social studies	3.9	6.6	-2.7	-0.10	0.275
World languages	0.7	2.9	-2.2	-0.13	0.127
Other	1.2	3.5	-2.3	-0.13	0.272
School					
School size and type					
Total enrollment	763.27	699.10	64.16	0.11	0.617
Title I school (%)	95.7	96.7	-1.0	-0.05	0.778
Magnet school (%)	25.5	6.7	18.82 *	0.51	0.054
Charter school (%)	23.2	33.0	-9.8	-0.21	0.421
Staffing (%)					
First-year teachers at the school	36.3	28.3	8.0	0.28	0.388
Teacher absences (10 or more per year)	37.7	24.8	12.9 *	0.52	0.059
Certified	97.3	90.3	7.0	0.37	0.141
School location (%)					
Urban	58.3	70.6	-12.3	-0.25	0.404
Suburban	33.9	19.2	14.8	0.34	0.181
Town	2.0	5.5	-3.5	-0.16	0.583
Rural	5.2	4.7	0.5	0.02	0.914

(continued)

Table 2.5 (continued)

Placement Characteristic	Program Regions	Comparison Regions	Estimated Difference	Effect Size	P-Value
Students (%)					
Chronically absent	13.9	21.4	-7.5	-0.35	0.152
Retained	3.6	5.7	-2.1 *	-0.33	0.091
One or more in-school suspensions	7.6	9.2	-1.5	-0.12	0.472
One or more out-of-school suspensions	11.4	16.6	-5.2	-0.36	0.102
Free or reduced price lunch	81.5	86.3	-4.9	-0.28	0.222
English as a second language	16.1	15.4	0.7	0.04	0.896
Individualized Education Plan	10.8	12.1	-1.4	-0.19	0.416
Race and ethnicity					
Hispanic	33.7	42.4	-8.7	-0.24	0.537
Black	46.0	46.1	-0.1	0.00	0.996
White	12.9	6.4	6.5	0.54	0.118
Asian	2.6	2.2	0.5	0.09	0.771
Other	4.6	2.9	1.8	0.36	0.316
Female	48.6	49.0	-0.5	-0.07	0.429
Number of regions	8	13			
Number of corps members	361	815			

SOURCES: Teach For America administrative records for corps members, Common Core of Data (school year 2015-2016), and Office of Civil Rights Data Collection (2013-2014).

NOTES: A cohort is defined as a group of TFA corps members (CMs) who joined TFA and attended the summer institute the same year. The findings in this table are for the 2016 cohort of CMs, who joined when the redesigned summer institute was launched in summer 2016. The number of CMs represents the total number of CMs in this cohort. CMs' placement characteristics are measured in the fall of their first year of teaching.

The values in the columns "Program Regions" and "Comparison Regions" are observed means. Values in the "Effect Size" column are the estimated difference divided by the standard deviation for the pooled sample. Rounding may cause slight discrepancies in calculating sums and differences.

A two-tailed t-test was applied to estimated deviations and estimated differences between program regions and comparison regions. The statistical significance of estimated differences is indicated as follows: *** = 1 percent; ** = 5 percent; * = 10 percent.

CMs in the program and comparison regions were both working in high-poverty disadvantaged schools. Second, earlier cohorts of CMs in the program and comparison regions had very similar retention rates and survey outcomes before the redesigned summer training was launched (as shown in Table 2.3).[25] This, in turn, makes it more likely that the comparison regions' deviations from their baseline trend may potentially provide a plausible counterfactual for the program regions' deviations from trend in the absence of the redesigned training.

[25]The program and comparison regions also have similar baseline *trends* in teacher outcomes. The size and direction of the slope for these baseline trends are similar for the two groups of regions.

Survey Sample (CITS Design)

Whereas retention records were available for all CMs, not all CMs completed the TFA surveys that the study team used to measure CMs' perceptions of the training and their commitment to teaching. Therefore, the team estimated the effect of the redesigned training on these outcomes based on the survey sample, defined as the subset of CMs for whom there were data on the two confirmatory survey outcomes for a given survey wave. Depending on the wave, the survey sample ranges from 6,305 to 7,325 CMs across all regions and across all five cohorts (an average of 60 to 70 teachers per region per cohort). These CMs represent 72 to 93 percent of the full study sample, depending on the wave and cohort.[26] As in the full study sample, the outcomes and characteristics of CMs in the survey sample were similar for the program and comparison regions, except for the already-noted differences in the characteristics of CMs' placement schools.[27]

Teacher Log Sample (Cross-Sectional Difference Design)

The teacher log sample was used to examine whether the redesigned training shows promise for increasing CMs' use of instructional strategies that are aligned with college- and career-readiness standards. The sample includes the subset of CMs in the 2016 cohort who were teaching ELA, mathematics, or general education and who completed at least one closed-ended log during the year. The log sample includes 334 CMs (119 CMs in the program regions and 215 CMs in the comparison regions). These CMs represent 40 percent of CMs in the study regions who were teaching ELA, mathematics, or general education (45 percent in the program regions and 40 percent in the comparison regions).[28]

For the purposes of the analysis, the study team divided the log sample by content area (ELA, mathematics). There are 225 CMs in the ELA sample and 231 CMs in the math sample; CMs teaching general education are included in both samples. In each subsample (ELA and math), an average of 11 CMs per region completed the logs, and each teacher completed eight logs in the relevant content area on average.[29]

When completing a closed-ended log, CMs were asked to focus on a randomly selected student and to report on the characteristics of the selected focal student. Table 2.6 shows that the

[26]The differential attrition rate among program regions does not exceed 10 percent for any cohort or survey wave. See Appendix B for survey response rates. For the survey outcomes, the study meets the What Works Clearinghouse standards for low attrition based on the liberal boundary used by the clearinghouse's review protocol for teacher training studies (What Works Clearinghouse, 2016).

[27]See Appendix B for the outcomes and characteristics of teachers in the survey sample.

[28]See Appendix C for further information on response rates for the logs. About 58 percent of eligible CMs consented to completing the logs; among CMs who consented, about 70 percent completed at least one closed-ended log and were included in the log sample.

[29]On average across the entire log sample, teachers completed 74 percent of the closed-ended logs they were sent during the year.

Table 2.6

Characteristics of the Focal Students, Teacher Log Sample

Characteristic	Program Regions	Comparison Regions	Estimated Difference	Effect Size	P-Value
Individualized Education Plan (%)	14.7	15.5	-0.8	-0.02	0.770
Received same instruction as peers today (%)	92.5	91.5	1.0	0.03	0.574
Performance level (%)					
Lower third	28.9	30.0	-1.1	-0.02	0.634
Middle third	39.6	38.5	1.1	0.02	0.665
Upper third	31.8	31.5	0.3	0.01	0.901
Number of minutes spent on focal subject today	74.98	70.25	4.73	0.13	0.233
Number of regions	8	13			
Number of corps members	119	215			
Number of logs	1,367	2,340			

SOURCE: Closed-ended teacher logs administered by MDRC.

NOTES: Online teacher logs were administered to corps members (CMs) in the 2016 cohort in the program and comparison regions. CMs teaching English Language Arts (ELA), mathematics, or general education in the first year of their teaching placement were eligible for the logs. Closed-ended logs were sent to eligible and consenting CMs every two weeks during the school year (15 weeks total). The instructional items in the logs differ by level (lower and upper grades) and by content area (ELA and mathematics). Secondary school teachers, as well as elementary school teachers in departmentalized schools, were sent the log for their content area and level every two weeks (about once per month). Elementary school teachers teaching both content areas (general education teachers) were sent the ELA log or the math log every two weeks, with the content area alternating between logs (about one log per content area per month). A CM is included in the log sample if the CM completed at least one log during the school year.

At the start of each log, teachers were given a randomly selected letter of the alphabet and asked to report on their instructional practices with a focal student whose name starts with that letter (or the closest letter) on a particular day that week. Teachers were also asked to report on the characteristics of the selected focal student. The values in the columns "Program Regions" and "Comparison Regions" are the observed mean characteristics of these students, as reported by teachers in the log sample. Values in the "Effect Size" column are the estimated difference divided by the standard deviation for the full sample. Rounding may cause slight discrepancies in calculating sums and differences.

A two-tailed t-test was applied to estimated differences between program regions and comparison regions. The statistical significance of estimated differences is indicated as follows: *** = 1 percent; ** = 5 percent; * = 10 percent.

focal students taught by the CMs in the program and comparison regions were very similar with respect to performance levels and academic needs.[30]

As in the full study sample, CMs in the program and comparison regions in the log sample also had similar demographic characteristics and experiences before joining TFA, but they differed with respect to some of the characteristics of the schools where they taught.[31] The statistical analysis controlled for these measurable differences in school characteristics, but the adjustment

[30]All baseline differences in this table are less than 0.25 as an effect size.

[31]See Appendix B for the outcomes and characteristics of teachers in the log sample.

may not have accounted for other unobserved differences in school context between the two groups of CMs that could have been confounded with the causal effect of the redesigned training. Thus, the findings pertaining to the effect of the redesigned training on CMs' use of standards-aligned strategies should be interpreted with caution.

A further limitation is that a high proportion of eligible CMs did not complete the logs.[32] CMs in the log sample differed from the CMs who did not complete the logs with respect to certain characteristics; if the effect of the redesigned training depends on these characteristics, then the findings for CMs in the log sample may not generalize to the effect for CMs who did not participate.[33]

[32]This is considered "high attrition" based on the What Works Clearinghouse standards (What Works Clearinghouse, 2017).

[33]CMs in the teacher log sample and CMs who were eligible for the logs but did not participate had similar application scores and they taught in schools with similar characteristics. However, CMs in the log sample were younger (25 years old, compared with 26 years old for nonparticipating CMs), less likely than nonparticipating CMs to have been a person of color (42 percent, compared with 57 percent), and less likely to have received a Pell grant (45 percent, compared with 53 percent).

Chapter 3

Implementation of the Redesigned Summer Training

This chapter examines the implementation of the redesigned summer training piloted in 2016 at the national institute in Tulsa. The first section starts by discussing how the lead instructors and coaches in both Tulsa and the comparison institutes were themselves taught to deliver training to corps members (CMs) and the extent to which they felt prepared for their roles. The second section discusses the degree to which the Learning Cycle, instructional activities, and the core practices were used by the lead instructors and coaches in Tulsa. As further context for interpreting the implementation findings, key differences between the experiences of CMs trained at the national institute in Tulsa and those trained at the other national institutes are included in the chapter. The following are the main findings discussed in this chapter:

- At the national institute in Tulsa, the lead instructors felt prepared to teach via the Learning Cycle, while the coaches did not feel prepared for their role. The lead instructors and coaches in the comparison institutes felt more prepared for their responsibilities than their counterparts in Tulsa.
- Lead instructors in Tulsa used the first two quadrants of the Learning Cycle to train CMs, but their use of the core practices was minimal.
- Similar to new teachers generally (not just TFA), CMs in Tulsa and in the comparison institutes felt that their training did not lead them to feel very prepared to teach in the fall.

The findings discussed in this chapter draw from observations that the study team conducted at TFA staff training conferences, during Learning Cycle and coaching sessions in Tulsa, and in CMs' summer school classrooms in Tulsa, as well as from focus groups with lead instructors, coaches, and CMs in both Tulsa and comparison institutes at the end of the summer program. See Table 3.1 for a breakdown of data collected.

Training for Lead Instructors and Coaches at the National Summer Institutes

The 2016 national institute in Tulsa was redesigned to enhance the rigor and the relevance of the summer training CMs received. For more than 20 years, the training at the national institutes consisted of training CMs in best teaching practices and then having them apply those best practices to teaching summer school independently. For instance, the best practice for giving students clear directions might have been explained or modeled to CMs before having them try it out in their classrooms. The redesigned training in Tulsa, on the other hand, called for CMs to learn how to give students clear directions by practicing how to do so in authentic situations. For example, CMs might have been asked to rehearse giving students clear directions within the context of a

Table 3.1

Number of Qualitative Research Activities Conducted, by Data Source

Research Activity	Tulsa Corps Members (CMs)	Comparison CMs	Tulsa Lead Instructors	Comparison Lead Instructors	Tulsa Coaches	Comparison Coaches
Focus groups	8	15	4	2	6	5
Observations	17	N/A	17[a]	N/A	5	N/A

SOURCE: Classroom observations conducted by MDRC.

NOTES: [a]Several other sessions were observed but they addressed special topics, and the Learning Cycle was not used.

lesson on Choral Counting. The redesigned training incorporated several "andragogies," or training methods and practices meant specifically for adult learners, as part of the Learning Cycle, while the student curricula that CMs received were aligned with rigorous content standards and grounded in culturally relevant pedagogy.

When TFA, in partnership with the University of Washington, introduced the redesigned training at the regional institute in Chicago in summer 2014,[1] expert staff members from both the university and TFA delivered the training. The circumstances at the national institute in Tulsa in 2016 were very different: TFA alone implemented the program, the program served many more CMs than the institute in Chicago, and the teacher training staff was made up mainly of part-time TFA alumni who also work as full-time teachers, principals, and other education staff, and who have limited experience coaching adults (as opposed to full-time TFA expert staff members).[2] Therefore, to prepare for the pilot in Tulsa, all summer staff members underwent extensive training from TFA, with much of the time devoted to internalizing the new curriculum and the instructional activities. For TFA, this meant both redesigning the training program for TFA staff in Tulsa and increasing the amount of training it provided.

The Role of Lead Instructors and Coaches

The Tulsa and the comparison institutes had lead instructors and coaches who were responsible for training CMs how to teach. The majority of them were TFA alumni. Typically, lead instructors have about three to six years of teaching experience, while coaches have about one to four years of experience. TFA sought out the strongest, most experienced lead instructors for Tulsa. To do so, TFA included more full-time TFA staff members in Tulsa than it typically does at other national institutes. It also made the job application process more rigorous than the process at the comparison institutes. Those applying for lead instructor positions in Tulsa were asked to submit a teaching demonstration video along with their application. TFA gave the applicants feedback about their videos, and as part of the selection process, TFA staff members discussed

[1]The methods used to train CMs in Chicago and later in Tulsa trace their origins to the Teacher Education by Design (TEDD), a project of the University of Washington's College of Education.

[2]The majority of the teacher training staff worked as teachers or in other roles in a school district during the school year.

this feedback with each applicant. In contrast, applicants for lead instructor positions at comparison institutes were not asked for such videos.

In Tulsa, lead instructors were responsible for the daily planning and facilitation of the Learning Cycle sessions. TFA offered lead instructors some guidance on particular sessions, but for the most part the lead instructors relied on their own training and expertise to run their sessions. Coaches floated between Learning Cycle sessions, jumping in to help where needed. For example, coaches helped lead instructors by playing the role of either the teacher or a student during the "modeling to decompose" component of Quadrant 1. Alternatively, the lead instructor and coach(es) spread out to give individual or group support to CMs who were planning or rehearsing a lesson.

In addition to their support role in Learning Cycle sessions, coaches were assigned a small group of CMs teaching the same grade and content, with whom they worked throughout the program. Coaches met with their groups at their assigned summer school teaching site when CMs were neither teaching nor attending social justice seminars. During the first week of the institute, TFA went over generally with coaches what topics or activities they should cover with CMs, but thereafter the coaches used the coaching time to respond to the groups' particular needs. For instance, CMs could use the time to plan and rehearse lessons, reflect on their teaching, and learn about topics such as lesson planning or using student data. In addition, coaches observed and videotaped CMs teaching their summer school classes both to identify where CMs needed support and to have recordings for the CMs to use to engage in analysis and reflection of their practice by watching videos during Quadrant 4 of the Learning Cycle.

The responsibilities of the lead instructors and coaches at the comparison institutes differed from those at the institute in Tulsa. For staff at the comparison institutes who had previously worked at a TFA institute, the model in 2016 was quite familiar to them and thus did not represent a radical departure in their training. The lead instructors at the comparison institutes were responsible for leading professional development sessions that CMs attended each afternoon on topics related to lesson planning, executing lessons, classroom management, and diversity and inclusiveness. TFA staff members in comparison institutes designed the professional development sessions in advance, but lead instructors customized them. Coaches, on the other hand, were expected to observe a group of CMs assigned to them in their summer school classes several times a week and provide written and verbal feedback to CMs in groups and one-on-one when they were not teaching or attending sessions, to provide additional support for their development.

Preparedness of Lead Instructors and Coaches

The Tulsa lead instructors participated in extensive training and completed assignments to prepare for their roles at the institute. All together, they attended five weekend-long training conferences instead of the usual four. Lead instructors began their training by attending a weekend conference in January 2016 specifically geared toward them, followed by a weekend leadership conference in March 2016. At these conferences, lead instructors had the opportunity to experience the instructional activities and the core practices through the Learning Cycle as learners (that is, they

role-played as CMs). TFA wanted lead instructors to begin their training in this way to deepen their learning by making it more authentic, believing it would instill in the instructors a sense of empathy for their future CMs and encourage them to take a more CM-centered approach to teaching. For the remainder of their training, lead instructors led Learning Cycle sessions in which coaches were the learners. This gave lead instructors the opportunity to further deepen their understanding of how to use the Learning Cycle to train CMs. This training occurred at conferences held on weekends in March (virtual and in addition to the previously mentioned March conference in March), April, and May, as well as during the week immediately before the summer institute in Tulsa began. Lead instructors were also asked to enact instructional activities in their own classrooms and to meet virtually with other lead instructors to discuss what they were learning to ensure a continual arc of development between conferences.

The four core practices central to the redesigned model were emphasized and discussed with the lead instructors throughout their training. Lead instructors were asked to incorporate these core practices into the lessons they were developing and practicing for teaching through the Learning Cycle. However, the lead instructors were allowed to use their own judgment about which and how many core practices to feature in their lessons.

The coaches attended the same training sessions as the lead instructors, beginning with the virtual conference in March. As mentioned above, the coaches experienced the Learning Cycle as learners, but they never had the opportunity to practice delivering the training. Thus, they did not have the benefit of the deeper learning that comes from that experience. The majority of coach-specific training was condensed to a brief window shortly before the institute training began.

Lead instructors and coaches at the comparison institutes also received training before the summer institute began. Lead instructors participated in training at four weekend-long conferences between March and May 2016, while coaches participated in two of these. Their training, as described by TFA, focused on "orienting staff to the vision of their roles, institute goals, and some skill-building and practice." The training was not redesigned as significantly as the Tulsa training was in 2016, although TFA makes improvements to institutes from one year to the next.

How Prepared the Lead Instructors and Coaches Felt

The lead instructors in Tulsa generally began the program feeling confident in their ability to train CMs via the Learning Cycle. They had had several opportunities to practice for their role in the months preceding the institute training. Soon after the institute training began, however, some of their confidence waned. The student curricula that they were training CMs to use was grounded in culturally relevant pedagogy. Although the lead instructors still felt good about using the Learning Cycle, they had not anticipated the issues and questions that came up organically during the session relevant to social injustices and inequity in education as a result of the curricula. They did not feel prepared for those conversations, as it was their understanding that culturally relevant pedagogy training was the responsibility of the social

justice seminar leaders.[3] Some of the lead instructors thought it might have been helpful to have time to plan with the social justice seminar leaders, but the organization of the institute made it logistically difficult for the lead instructors and the social justice seminar leaders to meet to plan together; seminar leaders taught CMs at the summer school teaching sites in the morning, while lead instructors taught at the institute hub in the afternoons.

Coaches in Tulsa were not satisfied with the training they received before the institute training began. Among the issues they raised were that they received too little training, too late, and that experiencing the new model from only the perspective of a learner was unhelpful as it did not allow them to practice coaching with the redesigned model in mind. They thought that the training did not equip them with coaching strategies, nor did it explain the benefit and intent of the model or what their role was within the model. They did, however, feel well versed in the instructional activities that they were taught by the lead instructors during the lead instructors' practice sessions and that CMs would rehearse and teach to their summer school students.

The majority of the coaches in Tulsa reported in focus groups that they were generally unclear about what their role at the institute was supposed to be, as the redesigned model brought about a shift in expectations. They had expected to work with CMs on a one-on-one basis, which was what coaches in the comparison institutes did. However, the redesigned training called for coaches to work with CMs in a group setting in which they would work in Quadrants 2 and 4 of the Learning Cycle and cover topics such as lesson planning, data management, and so on in the morning at the summer school sites. In the afternoons, they were asked to float between Learning Cycle sessions and assist the lead instructors as needed. As they had only experienced the Learning Cycle as learners, many did not feel prepared to lead group coaching related to Quadrants 2 and 4. They were also unsure about their role in the Learning Cycle sessions. In the comparison institutes, coaches worked all day at a summer school site and reported to a TFA school director, who held them to account. In Tulsa, however, they began their days at a summer school site and then had to travel to the TFA hub in the afternoons. They were not sure if the lead instructor, the school director, someone else, or all of them were holding them to account, and they did not know for what they were being held accountable. Similarly, lead instructors did not know what they could ask of the coaches. There was not enough time for lead instructors and coaches to plan together, so coaches were not sure how they were supposed to support the CMs or assist the lead instructors.

Because the Tulsa coaches had only experienced the Learning Cycle and the instructional activities with embedded core practices as learners and did not receive explicit instructions on how to support CMs in these areas, some coaches reverted to their own coaching methods, such as providing one-on-one coaching (as in the traditional model), sometimes in lieu of group coaching (as in the redesigned model).

[3]In fact, the content that was supposed to be covered in the social justice seminars was very different from what was supposed to be covered in sessions on culturally relevant pedagogy. This was a course of some confusion among many of the teacher training staff. Social justice seminars covered issues of identity and privilege while culturally relevant pedagogy was specific to the classroom.

Lead instructors and coaches in the comparison institutes did not express the same dissatisfactions with their training. It was very similar to previous trainings for those who had been lead instructors, coaches, or both for TFA in the past. And it was familiar to most of them, as they had received similar training when they were CMs.

Although the Tulsa coaches felt ill prepared, CMs who attended the institute in Tulsa often described their coach as their "go-to" person outside of class or as someone who helped them with specific teaching issues. CMs at the comparison institutes felt similarly about their coaches.

Implementation of the Learning Cycle, Instructional Activities, and Core Practices During the Institute Training in Tulsa

The study team visited the institute in Tulsa at two separate points in time during summer 2016 to conduct observations of Learning Cycle sessions and of CMs teaching their summer school classes.[4] A third visit, during the fifth (last) week of institute training, was devoted to focus groups.

Learning Cycle

The study team observed Learning Cycle sessions during the first and third weeks of the institute training. Of the 17 sessions observed, 11 occurred during the first week, and the remaining 6 occurred during the third week.[5] The team used six different versions of the Learning Cycle observation protocol — one for each of the andragogies within the quadrants.[6] The main function of the protocols was to capture the use of the andragogies, the instructional activities, and the core practices central to the model during each session. Any given session might have included several andragogies, and thus several versions of the observation protocols were used for each session. For instance, in a single session, an instructional activity could be modeled, modeled to decompose, and then rehearsed by CMs.

Table 3.2 shows the breakdown of the andragogies that were in play, by content, during the 17 Learning Cycle sessions the study team observed. While many Learning Cycle sessions were not observed at all, the results strongly suggest that Quadrant 4, in which CMs have an

[4]The study team observed too few coaching sessions to report on them. Since these sessions occurred during summer school classes, a choice was made to prioritize observations of CMs over those of coaches.

[5]During each Learning Cycle time slot, different sessions were held for CMs specializing in English language arts or math. Within these two subject areas, sessions were further divided by grade level wherein there were three elementary grade-level sessions and three secondary grade-level sessions happening concurrently, for a total of six different sessions occurring during any given time slot. MDRC observers split their time roughly evenly across subject areas and grade levels. The 17 observations occurred during 10 time slots. There were also science Learning Cycle sessions but they were not the focus of the study.

[6]These andragogies are modeling, modeling to decompose, planning, rehearsing, microteaching, and video professional learning community or analyzing student work.

Table 3.2
Pedagogies of Focus During the Learning Cycle Sessions, by Content Area

Pedagogy	ELA	Math	Total
Quadrant 1			
Modeling	5	3	8
Modeling to decompose	4	3	7
Quadrant 2			
Collaborative planning	4	2	6
Rehearsing or microteaching	4	3	7
Quadrant 4			
Video professional learning community	2	0	2
Analyzing student work	0	0	0
Number of observations	10	7	17

SOURCE: Classroom observations conducted by MDRC.

opportunity to reflect on their teaching, was rarely used in sessions. Lead instructors were rarely observed using video professional learning communities and never observed analyzing student work. However, lead instructors in the observed session spread their use of the andragogies evenly in the first two quadrants.[7]

Seventy percent of the time in Learning Cycle sessions was spent on the andragogies. The remainder of the time was spent discussing logistics (reviewing session agendas, clarifying when different activities were happening, and so on) and teaching-related topics (instructional concepts, the appropriateness of particular texts, and so on). As might be expected given the new program, many logistical issues and questions about the curricula arose and had to be worked out. Lead instructors, and especially coaches, reported that as the institute training went on, they increasingly made decisions about how to spend their time, during Learning Cycle sessions and other times, by responding to CMs' specific needs. For instance, while a lead instructor may have planned for CMs to model an instructional activity on a certain day, the instructor may have decided to let CMs use that time to refine the next day's lesson plans instead.

In general, CMs had mixed responses to the Learning Cycle. For example, CMs in five out of eight focus groups in Tulsa felt that modeling and rehearsing were not very helpful because they were not realistic. (That is, the CMs acting as "students" were actually well-behaved adults who did not act up or disrupt the classroom.) While CMs in comparison institutes were not explicitly exposed to the Learning Cycle concept, their lead instructors might have modeled lessons on occasion to teach specific skills if they thought that it would have been helpful, and in some

[7]Lead instructors and coaches had some conflicting ideas related to two of the andragogies (rehearsing and microteaching) in Quadrant 2.

cases, they may have asked CMs to rehearse in front of their peers. However, these activities were not a requirement of the comparison institute training. CMs in comparison focus groups did not share many of the same criticisms of modeling as CMs in the Tulsa group. For example, they generally did not describe the modeling they were exposed to as "unrealistic" and did not think too much time was spent on modeling. This is most likely a function of the greater amount of time that lead instructors in Tulsa spent modeling, as called for in the redesigned training.

Another reflection that CMs in some focus groups had about the Learning Cycle was that they often found themselves rehearsing for each other without the benefit of feedback from an experienced educator. At the same time, CMs reported that rehearsal was often useful simply for the opportunity to hear themselves verbalize the lesson.

TFA has operated without a major overhaul for more than 20 years, and many CMs arrive at the institute having heard from alumni and TFA about what to expect. A reason that the CMs did not embrace the Learning Cycle wholeheartedly may have been that the new training model did not match their expectations of what the institute would be like. Coaches, though more aware of the major changes, also seemed to struggle with the new model, as it was so different from what they had experienced as CMs (and in some cases as prior institute staff members), and they sometimes reverted to the traditional model. This wavering between models probably added to CMs' confusion. Instead of seeing the Learning Cycle as a platform for observing, practicing, and discussing teaching decisions that could apply in multiple circumstances, as some lead instructors understood it, some CMs reported in focus groups as seeing it more as a set of disconnected, not always relevant examples of what they might do in their own classrooms.

Instructional Activities

As described in Chapter 1, the instructional activities are content-based learning activities designed to be routine and simple enough for new teachers to implement while teaching to rigorous content standards. Their simplicity and rigor make them a strong vehicle for lead instructors and coaches to train CMs how to teach according to the core practices and for CMs to practice teaching in their summer school classes.

The study team observed lead instructors, coaches, and CMs using some of the instructional activities. Due to the timing of the majority of Learning Cycle observations and of summer school teaching (weeks 1 and 3 of the institute training), not all of the instructional activities were observed.[8] The most observed instructional activity was "the interactive read-aloud," in which teachers read text with a group of students at any grade level and have a focused discussion around the key ideas and details in the text. Other instructional activities observed included "Quick Images" in elementary mathematics and "4 Corners," in which the corners of the room represent Strongly Agree, Agree, Disagree, and Strongly Disagree and students have to align themselves

[8]TFA adopted several instructional activities from the University of Washington. These can be found at tedd.org

in reaction to a statement of opinion about a math problem or text.[9] In place of some content-specific instructional activities during the first week, CMs teaching elementary school students learned routines to begin and end each day with their students. CMs teaching in secondary school learned "Starting 7," a protocol that describes what teachers should be doing during the first seven minutes of class.

The study team's main interest in observing instructional activities was to see how TFA training staff members and CMs were able to use them as a vehicle for getting at the core practices. The following section describes how they did so.

Core Practices

The core practices consist of strategies and routines that teachers can use to establish a productive learning environment, to enhance classroom dialogue, to engage students in the content, and to help students make meaning of the content.[10] For instance, Choral Counting is an instructional activity in which the teacher leads children in counting aloud together by a given number. While this activity may appear to be straightforward, as described in Chapter 1, a lot goes into making sure that the activity provides an opportunity for all students to engage in high-level thinking and to discuss this thinking with the class. This is where the core practices come into play.

In the case of Choral Counting, the teacher must deliver clear directions (student-centered directions), a strategy that falls under the core practice *creating and maintaining a productive learning environment*. Students need to understand what it means to count by a certain number and what their role is in the activity — when to count, when to pause, and how to signal to the teacher that they know why a particular number comes next in the series when asked. When asked why a particular number comes next, students need to feel that they can make mistakes — take academic risks — and this can only happen if the teacher has nurtured a culture of respect and safety, for instance, by having them discuss their thoughts with a peer before answering a question in front of the whole class (*positioning students as competent sensemakers*). For the activity to proceed smoothly, the teacher needs to build excitement, while simultaneously recognizing and reinforcing positive student behaviors and making decisions about how and when to redirect students who are off task (*creating and maintaining a productive learning environment*). Thus, the core practices guide the many decisions that teachers make to support their students in the learning process.

Lead instructors were directed to use their own judgment regarding which and how many core practices to feature in any given session. In addition, the lead instructors were not given specific instructions about how and when to incorporate the core practices into their sessions. Thus, the use or mention of the core practices varied considerably across Learning Cycle sessions.

[9]Teacher Education by Design (n.d.)

[10]Core Practice Consortium (n.d.)

Table 3.3 shows, by content area, the extent to which the core practices were mentioned during Learning Cycle sessions. The use of the word "mentioned" in the above sentence is intentional — connections made between the teaching of the instructional activities and the core practices were quick or perfunctory and usually made only at the beginning of the sessions. Core practices rarely came up in relation to specific strategies or routines once the sessions began. As shown in Table 3.3, the core practices were not mentioned at all in slightly more than half of the 17 sessions observed, while in some sessions more than one core practice was mentioned. *Creating and maintaining a productive learning environment* came up twice as many times (eight) as any other core practice, and no mentions of *teaching with society in mind* were made.

Table 3.3

Extent to Which Core Practices Were Mentioned During Learning Cycle Lessons, by Content Area

Core Practice	ELA	Math	Total
Creating and maintaining a productive learning environment	6	2	8
Teaching toward an instructional goal	4	0	4
Positioning students as competent sensemakers	3	`1	3
Teaching with society in mind	0	0	0
No core practice mentioned	4	5	9
Number of observations	10	7	17

SOURCE: Classroom observations conducted by MDRC.

It is not surprising that *creating and maintaining a productive learning environment* was the most often mentioned core practice and that the mention of the core practice *teaching with society in mind* (promoting equitable participation in the classroom; challenging social narratives that marginalize and oppress; understanding how identity, power, and privilege affect interactions with students; and so on) was not observed. The strategies and routines that together help create a productive environment are the concrete actions that teachers readily understand and know they should do right away — reinforce positive student behaviors, learn how to pronounce students' names correctly, give students clear directions, and so on. It may not have been as clear to a novice instructor how one would promote equitable participation in the classroom or challenge social narratives. Additionally, lead instructors did not cover *teaching with society in mind* because they thought that that was the responsibility of the social justice seminar leaders. Lead instructors in some of the focus groups in Tulsa shared that they did not feel prepared to train CMs on issues of diversity and social inequity and how they affect teaching. They also noted that the structure of the institute did not allow for time when they could meet and plan together with the social justice seminar leaders, whom they saw as the ones responsible for covering those issues.

In addition to the low number of times that the core practices were mentioned in observed sessions, the way in which they were mentioned also turned out to be an issue. Rather than focusing on any one particular core practice during a Learning Cycle session, instructors, coaches, and CMs seemingly randomly referenced core practices as they came up organically in a session. Since teaching is complex, this meant that, in any given Learning Cycle session, instructors, coaches, and CMs hypothetically could have pointed out and discussed many examples of each of the core practices. Although, as noted, core practices were not mentioned frequently, when they were mentioned, it appears to have been at the discretion of the lead instructor and there was never a focus on any one particular core practice. In other words, CMs did not have the opportunity to delve into any one of the core practices and explore the various ways that it could be expressed. Instead, what they received were disconnected examples of strategies and routines that could lead to the enactment of a mix of core practices.

Below are some typical examples from classroom observations conducted by the study team of CMs using the core practices while teaching during the third week of institute training. As expected, CMs more often implemented practices associated with *creating and maintaining a productive learning environment,* and, when they did, they usually did so at a very concrete and basic level.

Creating and Maintaining a Productive Learning Environment

- *Recognizing the positive:* CMs regularly recognized and reinforced positive student behaviors and choices. CMs did use strategies such as giving points to groups that finished first, asking the class to applaud a student for coming up to the board to solve a problem, and verbally praising specific actions (for example, with statements such as "I like that you're working diligently" and "Jose is indicating he is done by giving a thumbs up"), though CMs more consistently gave nonspecific praise (for example, "that's exactly right," "perfect," "good job"). At times, the approach CMs took was to react to negative behaviors and choices (for example, saying "stop whistling" or "sit up"). CMs were also regularly seen giving positive recognition for academic achievement. In rare cases, CMs were so enthusiastic about praising students that they praised all responses, even when they were inaccurate.

- *Sharing procedural and academic expectations:* CMs varied in their attempts to let students know what was expected of them. When it came to giving procedural directions, some CMs were very clear. They asked students to raise their hands if they needed help, restate in their own words what exactly the class was supposed to be doing, provide a thumbs up, and so on. Other times, the directions were not as clear or were provided inconsistently or infrequently. However, when it came to setting out expectations about how to engage in academic tasks, such as discussing ideas with their "turn and talk" partner or looking for evidence in the text, the directions were very basic, and students did not have a sense of what was being asked of them. For example, in one

math classroom, a CM asked the students to "match the scenario with the equations." It was clear that some students were not listening when the directions were given, but the CM did not ask anyone to reiterate directions and did not follow up and make sure students understood what they were supposed to be doing. In another math classroom, a CM asked students to get up from the classroom rug and work at their desks, instructing them to "go measure objects in this drawing using cubes."

Teaching Toward an Instructional Goal

- *Verbalizing and/or posting lesson objectives and instructional goals:* This behavior was rarely observed. In a couple of classrooms, lesson objectives were posted on the smartboard either throughout the entire lesson or temporarily as the lesson began. In one classroom, a CM reminded students what they should have been doing during the lesson as it unfolded, but the information was not posted anywhere or clearly stated from the start of the lesson. In another classroom, just the name of the activity was posted with a timeframe for completion ("Assessment: 20 minutes").

- *Clear alignment between the lesson content, instructional activities, and instructional goals:* In most classes, there was no evidence of what the instructional goal was. In some classes, the CM may have known how the instructional goal, the lesson content, and instructional activities were aligned, but this connection was lost on the students. For example, in one class, although the instructional goal was not shared with the students, the CM shared the goal with the observing member of the study team before the lesson and there was clear alignment between the lesson content, the activity, and the goal from the observing team member's perspective. However, this connection was lost on the students, as they did not know what the goal was, nor did they appear to understand the content of the lesson or what they were being asked to do to complete the activity.

Positioning Students as Competent Sensemakers

- *More student talk than teacher talk:* CMs appeared more comfortable lecturing than facilitating, as the study team observed more teacher talk than student talk. For example, in one class, students were asked to do a turn-and-talk activity with a partner, once and for about 30 seconds; the students otherwise had no opportunities to share their thoughts. In another class, students were asked to complete an activity in pairs, about five minutes before the bell rang; other than that, the CM did all the talking. Students were most likely to speak when the CMs asked them questions, and the majority of questions that CMs asked could be answered in one word.

- *Preparing students for and asking them higher-order questions:* CMs teaching in secondary school made more attempts than CMs teaching in elementary school to set students up to be able to answer higher-level questions in a variety of ways. In one secondary school classroom, the CM assessed students' prior knowledge by giving a mini-lesson on previously covered material before delving into the lesson. In another class, a CM restated her question to make it more accessible to students. In elementary school classrooms, however, this technique was rarely seen. For example, students in one class were told to write the answers to questions such as "Why did Unhei decide to use her own name at school?" and "What parts of the story help explain why she kept her name?" This activity followed a quick first read of the book without any discussion, so the students were disengaged. It was thus difficult for the CM to elicit a higher level of thinking from the students. The CM instead received and accepted inaccurate and one-word responses.

Teaching with Society in Mind

- *Making decisions to highlight the voice of students:* Only CMs teaching secondary school made some attempts to help students connect personally to content (for example, asking them to give examples from their own lives). In general, however, CMs across secondary school classrooms did not incorporate content that highlighted students' backgrounds or communities.

- *Promoting equity of voice in the classroom:* Some CMs solicited answers from students who were either (1) not visibly engaged or (2) raising their hands. Others called on a range of students, and still others called on every student individually to respond to one question.[11]

When asked during focus groups in Tulsa what they thought were the most important ideas they got from the institute training, CMs in only two out of eight focus groups brought up the core practices organically. Also, in six of eight focus groups, CMs suggested that the instruction they received did not prepare them to implement effective classroom management strategies. (For example, they struggled with maintaining a productive learning environment.)

The fact that CMs did not mention the core practices does not necessarily mean that lead instructors were not exposing them to these important practices. The strategies and routines that make up the core practices are what good teachers do as a matter of course. The lead instructors

[11]As they were written, some of the "teacher actions" included within "teaching with society in mind" were difficult for the study team to actually observe (even if they were happening in classrooms). For example, when a teacher may have made a decision with the "voice of students and parents" in mind (or when they did not do so, either intentionally or not intentionally), it may not have been visible to the observing team member. Moreover, even if CMs had society in mind when they developed their lesson plans, they may have struggled to illuminate that connection once executing those lessons. For example, CMs sometimes struggled to get students to make "text-to-self" and "text-to-world" connections. (Though, they may have understood why making those connections is critical to engaging students in the material.)

were incorporating some of these practices in their sessions, but they did not always refer to them as core practices, nor did they always point them out when they were incorporating them. As discussed above, CMs may well have seen the core practices' strategies and routines as random, disconnected ideas that they could incorporate into their teaching.

The lack of consistency in exposure to the strategies and routines that make up the core practices is important because *together* they provide guidance to teachers about how to do things such as *creating and maintaining a productive learning environment*. Having positive expectations of students is a necessary strategy for creating and maintaining a productive learning environment but it is not sufficient: Teachers must also build positive relationships with students, know how to redirect students who are off task, and so on. Had lead instructors received instructions about how and when to incorporate the strategies and routines into their sessions (that is, how to package them) and had CMs had the time to see, practice, and reflect on different ways that core practices could be embedded in an instructional activity, they might have felt that they had a stronger foundation in classroom management and other aspects of teaching to guide them in making informed and effective decisions than they reported having.

Chapter 4

Teachers' Experience During the School Year

This chapter explores the experience during the school year of the first cohort of Teach for America (TFA) corps members (CMs) to participate in the redesigned training offered at the Tulsa national institute. The first section examines CMs' perceptions of the overall value of their summer training, which provides a broad measure of CMs' satisfaction with their summer training as preparation for their school-year teaching placement. The second section explores the extent to which CMs used instructional strategies that are aligned with college- and career-readiness standards and that are generally considered to be effective teaching.[1] The third section examines the extent to which CMs used the core practices, which were a focus of the redesigned training in Tulsa. The fourth section considers CMs' perspectives on their first year of teaching, including their views on the concepts from the institute that they found useful during their first year of teaching, their perceptions of the support they received, and how prepared they felt to teach for a second year.

The chapter considers these aspects of CMs' experience during the school year for CMs trained at the national institute in Tulsa (program regions), as well as for CMs trained at the other national institutes (comparison regions). Doing so makes it possible to better understand how the experiences of CMs in these two groups differed, and to assess whether the redesigned training affected CMs' teaching experiences. The findings from the first two sections — which focus on CMs' perceptions of the value of the training and their use of standards-aligned practices — are based on a large sample of CMs and provide the most representative picture of the differences or similarities between the two groups' experiences. The findings from the last two sections — which look at CMs' use of the core practices and their perspectives in their first year of teaching — are based on qualitative data from a smaller sample of CMs; these findings provide a more nuanced picture of CMs' experiences during the school year.

Overall, the findings suggest that CMs who attended the national institute in Tulsa had a similar first-year teaching experience as CMs who attended the comparison institutes:

- CMs who received the redesigned training did not have more positive perceptions of the value of their summer training than CMs who attended the comparison national institutes.
- Although CMs in the program regions used standards-aligned instructional strategies in the classroom, they did not use these strategies more often than CMs in the comparison regions.
- By and large, CMs in the program regions and comparison regions used teaching strategies and routines related to the core practices similarly.

[1]Internal to TFA, these standards are called the "core components of instruction."

- CMs reported that what they learned about diversity, equity, and inclusiveness was the most useful aspect of the summer training. This was true for CMs in the program regions and comparison regions.
- CMs reported receiving the most support from their TFA managers of teacher leadership and development (MTLDs) and their coworkers.[2] This was true for CMs in both groups.
- At the end of their first year of teaching, the majority of CMs who were interviewed across both groups felt prepared for their second year of teaching.

These results — which suggest that CMs who were trained at the national institute in Tulsa did not have an appreciably different experience during the school year from that of other CMs — are consistent with the challenges of implementing a new (and ambitious) training model. The remainder of this chapter describes these findings in greater detail.

Perceptions of the Value of the Training

This section explores whether the redesigned training offered in Tulsa affected CMs' perceptions of the overall value of their summer training. The study team obtained information about CMs' perceptions from surveys that TFA regularly administers at different points during CMs' first year of teaching (beginning, middle, and end) and the fall of their second year. As explained in Chapter 2, these data allowed the study team to use a comparative interrupted time series (CITS) design to look at trends in perceptions of the training for CMs in the program and comparison regions. The findings from the CITS design cannot be definitively attributed to the causal effect of the redesigned training, but they can be interpreted as being strongly suggestive of the presence or absence of a causal effect.

Table 4.1 presents the results from the CITS design.[3] (See Box 4.1 for an explanation of how to read the findings in this table.)[4] Overall, the findings suggest that the redesigned training may have had a statistically significant negative effect on CMs' perceptions of the value of the summer training during their first year of teaching, but this effect ultimately dissipated in the fall of CMs' second year. Figure 4.1 illustrates the trends over time in CMs' perceptions of the training in the fall of their second year of teaching — the latest time point at which CMs can be followed — across consecutive cohorts of CMs. Before the new training was launched, CMs in both groups were "neutral" to "somewhat" in agreement on average that the summer training was

[2]MTLDs are TFA coaches.

[3]"Value of the summer training" was measured using a survey item that asks CMs to rate their level of agreement with the following statement: "Overall, I believe the preparation I received from Teach For America prior to starting at my school was valuable in my efforts to become a successful teacher." Agreement was rated on a 7-point scale, from strongly disagree (1) to strongly agree (7).

[4]This analysis is based on the survey sample, which is defined as the subset of CMs for whom data are available on the key survey outcomes for a given survey wave. See Chapter 2 for more information.

Table 4.1

Estimated Effect of the Redesigned Training on Perceptions of the Value of the Summer Training, School-Year Surveys

Outcome	Mean Outcome (Baseline Cohort)		Deviation from Baseline Trend		Estimated Effect	Effect Size	P-Value for Estimated Effect
	Program Regions	Comparison Regions	Program Regions	Comparison Regions			
Fall of first year							
Value of the summer training (1-7)	4.72	4.80	-0.51 ††	0.12	-0.62 **	-0.35	0.020
Number of regions			8	13			
Number of corps members			2510	4811			
Middle of first year							
Value of the summer training (1-7)	4.62	4.65	-0.78 †††	0.23	-1.01 ***	-0.57	0.000
Number of regions			8	13			
Number of corps members			2401	4578			
End of first year							
Value of the summer training (1-7)	4.49	4.53	-0.46 ††	0.13	-0.58 **	-0.32	0.015
Number of regions			8	13			
Number of corps members			2286	4351			
Fall of second year							
Value of the summer training (1-7)	4.61	4.68	-0.37 ††	-0.03	-0.34	-0.19	0.114
Number of regions			8	13			
Number of corps members			2193	4112			

(continued)

Table 4.1 (continued)

SOURCE: Teach For America corps member surveys.

NOTES:
The "Value of the summer training" measure is based on CMs' agreement with the following question: "Overall, I believe the preparation I received from Teach For America prior to starting at my school was valuable in my efforts to become a successful teacher." Agreement is rated on a 7-point scale with 1 equal to "strongly disagree" and 7 equal to "strongly agree."

A cohort is defined as a group of TFA CMs who joined TFA and attended the summer institute the same year. The "baseline cohorts" (2012-2015) received their summer training before the launch of the redesigned institute in Tulsa. The "redesigned institute cohort" (2016) is the first cohort of CMs trained after the launch of the redesigned institute. The number of CMs represents the total number of CMs across all cohorts who are in the survey sample for a given follow-up period.

The values in the columns labeled "Mean Outcome (Baseline Cohort)" are the average outcomes of the CMs in the progam and comparison regions who were trained in summer 2015, before the new training was launched. The "Deviations from Baseline Trend" are the difference between the actual and predicted outcomes for the 2016 CMs in the progam and comparison regions. The values in the "Estimated Effect" column are the difference between the program regions and comparison regions with respect to their deviations from baseline trend. All values are regression-adjusted for between-region and between-cohort differences in the characteristics of CMs and their school placement. Values in the "Effect Size" column are the estimated effect divided by the standard deviation for the last baseline cohort. Rounding may cause slight discrepancies in calculating sums and differences.

A two-tailed t-test was applied to estimated deviations and estimated differences between program regions and comparison regions. The statistical significance of estimated deviations is indicated as follows: ††† = 1 percent; †† = 5 percent; † = 10 percent. The statistical significance of estimated effects is indicated as follows: *** = 1 percent; ** = 5 percent; * = 10 percent.

valuable (an average rating of 4 to 5), with satisfaction levels having decreased slightly over time (as indicated by the downward slope of the trend lines in Figure 4.1). After the redesigned training was launched, CMs in the 2016 cohort in the comparison regions rated the summer training as would be expected given prior trends. (The deviation from trend for this group is not statistically significant.) In contrast, CMs in the program regions rated the summer training as less valuable than had prior cohorts of CMs. However, the estimated effect of the training — defined as the difference between the deviation from trend for the program regions and comparison regions — is not statistically significant. Therefore, it cannot be concluded that the redesigned training affected CMs' perceptions of the value of the summer training in the fall of their second year.[5]

There are two possible explanations for this pattern of effects. First, as discussed in Chapter 3, CMs who attended the redesigned training at the national institute in Tulsa did not get as much explicit help with lesson planning and classroom management as did the CMs in the comparison institutes during the summer training, which may explain why they initially perceived the summer training as less valuable than did earlier cohorts of CMs and CMs in the comparison regions. However, by their second year, CMs may have figured out how to perform these tasks, which in turn may have improved their perceptions of the training. This statement is supported by interviews

[5]The statistical significance of estimated effects during CMs' first year is consistent across several sensitivity analyses. The reduction in the magnitude of this effect in CMs' second year (relative to their first year) is also consistent across sensitivity analyses. However, the estimated effect of the training in the fall of CMs' second year is statistically significant in one of several sensitivity analyses conducted, though only at the 10 percent level. See Appendix F for more information on these findings.

Box 4.1

How to Read the Tables in This Chapter That Present Estimated Effects on CMs' Perceptions, Commitment, and Retention

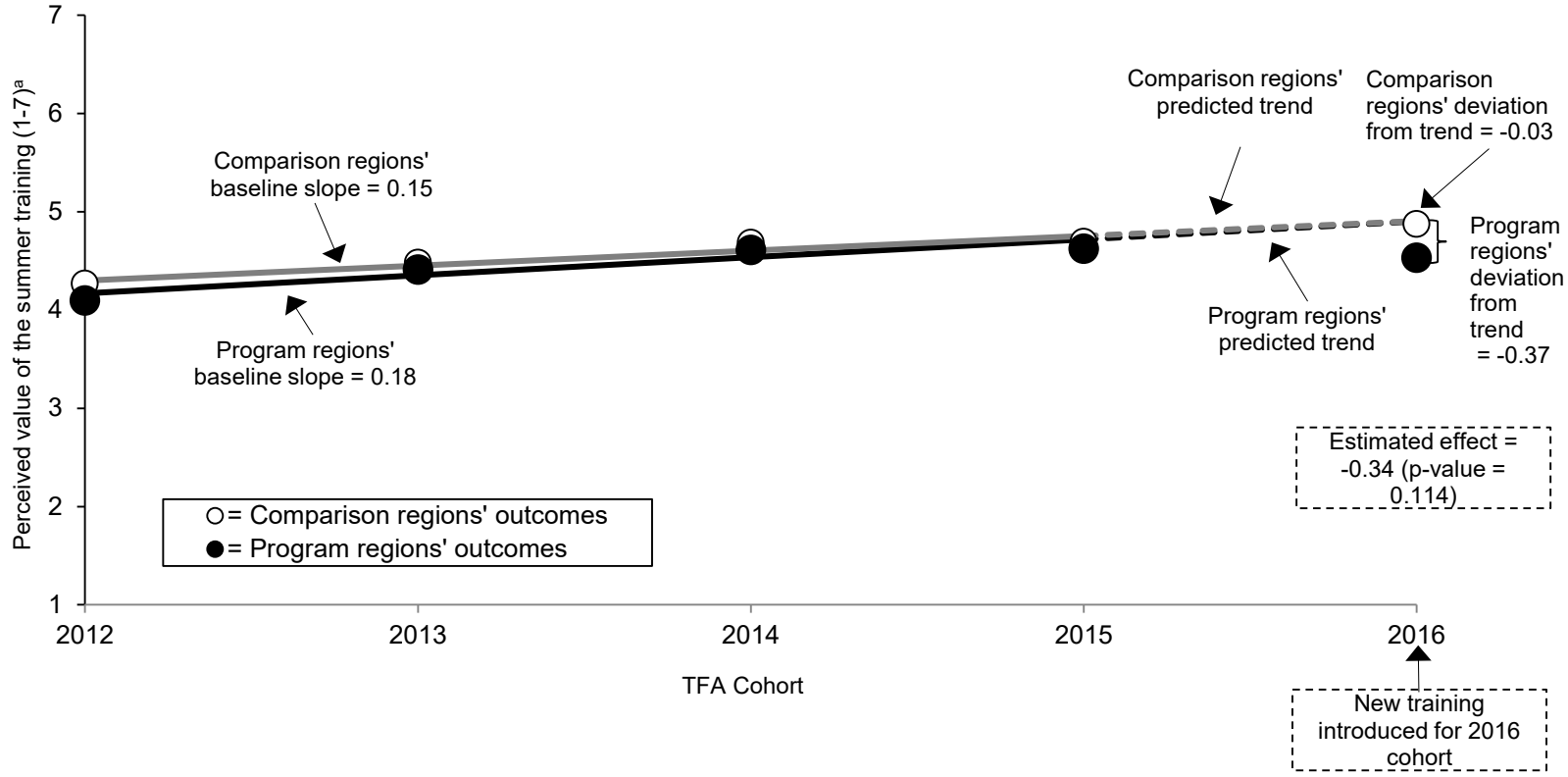
Some of the tables in this report show the estimated effect of the redesigned training on CMs' perceptions of the training, their commitment to teaching, and their retention. The *columns* in these tables present the following information:

- **Mean outcome (baseline cohort).** The first two columns of results show the mean outcomes for the 2015 cohort of CMs, who were trained in the summer before the redesigned training was launched. The mean outcomes are shown to provide context for interpreting the magnitude of the estimated of effects.
- **Deviation from baseline trend.** The next two columns of results show the estimated deviation from their baseline trend for the program regions and the comparison regions. This deviation is the difference between the observed (actual) outcomes and the predicted outcomes for the 2016 cohort of CMs in each group of regions. Positive deviations from trend indicate that CMs in the program regions or the comparison regions had more positive outcomes than expected given prior trends, while negative deviations from trend indicate that CMs had more negative outcomes than expected. The statistical significance of estimated deviations from trend is indicated for levels of 1 percent (†††), 5 percent (††), and 10 percent (†).
- **Estimated effect and effect size.** The next column of results is the estimated effect of the redesigned training piloted at the national institute in Tulsa. The estimated effect is the difference between the deviations from baseline trend for the 2016 cohort of CMs in the program regions and the comparison regions. For outcomes whose scale is more challenging to interpret, the estimated effect is also shown as an effect size. The effect size is equal to the estimated effect divided by the standard deviation of that outcome for the last baseline cohort of CMs (the 2015 cohort).
- **P-value for estimated effect.** The last column of the table shows the p-value for each estimated effect. The p-value indicates the probability of finding an estimated effect of that magnitude (or a larger effect) if there were in fact no difference between the program regions and the comparison regions with respect to their deviations from their trends. In this study, an estimated effect that has a p-value of 10 percent or less is considered "statistically significant" because it is unlikely that this estimated effect would be observed if the program regions and the comparison regions did not have different deviations from trend. The number of asterisks indicates whether the estimated effect is statistically significant at the 1 percent (***), 5 percent (**), or 10 percent (*) level.

The *rows* in the tables represent the outcome — and in some cases the follow-up period (for example, end of summer, fall of first year of teaching, end of first year of teaching, and fall of second year of teaching) — that is associated with each estimated deviation from baseline trend and each estimated effect.

Figure 4.1

Trends in the Perceived Value of the Summer Training, Fall of Second-Year Survey



SOURCE: Teach For America corps member surveys.

NOTES: [a]The reported outcome measures are based on survey items with a 7-point agreement scale. The "value of the summer training" measure is based on the following question: "Overall, I believe the preparation I received from Teach For America prior to starting at my school was valuable in my efforts to become a successful teacher."

conducted with CMs in the spring of their first year of teaching. By then, CMs rarely mentioned challenges related to lesson planning and classroom management, which suggests that they had become more comfortable with these tasks than they were initially.

Second, the coaching received by CMs in the program regions during the school year may have been misaligned with the redesigned summer training. As discussed later in this chapter, CMs received support from different TFA staff members during the school year. In general, however, this support was not categorically linked to the concepts presented at the institute in Tulsa, and this misalignment may have affected many CMs' perceptions of the value of their summer training.[6] By the fall of their second year, however, these negative effects had dissipated, perhaps because CMs started to rely more on what they learned about teaching during the school year, and less on what they learned during their summer training.

Instruction Aligned with College- and Career-Readiness Standards

This section examines CMs' use of instructional practices that were aligned with college- and career-readiness standards. As explained in Chapter 1, the redesigned training in Tulsa introduced CMs to content-specific instructional activities, whereas the training in the other national institutes was content neutral and more focused on classroom management and lesson planning. Therefore, one might expect the redesigned training to have increased CMs' use of instructional strategies aligned with college- and career-readiness standards. This section explores whether CMs in the program regions used standards-aligned strategies and whether they did so more often than CMs in comparison regions. As noted previously, these differences should not be attributed to the causal effect of the redesigned training because the study design cannot account for unobserved preexisting differences between the two groups of CMs. Instead, the findings should be used to identify patterns indicating that the redesigned training may have had an effect.

Information about CMs' instructional strategies comes from closed-ended logs that the study team administered every other week to CMs in the 2016 cohort teaching English language arts (ELA), mathematics, or general education during the first year of their placement.[7] The items in the closed-ended logs reflect strategies that are commonly recognized as effective instructional

[6]It is possible that regional TFA staff in some program regions may have made on-the-ground decisions to link training content during the school year with key concepts from the redesigned training in the summer, but these decisions would not have happened as a result of strategic linking or directives from national TFA staff.

[7]See Appendix C for further information about how the log items map onto different standards. As explained in Chapter 2, the analysis of these logs was based on the teacher log sample, which includes the subset of CMs in the 2016 cohort who were teaching ELA, mathematics, or general education and who completed at least one closed-ended log during the year. For the purposes of the analysis, the log sample is divided by content area (ELA and mathematics). There are 225 CMs in the ELA sample and 231 CMs in the math sample; CMs teaching general education are included in both samples.

Table 4.2

Estimated Effects on Self-Reported Instructional Strategies Aligned with College- and Career-Readiness Standards

Outcome	Program Regions	Comparison Regions	Estimated Difference	Effect Size	P-Value
English language arts (% of classes)					
Text used for the lesson featured an author, characters, and/or a community with similar background to focal student	38.3	49.2	-10.9	-0.21	0.118
Students were asked to cite text evidence in their writing[a]	75.8	81.2	-5.4	-0.12	0.447
Number of regions	8	12			
Number of corps members	87	138			
Number of logs	739	1,143			
Mathematics					
Math instructional practice composite (z-score)[b]	-0.15	-0.04	-0.11	-0.12	0.415
Number of regions	8	13			
Number of corps members	80	151			
Number of logs	628	1,197			

(continued)

strategies.[8] For the analysis, the study team prioritized three confirmatory instructional strategies that are most aligned with the redesigned training: using culturally responsive texts (cultural responsiveness domain), asking students to cite text evidence (ELA domain), and using standards-aligned math practices (a composite measure in the math domain).[9] Table 4.2 shows the extent to which CMs in the program and comparison regions used these standards-aligned instructional strategies during their first year of teaching.

Overall, it does not appear that the redesigned training increased CMs' use of standards-aligned instructional strategies.[10] CMs in the program regions did use teaching strategies aligned with college- and career-readiness standards; however, they did not use them more frequently than did CMs in the comparison regions. Across all three measures, CMs in the program and

[8]The math logs were designed based on the Common Core State Standards for mathematics, and the ELA log items were based on TFA's Core Components of Instruction, which themselves were based on the Common Core State Standards for ELA.

[9]Citing text evidence was selected as the primary measure for the ELA domain because it is the most aligned with comprehension.

[10]These findings are robust across several sensitivity analyses. See Appendix F for further information.

Table 4.2 (continued)

SOURCE: Closed-ended teacher logs administered by MDRC.

NOTES: Online teacher logs were administered to corps members (CMs) in the 2016 cohort in the program and comparison regions. CMs teaching English Language Arts (ELA), mathematics, or general education in the first year of their teaching placement were eligible for the logs. Secondary school teachers, as well as elementary school teachers in departmentalized schools, were sent the log for their content area and level every two weeks (about twice per month). Elementary school teachers teaching both content areas (general education teachers) were sent the ELA log or the math log every two weeks, with the content area alternating between logs (about one log per content area per month). At the start of each log, teachers were given a randomly selected letter of the alphabet and asked to report on their instructional practices with a focal student whose name starts with that letter (or the closest letter) on a particular day that week.

The values in the "Program Regions" column are the observed mean outcome for teachers in the program regions. The values in the "Comparison Regions" column are the regression-adjusted mean outcome for teachers in the comparison regions, using the mean covariate values for the program region as the basis for the adjustment. The values in the "Estimated Difference" column are the difference between the program and comparison regions, adjusted for differences in the characteristics of the focal students and teachers. Values in the "Effect Size" column are the estimated effect divided by the standard deviation for the sample.

A two-tailed t-test was applied to estimated differences. The statistical significance of estimated differences is indicated as follows: *** = 1 percent; ** = 5 percent; * = 10 percent.

Rounding may cause slight discrepancies in calculating sums and differences.

Sample sizes in this table are the number of teachers in the log sample. A CM is included in the log sample if they completed at least one log during the school year.

[a]This item is only available for upper elementary and secondary school teachers. The number of teachers included in this analysis is 60 in the program group and 93 in the comparison group.

[b]This composite combines the items in the math logs that map onto college- and career-readiness standards for math instruction. The scale of the items differs, so each item was first z-scored based on the mean and standard deviation of the comparison group logs; these z-scores were then averaged across items, and the average score was z-scored, again, based on the mean and standard deviation for the comparison logs. Because some of the log items differ across levels, the composite score was created separately for the lower elementary school and upper elementary or secondary school logs. The reliability (Cronbach's alpha) for the composite is 0.71 at the lower elementary school level and 0.75 at the upper elementary or secondary school level.

comparison regions did not differ in their use of these strategies by a statistically significant amount. For example, CMs in the program regions used culturally responsive texts in 38 percent of their classes, compared with 49 percent of classes that CMs in the comparison regions taught — a difference that was not statistically significant. As discussed earlier, CMs in the program regions struggled with lesson planning and classroom management when they started teaching in their regions, which may explain why they were not able to use standards-aligned instructional strategies with greater frequency than CMs in the comparison regions.[11] The next section explores these factors and others, as well as describes CMs' experiences during the school year using qualitative data collected for the study.

[11]Appendix C examines CMs' responses to additional closed-ended ELA log items, as well as CMs' responses to the individual log items included in the math composite measure. In general, these exploratory findings support the general conclusion that CMs in the program and comparison regions used standards-aligned teaching strategies at similar levels.

Use of Core Practices During the School Year

This section examines CMs' use of the core practices, which were a central component of the redesigned training in Tulsa. To capture their use of the core practices during their first year of teaching, the study team asked CMs in the 2016 cohort to complete open-ended teacher logs every other week during the school year.[12] Each open-ended log focused on a particular TFA core practice, asking CMs to describe the instructional practices they used that week that served that core practice.

Because the study team sent these logs to CMs in both program and comparison regions, the term "core practices" did not appear in the wording of any of the questions, and none of the questions explicitly mentioned or described any of the four core practices. Instead, similar to the observational tool that the study team used during the summer, the log items about core practices were phrased in terms of specific examples of strategies and routines used by TFA.

Overall, CMs' responses to the open-ended logs indicate that there was not a large difference between the CMs in the program and comparison regions in terms of their use of the core practices. There were, however, some exceptions for each core practice, which are discussed in the remainder of this section.

Creating and Maintaining a Productive Learning Environment

Teachers who *create and maintain a productive learning environment* have classrooms in which their students feel supported and cared for while being held accountable for learning. CMs in the program and comparison regions were similarly likely to report using strategies and routines that fell under the *creating and maintaining a productive learning environment* core practice.[13] When asked how they responded to unengaged students, CMs reported that they were most likely to do so verbally (46 percent of CMs on average across both groups). They also tried to reengage students by calling on them or standing close to them (21 percent), giving them special assignments (18 percent), and issuing warnings before delivering a consequence (17 percent). Only a small proportion of CMs tried to reengage students by changing how content was delivered, tweaking questions, or switching partners (13 percent). CMs in the program regions were more likely than CMs in the comparison regions to explain to students the reasons for giving them praise (38 percent versus 28 percent).

[12]These logs were sent to CMs who consented to participate in this data collection.

[13]Responses differed by 5 percent at most between CMs in the program and comparison regions. Statistical significance testing was not conducted because these findings are descriptive and exploratory. The log items pertaining to this core practice are the following: (1) "Please describe an example of your response to a student(s) (what you said or did) for a task(s) that was done well this week. Why did you respond in this way? What, if any, were the effects of responding in this way?" and (2) "Please describe an example of your response to a student(s) (what you said or did) who was not engaged in a task(s) you assigned this week, or the last week you taught. What, if any, were the effects of responding in this way? Why do you think this student(s) was unengaged in the task?"

Teaching Toward an Instructional Goal

Teachers who *teach toward an instructional goal* understand why the content they teach matters, both in terms of the larger purpose of their content area and in terms of their students' lives outside of school. To that end, CMs generally reported taking the same types of things into consideration when selecting content to teach, regardless of which summer institute they attended.[14]

The desire to promote specific skill sets or content most influenced CMs in their decisions about what content to teach (for example, their choice of a text or a problem to teach): 49 percent of CMs on average across both groups were motivated by this desire. For example, one CM reported that one of the main reasons for choosing a particular text was because it promoted reading skills: "This week I taught The Ugly Duckling because of its social-emotional lesson about how words can make people feel, and also because it can easily be adapted into a play that helps my students practice fluency, reading with expression, and fast reading."

Students' backgrounds, their communities, current events, or an interest in making content relevant and applicable to the real world also influenced CMs' decisions about what material to teach (33 percent of CMs on average across both groups).[15] For example, one CM said, "I chose to teach 'Separate Is Never Equal' because it's a story of school desegregation before *Brown v. Board of Education* and is about a Mexican family in California. I thought it was important history for the kids in my class, [who are] 98 percent Mexican, to know about."

Reasons less commonly cited for choosing content to teach included selecting material required by the school or district (19 percent of CMs on average across both groups) and differentiating or scaffolding instruction in areas where CMs had identified that their students needed extra help (17 percent).

Positioning Students as Competent Sensemakers

This core practice is in part about creating multiple opportunities and methods for students to think and express themselves.[16] Across both program and comparison regions, CMs reported using a variety of methods for *positioning students as competent sensemakers*, of which the following were the most common: facilitating whole class discussions (71 percent of CMs on average across both groups); asking students to share their ideas in writing for diverse purposes,

[14]The log item pertaining to this core practice is the following: "Please list all the ways/methods your students were able to share their responses and thoughts with you or with others in your class this week or the last week you taught."

[15]This category also included instances when CMs cited influence related to classroom culture. For example, one CM said "we read about fighting because there has been fighting in the classroom." It is worth noting that the week the study team asked this question was the week before Martin Luther King Day, so a number of CMs mentioned material related to Martin Luther King in their response.

[16]The log item pertaining to this core teaching practice is the following: "please think about one text or problem you incorporated into one of your lessons this week, or the last week you taught. what influenced your decision to select that text/problem?"

such as through graphic organizers, short essays, and so on (65 percent); and asking students to "turn and talk" with their partners (54 percent).

Overall, CMs in both program and comparison regions used similar methods for getting students to share responses.[17] Because the core practices as a whole were not strategically or systematically emphasized to CMs during the institute in Tulsa, it makes sense that CMs in the program regions did not use the core practices more frequently.

CMs' Perspectives on Their First Year of Teaching

This section examines CMs' perspectives on their first year of teaching, based on phone interviews that the study team conducted with CMs at the end of their first year of teaching. In total, the team interviewed 39 CMs (18 in the program regions and 21 in the comparison regions) by phone in spring 2017. The team asked CMs about the concepts, ideas, or practices from the institute they found useful in their first year of teaching; who supported them; and how prepared they thought they were to teach for a second year. The findings are discussed below.

What Concepts Learned at Institute Did CMs Find Most Useful?

During the phone interviews, CMs most often mentioned the social justice concepts of the training as being useful. CMs in both program and comparison regions (36 percent on average) indicated that what they learned about social justice was useful as a framework for lesson or curriculum planning or meeting students' academic needs, for working and interacting with students generally, and for working with students of color.

CMs rarely mentioned other aspects of the summer training during the interviews. During the summer focus groups, the two key challenges that CMs discussed most were related to lesson planning and classroom management. However, CMs rarely brought up these topics in interviews during the school year. On average across both groups, only 13 percent of CMs said that the main concept or skill they took from the institute was lesson planning. A slightly greater proportion of CMs in the program regions reported feeling unprepared to write their own lesson plans during the school year, relative to CMs in the comparison regions. As explained in Chapter 1, CMs in Tulsa wrote their own summer school lesson plans only during the last week of the institute training, whereas comparison CMs wrote their own lesson plans for the whole time that they taught summer school (weeks 2 to 5). Similarly, very few CMs in either program or comparison regions indicated that the overall training they received in classroom management informed their practices throughout the year. That said, however, 29 percent of CMs in the comparison regions said that positive narration, a specific strategy related to classroom management they had learned, was useful.[18] (CMs in Tulsa were not taught this strategy, as they learned the core teaching practices instead.) The CMs in the program regions also rarely mentioned the core teaching practices as something that stuck with them.

[17]Differences between CMs in program and comparison regions are between 1 and 9 percent.

[18]Positive narration involves pointing out positive behaviors, as opposed to negative behaviors.

One reason why many of the concepts from summer institute may not have stuck with CMs is because once CMs went to their regions, they entered a new and immediate reality, one that required them to meet different demands and expectations than those at institute (for example, writing lesson plans with specific state and district standards in mind). Expectations were different about what they should teach and how they should teach it, and they received different types and levels of support in their regions. Indeed, during end-of-year interviews, 41 percent of CMs across both groups reported that they felt that there was no alignment between the expectations at the institute and those between TFA and district staff in their regions, and only 28 percent thought there was alignment.[19]

Who Supported Them?

CMs in both program and comparison regions were most appreciative of the support they received from their TFA manager of teacher leadership and development (MTLD) during their first year of teaching (54 percent on average across both groups), as opposed to other district and school-specific staff members such as principals or school-based coaches. MTLDs provide individualized coaching, support, and guidance throughout CMs' two-year teaching commitment. CMs described their MTLDs as responsive and as providers of good advice and helpful techniques. They were also emotionally supportive. Other than MTLDs, CMs found their coworkers to be helpful (41 percent). Twenty-three percent of CMs indicated that their administrator or principal was supportive, and only 21 percent mentioned a non-TFA school coach or mentor as being supportive.

How Prepared Did CMs Feel for Their Second Year of Teaching?

At the end of the first year of teaching, the majority of CMs said that they felt prepared for their second year (62 percent on average across both groups) because they had a year of teaching under their belts, had developed a good foundation on which to build, and were familiar with the content or curriculum they would use. CMs in both groups expressed similar feelings of preparedness.

[19]A few CMs reported that they learned broad concepts at institute and that they learned specific strategies on how to apply the broader concepts in their regions. Other CMs shared that their schools or their teaching assignment in their schools were too different from what they experienced at institute for the latter to have been relevant.

Chapter 5

Corps Members' Commitment to Teaching and Their Retention

This chapter examines whether the redesigned summer training piloted at the national institute in Tulsa shows promise for improving the longer-term outcomes of the first cohort of Teach For America (TFA) corps members (CMs) to participate in the new training. The first section looks at CMs' commitment to teaching and educational equity and the second section examines CMs' retention rates in the TFA program. The effect of the redesigned training on these outcomes was estimated using a comparative interrupted time series (CITS) design. As explained in Chapter 2, the findings from the CITS design cannot be definitively attributed to the causal effect of the redesigned training, but they can be interpreted as being strongly suggestive of the presence or absence of a causal effect.

The key findings discussed in this chapter are that the *redesigned training does not appear to have improved CMs' commitment to teaching and educational equity. Nor does it appear to have improved their retention rates in the program*. These results are consistent with the challenges of implementing a new training model. As discussed in previous chapters, the content-focused approach piloted at the national institute in Tulsa was not only ambitious in scope but also a notable departure from TFA's traditional approach to training its CMs. The implementation challenges encountered — and the associated lack of effects on CMs' outcomes — are consistent with the complexity of the new training model and the fact that it was being tested for the first time in summer 2016. Viewed through a longer-term lens, the findings in this chapter are encouraging because they suggest that TFA was able to radically change its training approach without adversely affecting the first cohort of CMs to experience it, and, in theory, this training may yet produce positive outcomes for future cohorts of CMs who will receive a more mature version of the redesigned training.

Commitment to Teaching and Educational Equity

This section examines CMs' self-reported commitment to teaching and educational equity, based on surveys that TFA administers regularly to its CMs at different points during their first and second years of teaching. CMs' commitment to teaching and equity is a composite measure created from eight items about their beliefs in TFA's mission of ensuring that each child has an opportunity to receive an excellent education, and their beliefs in their own capacity to effectively contribute to that mission. The composite scale is based on a seven-point agreement scale. Table 5.1 looks at

Table 5.1

Estimated Effect of the Redesigned Training on Commitment to Teaching and Educational Equity, School-Year Surveys

Outcome	Mean Outcome (Baseline Cohort)		Deviation from Baseline Trend		Estimated Effect	Effect Size	P-Value for Estimated Effect
	Program Regions	Comparison Regions	Program Regions	Comparison Regions			
Fall of first year							
Commitment to teaching and equity (1-7)	5.83	5.85	-0.06	-0.04	-0.01	-0.02	0.920
Number of regions			8	13			
Number of corps members			2,510	4,811			
Middle of first year							
Commitment to teaching and equity (1-7)	5.67	5.78	0.21	0.21 †	0.00	0.00	0.999
Number of regions			8	13			
Number of corps members			2,401	4,578			
End of first year							
Commitment to teaching and equity (1-7)	5.82	5.89	-0.02	0.07	-0.09	-0.11	0.493
Number of regions			8	13			
Number of corps members			2,286	4,351			
Fall of second year							
Commitment to teaching and equity (1-7)	5.91	5.93	-0.08	-0.02	-0.06	-0.07	0.641
Number of regions			8	13			
Number of corps members			2,193	4,112			

(continued)

Table 5.1 (continued)

SOURCE: Teach For America corps member surveys.

NOTES:
The commitment to teaching and equity measure is a composite created from 8 items related to corps members' (CM) beliefs in TFA's mission of ensuring that each child has an opportunity to receive an excellent education, and CMs' beliefs in their own capacity to effectively contribute to this mission. The items and the composite scale are based on a 7-point agreement scale with 1 equal to "strongly disagree" and 7 equal to "strongly agree."

A cohort is defined as a group of TFA CMs who joined TFA and attended the summer institute the same year. The "baseline cohorts" (2012-2015) received their summer training before the launch of the redesigned institute in Tulsa. The "redesigned institute cohort" (2016) is the first cohort of CMs trained after the launch of the redesigned institute. The number of CMs represents the total number of CMs across all cohorts who are in the survey sample for a given follow-up period.

The values in the columns labeled "Mean Outcome (Baseline Cohort)" are the average outcomes of the CMs in the progam and comparison regions who were trained in summer 2015, before the new training was launched. The "Deviations from Baseline Trend" are the difference between the actual and predicted outcomes for the 2016 CMs in the progam and comparison regions. The values in the "Estimated Effect" column are the difference between the program regions and comparison regions with respect to their deviations from baseline trend. All values are regression-adjusted for between-region and between-cohort differences in the characteristics of CMs and their school placement. Values in the "Effect Size" column are the estimated effect divided by the standard deviation for the last baseline cohort. Rounding may cause slight discrepancies in calculating sums and differences.

A two-tailed t-test was applied to estimated deviations and estimated differences between program regions and comparison regions. The statistical significance of estimated deviations is indicated as follows: ††† = 1 percent; †† = 5 percent; † = 10 percent. The statistical significance of estimated effects is indicated as follows: *** = 1 percent; ** = 5 percent; * = 10 percent.

the estimated effect of the redesigned training on this measure, during CMs' first year of teaching (beginning, middle, and end) and in the fall of their second year.[1]

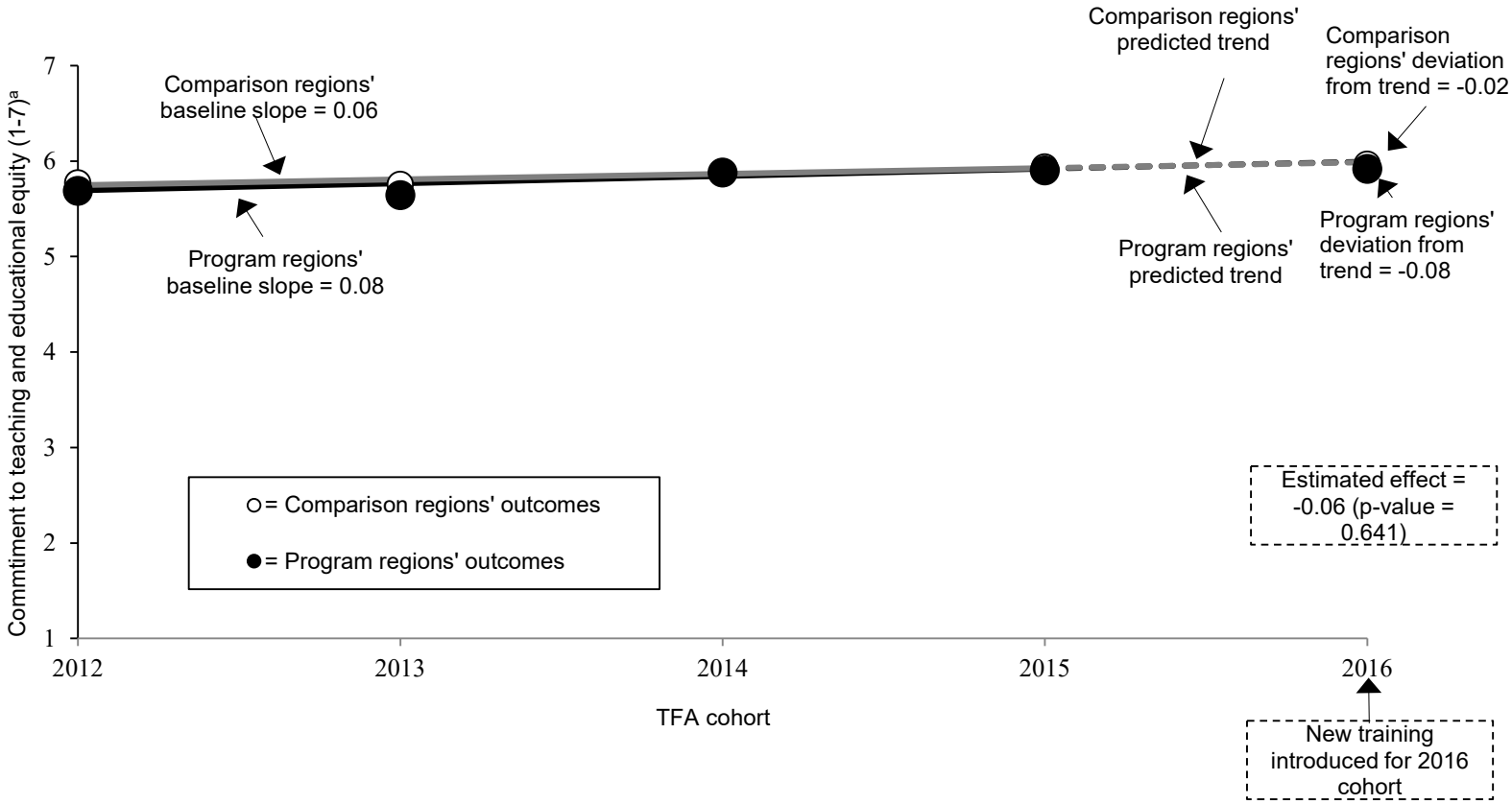
Overall, the findings suggest that the redesigned training did not have an effect on CMs' commitment to teaching and educational equity, either during their first year of teaching or in the fall of their second year. Figure 5.1 shows the trends over time in CMs' commitment in the fall of the second year of their placement, across consecutive cohorts. Before the new training was launched (2012 to 2015 cohorts), CMs in both groups generally agreed on average that they were committed to teaching and that the summer training was valuable (an average rating of almost 6), with satisfaction levels having increased slightly over time (as indicated by the upward slope of the trend lines in Figure 5.1). After the redesigned summer training was launched, CMs in the 2016 cohort in both the program and comparison regions rated their commitment to teaching and equity at a similar level as would be expected given prior trends. (The deviation from trend for this group is not statistically significant.) Thus, it cannot be concluded that the redesigned training affected CMs' commitment to teaching and equity. This is true at all four time points when this outcome was measured.[2]

[1]As discussed in Chapter 2, this analysis was based on the survey sample, which is defined as the subset of CMs for whom data were available on the two key survey outcomes for a given survey wave. Given the timeline for the project, CMs in all cohorts could be followed to the fall of their second year of teaching.

[2]These findings are robust across several sensitivity analyses. See Appendix F for further information.

Figure 5.1

Trends in Commitment to Teaching and Educational Equity, Fall of Second-Year Survey



SOURCE: Teach For America corps member surveys.

NOTES: [a]The reported outcome measures are based on survey items with a 7-point agreement scale. The commitment to teaching and equity measure is a composite created from 8 items related to CMs' beliefs in TFA's mission of ensuring that each child has an opportunity to receive an excellent education, and CMs' beliefs in their own capacity to effectively contribute to this mission.

Table 5.2

Estimated Effect of the Redesigned Training on Retention Rates

Follow-up Point	Mean Outcome (Baseline Cohort)		Deviation from Baseline Trend		Estimated Effect	P-Value for Estimated Effect
	Program Regions	Comparison Regions	Program Regions	Comparison Regions		
Fall of first year (%)	93.7	96.5	-1.3	-0.7	-0.7	0.801
End of first year (%)	83.8	87.5	0.0	-2.5	2.5	0.518
Fall of second year (%)	80.2	82.1	0.5	-1.6	2.1	0.641
Number of regions			8	13		
Number of corps members			2,819	5,505		

SOURCE: Teach For America administrative records for corps members.

NOTES:

A cohort is defined as a group of TFA corps members (CMs) who joined TFA and attended the summer institute the same year. The "baseline cohorts" (2012-2015) received their summer training before the launch of the redesigned institute in Tulsa. The "redesigned institute cohort" (2016) is the first cohort of CMs trained after the launch of the redesigned institute. The number of CMs represents the total number of CMs across all cohorts.

The values in the columns labeled "Mean Outcome (Baseline Cohort)" are the average outcomes of the CMs in the progam and comparison regions who were trained in summer 2015, before the new training was launched. The "Deviations from Baseline Trend" are the difference between the actual and predicted outcomes for the 2016 CMs in the progam and comparison regions. The values in the "Estimated Effect" column are the difference between program regions and comparison regions with respect to their deviations from baseline trend, adjusted for between-region and between-cohort differences in the characteristics of CMs and their school placement. Rounding may cause slight discrepancies in calculating sums and differences.

A two-tailed t-test was applied to estimated deviations and estimated differences between program regions and comparison regions. The statistical significance of estimated deviations is indicated as follows: ††† = 1 percent; †† = 5 percent; † = 10 percent. The statistical significance of estimated effects is indicated as follows: *** = 1 percent; ** = 5 percent; * = 10 percent.

Corps Member Retention

This study also examines whether the redesigned summer training had an effect on CMs' retention rates in the TFA program. (See Table 5.2.) Whether a CM is still a TFA teacher is measured at three points: the fall of CMs' first year of teaching, the end of their first year, and the fall of their second year.[3] Overall, the findings suggest that the redesigned summer training does not appear to have had an effect on CMs' retention rates. Figure 5.2, for example, shows that before the redesigned training was launched (2012 to 2015 cohorts), second-year retention rates had been generally declining in the program and comparison regions. In both groups, the 2016 cohort's retention rates did not deviate visibly or statistically from this baseline trend.[4] This is true at all three time points when retention was measured.[5]

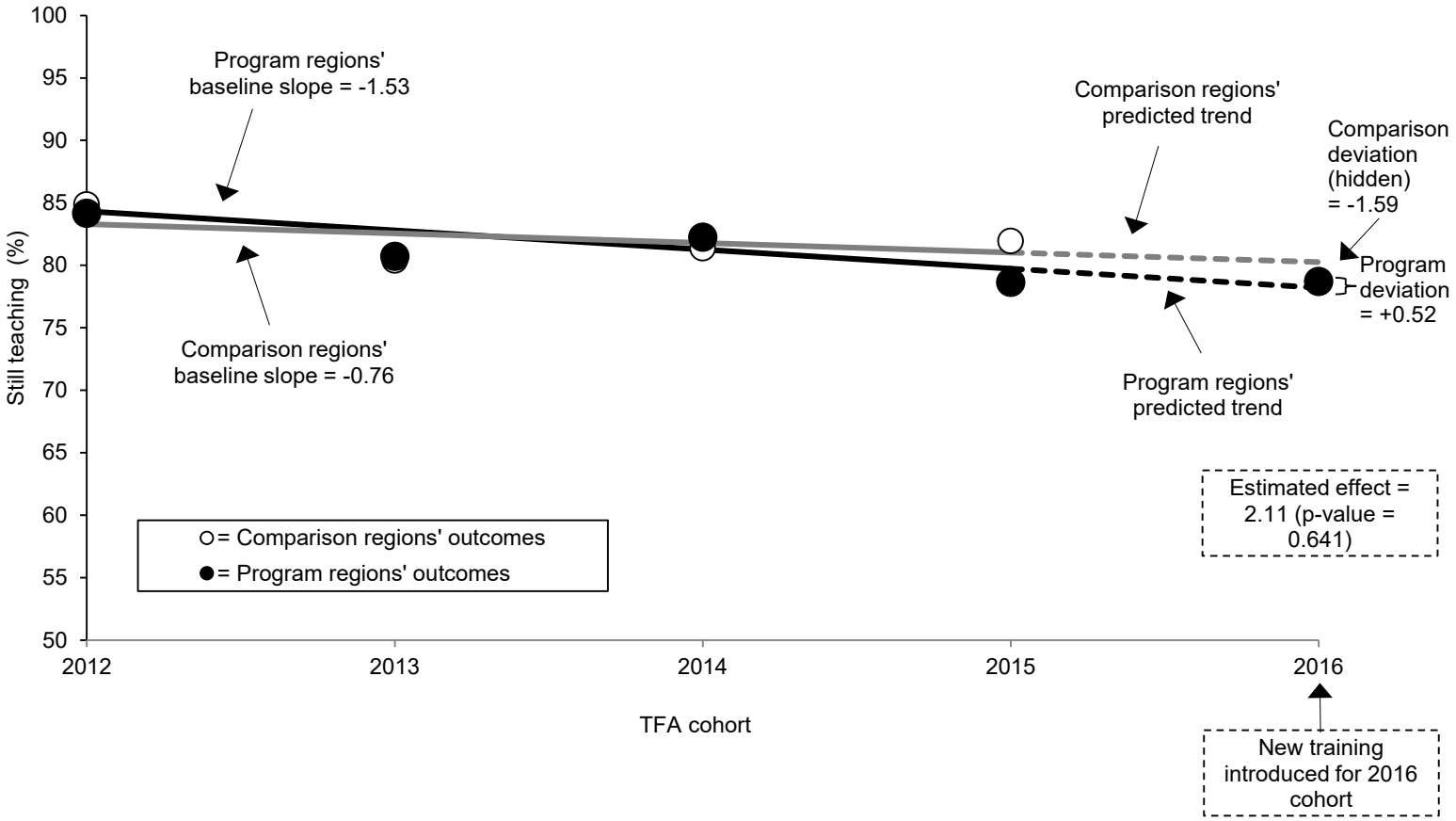
[3]These outcomes were measured using TFA administrative records. TFA tracks CM retention up to the *end* of CM's second year, but this was not examined as an outcome because, at the time of the analysis, the 2016 cohort of CMs had not yet had a chance to complete their second year.

[4]These findings are robust across several sensitivity analyses. See Appendix F for further information.

[5]For CMs in the three program regions located in Florida, data were also available on value-added scores. See Appendix G for a time series analysis of these scores.

Figure 5.2

Trends in Corps Member Retention Rates, Fall of Second Year



SOURCE: Teach For America administrative records.

Chapter 6

Looking Ahead

The decision by Teach For America (TFA) to redesign its summer training was a bold one. It was a major overhaul of something that, one could argue, had been working quite well for more than 20 years. Three large randomized studies showed not only that corps members (CMs) were equally as effective as new and certified comparison teachers but that in some instances they were more effective. As noted in Chapter 1, however, one of TFA's core values is to learn continuously, and TFA recognized an opportunity to do better by adopting and adapting the University of Washington's Teacher Education by Design model.

The implementation challenges encountered and the associated lack of effects on CM outcomes should not come as a surprise. It was the first year that TFA independently sought to implement the redesign. Everything, from the preparation of the CM training staff to the logistics of the institute, was new and very different from anything TFA had done before. TFA has learned from that first year and is on the way to improving its implementation of the redesign.

Since 2016, TFA has scaled up the redesigned training model to all of its national summer institutes, and it has just begun to align its school-year training with the summer training. MDRC is evaluating this alignment effort. This has allowed MDRC to once again observe how TFA is preparing its summer trainers and implementing the redesign, now in its third year. Early qualitative findings indicate that TFA has indeed learned from the challenges encountered in Tulsa in summer 2016. The remainder of this chapter describes some of the improvements TFA has made as observed by MDRC before and during the summer institute training in 2018. (More detailed findings on the alignment of the summer and school-year training will be published in an upcoming report.)

The Promise of Improvement

Integration of the Core Practices

TFA's biggest shortcoming in implementing the redesigned training in summer 2016 was that lead instructors and coaches were never explicitly trained on how to embed the core practices into the instructional activities during the Learning Cycle session. When and how CMs were to be trained on the core practices was left to the judgment of the lead instructors and were not part of any deliberate plan to emphasize the core practices' overall function relative to student dialogue, classroom management, lesson planning, content, and student equity.

Since 2016, TFA has taken a much more deliberate approach to training lead instructors and coaches on the way to teach CMs about the core practices. Before summer 2018, the core practices were strategically taught to TFA staff members during a training conference that MDRC

observed.[1] This approach is clearer and more specific about the strategies and routines that together lead to the enactment of the core practices. For instance, there are many ways to create and maintain a productive learning environment (one of the core practices): CMs can explain a task to their students using student-centered directions, or they can engage the class in a lesson by noticing and naming students who are on task. To deepen the staff's understanding of how to create and maintain a productive learning environment, there are now written guides that focus specifically on the strategies and routines that go into enacting this core practice. For example, there is a guide for giving student-centered directions. The guide defines what student-centered directions are (directions that make clear to all learners the "what and how" of engaging in a task); explains when to use them (for example, at the beginning of a lesson, during transitions, or when differentiating instruction for students or checking for understanding); describes how to use them (that is, be concise, positive, concrete, consistent, and engaging); and what these directions might look like in practice (for example, turn and face your neighbor, whisper, and write a prediction about this book). There are similar guides for using strategies such as giving sincere and authentic praise, building relationships with students, engaging students, and so on. These guides were a part of the training for lead instructors and for coaches, and they were later shared with CMs.

Every Learning Cycle session was also planned to build CMs' knowledge, skills, and judgment to enable them to enact specific core practices. (That is, lead instructors are no longer directed to use their own judgment regarding which and how many core practices to feature in any given session.) For example, the first two sessions are about *creating and maintaining a productive learning environment*. Within each session, one or two pieces of evidence that make up a part of one of the core practices is emphasized. For example, Learning Cycle Session 1 emphasizes *having an engaging presence* and *student-centered directions,* and Session 2 emphasizes *sincere and authentic praise* and *checking in with students.* Sessions 4 and 5 feature evidence that makes up *teaching toward an instructional goal*. In addition, the core practices are introduced gradually, and once introduced, core practices are revisited and reinforced. For instance, Session 6 revisits the learning environment core practice.

There is a guide for each of the other core practices. These have not been further broken down into specific guides for each type of evidence that together make up that core practice.[2] However, there is a version of these guides by grade level and content area. In other words, teachers of secondary mathematics use a different *positioning students as competent sensemakers* guide than do teachers of elementary mathematics or secondary history. This helps by giving CMs grade level- and content-specific instructions and examples for each of the core practices.

Eliciting and responding to student thinking replaced *teaching with society in mind.* This core practice enables lead instructors to provide targeted support to CMs that focuses on (1) eliciting student thinking on prior experiences, current understandings, interests, and needs, so that students are not just recalling facts or finding correct answers; (2) monitoring student

[1]Lead instructors and coaches were given much more robust, focused training related to teaching the different andragogies that are a part of the Learning Cycle.

[2]There is one exception. TFA adopted a new core practice called eliciting and *responding to student thinking*, and there is a "talk move" guide that supplements the main guide for that core practice.

problem solving to leverage key ideas and guide students toward the instructional goal; (3) pressing students for evidence-based explanations; and (4) challenging student ideas and encouraging more complex thinking for the purpose of deepening conceptual understandings. TFA made the decision to shift the content associated with *teaching with society in mind* fully to the Diversity, Equity, and Inclusion component of the institute training (replacing the previous social justice training).

In addition to specifying the core practices and evidence of them in each Learning Cycle session, a detailed script has been developed for each Learning Cycle session by grade and content area. The scripts specify the content, core practice, and andragogies to be used during the sessions, as well as the intended session objectives, session outputs, expected CM and student outcomes, and ideas for customizing the sessions. In observations, lead instructors sometimes decided to stray from the script based on their reading of the session. (The scripts also indicate where in the sessions instructors can improvise or veer off script slightly and what portions must be implemented with fidelity as they appear in the session plan.) For instance, the lead instructors were at times observed verbally explaining something instead of modeling it, or allowing CMs more planning time when they appeared to need it. Even when lead instructors took these liberties, they seemed to stay focused on the content; unlike in summer 2016, there was little evidence of time being used on off-topic conversations such as institutional logistics.

Perhaps as a result of TFA giving much more thought to each Learning Cycle session, there were clear indications that the CMs were attempting to embed some of the core practices in their teaching, based on observations of CMs' summer school teaching. Evidence was most often seen of CMs *creating and maintaining a productive learning environment*. CMs were often observed giving student-centered directions, redirecting unengaged students, and stating what students were doing, as directed, as a way of enforcing positive behaviors. Sometimes this went well for CMs, and lessons proceeded as planned. Other times, CMs did not implement these strategies and routines effectively, and lessons could not get off the ground. For example, CMs might not have gotten students' attention before giving student-centered directions, so directions went unheeded, or they often called attention to the positive behaviors of a few students in such a way that it did not appear to have an impact on the behavior of students who were not following directions. Another result of the refined implementation of the Learning Cycle sessions was that CMs did not report, as they had in 2016, that they were unprepared to implement effective classroom management strategies. For the most part, they felt that the institute training had, at the very least, prepared them for that. There was also evidence that CMs were stronger in *positioning students as competent sensemakers* in 2018 than they were in 2016, but there was little evidence of CMs *eliciting and responding to student thinking*.

Lesson Planning

In summer 2016, CMs in Tulsa wrote their own summer school lesson plans in the last week of institute training, whereas comparison CMs wrote their own lesson plans during the whole time that they taught summer school (weeks 2 to 5). Though the methods used to teach

lesson planning were different across the institute in Tulsa and comparison institutes, neither group felt that lesson planning was the main concept or skill they took from the institute.

In summer 2018, the training model took into account what TFA had learned since 2016 about CMs' readiness to plan lessons after their institute training. While TFA still provides lesson plans to CMs, the approach is now one of gradual release. That is, the first five lessons are fully scripted. The next five are semi-scripted, with space for CMs to add in detail. And the rest of the lesson plans include the lesson outcomes but are unscripted. Throughout the lesson plans, there are prompts for CMs to consider. For example, there are prompts to consider the learning environment. These prompts may be yet another reason why CMs were observed attempting to embed some of the core practices in their teaching.

Culturally Relevant Pedagogy

What culturally relevant pedagogy (CRP) is and who was supposed to train CMs to use it was an issue of much confusion in summer 2016. The student curricula that lead instructors were training CMs to use was grounded in CRP, but there was no explicit training about what that meant in practice. As described in Chapter 3, it was also lead instructors' understanding that CRP training was the responsibility of the social justice seminar leaders.

The 2018 training of lead instructors and coaches for the institute made it clear that CRP was the bridge between the self-reflective work in which CMs engage as part of the Diversity, Equity, and Inclusiveness (DEI) component (what were called "social justice seminars" and sometimes CRP training in 2016) and the teacher training that focused on the Learning Cycle and core practices. The purpose of CRP is to help CMs become culturally competent teachers by holding their students to high and transparent academic expectations; building affirming, authentic relationships with their students; and using these relationships and knowledge to make the lessons they are teaching more relevant. DEI, on the other hand, undergirds CRP by promoting the development of CMs' identity and cultural competence.[3] The 2018 training also makes it clear that both CRP and DEI are at the core of TFA's mission.

TFA is working to revamp the most salient aspects of its summer training to help CMs become stronger teachers and leaders who are prepared to lead highly effective classrooms and advance the cause of educational equity. The 2016 redesign was just the first step in that process. Starting in 2018, TFA also began to redesign the support offered to CMs throughout their two-year commitment, so that it is more aligned with the redesigned summer institute training. MDRC is conducting an evaluation of this effort. Findings from this evaluation will be published in an upcoming report.

[3]Ladson-Billings (1995).

Appendix A

Qualitative Implementation Research Activities

Research on the implementation of the program included both qualitative and quantitative data-collection activities. On the qualitative side, the study team conducted site visits to the program training institute that included observations of and focus groups with lead instructors, coaches, and corps members (CMs). At two of the comparison institutes, the study team conducted focus groups with the lead instructors, coaches, and CMs. In addition, at the end of the CMs' first year of teaching, the team conducted phone interviews with both the program and comparison CMs.

Observations

- A sample of **CMs' summer school teaching** in the program was observed. The observations were of both English language arts (ELA) and math lessons across Grades 1 through 12 in the middle week of the institute.
- **Lead instructor and coaching sessions** were observed at two points in time (beginning and middle) during institute to assess the fidelity of implementation of the Learning Cycle and of how core practices were embedded in the model. The study team observed ELA and math sessions in Grades 1 to 12 from beginning to end. The team observed most lead instructor sessions at least once and also various coaching sessions.

Focus Groups

- The study team conducted 23 focus groups with CMs in the last week of the program and comparison institutes. During the process of obtaining their consent, CMs indicated their interest in participating in a focus group. Of those who were interested, the team chose a sample that was racially diverse and included both men and women. The size of focus groups ranged from 6 to 10 CMs.
- The study team conducted 17 focus groups with teacher educators (lead instructors and coaches separately) during the last week of the program and comparison institutes. All lead instructors for ELA, mathematics, and elementary school, as well as a portion of coaches, across the three institutes volunteered.

Follow-Up Interviews

- In spring 2017, the study team interviewed by phone a sample of 39 CMs who participated in focus groups at the end of the summer institute in 2016. CMs indicated their interest and consented to participate in the spring interview at the start of the summer focus groups. From those who were interested, the team chose a sample that was racially diverse and included both men and women.

- The interviews with CMs followed up on the topics discussed during the focus groups at summer institute using a discussion format. These topics included key takeaways from the institute that stuck with CMs during the school year, how prepared CMs were when they first started teaching and how prepared they felt going into the second year of teaching, and the support that CMs received during the school year and how aligned it was with the institute.

What Was Measured and How

- The main goal of the redesigned training was to teach CMs how to use the core practices in their teaching. To measure their use by CMs, the study team developed rubrics designed to capture evidence of the core practices in CMs' summer school teaching. See Appendix Table A.1.

- The role of lead instructors and coaches was to teach CMs how to use the core practices in their teaching via the Learning Cycle. The study team developed checklist observation protocols to keep track of the Learning Cycle andragogies that were taught to CMs, the instructional activities used, and the core practices that were embedded in the instructional activities.

- The study team conducted focus groups with CMs using a discussion format. The interviewer introduced two specific broad topics related to the key ideas, practices, or experiences that CMs took from their training and how prepared to teach they felt, to organically draw out from interviewees their most salient thoughts regarding these topics. The study team used probes to elicit from CMs why these ideas, practices, and experiences were key to their training, how they were introduced to them, and whether there were ideas, practices, or experiences introduced during training that they thought were unimportant. The probes related to CM preparedness included why they felt prepared and what aspects of teaching they felt most and least prepared to carry out in their classrooms. The study team intentionally did not ask respondents to comment on specific aspects of the institute such as the Learning Cycle, their sessions, support staff, or core practices, with the idea that what was important would rise to the surface.

- The focus groups with lead instructors and coaches followed the same structure as the focus groups with CMs, and the broad topics were the same, except from the lead instructors' and coaches' perspectives. For instance, they were asked to identify the key ideas, practices, or experiences that were covered in training and how prepared they thought CMs were to teach. The study team used similar probes to facilitate the discussion. Additionally, the team asked the instructors and coaches about the training they received to teach the CMs and how prepared they felt to teach the CMs. The team also asked them to comment on their institute experience in terms of the institute's organization.

- The follow-up interviews with CMs followed the same structure as the focus groups. The broad topics used were meant to follow up on the topics covered in the focus groups. These included questions about the key ideas, practices, or experiences that CMs took from their training and those that were unimportant. The study team also asked them about how prepared they were to teach in the fall and how prepared they were feeling about teaching a second year. They were also asked about the support they received during the school year. As with the focus groups, the interviewer introduced topics to organically draw out from interviewees their most salient thoughts regarding these topics. The study team used probes to elicit from CMs why these ideas, practices, and experiences were or were not key to their training, and about supports during the school year. The probes related to CM preparedness included why they felt prepared and what aspects of teaching they felt most and least prepared to carry out in their classrooms. The study team intentionally kept the topics broad.

Appendix Table A.1

Core Practice Summer Tool

CORPS MEMBER EVIDENCE	
Core Practice	**Evidence to Look For:**
Creating and Maintaining a Productive Learning Environment	• CMs develop and consistently implement a learning environment plan that includes expectations, positive recognition, support for unengaged learners, policies and procedures, and a room arrangement that foster an inclusive and equitable classroom community in which there is deep engagement in meaningful content and relationships thrive. • CMs know the name (and how to correctly pronounce it) of every student in their class, and also know some key pieces of information about each student's life and what matters to her/him. • CMs have a positive presence and communication style: warm, confident voice; positive facial expressions; and open body language (that is void of any fear or domination). • CMs deliver student-centered directions that provide clarity on the "what and how" of the task at hand. • CMs recognize and reinforce positive student behaviors and choices. • CMs use what they know about students as people and learners to redirect and hold them accountable when students are off-task or unengaged. • CMs promote relationships in which students are seen and heard as individuals, through checking in with students to see how they are as people, expressing genuine curiosity and truly listening when students share about themselves, letting students know they care, and responding to a student's needs in an individualized manner. • CMs work toward increasing student autonomy and prioritize and foster meaningful student-student communication and relationships. • CMs encourage academic risk-taking by actively nurturing an environment in which students are affirmed for who they are and celebrated for asking questions, making and learning from mistakes, positing ideas, and sharing connections to their life. • CMs build excitement and investment about the instructional goals and upcoming learning experiences. • CMs connect with an adult in the family of each student at least twice during the course of institute.

(continued)

Appendix Table A.1 (continued)

CORPS MEMBER EVIDENCE	
Core Practice	**Evidence to Look For:**
Teaching Toward an Instructional Goal	• CMs can identify what they are teaching and why it matters (to both the vision of one day in their content area and in students' lives) and what success looks like at the end of the lesson cycle. • CMs execute lessons where the majority of opportunities for students to dialogue, read, write, or express themselves are clearly aligned to the instructional goal. • CMs make decisions and look for ways in the classroom to balance honoring student thinking, culture, and maintaining the rigor and goal of the lesson at hand because they understand that highly engaging lessons are connected to socio-political context and cultural and linguistic practices of students. • CMs understand that students should always experience the joy and wonder of learning through content that is deeply engaging (because they see and believe it has relevance to and purpose in their lives) and appropriately challenging (and feel properly supported to meet those challenges). • CMs begin to understand that their impact on students is measured by the assessment of learning and measurable results with students. • CMs are personally excited about the content and invested in their instructional goals.
Positioning Students as Competent Sense-makers	• CMs see students as competent and *begin* to understand the ways in which funds of knowledge and cultural wealth can be/should be utilized in informing instructional planning. • CMs begin to see the humanity, individuality, and dignity in their students *because of*, not in spite of, who they are as people and as learners — and signal to students that they do (through their language, tone, instructional choices, expressions of care, orientation towards curiosity, etc.). • CMs release the majority of responsibility of learning to students, positioning themselves as a facilitator of student dialogue. • CMs create multiple opportunities and methods for students to think, process, and share their voice. • CMs ask questions that students can answer and set them up throughout the lesson to answer said questions; questions go beyond the factual /recall and elicit student thinking around new ideas or opinions; CMs offer space for students to make personal connections to the work. • CMs express curiosity in all student ideas by validating students' thoughts and experiences — regardless of whether their thought is inaccurate, unfinished, or imprecise.

(continued)

Appendix Table A.1 (continued)

CORPS MEMBER EVIDENCE	
Core Practice	**Evidence to Look For:**
Teaching with Society in Mind	• CMs understand their own identity, power, and privileges and how these have an impact on the way in which they interact with their students, parents, and families inside and outside the classroom. • CMs are aware of the way in which social assumptions and privileges slip into daily interactions and believe that it is their job as a teacher to challenge social narratives that marginalize and oppress. • CMs make decisions that highlight the voice of students, parents, families, and the community in Tulsa. • CMs promote equity of voice in the classroom and are able to analyze their classroom for equity including, but not limited to gender, race, language, ability, or other social categorizations. • CMs use inclusive language when speaking about students, parents, families and communities (for example, when discussing relationships, being mindful not to privilege heteronormative language). • CMs challenge socially offensive language in the classroom and engage in a discussion with students as to why the language was challenged.

Appendix B

Teacher Surveys

This appendix provides additional information related to the teacher surveys that were used to explore whether the redesigned training shows promise for improving corps members' (CM) perceptions of the training and their commitment to teaching. The first section describes the two survey measures used in the analysis. The second section presents additional information on the proportion of CMs in the program and comparison regions who are included in the survey sample, as well as the characteristics of these CMs.

Survey Measures

As explained in Chapter 2, Teach For America (TFA) regularly administers surveys to monitor CMs' expectations, their perceptions of the support they receive from TFA, their relationship with the TFA community, and their commitment to teaching, equity, and TFA's vision. TFA administers the surveys online three times per year during the school year (eight weeks into the school year, mid-year, and at the end of year). The study team used the survey waves up to the fall of CMs' second year of teaching; surveys administered at the end of CMs' second year of teaching were not used because the 2016 cohort had not yet completed their second year of teaching at the time of the analysis.

The study team measured the following two confirmatory outcomes at each time point, both of which are based on a 7-point agreement scale (1=strongly disagree, 2=disagree, 3=somewhat disagree, 4=neutral, 5=somewhat agree, 6=agree, 7=strongly agree):

- **Value of the summer training:** This measure is based on CM's level of agreement with the following survey item: "Overall, I believe the preparation I received from Teach For America prior to starting at my school was valuable in my efforts to become a successful teacher."
- **Commitment to teaching and educational equity:** This measure is a composite based on eight items from the teacher survey that capture this construct.[1] CMs were asked to rate their level of agreement with the following statements:
    - I am confident that one day all children will have the opportunity to attain an excellent education.
    - I feel part of a larger movement working to ensure that all children have the opportunity to attain an excellent education.
    - I am working in partnership with families and community members to ensure that all children have the opportunity to attain an excellent education.

[1]TFA also uses a version of this composite internally for monitoring purposes; it is called the Corps and Alumni Learning Index.

- I have a working theory of what it will take for all children to have the opportunity to attain an excellent education.
- I know what role I want to play in ensuring that all children have the opportunity to attain an excellent education.
- I am committed to working, now and in the future, to ensure that all children have the opportunity to attain an excellent education.
- I am growing the capabilities I need (e.g., knowledge, skills, relationships, etc.) to play a valuable role in the movement to ensure that all children have the opportunity to attain an excellent education.
- I am effectively contributing to the effort to ensure that all children have the opportunity to attain an excellent education.

A CM's score on the composite is their average response across these items. The internal reliability of the scale (Cronbach's alpha) ranges from 0.85 to 0.89, depending on the cohort and survey wave. (See Appendix Table B.1.)

Appendix Table B.1

Internal Reliability of the Commitment to Teaching and Educational Equity Measure

Survey Wave	2012 Cohort	2013 Cohort	2014 Cohort	2015 Cohort	2016 Cohort
Fall of first year	NA	0.88	0.89	0.87	0.87
Middle of first year	0.84	0.89	0.89	0.87	0.88
End of first year	0.87	0.88	0.87	0.87	0.88
Fall of second year	0.88	0.89	0.87	0.87	0.88

SOURCE: Teach For America corps member surveys.

NOTE: A cohort is defined as a group of TFA corps members (CMs) who joined TFA and attended the summer institute the same year. The "baseline cohorts" (2012-2015) received their summer training prior to the launch of the redesigned institute in Tulsa. The "redesigned institute cohort" (2016) is the first cohort of CMs trained after the launch of the redesigned institute.Values in the table reflect the standardized Cronbach's alpha for the commitment to teaching and equity measure, which is a composite created from eight items related to CMs' beliefs in TFA's mission of ensuring that each child has an opportunity to receive an excellent education, and CMs' beliefs in their own capacity to effectively contribute to this mission. Each item is based on a 7-point agreement scale.

Survey Sample

This section provides additional information on the survey sample that was used to estimate the effect of the redesigned training on the survey outcomes, based on a comparative interrupted time series (CITS) design. In general, the survey sample is defined as all CMs in the 2012 to 2016 cohorts for whom, in a given survey wave, data were available for *both* of the survey measures

previously described.[2] However, as can be seen in Appendix Table B.1, items measuring CMs' commitment to teaching and educational equity were not included in the surveys administered in fall 2012. Therefore, for this survey wave, the survey sample for the 2012 cohort is defined as CMs for whom data were available on CMs' perceptions of the training only (the first survey outcome). More generally, the fact that CMs' commitment to teaching was not measured for this survey wave means that the survey sample for the fall of CMs' first year does not include the 2012 cohort of CMs. Relatedly, it also means that there are three baseline cohorts (and not four) in the CITS design when estimating the effect of the redesigned training on CMs' perceptions in the fall of their first year.

Response Rates

For the purposes of measuring response rates, the target population for the surveys includes all CMs in the study regions and in the 2012 to 2016 cohorts (the full study sample). Appendix Table B.2 shows the percentage of these CMs who are included in the survey sample, by cohort and for each group of regions (program and comparison). This table shows that the percentage of CMs included in the survey sample generally decreases across survey waves. For example, among all CMs in the 2016 cohort, 88 percent of CMs completed the survey in the fall of their first year of teaching, whereas 72 percent completed the survey in the fall of their second year of teaching. This reflects the fact that some CMs were no longer with TFA by this time. There are fewer regions in the 2012 to 2014 cohort study sample (15 regions rather than 18 regions) because two of the program regions did not exist before 2015, so these two program regions and their matched comparison regions are not included in the study sample or the survey sample for these earlier cohorts.[3]

Appendix Table B.2 also shows that although the differential attrition rate between program regions is sometimes statistically significant, it does not exceed 10 percent for any cohort or survey wave. Thus, the study meets the What Works Clearinghouse (WWC) standards for "low attrition" based on the liberal boundary used by the WWC review protocol for teacher training and professional development studies.[4]

When using a CITS design, the response rates should also be similar across cohorts, and any deviations over time in response rates should be similar across the program and comparison regions. To examine this issue, the study team used a CITS design to look at the "effect" of the redesigned training on survey response rates. An "effect" on response rates would suggest that the estimated effect of the new training on CMs' outcomes could be confounded with differential

[2]There are very few CMs for whom data was available for one outcome but not the other, so restricting the sample to include CMs with both measures did not result in the exclusion of many additional CMs from the survey sample.

[3]The statistical model for the CITS design can accommodate the fact that these regions had fewer baseline cohorts.

[4]What Works Clearinghouse (2017).

Appendix Table B.2

Percentage of Corps Members Included in the Survey Sample

Inclusion Rate	Program Regions	Comparison Regions	Estimated Difference	P-Value
2012 cohort				
Fall of first year (%)	90.2	90.2	0.0	0.997
Middle of first year (%)	87.4	85.1	2.2	0.534
End of first year (%)	82.6	81.5	1.1	0.798
Fall of second year (%)	81.8	78.8	3.0	0.247
Number of regions	6	11		
Number of corps members	738	1,170		
2013 cohort				
Fall of first year (%)	89.7	85.1	4.6 *	0.084
Middle of first year (%)	84.5	81.8	2.7	0.491
End of first year (%)	79.1	76.0	3.1	0.476
Fall of second year (%)	76.4	72.6	3.8	0.323
Number of regions	6	11		
Number of corps members	694	1,061		
2014 cohort				
Fall of first year (%)	88.6	87.8	0.8	0.786
Middle of first year (%)	87.0	84.2	2.8	0.354
End of first year (%)	82.3	79.2	3.0	0.401
Fall of second year (%)	77.3	75.5	1.7	0.601
Number of regions	6	11		
Number of corps members	606	973		
2015 cohort				
Fall of first year (%)	85.4	87.6	-2.2	0.432
Middle of first year (%)	78.2	83.9	-5.6	0.130
End of first year (%)	77.5	80.7	-3.2	0.377
Fall of second year (%)	77.7	74.9	2.8	0.459
Number of regions	8	13		
Number of corps members	420	887		

(continued)

Appendix Table B.2 (continued)

Inclusion Rate	Program Regions	Comparison Regions	Estimated Difference	P-Value
2016 cohort				
Fall of first year (%)	88.3	89.3	-1.0	0.758
Middle of first year (%)	83.7	83.1	0.6	0.882
End of first year (%)	82.1	80.3	1.8	0.672
Fall of second year (%)	74.0	70.9	3.1	0.577
Number of regions	8	13		
Number of corps members	361	815		

SOURCE: Teach For America corps member surveys.

NOTES: A cohort is defined as a group of TFA corps members (CMs) who joined TFA and attended the summer institute the same year. The baseline cohorts (2012-2015) received their summer training before the launch of the redesigned institute in Tulsa. The 2016 cohort is the first cohort of CMs trained after the launch of the redesigned institute. The number of CMs represents the total number of CMs in the cohort in the full study sample.

A CM is included in the survey sample if there are data available for the two confirmatory survey measures in a given survey wave. There are fewer regions in the 2012-2014 samples because two of the program regions (Idaho and Orlando) did not yet exist at this time. Therefore, these regions and their comparison regions are not included in the 2012-2014 samples. Rounding may cause slight discrepancies in calculating sums and differences.

A two-tailed t-test was applied to estimated deviations and estimated differences between program regions and comparison regions. The statistical significance of estimated differences is indicated as following: *** = 1 percent; ** = 5 percent; * = 10 percent.

changes over time in the composition of the survey sample. However, in fact, the redesigned training did *not* have a statistically significant effect on survey response rates, which gives further credibility to the CITS findings presented in this study.

Sample Characteristics

This section examines the characteristics of the CMs in the survey sample, in both the program and comparison regions. As explained in Chapter 2, a useful rule of thumb is that differences should not exceed an effect size of 0.25 standard deviations. An effect size is the difference in values for an outcome measure or characteristic expressed as a proportion of the standard deviation for that outcome measure or characteristic.[5]

Appendix Tables B.3 to B.6 show the characteristics of CMs in the survey sample, for each of the four survey waves. These tables show that the characteristics of CMs in the survey samples were similar to those of CMs in the full study sample (Table 2.4). In addition, CMs in the program and comparison regions had similar levels of education and preparation before joining TFA, as

[5]Effects sizes in this section are based on the standard deviation for the pooled sample (program and comparison regions). See Appendix E for standard deviations by sample and group.

well as similar demographic characteristics. All baseline differences in these tables are less than 0.25 as an effect size.

Appendix Tables B.7 to B.10 present the characteristics of the teaching placements of CMs in the survey sample for each of the four survey waves. Similar to the full study sample, CMs in the survey sample were also teaching in schools with high levels of need, and this was true for CMs in both the program regions and the comparison regions. However, the two groups of CMs differed somewhat with respect to some of the characteristics of their schools, and these differences exceed 0.25 in magnitude. For example, students in the program CMs' schools were less likely to be eligible for free or reduced price lunch and less likely to be chronically absent.

However, as discussed in Chapter 2, the comparison regions may still provide a credible counterfactual for the program regions, for two reasons. First, the CMs in both the program and comparison regions were working in high poverty disadvantaged schools. Second, the program and comparison regions were similar to each other on average with respect to the survey outcomes of a prior cohort of CMs (Table 2.3).

Additionally, as a sensitivity analysis, a CITS design was used to look at the "effect" of the redesigned training on the characteristics of CMs and their school placements. An "effect" on these characteristics would suggest that the estimated effect of the redesigned training on CMs' outcomes could be confounded with differential changes in CMs' characteristics or their school context that happened for the 2016 cohort relative to prior cohorts. However, the results of this sensitivity analysis show that the estimated effect of the redesigned training does not exceed 0.25 (as an effect size) for any characteristic, and it is not statistically significant for any characteristic except whether a CM had prior training in education. For this latter characteristic, the effect size does not exceed 0.25. These findings lend further credibility to the survey findings presented in this report.

Appendix Table B.3

Characteristics of the 2016 Cohort, Fall Survey Sample, First Year of Placement

Characteristic	Program Regions	Comparison Regions	Estimated Difference	Effect Size	P-Value
Preparation and skills					
Composite application score (1-5)	3.63	3.63	0.00	-0.02	0.845
Educational attainment (%)					
Bachelor	86.7	87.1	-0.4	-0.01	0.920
Masters	10.8	10.7	0.1	0.00	0.984
Doctorate	1.3	2.2	-0.9	-0.06	0.373
Other graduate degree	0.9	0.2	0.7	0.10	0.129
Major or minor in Education (%)	15.1	16.7	-1.5	-0.04	0.577
Prospect type (%)					
Undergraduate	57.2	61.7	-4.5	-0.09	0.474
Graduate	6.8	5.5	1.2	0.05	0.529
Professional	35.6	32.9	2.6	0.05	0.641
Demographic					
Age at entry into TFA	25.21	24.81	0.39	0.07	0.585
First in family to attend college (%)	35.5	33.9	1.7	0.03	0.791
Received a Pell grant (%)	47.8	51.3	-3.5	-0.07	0.554
Race and ethnicity (%)					
Hispanic	11.5	17.1	-5.6	-0.15	0.466
Black	20.0	21.0	-1.0	-0.02	0.900
White	44.3	48.9	-4.6	-0.08	0.593
Asian	6.8	4.2	2.7	0.12	0.202
Other	15.3	8.9	6.4 ***	0.19	0.009
Person of color (%)	52.1	49.9	2.2	0.04	0.788
Female (%)	72.8	76.5	-3.7	-0.08	0.386
Number of regions	8	13			
Number of corps members	316	714			

SOURCE: Teach For America administrative records for corps members.

NOTES: A cohort is defined as a group of TFA corps members (CMs) who joined TFA and attended the summer institute the same year. The findings in this table are for the 2016 cohort of CMs, who joined TFA in the year the redesigned summer institute was launched. The number of CMs represents the total number of CMs in this cohort who responded to the survey (that is, for whom data are available on the two primary survey measures). CMs' characteristics are measured during the application process to TFA. The composite application score is based on a 5-point scale measuring CMs' skills in several areas, using information from their application materials and interviews.

The values in the columns "Program Regions" and "Comparison Regions" are observed means.

A two-tailed t-test was applied to estimated deviations and estimated differences between program regions and comparison regions. The statistical significance of estimated differences is indicated as follows: *** = 1 percent; ** = 5 percent; * = 10 percent.

Rounding may cause slight discrepancies in calculating sums and differences.

Appendix Table B.4

Characteristics of the 2016 Cohort, Mid-Year Survey Sample, First Year of Placement

Characteristic	Program Regions	Comparison Regions	Estimated Difference	Effect Size	P-Value
Preparation and skills					
Composite application score (1-5)	3.62	3.63	-0.01	-0.05	0.610
Educational attainment (%)					
Bachelor	86.8	87.3	-0.5	-0.01	0.912
Masters	10.5	11.0	-0.5	-0.01	0.901
Doctorate	1.3	2.0	-0.6	-0.04	0.521
Other graduate degree	1.0	0.0	1.0 **	0.18	0.020
Major or minor in Education (%)	15.6	17.2	-1.6	-0.04	0.603
Prospect type (%)					
Undergraduate	56.2	62.5	-6.3	-0.12	0.357
Graduate	6.8	5.1	1.7	0.07	0.405
Professional	36.5	32.7	3.9	0.08	0.531
Demographic					
Age at entry into TFA	24.99	24.78	0.21	0.04	0.764
First in family to attend college (%)	33.9	34.7	-0.7	-0.01	0.913
Received a Pell grant (%)	48.4	51.0	-2.6	-0.05	0.685
Race and ethnicity (%)					
Hispanic	11.3	17.2	-6.0	-0.16	0.438
Black	20.6	20.7	-0.1	0.00	0.990
White	44.8	48.5	-3.7	-0.07	0.661
Asian	6.5	4.0	2.4	0.11	0.243
Other	14.8	9.6	5.1 **	0.15	0.035
Person of color (%)	51.7	50.0	1.7	0.03	0.828
Female (%)	73.5	76.3	-2.8	-0.06	0.482
Number of regions	8	13			
Number of corps members	300	663			

SOURCE: Teach For America administrative records for corps members.

NOTES: A cohort is defined as a group of TFA corps members (CMs) who joined TFA and attended the summer institute the same year. The findings in this table are for the 2016 cohort of CMs, who joined TFA in the year the redesigned summer institute was launched. The number of CMs represents the total number of CMs in this cohort who responded to the survey (that is, for whom data are available on the two primary survey measures). CMs' characteristics are measured during the application process to TFA. The composite application score is based on a 5-point scale measuring CMs' skills in several areas, using information from their application materials and interviews.

The values in the columns "Program Regions" and "Comparison Regions" are observed means.

A two-tailed t-test was applied to estimated deviations and estimated differences between program regions and comparison regions. The statistical significance of estimated differences is indicated as follows: *** = 1 percent; ** = 5 percent; * = 10 percent.

Rounding may cause slight discrepancies in calculating sums and differences.

Appendix Table B.5

Characteristics of the 2016 Cohort, End-of-Year Survey Sample, First Year of Placement

Characteristic	Program Regions	Comparison Regions	Estimated Difference	Effect Size	P-Value
Preparation and skills					
Composite application score (1-5)	3.62	3.63	0.00	-0.01	0.866
Educational attainment (%)					
Bachelor	86.1	87.9	-1.9	-0.05	0.668
Masters	11.2	10.3	0.9	0.03	0.809
Doctorate	1.4	2.0	-0.7	-0.05	0.506
Other graduate degree	1.0	0.0	1.0 **	0.18	0.018
Major or minor in Education (%)	15.9	17.3	-1.5	-0.04	0.644
Prospect type (%)					
Undergraduate	56.6	63.0	-6.4	-0.12	0.342
Graduate	7.3	4.8	2.4	0.10	0.236
Professional	35.6	32.5	3.1	0.06	0.613
Demographic					
Age at entry into TFA	25.10	24.73	0.37	0.07	0.641
First in family to attend college (%)	33.2	33.9	-0.8	-0.02	0.909
Received a Pell grant (%)	48.5	50.7	-2.2	-0.04	0.733
Race and ethnicity (%)					
Hispanic	11.5	16.9	-5.4	-0.14	0.474
Black	21.5	20.6	0.9	0.02	0.908
White	44.0	49.1	-5.1	-0.10	0.559
Asian	6.6	4.2	2.4	0.11	0.257
Other	14.3	9.0	5.2 *	0.16	0.057
Person of color (%)	52.4	49.6	2.8	0.05	0.736
Female (%)	72.4	76.1	-3.7	-0.08	0.364
Number of regions	8	13			
Number of corps members	295	640			

SOURCE: Teach For America administrative records for corps members.

NOTES: A cohort is defined as a group of TFA corps members (CMs) who joined TFA and attended the summer institute the same year. The findings in this table are for the 2016 cohort of CMs, who joined TFA in the year the redesigned summer institute was launched. The number of CMs represents the total number of CMs in this cohort who responded to the survey (that is, for whom data are available on the two primary survey measures). CMs' characteristics are measured during the application process to TFA. The composite application score is based on a 5-point scale measuring CMs' skills in several areas, using information from their application materials and interviews.

The values in the columns "Program Regions" and "Comparison Regions" are observed means.

A two-tailed t-test was applied to estimated deviations and estimated differences between program regions and comparison regions. The statistical significance of estimated differences is indicated as follows: *** = 1 percent; ** = 5 percent; * = 10 percent.

Rounding may cause slight discrepancies in calculating sums and differences.

Appendix Table B.6

Characteristics of the 2016 Cohort, Fall Survey Sample, Second Year of Placement

Characteristic	Program Regions	Comparison Regions	Estimated Difference	Effect Size	P-Value
Preparation and skills					
Composite application score (1-5)	3.63	3.64	0.00	-0.01	0.879
Educational attainment (%)					
Bachelor	85.7	88.6	-2.9	-0.08	0.517
Masters	11.8	10.5	1.3	0.04	0.743
Doctorate	1.1	1.0	0.1	0.01	0.877
Other graduate degree	0.7	0.0	0.7 **	0.15	0.046
Major or minor in Education (%)	16.0	17.9	-1.9	-0.05	0.586
Prospect type (%)					
Undergraduate	57.8	62.5	-4.7	-0.09	0.497
Graduate	7.7	4.9	2.8	0.11	0.190
Professional	33.6	32.7	0.9	0.02	0.879
Demographic					
Age at entry into TFA	25.21	24.62	0.59	0.11	0.473
First in family to attend college (%)	32.7	33.3	-0.5	-0.01	0.936
Received a Pell grant (%)	48.3	50.3	-2.1	-0.04	0.755
Race and ethnicity (%)					
Hispanic	11.1	16.6	-5.5	-0.15	0.469
Black	21.1	20.7	0.5	0.01	0.954
White	44.0	49.1	-5.1	-0.09	0.575
Asian	6.5	4.6	1.9	0.08	0.386
Other	14.9	8.8	6.1 **	0.18	0.031
Person of color (%)	52.8	49.9	2.8	0.05	0.748
Female (%)	74.0	75.5	-1.6	-0.03	0.717
Number of regions	8	13			
Number of corps members	267	561			

SOURCE: Teach For America administrative records for corps members.

NOTES: A cohort is defined as a group of TFA corps members (CMs) who joined TFA and attended the summer institute the same year. The findings in this table are for the 2016 cohort of CMs, who joined TFA in the year the redesigned summer institute was launched. The number of CMs represents the total number of CMs in this cohort who responded to the survey (that is, for whom data are available on the two primary survey measures). CMs' characteristics are measured during the application process to TFA. The composite application score is based on a 5-point scale measuring CMs' skills in several areas, using information from their application materials and interviews.

The values in the columns "Program Regions" and "Comparison Regions" are observed means.

A two-tailed t-test was applied to estimated deviations and estimated differences between program regions and comparison regions. The statistical significance of estimated differences is indicated as follows: *** = 1 percent; ** = 5 percent; * = 10 percent.

Rounding may cause slight discrepancies in calculating sums and differences.

Appendix Table B.7

Characteristics of the First-Year Teaching Placements of the 2016 Cohort, Fall Survey Sample (First Year of Placement)

Placement Characteristic	Program Regions	Comparison Regions	Estimated Difference	Effect Size	P-Value
Grade and subject of placement					
Grade level					
Lower elementary school	18.1	16.8	1.3	0.03	0.736
Upper elementary school	29.3	22.2	7.0	0.15	0.167
Middle school	28.6	29.9	-1.3	-0.03	0.797
High school	23.8	31.3	-7.5	-0.15	0.183
Subject taught					
General education	38.4	26.8	11.6	0.24	0.148
ELA	20.6	23.1	-2.5	-0.06	0.608
Mathematics	19.0	21.2	-2.2	-0.05	0.462
Science	15.5	15.8	-0.3	-0.01	0.895
Social studies	4.3	6.3	-2.0	-0.08	0.453
World languages	0.8	3.1	-2.3	-0.14	0.122
Other	1.4	3.5	-2.1	-0.12	0.317
Characteristics of placement schools					
School size and type					
Total enrollment	757.07	695.54	61.53	0.11	0.633
Title I school (%)	95.8	96.6	-0.8	-0.04	0.807
Magnet school (%)	26.3	7.0	19.3 *	0.51	0.052
Charter school (%)	22.2	33.2	-11.0	-0.23	0.368
Staffing					
First-year teachers at the school (%)	35.8	28.4	7.5	0.26	0.426
Teacher absences (% with 10 or more per year)	38.3	24.5	13.9 **	0.56	0.046
Certified (%)	97.4	90.6	6.8	0.38	0.142
School location (%)					
Urban	59.4	70.3	-10.9	-0.23	0.458
Suburban	32.7	19.6	13.1	0.30	0.241
Town	2.0	5.4	-3.3	-0.15	0.593
Rural	5.4	4.6	0.7	0.04	0.873
Characteristics of students in placement schools					
Chronically absent students (%)	13.9	21.6	-7.7	-0.36	0.145
Students retained (%)	3.6	5.8	-2.2 *	-0.33	0.087

(continued)

Appendix Table B.7 (continued)

Placement Characteristic	Program Regions	Comparison Regions	Estimated Difference	Effect Size	P-Value
Students with 1 or more in-school suspensions (%)	7.9	9.2	-1.3	-0.11	0.523
Students with 1 or more out-of-school suspensions (%)	11.6	17.0	-5.3	-0.36	0.102
Free or reduced price lunch students (%)	82.4	86.3	-3.9	-0.23	0.320
Students with English as a second language (%)	16.3	15.2	1.1	0.07	0.839
Students with an Individualized Education Plan (%)	10.8	12.1	-1.3	-0.19	0.459
Racial/ethnic composition of students (%)					
Hispanic	33.5	42.4	-8.9	-0.25	0.533
Black	46.5	46.2	0.4	0.01	0.981
White	12.6	6.5	6.1	0.51	0.143
Asian	2.6	2.2	0.5	0.08	0.776
Other	4.6	2.7	1.8	0.38	0.282
Female students (%)	48.4	49.0	-0.5	-0.08	0.392
Number of regions	8	13			
Number of corps members	316	714			

SOURCE: Teach For America administrative records for corps members, National Center for Education Statistics Common Core of Data (2015-2016), Private School Survey (2015-2016), and Office of Civil Rights Data Collection (2013-2014).

NOTES: A cohort is defined as a group of TFA corps members (CMs) who joined TFA and attended the summer institute the same year. The findings in this table are for the 2016 cohort of CMs, who joined TFA in the year the redesigned summer institute was launched. The number of CMs represents the total number of CMs in this cohort who responded to the survey (that is, for whom data are available on the two primary survey measures). CMs' placement characteristics are measured in the fall of their first year of teaching.

The values in the columns "Program Regions" and "Comparison Regions" are observed means. Values in the "Effect Size" column are the estimated difference divided by the standard deviation for the pooled sample. Rounding may cause slight discrepancies in calculating sums and differences.

A two-tailed t-test was applied to estimated deviations and estimated differences between program regions and comparison regions. The statistical significance of estimated differences is indicated as follows: *** = 1 percent; ** = 5 percent; * = 10 percent.

Appendix Table B.8

Characteristics of the First-Year Teaching Placements of the 2016 Cohort, Mid-Year Survey Sample (First Year of Placement)

Placement Characteristic	Program Regions	Comparison Regions	Estimated Difference	Effect Size	P-Value
Grade and subject of placement					
Grade level					
Lower elementary school	17.6	16.9	0.7	0.02	0.865
Upper elementary school	30.6	23.0	7.6	0.16	0.125
Middle school	26.6	29.2	-2.7	-0.05	0.608
High school	25.0	31.1	-6.1	-0.13	0.308
Subject taught					
General education	38.8	27.0	11.9	0.24	0.155
ELA	21.1	23.7	-2.7	-0.06	0.604
Mathematics	18.9	21.1	-2.1	-0.05	0.473
Science	15.3	15.6	-0.3	-0.01	0.916
Social studies	4.2	6.1	-1.9	-0.08	0.429
World languages	0.4	3.1	-2.8 **	-0.16	0.042
Other	1.4	3.3	-1.9	-0.11	0.408
Characteristics of placement schools					
School size and type					
Total enrollment	759.17	702.37	56.80	0.10	0.667
Title I school (%)	95.1	96.7	-1.7	-0.09	0.653
Magnet school (%)	26.5	7.0	19.5 **	0.52	0.049
Charter school (%)	21.5	31.9	-10.5	-0.22	0.383
Staffing					
First-year teachers at the school (%)	36.4	28.3	8.1	0.28	0.398
Teacher absences (% with 10 or more per year)	38.3	25.1	13.1 *	0.54	0.064
Certified (%)	97.8	90.1	7.7	0.41	0.125
School location (%)					
Urban	58.8	70.8	-12.0	-0.25	0.416
Suburban	32.5	19.0	13.6	0.32	0.215
Town	2.2	5.5	-3.3	-0.15	0.602
Rural	5.8	4.7	1.2	0.05	0.805
Characteristics of students in placement schools					
Chronically absent students (%)	14.0	21.5	-7.6	-0.35	0.156
Students retained (%)	3.6	5.9	-2.3 *	-0.34	0.077

(continued)

Appendix Table B.8 (continued)

Placement Characteristic	Program Regions	Comparison Regions	Estimated Difference	Effect Size	P-Value
Students with 1 or more in-school suspensions (%)	7.9	9.0	-1.1	-0.09	0.592
Students with 1 or more out-of-school suspensions (%)	11.8	16.7	-4.9	-0.34	0.126
Free or reduced price lunch students (%)	82.7	86.5	-3.8	-0.24	0.315
Students with English as a second language (%)	16.1	15.5	0.6	0.03	0.916
Students with an Individualized Education Plan (%)	10.8	12.1	-1.3	-0.19	0.437
Racial/ethnic composition of students (%)					
Hispanic	33.8	42.7	-8.9	-0.25	0.532
Black	46.6	46.0	0.5	0.01	0.973
White	12.5	6.3	6.2	0.52	0.128
Asian	2.6	2.2	0.4	0.07	0.821
Other	4.5	2.8	1.7	0.33	0.330
Female students (%)	48.4	48.9	-0.5	-0.07	0.377
Number of regions	8	13			
Number of corps members	300	663			

SOURCE: Teach For America administrative records for corps members, National Center for Education Statistics Common Core of Data (2015-2016), Private School Survey (2015-2016), and Office of Civil Rights Data Collection (2013-2014).

NOTES: A cohort is defined as a group of TFA corps members (CMs) who joined TFA and attended the summer institute the same year. The findings in this table are for the 2016 cohort of CMs, who joined TFA in the year the redesigned summer institute was launched. The number of CMs represents the total number of CMs in this cohort who responded to the survey (that is, for whom data are available on the two primary survey measures). CMs' placement characteristics are measured in the fall of their first year of teaching.

The values in the columns "Program Regions" and "Comparison Regions" are observed means. Values in the "Effect Size" column are the estimated difference divided by the standard deviation for the pooled sample. Rounding may cause slight discrepancies in calculating sums and differences.

A two-tailed t-test was applied to estimated deviations and estimated differences between program regions and comparison regions. The statistical significance of estimated differences is indicated as follows: *** = 1 percent; ** = 5 percent; * = 10 percent.

Appendix Table B.9

Characteristics of the First-Year Teaching Placements of the 2016 Cohort, End-of-Year Survey Sample (First Year of Placement)

Placement Characteristic	Program Regions	Comparison Regions	Estimated Difference	Effect Size	P-Value
Grade and subject of placement					
Grade level					
Lower elementary school	17.7	16.9	0.8	0.02	0.837
Upper elementary school	29.6	22.6	7.0	0.15	0.177
Middle school	26.8	29.5	-2.6	-0.05	0.616
High school	25.7	31.3	-5.7	-0.12	0.340
Subject taught					
General education	38.5	27.2	11.3	0.23	0.175
ELA	21.5	24.1	-2.6	-0.06	0.603
Mathematics	19.9	21.4	-1.4	-0.03	0.637
Science	14.5	15.4	-0.9	-0.02	0.737
Social studies	4.3	5.8	-1.5	-0.06	0.577
World languages	0.4	2.9	-2.5 *	-0.15	0.082
Other	1.1	3.1	-2.0	-0.12	0.352
Characteristics of placement schools					
School size and type					
Total enrollment	759.51	697.97	61.55	0.11	0.633
Title I school (%)	95.2	96.7	-1.5	-0.07	0.686
Magnet school (%)	27.6	7.0	20.6 **	0.54	0.046
Charter school (%)	21.3	31.7	-10.4	-0.22	0.383
Staffing					
First-year teachers at the school (%)	36.4	28.3	8.1	0.28	0.404
Teacher absences (% with 10 or more per year)	38.2	24.8	13.4 **	0.55	0.046
Certified (%)	97.8	89.8	8.0	0.42	0.110
School location (%)					
Urban	58.8	70.8	-12.0	-0.25	0.409
Suburban	32.8	19.2	13.7	0.31	0.210
Town	2.1	5.2	-3.1	-0.15	0.608
Rural	5.7	4.8	0.9	0.04	0.848
Characteristics of students in placement schools					
Chronically absent students (%)	14.2	21.3	-7.1	-0.33	0.185
Students retained (%)	3.6	5.8	-2.2	-0.32	0.104

(continued)

Appendix Table B.9 (continued)

Placement Characteristic	Program Regions	Comparison Regions	Estimated Difference	Effect Size	P-Value
in-school suspensions (%)	8.0	9.3	-1.3	-0.10	0.545
Students with 1 or more out-of-school suspensions (%)	12.0	16.5	-4.4	-0.31	0.164
Free or reduced price lunch students (%)	82.2	86.6	-4.4	-0.27	0.253
Students with English as a second language (%)	16.0	15.8	0.3	0.01	0.963
Students with an Individualized Education Plan (%)	10.8	12.3	-1.5	-0.21	0.382
Racial/ethnic composition of students (%)					
Hispanic	32.9	43.1	-10.2	-0.28	0.476
Black	47.5	45.9	1.5	0.04	0.924
White	12.5	6.1	6.4	0.53	0.121
Asian	2.6	2.2	0.4	0.06	0.833
Other	4.5	2.7	1.9	0.38	0.273
Female students (%)	48.3	49.0	-0.6	-0.10	0.204
Number of regions	8	13			
Number of corps members	295	640			

SOURCE: Teach For America administrative records for corps members, National Center for Education Statistics Common Core of Data (2015-2016), Private School Survey (2015-2016), and Office of Civil Rights Data Collection (2013-2014).

NOTES: A cohort is defined as a group of TFA corps members (CMs) who joined TFA and attended the summer institute the same year. The findings in this table are for the 2016 cohort of CMs, who joined TFA in the year the redesigned summer institute was launched. The number of CMs represents the total number of CMs in this cohort who responded to the survey (that is, for whom data are available on the two primary survey measures). CMs' placement characteristics are measured in the fall of their first year of teaching.

The values in the columns "Program Regions" and "Comparison Regions" are observed means. Values in the "Effect Size" column are the estimated difference divided by the standard deviation for the pooled sample. Rounding may cause slight discrepancies in calculating sums and differences.

A two-tailed t-test was applied to estimated deviations and estimated differences between program regions and comparison regions. The statistical significance of estimated differences is indicated as follows: *** = 1 percent; ** = 5 percent; * = 10 percent.

Appendix Table B.10

Characteristics of the First-Year Teaching Placements of the 2016 Cohort, Fall Survey Sample (Second Year of Placement)

Placement Characteristic	Program Regions	Comparison Regions	Estimated Difference	Effect Size	P-Value
Grade and subject of placement					
Grade level					
Lower elementary school	18.9	14.4	4.5	0.11	0.256
Upper elementary school	26.5	22.4	4.0	0.09	0.465
Middle school	27.9	30.1	-2.2	-0.04	0.704
High school	26.2	33.4	-7.2	-0.15	0.250
Subject taught					
General education	27.7	26.3	1.4	0.03	0.871
ELA	24.9	25.4	-0.5	-0.01	0.914
Mathematics	22.4	20.4	2.0	0.05	0.529
Science	15.0	16.9	-1.9	-0.05	0.576
Social studies	4.9	5.8	-0.9	-0.04	0.741
World languages	0.4	2.4	-2.0	-0.13	0.155
Other	6.1	3.6	2.5	0.12	0.508
Characteristics of placement schools					
School size and type					
Total enrollment	712.75	714.86	-2.11	0.00	0.987
Title I school (%)	96.2	97.6	-1.4	-0.08	0.633
Magnet school (%)	26.1	4.5	21.6 **	0.59	0.034
Charter school (%)	22.4	30.1	-7.8	-0.17	0.495
Staffing					
First-year teachers at the school (%)	34.9	27.6	7.3	0.25	0.416
Teacher absences (% with 10 or more per year)	36.6	25.7	10.9	0.45	0.173
Certified (%)	98.0	90.3	7.7	0.44	0.115
School location (%)					
Urban	59.2	70.3	-11.0	-0.23	0.460
Suburban	29.5	19.7	9.8	0.22	0.355
Town	3.3	5.2	-2.0	-0.09	0.744
Rural	7.5	4.7	2.8	0.13	0.609
Characteristics of students in placement schools					
Chronically absent students (%)	14.1	21.9	-7.8	-0.37	0.150
Students retained (%)	3.9	6.2	-2.3 *	-0.37	0.082

(continued)

Appendix Table B.10 (continued)

Placement Characteristic	Program Regions	Comparison Regions	Estimated Difference	Effect Size	P-Value
Students with 1 or more in-school suspensions (%)	8.0	9.6	-1.6	-0.12	0.429
Students with 1 or more out-of-school suspensions (%)	11.7	17.1	-5.4 *	-0.36	0.082
Free or reduced price lunch students (%)	81.4	86.1	-4.7	-0.27	0.272
Students with English as a second language (%)	15.9	15.5	0.4	0.02	0.941
Students with an Individualized Education Plan (%)	11.1	13.2	-2.1	-0.26	0.349
Racial/ethnic composition of students (%)					
Hispanic	31.1	43.5	-12.4	-0.35	0.382
Black	48.1	44.6	3.5	0.09	0.829
White	13.4	7.0	6.3	0.48	0.224
Asian	2.7	2.3	0.5	0.08	0.780
Other	4.7	2.6	2.1	0.45	0.201
Female students (%)	48.5	48.8	-0.3	-0.06	0.583
Number of regions	8	13			
Number of corps members	267	561			

SOURCE: Teach For America administrative records for corps members, National Center for Education Statistics Common Core of Data (2015-2016), Private School Survey (2015-2016), and Office of Civil Rights Data Collection (2013-2014).

NOTES: A cohort is defined as a group of TFA corps members (CMs) who joined TFA and attended the summer institute the same year. The findings in this table are for the 2016 cohort of CMs, who joined TFA in the year the redesigned summer institute was launched. The number of CMs represents the total number of CMs in this cohort who responded to the survey (that is, for whom data are available on the two primary survey measures). CMs' placement characteristics are measured in the fall of their first year of teaching.

The values in the columns "Program Regions" and "Comparison Regions" are observed means. Values in the "Effect Size" column are the estimated difference divided by the standard deviation for the pooled sample. Rounding may cause slight discrepancies in calculating sums and differences.

A two-tailed t-test was applied to estimated deviations and estimated differences between program regions and comparison regions. The statistical significance of estimated differences is indicated as follows: *** = 1 percent; ** = 5 percent; * = 10 percent.

Appendix C

Closed-Ended Teacher Logs

This appendix provides additional information about the closed-ended teacher logs that the study team used to examine corps members' (CM) use of instructional strategies aligned with college- and career-readiness standards. The first section describes the format and content of the teacher logs and the measures that the team constructed from them. The second section presents additional information on the proportion of CMs in the program and comparison regions who are included in the teacher log sample, as well as the characteristics of these CMs.

Data Collection and Measures

As explained in Chapter 2, the study team used weekly online teacher logs to measure the self-reported instructional practices used by CMs in the 2016 cohort during their first year of teaching. The logs were sent to CMs in the program and comparison regions who were teaching English Language Arts (ELA), mathematics, or general education, and who consented to participate in this data collection. The format of the log (open-ended versus closed-ended) alternated across weeks.

The closed-ended logs — which were used to look at the effect of the redesigned training on CMs' use of instructional strategies aligned with college- and career-readiness standards — asked CMs to report on how often they used different type of strategies. At the start of each log, CMs received a randomly selected letter of the alphabet and asked to report on their instructional practices with a focal student whose name starts with that letter (or the closest letter).

The items in the closed-ended logs were aligned with TFA's college- and career-readiness standards, which are subject-specific instructional standards that undergird the training model used at the national institute in Tulsa and that are aligned with the common core state standards.[1] Because the standards differ across subjects and grade levels, the closed-ended log items differed across levels and subjects as well. Appendix Tables C.1 and C.2 list the items included in the closed-ended logs for ELA and mathematics, respectively, and Appendix Boxes C.1 to C.4 show how the items maps onto the college- and career-readiness standards by content area and level.

The study team asked each CM to complete 15 closed-ended logs during the year. The team sent CMs teaching secondary school, as well as CMs teaching elementary school in departmentalized schools, the closed-ended log for their content area and level every two weeks (about twice per month). The team sent CMs in elementary schools teaching both content areas (CMs teaching general education) the ELA log *or* the math log every two weeks, with the content area alternating across weeks (about one log per content area per month). Each log was intended to take no more than 5 minutes to complete and CMs received a $5 gift card for each completed log.

[1]The standards are called the "core components of instruction" internally at TFA. The standards map onto several content-specific instructional activities that were introduced at the redesigned training at the national institute in Tulsa.

Appendix Table C.1

Closed-Ended Log Items, by Level, English Language Arts

Log Item	Lower	Upper
How many minutes did the focal student spend in your ELA/math lesson today?	x	x
Did the focal student receive the same instruction as the majority of the students in the class during today's ELA lesson? (yes, no)	x	x
Does the focal student have an Individualized Education Program (IEP)? (yes, no)	x	x
To what extent was the focal student expected to work on the following activities in your ELA lesson today? (A focus of instruction, touched on briefly, not taught today)		
Developing skills in phonics	x	
Developing skills in fluency	x	
Developing skills in vocabulary	x	
Developing skills in writing	x	
Listening to text read aloud	x	x
Reading with a partner or group	x	x
Reading independently	x	x
Taking a test or quiz	x	x
The focal student received instruction that (no time, some of the time, most of the time, all of the time)		
Addressed gaps in focal student's reading performance	x	
Addressed gaps in focal student's writing performance	x	
Was the focal student asked the following types of questions explicitly or as part of a whole group (yes, no)		
Text-dependent questions related to text structure		x
Text-dependent questions related to character perspective		x
Text-dependent questions related to vocabulary		x
Text-dependent questions related to author's choices		x
Text-dependent questions related to theme		x
Text-dependent questions related to plot		x
Non-text-dependent questions related to the text		x
To cite text evidence in their writing activities and/or written responses		x
Did the focal student's work today include a text that featured an author, characters, and/or a community with a similar background to this student? (yes, no)	x	x
Today's lesson was related to a text that is… (fiction, non-fiction, both)	x	x
The focal student received instruction that integrated ideas, themes, or information related to other content areas (no time, some, most, all of the time)		x
The main text featured in today's lesson is… (far below reading level, slightly below reading level, at reading level, slightly above reading level, far above reading level)	x	x
Relative to other students in this class, how would you rank the focal student's overall ELA/math performance? (lower, middle, upper third)	x	x
What proportion of the time during the ELA lesson was the focal student engaged in the lesson? (no time, some of the time, most of the time, all of the time)	x	x

Appendix Table C.2

Closed-Ended Log Items, by Level, Mathematics

Log Item	Lower	Upper
How many minutes did the focal student spend in your ELA/math lesson today?	x	x
Did the focal student receive the same instruction as the majority of the students in the class during today's math lesson? (yes, no)	x	x
Does the focal student have an Individualized Education Program (IEP)? (yes, no)	x	x
How many mathematical problems or exercises was the focal student assigned to work on during today's math lesson*?* [one or two–more than 10]	x	x
To what extent was the focal student expected to work on the following activities in your math lesson today? (A focus of instruction, touched on briefly, not taught today)		
Solving a math problem(s) with a given algorithm	x	x
Using concrete manipulatives to solve a math problem (e.g., base 10 blocks, pattern blocks, real life objects, etc.)	x	
Using tool(s) to solve a math problem (e.g., manipulatives, technology, real life objects, etc.)		x
Representing and analyzing relationships using pictorial illustrations (e.g., tables, charts, number lines, etc.)	x	x
Estimating solutions to a math problem(s)	x	
Trying alternative solutions to a math problem(s)	x	x
Critiquing the mathematical reasoning of others	x	x
Discussing math with a peer(s)	x	x
Justifying responses to a math problem(s)	x	x
Writing down mathematical thinking	x	x
Applying math concepts to a "real-world" problem(s)	x	x
Working in a math team(s) or pair(s)	x	x
Investigating a multistep problem(s)	x	x
Developing hypotheses		x
Testing hypotheses		x
Practicing or relearning prerequisite concepts	x	x
Receiving remediation	x	x
Continuing to persevere on a challenging problem	x	x
Taking a test or quiz	x	x
Relative to other students in this class, how would you rank the focal student's overall math performance? (upper, middle, lower third)	x	x
What proportion of the time during the math lesson was the focal student engaged in the lesson? (no time, some of the time, most of the time, all of the time)	x	x

Appendix Box C.1

Closed-Ended Items, by Career- and College-Readiness Standard, Log for Lower Elementary School English Language Arts

Rich and high-quality complex texts

- Today's lesson was related to a text that is… [fiction/non-fiction/both]

Balanced literacy

- Developing skills in phonics
- Developing skills in fluency
- Developing skills in vocabulary

Miles on the page

- Developing skills in writing
- Listening to text read aloud
- Reading with a partner or in a group
- Reading independently

Scaffolding and differentiation

- Addressed gaps in focal student's reading performance
- Addressed gaps in focal student's writing performance

The study team examined the effect of the redesigned training on the following three confirmatory measures related to CMs' use of standards-aligned strategies in three domains of instruction:

- **Use of culturally responsive text (domain: culturally responsive instruction):** This measure is based on CMs' response to the following item from the ELA logs: "Did the focal student's work today include a text that featured an author, characters, and/or a community with a similar background to this student? (yes, no)." The study team selected the use of culturally responsive text as a primary outcome because of the importance for TFA of reducing educational inequity.
- **Citing text evidence (domain: ELA instruction):** This measure, which is available for CMs at the upper elementary and secondary levels only, is based on CMs' responses to the following items from the ELA log: "Was the focal student asked to cite text evidence in their writing activities and/or written responses? (yes, no)." The team selected citing text evidence as the primary measure for ELA instruction because it was the most aligned with comprehension.

Appendix Box C.2

Closed-Ended Items, by Career- and College-Readiness Standard, Log for Upper Elementary or Secondary School English Language Arts

Evidence-based discussions and writing [UE] / Reading, writing, and speaking are grounded in evidence from text, both literary and informational [SEC]

- Text-dependent questions related to text structure
- Text-dependent questions related to character perspective
- Text-dependent questions related to vocabulary
- Text-dependent questions related to author's choices
- Text-dependent questions related to theme
- Text-dependent questions related to plot
- Non-text-dependent questions related to the text
- To cite text evidence in their writing activities and/or written responses

Literacy fosters identity, community, and awareness of the broader world

- Did the focal student's work today include a text that featured an author, characters, and/or a community with a similar background to this student? (yes, no)

Miles on the page [UE] / Independent reading [SEC]

- Reading independently

Integration of Literacy and content areas [UE] / Read complex, grade-appropriate texts worth reading [SEC]

- The focal student received instruction that integrated ideas, themes, or information related to other content areas (no time, some, most, all of the time)

Textual analysis of complex texts [UE] / Read complex, grade-appropriate texts worth reading [SEC]

- Today's lesson was related to a text that is… [fiction/non-fiction/both]

Children bear the burden and joy of the reading, writing, and thinking (SEC]

- Listening to text read aloud
- Reading with a partner or group

[UE] = Name of upper elementary core components
[SEC] = Names of the secondary core component

Appendix Box C.3

Closed-Ended Items, by Career- and College-Readiness Standard, Log for Lower Elementary School Mathematics

Multiple methods, representations, and models

- Solving a math problem(s) with a given algorithm
- Using concrete manipulatives to solve a math problem (e.g., base 10 blocks, pattern blocks, real life objects, etc.)
- Representing and analyzing relationships using pictorial illustrations (e.g., tables, charts, number lines, etc.)
- Trying alternative solutions to a math problem(s)
- Investigating a multistep problem(s)

Emphasis on number sense

- Estimating solutions to a math problem(s)

Real and meaningful discourse

- Critiquing the mathematical reasoning of others
- Discussing math with a peer(s)
- Justifying responses to a math problem(s)
- Writing down mathematical thinking

Learning builds on what students know

- Applying math concepts to a "real-world" problem(s)

Cultivating growth mindset and team

- Working in a math team(s) or pair(s)
- Continuing to persevere on a challenging problem

- **Math total score (domain: math instruction):** This measure is a composite scale that combines the math log items that map onto the college- and career-readiness standards for math instruction. (See Appendix Boxes C.3 and C.4.) Because the response scale differs across items, each item was first z-scored based on the mean and standard deviation of the comparison group logs; these z-scores were then averaged across items, and the average score was z-scored, again, based on the mean and standard deviation for the comparison group logs. Because some of the log items differed across levels, the composite

Appendix Box C.4

Closed-Ended Items, by Career- and College-Readiness Standard, Log for Upper Elementary or Secondary School Mathematics

Flexible engagement with problems

- Solving a math problem(s) with a given algorithm
- Using tool(s) to solve a math problem (e.g., manipulatives, technology, real life objects, etc.)
- Representing and analyzing relationships using pictorial illustrations (e.g., tables, charts, number lines, etc.)
- Trying alternative solutions to a math problem(s)
- Investigating a multistep problem(s)
- Developing hypotheses
- Testing hypotheses

Real and meaningful discourse

- Critiquing the mathematical reasoning of others
- Justifying responses to a math problem(s)
- Writing down mathematical thinking

Learning builds on what students know

- Applying math concepts to a "real-world" problem(s)

Support for all learners with grade-level content

- Practicing or relearning prerequisite concepts
- Receiving remediation
- Continuing to persevere on a challenging problem

score was created separately for the lower elementary school and upper elementary or secondary school logs, and then the scores were pooled across levels for the analysis. The reliability (Cronbach's alpha) for the composite is 0.71 at the lower elementary school level (13 items) and 0.75 at the upper elementary or secondary school level (14 items).

The study team also considered creating a composite score for ELA instruction; however, the reliability of such a composite would have been too low (Cronbach's alpha < 0.60).

Teacher Log Sample

This section provides additional information on the teacher log samples that the study team used to examine the effect of the redesigned training on self-reported instructional practices. As explained in Chapter 2, a CM was included in the log sample if they responded to at least one closed-ended log during the school year.

Response Rates

The target population for the logs included all CMs teaching mathematics, ELA, or general education in the 2016 cohort who were in the study regions. Appendix Table C.3 shows the percentage of these CMs who consented to participating in the logs and who are included in the log sample, in each group of regions (program and comparison). Among eligible CMs in the study sample, about 58 percent of CMs consented to participate in the logs, and among those who consented, about 70 percent on average completed at least one log. (In one comparison region, no ELA logs were completed.) Thus, overall, about 40 percent of eligible CMs are included in the log sample. Given this high level of non-response, findings related to CMs' use of standards-aligned instructional strategies should be interpreted with caution. Moreover, CMs in the log sample differed from the CMs who did not participate in the logs with respect to

Appendix Table C.3

Consent and Response Rates for the Teacher Logs, 2016 Cohort

Response Rate	Program Regions	Comparison Regions	Estimated Difference	P-Value
Among eligible teachers in the study regions				
Teachers who consented (%)	67.6	56.1	11.5	0.181
Teachers in the log sample (%)	44.8	39.2	5.6	0.340
Number of regions	8	13		
Number of corps members	275	567		
Among eligible teachers in the study regions who consented				
Teachers in the log sample (%)	67.6	71.1	-3.4	0.456
Number of regions	8	13		
Number of corps members	176	301		

SOURCE: Closed-ended teacher logs administered by MDRC.

NOTES: The 2016 cohort is the first cohort of TFA corps members (CMs) trained after the launch of the redesigned institute in Tulsa. To be eligible to complete the logs, CMs had to be teaching English Language Arts, mathematics, or general education. A CM is defined as being in the log sample if the CM completed at least one of the closed-ended logs.

A two-tailed t-test was applied to estimated deviations and estimated differences between program regions and comparison regions. The statistical significance of estimated differences is indicated as follows: *** = 1 percent; ** = 5 percent; * = 10 percent.

certain characteristics; if the effect of the redesigned training depends on these characteristics, then the findings for CMs in the log sample may not generalize to CMs who did not participate.[2]

Sample Characteristics

This section examines the characteristics of the CMs in the log sample in the program and comparison regions. As explained in Chapter 2, a useful rule of thumb is that differences should not exceed and effect size of 0.25 standard deviations. An effect size is the difference in values for an outcome measure or characteristic expressed as a proportion of the standard deviation for that outcome measure or characteristic.[3]

Appendix Tables C.4 and C.5 show the characteristics of CMs in the log sample by content area (ELA and mathematics). These tables show that the characteristics of CMs in the log samples were generally similar to those of CMs in the full study sample (Table 2.4). In addition, CMs in the program and comparison regions had similar levels of education and preparation before joining TFA, as well as similar demographic characteristics. All baseline differences in these tables are less than 0.25 as an effect size, with the following exceptions: CMs teaching ELA in the program regions were less likely than the comparison region CMs to be black, and CMs teaching mathematics in the program regions were more likely than the comparison group CMs to classify their race as "other."

Appendix Tables C.6 and C.7 show the characteristics of the teaching placements of CMs in the log sample. Similar to the full study sample, CMs in the log sample were also teaching in schools with high levels of need, and this is true for CMs in both the program and comparison regions. However, the two groups of CMs differed somewhat with respect to some of the characteristics of their schools, and these differences exceed 0.25 in magnitude. For example, the CMs in the program regions were more likely to be in magnet schools, and in schools with higher teacher absenteeism rates. The schools in the program regions also had a higher proportion of white students. As discussed in Chapter 2, the study design that was used to estimate the effects of the redesigned summer training on instructional practices is less rigorous, so these differences in school characteristics between the program and comparison regions are more likely to influence the results. Thus, estimated effects on standards-aligned instructional strategies should not be interpreted as the causal effect of the redesigned training.

Appendix Table C.8 and C.9 present the characteristics of the focal students that CMs were asked to consider when reporting on their instructional practices, based on items about these students in the logs.[4] These tables show that the focal students taught by CMs in the program and

[2]CMs in the teacher log sample and CMs who were eligible for the logs but did not participate had similar application scores and they taught in schools with similar characteristics. However, CMs in the log sample were younger (25 years old, compared with 26 years old for the CMs who did not participate), less likely than CMs who did not participate to be a person of color (42 percent, compared with 57 percent), and less likely to have received a Pell grant (45 percent, compared with 53 percent).

[3]Effects sizes in this section are based on the standard deviation for the pooled sample (program and comparison regions). See Appendix E for standard deviations by sample and group.

[4]These tables are like Table 2.6, but shown separately for each content area (ELA and math).

comparison regions were very similar with respect to their levels of academic performance and academic needs. All baseline differences in these tables are less than 0.25 as an effect size.

Appendix Table C.4

Characteristics of the 2016 Cohort, Teacher Log Sample (English Language Arts)

Characteristic	Program Regions	Comparison Regions	Estimated Difference	Effect Size	P-Value
Preparation and skills					
Composite application score (1-5)	3.61	3.62	-0.01	-0.04	0.754
Educational attainment (%)					
Bachelor	88.4	85.1	3.4	0.09	0.582
Masters	9.0	13.4	-4.4	-0.13	0.448
Doctorate	1.2	1.7	-0.5	-0.04	0.787
Other graduate degree	1.1	0.0	1.1	0.17	0.205
Major or minor in Education (%)	20.7	16.1	4.6	0.11	0.406
Prospect type (%)					
Undergraduate	62.0	62.8	-0.8	-0.02	0.934
Graduate	5.6	4.6	1.1	0.05	0.775
Professional	31.8	32.6	-0.8	-0.02	0.912
Demographic					
Age at entry into TFA	25.12	24.64	0.48	0.08	0.604
First in family to attend college (%)	26.5	30.7	-4.2	-0.09	0.613
Received a Pell grant (%)	35.9	47.8	-11.9	-0.22	0.233
Race/ethnicity (%)					
Hispanic	8.7	8.5	0.2	0.01	0.958
Black	7.6	21.0	-13.4	-0.32	0.189
White	64.4	55.5	8.9	0.17	0.398
Asian	2.3	1.7	0.6	0.04	0.771
Other	16.0	11.6	4.5	0.13	0.499
Person of color (%)	32.1	43.1	-10.9	-0.21	0.288
Female (%)	86.2	88.9	-2.6	-0.08	0.577
Number of regions	8	12			
Number of Corps Members	87	138			

(continued)

Appendix Table C.4 (continued)

SOURCE: Teach For America administrative records for corps members.

NOTES: Online teacher logs were administered to corps members (CMs) in the 2016 cohort in the program and comparison regions. CMs teaching English Language Arts (ELA), mathematics, or general education, in the first year of their teaching placement were eligible for the logs. Closed-ended logs were sent to eligible and consenting CMs every two weeks during the school year (15 weeks total). The instructional items in the logs differ by level (lower versus upper grades) and by content area (ELA and math). Secondary school teachers, as well as elementary school teachers in departmentalized schools, were sent the log for their content area and level every two weeks (about once per month). Elementary school teachers teaching both content areas (general education teachers) were sent the ELA log or the math log every two weeks, with the content area alternating between logs (about one log per content area per month).

A CM is included in the log sample if the CM completed at least one log during the school year. The values in the columns "Program Regions" and "Comparison Regions" are the observed mean characteristics of these CMs. Values in the "Effect Size" column are the estimated difference divided by the standard deviation for the pooled sample.

A two-tailed t-test was applied to estimated differences between program regions and comparison regions. The statistical significance of estimated differences is indicated as follows: *** = 1 percent; ** = 5 percent; * = 10 percent. Rounding may cause slight discrepancies in calculating sums and differences.

CMs' characteristics are measured during the application process to TFA. The composite application score is based on a 5-point scale measuring CMs' skills in several areas, using information from their application materials and interviews.

Appendix Table C.5

Characteristics of the 2016 Cohort, Teacher Log Sample (Mathematics)

Characteristic	Program Regions	Comparison Regions	Estimated Difference	Effect Size	P-Value
Preparation and skills					
Composite application score (1-5)	3.61	3.61	0.00	0.02	0.903
Educational attainment (%)					
Bachelor	87.4	89.0	-1.6	-0.05	0.770
Masters	9.8	10.3	-0.5	-0.01	0.929
Doctorate	1.3	1.0	0.3	0.03	0.853
Other graduate degree	1.3	0.0	1.3	0.19	0.169
Major or minor in Education (%)	26.3	22.3	4.0	0.09	0.526
Prospect Type (%)					
Undergraduate	61.2	64.3	-3.2	-0.06	0.759
Graduate	7.5	4.7	2.8	0.11	0.401
Professional	32.1	31.3	0.8	0.02	0.931
Demographic					
Age at entry into TFA	25.43	24.21	1.21	0.19	0.162
First in family to attend college (%)	32.6	31.4	1.3	0.03	0.872
Received a Pell grant (%)	43.8	44.8	-0.9	-0.02	0.921
Race/ethnicity (%)					
Hispanic	10.7	14.1	-3.4	-0.10	0.618
Black	10.2	16.4	-6.2	-0.15	0.533
White	48.9	55.4	-6.5	-0.12	0.600
Asian	7.5	6.2	1.3	0.05	0.726
Other	18.7	7.5	11.2 **	0.34	0.039
Person of color (%)	45.6	42.4	3.2	0.06	0.785
Female (%)	79.4	79.7	-0.3	-0.01	0.959
Number of regions	8	13			
Number of corps members	80	151			

(continued)

Appendix Table C.5 (continued)

SOURCE: Teach For America administrative records for corps members.

NOTES: Online teacher logs were administered to Corps Members (CMs) in the 2016 cohort in the program and comparison regions. CMs teaching English Language Arts (ELA), mathematics, or general education, in the first year of their teaching placement were eligible for the logs. Closed-ended logs were sent to eligible and consenting CMs every two weeks during the school year (15 weeks total). The instructional items in the logs differ by level (lower versus upper grades) and by content area (ELA and math). Secondary school teachers, as well as elementary school teachers in departmentalized schools, were sent the log for their content area and level every two weeks (about once per month). Elementary school teachers teaching both content areas (general education teachers) were sent the ELA log or the math log every two weeks, with the content area alternating between logs (about one log per content area per month).

A CM is included in the log sample if the CM completed at least one log during the school year. The values in the columns "Program Regions" and "Comparison Regions" are the observed mean characteristics of these CMs. Values in the "Effect Size" column are the estimated difference divided by the standard deviation for the pooled sample.

A two-tailed t-test was applied to estimated differences between program regions and comparison regions. The statistical significance of estimated differences is indicated as follows: *** = 1 percent; ** = 5 percent; * = 10 percent. Rounding may cause slight discrepancies in calculating sums and differences.

CMs' characteristics are measured during the application process to TFA. The composite application score is based on a 5-point scale measuring CMs' skills in several areas, using information from their application materials and interviews.

Appendix Table C.6

Characteristics of the First-Year Teaching Placements of the 2016 Cohort, Teacher Log Sample (English Language Arts)

Placement Characteristic	Program Regions	Comparison Regions	Estimated Difference	Effect Size	P-Value
Grade and subject of placement					
Grade level (%)					
Lower elementary school	28.8	33.7	-4.9	-0.10	0.551
Upper elementary school	47.4	31.5	16.0 *	0.31	0.071
Middle school	15.4	15.8	-0.5	-0.01	0.940
High school	8.3	19.1	-10.8 *	-0.28	0.096
Subject taught (%)					
General education	64.9	51.2	13.6	0.26	0.239
ELA	34.9	44.4	-9.6	-0.18	0.382
Mathematics	0.0	4.2	-4.2	-0.26	0.229
Characteristics of placement schools					
School size and type					
Total enrollment	671.48	630.50	40.98	0.10	0.716
Magnet school (%)	16.6	6.3	10.4	0.34	0.290
Charter school (%)	20.9	35.6	-14.7	-0.29	0.335
Title I school (%)	95.2	98.7	-3.5	-0.21	0.186
Staffing					
First-year teachers at the school (%)	36.4	29.9	6.5	0.25	0.481
Teacher absences (% with 10 or more per year)	39.9	21.1	18.8 ***	0.77	0.009
Certified (%)	97.4	92.8	4.6	0.27	0.288
School location (%)					
Urban	60.2	68.8	-8.6	-0.17	0.601
Suburban	32.7	21.1	11.6	0.26	0.351
Town	2.3	4.6	-2.3	-0.11	0.705
Rural	3.1	5.7	-2.6	-0.12	0.597

(continued)

Appendix Table C.6 (continued)

Placement Characteristic	Program Regions	Comparison Regions	Estimated Difference	Effect Size	P-Value
Characteristics of students in placement schools					
Chronically absent students (%)	13.6	16.9	-3.2	-0.19	0.419
Students retained (%)	3.7	4.4	-0.7	-0.16	0.598
Students with 1 or more in-school suspensions (%)	4.8	7.3	-2.5	-0.23	0.147
Students with 1 or more out-of-school suspensions (%)	9.8	13.6	-3.8	-0.25	0.377
Free or reduced price lunch students (%)	85.9	89.3	-3.5	-0.29	0.254
Students with English as a second language (%)	17.9	18.7	-0.8	-0.04	0.901
Students with an Individualized Education Plan (%)	10.9	11.0	0.0	0.00	0.986
Racial/ethnic composition of students (%)					
Hispanic	35.5	48.6	-13.1	-0.35	0.374
Black	44.8	41.9	2.9	0.07	0.859
White	11.5	4.7	6.7 **	0.68	0.035
Asian	2.7	2.8	-0.1	-0.01	0.976
Other	4.7	2.0	2.7	0.49	0.133
Female students (%)	48.2	48.8	-0.7	-0.12	0.448
Number of regions	8	12			
Number of corps members	87	138			

SOURCE: Teach For America administrative records for corps members, Common Core of Data (school year 2015-2016), and Office of Civil Rights Data Collection (school year 2013-2014).

NOTES: Online teacher logs were administered to corps members (CMs) in the 2016 cohort in the program and comparison regions. CMs teaching English Language Arts (ELA), mathematics, or general education, in the first year of their teaching placement were eligible for the logs. Closed-ended logs were sent to eligible and consenting CMs every two weeks during the school year (15 weeks total). The instructional items in the logs differ by level (lower versus upper grades) and by content area (ELA and math). Secondary school teachers, as well as elementary school teachers in departmentalized schools, were sent the log for their content area and level every two weeks (about once per month). Elementary school teachers teaching both content areas (general education teachers) were sent the ELA log or the math log every two weeks, with the content area alternating between logs (about one log per content area per month).
A CM is included in the log sample if the CM completed at least one log during the school year. The values in the columns "Program Regions" and "Comparison Regions" are the observed mean characteristics of these CMs. Values in the "Effect Size" column are the estimated difference divided by the standard deviation for the pooled sample.

A two-tailed t-test was applied to estimated differences between program regions and comparison regions. The statistical significance of estimated differences is indicated as follows: *** = 1 percent; ** = 5 percent; * = 10 percent. Rounding may cause slight discrepancies in calculating sums and differences.

Appendix Table C.7

Characteristics of the First-Year Teaching Placements of the 2016 Cohort, Teacher Log Sample (Mathematics)

Placement Characteristic	Program Regions	Comparison Regions	Estimated Difference	Effect Size	P-Value
Grade and subject of placement					
Grade level					
Lower elementary school	26.4	30.2	-3.8	-0.08	0.666
Upper elementary school	34.4	29.7	4.7	0.09	0.646
Middle school	23.1	22.0	1.1	0.02	0.896
High school	13.8	18.1	-4.3	-0.11	0.440
Subject taught					
General education	57.5	46.5	11.0	0.20	0.347
ELA	0.9	6.6	-5.7	-0.21	0.243
Mathematics	39.5	46.5	-7.0	-0.13	0.486
Characteristics of placement schools					
School size and type					
Total enrollment	765.86	624.17	141.69	0.29	0.319
Magnet school (%)	23.6	5.4	18.2 *	0.58	0.061
Charter school (%)	20.7	31.2	-10.5	-0.22	0.482
Title I school (%)	94.2	100.0	-5.8	-0.43	0.102
Staffing					
First-year teachers at the school (%)	35.8	29.0	6.9	0.23	0.515
Teacher absences (% with 10 or more per year)	38.1	22.9	15.2 **	0.65	0.038
Certified (%)	98.1	94.4	3.7	0.28	0.344
School location (%)					
Urban	59.9	67.9	-7.9	-0.16	0.647
Suburban	31.5	20.7	10.8	0.25	0.416
Town	3.1	4.9	-1.8	-0.08	0.767
Rural	3.1	6.5	-3.4	-0.16	0.597

(continued)

Appendix Table C.7 (continued)

Placement Characteristic	Program Regions	Comparison Regions	Estimated Difference	Effect Size	P-Value
Characteristics of students in placement schools					
Chronically absent students (%)	14.5	16.8	-2.4	-0.14	0.657
Students retained (%)	3.9	5.5	-1.6	-0.23	0.295
Students with 1 or more in-school suspensions (%)	6.8	7.5	-0.7	-0.06	0.761
Students with 1 or more out-of-school suspensions (%)	12.5	14.3	-1.7	-0.11	0.688
Free or reduced price lunch students (%)	84.3	88.3	-4.0	-0.29	0.291
Students with English as a second language (%)	18.9	19.2	-0.3	-0.01	0.968
Students with an Individualized Education Plan (%)	11.4	12.3	-0.9	-0.11	0.661
Racial/ethnic composition of students (%)					
Hispanic	34.6	46.6	-12.0	-0.33	0.432
Black	47.5	43.6	3.8	0.10	0.821
White	10.3	5.5	4.8 *	0.43	0.091
Asian	2.0	2.3	-0.3	-0.06	0.854
Other	4.7	2.2	2.5	0.50	0.124
Female students (%)	48.7	49.7	-1.0	-0.17	0.236
Number of regions	8	13			
Number of corps members	80	151			

SOURCE: Teach For America administrative records for corps members, Common Core of Data (school year 2015-2016), and Office of Civil Rights Data Collection (school year 2013-2014).

NOTES: Online teacher logs were administered to corps members (CMs) in the 2016 cohort in the program and comparison regions. CMs teaching English Language Arts (ELA), mathematics, or general education, in the first year of their teaching placement were eligible for the logs. Closed-ended logs were sent to eligible and consenting CMs every two weeks during the school year (15 weeks total). The instructional items in the logs differ by level (lower versus upper grades) and by content area (ELA and math). Secondary school teachers, as well as elementary school teachers in departmentalized schools, were sent the log for their content area and level every two weeks (about once per month). Elementary school teachers teaching both content areas (general education teachers) were sent the ELA log or the math log every two weeks, with the content area alternating between logs (about one log per content area per month).

A CM is included in the log sample if the CM completed at least one log during the school year. The values in the columns "Program Regions" and "Comparison Regions" are the observed mean characteristics of these CMs. Values in the "Effect Size" column are the estimated difference divided by the standard deviation for the pooled sample.

A two-tailed t-test was applied to estimated differences between program regions and comparison regions. The statistical significance of estimated differences is indicated as follows: *** = 1 percent; ** = 5 percent; * = 10 percent. Rounding may cause slight discrepancies in calculating sums and differences.

Appendix Table C.8

Characteristics of the Focal Students, Teacher Log Sample (English Language Arts)

Characteristic	Program Regions	Comparison Regions	Estimated Difference	Effect Size	P-Value
Student has an Individualized Education Plan (%)	14.4	16.1	-1.7	-0.04	0.611
Student received same instruction as peers today (%)	91.8	89.5	2.3	0.07	0.434
Student's performance level					
Lower third	29.4	31.5	-2.1	-0.04	0.520
Middle third	40.4	40.4	0.0	0.00	0.990
Upper third	30.3	28.4	1.9	0.04	0.474
Number of minutes spent on focal subject today	79.93	73.56	6.38	0.16	0.143
Number of regions	8	12			
Number of corps members	87	138			
Number of logs	739	1143			

SOURCE: Closed-ended teacher logs administered by MDRC.

NOTES: Online teacher logs were administered to corps members (CMs) in the 2016 cohort in the program and comparison regions. CMs teaching English Language Arts (ELA), mathematics, or general education, in the first year of their teaching placement were eligible for the logs. Closed-ended logs were sent to eligible and consenting CMs every two weeks during the school year (15 weeks total). The instructional items in the logs differ by level (lower versus upper grades) and by content area (ELA and math). Secondary school teachers, as well as elementary school teachers in departmentalized schools, were sent the log for their content area and level every two weeks (about once per month). Elementary school teachers teaching both content areas (general education teachers) were sent the ELA log or the math log every two weeks, with the content area alternating between logs (about one log per content area per month). A CM is included in the log sample if the CM completed at least one log during the school year.

At the start of each log, teachers were given a randomly selected letter of the alphabet and asked to report on their instructional practices with a focal student whose name starts with that letter (or the closest letter) on a particular day that week. Teachers were also asked to report on the characteristcis of the selected focal student. The values in the columns "Program Regions" and "Comparison Regions" are the observed mean characteristics of these students, as reported by teachers in the log sample. Values in the "Effect Size" column are the estimated difference divided by the standard deviation for the full sample. A two-tailed t-test was applied to estimated differences between program regions and comparison regions. The statistical significance of estimated differences is indicated as follows: *** = 1 percent; ** = 5 percent; * = 10 percent. Rounding may cause slight discrepancies in calculating sums and differences.

Appendix Table C.9

Characteristics of the Focal Students, Teacher Log Sample (Mathematics)

Characteristic	Program Regions	Comparison Regions	Estimated Difference	Effect Size	P-Value
Student has an Individualized Education Plan (%)	13.0	15.4	-2.4	-0.06	0.460
Student received same instruction as peers today (%)	92.6	91.9	0.8	0.03	0.718
Student's performance level					
Lower third	29.4	29.4	0.0	0.00	0.995
Middle third	37.1	36.1	1.0	0.02	0.748
Upper third	33.6	34.4	-0.8	-0.01	0.835
Number of minutes spent on focal subject today	65.19	68.51	-3.32	-0.12	0.392
Number of regions	8	13			
Number of corps members	80	151			
Number of logs	628	1197			

SOURCE: Closed-ended teacher logs administered by MDRC.

NOTES: Online teacher logs were administered to corps members (CMs) in the 2016 cohort in the program and comparison regions. CMs teaching English Language Arts (ELA), mathematics, or general education, in the first year of their teaching placement were eligible for the logs. Closed-ended logs were sent to eligible and consenting CMs every two weeks during the school year (15 weeks total). The instructional items in the logs differ by level (lower versus upper grades) and by content area (ELA and math). Secondary school teachers, as well as elementary school teachers in departmentalized schools, were sent the log for their content area and level every two weeks (about once per month). Elementary school teachers teaching both content areas (general education teachers) were sent the ELA log or the math log every two weeks, with the content area alternating between logs (about one log per content area per month). A CM is included in the log sample if they completed at least one log during the school year.

At the start of each log, teachers were given a randomly selected letter of the alphabet and asked to report on their instructional practices with a focal student whose name starts with that letter (or the closest letter) on a particular day that week. Teachers were also asked to report on the characteristcis of the selected focal student. The values in the columns "Program Regions" and "Comparison Regions" are the observed mean characteristics of these students, as reported by teachers in the log sample. Values in the "Effect Size" column are the estimated difference divided by the standard deviation for the full sample. A two-tailed t-test was applied to estimated differences between program regions and comparison regions. The statistical significance of estimated differences is indicated as: follows: *** = 1 percent; ** = 5 percent; * = 10 percent. Rounding may cause slight discrepancies in calculating sums and differences.

Appendix D

Selection of the Study Regions

This appendix describes the selection of the TFA regions in the study sample. In total, 21 TFA regions were included in the analyses: 8 program regions and 13 comparison regions. The first section in this appendix discusses the selection of the program regions. The second section provides information on the selection of the comparison regions.

Program Regions

As explained in Chapter 2, when corps members (CMs) are recruited into TFA, they are assigned to a TFA region that determines the location of their placement. In turn, CMs' regions determine where they will be trained. Some regions hold their own "regional" summer institute for their CMs, whereas other regions send their CMs to one of the national summer institutes. In Summer 2016, there were six national summer institutes: Tulsa, Oklahoma; Philadelphia, Pennsylvania; Houston, Texas; Cleveland, Ohio; Greater Delta (comprising Mississippi and Arkansas); Phoenix, Arizona; and Atlanta, Georgia.

The program regions in this study are the eight TFA regions that were trained at the Tulsa national institute where the redesigned training was piloted. The eight program regions are Detroit, Idaho, Jacksonville, Miami-Dade, Greater Tulsa, Orlando, Washington (state), and the Bay Area. For the purposes of the comparative interrupted time series (CITS) analysis, it is important that these regions be similarly defined across TFA cohorts. Therefore, the study team modified the definition of two of the regions for the study:

- **Bay Area:** In the Bay Area region, some CMs went to the national institute in Tulsa in summer 2016, whereas other CMs attended the regional Bay Area institute. The following Bay Area CMs were trained in Tulsa: (1) CMs whose summer teaching placement was in mathematics, science, or English; and (2) CMs in other subject areas who applied to TFA under later or special application deadlines. Thus, for the purposes of this study, the "Bay Area" region only includes the CMs that met these criteria, and these criteria were applied to all TFA cohorts.
- **Oklahoma:** Before 2016, the Tulsa region was part of a larger Oklahoma region that also included Oklahoma City. Therefore, to make the "Tulsa" region consistent over time, the study team excluded from the sample CMs in the 2012 to 2015 cohorts who were in the Oklahoma region and placed in Oklahoma City, and redefined the Oklahoma region for these earlier cohorts as the Tulsa region.

Comparison Regions

The study team selected comparison regions from among the TFA regions that were trained at one of the other five national institutes in summer 2016. The study team excluded the following TFA regions from the comparison pool:

- CMs in TFA regions that received their training at a regional institute in summer 2016.
- CMs from TFA regions that were trained at a national institute but whose context was very different from that of the program regions. This included New Orleans (because all schools are charters); Appalachia, North Carolina Piedmont Triad, and Los Angeles (because no CMs or a very small fraction of CMs had been placed in elementary schools in the past few years); South Dakota (because it is entirely rural and unique in terms of the race-ethnicity of students in the placement schools); New Mexico (because it is distinctive in terms of the race-ethnicity of students in the placement schools); and Hawaii (because of its geographical distinctiveness).

The resulting comparison pool included 28 TFA regions.

The study team then chose the comparison regions for the study by matching each of the eight program regions to the two TFA regions in the comparison pool whose 2015 or 2014 cohort of CMs had the most "similar" outcomes and characteristics. This is called nearest-neighbor matching. In this study, the "similarity" between a program region and its potential comparison regions was measured using the Euclidian distance, which is a metric that captures the distance between two regions across several matching variables measured at the region level. The Euclidian distance across M matching variables is defined as: $D = \sqrt{\sum_M (T_m - C_m)^2}$, where T_m is the value of characteristic m for a program region and C_m is the value of characteristic m for the potential comparison region. The smaller is the Euclidian distance between two regions, the more similar they are to each other based on the matching variables.

Given the small number of comparison regions available for matching, the matching characteristics were limited to the following CM placement characteristics and outcomes measured at the region level:

- The retention rate in the fall of CMs' second year of teaching (2014 cohort);[1]
- The retention rate in the fall of CMs' first year of teaching (2015 cohort);
- CMs' average satisfaction with the summer training measured at the end of institute (2015 cohort);
- The percentage of CMs teaching in a suburb (2015 cohort);
- The percentage of CMs teaching in lower elementary school (2015 cohort).

These matching characteristics produced the most optimal set of matches, as indicated by the fact that the inclusion of additional characteristics in the matching process made the

[1]The 2014 cohort was used for this outcome (rather than the 2015 cohort) because data were not available for the 2015 cohort on this outcome at the time of the matching.

program and comparison regions less similar with respect to the outcomes of the 2014 or 2015 cohorts of CMs.

Because the comparison pool was small, the study team chose comparison regions with replacement — meaning that a comparison region could be chosen as the match for more than one program region. Though matching with replacement reduced the number of unique comparison regions in the study, it maximized the similarity of the comparison regions and the program regions with respect to the matching variables.

The 13 unique comparison regions were Charlotte, Northeast Ohio (Cleveland), Southwest Ohio, Atlanta, Mississippi, Rio Grande Valley, San Antonio, Indianapolis, California Capital Valley, District of Columbia, Rhode Island, Phoenix, and Colorado. The three comparison regions that were chosen more than once were given a greater weight in the analyses.[2]

[2]Three comparison regions were each matched to two comparison regions: Rhode Island, Indianapolis, and Atlanta. See Appendix E for further information on the weighting.

Appendix E

Statistical Models, Minimum Detectable Effect Sizes, and Standard Deviations

This appendix provides information on various technical aspects of the analysis. The first section describes the statistical models used to estimate the effect of the redesigned summer training that was piloted at the national institute in Tulsa in summer 2016. The second section discusses the minimum detectable effect sizes for the study's outcomes. The final section provides the standard deviations for the outcome measures in this study, which can be used to convert the estimated effects in this report into effect sizes.

Statistical Models

As explained in Chapter 2, two quasi-experimental study designs were used to estimate the effect of the redesigned summer training. This section presents the statistical models used for each study design.

Comparative Interrupted Time Series Design (CITS)

In this study, the effects of the redesigned training on corps members' (CM) perceptions of the training, their commitment to teaching, and their retention were estimated using a comparative interrupted time series (CITS) design.

To account for the clustering of CMs within regions and cohorts, the effect of the redesigned training was estimated by fitting a two-level model to CM data on the outcomes of interest. The first level is a model the variation in CMs' outcomes between CMs within a region and cohort, and the second level is a model of the variation in outcomes between regions and cohorts:

Level 1 (CMs)

$$Y_{irc} = \lambda_{rc} + \sum_{k=1}^{K} \psi_k Z_{k,irc} + \varepsilon_{irc}$$

Level 2 (region/cohort)

$$\lambda_{rc} = \sum_m \delta_m REG_m + \sum_m \phi_m REG_m * RELYR_c + \alpha POST_c + \beta PROGRAM_r * POST_c + u_{rc}$$

where *i* denotes CMs, *r* is regions, and *c* is cohorts. In a CITS design, time (as represented by the TFA cohorts) is measured relative to the intervention. Therefore, *c* is equal to 1 for CMs in the 2016 cohort (some of whom attended the new training at the national institute in Tulsa), and it is equal to 0 for the 2015 cohort, -1 for the 2014 cohort, -2 for the 2013 cohort, and -3 for the 2012 cohort.[1]

The variables in the model are defined as follows:

Y_{irc} = Outcome for CM *i* in cohort *c* and region *r*

[1]Two of the program regions (Idaho and Orlando) are newer TFA regions that have only existed for the 2015 cohort onward. (So, there is only one baseline cohort for these regions.) The statistical model can accommodate this variability in the length of the baseline period across regions.

REG_m = A set of M dichotomous indicators for region (=1 for region *m* and 0 otherwise).

$RELYR_c$ = Continuous variable for time (cohort) centered at the last baseline (pre-intervention) cohort, such that RELYR=0 for the 2015 cohort

$PROGRAM_r$ = Dichotomous indicator for whether region *r* is a program region (=1 if a region was trained at the national institute in Tulsa; =0 for the comparison regions)

$POST_c$ = Dichotomous indicator for the first cohort to have an opportunity to receive the new training (=1 if 2016 cohort; 0 otherwise)

$Z_{k,ir}$ = A set of K characteristics for CM *i*, including their prior experience, their demographic characteristics, and the characteristics of their placement.[2]

ε_{irc} = Unexplained variation in Y within regions and cohorts

u_{rc} = Unexplained variation in Y between regions and cohorts

From the model, the following quantities of interest can be obtained:

δ_j = Intercept of the baseline trend for each program and comparison region in the study[3]

ϕ_j = Slope of the baseline trend for each program and comparison region in the study[4]

[2]The model controlled for the following characteristics of CMs: their age; race or ethnicity; sex; their pre-TFA status (professional, graduate school, or undergraduate); whether they were the first in their family to attend college; whether they were a Pell grant recipient; their highest educational attainment; whether their degree was in an education-related field; and their overall application score when applying to TFA (a composite of their scores measuring different types of skills). The model also controlled for the characteristics of CMs' placement: grade level and subject taught; whether the CM's school was a Title I school; whether their school was a charter school; its total enrollment, the school's level (elementary, middle, high); the school's location (urban, suburban, town, rural); its pupil-teacher ratio; the percent of certified teachers at the school; the percent of teachers at the school absent for 10 or more days; the percent of teachers at the school in their first or second year (novice); the percent of English as a second language students; the percent of students eligible for free or reduced price lunch; the racial or ethnic composition of students; the percent of girls; the percent of students with an individualized education program (IEP); the percent of students who were chronically absent; the percent of students retained; the percent of students with one or more in-school suspensions; and the percent of students with one or more out-of-school suspensions. The characteristics of a CM's school were measured using the most recent data available (the 2013-2014 or the 2015-2016 school year, depending on the variable).

[3]The intercept can also be thought of as the predicted outcome for the 2015 cohort of CMs in a region.

[4]In this model, the baseline trend was separately estimated for each region in the study. This strategy was used to account for the longitudinal nature of the data, whereby cohorts are nested within regions.

α = Deviation from baseline trend for the 2016 cohort in the comparison regions

$\alpha + \beta$ = Deviation from baseline trend for the 2016 cohort in the program regions

β = Estimated effect of the redesigned training on the 2016 cohort of CMs

Importantly, in this model, β represents the estimated effect of the redesigned summer training — that is, the deviation from baseline trend for the program regions minus the deviation from trend for the comparison regions. As explained in Chapter 2, some comparison regions were chosen as the "match" for *two* different program regions.[5] These comparison regions were given a proportionally greater weight in the analysis.[6]

The main model used for the analysis also controlled for the characteristics of CMs and their schools. Missing data on these covariates were imputed using multiple imputation (MI) with five imputed datasets. The analyses were estimated for each of the five imputed datasets, and then combined across the five sets of results using appropriate methods.[7] As a sensitivity analysis, however, effects were also estimated using a "complete case" sample of CMs with no missing data on any of the covariates; the results from this sensitivity analysis are similar to the findings based the imputed datasets, which suggests that using multiple imputation did not alter or bias the results. (See Appendix F.)

Cross-Sectional Program-Comparison Group Design

Because data on CMs' use of instructional strategies aligned to college- and career-readiness standards were available for the 2016 cohort only, the effects of the redesigned training on this outcome were estimated by comparing the strategies used by CMs in the program and comparison regions. As explained in Chapter 2, each CM completed up to 15 closed-ended logs on the practices they used with different focal students. A three-level model was used to account for the clustering of logs within CMs, and of CMs within regions. The first level is a model of the variation in a CM's use of standards-aligned instructional strategies across logs; the second level represents the variation in instructional strategies across CMs within regions; and the third level is a model of the variation in strategies between regions:

[5]The following comparison regions were chosen as the match for two program regions: Rhode Island, Indianapolis, and Atlanta.

[6]The twice-selected comparison regions were given a weight of 2, and the other comparison regions (those chosen once) and the program regions were given a weight of 1. The weights were then normalized so that their sum equals to the number of CMs in the analysis. As a sensitivity check, the analysis was also conducted by including the twice-selected comparison regions in the dataset twice as two unique observations, and estimating an unweighted model with cluster-robust standard errors to account for the duplicated data. The findings from this sensitivity analysis were very similar to those from the main (weighted) analysis.

[7]The imputation and combining of results were performed using PROC MI and PROC MIANALYZE in SAS.

Level 1 (logs)

$$Y_{gir} = \lambda_{ir} + \sum_{j=1}^{J} \phi_j X_{j,gir} + \varepsilon_{gir}$$

Level 2 (CMs)

$$\lambda_{ir} = \rho_r + \sum_{k=1}^{K} \psi_k Z_{k,ir} + \eta_{ir}$$

Level 3 (regions)

$$\rho_r = \delta + \beta PROGRAM_r + u_r$$

where *g* denotes log, *i* denotes CMs, and *r* is regions.

The variables in the model are defined as follows:

Y_{gir}	=	Instructional strategy reported in log *g* by CM *i* in region *r*
$X_{j,gr}$	=	A set of J characteristics for the focal student that is the focus of log *g*[8]
$Z_{k,ir}$	=	A set of K characteristics for CM *i*, including their prior experience, their demographic characteristics, and the characteristics of their placement[9]
$PROGRAM_r$	=	Dichotomous indicator for whether region *m* is a program region (=1 if a region was trained at the national institute in Tulsa; =0 for the comparison regions)
$\boldsymbol{\varepsilon_{gir}}$	=	Unexplained variation in Y between logs within CMs
η_{ir}	=	Unexplained variation in Y between CMs within regions
u_r	=	Unexplained variation in Y between regions

From the model, the following quantities of interest can be obtained:

δ	=	Average adjusted mean outcome for the comparison regions
$\delta + \beta$	=	Average adjusted mean outcome for the program regions
β	=	Estimated effect of the redesigned training

The key parameter of interest is β, which represents the estimated effect of the redesigned summer training — that is, the adjusted difference between the program and comparison regions with respect to CMs' average use of an instructional strategy. Similar to the CITS analysis, this

[8]The model controlled for the following characteristics of the focal students: whether the focal student had an IEP, number of minutes the student was in the class, and the student's performance in ELA or mathematics relative to other students in the class (bottom, middle, or lower third).

[9] See footnote 22 for a list of characteristics.

cross-sectional analysis was also weighted to account for the fact that some comparison regions were chosen more than once.

Similar to the CITS analysis, the cross-sectional analysis controlled for the characteristics of focal students, of CMs, and of their placements. Controlling for these characteristics is especially important in a cross-sectional analysis, because it can help minimize the extent to which differences in pre-existing factors (such as the characteristics of CMs' schools) are confounded with the effect of the redesigned summer training. Similar to the analysis for the CITS design, missing data on the covariates included in the model were imputed using multiple imputation with five imputed datasets. The findings from a sensitivity analysis based on a "complete case" sample of CMs with no missing data on any of the covariates are similar to the findings based the imputed datasets, which suggests that using multiple imputation did not alter or bias the results. (See Appendix F.)

Minimum Detectable Effects

Appendix Tables E.1 and E.2 show the minimum detectable effects (MDE) for the estimated effect of the redesigned training on CMs' outcomes.[10] Formally, the MDE is the smallest true program effect that can be detected with a reasonable degree of power (in this case, 80 percent) for a given level of statistical significance (in this case, 10 percent for a two-tailed test). The MDE were obtained by multiplying the standard error of the estimated effect by a multiplier of 2.5 (for survey outcomes and retention) or 2.6 (for standards-aligned instructional strategies).[11] The minimum detectable effect size (MDES) is the MDE scaled as an effect size.

The MDES for CMs' perceptions of the training ranges from 0.30 to 0.36, depending on the survey wave. The MDES for CMs' commitment to teaching is 0.39 to 0.51. It is reasonable to expect the redesigned training to have effects of this size because these measures are well aligned with the training. The MDE for the CMs' use of culturally responsive text and the citing of text evidence is 17 to 18 percentage points (0.33 to 0.40) and the MDES for self-reported math instruction is 0.37. These MDES are in the range of effects found in prior studies of teacher professional development interventions.[12]

The MDE for CM retention is 6.5 to 11.3 percentage points (effect size = 0.27 to 0.30), depending on the time point. Thus, the estimated effect of the training would have to be quite large to be statistically significant. This raises the possibility that there may have been a small but positive effect, but that the study was not able to reliably detect it. However, an advantage of the CITS design is that trends can be visually examined to look for signs of an effect. Yet as discussed in Chapter 5, visually it does not appear as though the redesigned training had an effect on the retention rates of the first cohort of CMs to receive it.

[10]The table for instructional strategies focuses on the MDES for the three confirmatory outcomes.

[11]The multiplier is larger for the instructional outcomes because of the smaller sample size.

[12]Prior MDRC studies have found effect sizes as high as 0.53 on teacher outcomes. For example, see Garet et al. (2008).

Appendix Table E.1

Minimum Detectable Effect (MDE) and Minimum Detectable Effect Size (MDES), Survey Outcomes and Retention

Outcome	MDE	MDES
Fall of first year		
Value of the summer training (1-7)	0.65	0.36
Commitment to teaching and equity (1-7)	0.37	0.45
Retention (%)	6.48	0.30
Middle of first year		
Value of the summer training (1-7)	0.63	0.35
Commitment to teaching and equity (1-7)	0.43	0.51
End of first year		
Value of the summer training (1-7)	0.58	0.32
Commitment to teaching and equity (1-7)	0.33	0.39
Retention (%)	9.64	0.27
Fall of first year		
Value of the summer training (1-7)	0.53	0.30
Commitment to teaching and equity (1-7)	0.30	0.39
Retention (%)	11.33	0.28

SOURCE: Teach For America corps member surveys and administrative records.

NOTES: The MDE in this table is calculated by multiplying the standard error of the estimated effect by 2.5. A statistical significance level of 10 percent is assumed. The MDES is the MDE divided by the standard deviation for the last baseline cohort of corps members.

The survey outcome measures are based on survey items with a 7-point agreement scale. The "value of the summer training" measure is based on the following question: "Overall, I believe the preparation I received from Teach For America prior to starting at my school was valuable in my efforts to become a successful teacher." The commitment to teaching and equity measure is a composite created from eight items related to CMs' beliefs in TFA's mission of ensuring that each child has an opportunity to receive an excellent education, and CMs' beliefs in their own capacity to effectively contribute to this mission.

Appendix Table E.2

Minimum Detectable Effect (MDE) and Minimum Detectable Effect Size (MDES), Self-Reported Standards-Aligned Instructional Strategies

Outcome	MDE	MDES
Text for the lesson featured author/characters/community with similar background to student (% of classes)	17.29	0.33
Student was asked to cite text evidence in their writing (% of classes)[a]	17.87	0.40
Math instructional practice composite (z-score)[b]	0.36	0.37

SOURCE: Closed-ended teacher logs administered by MDRC.

NOTES: The MDE in this table is calculated by multiplying the standard error of the estimated effect by 2.6. A statistical significance level of 10 percent is assumed. The MDES is the MDE divided by the standard deviation for the sample.

[a]This item is only available for upper elementary and secondary school teachers.

[b]This composite combines the items in the math logs that map onto college and career-readiness standards for math instruction. The scale of the items differs, so each item was first z-scored based on the mean and standard deviation of the comparison group logs; these z-scores were then averaged across items, and the average score was z-scored, again, based on the mean and standard deviation for the comparison logs. Because some of the log items differ across levels, the composite score was created separately for the lower elementary school or upper elementary or secondary school logs. The reliability (Cronbach's alpha) for the composite is 0.71 at the lower elementary school level and 0.75 at the upper elementary or secondary school level.

Standard Deviations

The tables in this section provide the standard deviation for the outcomes of the CMs in the study, by study group (program and comparison regions) and for both regions pooled together. Appendix Tables E.3 to E.5 present the standard deviations for CMs' retention in the TFA program, their perceptions of the training, their commitment to teaching, and their use of standards-aligned instructional strategies, respectively.[13]

Note that when calculating the effect sizes presented in Chapters 4 and 5 of this report, the standard deviation for the pooled sample (program and comparison) was used. For estimated effects on CMs' perceptions of the training, commitment to teaching, and retention, the standard deviation for the last baseline cohort (2015 cohort) was used; for estimated effects on standards-aligned instructional strategies, the standard deviation for the 2016 cohort was used.

[13]The table for instructional strategies focuses on outcomes that are confirmatory, available at all grade levels, or both.

Appendix Table E.3

Standard Deviations (and Number of Corps Members) for Retention Outcome

	2015 Cohort			2016 Cohort		
Follow-up time point	Program Regions	Comparison Regions	Pooled	Program Regions	Comparison Regions	Pooled
Fall of first year (%)	24.56 (420)	19.81 (887)	21.49 (1,307)	24.46 (361)	20.37 (815)	21.74 (1,176)
End of first year (%)	36.44 (420)	36.12 (887)	36.24 (1,307)	35.16 (361)	38.67 (815)	37.61 (1,176)
Fall of second year (%)	39.69 (420)	41.78 (887)	41.11 (1,307)	39.80 (361)	44.31 (815)	42.96 (1,176)

SOURCE: Teach For America administrative records for corps members.

NOTES: A cohort is defined as a group of TFA corps members (CMs) who joined TFA and attended the summer institute the same year. The 2015 cohort received their summer training before the launch of the redesigned institute in Tulsa. The 2016 cohort is the first cohort of CMs trained after the launch of the redesigned institute. The values in the table show the standard deviation (and number of teachers) by cohort for the program regions, the comparison regions, and both groups pooled together.

Appendix Table E.4

Standard Deviations (and Number of Corps Members) for Survey Outcomes

Survey Wave and Outcome	2015 cohort			2016 cohort		
	Program Regions	Comparison Regions	Pooled	Program Regions	Comparison Regions	Pooled
Fall of first year						
Value of the summer training (1-7)	1.65 (360)	1.86 (781)	1.79 (1,141)	1.82 (317)	1.84 (721)	1.85 (1,038)
Commitment to teaching and equity (1-7)	0.74 (360)	0.85 (781)	0.82 (1,141)	0.86 (317)	0.93 (721)	0.91 (1,038)
Middle of first year						
Value of the summer training (1-7)	1.64 (335)	1.83 (746)	1.77 (1,081)	1.81 (301)	1.83 (670)	1.87 (971)
Commitment to teaching and equity (1-7)	0.73 (335)	0.90 (746)	0.85 (1,081)	0.76 (301)	0.94 (670)	0.89 (971)
End of first year						
Value of the summer training (1-7)	1.76 (331)	1.84 (715)	1.82 (1,046)	1.88 (296)	1.95 (647)	1.95 (943)
Commitment to teaching and equity (1-7)	0.78 (331)	0.87 (715)	0.84 (1,046)	0.85 (296)	0.92 (647)	0.90 (943)
Fall of second year						
Value of the summer training (1-7)	1.67 (328)	1.83 (660)	1.77 (988)	1.33 (268)	1.52 (568)	1.47 (836)
Commitment to teaching and equity (1-7)	0.68 (328)	0.82 (660)	0.77 (988)	0.69 (268)	0.85 (568)	0.80 (836)

SOURCE: Teach For America corps member surveys.

NOTES: A cohort is defined as a group of TFA corps members (CMs) who joined TFA and attended the summer institute the same year. The 2015 cohort received their summer training before the launch of the redesigned institute in Tulsa. The 2016 cohort is the first cohort of CMs trained after the launch of the redesigned institute. The values in the table show the standard deviation (and number of teachers) by cohort for the program regions, the comparison regions, and both groups pooled together.

The reported outcome measures are based on survey items with a 7-point agreement scale. The "value of the summer training" measure is based on the following question: "Overall, I believe the preparation I received from Teach For America prior to starting at my school was valuable in my efforts to become a successful teacher." The commitment to teaching and equity measure is a composite created from eight items related to CMs' beliefs in TFA's mission of ensuring that each child has an opportunity to receive an excellent education, and CMs' beliefs in their own capacity to effectively contribute to this mission.

Appendix Table E.5

Standard Deviations (and Number of Logs) for Self-Reported Instructional Practices

Outcome	Program Regions	Comparison Regions	Pooled
English Language Arts			
Text used for the lesson featured an author, characters, and/or a community with similar background to focal student (% of classes)	48.6 (625)	55.4 (919)	52.8 (1,544)
Student was asked to cite text evidence in their writing (% of classes)[a]	41.3 (546)	46.7 (828)	44.7 (1,374)
These areas were a focus of instruction (% of classes):			
Listening to a text read aloud	49.7 (733)	55.1 (1,128)	53.4 (1,861)
Reading with a partner or group	44.4 (729)	49.6 (1,123)	47.6 (1,852)
Reading independently	49.4 (732)	53.4 (1,127)	52.0 (1,859)
Mathematics			
Math instructional practice composite (z-score)[b]	0.86 (621)	1.00 (1,177)	0.96 (1,798)
These areas were a focus of instruction (% of classes):			
Solving a math problems with an algorithm	49.4 (618)	56.4 (1,174)	54.3 (1,792)
Representing/analyzing relationships using illustrations	50.0 (619)	56.4 (1,174)	54.3 (1,793)
Critiquing the mathematical reasoning of others	35.4 (613)	39.9 (1,167)	38.4 (1,780)
Discussing math with a peer(s)	49.6 (616)	55.6 (1,173)	53.6 (1,789)
Justifying responses to a math problem(s)	48.3 (618)	55.4 (1,172)	53.1 (1,790)
Writing down mathematical thinking	48.9 (618)	54.9 (1,174)	52.9 (1,792)
Applying math concepts to a "real-world" problem(s)	47.7 (619)	54.8 (1,173)	52.4 (1,792)
Working in a math team(s) or pair(s)	49.1 (617)	55.6 (1,172)	53.4 (1,789)
Investigating a multistep problem(s)	48.8 (619)	56.0 (1,171)	53.7 (1,790)
Practicing or relearning prerequisite concepts	48.4 (617)	53.6 (1,174)	51.9 (1,791)

(continued)

Appendix Table E.5 (continued)

Outcome	Program Regions		Comparison Regions		Pooled	
Receiving remediation	33.2	(610)	42.1	(1160)	39.3	(1770)
Continuing to persevere on a challenging problem	47.9	(616)	54.0	(1172)	52.0	(1788)

SOURCE: Closed-ended teacher logs administered by MDRC.

NOTES: Online teacher logs were administered to corps members (CMs) in the 2016 cohort in the program and comparison regions. CMs teaching English Language Arts (ELA), mathematics, or general education in the first year of their teaching placement were eligible for the logs. Secondary school teachers, as well as elementary school teachers in departmentalized schools, were sent the log for their content area and level every two weeks (about twice per month). Elementary school teachers teaching both content areas (general education teachers) were sent the ELA log or the math log every two weeks, with the content area alternating between logs (about one log per content area per month). A CM is included in the log sample if they completed at least one log during the school year. At the start of each log, teachers were given a randomly selected letter of the alphabet and asked to report on their instructional practices with a focal student whose name starts with that letter (or the closest letter) on a particular day that week. A CM is included in the log sample if the CM completed at least one log during the school year.

The values in the table show the standard deviation (and the number of logs) by outcome for the program regions, the comparison regions, and both groups pooled together.

[a]This item is only available for upper elementary and secondary school teachers. The number of teachers included in this analysis is 60 in the program group and 93 in the comparison group.

[b] This composite combines the items in the math logs that map onto the TFA core components. The scale of the items differs, so each item was first z-scored based on the mean and standard deviation of the comparison group logs; these z-scores were then averaged across items, and the average score was z-scored, again, based on the mean and standard deviation for the comparison logs. Because some of the log items differ across levels, the composite score was created separately for the lower elementary school and upper elementary or secondary school logs. The reliability (Cronbach's alpha) for the composite is 0.71 at the lower elementary school level and 0.75 at the upper elementary or secondary school level.

Appendix F

Sensitivity Analyses

This appendix discusses several analyses that the study team conducted to examine whether estimated effects on corps members' (CM) outcomes are sensitive to how the analysis was conducted. Because the study designs used in this evaluation are quasi-experimental, the findings in this report cannot be interpreted as the true causal effect of the redesigned training on CMs' outcomes. However, when the findings are robust across sensitivity analyses — as they are in this study — this makes it more likely that the results were due to the effect of the new training. The appendix describes the types of sensitivity analyses that the study team conducted, followed by a presentation and overview of the results.

Overview of Analyses

In this study, estimated effects on CMs' retention in the TFA program, their perceptions of the training, and their commitment to teaching were based on a comparative interrupted time series (CITS) design, which is the more rigorous of the two study designs used in this evaluation. The study team conducted several sensitivity analyses to explore various factors that could compromise the credibility of the comparison group in the CITS design:

- **Design validation:** One of the advantages of having access to data for several cohorts of CMs is that the results of the CITS design can be validated by looking at the estimated "effect" of the redesigned training on the last baseline cohort of CMs (the 2015 cohort). By definition, this effect should be zero because the redesigned training had not yet been launched, so this provides a useful benchmark for validating the CITS design and the comparison regions. If the effect of the redesigned training on CMs' outcomes is not statistically significant, then this lends greater credibility to the CITS design and the main findings presented in the report. In practice, this sensitivity analysis was set up by pretending that the 2015 cohort of CMs is the intervention cohort in the CITS design. The selection of comparison regions was then conducted, again, using the newly defined baseline period (which now excluded the 2015 cohort) and the effects of the redesigned training on CMs' outcomes were re-estimated.[1]

- **Complete case analysis:** As explained in Appendix E, missing data on the baseline covariates (CMs' demographic and placement characteristics) were imputed. As a sensitivity analysis, estimated effects were re-estimated using only CMs for whom there were complete data on these characteristics, to verify that the imputation procedures were not causing bias in the results. The results from this analysis are not perfectly comparable to the main findings, because differences in the findings could be due to variation in the true effect of the training across different subgroups of CMs (that is, effects could have been different for the subgroup of CMs with complete data), as opposed to

[1]This analysis excluded two program regions — Orlando and Idaho — and their comparison regions, because these regions are newer and did not have enough baseline cohorts to be included in this sensitivity analysis.

indicating a problem with the imputation. However, this remains a useful sensitivity analysis.

- **1-to-1 match:** As explained in Chapter 2 and Appendix D, each program region was matched to the two comparison regions whose prior cohorts of CMs had the most similar outcomes and characteristics. This is called nearest-neighbor matching. In practice, there is an inherent trade-off with respect of the number of "neighbors" chosen for each program region. The greater the number of matches, the larger the sample is, and the greater is the study's ability to statistically detect effects. However, the more matches that are chosen, the more dissimilar the matches are likely to be to the program regions. For the primary analysis, two comparison regions (a 1:2 match) was used to maximize statistical power while also maintaining comparable groups. However, these groups could be made even more closely comparable by only including the "best" match for each program region. Therefore, in the first sensitivity analysis, only the most similar of each program region's matches was included in the comparison group. This analysis was based on a smaller sample, so effects are less likely to be statistically detectable; however, this analysis is still useful for looking at consistency in the pattern of estimated effects.

- **Propensity score matching:** As explained in Appendix D, the main analysis in this study used the Euclidian distance as a composite measure of the "similarity" between the program regions and the potential comparison regions for the purposes of matching. An alternative measure of distance is the propensity score.[2] The literature on matching indicates that the choice of matching method should not matter, as long as the pool of comparison units (for example, regions) from which to choose is suitable and that pre-intervention data on the outcomes of interest (for example, from prior cohorts) are available.[3] Thus, examining the sensitivity of the findings to the choice of distance metric is a useful analysis. If the estimated effect of the redesigned training *is* sensitive to how the comparison group was chosen, then this would indicate that the comparison regions may not provide a credible counterfactual for the program regions schools. Therefore, for the purposes of the sensitivity analysis, comparison regions were chosen using propensity score matching.[4] For each program

[2]Rosenbaum and Rubin (1983).

[3]Cook, Shadish, and Wong (2008); Somers et al. (2013).

[4]The dependent variable in the propensity score model was an indicator for whether a region went to the national institute in Tulsa; the independent variables were the same matching characteristics as were used in the main matching approach. (See Appendix D.) A logistic regression was used to fit this model to the data. The regression coefficients on the independent variables were then used to predict each region's likelihood of being a program region, given the characteristics and outcomes of prior cohorts of CMs. This predicted probability is defined as the *propensity score*. In practice, the *logit* of the propensity score was used for matching (Rubin 2001).

region, the two regions in the comparison pool with the most similar propensity score were included in the comparison group.

These analyses were used to explore the sensitivity of the effect of the redesigned training on retention as well as CMs' perceptions of the training and commitment to teaching. For the latter domains, an additional sensitivity analysis was conducted:

- **Stable sample:** In the main analysis, estimated effects on CMs' perceptions of the training and their commitment to teaching are based on the survey sample, whose size varies across survey waves (it is smaller for later survey waves as CMs leave TFA). This means that differences in the estimated effect of the redesigned training across survey waves could be due to shifts in the sample. Therefore, estimated effects were re-estimated based on a stable survey sample, which included only the CMs who were included in *all* survey waves.

Sensitivity analyses were also conducted to explore the robustness of findings related to CMs' use of instructional strategies aligned with college- and career-readiness standards. Given the study design and available data, two of the sensitivity analyses described above were applicable: the complete case analysis and the 1-to-1 match analysis.

Findings

Appendix Tables F.1 to F.3 present the sensitivity analysis results for the estimated effect of the redesigned training on retention rates, CMs' perceptions of the training and their commitment to teaching, and CMs' use of standards-aligned instructional practices, respectively. As a reference point, the first column in each table shows the estimated effects presented in the report. The next set of columns show the estimated effect from each sensitivity analysis. Ideally, the results in all columns should be similar to the main findings in the first column, with the exception of the results of the "design validation" analysis, which should not be statistically significant because there should be no "effects" for the 2015 cohort.

As shown in Appendix Tables F.1 and F.2, which pertain to effects on retention rates and CMs' perceptions and commitment, the findings from the "validation analysis" confirm that estimated effects for the 2015 cohort are not statistically significant, as expected. Additionally, none of the other sensitivity analyses that were conducted contradict the key finding that the redesigned training does not appear to have improved CMs' retention rates, their perceptions of the training, or their commitment to teaching. Related to CMs' perceptions of the training, the sensitivity findings consistently indicate that the new training may have had a negative effect on CMs' perceptions during their first year of teaching, but that this effect appears to dissipate in the fall of their second year. At that point in time, estimated effects on CMs' perceptions of the value of the summer training is not statistically significant, with the exception of one analysis (1:1 matching) where the estimated effect is statistically significant at the 10 percent level. However, across

Appendix Table F.1
Estimated Effects on Retention Outcomes, by Sensitivity Analysis

Outcome and Follow-Up Year	Main Sample[a]	Design Validation[b]	Complete Case[c]	1:1 Matching[d]	Propensity Score Matching[e]
Retention (%)					
Fall of first year	-0.65	-2.94	-0.22	1.01	1.07
End of first year	2.49	-6.69	5.10	5.17	5.40
Fall of second year	2.11	-3.65	7.32	3.81	5.88
Number of corps members	8,324	4,840	4,801	6,068	4,349

SOURCES: Teach For America Administrative Records for corps members.

NOTES: A two-tailed test was applied to estimated effects. The statistical significance of estimated effects is indicated as follows: *** = 1 percent; ** = 5 percent; * = 10 percent.

[a]The estimated effects in this colum are the main findings presented in the report.

[b]To estimate the effects in this column, the selection of comparison regions was conducted using the same process as was used for the main study sample, but in this case pretending that the last baseline cohort (2015 cohort) was the first cohort to participate in the redesigned training.

[c]Estimated effects in this column are based on teachers with complete (non-missing) data on the characteristics used as covariates in the analysis.

[d]Estimated effects in this column are based on a set of comparison regions that includes only the best match for each program region (1-to-1 nearest neighbor), rather than the two best matches (1-to-2 nearest neighbor).

[e]Estimated effects in this column are based on comparison regions that were chosen using propensity scores as an overall measure of the similarity between regions, rather than Euclidian distance.

all analyses, the magnitude of the estimated effect is numerically smaller in CM's second year than in their first year. Taken together, the consistency of these findings strongly suggests that the redesigned training did not improve these outcomes.

Appendix Table F.3, which focuses on CMs' use of the three confirmatory standards-aligned instructional practices, also shows that the findings from the main analysis are robust to sample definitions. Across all analyses, CMs in the program regions do not appear to have used these practices more often than CMs in the comparison regions.

Appendix Table F.2

Estimated Effects on Survey Outcomes, by Sensitivity Analysis

Outcome	Main Sample	Design Validation[b]	Complete Case[c]	1:1 Matching[d]	Propensity Score Matching[e]	Stable Sample[f]
Fall of first year						
Value of the summer training (1-7)	-0.62 **	-0.05	-0.70 *	-0.73 **	-0.67 *	-0.46
Commitment to teaching and equity (1-7)	-0.01	-0.04	-0.08	0.01	-0.08	0.04
Number of regions	21	16	21	15	15	21
Number of corps members	7321	4246	4375	5355	3857	5313
Middle of first year						
Value of the summer training (1-7)	-1.01 ***	0.16	-1.11 ***	-1.11 ***	-1.08 ***	-0.98 ***
Commitment to teaching and equity (1-7)	0.00	-0.04	-0.01	0.10	0.00	0.01
Number of regions	21	16	21	15	15	21
Number of corps members	6979	4050	4185	5113	3675	5313
End of first year						
Value of the summer training (1-7)	-0.58 **	0.08	-0.70 **	-0.76 ***	-0.80 **	-0.68 ***
Commitment to teaching and equity (1-7)	-0.09	-0.17	-0.11	-0.04	-0.13	-0.06
Number of regions	21	16	21	15	15	21
Number of corps members	6637	3876	3983	4853	3525	5313
Fall of second year						
Value of the summer training (1-7)	-0.34	0.03	-0.37	-0.43 *	-0.35	-0.31
Commitment to teaching and equity (1-7)	-0.06	-0.07	-0.03	-0.01	-0.11	0.01
Number of regions	21	16	21	15	15	21
Number of corps members	6305	3746	3874	4619	3345	5313

(continued)

Appendix Table F.2

SOURCES: Teach For America corps member surveys.

NOTES: A two-tailed test was applied to estimated effects. The statistical significance of estimated effects is indicated as follows: *** = 1 percent; ** = 5 percent; * = 10 percent.

[a]The estimated effects in this colum are the main findings presented in the report.

[b]To estimate the effects in this column, the selection of comparison regions was conducted using the same process as was used for the main study sample, but in this case pretending that the last baseline cohort (2015 cohort) was the first cohort to participate in the redesigned training.

[c]Estimated effects in this column are based on teachers with complete (non-missing) data on the characteristics used as covariates in the analysis.

[d]Estimated effects in this column are based on a set of comparison regions that includes only the best match for each program region (1-to-1 nearest neighbor), rather than the two best matches (1-to-2 nearest neighbor)

[e]Estimated effects in this column are based on comparison regions that were chosen using propensity scores as an overall measure of the similarity between regions, rather than Euclidian distance.

[f]Estimated effects in this column are based on a stable sample of teachers who responded to the TFA survey in all survey waves.

Appendix Table F.3
Estimated Effects on Self-Reported Instructional Strategies Aligned with College- and Career-Readiness Standards, by Sensitivity Analysis

Outcome	Main Sample[a]	Complete Case[b]	1:1 Matching[c]
English Language Arts			
Text used for the lesson featured an author, characters, and/or a community with similar background to focal student (% of classes)	-10.9	-11.4	-2.2
Student was asked to cite text evidence in their writing (% of classes)	-5.4	-4.3	-4.6
Number of regions	20	20	15
Number of corps members	225	136	172
Number of logs	1882	1170	1439
Mathematics			
Math instructional practice composite (z-score)	-0.11	-0.25	-0.11
Number of regions	21	19	15
Number of corps members	231	136	155
Number of logs	1825	1070	1158

SOURCES: Closed-ended teacher logs administered by MDRC.

NOTES: A two-tailed test was applied to estimated effects. The statistical significance of estimated effects is indicated as follows: *** = 1 percent; ** = 5 percent; * = 10 percent.

[a]The estimated effects in this colum are the main findings presented in the report.

[b]Estimated effects in this column are based on teachers with complete (non-missing) data on the characteristics used as covariates in the analysis.

[c]Estimated effects in this column are based on a set of comparison regions that includes only the best match for each program region (1-to-1 nearest neighbor), rather than the two best matches (1-to-2 nearest neighbor).

Appendix G

Supplemental Findings

This appendix presents supplemental findings related to corps members' (CM) outcomes. The first section examines CMs' use of additional standards-aligned instructional strategies that were not discussed in the report. The second section presents a descriptive time series analysis of value-added scores for CMs in three regions in Florida.

Use of Standards-Aligned Instructional Strategies

As discussed in Chapter 2 and Appendix C, CMs' use of instructional strategies aligned with college and career readiness standards were measured using closed-ended logs. These logs were administered to CMs in the 2016 cohort teaching English Language Arts (ELA), mathematics, and general education during the first year of their teaching placement.

Chapter 4 examined the difference between CMs in the program and comparison regions with respect to three key confirmatory instructional practices that were most aligned with the redesigned training: CMs' use of culturally responsive texts (cultural responsiveness domain), the extent to which CMs asked students to cite text evidence (English Language Arts domain), and a composite measure of CMs' use of standard-aligned math practices (mathematics domain). As discussed in Chapter 4, although CMs in the program regions used these strategies, they did not use them more often than CMs in the comparison regions.

As an exploratory analysis, CMs' use of other standards-aligned instructional strategies was also examined:

- Appendix Table G.1 examines several ELA instructional strategies for CMs in the program and comparison regions. For items that were included in both the lower elementary school and upper elementary or secondary school logs, the findings are shown for all CMs; for items that were only included in the lower elementary school or upper elementary and secondary school logs, the findings are shown by level. The results indicate that CMs in the program regions had students listen to a text read aloud more frequently than the comparison regions CMs (58 percent of classes versus 42 percent of classes), and this difference is statistically significant. All other ELA practices were used with statistically similar frequency across the two groups of CMs.
- Appendix Table G.2 examines specific math instructional practices, for CMs in the program and comparison regions.[1] These findings show that CMs in the program regions had students solve math problems using an algorithm less frequently than CMs in the comparison regions (41 percent of classes versus

[1]These items were included in the confirmatory composite measure of standards-aligned math instruction examined in Chapter 4.

Appendix Table G.1

Estimated Effects on Supplemental Self-Reported Standards-Aligned Instructional Strategies, English Language Arts

Outcome	Program Regions	Comparison Regions	Estimated Difference	Effect Size	P-Value
All levels					
These areas were a focus of instruction (% of classes):					
Listening to a text read aloud	57.7	41.9	15.8 **	0.30	0.030
Reading with a partner or group	26.5	31.0	-4.5	-0.09	0.391
Reading independently	41.1	38.3	2.9	0.06	0.725
Number of regions	8	12			
Number of corps members	87	138			
Number of logs	739	1143			
Lower elementary school					
These areas were a focus of instruction (% of classes):					
Developing skills in phonics	57.3	84.7	-27.5	-0.50	0.328
Developing skills in fluency	32.4	40.0	-7.6	-0.14	0.858
Developing skills in vocabulary	34.3	52.0	-17.7	-0.33	0.725
Developing skills in writing	32.4	60.6	-28.3	-0.51	0.584
Extent to which students received instruction that[a]					
Addressed gaps in reading performance (1-4)	2.35	2.22	0.14	0.13	0.871
Addressed gaps in writing performance (1-4)	1.89	2.75	-0.86	-0.86	0.351
Number of regions	5	10			
Number of corps members	16	32			
Number of logs	117	216			

(continued)

Appendix Table G.1 (continued)

Outcome	Program Regions	Comparison Regions	Estimated Difference	Effect Size	P-Value
Upper elementary/secondary school					
Student was asked these types of questions (% of classes):					
Text-dependent questions related to text structure	62.8	52.5	10.3	0.20	0.230
Text-dependent questions related to character perspective	43.9	48.9	-5.0	-0.10	0.458
Text-dependent questions related to vocabulary	69.2	67.2	2.0	0.04	0.802
Text-dependent questions related to author's choices	41.3	44.5	-3.2	-0.06	0.678
Text-dependent questions related to theme	41.4	40.6	0.8	0.02	0.922
Text-dependent questions related to plot	44.3	49.6	-5.3	-0.10	0.482
Non-text-dependent questions related to the text	60.4	60.4	0.0	0.00	0.999
Extent to which students received instruction that is integrated with other content areas (1-4)[a]	2.43	2.49	-0.06	-0.06	0.727
Number of regions	8	12			
Number of corps members	71	106			
Number of logs	622	927			

SOURCE: Closed-ended teacher logs administered by MDRC.

NOTES: Online teacher logs were administered to corps members (CMs) in the 2016 cohort in the program and comparison regions. CMs teaching English Language Arts (ELA), mathematics, or general education in the first year of their teaching placement were eligible for the logs. Secondary school teachers, as well as elementary school teachers in departmentalized schools, were sent the log for their content area and level every two weeks (about twice per month). Elementary school teachers teaching both content areas (general education teachers) were sent the ELA log or the math log every two weeks, with the content area alternating between logs (about one log per content area per month). At the start of each log, teachers were given a randomly selected letter of the alphabet and asked to report on their instructional practices with a focal student whose name starts with that letter (or the closest letter) on a particular day that week.

The values in the "Program Regions" column are the observed mean outcome for teachers in the program regions. The values in the "Comparison Regions" column are the regression-adjusted mean outcome for teachers in the comparison regions, using the mean covariate values for the program region as the basis for the adjustment. The values in the "Estimated Difference" column are the difference between the program and comparison regions, adjusted for differences in the characteristics of the focal students and teachers. Values in the "Effect Size" column are the estimated effect divided by the standard deviation for the sample. A two-tailed t-test was applied to estimated differences. The statistical significance of estimated differences is indicated as follows: *** = 1 percent; ** = 5 percent; * = 10 percent. Rounding may cause slight discrepancies in calculating sums and differences.

Sample sizes in this table are the number of teachers in the log sample. A CM is included in the log sample if the CM completed at least one log during the school year.

[a]These items are based on a 4-point frequency scale (no time, some of the time, most of the time, all of the time).

Appendix Table G.2

Estimated Effects on Supplemental Self-Reported Standards-Aligned Instructional Strategies, Mathematics

Outcome	Program Regions	Comparison Regions	Estimated Difference	Effect Size	P-Value
All levels					
These areas were a focus of instruction (% of classes):					
Solving a math problems with an algorithm	41.1	53.3	-12.2 *	-0.23	0.074
Representing/analyzing relationships using illustrations	46.9	42.1	4.8	0.09	0.430
Critiquing the mathematical reasoning of others	13.6	12.7	0.9	0.02	0.834
Discussing math with a peer(s)	43.7	35.5	8.2	0.15	0.244
Justifying responses to a math problem(s)	35.6	35.9	-0.4	-0.01	0.956
Writing down mathematical thinking	39.1	38.7	0.4	0.01	0.959
Applying math concepts to a "real-world" problem(s)	35.0	33.5	1.5	0.03	0.799
Working in a math team(s) or pair(s)	40.5	37.1	3.4	0.06	0.612
Investigating a multi-step problem(s)	37.3	40.5	-3.2	-0.06	0.594
Practicing or relearning prerequisite concepts	35.1	29.4	5.7	0.11	0.381
Receiving remediation	11.2	15.5	-4.3	-0.11	0.365
Continuing to persevere on a challenging problem	32.9	31.9	1.0	0.02	0.889
Number of regions	8	13			
Number of corps members	80	151			
Number of logs	628	1197			

(continued)

Appendix Table G.2 (continued)

Outcome	Program Regions	Comparison Regions	Estimated Difference	Effect Size	P-Value
Lower elementary schools					
These areas were a focus of instruction (% of classes):					
Using manipulatives to solve math problems	24.7	35.2	-10.5	-0.20	0.529
Estimating solutions to a math problem(s)	10.8	6.7	4.1	0.13	0.624
Number of regions	6	12			
Number of corps members	44	71			
Number of logs	240	391			
Upper elementary/secondary schools					
These areas were a focus of instruction (% of classes):					
Using tool(s) to solve a math problem	25.8	33.6	-7.8	-0.15	0.374
Developing hypotheses	7.1	4.1	3.1	0.11	0.666
Testing hypotheses	6.4	4.9	1.5	0.06	0.839
Number of regions	8	13			
Number of corps members	37	82			
Number of logs	388	806			

(continued)

Appendix Table G.2 (continued)

SOURCE: Closed-ended teacher logs administered by MDRC.

NOTES: Online teacher logs were administered to corps members (CMs) in the 2016 cohort in the program and comparison regions. CMs teaching English Language Arts (ELA), mathematics, or general education in the first year of their teaching placement were eligible for the logs. Secondary school teachers, as well as elementary school teachers in departmentalized schools, were sent the log for their content area and level every two weeks (about twice per month). Elementary school teachers teaching both content areas (general education teachers) were sent the ELA log or the math log every two weeks, with the content area alternating between logs (about one log per content area per month). At the start of each log, teachers were given a randomly selected letter of the alphabet and asked to report on their instructional practices with a focal student whose name starts with that letter (or the closest letter) on a particular day that week.

The values in the "Program Regions" column are the observed mean outcome for teachers in the program regions. The values in the "Comparison Regions" column are the regression-adjusted mean outcome for teachers in the comparison regions, using the mean covariate values for the program region as the basis for the adjustment. The values in the "Estimated Difference" column are the difference between the program and comparison regions, adjusted for differences in the characteristics of the focal students and teachers. Values in the "Effect Size" column are the estimated effect divided by the standard deviation for the sample. A two-tailed t-test was applied to estimated differences. The statistical significance of estimated differences is indicated as follows: *** = 1 percent; ** = 5 percent; * = 10 percent. Rounding may cause slight discrepancies in calculating sums and differences.

Sample sizes in this table are the number of teachers in the log sample. A CM is included in the log sample if the CM completed at least one log during the school year.

.

53 percent of classes) and this difference is statistically significant. All other math practices were used with statistically similar frequency across the two groups of CMs.

However, it is important to note that 34 different instructional strategies are examined in this exploratory analysis, so the two statistically significant results could be spurious and due to chance (a "false positive").[2] In particular, when a 10 percent statistical significance level is used, one would expect to see a false positive for every 10 hypothesis tests conducted, or 3.4 false positives among the 34 outcomes examined in this analysis. This means that the two statistically significant results in the exploratory analysis could have been due to chance. Therefore, the findings from the exploratory analysis support the overall result that the redesigned training piloted in Tulsa does not appear to have increased CMs' use of standards-aligned instructional strategies.

Florida Value-Added Scores

Value-added scores are used by the state of Florida to measure teacher effectiveness. These data were used to look at trends over time in value-added scores for CMs in the three TFA regions located in Florida (Miami-Dade, Duval, and Orlando), all of which were trained at the national institute in Tulsa. The purpose of this analysis was to explore whether CMs in Florida in the 2016 cohort (who participated in the redesigned training) had higher value-added scores than prior cohorts of CMs in these regions. This section provides an overview of the data, analytic methods, and the sample of CMs used in the analysis, as well as a discussion of the findings.

Value-Added in Florida

The Florida Department of Education measures teacher effectiveness using evaluations of a teacher's instructional practice (based on classroom observations) and their value-added scores.[3] The latter scores are publicly available for all CMs in the state who teach ELA or mathematics in tested grades, who have 10 or more students in their class, and who are not protected by public disclosure laws.

In Florida, a teacher's value-added score is defined as the standardized difference between the *actual* spring state test score of his or her students and their *expected* score given their characteristics and prior achievement.[4] In recognition of the fact that the school can have an impact on student learning — through pathways outside of teachers' control — a teacher's value-

[2]This is called a Type I error, which is the risk of mistakenly concluding that the training had a statistically significant effect on a CM outcome, when in fact the training did not affect this outcome.

[3]For a general overview of teacher evaluations and value-added scores in Florida, see www.fldoe.org/teaching/performance-evaluation.

[4]The following information was used to predict a student's score: the number of subject-relevant courses in which the student was enrolled; up to two prior years of state test scores, students with disabilities status, English language learner status, gifted status, attendance, mobility (number of transitions), the difference from modal age in grade (as an indicator of retention), class size, and homogeneity of prior spring state test scores for students in the class. See Florida Department of Education (2011).

added score is the sum of two components: how the *teacher's* students performed relative to similar students at the *school* ("teacher component") and how the *school's* students performed relative to similar students in the *state* ("school component"). Thus, a teacher's value-added score measures his or her students' performance relative to similar students in the state. Value-added scores are scaled as a percentage. If a teacher's value-added score is +0.10, this means that the teacher's students scored 10 percent higher than expected relative to similar students in the state for that grade and subject. Negative value-added scores indicate that a teacher's students had lower state test scores on average than expected.

These data are publicly available for all Florida teachers who teach ELA or mathematics in tested grades. Therefore, for the purposes of this study, value-added scores were obtained for CMs in the 2012 to 2016 cohorts in Florida who were teaching mathematics or ELA (or general education in upper elementary school grades) during the first year of their teaching placement.

Interrupted Time Series Design

All three Florida regions were trained at the national institute in Tulsa in summer 2016 and received the redesigned training. This means that a comparative interrupted time series (CITS) design could not be used to estimate the effect of the training on value-added scores. Instead, the analysis was based on an interrupted time series (ITS) design, which is a CITS design without a comparison group. Using an ITS design, the effect of the redesigned training was estimated by looking at how much the value-added scores of the 2016 CMs in the Florida regions deviated from prior trends based on earlier cohorts in the same regions. A positive deviation from trend would suggest that the redesigned training may have had a positive effect on the value-added scores of CMs in the 2016 cohort in Florida.

The statistical model used to implement the ITS design is similar to the model used for a CITS design (described in Appendix E), but without a comparison group. The ITS analysis accounts for the clustering of CMs within regions and cohorts using a two-level model. The first level is a model of the variation in value-added scores between CMs within a region and cohort, and the second level is a model of the variation in value-added scores between regions and cohorts:

Level 1 (CMs)

$$\boldsymbol{Y_{irc} = \lambda_{rc} + \sum_{k=1}^{K} \psi_k Z_{k,ir} + \varepsilon_{irc}}$$

Level 2 (region/cohort)

$$\boldsymbol{\lambda_{rc} = \sum_m \delta_m REG_m + \sum_m \phi_m REG_m * RELYR_c + POST_c + u_{rc}}$$

where *i* denotes CMs, *r* is regions, and *c* is cohorts. The variables in the model are defined as follows:

Y_{irc} = Outcome for CM *i* in cohort *c* and region *r*

REG_m = A set of M dichotomous indicators for region (=1 for region *m* and 0 otherwise).

$RELYR_c$ = Continuous variable for time (cohort) centered at the last baseline (pre-intervention) cohort, such that RELYR=0 for the 2015 cohort

$POST_c$ = Dichotomous indicator for the first cohort to have an opportunity to receive the new training (=1 if 2016 cohort; 0 otherwise)

$Z_{k,ir}$ = A set of K characteristics for CM *i*, including their prior experience, their demographic characteristics, and the characteristics of their placement.[5]

ε_{irc} = Unexplained variation in Y within regions and cohorts

u_{rc} = Unexplained variation in Y between regions and cohorts

From the model, the following quantities of interest can be obtained:

δ_j = Intercept of the baseline trend for each of the three program regions in the analysis[6]

ϕ_j = Slope of the baseline trend for each the three program regions in the analysis[7]

β = Deviation from baseline trend for the 2016 cohort in the program regions, which is also the estimated effect of the redesigned training for these CMs

A key limitation of the ITS design, however, is that the deviation from trend for the program group could be confounded with other policy events or shifts that happened in 2016 in

[5]The model controlled for the following characteristics of CMs: their age; race/ethnicity; sex; their pre-TFA status (professional, graduate school, or undergraduate); whether they were the first in their family to attend college; whether they were a Pell grant recipient; their highest educational attainment; whether their degree was in an education-related field; and their overall application score when applying to TFA (a composite or scores on different types of skills). The model also controlled for the characteristics of CMs' placement: grade level and subject taught; whether the CM's school was a Title I school; whether their school was a charter school; its total enrollment; the school's level (elementary, middle, high); the school's location (urban, suburban, town, rural); its pupil-teacher ratio; the percent of certified teachers at the school; the percent of teachers at the school absent for 10 or more days; the percent of teachers at the school in their first or second year (novice); the percent of English as a second language students, the percent of students eligible for free or reduced price lunch; the racial or ethnic composition of students; the percent of girls; the percent of students with an individualized education program; the percent of students who were chronically absent; the percent of students retained; the percent of students with 1 or more in-school suspensions, and the percent of students with 1 or more out-of-school suspensions. The characteristics of a CM's school were measured using the most recent data available (the 2013-2014 or the 2015-2016 school year, depending on the variable).

[6]The intercept can also be thought of as the predicted outcome for the 2015 cohort of CMs in a region.

[7]In this model, the baseline trend was separately estimated for each region in the study. This strategy was used to account for the longitudinal nature of the data, whereby cohorts are nested within regions.

Florida. Therefore, the findings from the ITS design can less plausibly be attributed to the causal effect of the redesigned training.

Value-Added Sample

For the purposes of the analysis, a CM was included in the value-added sample if they had a value-added score for their first year of teaching and it was possible to identify them in the publicly available dataset from the state of Florida. The sample included CMs from the 2012 to 2016 cohorts in the three program regions in Florida (Orlando, Duval, and Miami-Dade).

Appendix Table G.3 shows the percentage of CMs in the Florida regions who are included in the value-added sample by content area (ELA and mathematics) and by cohort. For the purposes of measuring response rates, the target population for the value-added scores includes all CMs in the three Florida regions who taught ELA, mathematics, or general education in the upper elementary school grades; ELA and mathematics in middle school grades CMs; and ELA in high school grades.[8]

Appendix Table G.3

Percentage of Corps Members Included in the Value-Added Sample

	English Language Arts		Mathematics	
Inclusion Rate	Number in Sample	% in Sample	Number in Sample	% in Sample
2012 cohort	116	60.3	36	66.7
2013 cohort	114	57.9	44	63.6
2014 cohort	89	61.8	39	61.5
2015 cohort	73	50.7	36	52.8
2016 cohort	72	50.0	41	48.8
All cohorts	464	56.9	196	58.7

SOURCE: Florida Department of Education value-added measures.

NOTES: A cohort is defined as a group of TFA corps members (CMs) who joined TFA and attended the summer institute the same year. The "baseline cohorts" (2012-2015) received their summer training before the launch of the redesigned institute in Tulsa. The 2016 cohort is the first cohort of CMs trained after the launch of the redesigned institute. The number of CMs in the ELA column represents the total number of CMs in the three Florida program regions who are teaching ELA at the high school or middle school level, as well as CMs who are teaching ELA or general education at the upper elementary level. The number of CMs in the math column represents the total number of CMs in the three Florida program regions who are teaching math at the middle school level, as well as CMs who are teaching math or general education at the upper elementary level. A CM is included in the value added measure sample if they have a value-added score.

[8]Value-added scores are not used by Florida for CMs teaching high school mathematics.

Across all cohorts, 57 percent of eligible CMs teaching ELA were included in the value-added sample, and 59 percent of eligible CMs teaching mathematics were included in the sample. A CM may have been excluded from the sample because of missing data for several reasons. First, a CM may have a value-added score but be listed under a different name or be teaching in a different school relative to the information available to the study team, thereby making it impossible to identify the individual in the Florida database. Second, a CM could have been protected by a public disclosure law, in which case the CM's score was reported in the Florida database but the person's name was suppressed. Finally, a CM could have had fewer than 10 students in his or her class, in which case a CM's name would also have been suppressed.

When an ITS design is used to estimate effects, ideally the characteristics of sample members should be similar across cohorts. Accordingly, the characteristics of CMs in the value-added sample were examined, focusing on CMs in the 2015 cohort (baseline cohort) and the 2016 cohort (follow-up cohort). As explained in Chapter 2, a useful rule of thumb is that differences should not exceed an effect size of 0.25 standard deviations. An effect size is the difference in values for an outcome measure or characteristic expressed as a proportion of the standard deviation for that outcome measure or characteristic.[9] Overall, there were several notable differences between cohorts:

- Appendix Tables G.4 and G.5 show the characteristics of CMs in the 2015 and 2016 cohorts with value-added scores, by content area. For the ELA sample, CMs in the 2016 cohort sample were more likely than the 2015 cohort to have been professionals before TFA (as opposed to undergraduate students). For the math sample, CMs in the 2016 cohort had higher application scores on average than the 2015 cohort and were more likely to be a person of color. Some of these baseline differences are larger than 0.25 as an effect size.

- Appendix Tables G.6 and G.7 present the characteristics of the teaching placements of CMs in the value-added sample, for the 2015 and 2016 cohort. The two cohorts of CMs differed with respect to some of the characteristics of their schools, and these differences exceed 0.25 in magnitude. For example, in the math value-added sample, CMs in the 2016 cohort were more likely than the 2015 cohort to be teaching in lower grade levels.

Additionally, as a sensitivity analysis, an ITS design was used to look at the "effect" of the redesigned training on the characteristics of CMs and their school placements. This effect was greater than 0.25 as an effect size for several characteristics. This means that the estimated effect of the redesigned training on value-added scores could be confounded with changes in the characteristics or school context of the 2016 cohort relative to prior cohorts.

[9]Effects sizes in this section are based on the standard deviation for the pooled sample (program and comparison regions).

Appendix Table G.4

Characteristics of the 2016 Cohort (Program) and the 2015 Cohort (Baseline), Value-Added Sample (English Language Arts)

Characteristic	2016 Cohort	2015 Cohort	Estimated Difference	Effect Size	P-Value
Preparation and skills					
Composite application score (1-5)	3.66	3.60	0.05	0.21	0.367
Educational attainment (%)					
Bachelor	91.7	91.9	-0.2	-0.01	0.973
Masters	8.3	8.1	0.2	0.01	0.973
Major or minor in Education (%)	22.1	16.2	5.9	0.15	0.526
Prospect Type (%)					
Undergraduate	52.8	67.6	-14.8	-0.30	0.202
Graduate	5.6	5.4	0.2	0.01	0.978
Professional	41.7	27.0	14.6	0.31	0.193
Demographic					
Age at entry into TFA	24.25	23.92	0.33	0.10	0.667
First in family to attend college (%)	46.0	47.8	-1.8	-0.04	0.875
Received a Pell grant (%)	52.8	54.1	-1.3	-0.03	0.914
Race/ethnicity (%)					
Hispanic	3.8	24.3	-20.5 **	-0.55	0.011
Black	32.0	25.1	6.8	0.14	0.530
White	55.7	48.2	7.6	0.15	0.510
Asian	2.9	0.0	2.9	0.24	0.306
Other	8.6	5.4	3.2	0.12	0.603
Person of color (%)	42.4	50.4	-8.0	-0.16	0.482
Female (%)	83.3	94.6	-11.3	-0.36	0.127
Number of regions	3	3			
Number of corps members	36	37			

(continued)

Appendix Table G.4 (continued)

SOURCE: Teach For America administrative records for corps members.

NOTES: A cohort is defined as a group of TFA corps members (CMs) who joined TFA and attended the summer institute the same year. The 2016 cohort of CMs joined TFA when the redesigned summer institute was launched in summer 2016. The 2015 cohort joined TFA before the launch of the new summer training. The number of CMs represents the total number of CMs in each cohort. CMs' characteristics are measured during the application process to TFA. The composite application score is based on a 5-point scale measuring CMs' skills in several areas, using information from their application materials and interviews.

The values in the columns "2016 Cohort" and "2015 Cohort" are observed means. Values in the "Effect Size" column are the estimated difference divided by the standard deviation for the pooled sample (2015 and 2016 cohorts). Rounding may cause slight discrepancies in calculating sums and differences.

A two-tailed t-test was applied to estimated deviations and estimated differences between program regions and comparison regions. The statistical significance of estimated differences is indicated as follows: *** = 1 percent; ** = 5 percent; * = 10 percent.

Overall, given the limitations of the study design, the substantial amount of missing data, and differences in CMs' characteristics across cohorts, the value-added findings in the next section should not be interpreted as the causal effect of the redesigned training.

Findings

Appendix Figure G.1 plots the trend over time in the ELA value-added scores of CMs in Florida during their first year of teaching. Before the redesigned training was launched (2012 to 2015 cohorts), ELA value-added scores had been fluctuating substantially across cohorts, though generally negative on average and declining over time. The value-added scores for CMs in the 2016 cohort who participated in the new summer training were higher than predicted relative to the scores of prior cohorts of CMs in the same regions, but not by a statistically significant amount.

Similarly, the redesigned training does not appear to have improved math value-added scores for CMs in Florida. Appendix Figure G.2 plots the trends over time in these scores. Before the new training was launched, value-added scores in math were negative on average and they had been steadily declining (2012 to 2015 cohorts). Math value-added scores continued to decline, as predicted, for the 2016 cohort. The deviation from trend for these CMs was small in magnitude and not statistically significant.

Overall, these findings suggest that the redesigned training piloted at the national institute in Tulsa did not improve the value-added scores of the first cohort of CMs in Florida who received it. However, as noted earlier, these findings should be interpreted with caution due to limitations with the study design, the amount of missing data, and the fact that the analysis only included CMs from three regions.

Appendix Table G.5

Characteristics of the 2016 Cohort (Program) and the 2015 Cohort (Baseline), Value-Added Sample (Mathematics)

Characteristic	2016 Cohort	2015 Cohort	Estimated Difference	Effect Size	P-Value
Preparation and skills					
Composite application score (1-5)	3.8	3.6	0.2 ***	0.92	0.003
Educational attainment (%)					
Bachelor	95.0	94.7	0.3	0.01	0.971
Masters	5.0	5.3	-0.3	-0.01	0.971
Major or minor in Education (%)	13.7	10.1	3.5	0.10	0.748
Prospect Type (%)					
Undergraduate	65.0	68.4	-3.4	-0.07	0.827
Graduate	-0.1	5.3	-5.4	-0.34	0.299
Professional	35.0	26.3	8.7	0.19	0.569
Demographic					
Age at entry into TFA	24.3	23.9	0.4	0.11	0.718
First in family to attend college (%)	40.0	36.8	3.2	0.06	0.845
Received a Pell grant (%)	60.0	36.8	23.2	0.46	0.157
Race/ethnicity (%)					
Hispanic	21.9	18.6	3.2	0.08	0.804
Black	40.0	31.6	8.4	0.17	0.595
White	16.3	47.5	-31.2 **	-0.67	0.022
Asian	6.3	2.1	4.3	0.27	0.403
Other	15.2	3.7	11.4	0.37	0.251
Person of color (%)	83.7	52.5	31.2 **	0.67	0.022
Female (%)	86.2	89.4	-3.2	-0.09	0.778
Number of regions	3	3			
Number of corps members	20	19			

(continued)

Appendix Table G.5 (continued)

SOURCE: Teach For America administrative records for corps members.

NOTES: A cohort is defined as a group of TFA corps members (CMs) who joined TFA and attended the summer institute the same year. The 2016 cohort of CMs joined TFA when the redesigned summer institute was launched in summer 2016. The 2015 cohort joined TFA before the launch of the new summer training. The number of CMs represents the total number of CMs in each cohort. CMs' characteristics are measured during the application process to TFA. The composite application score is based on a 5-point scale measuring CMs' skills in several areas, using information from their application materials and interviews.

The values in the columns "2016 Cohort" and "2015 Cohort" are observed means. Values in the "Effect Size" column are the estimated difference divided by the standard deviation for the pooled sample (2015 and 2016 cohorts). Rounding may cause slight discrepancies in calculating sums and differences.

A two-tailed t-test was applied to estimated deviations and estimated differences between program regions and comparison regions. The statistical significance of estimated differences is indicated as follows: *** = 1 percent; ** = 5 percent; * = 10 percent.

.

Appendix Table G.6

Characteristics of the First-Year Teaching Placements of the 2016 Cohort (Program) and 2015 Cohort (Baseline), Value-Added Sample (English Language Arts)

Placement Characteristic	2016 Cohort	2015 Cohort	Estimated Difference	Effect Size	P-Value
Grade and subject of placement					
Grade level					
Upper elementary school	22.2	13.5	8.7	0.23	0.338
Middle school	43.3	48.0	-4.7	-0.09	0.689
High school	32.5	36.1	-3.6	-0.07	0.752
Subject taught					
General education	7.3	4.5	2.8	0.14	0.546
English Language Arts	92.7	95.5	-2.8	-0.14	0.546
Characteristics of Placement Schools					
School size and type					
Total enrollment	1,033.5	1,056.1	-22.5	-0.04	0.871
Title I school (%)	100.0	100.0	0.0	0.00	N/A
Magnet school (%)	41.4	49.0	-7.7	-0.15	0.495
Staffing					
First-year teachers at the school (%)	40.7	36.7	4.1	0.27	0.191
Teacher absences (% with 10 or more per year)	55.1	55.5	-0.4	-0.03	0.894
Certified (%)	97.1	96.4	0.7	0.28	0.230
School location (%)					
Urban	49.3	48.5	0.9	0.02	0.913
Suburban	48.5	52.2	-3.7	-0.08	0.621

(continued)

Appendix Table G.6 (continued)

Placement Characteristic	2016 Cohort	2015 Cohort	Estimated Difference	Effect Size	P-Value
Characteristics of students in placement schools					
Chronically absent students (%)	7.1	6.8	0.3	0.02	0.909
Students retained (%)	3.2	3.7	-0.5	-0.22	0.363
Students with 1 or more in-school suspensions (%)	12.3	13.8	-1.4	-0.13	0.536
Students with 1 or more out-of-school suspensions (%)	11.8	10.5	1.3	0.18	0.449
Free or reduced price lunch students (%)	87.2	86.2	1.0	0.07	0.647
Students with English as a second language (%)	10.0	12.4	-2.4	-0.28	0.119
Students with an Individualized Education Plan (%)	13.7	13.8	-0.1	-0.02	0.922
Racial/ethnic composition of students (%)					
Hispanic	25.0	29.7	-4.6	-0.23	0.202
Black	63.4	59.9	3.5	0.16	0.488
White	9.6	8.0	1.5	0.12	0.526
Asian	1.2	1.5	-0.3	-0.24	0.249
Other	2.1	2.3	-0.2	-0.10	0.447
Female students (%)	48.5	48.6	-0.1	-0.04	0.866
Number of regions	3	3			
Number of corps members	36	37			

NOTES: A cohort is defined as a group of TFA corps members (CMs) who joined TFA and attended the summer institute the same year. The 2016 cohort of CMs joined TFA when the redesigned summer institute was launched in summer 2016. The 2015 cohort joined TFA before the launch of the new summer training. The number of CMs represents the total number of CMs in each cohort. CMs' characteristics are measured during the application process to TFA. The values in the columns "2016 Cohort" and "2015 Cohort" are observed means. Values in the "Effect Size" column are the estimated difference divided by the standard deviation for the pooled sample (2015 and 2016 cohorts). Rounding may cause slight discrepancies in calculating sums and differences.

A two-tailed t-test was applied to estimated deviations and estimated differences between program regions and comparison regions. The statistical significance of estimated differences is indicated as follows: *** = 1 percent; ** = 5 percent; * = 10 percent.

Appendix Table G.7

Characteristics of the First-Year Teaching Placements of the 2016 Cohort (Program) and 2015 Cohort (Baseline), Value-Added Sample (Mathematics)

Placement Characteristic	2016 Cohort	2015 Cohort	Estimated Difference	Effect Size	P-Value
Grade and subject of placement					
Grade level					
Upper elementary school	27.5	14.3	13.3	0.34	0.285
Middle school	72.5	85.7	-13.3	-0.34	0.285
Subject taught					
General education	12.3	10.4	1.9	0.07	0.818
Mathematics	87.7	89.6	-1.9	-0.07	0.818
Characteristics of placement schools					
School size and type					
Total enrollment	637.7	728.8	-91.1	-0.34	0.301
Title I school (%)	100.0	100.0	0.0	0.00	N/A
Magnet school (%)	35.0	36.8	-1.8	-0.04	0.908
Charter school (%)	6.3	2.1	4.3	0.27	0.403
Staffing					
First-year teachers at the school (%)	45.7	40.2	5.5	0.30	0.344
Teacher absences (% with 10 or more per year)	47.8	48.4	-0.6	-0.04	0.902
Certified (%)	95.4	97.8	-2.4 **	-0.66	0.021
School location (%)					
Urban	61.8	59.8	2.0	0.04	0.872
Suburban	38.2	40.2	-2.0	-0.04	0.872

(continued)

Appendix Table G.7 (continued)

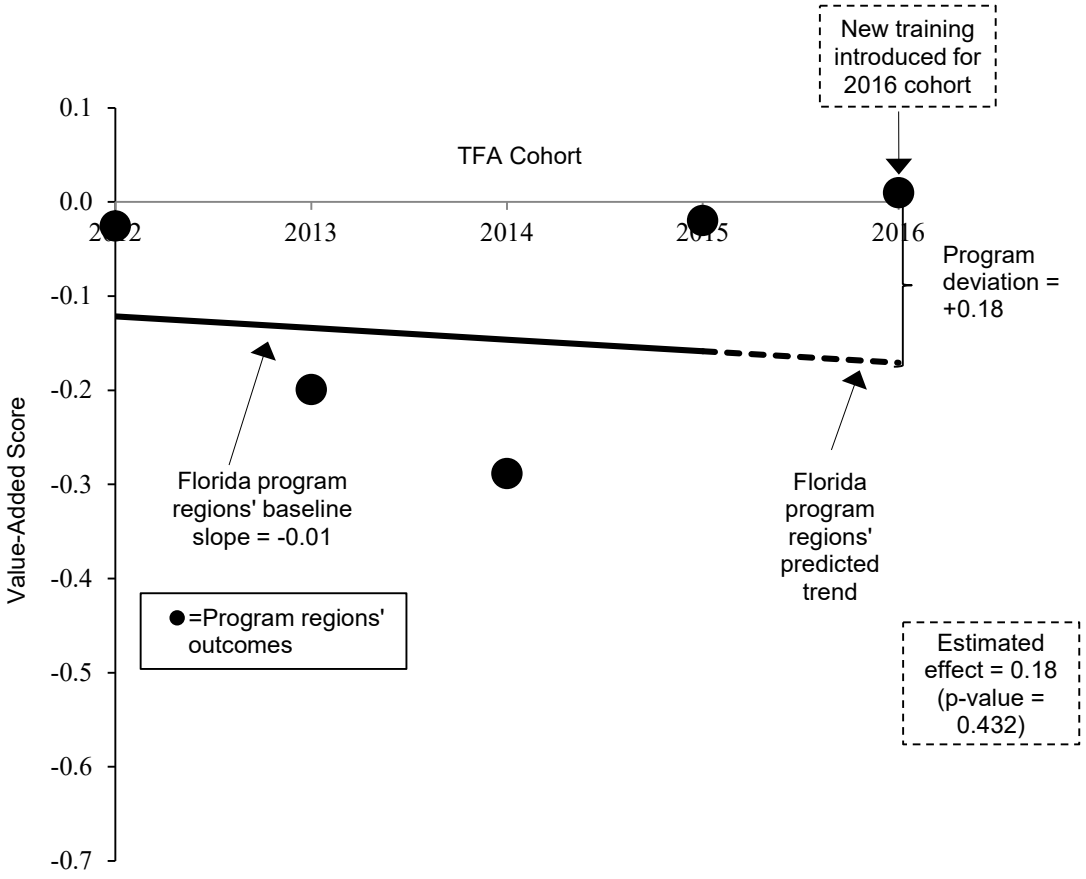
Placement Characteristic	2016 Cohort	2015 Cohort	Estimated Difference	Effect Size	P-Value
Characteristics of students in placement schools (%)					
Chronically absent students (%)	0.8	1.1	-0.3	-0.34	0.207
Students retained (%)	3.3	3.2	0.2	0.08	0.822
Students with 1 or more in-school suspensions (%)	15.7	15.9	-0.2	-0.02	0.949
Students with 1 or more out-of-school suspensions (%)	16.0	16.1	-0.1	-0.01	0.979
Free or reduced price lunch students (%)	81.7	84.2	-2.5	-0.14	0.640
Students with English as a second language (%)	9.6	12.9	-3.4	-0.36	0.202
Students with an Individualized Education Plan (%)	12.2	12.8	-0.6	-0.12	0.720
Racial/ethnic composition of students (%)					
Hispanic	19.3	27.1	-7.8	-0.42	0.128
Black	71.3	62.5	8.7	0.42	0.192
White	7.2	8.9	-1.7	-0.18	0.415
Asian	1.2	1.5	-0.3	-0.16	0.604
Other	1.9	2.2	-0.3	-0.15	0.347
Female students (%)	48.5	47.7	0.7	0.22	0.455
Number of regions	3	3			
Number of corps members	20	19			

NOTES: A cohort is defined as a group of TFA Corps Members (CMs) who joined TFA and attended the summer institute the same year. The 2016 cohort of CMs joined TFA when the redesigned summer institute was launched in summer 2016. The 2015 cohort joined TFA before the launch of the new summer training. The number of CMs represents the total number of CMs in each cohort. CMs' characteristics are measured during the application process to TFA. The values in the columns "2016 Cohort" and "2015 Cohort" are observed means. Values in the "Effect Size" column are the estimated difference divided by the standard deviation for the pooled sample (2015 and 2016 cohorts). Rounding may cause slight discrepancies in calculating sums and differences.

A two-tailed t-test was applied to estimated deviations and estimated differences between program regions and comparison regions. The statistical significance of estimated differences is indicated as follows: *** = 1 percent; ** = 5 percent; * = 10 percent.

.

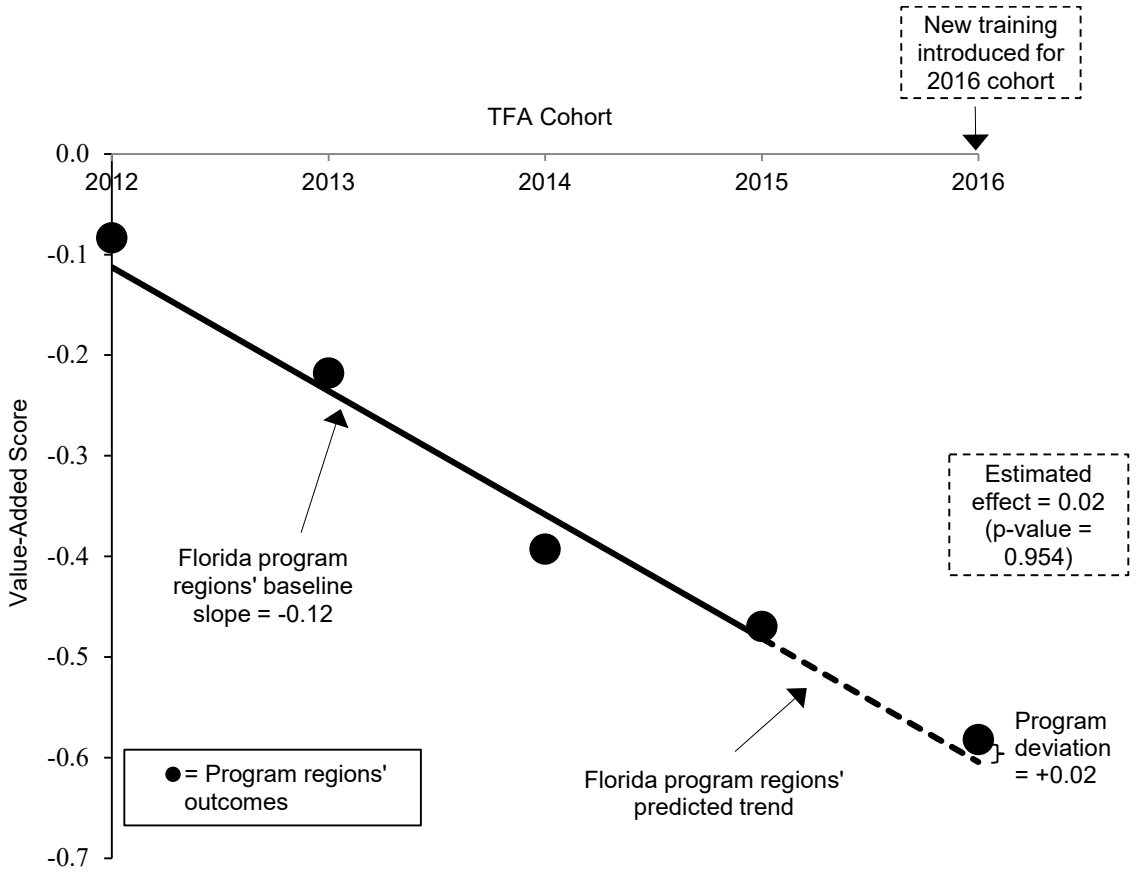
Appendix Figure G.1
Trends in English Laguage Arts Value-Added Scores, Florida Program Regions



SOURCE: Florida Department of Education value-added measures.

NOTES: The data in this table reflect value-added scores for 263 corps members across 3 regions.

Appendix Figure G.2
Trends in Mathematics Value-Added Scores, Florida Program Regions



SOURCE: Florida Department of Education value-added measures.

NOTES: The data in this table reflect value-added scores for 115 corps members across 3 regions.

References

Adamson, Frank, and Linda Darling-Hammond. 2011. *Addressing the Inequitable Distribution of Teachers: What It Will Take to Get Qualified, Effective Teachers in All Communities.* Stanford, CA: Stanford Center for Opportunity Policy in Education.

Ballart, Xavier, and Clara Riba. 1995. "Impact of Legislation Requiring Moped and Motorbike Riders to Wear Helmets." *Evaluation and Program Planning* 18, 4: 311-320.

Bloom, Howard S. 2003. "Using 'Short' Interrupted Time-Series Analysis To Measure The Impacts Of Whole-School Reforms: With Applications to a Study of Accelerated Schools." *Evaluation Review* 27, 3: 3-49.

Bloom, Howard S., and James A. Riccio. 2005. "Using Place-Based Random Assignment and Comparative Interrupted Time-Series Analysis to Evaluate the Jobs-Plus Employment Program for Public Housing Residents." *Annals of the American Academy of Political and Social Science* 599, 1: 19-51.

Campbell, Donald T., and H. Laurence Ross. 1968. "The Connecticut Crackdown on Speeding: Time-Series Data in Quasi-Experimental Analysis." *Law and Society Review* 3, 1: 33-53.

Clark, Melissa, Eric Isenberg, Albert Liu, Libby Makowsky, and Marykate Zukiewicz. 2017. *Impacts of the Teach For America Investing in Innovation Scale-Up.* Princeton, NJ: Mathematica Policy Research.

Clark, Melissa, Hanley S. Chiang, Tim Silva, Sheena McConnell, Kathy Sonnenfeld, Anastasia Erbe, and Michael Puma. 2013. *The Effectiveness of Secondary Math Teachers from Teach For America and the Teaching Fellows Programs*. Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Cook, Thomas D., William R. Shadish, and Vivian C. Wong. 2008. "Three Conditions under which Experiments and Observational Studies Produce Comparable Causal Estimates: New Findings From Within-Study Comparisons." *Journal of Policy Analysis and Management* 27, 4: 724-750.

Core Practice Consortium. n.d. "Core Practice." Accessed October 16, 2018. Website: www.corepracticeconsortium.com.

Corrin, William, and Thomas D. Cook. 1998. "Design Elements of Quasi-Experiments." Pages 35-58 in Arthur J. Reynolds and Herbert J. Welber (eds.), *Advances in Educational Productivity*. Bingley, UK: Emerald Group Publishing Limited.

Decker, Paul T., Daniel P. Mayer, and Steven Glazerman. 2004. *The Effects of Teach For America on Students: Findings from a National Evaluation*. Princeton, NJ: Mathematica Policy Research.

Dee, Thomas S., and Brian Jacob. 2011. "The Impact of No Child Left Behind on Student Achievement." *Journal of Policy Analysis and management* 30, 3: 418-446.

Ericsson, K. Anders, Ralf Th. Krampe, and Clemens Tesch-Römer. 1993. "The Role of Deliberate Practice in the Acquisition of Expert Performance." *Psychological Review* 100, 3: 363-406.

Florida Department of Education. (2011). "Value Added Model Recommendation to the Commissioner of Education from the Student Growth Implementation Committee." Website: www.fldoe.org/core/fileparse.php/7566/urlt/0075074-summaryfinalrecommendation.pdf.

Fretheim, Atle, Stephen B. Soumerai, Fang Zhang, Andrew D. Oxman, and Dennis Ross-Degnan. 2013. "Interrupted Time-Series Analysis Yielded an Effect Estimate Concordant with the Cluster-Randomized Controlled Trial Result." *Journal of Clinical Epidemiology* 66, 8: 883-887.

Garet, Michael S., Stephanie Cronen, Marian Eaton, Anja Kurki, Meredith Ludwig, Wehmah Jones, Kazuaki Uekawa, Audrey Falk, Howard Bloom, Fred Doolittle, Pei Zhu, and Laura Sztejnberg. *The Impact of Two Professional Development Interventions on Early Reading Instruction and Achievement*. Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Grossman, Pamela, Christa Compton, Danielle Igra, Matthew Ronfeldt, Emily Shahan, and Peter Williamson. 2009. "Teaching Practice: A Cross-Professional Perspective." *Teachers College Record* 111, 9: 2055-2100.

Haycock, Kati, and Heather G. Peske. 2006. "Teaching Inequality: How Poor and Minority Students Are Shortchanged on Teacher Quality: A Report and Recommendations by the Education Trust." Washington, DC: Education Trust.

Ho, Daniel, Kosuke Imai, Gary King, and Elizabeth A. Stuart. 2007. "Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference." *Political Analysis* 15, 3: 199-236.

Kemple, James J., Corinne M. Herlihy, and Thomas J. Smith. 2005. Making Progress Toward Graduation: Evidence from the Talent Development High School Model. New York: MDRC.

Ladson-Billings, Gloria. 1994. *The Dreamkeepers: Successful Teachers of African American Children.* San Francisco: Jossey-Bass.

Ladson-Billings, Gloria. 1995. "Toward a Theory of Culturally Relevant Pedagogy." *American Educational Research Journal* 32, 3: 465-491.

Mulford, Harold A, Johannes Ledolter, and JL Fitzgerald. 1992. "Alcohol Availability and Consumption: Iowa Sales Data Revisited." *Journal of Studies on Alcohol and Drugs* 53, 5: 487-494.

Rosenbaum, Paul R., and Donald R. Rubin. 1983. "The Central Role of the Propensity Score in Observational Studies for Causal Effects." *Biometrika* 70, 1: 41-55.

Rowan, Brian, Eric Camburn, and Richard Correnti. 2004. "Using Teacher Logs to Measure the Enacted Curriculum in Large-Scale Surveys: A Study of Literacy Teaching in 3rd Grade Classrooms." *Elementary School Journal* 105, 1: 75-101.

Rowan, Brian, Richard Correnti, Robert J. Miller, and Eric M. Camburn. 2009. *School Improvement by Design: Lessons from a Study of Comprehensive School Reform Programs.* Philadelphia: Consortium for Policy Research in Education, University of Pennsylvania.

Shadish, William R., Thomas D. Cook, and Donald T. Campbell. 2002. *Experimental and Quasi Experimental Designs for Generalized Causal Inference*. Boston: Houghton Mifflin.

Somers, Marie-Andree, Pei Zhu, Robin Tepper Jacob, and Howard S. Bloom. 2013. "The Validity and Precision of the Comparative Interrupted Time Series Design and the Difference-in-Difference Design in Educational Evaluation." Working paper. New York: MDRC.

St. Clair, Travis, Thomas D. Cook, and Kelly Hallberg. 2014. "Examining the Internal Validity and Statistical Precision of the Comparative Interrupted Time Series Design by Comparison With a Randomized Experiment." *American Journal of Evaluation* 35, 3: 311-327.

Sutcher, Leib, Linda Darling-Hammond, and Desiree Carver-Thomas. 2016. *A Coming Crisis in Teaching? Teacher Supply, Demand, and Shortages in the U.S.* Palo Alto, CA: Learning Policy Institute.

Teacher Education by Design. n.d.(a). "Choral Counting." Accessed October 16, 2018. Website: tedd.org.

Teacher Education by Design. n.d.(b). "Content Areas." Accessed October 16, 2018. Website: tedd.org.

Teach For America. n.d. "Our Values," Accessed October 16, 2018. Website: www.teachforamerica.org.

What Works Clearinghouse. 2016. *WWC Evidence Review Protocol for Teacher Training, Evaluation, and Compensation Version 3.2*. Washington, DC: What Works Celaringhouse, Institute of Education Sciences, U.S. Department of Education.

What Works Clearinghouse. 2017. *What Works Clearinghouse Standards Handbook, Version 4.0.* Washington, DC: What Works Celaringhouse, Institute of Education Sciences, U.S. Department of Education.

Wilgoren, Jodi. 2000. "Wendy Kopp, Leader of Teach For America." *New York Times* (November 12), p. A-4.

Wong, Manyee, Thomas D. Cook, and Peter M. Steiner. 2009. "No Child Left Behind: An Interim Evaluation of Its Effects on Learning Using Two Interrupted Time Series Each with Its Own Non-Equivalent Comparison Series." Working Paper IPR-WP-09-11. Evanston, IL: Institute for Policy Research, Northwestern University.

About MDRC

MDRC is a nonprofit, nonpartisan social and education policy research organization dedicated to learning what works to improve the well-being of low-income people. Through its research and the active communication of its findings, MDRC seeks to enhance the effectiveness of social and education policies and programs.

Founded in 1974 and located in New York; Oakland, California; Washington, DC; and Los Angeles, MDRC is best known for mounting rigorous, large-scale, real-world tests of new and existing policies and programs. Its projects are a mix of demonstrations (field tests of promising new program approaches) and evaluations of ongoing government and community initiatives. MDRC's staff members bring an unusual combination of research and organizational experience to their work, providing expertise on the latest in qualitative and quantitative methods and on program design, development, implementation, and management. MDRC seeks to learn not just whether a program is effective but also how and why the program's effects occur. In addition, it tries to place each project's findings in the broader context of related research — in order to build knowledge about what works across the social and education policy fields. MDRC's findings, lessons, and best practices are shared with a broad audience in the policy and practitioner community as well as with the general public and the media.

Over the years, MDRC has brought its unique approach to an ever-growing range of policy areas and target populations. Once known primarily for evaluations of state welfare-to-work programs, today MDRC is also studying public school reforms, employment programs for ex-prisoners, and programs to help low-income students succeed in college. MDRC's projects are organized into five areas:

- Promoting Family Well-Being and Children's Development
- Improving Public Education
- Raising Academic Achievement and Persistence in College
- Supporting Low-Wage Workers and Communities
- Overcoming Barriers to Employment

Working in almost every state, all of the nation's largest cities, and Canada and the United Kingdom, MDRC conducts its projects in partnership with national, state, and local governments, public school systems, community organizations, and numerous private philanthropies.