

No Guarantees:

Is it Possible to Ensure Teachers Are Ready on Day One?

Chad Aldeman and Ashley LiBetti Mitchel



BELLWETHER
EDUCATION PARTNERS
IDEAS | PEOPLE | RESULTS



Table of Contents

Introduction	3
We Don't Know How to Train Good Teachers	5
The Look to Outcomes	10
Following the Existing Research	17
Embracing This Agenda	27
Endnotes	28
Acknowledgments	32

Introduction

This concept—that we can't identify a great teacher on the basis of her résumé—should be liberating. But for some reason, it's largely ignored in the education field.

The single best predictor of who will be a great teacher next year is who was a great teacher this year.

The second best predictor is... Well, there really isn't one that's close. Some great teachers are seasoned veterans, while some are new to the field. Experience matters, of course, particularly in the early years of a teaching career, but it's no guarantee of teacher effectiveness; there is a wide variation in performance, even among teachers with the same amount of experience. Some great teachers went through a traditional preparation program with a standard student-teaching experience. Others perform just as well with barely any prior training. Some great teachers are Ivy League grads, but others come from their local state colleges. The list could go on.

This concept—that we can't identify a great teacher on the basis of her résumé—should be liberating. But for some reason, it's largely ignored in the education field. Instead, without worrying too much about whether requirements will make teachers more effective, each state imposes myriad conditions on potential teachers and the programs they must attend to be eligible to teach.

The result is requirements that amount to little more than barriers—new ways to keep potential teachers out of schools. These requirements, from licensure tests to minimum GPAs and SAT scores, from rules on student teaching to teacher-performance assessments, do not guarantee effective teachers. For some of these requirements, the research reports mixed results. Other requirements are too nascent for definitive conclusions, while still others are simply useless.

Candidates spend billions of dollars and thousands of hours on teacher preparation requirements, but there's little evidence to justify them.

What's more, even if we knew the right inputs—the right requirements that candidates need to meet to become a teacher—we couldn't be sure they were well implemented at each of the 26,589 teacher preparation programs in 2,171 colleges, universities, and other education providers across the country.¹ That dispersion makes it challenging to create effective policies for teacher preparation.

In response, the latest efforts to improve preparation shift away from input requirements for candidates and their programs and toward each program's *outcomes*. Policies then become less about defining the "how" or "who" of teacher preparation and more about defining the end result—student learning growth, for example, or teacher placement rates. This shift is an improvement, but it is fraught with challenges when it comes to deciding which measures determine success and if those measures can distinguish programs from one another.

This paper attempts to trace what we know about identifying and training successful teachers. It starts with a sobering conclusion: Although candidates spend billions of dollars and thousands of hours on teacher preparation courses, we don't yet have a body of evidence justifying those requirements. Nor do we know how to measure and define a successful teacher training program. There is not yet any magic cocktail of program and candidate requirements that would ensure all teachers are great before they begin teaching. As such, policymakers should invest much more time and resources into learning about the science of teaching and how individual teachers develop their skills. And in the meantime, policies should reflect the fact that we know far more about a teacher after they enter the classroom than beforehand.

We Don't Know How to Train Good Teachers

The best available evidence suggests that we don't know how to train good teachers, no matter what stage they're at in their careers.

As a country, we make an enormous investment in training teachers. We estimate it costs \$24,250 to train the average teacher candidate, and each candidate will spend an average of 1,512 hours in training²—more hours than the typical American teacher works over the course of a school year—to meet the entry requirements to become a teacher.³ Collectively, we estimate that over the course of just one single year, new teachers will invest \$4.85 billion and 302 million hours on their preparation. The real cost is even higher because these estimates don't factor in the opportunity cost. Nor do these estimates consider training delivered to teachers once they begin their careers, which the nonprofit TNTP suggests can amount to another \$8 billion annually.⁴

That's a lot of time and money, and we make these investments year after year for all of the nearly 200,000 teachers certified each year.⁵ This investment *should* be producing effective teachers, but it's not. The best available evidence suggests that we simply don't know how to train good teachers, no matter what stage they're at in their careers.

Historically, the education sector has relied on inputs—things like admissions criteria, required coursework, certification tests, years of teaching, and advanced degrees—to identify good, “highly qualified” teachers. But relying on inputs is risky, unless we're sure that they are true indicators of quality. If not, candidates who may have been successful will be filtered out, and candidates who won't ever be successful are allowed in.

[5] No Guarantees: Is it Possible to Ensure Teachers Are Ready on Day One?

What’s more, the inputs used for teacher training are unproven. Admissions criteria, required coursework, advanced degrees, and certification requirements tell us little about who is likely to be an effective teacher. These are all barriers to entry—put into place by states, preparation programs, accreditors, and the schools hiring teachers. But regardless of who set the requirements, no requirement ensures that only good teachers will reach students. And even after a candidate jumps through these hoops, we don’t know how to provide the professional development and training that will improve her practice. Additional research, particularly into new measures such as teacher performance assessments like the new edTPA, may provide more insight into what works. But the research we have today doesn’t support the current slate of requirements.

It’s a sobering, even humble, conclusion: At every stage of a teacher’s career we simply don’t know how to help her improve.

We don’t know which candidates to admit

Admissions criteria are the first barriers for most teacher candidates. There’s a large national push for states to require high GPAs and SAT scores from students looking to enter teaching programs. The Council for Accreditation of Educator Preparation (CAEP), the accreditation body for educator-preparation programs, requires that providers set minimum GPA and admissions-test requirements for candidates.⁶ CAEP’s minimum admissions scores are designed to rise over time, reflecting a sense that teachers should be culled from the “best and brightest” candidates.

But while there’s some evidence that these screens can predict teacher effectiveness, the differences in effectiveness between candidates who meet the requirements and candidates who don’t are small, and the research is conflicting.⁷ Using data from New York City, researchers found a positive correlation between a teacher’s SAT math score and student achievement, but they found no relationship between student growth and a teacher’s SAT verbal score.⁸ A 2015 study of a large urban district in the South did find positive effects for a teacher’s undergraduate GPA and attendance at a competitive college but concluded that “these characteristics together account for only a small portion—less than 2 percent—of the total variation in teachers’ value-added estimates, meaning that they are correlated with teacher value-added but only weakly.”⁹

While the evidence tilts toward greater teacher effectiveness for candidates with stronger incoming academic credentials (all else being equal), the value of those credentials is relatively small, and they are not a guarantee for any individual. This research also gives states no guidelines on where to set setting appropriate minimal standards that each candidate must meet.

There’s some evidence that GPA and SAT screens can predict teacher effectiveness, but the value of those credentials is relatively small, and they are not a guarantee for any individual.

The “best practices” of teacher preparation often make intuitive sense, but aren’t backed by research.

We don’t know what coursework—if any—to require

After being admitted to a preparation program, most candidates have to complete substantial coursework before they can teach. States require candidates to pass certain courses before they can be certified, but the courses vary from state to state, and the actual content varies widely—even among different course sections within the same program and among separate programs within the same university. There’s little research on how much or what type of coursework would help candidates become effective teachers, so programs generally rely on best practices, as defined by the field. These best practices often make intuitive sense, but they’re not backed by research. It’s logical to assume, for example, that additional coursework and more clinical experience should produce better teachers. But that’s not necessarily the case.

There is no evidence to show that the length of a preparation program is linked to better student outcomes. In a rigorous Institute of Education Sciences (IES) study, researchers from Mathematica Policy Research divided alternatively certified teachers into two subgroups, high coursework teachers and low coursework teachers, according to the number of coursework hours that were required by each teacher’s certification program. The researchers found no evidence that students of high-coursework teachers performed differently than students taught by teachers with less coursework. And neither of these groups performed any differently than traditionally certified teachers.¹⁰ Another study found that Florida teachers who completed an alternative-certification path that required a bachelor’s degree but no additional coursework were more effective than traditionally prepared teachers. The study also found that candidates with higher levels of academic achievement were more likely to pick shorter training programs, which may create a false picture of how well these programs train future teachers.¹¹ Regardless of the explanation, there is no clear correlation between the number of courses a future teacher completes and her eventual success as a teacher.

This finding largely holds true for different types of coursework. The Mathematica researchers also looked at the effects of additional content coursework, such as math or English courses, and additional education-specific coursework, such as classes in classroom management, child development, or student assessment. The researchers found no relationship between additional content or education coursework and student achievement. This isn’t to say that these skills don’t matter, but it does suggest that we don’t yet know how to teach them. Other studies have found similar results. Using data from San Diego Unified School District¹² and Chicago Public Schools,¹³ researchers found no link between a teacher’s undergraduate major or minor and student performance, even if the teacher earned a degree in math or English. There are some exceptions—research shows, for example, that secondary math and science teachers with more content expertise are more effective¹⁴—but overall, coursework requirements are not correlated with student achievement.

One solution to the problem of more coursework not producing better teachers is to change how candidates spend their time. They can spend less time on theory or content learning and more time on the practice of actual teaching skills. But this, too, isn't as simple as it may seem. Merely adding additional clinical fieldwork experiences, or increasing the time a teacher spends in clinical practice during her preparation program, is also not associated with increases in student achievement. From some research, such as a 2008 study of New York City,¹⁵ we know that some student teaching experience is better than none. But more is not always beneficial. In the IES study, the Mathematica researchers found that the number of required fieldwork courses had no effect on student achievement.¹⁶ In another study, researchers analyzed test data from Florida for all students in 3rd through 10th grades between 2000 and 2005. The research showed no relationship between additional classroom-practice requirements and student achievement. There's other research¹⁷ that student teaching programs that are controlled by the preparation program, rather than the K-12 school, produce higher quality experiences and more effective teachers. But again—that evidence is an argument for quality, not quantity.

No one knows exactly *why* these inconsistencies exist. It could be that the training itself is valuable but candidates select programs that better fit their needs and goals. Highly motivated candidates may choose faster and less restrictive alternative-route programs, while other candidates may benefit from the longer training offered by traditional programs. It would be impossible to design a randomized controlled trial to examine this phenomenon. But regardless of the effects of program selection, there are trade-offs between requiring more training and allowing candidates to enter the teaching field with less training. So far, there's no clear indication that more training is always better than less.

State certification is the sum of the required components of teacher preparation. It should predict teacher effectiveness, but it doesn't.

We don't know what the right certification requirements are

Before entering the classroom as teachers, candidates must meet state certification requirements. States offer different types of certification on the basis of the candidate's background and future teaching subject. For an initial teacher's certificate, states generally require candidates to complete an approved teacher preparation program; pass content examinations, such as Praxis Subject Assessments; and undergo a background check. Some states also offer alternative and professional certifications for candidates who are less or more qualified.

State certification is the sum of the required components of teacher preparation: coursework, clinical practice, and assessments. It's reasonable to think that certification may capture something meaningful that predicts teacher effectiveness. But it doesn't. Research shows that certification has no relationship to teaching ability.

Historically, there's been a bias against short alternative-preparation programs. After all, how can teachers be effective without comprehensive training? But teachers who spend extra time and money on traditional teacher preparation programs are not necessarily better teachers than those who start teaching and get training along the way.¹⁸ A 2006 study of 9,400 Los

Angeles teachers found no statistically significant differences among traditionally certified, alternatively certified, and uncertified teachers.¹⁹ In another study, researchers examined the effects of initial teacher-certification status (certified, uncertified, and alternatively certified) on student achievement, using data from New York City. The researchers found that the differences *among* different certification statuses were minimal; average traditionally certified, alternatively certified, and uncertified teachers had similar effects on student achievement. Instead, the researchers found substantial differences *within* certification status. For elementary math teachers, the differences between the best and worst teachers who had taken the same route to certification was almost ten times as large as any differences among teachers who were sorted by certification type.²⁰

In the IES study mentioned above, researchers conducted mini-experiments in which students in a single grade at one school were randomly assigned to either a teacher who was trained in an alternative program or a teacher who was trained in a traditional program. At the end of the trial, there were no statistically significant differences in the students' test-score gains. The researchers concluded that it would be "very difficult to predict, based solely on route of certification, the outcome of students placed with a particular teacher."²¹

We don't know how to help teachers improve once they begin teaching

Once candidates enter the classroom as teachers, we don't know what on-the-job professional development improves their practice.

A substantial body of research shows that teachers' effectiveness tends to improve during their first few years of teaching but then levels off considerably.²² On average, teachers with 20 years of experience are more effective than teachers with no experience, but they're only slightly more effective than teachers with five years of experience.²³ And effectiveness varies within experience levels. There is substantial overlap in the effectiveness of teachers with no experience, some experience, and a lot of experience.²⁴

Once teachers enter the classroom, we don't know what professional development improves their practice.

This evidence aligns with what researchers have concluded about professional development for teachers. Recent research from TNTP found that districts spend about \$18,000 per teacher per year on professional development, an expense that adds up to \$8 billion annually for just the 50 largest school districts in the country. Despite that massive investment, only 30 percent of teachers show noticeable improvement, and TNTP cannot pinpoint the professional development elements that led those teachers to improve. The authors note that "no type, amount or combination of development activities appears to be more likely than any other to help teachers improve substantially, including the 'job-embedded,' 'differentiated' variety that we and many others believed to be promising."²⁵ High-quality, randomized controlled studies conducted by IES have come to similar conclusions about professional development in reading,²⁶ mathematics,²⁷ and science.²⁸

It's depressing but true: We simply don't know how to train teachers to teach, no matter what stage they are at in their careers.

The Look to Outcomes

States, programs, and schools have long focused on inputs, which were thought to predict teacher effectiveness and were often the best option available. But in the mid-2000s, it became possible to evaluate preparation programs through the outcomes of their graduates.²⁹ No longer would policymakers have to impose rules that were essentially best guesses about what would make an effective teacher. Policymakers could measure which teachers were effective and then use the data to shape policy.

Louisiana and Tennessee were the first states to try out this idea. In 2000, Louisiana started looking at preparation programs through their outcomes data. Between 2003 and 2006, the state began linking preparation programs to the student-learning data of their recent graduates, making the data available to the public in 2007.³⁰ Louisiana's work suggested that it was possible to discern program quality from completer outcomes.³¹ Tennessee began a similar initiative in 2007, when the state passed legislation that required an annual report on preparation-program outcomes. Louisiana's and Tennessee's efforts laid the foundation for national interest in linking outcomes to preparation programs.

The U.S. Department of Education officially endorsed linking programs to outcomes in the summer of 2009, when Secretary of Education Arne Duncan and President Barack Obama announced the Race to the Top (RTT) competition. The \$4.35 billion RTT program awarded points to states for a range of accomplishments, including 14 points for improving the effectiveness of teacher and principal preparation programs. Toward the end of 2009, Secretary Duncan gave a forceful speech at Columbia University in which he called out preparation programs for "doing a mediocre job of preparing teachers for the realities

of the 21st century classroom.”³² In his speech, Duncan lauded Louisiana’s work linking completer outcomes to preparation programs.

RTT prompted a number of states to begin linking programs to outcomes. Forty states and Washington, D.C., applied for RTT funding in the first phase alone,³³ and nineteen states received grants throughout the three phases of the competition.³⁴ Not all of those states promised teacher preparation reforms along the lines of Louisiana and Tennessee, but many of them did, including Rhode Island, Florida, Ohio, North Carolina, Massachusetts, and Georgia.

In 2011, the U.S. Department of Education took another step to encourage states to link preparation programs to outcomes. The department announced that it would begin the process of regulating Title II and Title IV of the Higher Education Act ³⁵ to address teacher preparation accountability and reporting. Title II affects how states and institutions report on the quality of preparation programs and requires states to identify their low-performing programs. During the rulemaking process, the department pushed to include completer outcomes in states’ definitions of program quality.

The regulation, which is still making its way toward a final rule, would require states to assess preparation programs on three performance outcomes: student learning (measured by student-growth or teacher evaluation results); employment (placement and retention rates, especially in high-need schools); and survey outcomes (of completers and employers). Although the rule is pending, if the final rule looks like the proposed version, all states will be required to link completer outcomes to preparation programs, beginning in April 2019, and to report the data publicly.

The Promise of Outcomes

Completer outcomes seemed to be the solution to at least a few teacher quality issues. The idea was—and is still—appealing: States could loosen preparation program requirements, give providers more freedom to design their programs as they saw fit, and then make decisions about programs on the basis of the success of their completers. If a program didn’t produce effective teachers, or produced teachers who didn’t get jobs, the state could warn the provider that it needed to fix the problem. If the provider couldn’t fix the problem, the state could close the low-performing program. States could learn from the highest performing programs, and target supports and consequences to gradually shift the teacher preparation market so that it offered only the best providers.

Even if, at first, states can’t define productive interventions, outcomes allow them to differentiate among poor, satisfactory, and excellent alternative and traditional certification programs in a way that inputs do not. Theoretically, the data would reveal which programs produce the best teachers, and “consumers”—prospective teachers and potential employers—could act on that information.

Comparing outcomes would provide real evidence of results, rather than relying on anecdotal and largely unproven “best practices.”

Perhaps the most important short-term use for outcomes is providers' own continuous improvement efforts. Comparing outcomes across similar providers would give providers the evidence they need to change their program practice, rather than rely on anecdotal and largely unproven "best practices." Tracking and publicly reporting outcomes would increase transparency and reduce information gaps so that providers could learn from others in the state.

Bringing Outcomes Back to Reality

For all the reasons mentioned above, linking outcomes to preparation programs was and remains a promising approach to improving educator preparation and the quality of the teaching force. But the initial enthusiasm for this work must be tempered by new research and ongoing implementation challenges.

Recent studies from Missouri and Texas suggest that completer outcomes may not differentiate providers as distinctly as hoped. Researchers attempted, and failed, to use outcomes data to differentiate programs by the effectiveness of their completers. But there was no pattern in the types of teachers these programs produced. The implication can't be overstated: If there are no meaningful differences in teacher effectiveness between programs, we're back at square one. Outcomes work assumes that a teacher's effectiveness is determined in large part by the preparation program she attended, and that preparation programs produce teachers who are at a consistent level of effectiveness. In other words, a great preparation program is one that produces highly effective teachers, a good preparation program produces effective teachers, and so on. But if that's not the case, as the Missouri and Texas research suggests, then policymakers have no reliable information to guide them. States can't reward or learn from great programs—those programs that produce the most effective teachers—if they can't identify great programs. Without meaningful differences in outcomes among programs, school districts and potential candidates (the consumers in the preparation market) have little information about program quality.

The Missouri and Texas findings also conflict with the research from Louisiana that fueled much of the original interest in outcomes. The researchers in Missouri used a different sampling method from the one employed in Louisiana, and they warn in their paper that the Louisiana research may have "overstated differences in teacher performance across preparation programs."³⁶ Although Texas has a wide variety of preparation programs, including some run by school districts that require very little formal training, researchers were unable to find any meaningful differences in outcomes among 100 programs of all types and durations.³⁷ Combined, the Missouri and Texas findings suggest that policymakers should be cautious and temper their expectations for actionable results in their states.

If there are no meaningful differences in teacher effectiveness between programs, then states, policymakers, and consumers have no reliable information to guide them.

Two other studies, out of North Carolina³⁸ and Washington State,³⁹ found some evidence of variation in preparation program outcomes, but suggest that the variations in outcomes within preparation programs are larger than those among them. The North Carolina study did not look at individual preparation programs, but it did compare student achievement across four groups of teachers: (1) traditionally certified teachers, alternatively certified teachers, and Teach For America corps members; (2) traditionally certified teachers from in-state and out-of-state programs; (3) traditionally certified teachers from in-state public and private universities; and (4) teachers who began teaching with undergraduate and graduate degrees. The researchers found small, statistically significant differences among certain groups of teachers for certain grades and subjects, but there was also a large variation in teacher performance within every preparation route, as shown in figure 1. The differences among category averages are distinguishable statistically, but the ranges show the wide disparities among teachers from the same program.⁴⁰ At best, the differences are small. At worst, the overlap among preparation routes makes any average differences meaningless for any individual candidate. This research, combined with the failure to replicate the Louisiana findings, is a warning that outcomes are not yet much better than inputs in identifying effective preparation programs.

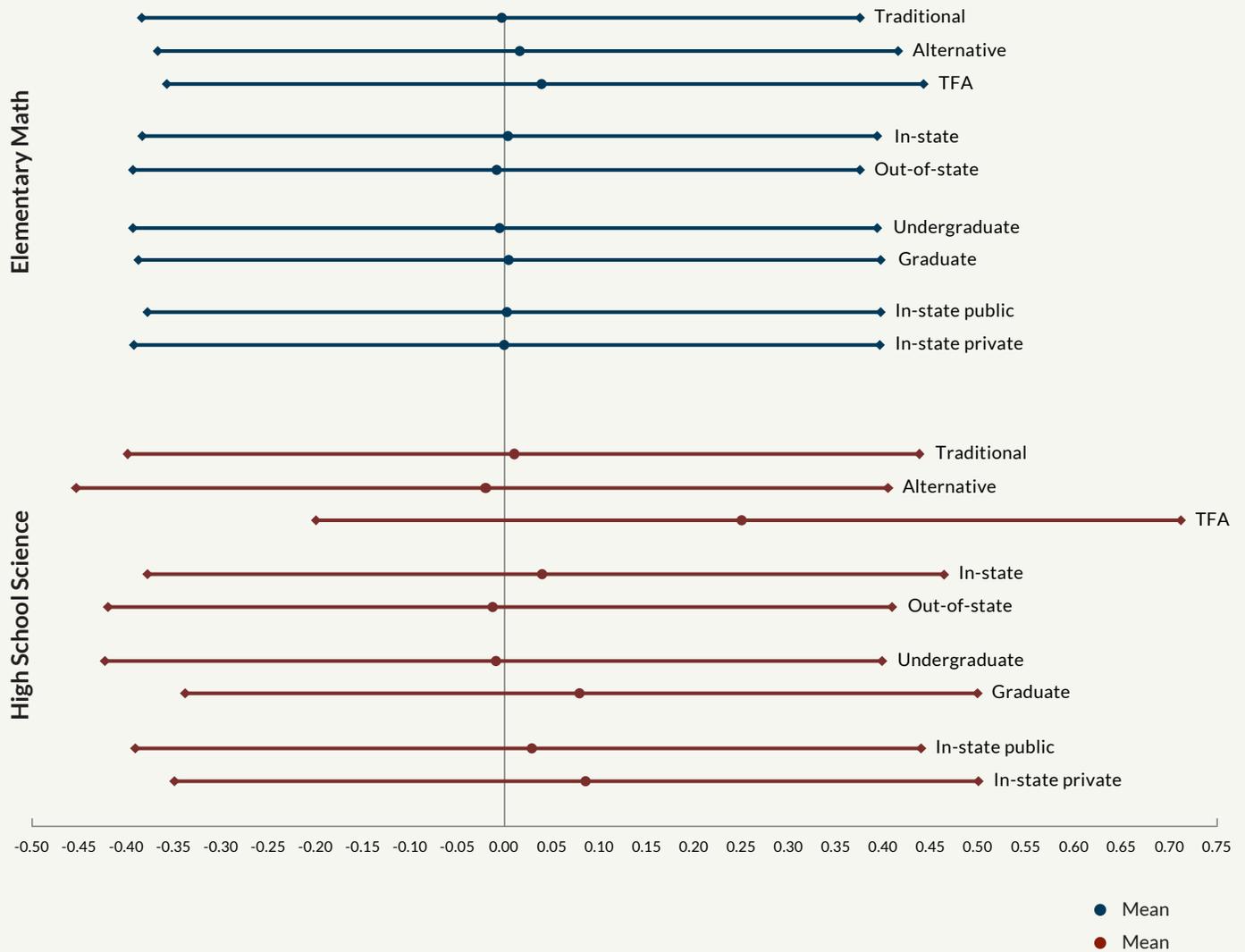
Implementation issues also bog down the outcomes-based approach as states confront difficult questions and trade-offs. To begin with, states may have access to their own data, but they're unable to continue collecting data on teachers who cross state lines. There's also the question of how to interpret results. If a preparation program produces great teachers, does that mean the training program is great, or did the program just do a great job of selecting candidates? The cause may not matter to the school that hires the teacher, but it does matter to the state or preparation program trying to make sense of the results.

States have also struggled with more technical questions. For example, what's the right group size, or "n-size," for reporting or accountability purposes? (In the context of teacher preparation, n-size is the minimum number of completers who can be included in a statistical analysis of program effectiveness.) The larger the n-size, the more confident states can be that the results truly reflect the program and are not just random noise.

From a statistical perspective, a larger n-size is better, but it's not always possible. Many programs and institutions produce a small number of completers every year, and only a fraction of those completers find jobs. Certain types of outcomes may also limit the sample of completers. (Only completers in specific grades and subjects, for example, will have student-learning data.) So setting a higher minimum n-size allows for statistical analysis and protects completer privacy, but it may incentivize providers to produce fewer completers in order to avoid public reporting. States deal with n-size issues by "rolling up" completers across multiple years or similar types of programs, or by linking outcomes to whole institutions rather than to each program. But an n-size that forces providers to roll up completer cohorts may undermine the statistical conclusions and reduce transparency.⁴¹

Figure 1

Distribution of Teacher Effectiveness in Elementary Mathematics and High School Science



Note: Figure depicts the distribution of teacher effectiveness (25th to 75th percentile) for nine teacher categories in elementary mathematics and high school science.

Source: Gary T. Henry et al., "Teacher Preparation Policies and Their Effect on Student Achievement," *Education Finance and Policy* 9 no. 3 (Summer 2014), 264-303.

Implementation challenges are not just limited to n-sizes. Some states have struggled to quickly link data systems, while others lack staff members who can analyze the data. And all the technical and logistical questions are intertwined with complex politics. Because of the high-stakes nature of this work, for example, states convene committees of pre-k through 12th grade, and higher, education stakeholders to decide which outcomes to use and for what purpose. These decisions can quickly become part of political battles over testing and teacher evaluations.

The biggest obstacle to outcomes-based accountability is that states are struggling to find meaningful differences in outcomes across programs for a range of variables.

Perhaps the biggest obstacle to outcomes-based accountability is that states are struggling to find meaningful differences in outcomes among programs for a range of variables, not just student growth. For example, some states would prefer to link preparation programs to the overall evaluation results of their completers, rather than only to test scores. But when those underlying evaluation results fail to show any meaningful differences among teachers, it is impossible to find meaningful differences among teacher preparation programs. If evaluations better assessed teacher quality, then we could more precisely differentiate teachers and, therefore, providers. But outcomes-based accountability relies on the quality of measures that go into it.

All states doing this work planned to use outcomes in their accountability processes, but because of the challenges we've discussed, few actually do. Florida, Massachusetts, and Ohio are the only states we know of that link completer outcomes to program approval, and the degree to which they actually integrate outcomes into their approval processes varies. Florida is the most objective state. It set thresholds that differentiate program performance on four levels, and it developed a formula to roll up five years of outcomes data into one summative rating. On the other end of the spectrum, Ohio allows the chancellor to determine whether and how outcomes are used in the program approval process. Massachusetts is somewhere in the middle; the state created a rubric to guide how reviewers assess programs using input data, completer outcomes, and on-site visits, but the rubric does not set minimum requirements for completer outcomes.⁴²

All other states doing this work only use outcomes for continuous improvement and public reporting. Some states made plans to include outcomes in their accountability processes—Rhode Island, for example, plans to include outcomes but has not finalized its rules for how it will do so. And Delaware plans to attach consequences to its completers' outcomes in 2016.⁴³ But it's telling that even the pioneers in this work, Louisiana and Tennessee, are still figuring out how to best use outcomes for accountability.

Without the ability to observe and act on differences across programs, states must bet that transparency alone will shape supply and demand in the market for new teachers. In an ideal world, potential educators would use data to inform their enrollment decisions, and school districts would integrate outcomes data into their hiring practices. But these goals depend on reliable, easy-to-understand data. So far, states are struggling to make the data meaningful to the public. When it comes to making decisions about what data to report or how to report them, states often prioritize making data accessible to preparation programs over other potential purposes. Preparation programs should, indeed, be one of the main users of the data, but the idea that these programs will use the data to voluntarily improve their training efforts does not reflect the full theory of action behind the push for outcomes-based accountability.

Following the Existing Research

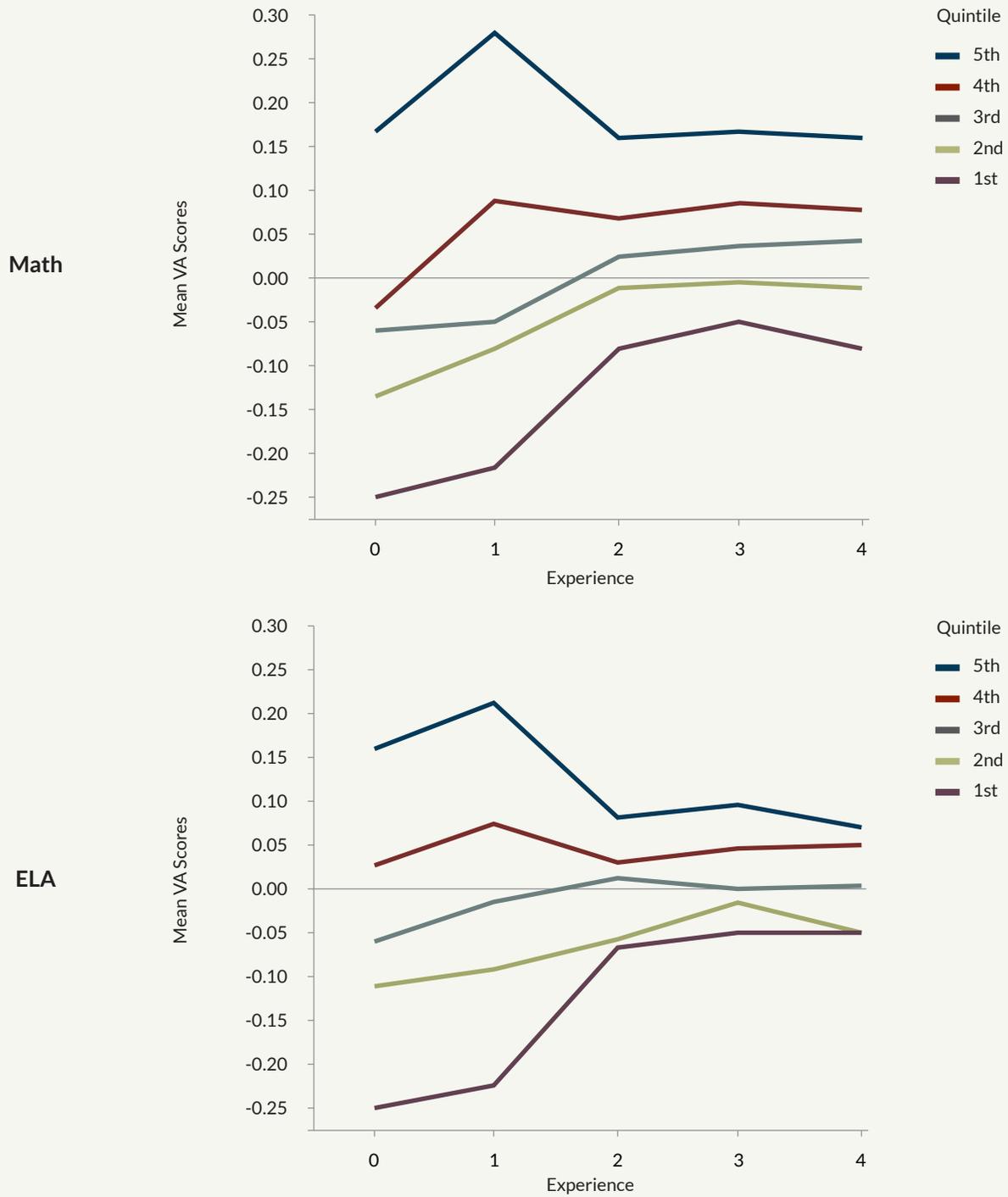
All of this means that policymakers are still looking for the right way to identify effective teacher preparation and predict who will be an effective teacher. Nothing tried so far guarantees effective teachers. Yet there are breadcrumbs that could lead to a better approach.

One starting point would be to embrace the research on initial teacher effectiveness. In one Los Angeles study, researchers found that a teacher's initial classroom effectiveness better predicted her future classroom effectiveness than anything schools and districts used at the time. The study divided teachers into quartiles on the basis of effectiveness and tracked their students' progress over three years. Students assigned to teachers in the top quartile outpaced peers with similar baseline scores and demographics, while students taught by bottom-quartile teachers fell behind their peers. It's commonly noted that teachers improve with experience, and that was also true here. But even though all teachers improve *on average*, teachers tend to stay as relatively effective or ineffective as they were when they started teaching.⁴⁴

A 2013 study of New York City teachers found similar results. On average, teachers continued to perform at the same relative performance level, and, while the lowest-performing teachers did improve over time, they never caught up to their more effective peers, even after several years of teaching. Worse, while all teachers tended to improve, the teachers who were the lowest performers from the outset never caught up, even to the average *new* teacher. Figure 2 illustrates these findings.

Figure 2

Mean Value-Added Scores by Subject, Quintile of Initial Performance, and Years of Experience



Note: Mean value-added (VA) scores, by subject (math or English language arts), quintile of initial performance, and years of experience for elementary school teachers with VA scores in at least first 5 years of teaching. Each year follows the same sample of teachers. Sample includes elementary and high school teachers with VA scores in at least their first five years of teaching.

Source: Allison Atteberry, Susanna Loeb, and James Wyckoff, "Do First Impressions Matter? Predicting Early Career Teacher Effectiveness," *AERA Open* 1 no. 4 (October-December 2015), 1-23.

Policymakers should be less focused on defining the perfect mix of requirements for incoming teachers and more focused on measuring and acting on teacher effectiveness in a teacher's first years on the job.

The results from a teacher's first few years of teaching are much more informative than anything we know about a teacher before she begins teaching. Taken together, the type of preparation program teachers attend, their credentialing and SAT scores, the competitiveness of their undergraduate institution, their race and ethnicity, and their gender explain less than 3 percent of the variation in classroom effectiveness. In contrast, a teacher's initial value-added results, measured in her first years on the job, explain 21 percent of her future effectiveness in reading and 28 percent in math. In other words, with information collected in just a teacher's first year on the job, districts know 7-9 times more about their employee than they did before the year began.

The current research suggests that a teacher's preparation is a relatively small factor in her eventual effectiveness. It's possible that future research will challenge these findings. But if we take the current findings to their logical extent, they suggest policymakers should be less focused on defining the perfect mix of requirements for incoming teachers and more focused on measuring and acting on teacher effectiveness in a teacher's first years on the job.

To accomplish this goal, policymakers need to shift their priorities. They should roll back burdensome and ineffective teaching requirements, rethink licensure, create systems to make preparation-pathway data accessible to the public, and create the conditions for alternative pathways to teaching. We propose four strategies for ensuring that schools and students have access to the best teachers possible:

1. Make it less risky to try teaching

In the current preparation system, becoming a teacher is laden with risk. States force candidates to spend thousands of dollars and thousands of hours on fulfilling requirements that won't make the candidates more effective teachers. When a candidate finally begins teaching, she'll waste more time on questionable professional development. If she wants a salary increase, she'll pay for, and sit through, graduate coursework to earn an advanced degree that won't likely help her improve her practice. And these are just the risks associated with teacher training—there is also the challenging and high-stakes work, the generally low pay, and the lack of advancement opportunities that come with teaching itself. The result is a profession in which potential candidates' actual and opportunity costs are extremely high.

There are a number of ways to reduce the risks associated with teaching and make the profession more appealing. Reducing barriers to entry is one place to start.

States should make it relatively easy for teacher candidates to enter the field. This may sound counterintuitive, but current barriers to entry do not guarantee quality teachers, and they keep out candidates who could be effective. States should stop writing the rules, down to specific minimum test scores and course requirements, for who gets to be a teacher. To impose a minimum level of quality, states should require candidates

to possess a bachelor's degree from an accredited college or university. But with the possible exception of secondary math and science teachers—for whom research has found a definitive link between training and outcomes—states should not impose additional academic requirements on candidates.

Rather than layering on requirements, states should get out of the way and rely on local school districts to hire the best candidate for the job. School districts should have the flexibility to screen candidates for majors, such as science or English, for certain teaching positions. While school districts face their own challenges in hiring and evaluating teachers, it's not the state's role to determine the requirements teachers must meet, certainly not when the evidence is inconclusive on the merits of those requirements.

This would not eliminate the demand for existing teacher preparation programs. Candidates would be free to enroll in these programs, particularly if the programs could sell themselves as adding value. Districts may wish to continue hiring from these programs. But there's no reason that states should require this monopoly on teacher training to continue. In opening the door to the profession in other ways, states should make a distinction between allowing someone to try teaching and allowing her to become a licensed teacher with full responsibility for student learning. See recommendation 2 for more on this distinction.

Objections to this approach

Untrained teachers shouldn't be allowed to teach

This argument is an emotional appeal that hinges on the assumption that today's barriers to entry prevent unqualified people from teaching our children. But in reality, today's barriers are little more than veils, not guarantees of quality. We don't know the right combination of inputs for training effective teachers, and it's impossible to guarantee that all new teachers will be good teachers, regardless of the type of training they receive.

This objection also assumes that with the right training, an incoming teacher should be given the same responsibility as veterans with many more years of experience. That approach, used in most schools across the country, does not align with how teachers learn on the job. Rather than trying to fit new teachers into a system they're not prepared for, states should change an incoming teacher's responsibilities so she does not have direct and sole responsibility for educating children until she has a record of effectiveness. The current system blindly assumes that all incoming teachers have that track record, but they don't.

Allowing anyone to become a teacher will lower the status of the teaching profession

Let us be the first to admit it: There could be ramifications if we lower the barriers to teaching. If those barriers shape how society values the teaching profession—if the barriers give teachers status—society's perception of teachers may change.

But this is a slippery objection, because there's no objective measure of status or what defines it. It's also unclear which factors are the most important determinants of status—training requirements, salaries, the responsibilities of the job?

For certain people, status is only conferred on professions like law or medicine that have strong barriers to entry. This camp tends to ignore the fact that law and medicine are much smaller professions with much higher salaries. Teaching is by far the largest profession made up of college-educated individuals, and any questions about barriers to entry must take into account the fact that schools simply need to hire lots of people. And states choosing to layer on additional requirements may drive away talented applicants who might otherwise be willing to try teaching. Ultimately, a profession's status is marked by the quality of its practitioners, not just how they came to the field.

To be clear, there are trade-offs inherent in this question. But status is not reason enough to maintain burdensome barriers to entry, particularly if those barriers are meaningless.

2. Give schools and districts, not preparation programs, responsibility for recommending a candidate for licensure, and require that recommendation to be based on a track record of effectiveness

In the current system, once a candidate meets the state requirements, her teacher preparation program recommends her for licensure. This is a flawed arrangement. Most preparation programs make recommendations on the basis of the completer's academic performance and a limited amount of (perhaps undersupervised) student teaching experience. But, as noted above, there's no guarantee that these experiences create a teacher who is prepared to be effective on Day One. Moreover, the current system encourages schools to treat all licensed teachers as interchangeable once they enter the classroom, with identical workloads, evaluation systems, and development opportunities.

A better system would base licensure on actual candidate performance. To that end, states should strip the power to grant teacher licenses from preparation programs and give that responsibility to the districts where candidates teach. K-12 school leaders are the closest observers of actual candidate performance—much closer observers than preparation programs are—and they have access to data that preparation programs don't, including student-level learning results. Instead of relying on a checklist of meaningless inputs, states should allow individual candidates to apply for full-time licensure only after they have a track record of performance. States should set their own rules for what a track record of performance looks like, but at the very least, licensure decisions should require observations of candidates' performance in real-time classroom settings and demonstrated effectiveness in supporting students' academic growth.

In this scenario, candidates would enter the classroom in a relatively low-stakes environment. Unlike today's typical teacher, who, for her first full-time teaching job, is thrust in front of a full class of students at the beginning of the school year, a candidate should have opportunities for lower-stakes interactions with students, such as summer school and after-school and tutoring programs. Over time, the candidate would play a substantial role in the classroom, to the school's benefit and her own. There are various models districts could follow, but as far as the state is concerned, a candidate should become an official "teacher of record" only after she has demonstrated her effectiveness and earned her license. This puts the focus on each teacher's accomplishments rather than on her credentials or her preparation pathway. Programs such as the Teacher Advancement Program,⁴⁵ the Opportunity Culture initiative,⁴⁶ and Ohio's Resident Educator Program⁴⁷ focus on teacher accomplishments. TNTP started down a similar path almost four years ago.⁴⁸

Withholding licensure has three effects: It ensures that only the most effective candidates earn full teacher licenses, it increases the prestige associated with becoming a fully licensed teacher, and it encourages ineffective candidates to self-select out of the profession. To be clear, this process does not rely on districts firing ineffective candidates. Recent research from New York City suggests that merely identifying low-performing teachers and delaying decisions on their tenure can encourage weaker teachers to leave the profession.⁴⁹ Districts should determine their own processes for making decisions—for example, candidates who do not earn their licenses could stay on as aides, small-group instructors, or co-teachers. But it is crucial that, as we lower the barriers to teaching, we raise the status and the standards for becoming a fully licensed teacher.

Objections to this approach

There's no objective way of knowing who's a great teacher

This objection suggests it's simply impossible to quantify a teacher's contributions to the classroom, and so it would be impossible for teachers to ever demonstrate a record of effectiveness. We think this line of reasoning is wrong—high-quality observations, measures of student growth, and other indicators can be combined into a reliable estimate of teacher quality. It's also easy to turn this objection back on the speaker. If it's impossible to know who is a good teacher, it's also impossible to put in place a training program that will ensure that candidates are effective. After all, how would we ever know that the program is working?

This objection has implications for the entire teaching profession, not just for teacher preparation. If we need better measures of teacher quality, those measures are even more important for current teachers' concerns—evaluation, compensation, professional development, retention decisions—than for teacher preparation.

Schools would need to change before preparation can change

It's tempting to suggest that everything about how schools are run must change in order to change the way teachers are prepared. But while it would behoove states and schools to do business in a new way, it isn't required.

In fact, many places are already working toward our vision of a diverse network of preparation providers. Twenty-five years ago, there were no alternative-route preparation programs, and all teachers were trained in traditional, university-based programs. Today, one-quarter of new teachers arrive at the profession through alternative routes, and the largest teacher preparation program in the country is Teach For America.⁵⁰ TFA provides incoming candidates with just six weeks of training, yet its teachers are no less effective than veteran teachers, and they significantly outperform other new teachers.⁵¹ In Florida, teachers entering the profession through a self-paced, online learning program that can be completed in a few months outperform other new teachers in the state.⁵² And some districts and charter-management organizations already operate their own training programs. Without changing any other policies, states should learn from these examples and slowly offer new routes to teaching. If the outcomes from these new approaches look similar to today's outcomes, states could feel confident about expanding the new programs.

3. Measure and publicize results

There's a large market for new teachers—school districts hire between 75,000 and 90,000 newly trained teachers a year—but that market has been plagued by a lack of reliable information. Candidates have little information about which teacher preparation programs will most likely lead to a job or help them become effective teachers once they enter the profession. When hiring a new teacher, districts typically look at only the candidate's GPA, alma mater, and letters of recommendation. Some districts also apply their own screening tools, and a few districts ask candidates to teach a sample lesson.

Districts are partly to blame for this market failure. All districts could ask for more information from candidates or require them to demonstrate their teaching skills. And even in the current environment, there's nothing stopping districts from hiring teachers before the school year begins, such as for summer school, so they can start off in a low-stakes environment. Districts could also do a much better job of training new teachers and ensuring that they receive professional development tailored to their needs.

But states are mostly to blame for the lack of reliable information about the performance of preparation programs. States are the only entities that could have enough data to objectively assess candidate performance, placement, and retention. Candidates will

never have this information unless the states collect and provide it. Districts will never see beyond their own hiring practices unless their state collects information from all schools and aggregates the results.

In the world we envision, states would do a much better job of collecting and reporting on this information. They would collect and publish program-level data on teacher effectiveness, retention, placement, and years to licensure. And they would invest substantial time and effort in making the data accessible to the public.

Here, the word “program” is not limited to a traditional preparation program at a college or university. It includes other providers in the education field, districts that offer their own training, and even “no program” candidates who enter the classroom as teachers without attending any program at all.

By opening the door to new providers, states would be introducing new competition to an old market. Existing providers would be required to compete with alternative programs to show that the time and expense they require of candidates are worthwhile. With new data tracking outcomes, candidates would be equipped with reliable information about their options and would be free to choose their own pathways. Districts would not be required to select candidates from less intensive preparation programs; they’d be free to choose their own candidates, just as they do now. The majority of candidates could continue going to existing preparation programs, but those programs would have to compete for market share, rather than corner the market by default.

In the short term, there’s reason to be cautious about what these outcomes can tell us. But in the long term, collecting and reporting outcomes information is the only way teacher preparation can ever improve. Without a systematic way to track outcomes, we cannot know what makes an effective teacher and what does and does not matter for kids. If states track outcomes, they have a path forward. If they don’t, they’re left blindly wandering from input to input.

Objections to this approach

If completer outcomes data aren’t meaningful, it’s not worth the time and effort to track and publish them

There’s been very little evidence that preparation programs can be distinguished from one another on the basis of completer outcomes for things like student-growth scores or teacher-evaluation ratings. This objection assumes that, given the lack of actionable information, there’s no value in measuring and publishing completer outcomes.

While it's true that the data may not yet be meaningful from a statistical perspective, collecting, analyzing, and publishing outcomes data are the only ways preparation programs can ever know if they're improving. Measuring and publishing completer outcomes data bolsters programs' continuous improvement efforts, giving them deeper insights into the information they already have. Working with the data also builds technical capacity and allows researchers to study the policies of teacher preparation programs and gauge their effectiveness over time. And in the absence of rigorous state accountability systems, public completer outcomes data give potential candidates and employers useful information that they can use to choose programs and hire teachers.

Despite the evidence from Missouri and Texas, it's possible that research in other states will find meaningful differences in program quality. That hope doesn't justify blindly crafting an accountability system around outcomes, but it is enough reason to require all states to measure and publish completer data. We also remain open to the possibility that future assessments, whether the new Common Core-aligned assessments or another iteration, will reveal more variability among teachers and, consequently, preparation programs.

Finally, although some outcomes data may not give states the information they need to change their path to the teaching profession, the data may provide useful information to schools and future teachers. For example, prospective teachers today have little information about which programs are most successful at placing teachers in full-time teaching positions and which preparation programs graduate teachers who stay in the profession. Research from Washington State found that teachers from some preparation programs had retention rates equal to four to five percentage points higher than their peers. This sort of information may be useful to both prospective teachers and their future employers.⁵³

4. Unpack the black box of good teaching

We don't know how to identify or train good teachers. To make matters worse, as a field we tend to recklessly embrace faddish "best practices," whether or not there's research to back them up. When a new idea comes along, it's reasonable to first try it on a small scale, measure the results, and then scale it up only if the research says it's effective. But that's not what states have done with teacher preparation. A number of ineffective requirements—from higher minimum GPAs to more clinical coursework hours to better teacher-performance assessments—started off as good ideas that states and programs codified as policy before the ideas were sufficiently tested. Instead, states, the federal government, and private philanthropy organizations should invest strategically in research on what makes a good teacher and use that research to make policy.

We see several possibilities for research. One is analyzing existing measures and inputs to see whether they matter for a wider range of outcomes. Most of the studies cited in this paper look at whether various inputs lead to higher student test scores. But these analyses use results from state tests that measure relatively low-level skills. By using new assessments that measure higher-order thinking skills, it's possible we will gain new insights into teacher preparation programs. The same might be true if researchers look at a wider range of outcomes, rather than just test scores.

To learn more about teacher training, we also need to encourage more variability in the programs themselves. Rather than try to standardize teacher training before we know what works, we need to do the opposite: more experimentation and more research about how the unique features of programs affect outcomes. Researchers should try to penetrate the black box of what makes a great teacher and how districts can select or train for those characteristics. A study of Teach For America, for example, found that its screening mechanism had some predictive power, particularly in terms of how much “grit” or “stick-to-itiveness” the candidates demonstrated.⁵⁴ If the vast majority of differences manifest *within* individual preparation programs, rather than across them, it would be instructive to understand why a single program produces teachers with very different levels of effectiveness.

Objections to this approach

Research is slow, and it doesn't trickle down to schools

One criticism of research is that it's too expensive and time consuming, and that policymakers don't use it when making decisions. The current teacher preparation requirements justify this objection: States have implemented a wide range of requirements for candidates and preparation programs, despite the lack of research linking those requirements with teacher effectiveness.

We are also concerned about the expense, timeliness, and usability of research. But without more and better research, policymakers will never have good information on what levers they can pull to improve teacher performance. After all, the money spent on teachers is the largest expenditure in education, and for schools, teachers are the most important, controllable factor in shaping student growth. In the last year alone, schools around the country spent \$320 billion on teachers.⁵⁵ It makes sense to invest in maximizing that investment and finding ways to help educators thrive.

Embracing This Agenda

Current methods of teacher training rely on imposing barriers to the profession, but there's little evidence that the system is worth the investment of time or money. Policymakers have pushed for these barriers to entry in the mistaken belief that the status of a professional is determined by the length of her training. They've come up with new ways to make it more difficult and more expensive to become a teacher. Our proposal would stop this downward spiral.

That's not to say that ours is the only solution. In fact, we will happily be proven wrong. If research finds that new, different assessments could predict a teacher's future effectiveness, for example, we would gladly endorse it as a method for screening candidates. If the Common Core-aligned assessments uncover consistent variations among preparation programs, it will be easier to know how to improve teacher preparation pathways. But that's not what's happening right now.

As for the objections to the vision we've outlined here, we acknowledge that these obstacles are real. They are not prohibitive, however, and not an excuse to be content with the status quo.

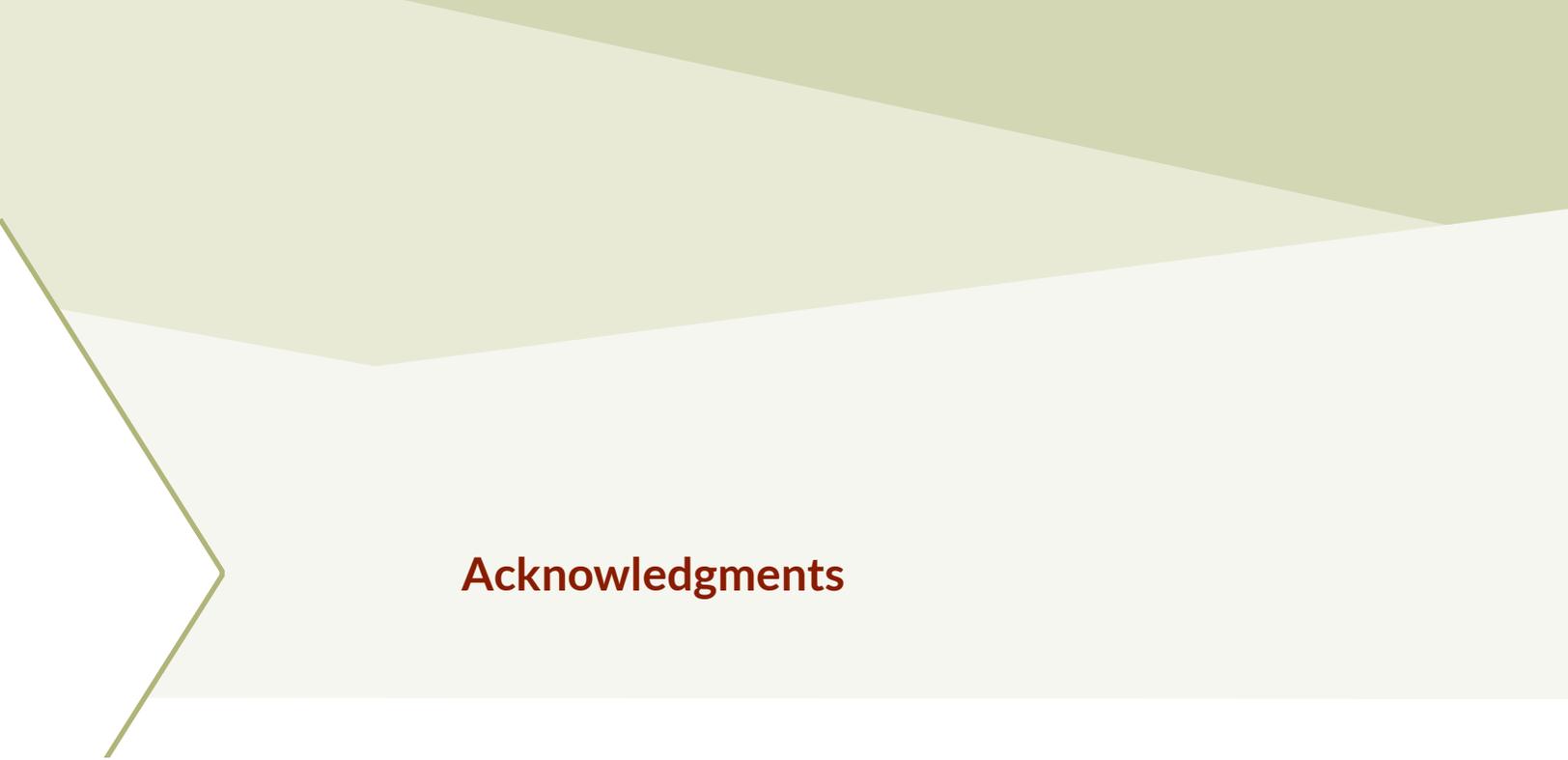
Endnotes

- 1 These are the latest numbers, as of November 5, 2015, from <https://title2.ed.gov>.
- 2 OECD, "Indicator D4: How Much Time Do Teachers Spend Teaching?" in *Education at a Glance 2014: OECD Indicators*, accessed November 1, 2015, <http://www.oecd.org/edu/EAG2014-Indicator%20D4%20%28eng%29.pdf>.
- 3 Total hours of teacher preparation are the sum of coursework hours, student-teaching hours, and testing hours. Total cost of teacher preparation is the sum of coursework costs, testing fees, and certification fees. Students aren't paying the full amount because the cost of attendance is defrayed by state and federal financial aid, but students, the federal government, states, and institutions of higher education combined are spending at least this amount. A complete analysis, with sources, is available upon request.
- 4 TNTP, *The Mirage: Confronting the Hard Truth About Our Quest for Teacher Development*, accessed October 1, 2015, http://tntp.org/assets/documents/TNTP-Mirage_2015.pdf.
- 5 "2014 Title II Reports: National Teacher Preparation Data At-a-Glance," US Department of Education, accessed July 30, 2015, <https://title2.ed.gov/Public/Home.aspx>.
- 6 Scott Jackson Dantley and Emerson J. Elliott, Council for the Accreditation of Educator Preparation, "CAEP Standard 3: Candidate quality, recruitment, and selectivity," accessed July 10, 2015, http://cms.bsu.edu/-/media/WWW/DepartmentalContent/Teachers/UnitAssess/PDFs/CAEP%20Resource%20files/CAEP_PowerPoint_Standard_3.pdf
- 7 Douglas N. Harris and Tim R. Sass, "Teacher Training, Teacher Quality and Student Achievement," *Journal of Public Economics* 95, nos. 7–8 (August 2011): 798–812; Matthew M. Chingos and Paul E. Peterson, "It's Easier to Pick a Good Teacher than to Train One: Familiar and New Results on the Correlates of Teacher Effectiveness," (paper prepared for a symposium, Harvard Kennedy School, 2010), http://www.hks.harvard.edu/pepg/PDF/Papers/2010-22_PEPG_Chingos_Peterson.pdf; Dan Goldhaber, "Everyone's Doing It, But What Does Teacher Testing Tell Us About Teacher Effectiveness?" (paper, University of Washington and the Urban Institute, 2006), http://public.econ.duke.edu/~staff/wrkshop_papers/2006-07Papers/Goldhaber.pdf; Jill Constantine et al., *An Evaluation of Teachers Trained Through Different Routes to Certification: Final Report*, accessed July 30, 2015, <http://ies.ed.gov/ncee/pubs/20094043/pdf/20094043.pdf>.
- 8 Daniel Boyd, "The Narrowing Gap in New York City Teacher Qualifications and Its Implications for Student Achievement in High-Poverty Schools," *Journal of Policy Analysis and Management* 27, no. 4 (2008): 793–818. See table 4. Both verbal and math SAT effects are statistically significant.
- 9 Jennifer L. Steele et al., "The Distribution and Mobility of Effective Teachers: Evidence from a Large, Urban School District," *Economics of Education Review* 48 (October 2015): 86–101.
- 10 Jill Constantine et al., *An Evaluation of Teachers Trained Through Different Routes to Certification: Final Report*, accessed July 30, 2015, <http://ies.ed.gov/ncee/pubs/20094043/pdf/20094043.pdf>.
- 11 Tim R. Sass, "Licensure and Worker Quality: A Comparison of Alternative Routes to Teaching" (paper, Georgia State University, 2014), <http://www2.gsu.edu/~tsass/pdfs/Alternative%20Certification%20and%20Teacher%20Quality%2013.pdf>.
- 12 Julian R. Betts, Andrew C. Zau, and Lorien A. Rice, *Determinants of Student Achievement: New Evidence from San Diego* (San Francisco: Public Policy Institute of California, 2003), http://www.ppic.org/content/pubs/report/R_803JBR.pdf.
- 13 Daniel Aaronson, Lisa Barrow, and William Sander, "Teachers and Student Achievement in the Chicago Public High Schools," *Journal of Labor Economics* 25, no. 1 (2007): 95–135, <http://faculty.smu.edu/millimet/classes/eco7321/papers/aaronson%20et%20al.pdf>.
- 14 Robert Floden and Marco Meniketti, "Research on the Effects of Coursework in the Arts and Sciences and in the Foundations of Education," in *Studying Teacher Education: The Report of the AERA Panel on Research and Teacher Education* (Mahwah, NJ: American Educational Research Association, 2006), 261–308; Douglas N. Harris and Tim R. Sass, "Teacher Training, Teacher Quality and Student Achievement," *Journal of Public Economics* 95, nos. 7–8 (August 2011): 798–812.
- 15 Daniel Boyd et al., *Teacher Preparation and Student Achievement*, accessed July 30, 2015, <http://www.nber.org/papers/w14314>.

- 16 Jill Constantine et al., *An Evaluation of Teachers Trained Through Different Routes to Certification: Final Report*, accessed July 30, 2015, <http://ies.ed.gov/ncee/pubs/20094043/pdf/20094043.pdf>.
- 17 Daniel Boyd et al., *Teacher Preparation and Student Achievement*, accessed July 30, 2015, <http://www.nber.org/papers/w14314>.
- 18 Thomas J. Kane, Jonah E. Rockoff, and Douglas O. Staiger, "Photo Finish: Certification Doesn't Guarantee a Winner," *Education Next* 7, no.2 (Winter 2007): 60–67; Donald Boyd et al., "Teacher Preparation and Student Achievement" (working paper 14314, National Bureau of Economic Research, 2008), <http://www.nber.org/papers/w14314.pdf>; Thomas J. Kane, Jonah E. Rockoff, Douglas A. Staiger, "What Does Certification Tell Us About Teacher Effectiveness? Evidence from New York City," <https://www0.gsb.columbia.edu/faculty/jrockoff/certification-final.pdf>; Donald Boyd et al., "Surveying the Landscape of Teacher Education in New York City: Constrained Variation and the Challenge of Innovation," *Educational Evaluation and Policy Analysis* 30, no. 4 (2008): 319–43; Brian Jacob, "The Challenges of Staffing Urban Schools with Effective Teachers," *Future Child* 17, no. 1 (Spring 2007), 129–53.
- 19 Robert Gordon, Thomas J. Kane, and Douglas O. Staiger, "Identifying Effective Teachers Using Performance on the Job" (discussion paper 2006-01, The Brookings Institution, 2006), http://www.brookings.edu/views/papers/200604hamilton_1.pdf.
- 20 Thomas J. Kane, Jonah E. Rockoff, and Douglas O. Staiger, "What Does Certification Tell Us About Teacher Effectiveness? Evidence from New York City," *Economics of Education Review* 27, no. 6 (December 2008): 615–31.
- 21 Constantine et al., *Evaluation of Teachers* (Washington, DC).
- 22 See, for example, Charles T. Clotfelter, Helen F. Ladd, and Jacob L. Vigdor, "How and Why Do Teacher Credentials Matter for Student Achievement?" (CALDER Working Paper 2, The Urban Institute, 2007); Helen F. Ladd, "Value-Added Modeling of Teacher Credentials: Policy Implications," (paper presented at the 2nd Annual CALDER Research Conference, Washington, DC, October 4, 2008); Tim R. Sass, "The Determinants of Student Achievement: Different Estimates for Different Measures" (paper presented at the 1st Annual CALDER Research Conference, Washington, DC, October 4, 2007).
- 23 Ladd, "Value-Added Modeling of Teacher Credentials."
- 24 Jennifer King Rice, "Learning from Experience? Evidence on the Impact and Distribution of Teacher Experience and the Implications for Teacher Policy," *Education Finance and Policy* 8, no. 3 (Summer 2013): 332–348.
- 25 TNTP, *The Mirage*.
- 26 Institute of Education Sciences and National Center for Education Evaluation and Regional Assistance, *Impact of Two Professional Development Interventions on Early Reading Instruction and Achievement: Executive Summary*, accessed July 30, 2015, http://ies.ed.gov/ncee/pubs/20084030/summ_e.asp.
- 27 Institute of Education Sciences and National Center for Education Evaluation and Regional Assistance, *Middle School Mathematics Professional Development Impact Study Findings After the Second Year of Implementation: Executive Summary*, accessed July 30, 2015, <http://ies.ed.gov/ncee/pubs/20114024/index.asp>.
- 28 Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, and US Department of Education, *Effects of Making Sense of SCIENCE™ Professional Development on the Achievement of Middle School Students, Including English Language Learners*, accessed July 30, 2015, http://ies.ed.gov/ncee/edlabs/regions/west/pdf/REL_20124002.pdf.
- 29 For a review of the evolution of teacher quality reforms, see http://bellwethereducation.org/sites/default/files/JOYCE_Teacher%20Effectiveness_web.pdf.
- 30 Dr. Jeanne Burns, interview with the author, July 29, 2015.
- 31 George H. Noell, Bethany A. Porter, and R. Maria Patt, *Value Added Assessment of Teacher Preparation in Louisiana: 2004-2006*, accessed June 10, 2015, <http://www.regents.la.gov/assets/docs/2013/09/VAATPPTechnicalReport10-24-2007-Yr4.pdf>.
- 32 Arne Duncan, "Teacher Preparation: Reforming the Uncertain Profession" (speech, Teachers College, Columbia University, New York, NY, October 22, 2009), <http://www.ed.gov/news/speeches/teacher-preparation-reforming-uncertain-profession>.

- 33 "40 States, D.C., Submit Applications in Phase 1 Race to the Top Competition," US Department of Education, last modified January 19, 2010, <http://www.ed.gov/news/press-releases/40-states-dc-submit-applications-phase-1-race-top-competition>.
- 34 "Awards: Scopes of Work Decision Letters," US Department of Education, last modified March 18, 2013, <http://www2.ed.gov/programs/racetothetop/awards.html>.
- 35 "Department of Education; 34 CFR Chapter VI; Negotiated Rulemaking Committee, Negotiator Nominations and Schedule of Committee Meetings—Teacher Preparation and TEACH Grant Programs," 76 Federal Register 207 (26 October 2011), pp. 66248-66249, <http://www.gpo.gov/fdsys/pkg/FR-2011-10-26/pdf/2011-27719.pdf>.
- 36 Cory Koedel et al., "Teacher Preparation Programs and Teacher Quality: Are There Real Differences Across Programs?" (working paper, University of Missouri, 2012), https://economics.missouri.edu/working-papers/2012/wp1204_koedel_et_al.pdf.
- 37 Paul T. Von Hippel et al., "Teacher Quality Differences Between Teacher Preparation Programs: How Big? How Reliable? Which Programs Are Different?," *Social Science Research Network* (October 7, 2014), http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2506935.
- 38 Charles T. Clotfelter, Helen F. Ladd, and Jacob L. Vigdor, "How and Why Do Teacher Credentials Matter for Student Achievement?" (CALDER Working Paper 2, The Urban Institute, 2007).
- 39 Dan Goldhaber and Stephanie Liddle, *The Gateway to the Profession: Assessing Teacher Preparation Programs Based on Student Achievement*, accessed July 1, 2015, [http://www.cedr.us/papers/working/CEDR%20WP%202011-2%20Teacher%20Training%20\(9-26\).pdf](http://www.cedr.us/papers/working/CEDR%20WP%202011-2%20Teacher%20Training%20(9-26).pdf).
- 40 Gary T. Henry et al., "Teacher Preparation Policies and Their Effect on Student Achievement," *Education Finance and Policy* 9 no. 3 (Summer 2014), 264-303.
- 41 Other examples of state trade-offs, and a discussion of the decisions select states have made, is forthcoming.
- 42 Massachusetts Department of Elementary & Secondary Education, *Review Evaluation Tool*, accessed November 1, 2016, <http://www.doe.mass.edu/edprep/evaltool/Overview.pdf>.
- 43 Avi Wolfman-Arent, "Grading the Graders: Delaware Releases New Scorecards on Teacher Prep Programs," *NewsWorks*, October 23, 2015, <http://www.newsworks.org/index.php/local/delaware/87528-grading-the-graders-delaware-releases-new-scorecards-on-teacher-prep-programs>.
- 44 Gordon, Kane, and Staiger, "Identifying Effective Teachers."
- 45 The Teacher Advancement Program (TAP), created in 1999, is an initiative of the Milken Family Foundation. TAP is a whole-school reform whose goal is to recruit, motivate, develop, and retain high-quality teachers using four strategies: offering multiple career paths for teachers, ongoing applied professional development, instructionally focused accountability, and performance-based compensation. More information is available at http://www.infoagepub.com/products/downloads/tap_overview.pdf.
- 46 The Opportunity Culture initiatives were created by Public Impact, a North Carolina-based education consulting firm. The goal of an Opportunity Culture is to extend the reach of excellent teachers to more students, for more pay, within recurring budgets. In an Opportunity Culture, the reach of excellent teachers is extended through strategic teaching structures; teachers receive higher salaries from existing budgets rather than from temporary grants, and educators are held accountable for the performance of the students within their reach. More information is available at <http://opportunityculture.org/opportunity-culture/>.
- 47 Ohio began its Resident Educator program in 2011. The program is a four-year induction system for new teachers, at the end of which teachers are eligible to apply for a five-year professional-educator license. Only teachers who complete the residency program are eligible for this license. During their residency, new teachers are paired with mentor teachers. More information is available at <http://education.ohio.gov/Topics/Teaching/Resident-Educator-Program/Resident-Educator-Overview>.
- 48 TNTP developed the Assessment of Classroom Effectiveness (ACE) to assess the performance of its Teaching Fellows after their first year in the classroom. TNTP uses ACE to determine who meets its standards for first-year performance, and only teachers who meet the standards are recommended for licensure. More information is available at <https://medium.com/tntp-ideas-research-and-opinion/teacher-prep-what-s-data-got-to-do-with-it-1c3ccd44e388#.j4xnmvdhe>.

- 49 Susanna Loeb, Luke C. Miller, and James Wyckoff, "Performance Screens for School Improvement: The Case of Teacher Tenure Reform in New York City," *Educational Researcher* 44, no. 4 (2015), 199-212..
- 50 US Department of Education and National Center for Education Statistics, "First Through Fifth Wave Data File" in *Beginning Teacher Longitudinal Study, 2007-08, 2008-09, 2009-10, 2010-11, and 2011-12*; Sara Mead, Carolyn Chuong, and Caroline Goodson, *Exponential Growth, Unexpected Challenges: How Teach For America Grew in Scale and Impact* (Washington, DC: Bellwether Education Partners, February 2015).
- 51 For an overview of the latest research on Teach for America, see <http://www.brookings.edu/blogs/brown-center-chalkboard/posts/2015/10/27-new-evidence-teach-for-america-impact-student-learning-hansen>.
- 52 Sass, "Licensure and Worker Quality,"
- 53 Dan Goldhaber and James Cowan, "Excavating the Teacher Pipeline: Teacher Training Programs and Teacher Attrition" (CEDR Working Paper 2013-5, University of Washington, 2014).
- 54 Angela Lee Duckworth et al., "Positive Predictors of Teacher Effectiveness," *The Journal of Positive Psychology* 4, no. 6 (November 2009): 540-47.
- 55 United States Census Bureau, Education Finance Branch, *Public Education Finances: 2013, Table 6*, accessed October 1, 2015, <http://www2.census.gov/govs/school/13f33pub.pdf>.



Acknowledgments

The authors would like to thank the individuals who contributed their time and expertise during the crafting of this paper, particularly Allison Atteberry, Stephanie Banchemo, Michael Dannenberg, Julie Greenberg, Cory Koedel, Sara Mead, George Noell, Ben Riley, Andy Rotherham, and Jason Quiara. The authors would also like to thank the Joyce Foundation for providing funding for this paper. As always, the findings and conclusions are those of the authors alone and do not necessarily reflect the opinions of the foundation or anyone we interviewed.



About the Authors

Chad Aldeman

Chad Aldeman is an associate partner at Bellwether Education Partners. He can be reached at chad.aldeman@bellwethereducation.org.



Ashley LiBetti Mitchel

Ashley LiBetti Mitchel is a senior analyst at Bellwether Education Partners. She can be reached at ashley.mitchel@bellwethereducation.org.



BELLWETHER
EDUCATION PARTNERS
IDEAS | PEOPLE | RESULTS

About Bellwether Education Partners

Bellwether Education Partners is a nonprofit dedicated to helping education organizations in the public, private, and nonprofit sectors become more effective in their work and achieve dramatic results, especially for high-need students. To do so, we provide a unique combination of exceptional thinking, talent, and hands-on strategic support.

© 2016 Bellwether Education Partners



This report carries a Creative Commons license, which permits noncommercial re-use of content when proper attribution is provided. This means you are free to copy, display and distribute this work, or include content from this report in derivative works, under the following conditions:



Attribution. You must clearly attribute the work to Bellwether Education Partners, and provide a link back to the publication at <http://bellwethereducation.org/>.



Noncommercial. You may not use this work for commercial purposes without explicit prior permission from Bellwether Education Partners.



Share Alike. If you alter, transform, or build upon this work, you may distribute the resulting work only under a license identical to this one.

For the full legal code of this Creative Commons license, please visit www.creativecommons.org. If you have any questions about citing or reusing Bellwether Education Partners content, please contact us.