



THE Teach for America EVALUATION

*Herewith, the first evidence on its recruits' actual performance
in the classroom* by MARGARET RAYMOND AND STEPHEN FLETCHER



SINCE 1990 THE NEW YORK-BASED TEACH FOR America (TFA) program has placed more than 7,000 teachers in some of the nation's most challenging school districts. The nonprofit organization recruits high-achieving seniors from top colleges and asks them to commit themselves to two years of teaching in inner-city or rural schools. TFA currently supplies teachers to 18 districts across the country, including Newark, New Jersey; rural districts in the Mississippi Delta; Phoenix, Arizona; and Houston, Texas. Most TFA recruits serve in schools that qualify for funding under Title I of the Elementary and Secondary Education Act due to their high concentrations of students living in poverty. These schools often find it difficult to fill teaching positions with strong candidates.

In number of teachers trained each year, TFA is one of the nation's largest suppliers of teachers. According to the American Association of Colleges for Teacher Education, which represents university- or college-based teacher-preparation programs, only 10 percent of programs offering undergraduate-level training produce more teachers each year. TFA recruits clearly help to alleviate teacher shortages in the school districts in which TFA operates. But how do these recruits perform once they are hired? The evaluation results reported here provide the first evidence of TFA recruits' actual performance in the classroom.

TFA's underlying idea is that many smart, accomplished college graduates would make significant contributions to public education—if it weren't for the fact that they must have majored in education or be willing to spend one to two more years (and significantly more in tuition) earning a master's in education in order to get certified. Even if these bright, young, committed individuals don't remain in teaching forever, TFA posits that they can make a difference in some of the nation's most challenging

schools. TFA gives what the program considers potentially excellent teachers—those with both solid knowledge of content and strong leadership skills—a low-cost opportunity to earn a teaching credential while discovering whether they have found a calling. The emphasis on knowledge and leadership is reflected in TFA's recruiting: a typical TFA corps member earned a grade-point average of 3.4 out of 4.0, and 87 percent of recruits have leadership experience.

TFA recruits undergo five weeks of training during the summer. They complete intensive preservice coursework and spend the balance of the summer in classrooms as student teachers or team teachers. Once school starts in the fall, TFA recruits must participate in weekly TFA-sponsored professional development workshops—more than the typical new teacher. This is in addition to any coursework they must complete under their district or state's alternative certification program. In Houston, the TFA site examined for this study, all uncertified teachers must enroll in the district's alternative certification program at the beginning of their first year. It takes a year to earn a credential.

Advocates of alternative certification hail TFA for its creative response to teacher shortages and for its efforts to recruit star college graduates who otherwise might not have found their way to teaching. Many corps members leave teaching after fulfilling their two-year commitment, but they often remain in the education field—as school administrators, in government, or with nonprofits that are involved in education reform. In April 2001 *Education Week* reported, "It's becoming clear that Teach For America is channeling a wealth of talent, energy, and creativity into educational leadership that might otherwise have wound up in such fields as medicine, law, or business."

But TFA has its detractors as well. They point to the lack of

pedagogical training and knowledge of child development theories among TFA teachers. They also tend to believe that TFA demeans teaching by treating it as a Peace Corps-style rescue mission rather than a true profession, with salaries appropriate to attracting solid candidates. "A frankly missionary program," wrote Stanford professor Linda Darling-Hammond in an oft-cited 1994 *Phi Delta Kappan* article, "TFA has recruiters and advocates who have focused much of their attention on the advantaged college graduates for whom TFA serves as something useful to do on their way to their 'real jobs' in law, medicine, or business."

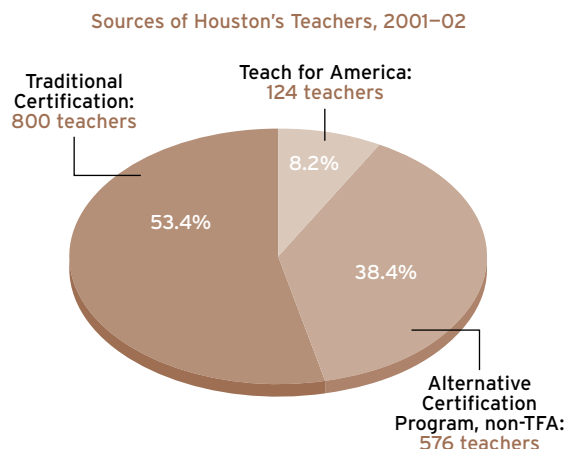
Into the Breach

So far, these battle lines have been drawn with only rhetoric as a weapon, for evidence on the performance of TFA teachers in the classroom has been nonexistent. To help fill this vacuum, CREDO (formerly known as the Center for Research on Education Outcomes), an independent nonpartisan research group at the Hoover Institution of Stanford University, evaluated the performance of TFA teachers relative to both new teachers and all teachers in the Houston Independent School District. Houston is the 7th-largest school district in the United States, encompassing 186 elementary schools and 34 middle schools. Houston hires approximately 350 new elementary- and middle-school teachers a year; since TFA began supplying teachers to the district, its proportion of new teachers has ranged between 5 and 10 percent (see Figure 1). Latino (54 percent) and African-American (33 percent) students make up the majority of the district's student population. The majority of students live in relatively poor households, with 75 percent receiving free or reduced-cost lunches.

To perform the analysis, the year-end learning gains of students with TFA teachers were measured against the gains

Another Source of New Teachers (Figure 1)

In the 2001–02 school year, Teach for America supplied 8 percent of the Houston school district's new hires.



SOURCES: Teach for America; Houston Independent School District

of students with non-TFA teachers. Individual scores on the annual Texas Assessment of Academic Skills (TAAS) in mathematics and reading and English language arts were used as the measure of student performance. Texas has annually tested students in grades 3 through 8 since 1993. Having multiple years of test scores for the same student allowed us to estimate the annual progress of individ-

Many corps members leave teaching after fulfilling their two-year commitment, but they often remain in the education field.

ual students by controlling for their previous year's TAAS score.

Linked student- and teacher-level data were obtained through the Texas Schools Project at the University of Texas at Dallas. Two groups of students were studied: elementary-school students in grades 3

through 5 (where most students have a single teacher throughout the year) and middle schoolers in grades 6 through 8 (where students tend to have a different teacher in each subject). This article presents the findings for the elementary grades only. The analysis of middle-school teachers produced similar results, but faced methodological challenges arising from students' having had multiple teachers. (For a description of the middle-school analysis, see the full report at: http://credo.stanford.edu/working_papers.htm.)

Our study was framed around two separate but related questions. First, we wanted to know whether TFA recruits were effective teachers. So we compared the performance of all TFA teachers with that of all non-TFA teachers in Houston, regardless of how experienced they were. The relevant comparison group was all Houston teachers in grades 3 through 5 from 1996 to 2000.

The second analysis was focused on the question that districts must ask when considering whether to work with TFA: How do TFA recruits measure up against other potential new hires? In other words, if a district had one remaining position to fill, would it be better off with a new TFA teacher or another candidate? For this analysis, we compared TFA teachers with one or two years of experience with other teachers in the district with the same amount of experience.

It is important to note that the group of new teachers with which TFA teachers were compared included teachers with other forms of alternative certification along with a number who were not fully certified. Such teachers need to be included because the purpose of this study is to compare TFA teachers with the pool of potential hires available to the district. As noted above, in both the comparisons of TFA teachers with all teachers and of TFA teachers with only new teachers, TFA teachers represented

a small proportion of the total number of teachers. Achieving statistically significant results requires very large impacts under such circumstances, and so our findings, though stable, do not always reach a level that passes the conventional significance threshold.

Our study examined the performance of TFA and non-TFA teachers from two perspectives. First, we made a straightforward comparison of the average test-score gains in classrooms run by TFA and non-TFA teachers, controlling for a variety of factors known to influence academic achievement, including students' backgrounds, the students' previous performance on the TAAS, characteristics of their schools, and characteristics of their classmates. In this way, we could isolate the typical effect of having a TFA teacher versus a non-TFA teacher. Second, we wanted to examine the distribution of performance among TFA and non-TFA teachers. For instance, how do high-performing TFA teachers compare with high-performing non-TFA teachers? How bad are the worst TFA teachers? The worst non-TFA teachers? Do the TFA teachers vary much in quality, or are they fairly similar? To address these questions, we estimated how much of the achievement gains in each classroom was attributable to that classroom's teacher.

Our study analyzed and controlled for the following factors:

Teacher Characteristics

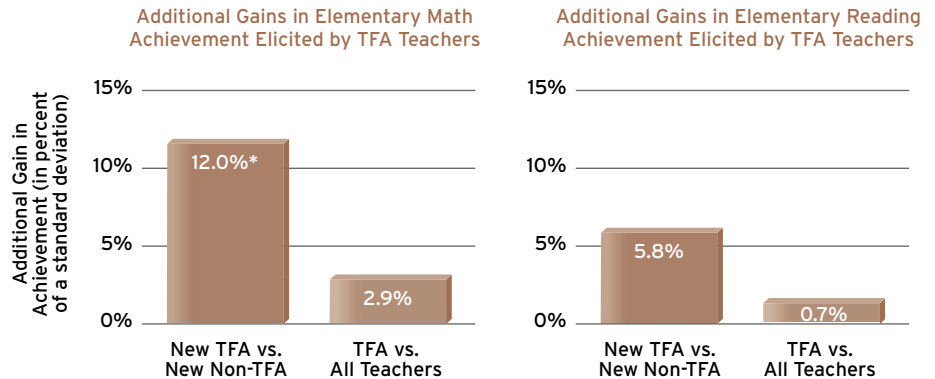
- TFA status
- Years of teaching experience

Student Characteristics

- The school attended
- School year
- Minority status
- Eligibility for free or reduced-cost lunch (a proxy for low socioeconomic status)
- Date of birth
- Assigned teacher
- English language proficiency
- Testing exemption status

Students of Teach for America Recruits Learn More (Figure 2)

New TFA recruits elicited gains in student achievement that were 12 percent of a standard deviation higher in elementary-school math and 5.8 percent higher in reading than those of students of other new teachers.



*Denotes statistical significance at $p < .05$

SOURCE: Authors' calculations

- Test scores on the annual TAAS exam in each year in which the student was enrolled in a Texas school

School Characteristics

- Share of students who are African-American
- Share of students who are Latino
- Share of students receiving free or reduced-cost lunch

Classmate Characteristics

- Share of a class receiving free or reduced-cost lunch
- Share of a class that scored below the state mean on the previous year's TAAS exam

Findings

Before the analysis was concluded, students' raw scores on the TAAS exam were standardized to a mean of 0 and a standard deviation of 1. This procedure was performed separately for each year in order to account for variations in the TAAS test from year to year. The average score for a given year was transposed to 0, and all the other scores were distributed around that point in a standard normal distribution. Gains

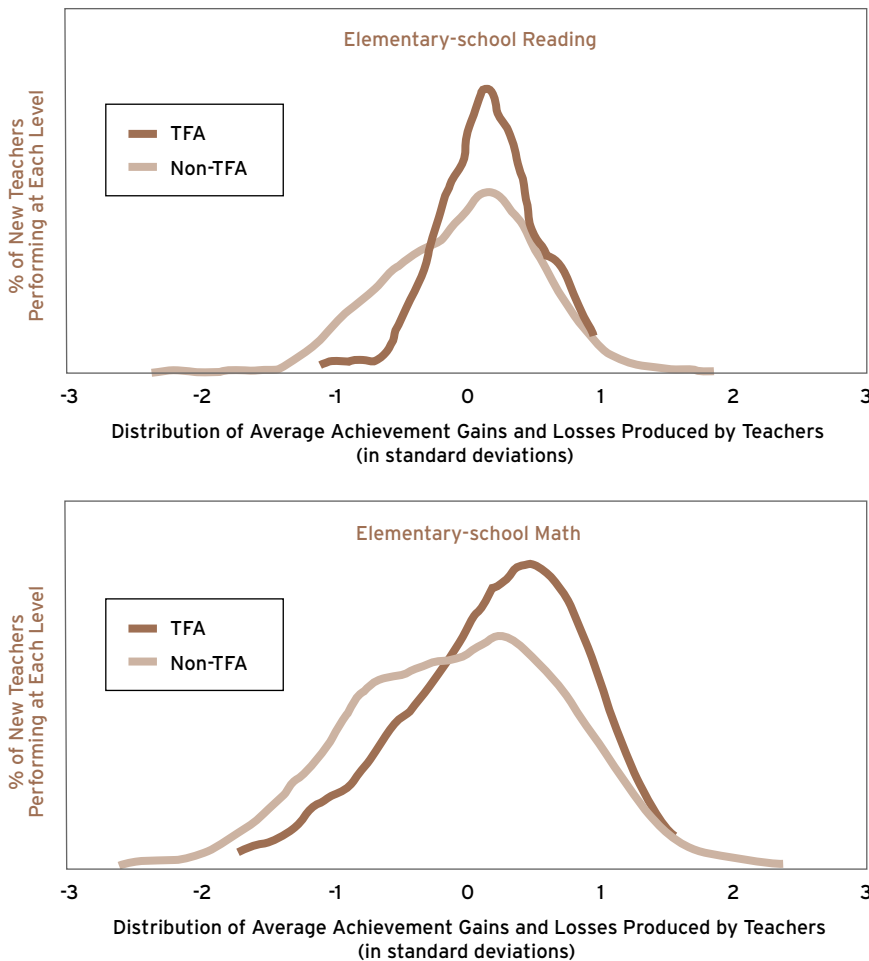
are therefore expressed in terms of standard deviations. A gain of 5 percent of a standard deviation in one year may be considered large, if it is thought that students will continue to experience similar annual gains through the elementary and secondary years. (A full standard deviation is approximately the difference between the test scores of black and white students nationally. More than half of the black-white test-score gap could be erased by steady gains of 5 percent a year among black students.)

Reading. In the elementary grades 3 through 5, students of new Teach for America teachers gained an average of 5.8 percent of a standard deviation more on the TAAS reading exam than did students with other new teachers, a difference that fell just short of statistical significance (see Figure 2). The students of the entire sample of Teach for America teachers made gains that were essentially similar to those made by students of all teachers in the Houston district.

There tended to be less variation in the performance of TFA teachers than there was among non-TFA teachers.

Better and More Consistent (Figure 3)

Of new Teach for America recruits, 63 percent elicited achievement gains in elementary-school reading that were higher than the median new non-TFA teacher was able to produce. The share was 64 percent in elementary-school math. The performance of TFA recruits was also more bunched up around the median, meaning that they were also more consistent in their results, making them less risky to hire.



SOURCE: Authors' calculations

More than 63 percent of TFA teachers generated achievement gains in reading that were higher than the median achievement gains for new non-TFA teachers (see Figure 3). More than 60 percent of TFA teachers did better than the median performance of all teachers. Moreover, TFA teachers tended to be more bunched up around the median than their non-TFA peers. In other words, there were fewer extremely low- and high-performing teachers among the Teach for America recruits; they

were more consistent in eliciting achievement gains. The difference between the two distributions was found to be statistically significant in both cases.

Mathematics. Elementary students benefited from having a TFA teacher in math as well. The achievement gains of students with new TFA teachers were 12 percent of a standard deviation higher than those with other new teachers, a result that was statistically significant (see Figure 2). Students of all

TFA teachers gained 2.9 percent of a standard deviation more in math than did students of all teachers in the Houston district, a difference that was not statistically significant.

Let's look again at the distribution of performance among TFA and non-TFA teachers. Of new TFA teachers, 64 percent generated achievement gains among their students that were higher than the median achievement gains for new non-TFA teachers (see Figure 3). Fifty-eight percent of TFA teachers performed above the median for all teachers. The lowest-performing new teachers were much worse than the lowest-performing TFA teacher. The best-performing TFA teachers were better than virtually all the teachers in the district. In other words, if you were choosing between two math teachers, and the only thing you knew about them was that one was a TFA member and one was not, you would choose the TFA member. This would give you the best chance of selecting a good teacher.

Our results also suggest that years of teaching experience significantly affect student performance. The largest gains in teacher effectiveness occur early in a teaching career and diminish thereafter until flattening out at around 8 years for math and 11 years for reading. The effect of experience operates independently of teachers' characteristics, so by isolating these effects, we provided a truer test of differences in teachers' backgrounds.

More and Better

TFA appears to be a viable alternative source of high-quality teachers, compared with both the pool of new teachers available for hire in Houston and the district's entire workforce of current teachers. All our results show the average TFA teacher improving her students' performance by more than new teachers and at least as much as all teachers in Houston. The advantage of having a TFA teacher was largest in math, though TFA teachers also generated

gains stronger than those of non-TFA teachers in reading. This dispels the notion that TFA is inferior to other sources of teachers. The findings in reading and the comparison between TFA teachers and all teachers in math were not statistically significant, owing largely to the small numbers of TFA teachers in proportion to the respective comparison groups. However, the fact that across many comparisons all the findings pointed in one direction is a strong indication that the general pattern is correct: that TFA teachers produce results as strong as or better than their peers.

Of course, as with any program, there were some TFA recruits who did not perform well in the classroom. This is likely to continue. However, the distribution of performance among teachers clearly shows that the lowest-performing teachers were consistently not TFA teachers and that TFA teachers make up a disproportionate number of high performers. Moreover, TFA teachers were more consistent in their performance than their non-TFA peers, which means that schools face lower risks in hiring them. The fact that many of the highest-performing teachers in the entire district were supplied by TFA is especially meaningful in light of the program's small contribution to the size of the district's workforce compared with other sources of new teachers.

Several caveats are in order. First, this study was based on data from a highly dynamic time in the Houston school district—and in public education more generally. Schools and teachers in Houston were subject to a wide variety of other programs, constraints, and opportunities that no analysis could possibly capture.

Furthermore, these findings cannot tell us what aspects of TFA and non-TFA teachers accounted for the differences in their students' performance. Is the effectiveness of TFA teachers due to the type of people being recruited, the difference in their academic backgrounds, the support provided by TFA,



Teach for America teachers make up a disproportionate number of the high-performing teachers in the Houston Independent School district.

participation in the district's alternative-certification program, or a combination of factors? We encourage further study of these questions.

A third caveat relates to the criticisms levied against us by the National Commission on Teaching & America's

not present data indicating how TFA teachers performed compared to teachers who came into the profession fully qualified and certified to teach. . . . To say that TFA teachers do just about as well as other new teachers [in Houston], an extraordinary number of whom are extraordinarily underqualified, is a weak endorsement at best."

First, this study was not designed to test the benefits of the various teacher-preparation programs or of certification. We do not have the ability, for example, to isolate the effect of Houston's alternative certification program, because an appropriate compar-

ison group was not available. Instead, our evaluation was explicitly intended to assess Teach for America relative to all other sources of new teachers currently available to school districts like Houston. The Houston district draws from a pool of fully certified, tradi-

If you were choosing between two math teachers, and the only thing you knew about them was that one was a TFA member and one was not, you would choose the TFA member.

Future (NCTAF) after our report's initial release. NCTAF, a strong promoter of traditional teacher-education programs and more stringent certification requirements, writes, "[CREDO] does

tionally prepared teachers and those who agree to work toward certification through the district's alternative certification program. Our results show that TFA teachers perform modestly better

than those from other sources of new teachers. Studies of the relative merits of various certification policies and programs are clearly important to the continuing policy debates concerning teacher preparation and teacher quality, but they ask questions that are fundamentally different from those of the study described here.

Second, we didn't just compare TFA teachers with the Houston district's other new hires, a fair share of whom are uncertified and didn't attend a traditional education school (though not nearly as many as NCTAF claims). We also compared TFA teachers with all the teachers in the district, the vast majority of whom are certified. The TFA recruits performed just as well as—if not better than—this pool of teachers; in fact, at least in elementary-school math, the best TFA teachers performed better than the district's best teachers after the data were adjusted for the effects of teaching experience.

Informing Policy

While these findings cannot speak to the effectiveness of various certification policies, they at least dispel the notion that only the traditional route to teaching can produce good teachers. What needs to be studied is the degree to which subject-matter skills and strong performance in college can substitute for pedagogical training. TFA corps members are an admittedly select group of college graduates, culled from the finest universities and often performing near the top of their class. These are not the types of students who ordinarily go into teaching. It's possible that traditional certification programs and pedagogical training are less necessary for them than they are for the typical teacher.

The implications of the findings extend beyond teaching preparation. Earlier research has identified strong returns to experience in the early years of teachers' careers. The year-to-year growth in teachers' effectiveness, as mea-

sured by student achievement, is largest in the early years of their careers and then quickly tapers off. In the past, attention has been paid to the average student gains for each year of teaching. The findings in this study suggest that the benefits of teaching experience may not be uniform—that new teachers may vary considerably in their initial starting points and in how quickly they improve. Therefore, differences in the value teachers glean from additional professional development may also be in play. These

What needs to be studied is the degree to which subject-matter skills and strong performance in college can substitute for pedagogical training.

questions cannot be answered without further study. The point is raised here, however, as a reminder that policy decisions concerning teacher training and certification cannot assume that a fixed set of requirements will make all teachers perform the same.

The present analysis of teacher effectiveness raises other questions beyond certification. One cherished tenet of the NCTAF position is that TFA is pernicious because the majority of TFA teachers leave after completing their two-year commitment. However, the mythical teacher-for-life doesn't exist. Across the four cohorts of teachers included in this study who completed two years or more of teaching, more than a third of TFA teachers remained past their commitment. The NCTAF criticism ignores the fact that 45 to 50 percent of other new teachers leave after two years. The

NCTAF assertion that no benefits can accrue to either students or the district from teachers who leave after a few years must apply to both groups or not at all. On the contrary, the evidence here points to the opportunity to identify early in teachers' careers whether they are effective, thus allowing those who are not effective to be encouraged to consider alternative occupations.

Once other states begin to develop the rich data resources found in Texas, this evaluation should be expanded to include other TFA districts. Corroborating and expanding the results obtained here would contribute to both the TFA program itself and the larger policy world in which TFA continues to grow. TFA has been shown to be a viable source of new teachers for Houston, both in number and in quality. With continuing attention to retention during their two years of service and with many TFA teachers choosing to remain after their commitment has been fulfilled, TFA is likely to create an enduring positive presence in the Houston Independent School District and elsewhere.

Expanding the supply of high-quality teachers is of the utmost importance, but there is little consensus on how to do so. The findings reported here suggest that multiple paths into teaching are both possible and desirable. They suggest that the way to ensure high-quality teaching is to set performance goals and hold teachers accountable for meeting them. The ultimate measure of a teacher-preparation program, traditional or alternative, is its ability to produce good teachers, not its ability to meet an unrelated set of requirements that are not grounded in empirical proof of student performance.

—Margaret Raymond is the director of CREDO, a nonpartisan research group at the Hoover Institution, Stanford University. Stephen Fletcher is the assistant director of CREDO.