

Educational Opportunity and Alternative Certification: New Evidence and New Questions

A SCOPE Policy Brief

by Linda Darling-Hammond

Educational Opportunity and Alternative Certification: New Evidence and New Questions

Linda Darling-Hammond

Over the last decade, a healthy debate has been launched about how to open up the teaching profession to a wider range of talented recruits, better prepare them, and create pipelines of candidates into hard-to-staff schools in poor urban and rural communities. Reforms of teacher education have raised standards for entry, focused on preparation to teach specific content, and created professional development school partnerships – much like teaching hospitals -- to ensure that student teachers and veterans can develop and learn best practices.¹ Alternative routes have also been launched in most states to enable mid-career recruits and other college graduates to enter teaching through programs other than the traditional four-year undergraduate models that have dominated since the 1950s.² These routes have helped to staff high-need schools, with providers forging tighter connections with the human resource offices of high need districts than universities typically have done. But key questions remain: To what extent do teacher education programs of different kinds – whether traditional or alternative – improve student achievement?

Alternative program models vary substantially. In some states, such programs are Master of Arts in Teaching (MAT) programs that provide all or most coursework and student teaching prior to their candidates' assuming a teaching position. In other cases, candidates receive several weeks of training in the summer and take on responsibilities as teacher of record in the fall, while they complete coursework for a credential – sometimes comparable to that completed by other recruits, and sometimes less – while they receive mentoring of varying amounts and quality, depending on the program. And there are many other variations in between.

Research on the outcomes of alternative programs has found mixed results. Four well-controlled longitudinal studies, using individual-level student data from Houston, Texas, New York City, and North Carolina, found that teachers who entered teaching before completing preparation—on temporary licenses or as alternative route candidates--were less effective than fully prepared beginning teachers working with similar students in their initial year or two on the job, especially in teaching elementary reading.³ However, these studies also found that, for those alternative route teachers who completed the education coursework for full certification and gained experience, there were few significant differences in effectiveness. Indeed, in some cases, the students of experienced teachers from selective alternative route programs had larger gains in mathematics than those of other similarly experienced teachers. However, higher attrition rates for these alternative routes raise questions about how entry paths affect both individual teacher effectiveness and the overall effectiveness of the teaching force, as the level of individual and collective teacher experience in a school strongly affects student achievement.⁴

Finally, findings from analyses of teacher effectiveness depend substantially on the nature of the comparison groups examined. For example, two studies of alternative route recruits found them to be as effective as other teachers in their school or district when the comparison group teachers, often hired on temporary or emergency permits, were even less likely to be trained and certified than the alternative candidates,⁵ but less effective when compared to a more qualified group of teachers.⁶ Thus, studies of the efficacy of various routes into teaching should examine teachers' experience, and carefully consider the nature of the comparison group.

The Mathematica Study of Alternate Routes to Certification

Recent findings from a Mathematica study comparing the performance of teachers prepared via alternative and traditional routes add additional evidence and raise more questions. The headline findings from the study, “An Evaluation of Teachers Trained through Different Routes to Certification,” suggested that students of teachers from alternative routes that allowed them to enter teaching before finishing their training performed no differently, statistically, than those of traditional-route teachers. Some have interpreted this finding to suggest that policymakers and practitioners should expand the use of fast-entry alternative routes and seek teachers trained through such programs, as they presumably perform as well in the classroom as any other teacher trained through traditional schools of education anywhere in the country.

This conclusion would be incorrect. In fact, both alternative and traditional teachers in the hard-to-staff schools selected for the study had less training than most teachers nationally, and neither group was highly effective. Teachers from the “low-coursework” alternative routes actually *lowered* their students’ achievement over the course of the year. Those from “high-coursework” programs, along with their traditional route counterparts, did somewhat better, but not very well, raising their students’ scores by only about 1 to 2 normal curve equivalent (NCE) points between fall and spring, with a slight advantage to traditional candidates. This is particularly discouraging given how far behind in achievement their low-income students of color already were.

If our goal is to provide access to high-quality public education to all students, particularly those placed at risk, this point is critical. Most of the alternative route teachers studied were not effective in boosting student achievement, and their traditional route counterparts did only marginally better. In fact, average student test scores dropped in the classrooms of those with the least training, leaving more of their students further behind. This negative impact has been masked in many studies when AC teachers are compared to other ill-prepared teachers – including emergency hires, temporary substitutes, and candidates from weak preparation programs -- who are also relatively ineffective with these struggling students.

Unfortunately, the Mathematica study provides little guidance for teacher education policy generally. Although the data tended to favor the traditional route teachers in the sample schools, these teachers still had much less preparation than most elementary teachers nationwide – on average about half the amount of coursework generally required by a majority of states. The study did not examine the effects on performance of the coursework candidates had actually taken, nor did it evaluate the effects of their student teaching, mentoring, or professional development experiences, so it did not illuminate the aspects of teacher preparation that might make a difference in teachers’ effectiveness.

The study offers little insight into the kinds of preparation that can help teachers significantly raise student achievement and close the achievement gap. Fortunately, other recent studies have examined the kinds of qualifications and training that produce stronger achievement gains for students. Those studies are discussed later in this brief.

What the Mathematica Study Did

The Sample. Mathematica sought out schools that hire large numbers of alternatively certified (AC) teachers from non-selective programs. (This meant, for example, that programs like Teach for America, which selects high-achieving college graduates, were excluded.) The researchers then drew a sample of 87 such teachers, matching them to traditionally certified (TC) teachers in the same buildings. This was not an easy sample to create, as such schools can only be found in a small number of states and districts that permit the hiring of alternative route elementary teachers. The researchers noted that they had to eliminate from consideration states with higher standards and those without such programs. Ultimately, they found only 12 states that had active AC programs for elementary teachers that met their criteria. In two of these – Michigan and Wisconsin – the programs operated in a single district (Detroit and Milwaukee, respectively), and not in the rest of the state, where well-trained teachers are plentiful.

All of the districts in the sample were majority “minority” and low income; most served more than 90 percent students of color. The schools in the study were even more economically disadvantaged and had more minority students than other schools in the same districts. Overall, more than 85 percent of students were racial/ethnic minorities and more than 65 percent were eligible for free or reduced price lunch. Most of the schools had hired more than 20 percent of their teachers from alternative routes. At four schools, AC teachers accounted for more than half of all staff.

In most of these states, schools cannot hire alternative route teachers unless traditionally-prepared teachers are not available, so the sample was drawn from the hardest-to-staff schools in jurisdictions with the least selective hiring standards where researchers have found that teachers are less well-qualified on nearly every dimension. For example, Lankford, Loeb, and Wyckoff have shown that on every measure of qualifications – certification, subject matter background, pedagogical training, selectivity of college attended, test scores, or experience – less qualified teachers are found in segregated low-income schools.⁷ Reviewing a number of such studies, Education Trust President Kati Haycock noted that the statistics on differentials in credentials actually *understate* the degree of the problem in the most impacted schools:

The fact that only 25% of the teachers in a school are uncertified doesn’t mean that the other 75% are fine. More often, they are either brand new, assigned to teach out of field, or low-performers on the licensure exam. . . . There are, in other words, significant numbers of schools that are essentially dumping grounds for unqualified teachers – just as they are dumping grounds for the children they serve.⁸

Indeed, the Mathematica study data show that even the traditional route teachers in these sites had completed only about half the amount of coursework generally expected for elementary teachers in most states. Thus, the study examines teachers in schools with some of the weakest teaching forces in the country. The sample of teachers is not representative of those in states with higher standards and cannot be generalized to districts and schools that draw from a better-prepared pool and hire much more selectively, or to the overall teaching pool nationally.

In the initial year of the study, researchers could not secure enough participants to create an adequate sample size, so they went back the following year, expanded the definition of “beginners” to those with up to five years of experience, and replaced the unusually large number of teachers who left in the middle of the year from these schools. They ended up including 14 teachers twice, using them as separate subjects in two separate years. (As described below, these 14 teachers were recorded as though they had the same amount of training in both years, even if they had completed more of their required coursework between the first and second year.) All of these problems are understandable, but they make the sample less useful for studying teacher education effects, since most of the teachers were no longer novices (the average experience level was three years across all the groups), and only those who had stayed in teaching were captured in the analyses.

Comparisons of Programs. The researchers divided the teacher sample into four groups: alternative-route teachers enrolled in “low-coursework” programs and their traditional-route matches in the same schools, and alternative-route teachers enrolled in “high-coursework” programs plus their traditional-route counterparts in the same schools.

Although the study purported to look at the effects of coursework on effectiveness, it did not actually analyze the coursework teachers had in fact taken, only their program category. Unfortunately, these categories did not describe what teachers had individually experienced. The requirements of AC and TC programs were highly overlapping. Some alternative programs actually required more hours of coursework than some traditional programs; and some candidates in AC programs had student teaching while some in TC programs had none. Furthermore, candidates were at various stages of program completion – and large proportions of both the AC and TC groups were still taking coursework -- so categorizing them according to their program’s requirements did not describe the kind or amount of coursework they had actually taken. In addition, the AC group received significantly more mentoring and much more extensive professional development coursework in methods of teaching than the TC group.

In their analyses regarding the “effects” of coursework, researchers did not record or use the amount of actual teacher education coursework, student teaching, mentoring, or additional professional development individuals received, despite the small size of the sample and the fact that the data were available. As a consequence, the study’s treatment group categories were badly contaminated. There were teachers in the alternate route category that received more teacher education, professional development coursework, and mentoring than some in the traditional route category – confounding the inferences the study sought to draw.

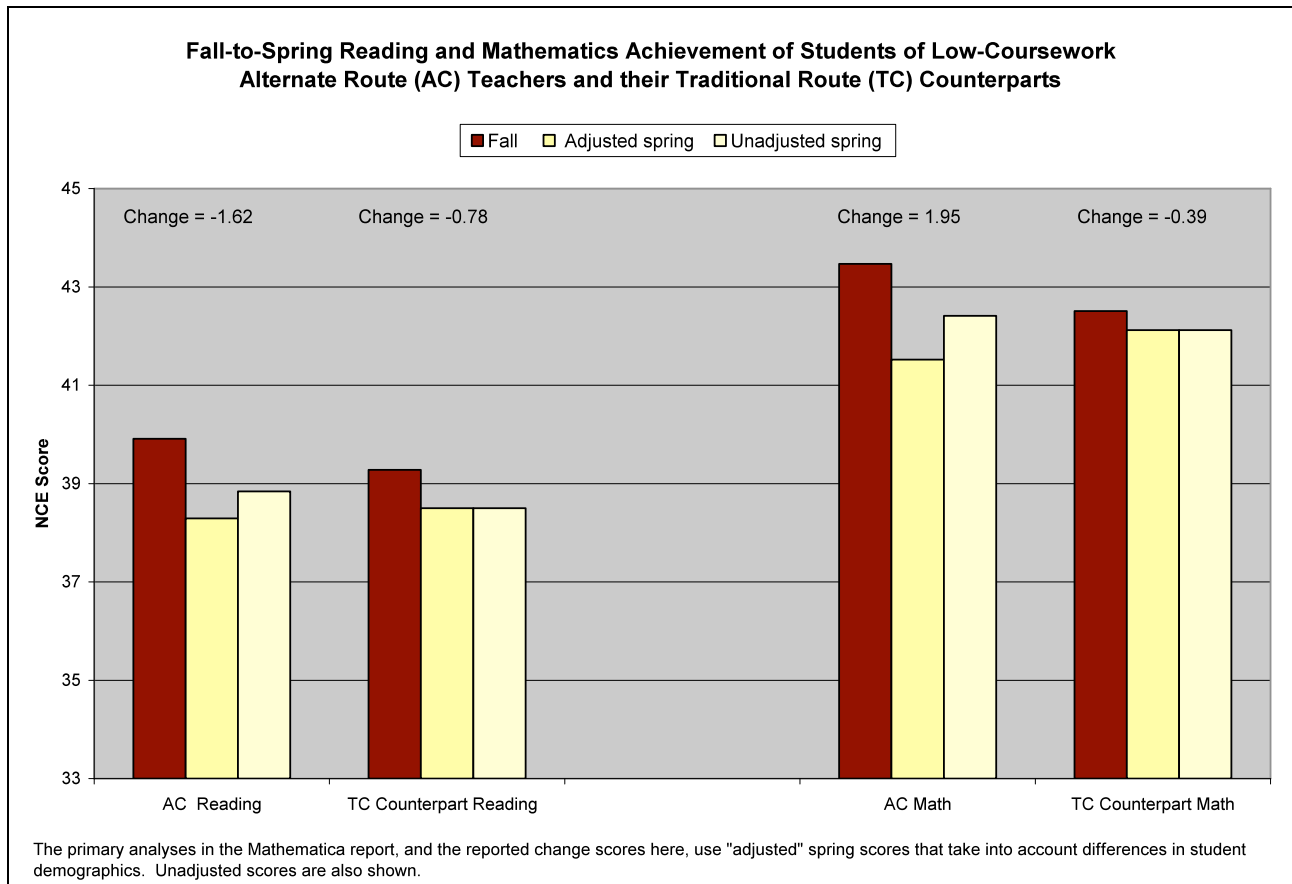
Contamination of treatment groups is a serious problem for an experimental study. The problems in this study would be the equivalent of a medical trial of two cancer treatments in which candidates in the treatment groups got varying doses of some of the same drugs, as well as some different ones, with the combination differently mixed for each patient. In addition, sometimes individuals received elements of the other group’s treatment. If the groups were not properly designed to represent clearly distinctive treatments, and the researchers did not keep careful track of who got what, ignoring the actual treatments patients received in their analysis, one would not expect the study to be useful in determining the effects of such badly defined

treatments. This is essentially the problem the Mathematica study faces, and why it is unlikely to find large differences in group effects.

The likelihood of finding significant effects was also reduced by the fact that, within an already small teacher sample, nearly all analyses were conducted for even smaller subsamples, divided into the “low-coursework” and “high-coursework” categories.

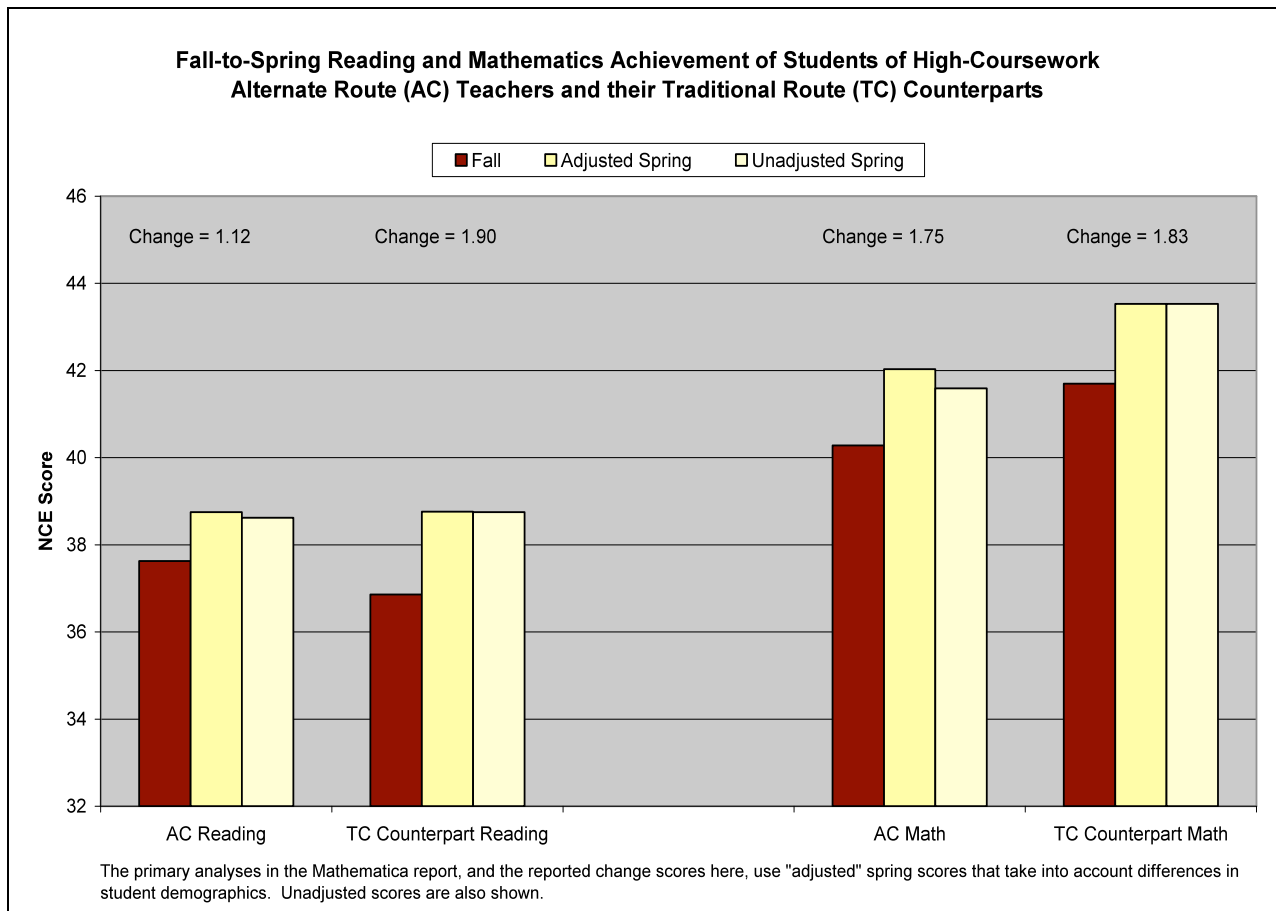
What the Study Found

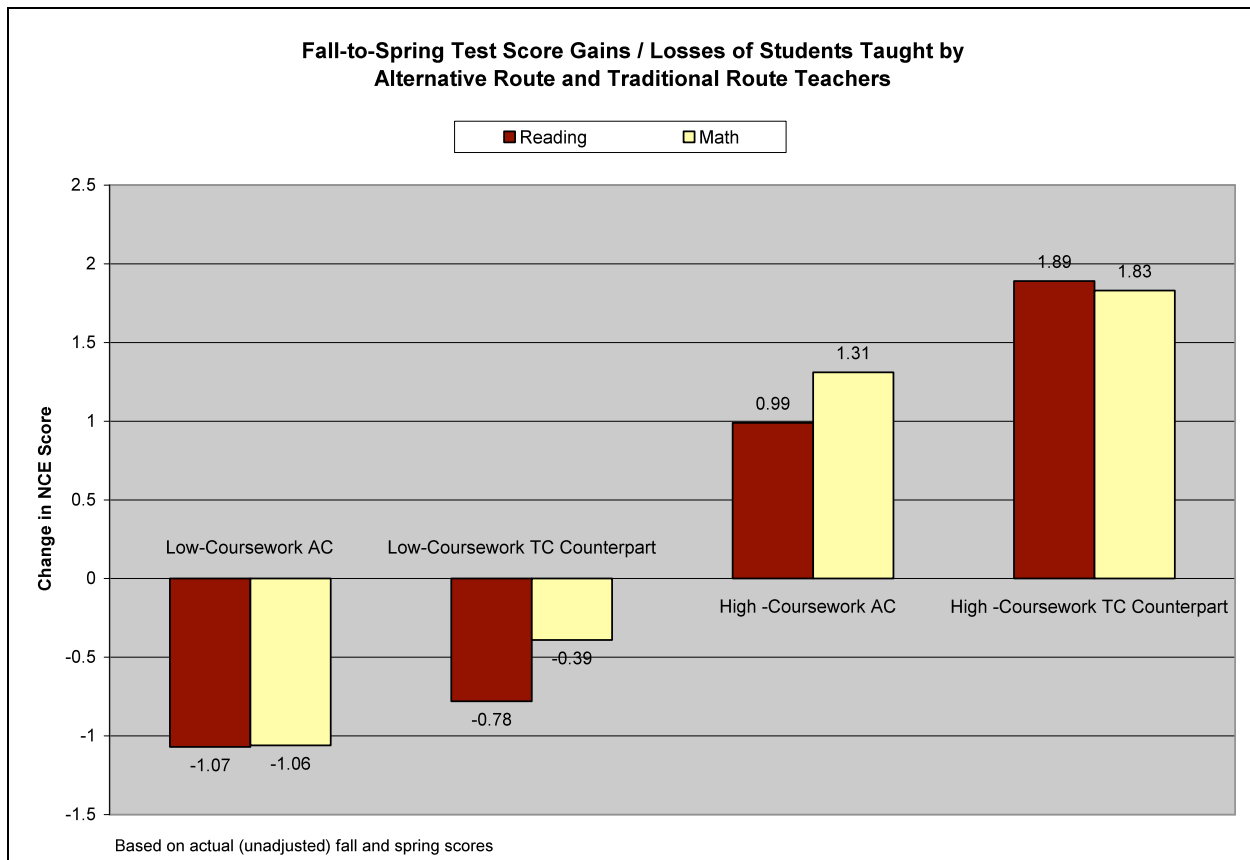
The study found few significant differences in spring scores for the students taught by AC and TC teachers. However, it did not find that these teachers are effective in producing student learning. In fact, a close look at the data shows that the students of teachers from what the study called “low-coursework” alternative programs (AC) actually *declined* by nearly two percentile points in their reading and math achievement between fall and spring. Students taught by their traditional route (TC) counterparts, who started somewhat further behind in the fall, declined by a smaller amount, and ended up roughly comparable to those of the AC teachers by spring. In the segregated high-minority, low-income schools that hire AC teachers, none of the teachers or their students did well.



Students of teachers in the “high-coursework” alternative programs and their traditional route counterparts did somewhat better, but did not make substantial gains. Their students improved by between 1 and 2 NCE points. Gains were slightly greater for the traditional route teachers than the AC teachers, but not nearly enough to begin closing the achievement gap, given how far behind their low-income students of color already were.

The analyses shown above use data reported in the study,⁹ however, the Mathematica researchers looked only at the spring scores of students in drawing their conclusions, rather than gain scores from fall to spring. Looked at over the course of the academic year, students of traditional route teachers started further behind in all but one case and either gained more or declined less, so they ended up scoring similarly or slightly ahead of the those of the AC teachers. Indeed, the students of the “high-coursework” AC and TC teachers also started behind those of the “low-coursework” group and produced greater gains, as shown below. Because the researchers did not report standard deviations around the mean scores in the fall and the spring, we cannot evaluate the significance levels of these gains.





The study discovered, as has previous research, that alternative candidates (AC) who were still taking courses while teaching were significantly less effective in both reading and mathematics than their traditional route counterparts (TC) who were not taking coursework.¹⁰ In addition, correlational analyses showed that, controlling for experience (defined as “certified experience” or “certified plus long-term sub experience”), TC teachers outperformed AC teachers in virtually all comparisons at marginally significant levels ($p=.08$ or $.09$) in mathematics. Had the small sample not been divided into subgroups for analysis, several of these comparisons would have been significant at the $p < .05$ level.¹¹ Students of AC candidates from California also did significantly less well in math than students of their traditional route counterparts. The researchers noted:

The general pattern of the negative difference in math scores for students with AC teachers compared to students of their TC counterparts persists across states and is statistically significant in California... The relative effect on math in California is negative and statistically significant (an effect size of -0.13 , nearly twice the overall effect size for high-coursework AC teachers from the basic experimental model).

Students of low-coursework alternative route candidates in their third and fourth years scored lower in both reading and mathematics than those of their TC counterparts. However, for most of those who completed their programs and stayed, the differences, while still generally negative, were small and statistically insignificant.

While one can read these findings as suggesting that the alternative route candidates who teach in these difficult contexts “do no more harm” – or at least not much more harm – than most other teachers in their schools, these poor outcomes are not an acceptable standard. They represent a race to the bottom for the students and schools in these communities, rather than the race to the top we need to create substantially higher levels of teacher effectiveness, especially for children who have been left furthest behind.

Preparation that Produces Positive Outcomes for Student Learning

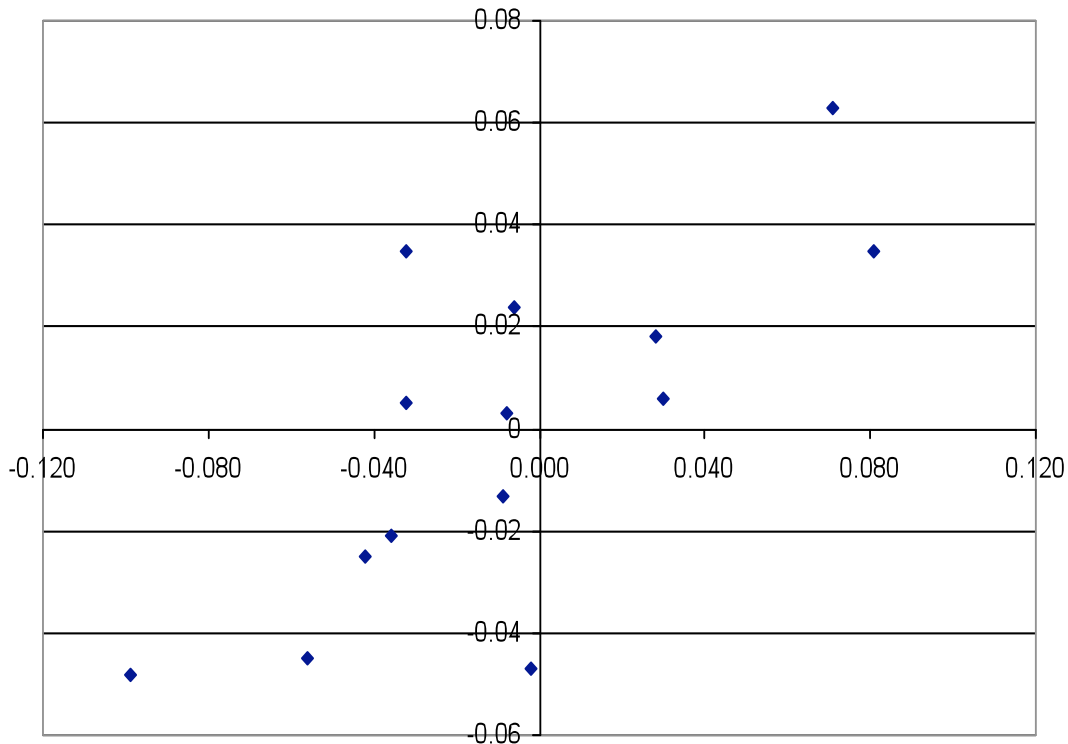
Fortunately, some states and districts have developed policies that recruit well-prepared teachers to high-need schools,¹² and studies have begun to identify the kinds of teacher qualifications and the features of teacher preparation programs that support gains large enough to begin to close the achievement gap.

A recent study of high school students in North Carolina, for example, found that students gain significantly if their teachers are fully prepared when they enter (rather than entering through a “lateral entry” alternative route requiring no prior teacher preparation), are certified in the specific field they teach, have higher scores on the teacher licensing test, have taught for more than two years, have graduated from a competitive college, and have become National Board Certified by completing a performance assessment documenting their teaching.¹³ While each of these variables has a significant effect on achievement in its own right, the influence of having a teacher with most of these qualifications – like many of the teachers in affluent suburbs -- as compared to one having few of them – like many in poor urban schools – is larger than the combined effects of race and parent education, e.g. the average difference in achievement between a typical white student with college-educated parents and a typical black student with high-school educated parents.

A similar study of teachers in New York City¹⁴ found that teachers’ certification status, pathway into teaching, teaching experience, graduation from a competitive college, and math SAT scores were significant predictors of teacher effectiveness in elementary and middle grades mathematics. Student achievement was most enhanced by having a fully certified teacher who had graduated from a university pre-service program, who had a strong academic background, and who had more than two years of teaching experience. Students’ achievement was hurt most by having an inexperienced teacher on a temporary license – again, a teaching profile most common in high-minority, low-income schools. In combination, improvements in these qualifications reduced the gap in achievement between the schools in deciles serving the poorest and most affluent student bodies by 25 percent.

Later analyses of these New York City data by the same research team found that some individual teacher education programs have much more positive effects than others, based on their graduates’ contributions to value-added student achievement.¹⁵ In the figure shown below, programs above the x-axis had graduates whose students experienced strong gains in English language arts or math; those in the upper right hand quadrant had positive outcomes in both subjects.

Program effects in Math (x-axis) and ELA (y-axis) for first-year teachers 2000-01 through 2005-06 (Institutions with 40 or more teachers with value-added estimates)



Source: Boyd, Grossman, Lankford, Loeb, Wyckoff (2009).

The researchers examined the features of programs that influenced their graduates' effectiveness, producing findings very similar to those from previous studies of exemplary programs. These features included:

- Candidates' student teaching experience and the match between the context of student teaching and their later teaching assignment
- Programs' careful oversight of the quality of candidates' field experiences
- A focus on helping candidates learn specific practices applied in clinical experiences
- The amount of coursework in content areas (math and reading) and in methods of teaching mathematics
- Candidates' opportunities to study the local district curriculum
- A capstone project (typically a portfolio of work done in classrooms with students)
- Programs' percentage of tenure-line faculty, which the researchers viewed as a possible proxy for institutional investment and program stability.

Much more work needs to be done to study and create preparation programs that can *significantly* increase teacher capacity and effectiveness, and to create recruitment and compensation policies that can provide such teachers to all of our schools, including and especially those that serve our highest-need children. To start a race to the top, we must identify those programs and those aspects of teacher preparation that produce highly effective practice, and then focus policy on replicating what works.

Endnotes

-
- ¹ See, e.g., L. Darling-Hammond & J. Bransford (2005). Preparing teachers for a changing world: What teachers should learn and be able to do. National Academy of Education. San Francisco: Jossey-Bass.
- ² For a summary, see E. Feistritzer & C. Haar (2006). Alternate Routes to Teaching. Pearson Education.
- ³ D. Boyd, P. Grossman, H. Lankford, S. Loeb, & J. Wyckoff (2006). How changes in entry requirements alter the teacher workforce and affect student achievement. Education Finance & Policy, *1* (2); 176-216; C. Clotfelter, H. Ladd, & J. Vigdor (2007). How and why do teacher credentials matter for student achievement? (NBER Working Paper 12828). Cambridge, MA: National Bureau of Economic Research.
- L. Darling-Hammond, D. Holtzman, S.J. Gatlin, & J.V. Heilig (2005). Does teacher preparation matter? Evidence about teacher certification, Teach for America, and teacher effectiveness. Education Policy Analysis Archives, *13* (42). <http://epaa.asu.edu/epaa/v13n42/>; T. J. Kane, J.E. Rockoff, & D.O. Staiger (2006). What does teacher certification tell us about teacher effectiveness? Evidence from New York City (NBER Working paper no. 12155). Cambridge, MA: National Bureau of Economic Research.
- ⁴ Ibid.
- ⁵ P.T. Decker, D. Mayer, & S. Glazerman (2004). The effects of Teach For America on students: Findings from a national evaluation. Washington, DC: Mathematica. Darling-Hammond et al., 2005.
- ⁶ Darling-Hammond et al. (2005).
- ⁷ H. Lankford, S. Loeb, & J. Wyckoff (2002). Teacher sorting and the plight of urban schools: A descriptive analysis. Educational Evaluation and Policy Analysis, *24*(1), 37-62.
- ⁸ K. Haycock (Spring 2000). No more settling for less. In Thinking K-16, *4* (1), p. 11. Washington DC: Education Trust.
- ⁹ Exhibit II.6 on p. 23 provides students' baseline pre-test scores. Exhibit IV.1 on p. 56 provides spring reading scores that are regression-adjusted for baseline test scores, eligibility for free or reduced-price lunch, gender, race/ethnicity, and teacher's years of experience. Exhibit IV.3 on p. 58 provides regression-adjusted spring mathematics scores. Exhibit A.7 on Appendix p. A21 provides unadjusted spring scores in reading and mathematics.
- ¹⁰ The reading results are reported differently in three different places in the study. On p. xxx, the report states: "Students of AC teachers who were taking coursework toward certification or a degree scored lower in reading (effect size -0.13) than did students of their TC counterparts who were not taking coursework." On pp. 71-72, the same finding is reported with an effect size of -0.12 . In footnote 72 on p. 71, the report states: "When the subgroup analysis from Exhibit IV.9 is restricted to those instances in which AC teachers are taking courses and TC teachers are not, the effect size is -0.06 ($p=0.23$) for reading and -0.11 ($p=0.04$) for math."
- ¹¹ These analyses are shown in Exhibit A.11 p. A25.
- ¹² L. Darling-Hammond and G. Sykes (2003). Wanted: A national teacher supply policy for education: The right way to meet the 'highly qualified teacher' challenge. Educational Policy Analysis Archives, *11* (33).
- ¹³ Clotfelter, Ladd, & Vigdor (2007).
- ¹⁴ D. Boyd, H. Lankford, S. Loeb, J. Rockoff, & J. Wyckoff, (2007, August). The narrowing gap in New York City teacher qualifications and its implications for student achievement in high-poverty schools. Urban Institute: National Center for Analysis of Longitudinal Data in Education Research (CALDER).
- ¹⁵ Boyd, D., Grossman, P., Lankford, H., Loeb, S., & Wyckoff, J. (September 2008). Teacher preparation and student achievement. NBER Working Paper No. W14314. National Bureau of Economic Research. Available at SSRN: <http://ssrn.com/abstract=1264576>.