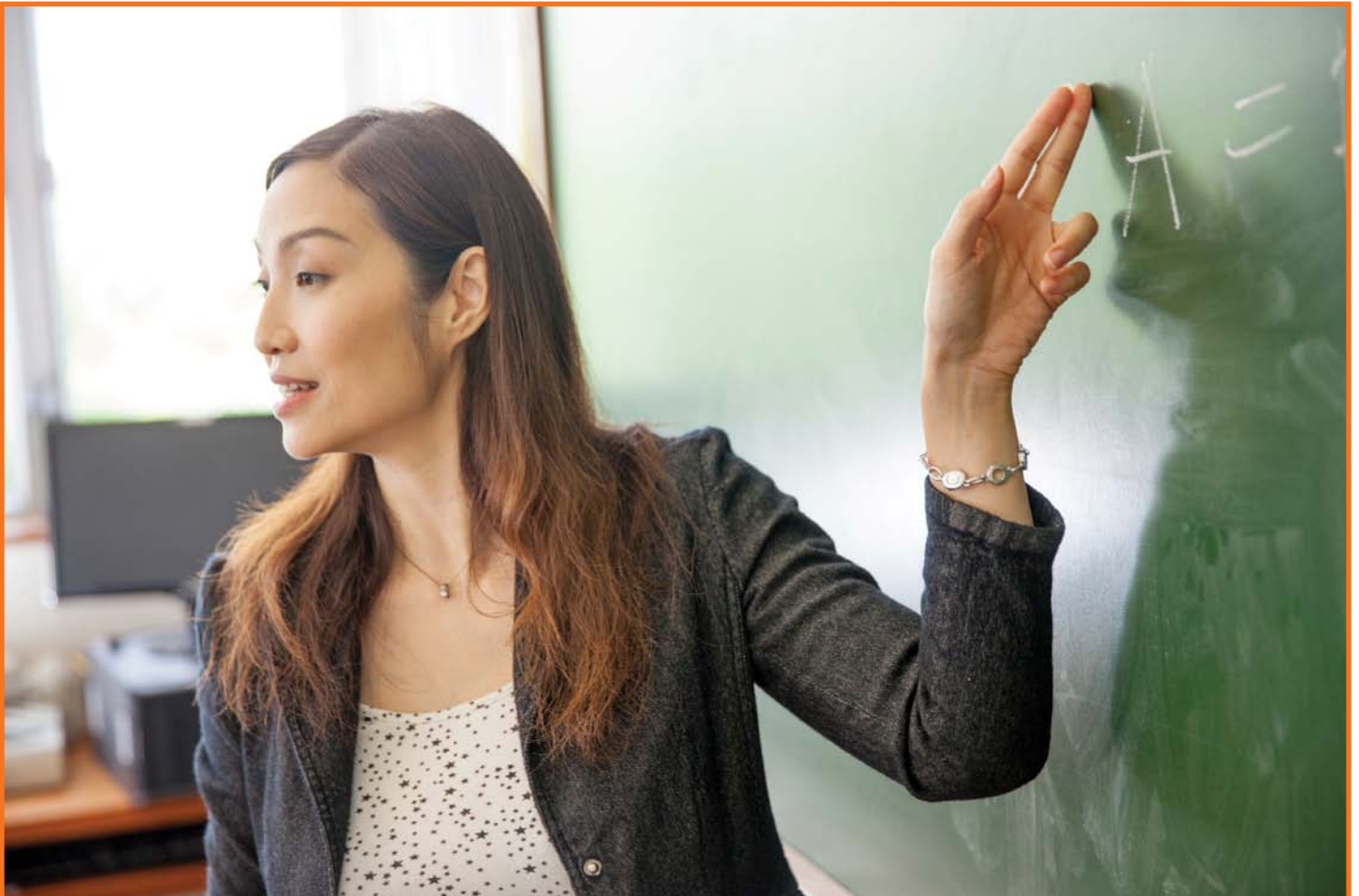


July 2016



**ACCOUNTABILITY IN TEACHER PREPARATION:
POLICIES AND DATA IN THE 50 STATES & DC**

THE COUNCIL OF CHIEF STATE SCHOOL OFFICERS

The Council of Chief State School Officers (CCSSO) is a nonpartisan, nationwide, nonprofit organization of public officials who head departments of elementary and secondary education in the states, the District of Columbia, the Department of Defense Education Activity, and five U.S. extra-state jurisdictions. CCSSO provides leadership, advocacy, and technical assistance on major educational issues. The Council seeks member consensus on major educational issues and expresses their views to civic and professional organizations, federal agencies, Congress, and the public.

COUNCIL OF CHIEF STATE SCHOOL OFFICERS

Tony Evers, President, Wisconsin
Chris Minnich, Executive Director

PREFACE

Our Commitment

Every student deserves a high quality education, and that means preparing and supporting educators to help students reach that expectation.

The Council of Chief State School Officers (CCSSO) in 2012 convened a Task Force on Transforming Educator Preparation and Entry into the Profession that issued a call to action for state chiefs in a report titled *Our Responsibility, Our Promise*. Utilizing three state levers—licensure, program approval, and data analysis and reporting—the report served as an invitation to colleagues in teacher preparation, and others interested in entry into the profession, to join CCSSO in supporting the implementation of the policy recommendations in the report.

CCSSO embarked on this effort by launching its Network for Transforming Educator Preparation (NTEP). Now with 14 states and support from philanthropic and other national partners, NTEP has mapped out innovative implementation strategies to use these levers and make comprehensive policy changes to ensure districts have the workforce to support each student on the pathway for success in career, college, and life.

Why focus on data?

Data is our best illuminator of the possible, the problem, and the array of solutions available.

NTEP's states and partners focus on developing governance structures that use outcome-oriented data collection, analysis and reporting of multiple measures that both support continuous improvement while improving the accountability of preparation programs.

As this report heads for publication, new regulations for the reporting requirements in the Higher Education Act (Title II) are being developed that, if finalized by the U. S. Department of Education (USED), may govern the required data on our nation's preparation programs reported to the USED. These new regulations call for a move away from information on "inputs" (e.g. faculty degrees) toward "outputs" that are stronger indicators of effectiveness of the programs and their graduates.

CCSSO commissioned this report through Teacher Preparation Analytics (TPA) in an attempt to provide all states a first look at the kinds of data that make up their outcome-oriented program review process.

We also asked TPA to offer insights and findings resulting from their 50-state scan on the national landscape for accountability in teacher preparation. This report aims to help constituents who are working together to create data systems that support the continuous improvement of every teacher preparation program, and ensure that candidates for licensure are effective educators beginning on "day-one" of their career.

TABLE OF CONTENTS

Preface	i
Section I.	
Introduction.....	1
Section II.	
About This Report.....	4
The State Program Approval Picture	4
The Status Quo and the Emerging Paradigm in State Preparation	5
The Role of the Key Effectiveness Indicators.....	6
State Information is Not Equally Available.....	7
Section III.	
Findings.....	9
1. Forward Churn.....	9
2. Emphasis on Program Performance.....	9
3. Gauging Consumer Satisfaction	11
4. Reliance on Summative and Licensure Assessments.....	12
5. Relevance of Candidate Cohort Characteristics.....	13
6. Reporting the Data: How Gross, How Granular?	15
7. Making the Indicators Count	16
Section IV.	
Conclusions and Implications	17
<i>EPP-Level vs. Individual Program-Level Data</i>	17
<i>Quality and Rigor of Measures</i>	18
<i>Putting the Measures to Use</i>	19
<i>The Present and the Future</i>	19
Acknowledgements and Author Notes	21
Sources.....	23
Appendix A. Key Effectiveness Indicators	25
Appendix B. TPA Data Collection Protocol	27
Appendix C. Basic State Program Review, Approval, and Reporting Features	59
Appendix D. State Program Review and/or Report Indicators.....	64

INTRODUCTION

How do we determine the effectiveness of our nation's educator preparation programs?

This, the principal question with which this report is concerned, is loaded with assumptions and prone to misunderstanding.

Ideally, we should be asking: *how effective are our nation's educator preparation programs?* However, there exists no universally accepted means of assessing how well our U.S. educator preparation programs are performing. In Teacher Preparation Analytics' (TPA) 2014 report, *Building an Evidence-Based System of Teacher Preparation*, the limits of evaluating the performance of its teacher preparation programs are detailed. These include "the absence of an adequate knowledge base and the lack of data that allow us to identify confidently what the essential characteristics of strong teacher preparation programs are"; the sheer number of preparation providers and the number of individual programs within each of those providers; and the limited capacity of states and providers to collect and analyze any data available.

National efforts are in place to evaluate and monitor educator preparation programs. There are federal reporting requirements under Title II in the Higher Education Act as well as national program accreditation through the Council for the Accreditation of Educator Preparation (CAEP). However, there is broad consensus that the Title II reporting requirements are not helpful in assessing the performance of educator preparation programs. National accreditation is voluntary and historically has been sought by only about half of the country's educator preparation providers (EPPs). Yet there is the uncomfortable reality that a significant number of non-accredited providers and their programs are considered to be as strong as or stronger than many accredited providers.

Over the last several years, the National Council on Teacher Quality (NCTQ) has undertaken several ambitious efforts to illuminate the strengths and weaknesses of teacher preparation programs. Its annual *State Policy Yearbook* is a valuable summary of important state policies that impact the quality of teacher preparation in the states. The more recent *Teacher Prep Review*, undertaken annually with U.S. News and World Report is a bold, if controversial, effort to provide a rating and ranking of all individual teacher preparation providers in the U.S.

Consistent with the recommendations in *Our Responsibility, Our Promise* (2012) the advances in various teacher assessments, econometric modeling, and especially the availability of many kinds of performance-related data that states are increasingly "warehousing" offers an unparalleled opportunity to use new measures as an evaluation tool in the educator preparation field. This is an unprecedented time to imagine the next stage of educator preparation accountability and effectiveness systems.

The most tenuous assumption in the opening question may be our nation's inability to date to construct a valid, compelling, and universally applied set of criteria for gauging the effectiveness of our educator preparation programs.

Educator preparation programs in the U.S. operate constitutionally under the auspices of the individual states. Institutions of higher education, nonprofit organizations, for-profit companies, or other agencies may fund and develop programs anywhere in the country. But for the graduates of any of those programs to be licensed in the states where the programs are offered, the programs have to be approved by an authority designated by state statute to serve that function. That authorizer is vested with the responsibility of developing and applying criteria for program approval that define acceptable program performance at the state level. No other requirements or criteria that might be imposed on educator preparation programs have equivalent policy power.

This report is an investigation of some of the essential policies and practices states employ to ensure that the educator preparation programs they approve adequately prepare new teachers who are effective from the first day they step into their classroom as the teacher of record. In particular, the report looks at those policies and practices related to the reporting of data concerning preparation program performance and the use of those data for program improvement and accountability.

A second important assumption of this report that relates to the central question above is that the most important measure of educator preparation programs—specifically *teacher* preparation programs—is ultimately how effective the teachers are that the programs produce. Is there evidence that candidates for licensure can provide engaging and capable instruction to P-12 students that results in academic progress deemed “adequate” for each student in their classroom? Our ultimate interest is in educator *outcomes*, while acknowledging this is not the only important measure of a preparation program’s success. The term “performance” is used rather than the common term “quality” as the target for our assessment of whether states are taking steps to ensure their teacher preparation programs meet our nation’s need for effective teachers.

The 2014 report noted the differences and inadequacies in the existing program performance measures employed by 15 sample states. Key Effectiveness Indicators (KEIs) are used to identify those inadequacies. TPA formulated KEIs to identify a limited, but strategic, set of measures of program performance that would be useful for three purposes: program improvement, state accountability, and public information.

The KEIs, found in [Appendix A](#) of this report, consist of 20 different program measures that yield data on 12 program indicators. These are grouped into four program performance categories: candidate selection profile; knowledge and skills for teaching; performance as classroom teachers; and contribution to state needs.

The choice of indicators and measures are based on findings from empirical research and, in some cases, on their proven utility for efforts to improve preparation programs. The KEIs track both program-related impacts and a core set of internal outcomes of program practices and policies. In most cases, the indicators and measures can be adopted by states that have adequate data systems. In a few cases, notably assessments of teaching promise and pedagogical content knowledge (or content knowledge for teaching), further research and development efforts are required to enable the indicators to be implemented. If adopted, the KEIs could be compelling for a broad range of stakeholders, reliably guide a state’s program accountability process, and provide a multi-dimensional assessment of programs that suggest directions for program improvement.

National interest emerges in outcome-based teacher preparation accountability

Prior to TPA's 2014 report, a renewed emphasis on improving teacher preparation was bubbling up around the country and gave rise to a multitude of activity. Efforts in Louisiana to link teachers' value-added outcomes to the preparation programs that trained them resulted in the U.S. Department of Education requiring many of its Race to the Top grant recipients to collect outcome-based performance data on their teacher preparation programs. The Council on Accreditation of Educator Preparation (CAEP) was established as a new national accrediting body for educator preparation that promised a more rigorous, outcomes-based approach to accreditation. The previously mentioned National Center on Teacher Quality/U.S. News & World report had recently been released and created consternation among many in the teacher preparation field, who argued the work was unfairly negative and skewed the current state of teacher preparation. At the same time, the U.S. Department of Education worked on new reporting requirements under HEA Title II that signaled the adoption of more outcomes-oriented performance measures at the program level, and established a rating system for programs that included the possible withdrawal of some federal funds for programs consistently rated as low-performing.

States Respond

As noted earlier, CCSSO engaged its membership to produce the report, *Our Responsibility, Our Promise: Transforming Educator Preparation and Entry into the Profession* (CCSSO, 2012). The report contains ten key policy recommendations for state officials to leverage with the authority they hold. Half of these recommendations focus on improving educator preparation program approval and related state data systems.

With the support of the philanthropic community (e.g. Bill & Melinda Gates Foundation and the Charles and Lynn Schusterman Family Foundation), CCSSO launched the Network for the Transformation of Educator Preparation (NTEP) – an initiative which includes 14 states working on developing and implementing the kinds of policies the report recommends. NTEP has enabled representatives from participating states to create a groundswell of activity related to state-level preparation program accountability policies that appear to not only boost the progress of the participants, but also encourage other states to undertake similar efforts. Teacher Preparation Analytics has partnered with this network since its inception.

Outside of the NTEP states and the network of Race to the Top funded states, there was not a complete, nationwide picture of state-level action in regard to improving effectiveness of teacher preparation programs. CCSSO engaged TPA to identify efforts in all 50 states (plus the District of Columbia) to strengthen state-level accountability for teacher preparation programs.

This report summarizes the findings of that research. We emphasize that this report is intended to observe and identify **states' efforts to evaluate the effectiveness of their preparation programs**. It is not intended to determine the actual performance of those programs – information (apart from “approved” or “accredited”) that relatively few states are currently able to provide to interested stakeholders and the public.

ABOUT THIS REPORT

This report is focused on the main program review and accountability policies and the related evaluation measures that state agencies use to make decisions for initial and continuing program approval. We concentrate specifically on teacher preparation, rather than educator preparation more generally, for four reasons:

1. Evaluation of teacher preparation is a more fully developed enterprise.
2. Richer data are available to evaluate teacher preparation programs than to evaluate other programs of educator preparation (e.g., principal preparation or school counselor preparation).
3. The effectiveness of teachers is currently a top concern for policy makers, as teachers play a very visible and central role in educating our nation's students.
4. An exclusive focus on teacher preparation would yield the greatest informational payoff for decision makers within the timeframe for the development of this report.

This report is intended to take an objective look at what states are (or are not) doing on the policy level to track the performance of teacher preparation programs and address deficiencies. The report does not judge the quality of these efforts. TPA acknowledges that some of what is specifically reported in this regard reflects both pragmatic limitations and TPA's own judgments about which state practices and policies are most likely to help move the field forward. The report relies on the assumption that the Key Effectiveness Indicators fall into that category. It should also be noted that the accessibility of information varied by state.

The State Program Approval Picture

As noted, this report focuses exclusively on the policies that govern the work of the state agency specifically charged with approval of teacher preparation programs leading to state licensure; this is where, in terms of responsibility for program performance, "the buck stops." In most states, this agency is the State Department of Education (K-12) or a subsidiary unit that reports to the State Board of Education and that is independent from other government agencies. In a small number of states, the agency for state education oversight is within the governor's office. Still other states have a completely independent board dedicated to professional education standards that approves preparation programs and sets standards for teacher licensure. Finally, program approval for some states is under the jurisdiction of a combined K-12/post-secondary agency.

In addition to the statutorily authorized program approving agency, a significant number of states also require a higher education commission, board of regents, or state university system to independently review teacher preparation providers and programs, prepare periodic reports on their performance, and request remediation of weak programs or, in extreme cases, withhold program funding. Such efforts on the part of higher education bodies can be an important part of the overall preparation program accountability and improvement process. However, teacher preparation program review by disparate governmental agencies also can be a source of tension.

The Status Quo and the Emerging Paradigm in State Preparation Program Accountability

Traditional state review and accountability structures have tended towards a multi-year, comprehensive re-approval determination—akin to and sometimes identical with national accreditation reviews that take place generally every 7 years. In addition, CAEP requires annual updates between long-term accreditation decisions. Program approval agencies in some states (e.g., California) have called for a more substantial biennial update for their public educator preparation programs.

All states and the District of Columbia must report various kinds of information annually to the U.S. Department of Education to comply with Title II of the Higher Education Act. A few elements of Title II data are used by states for accountability purposes, but the great majority is not.

It is noted whether or not each state has a multi-year program approval cycle and whether they require all of their teacher preparation programs to be CAEP-accredited. This is included because national accreditation is one commitment that many states and EPPs undertake in an effort to ensure program quality.

A growing number of states have begun to develop program “report cards” or “data dashboards” to indicate to state officials, preparation program providers, and the public how programs perform on various measures. Included are measures such as candidate completion rates, licensure examination pass rates, placement and retention rates, and various indicators of successful classroom teaching. These annual reports provide concise, largely quantitative information about programs that enables programs to be readily compared with one another and may point to weaknesses that can be addressed before their next scheduled long-term comprehensive review. Increasingly, states are giving greater weight to these interim reports, and research shows that some are considering completely replacing the comprehensive multi-year approval process with annual or biennial reviews.

These annual program reviews and reports are an important tool for program improvement, state accountability, and public information. Rigorous and strategic short-term reports can provide important, timely information to teacher educators and the state officials who monitor teacher preparation programs that can help reveal and address program weaknesses, as well as point to exemplary program practices. More frequent reports encourage a culture of continuous improvement among educator preparation programs. An accountability system that employs concise and frequent program performance reports is also readily adaptable for the diverse configuration of program providers that prepare teachers today.

Large EPPs that have significant capacity to do more exhaustive, qualitative reviews are increasingly joined by small, lean EPPs, some independent of larger parent institutions altogether. These smaller EPPs may be less likely to pursue national accreditation or to have the capacity or inclination to put together detailed and lengthy periodic accreditation reports. Concise, annual reports enable all EPPs and their programs in a state to be assessed and compared on equal footing on the basis of comparable data and identical measures with the same expectations for completeness and rigor. More streamlined reports that are easily comparable from one year to the next can provide salient,

meaningful, and transparent information to the public, elected officials, and other decision-makers—information that will provide a timely basis for assurance or concern about the performance of the state’s educator preparation programs. One of the goals of this report is to identify which states have an annual program review and reporting system that serves a state accountability function and to enumerate some of that system’s key characteristics.

NOTE: A few states issue annual public reports on their teacher preparation programs that are for information purposes only. Public information about program performance, with appropriate constraints on disclosing individual candidate or faculty data, is not only an obligation of a state regulatory agency but also provides an additional layer of accountability. However, if a state’s public report was not based on the same data as the state program approving agency’s review process and approval decisions, then those public data are not described in this report.

The Role of the Key Effectiveness Indicators

The Key Effectiveness Indicators (KEIs) are used as a screen for examining states’ program review and reporting processes. Although states are under no obligation to adopt the KEIs as part of their accountability structure, the KEIs are largely consistent with the indicators many states have already adopted. They are consistent with the annual reporting requirements specified by CAEP, though the KEIs are intended for individual program-level data and not only the provider-level data required by CAEP (i.e., data for institutions of higher education, or IHEs). The KEI list of indicators is also more extensive than the set of indicators many states employ, or are planning to employ. This is partly because in this report, performance is defined in a broad sense that includes not just the degree of success and impact program completers have in teaching P-12 students but indicators in three other categories. All of these are outcomes of various program policies and practices:

1. Indicators reflecting the composition of a program’s candidates.
2. Indicators reflecting what candidates have learned or can accomplish as the direct result of the program’s courses and completion requirements.
3. Indicators reflecting the contribution of a program’s completers to a state’s need for teachers.

Employing multi-dimensional indicators serves other purposes. It permits a kind of “triangulation” among the indicators as a check on the validity of the measures employed. For example, if a program’s completers scored comparatively low on their content knowledge assessment, it would not be surprising if they also scored low on indicators measuring P-12 student impact. It would likely be more unsettling, however, if the completers score high on content knowledge assessment and low on student impact. On the one hand, such a result might prompt questions about the validity of the measures used. On the other hand, this type of result can stimulate and guide efforts to find out whether or not there are deficiencies in the program that are responsible for such a result. This is the second benefit of multi-dimensional indicators. The bottom line is that multiple indicators can provide initial diagnostic guidance for program faculty and staff.

State Information Is Not Equally Available

TPA's experience with their 2014 report made it clear that it would be easier to obtain the information sought from some states than from others. It is not an accident that states receiving Race to the Top funding to strengthen preparation program accountability or states that participate in the NTEP initiative have realized significant progress in their efforts. States that are farther along in the process tend to have more resources—and more informative resources—available online than states that have done less work in the area. These states also tend to have more staff and financial resources specifically devoted to preparation program review and accountability, and those staff are likely to have been involved in the work long enough to become knowledgeable about the nuances of their state's relevant policies and practices.

What TPA did not anticipate from the earlier report, however, was the level of churn that currently seems to exist in states' program review and accountability practices and policies. This is one of the key findings discussed in more detail in the next section of the report. TPA attributes this principally to accelerated forward movement in many states, though a few states have slowed their progress over the last year. It also appears to be an inherent feature of teacher preparation program effectiveness that states will always seek to refine and improve whatever policies and practices they currently have. Continuous improvement attends not only to the programs that are measured, but even to the measures themselves.

The result is that it becomes exceedingly difficult in a number of states to precisely identify current preparation program review and accountability policies and practices. Indeed, in anticipation of adopting new policies in the near future, a few states, such as California, Connecticut and Louisiana, have suspended their incumbent system altogether and, in effect, have no discernable current review and accountability policies. In view of this situation, CCSSO honored Louisiana's request not to identify specific state policies for this report.

In other states, less sweeping but still important changes in program review and accountability structures appear to be imminent, so using the status quo as the basis of comparison for those states would be to use information that will soon be out-of-date. Some states have documented efforts to advance current program review and accountability practices and policies, but it is unclear to the researchers when these anticipated changes might occur. Using anticipated changes as the basis of comparison might mean attributing practices and policies that will never come to fruition. If TPA was confident from conversations with state officials that changes were well underway and would be substantially implemented by the fall of 2017, those changes were accepted as equivalent to "current" policy in an effort to be forward-looking in the analysis and postpone its obsolescence. For the larger database behind this report, TPA did its best to record the status quo as of May 1, 2016, and use that as a baseline for charting future progress. All of this indicates that gathering information about state policies and practices in this area is not an exact science and involves some guess work and judgment calls, sometimes even by state officials who know their preparation program review and accountability systems better than anyone.

Finally, there is a realistic constraint on the depth of information obtained through document searches and follow-up telephone conversations with state officials. States that are in comparatively early stages

of developing their preparation program review process may have reached agreement on specific procedural details and performance indicators, while not yet determining the specific measures and calculations they will use to score programs' actual performance. Furthermore, in states that have already implemented all or part of their process, officials who are expert in preparation program accountability policies and practices may not be familiar with the technical specifications of the measures they employ or all of the elements of their data system. This increases the data-gathering burden on both state officials and the researchers, making it difficult to obtain equally deep information from all states.

In spite of these limitations and preferences, we believe this report conveys a sufficiently accurate picture of the states' program review and accountability structures to ground solid comparisons between them and inform and stimulate continued efforts to strengthen those structures. For a few states, however, it was not possible to obtain information either from documents or interviews at a level of detail that provides confidence in its accuracy; these cases are denoted by an asterisk in the tables in Appendices C and D. Additionally, Wyoming – our least populous state – has only one provider in the state that offers licensure programs for teachers. Understandably, there is no need to develop an indicator-based state accountability system. For these reasons, readers should consider this report to be an initial analysis that merits periodic updating.

For most states TPA has gathered much more nuanced data than are presented in this report. This includes finer detail on the specific measures that some states employ. The full breadth and depth of the data sought from all 50 states is reflected in the TPA Data Collection Protocol, which can be found in Appendix B. The intent and hope is that all of these data, together with additional information, will be incorporated into a searchable, relational database that will be accessible online and updated periodically to reflect the inevitable changes that will occur over time. The database would allow users to find out current state policies and practices and track changes over time.

FINDINGS

The findings from this review of the teacher preparation policies and practices of the 50 states and the District of Columbia reflect the data summarized schematically in the two tables in Appendix C and Appendix D. Appendix C is a summary assessment of the basic features of the state program review, approval, and reporting process. Appendix D summarizes the indicators—the measured program features and outcomes—that each state employs in assessing program performance.

The section below discusses trends that emerged from this analysis including differences between the data and measures states employ, and the different kinds of responses states have made to some of the logistical and statistical challenges they face in system application.

1. Forward Churn

Forward churn was the most significant and surprising discovery in the review of state policies and practices regarding teacher preparation program review and accountability. TPA had not anticipated the number of states that are in the midst of reforming their practices and policies, nor the progress that many states had made since their 2014 research.

There are more states however, that lack a comprehensive program review system. Many of these states indicated that they aspire to acquire such a system, yet those aspirations have yet to move to concrete action. Officials in several states bluntly stated during telephone interviews that there was simply “no appetite” in their state for changing their educator preparation accountability system. This reveals a significant distance between states that are moving forward and those that remain part of the status quo. If all of the lagging states started today it would still take several years to make significant progress in closing this gap.

2. Emphasis on Program Performance

There is a growing recognition among states committed to reforming their preparation program review and accountability systems that the *performance* of programs, and not the features of various program components, is the true test of program quality. Programs need to devote a significant number of hours to candidates’ clinical preparation, forge strong partnerships with P-12 schools, and/or improve the alignment of their courses in order to ensure that they can adequately prepare their candidates. But it is candidate *outcomes*—the strength of their knowledge and skills, the amount of time they remain in the classroom and profession, and the impact they have on their students’ success—that are of increasing interest to policymakers, teacher educators, and teacher candidates themselves.

Among all of these, the impact of teachers on P-12 students’ success in school is the “holy grail” of measures of program performance. States that adopt this indicator accept the propositions that (a) teacher preparation programs are an important contributing factor in how completers of those programs perform as teachers, (b) stronger programs (with other factors being equal) will produce more effective

teachers than weaker programs, and (c) differences between programs can be validly and reliably measured. Louisiana was the first state to make the connection between teachers' classroom success, as measured by their value-added scores, and the teaching program that prepared them. They successfully used differences between very high-scoring and very low-scoring programs on value-added measures to diagnose and address specific weaknesses in the low-performing programs. State officials in Tennessee and North Carolina have undertaken similar diagnostic efforts using teachers' value-added scores.

TPA's scan indicates that a growing number of states are currently using—or have committed to implement within the next year—a teacher impact indicator as part of their teacher preparation program accountability system. Some states, such as Washington, have implemented a teacher impact indicator despite strong opposition from key stakeholders. In several states, the vehicle for the indicator will be a value-added system. Even more states will depend upon a student growth model that enables states to rely on a less-sophisticated data system but doesn't control for as many potentially confounding background variables as a value-added model does. Missouri recently attempted to implement a teacher impact indicator of preparation program quality based on a student growth model, but the effort revealed more variation in the scores of the teacher-completers from individual programs than between the teacher-completers from different programs. Similar results in other states may spur much-needed focus in the measurement and research community on advancing the sensitivity and reliability of the statistical models employed for these purposes. Of greater concern is an emerging retreat in using teacher impact data as part of performance evaluations of teachers, and this could result in making those data inaccessible for the assessment of program performance, as well.

Nevertheless, a number of states appear to be solidly committed to using teacher impact measures, and even more states employ other on-the-job measures as indicators of program performance. Some states that have a mandatory annual performance evaluation for all public school teachers use early-career teachers' scores on the evaluation as an indicator of the comparative strength of their preparation programs. Other states use the individual component scores on each section of the performance evaluation to construct several different program performance measures. These may include a classroom observation score, a portfolio assessment score, a teacher impact score, a self-assessment score, or other measures of teacher performance. Some states, however, prohibit the use of teachers' annual performance scores for any purpose other than their intended use for performance evaluations (even though individual teachers' scores would remain confidential).

Apart from their use in an annual performance evaluation, some states conduct a classroom observation assessment of all first-year teachers, often as part of a required induction program. They use this observation assessment to construct a measure of preparation program performance based on the scores of each program's completers.

Six states employ or will soon employ surveys of K-12 students' perceptions of their teachers' effectiveness. Several of these surveys have been developed over the years, and research has demonstrated their validity and their increased predictive accuracy when used together with value-added data. That is, teachers who receive high marks from their K-12 students for demonstrating qualities like caring, engagement, or helpfulness also tend to earn high value-added scores. The reverse is also true.

The majority of states include completer job placement and retention rates, and even states that do not otherwise have an annual or biennial program performance review or a regular preparation program report card generally track and publish program completion and completer placement and retention rates. Publication of these data is often part of a broader higher education productivity report that's regularly undertaken by a state higher education agency, as opposed to the preparation program approving authority. The state authorities who approve preparation programs may differ substantially in the data they use to measure teacher retention even though state productivity data are substantially the same from state to state. Kentucky, for example, collects five years of retention data, but the professional standards board looks only at three years of data in reviewing programs. Washington State also collects data on completers' retention in teaching within the state, but the professional standards board uses a completely different measure for its retention indicator. In order to account for differences in school factors that may influence retention rates unevenly, Washington's program accountability metric for retention is the mean difference over three years between the percentage of a program's completers who return for a second year to the school in which they began their teaching career and the second year return rate of other first-year teachers in the same school over that period.

3. Gauging Consumer Satisfaction

Many states administer one or more surveys to gauge the satisfaction of various program "consumers" with the program's efficacy in preparing new teachers. One obvious set of consumers are the program completers themselves, whom a number of states ask for their perceptions of how well prepared they were by the program for their classroom teaching responsibilities. In 16 states, these surveys are administered right after student teaching or an internship year to get responses from virtually all candidates; at that point candidates can still be located by the program and can be compelled to complete the survey as a requirement for graduation. Eighteen states attempt to survey program completers during their first year of teaching, believing the completers can provide a more valuable and accurate appraisal of their preparedness after they've had one or two semesters of experience as teachers of record. It becomes more difficult over time to locate completers and compel their participation once they've completed the program, but completers who teach in public schools in state can be easily located through employment records. Some states told us they compel participation in post-completion surveys as a precondition for receiving a second-stage teacher license.

The second consumer group includes school principals, district superintendents, or other school and district supervisory personnel. Fourteen states survey this audience about their satisfaction with candidate preparation. The assumption of these surveys is that the supervisors who complete them (usually on each first-year or second-year teacher) can judge the quality of a teachers' preparation through classroom observation, meetings, or other contact during the school year. However, the line between commenting on the skills and talents of the teacher and the implied strength of the program is a fuzzy one. Moreover, although a school principal or superintendent may be able to appraise teachers accurately on some aspects of their classroom practice, the supervisor may lack the content expertise or the familiarity with new teaching techniques to be able provide the most knowledgeable assessment. Some states request that the individual completing the survey be the new teacher's direct supervisor in order to reduce the likelihood that the supervisor will be out of touch with the teacher's field or teaching practices.

Another cautionary reality is that there are relatively few well-financed districts that can devote resources to effectively selecting and training teachers and administrators to conduct rigorous review and assessments of beginning teachers with any sort of inter-rater reliability. It is even more problematic in low-wealth districts. In order for feedback from employers and supervisors to contribute meaningfully and fairly to program improvement and accountability, appropriate support and procedures for ensuring reliability are imperative.

4. *Reliance on Summative and Licensure Assessments*

It may seem misleading to refer to teacher candidates' performance on in-program (or pre-service) assessments before they embark on their actual teaching careers as program "outcomes," and yet such assessments are precisely intended to be measures of the knowledge and skills candidates acquire from going through the program. These are critically important outcomes in that they can be attributed directly to the candidates' program and practice teaching experience or to the expectations programs have for candidates' knowledge and teaching skill. Candidate performance should thus be less influenced by extraneous factors. If anything about these assessments is questioned, it should be the validity of the specific measures some states use to gauge the performance of programs. With adequate metrics, however, pre-service candidate performance becomes a critical data point in the efforts of preparation program faculty and state officials to understand how a program ultimately impacts completers' success and, if necessary, to identify critical steps towards program improvement.

The most widely used indicator for assessments of preparation program performance is a candidate's score on state licensure exams. Most commonly, these scores are reported as program cohort pass rates because this is the most consequential federally-required reporting metric. Programs that don't meet an 80 percent pass rate may lose eligibility for federal funding, and many states will suspend admissions or impose other sanctions on programs that don't meet the passing mark.

The relevant licensure assessments generally include tests of content knowledge in candidates' teaching subjects, such as the Praxis II or NES/Pearson subject assessments, and some sort of assessment of teaching skill and knowledge. Many states still use the Praxis Principles of Learning and Teaching (PLT) or a similar pen-and-paper assessment of pedagogical proficiency, but an increasing number of states are requiring a passing score on more performance-based assessments of candidates' actual teaching practice such as the edTPA or the Praxis Performance Assessment for Teachers (PPAT) either for program completion or for licensure.

In the Key Effectiveness Indicators, TPA also includes as an indicator an assessment of "pedagogical content knowledge" (PCK), also called "content knowledge for teaching" (CKT). The edTPA and PPAT assess this kind of knowledge for the narrow slice of the subject field where a candidate teaches specifically for those assessments, but this provides no assurance that the candidate has the same pedagogical mastery of a much broader range of teaching content that a new teacher might be called upon to teach in his or her very first semester on the job. Currently, no state employs such a comprehensive assessment of PCK. TeachingWorks (University of Michigan) and ETS are currently piloting PCK assessments for elementary

grades as part of their National Observation of Teaching Exam (NOTE) designed to assess a much broader range of content-based teaching knowledge as well as key pedagogy tied to student success.

5. *Relevance of Candidate Cohort Characteristics*

In the face of the increasing emphasis on outcomes to gauge program performance, it may appear regressive to require programs to report on the intellectual, dispositional, or demographic characteristics of their candidate and completer cohorts. These, too, however, can be regarded as “outcomes.” One state official explained this well when she noted that if a teacher preparation program is held responsible for admitting or graduating a highly diverse candidate pool, the strategies the program uses to try to meet its target is not of much interest to the state. The degree of candidate diversity is precisely the outcome of the strategies—either adequate or inadequate—that the program employed. These cohort characteristics are important to the realization of several key stated goals of the teaching profession: to diversify the workforce in order to better reflect the growing diversity of the student population; to attract high-performing individuals into teaching for the benefit of students and the prestige of the profession; and to ensure that those who enter teaching are sufficiently committed that they will remain in the profession in spite of its many challenges.

Additional outcomes measured by some programs include the academic proficiency of entering candidates and program completers, entering candidates’ teaching promise, and the equivalence of the program completion rate of each candidate cohort disaggregated by gender and race/ethnicity. These outcomes, included in TPA’s Key Effectiveness Indicators, can have an important impact on how well candidates perform on their summative and licensure evaluations, and how well they may fare in the early stages of their teaching career. Additional measures that can provide important information about a program’s ability to ensure it meets the needs of its admitted candidates include the degree to which admission and completion rates of candidates by sub-group are equal, the overall performance of candidates by sub-group, academic proficiency, and fitness for teaching.

CANDIDATE/COMPLETER DIVERSITY

Virtually all 50 states report on the demographic diversity of their candidate pool or their completer pool, but this information is not always used as a measure of program performance. States employ a variety of different measures of diversity, and few states consider program diversity to be as consequential a measure for a program’s state review as other outcomes. Most states simply list the diversity of the entering pool of candidates and the diversity of the graduating cohort, providing no attempt to track the completion rate of each sub-group that began the program. Thus, there is no measure of how successfully the program prepares teachers from different demographic backgrounds.

Several states, however, do have more sophisticated diversity metrics. Washington, for example, seeks to determine how closely the diversity of candidates in teacher preparation matches the diversity of the university as whole. The state also measures the completion rate of admitted sub-groups of students as suggested in the KEIs.

TEACHING PROMISE

Teaching Promise – the extent to which candidates have various attributes identified with success in the classroom – is one of the least widely adopted indicators across the states. TPA included Teaching Promise as one of the KEI indicators in its 2014 report knowing that the indicator was not widely recognized and would take time to gain acceptance as an important measure of program performance. Assessing Teaching Promise contributes a potentially important dimension of selectivity for teacher preparation programs by identifying candidates who already have a commitment and proclivity for P-12 teaching. This indicator might also push programs to consider how they might identify the lesser-considered knowledge, skills, and attributes such as perseverance, grit, and cultural competency.

The profile of entering candidates is an important data point to consider in the interest of a full appraisal of program effectiveness and to facilitate program improvement. Several states require programs to administer an assessment of the habits, attitudes, or dispositions of entering candidates, but in all cases, the results are used internally to ensure the program identifies and addresses candidate deficiencies. Missouri uses a normed test developed by Pearson that is based on a previously-validated inventory of salutary work styles of other professions from the U.S. Department of Labor. Missouri's assessment scores candidates on 16 measures over six domains and could easily be used as a program performance indicator.

ACADEMIC PROFICIENCY

Measuring the academic proficiency of teacher candidates is currently one of the more controversial indicators of program performance. Its adoption as a KEI is supported first by a body of research indicating that academic proficiency is an important factor in degree completion for college students and is correlated with teacher candidates' scores on their content knowledge assessments for licensure. Second, there is research that shows that, other things being equal, academically strong teachers are more effective than academically weaker teachers—especially in teaching low-income and minority students who depend more heavily than other students on their teachers to make academic progress.

Many states and teacher preparation programs have independently raised the entrance requirements for teacher preparation candidates over the last decade. Research studies have shown that as a result, the overall academic proficiency of teacher candidates has risen as well. CAEP has also raised program entry standards to a 3.0 *mean* GPA for candidates in all EPP programs and to a *mean rank* in the national score distribution of at least the 50th percentile on the SAT, ACT, GRE, MAT, Praxis I, or other standardized assessment of college-level knowledge.

In general, most states require their teacher preparation programs to report limited information related to their candidates' or completers' academic proficiency. Many states don't use academic proficiency as a program performance indicator at all because they assume that individual candidates must meet a certain minimum benchmark by state statute. A number of states, however, report mean scores on academic assessments for program candidates as an indicator of comparative performance between programs—in some cases for individual certification fields. This is done more commonly using Praxis

Core (Praxis I) scores than SAT or ACT scores, but some states do report SAT, ACT, MAT, and GRE scores, sometimes in addition to Praxis Core scores. Some states also use GPA as an indicator of candidates' academic proficiency. This may be a mixture of high school GPA, GPA in general distribution courses, or GPA in courses within a candidate's college major. The GPA may be measured either upon program entry or program completion. Furthermore, several states report not only the mean assessment or GPA results but also pay attention to the distribution of candidates' assessment and/or GPA score. For example, Washington compares the GPA of teacher candidates in their major field to those of all university students in the major. The state also flags a program for review if more than one third of its entering candidates have a GPA below 3.0 or if one-third of completers in a given year have a GPA in their subject major below the 33rd percentile GPA for all majors in that field.

6. Reporting the Data: How Gross, How Granular?

The level of data collected and reported varies widely by state. "Program data" could denote data on a specific route to teacher certification in a specific field and at a specific school level. Or, it could denote data on the Education Program Provider (EPP) – the parent agency or "unit" that runs one or more specific teacher preparation programs for licensure.

All states gather and report data at the EPP level. Many states also collect at least some data on the individual program level, or at least at the subject certification level (e.g., on all completers from various programs at an EPP who are recommended for a license to teach elementary mathematics or secondary general science). Aggregation by teaching subject is often necessary in reporting teacher impact scores that depend upon statewide assessments. This is because the assessments are given in specific subjects and grade levels, and there may not be sufficient numbers of teachers to report on individual certification routes within a single EPP without breaching teacher and student privacy.

The number of teacher candidates or completers within each individual program on which states report is referred to as "n-size." States deal with this "n-size problem" in various ways. Some states simply refrain from reporting data on programs with too few candidates. This "too few" number varies by state, and can be anywhere from fewer than five to fewer than 25. Other states aggregate data on small programs that lead up to the certification subject, which may address the privacy problem and still enable program staff and faculty to use the data for diagnostic purposes.

It is also common for states to report all data (both EPP-level and individual program and subject-level) over a period of several years—most frequently three but up to five years. That means that the measures reported are essentially an average of all included cohorts, with the oldest cohort eliminated and a new cohort moving into the measured group the next year. This procedure has the advantage of solving the n-size problem, and likely provides a more stable indication of program performance over time. The risks of this strategy, however, are that it (1) may mask noteworthy fluctuations in performance from one year to the next and thus make it more difficult to track program improvement or decline over time, and (2) may require a complicated regression equation to account for the likelihood that the influence of the program on new teacher performance declines over time in comparison with P-12 school and other

factors. Kentucky employs this kind of corrective equation in calculating the measures for three-year cohorts. The state refers to this as “proportional accountability.”

7. Making the Indicators Count

States must not only define which indicators and measures they will use to assess program performance, but also determine how to use those indicators to evaluate, elevate or remediate programs. This is the ultimate purpose of the program accountability and improvement process, and states have adopted a variety of approaches to achieve this goal.

States with a high-stakes approach regard the measures as grades—or components of composite grades—that are assigned to programs to determine their status vis-à-vis accountability actions. Depending upon their composite score or rating or ranking, programs may be recognized as exemplary, designated for additional review, or sanctioned in some way. Some states assign different weights to the indicators; teacher impact indicators are most commonly assigned the largest weight. Only five states currently weight their indicators or have plans to do so soon, but other states have noted their intention to follow suit. States do not uniformly agree on the indicators that should be weighted.

States with a low-stakes approach regard the measures as signals that require further investigation before assigning grades to programs or taking any sort of corrective action. Missouri and Washington typify this approach. These states set performance benchmarks for programs on individual indicators; if a program score falls below the benchmark the state follows up with further review. Programs that continue to perform poorly or show problems are required to improve or face appropriate consequences.

A low program ranking or rating may incentivize programs to improve. In the absence of an overall rating or score, low scores on individual indicators may also be sufficient motivation for improvement.

The perceived churn in states’ efforts to revise their program accountability systems is mirrored by a similar churn over the last few years in national program accreditation. CAEP introduced its new standards in 2013 and continues to revise both the standards and various accreditation policies and procedures. CAEP offers state partnership agreements that can leverage a state’s program review process with the EPPs pursuit of national accreditation. However, partly because of the more stringent accountability and reporting requirements CAEP accreditation now entails, and partly because of the lack of finality to the standards and procedures, the timeline for some states to partner with CAEP (thereby using CAEP accreditation as equivalent to state approval) is quite long and at best tenuous in several states.

Among the states that are CAEP partners, some require all of their public programs to be CAEP accredited. Many more states, however, have adopted CAEP standards to guide their state program approval process and require all programs—whether CAEP accredited or not—to meet those standards or standards the states deem equivalent. A contribution of the KEIs can be the opportunity to create coherence in data collection for both purposes.

CONCLUSIONS AND IMPLICATIONS

Findings from this report indicate that more than half of states have a formal system structure—either an annual or biennial accountability-related review of their public and private teacher preparation programs—to support a continuous improvement process. Yet, there are significant shortcomings in the actual features of those systems. A number of states do not review their alternate route programs on an annual or biennial basis. In many states, data are reported only at the educator preparation provider level (e.g. the university), which limits the value of those data for improving individual preparation programs that are offered by a given provider. Many states that do report program-level data do not share these data with the public. Although there are 15 – 20 states that employ many of the program performance measures recommended by TPA for state accountability and program improvement efforts, many more states employ far fewer. Even in states that are farther ahead, conversations with state officials reveal that performance measures often have deficiencies or limitations.

States must decide how to use the program performance measures they've adopted to effectively drive and improve their approach and strategy. CCSSO and TPA will advocate for the utility and power that program performance measures can have as drivers of program change, while cautioning states not to give more credence to the measures than they warrant.

Educator Preparation Provider-Level vs. Individual Program-Level Data

The issue of data granularity as discussed in this report necessitates elaboration; as does the related issues of the accuracy and utility of measures at the provider vs. program level.

For some accountability purposes, EPP-level reporting may be satisfactory. The 80 percent licensure examination pass rate, for example, is most frequently reported by EPPs because this is what federal law, most state governments, and the national program accrediting agency have required for their accountability purposes. These EPP-level indicators are crude measures. Aggregating data on all programs operated by a particular provider can mask significant differences in performance between those programs. A program that may look satisfactory in the aggregate in reality is an indicator of the satisfactory rating of the provider—not the program. The provider may have a number of very high-performing programs that ought to be emulated, but there may be an even greater number of low-performing programs that should be targeted for closer scrutiny and possible intervention.

From the standpoint of real improvement of a teacher preparation program, only data on individual licensure programs—or at least data that permit programs to identify their individual performance scores—are ultimately useful. It remains difficult, however, to report meaningful program data on programs that graduate a small number of candidates in a given year. Data on such small populations lack predictive power, and risk violating the privacy of candidates or completers (and in some cases, their P-12 students). This is one reason why states or EPPs that collect program-level data refrain from reporting it publicly, though they may provide this information to individual

programs themselves because data on even a few candidates can be valuable for diagnostic purposes. It is noteworthy that the Department of Education will require reporting at the individual program level in its proposed new rules (now in final comment for a proposed late 2016 release) for educator preparation reports under HEA Title II.

Quality and Rigor of Measures

As important as it is for states to adopt the right indicators, it is even more important to adopt rigorous, valid, and informative measures to fill those indicators. A majority of states, for example, report pass rates of candidates on various pre-service or licensure assessments as a comparative measure of program performance. There is abundant research, however, showing that pass rates on teacher licensure examinations are poor predictors of teaching performance. Most state-level pass rates hover around 99% where the benchmark for successfully passing a licensure exam is a minimum threshold. While the passing scores are low and do not provide adequate information on future teaching effectiveness, that same research has shown that actual scores on licensure examinations do correlate with teaching success. This implies that using mean scores of candidate or completer populations would be a better measure than pass rates alone. Reporting data (e.g., by tercile) on the score distribution of a candidate or completer cohort is equally important because that data would reveal whether the cohort bunches around the mean, or whether there are significant numbers of very high-scoring or, more troubling, very low scoring individuals.

Delaware, Kentucky, Tennessee, and Washington all use distribution statistics as program performance measures. Kentucky notes the percentage of candidates who scored in the top 25 percent nationally on licensure examinations. Tennessee indicates what percentage of program completers scored in the top and bottom quintiles on its value-added data. Washington identifies a program for further review when more than one-third of its candidates or completers score below a certain threshold for some of its indicators.

Another limitation that all EPPs and states face is the inability to obtain data about the placement, retention, or performance of program completers who leave the state where they completed their program, or who teach in private schools. A few states are beginning to address this and a new effort from National Association of State Directors of Teacher Education and Certification (NASDTEC) could help resolve the problem on a national scale. NASDTEC, along with a few states, are piloting the Multi-State Educator Look-up System (MELS) that will allow states to determine if graduates from their educator preparation programs are employed as a licensed teacher in another state. Without this type of system, the reliability and value of program performance measures are severely limited.

A third measurement challenge is employing candidate or completer assessments that don't adequately distinguish between different skill or knowledge levels. States that use teacher impact measures in non-tested subjects, for example, have often found that teachers in those subjects—which may use teacher-designed or district-designed pre- and post-tests—tend to score significantly higher on average than teachers in state-tested subjects. Similarly, states that have annual teacher

performance assessments have found that the distribution of teacher scores is skewed towards the high end of the scale. Likewise, depending on where states set their “cut” or passing scores on licensure examinations, a similar consequence can occur; there is a large difference in the proficiency of candidates even if they all “pass”.

Putting the Measures to Use

As we noted above, states tend to approach the measures as either “high stakes” or “low stakes” consequences. High-stakes states assign a composite score yielding a ranking or rating to a program (often including a relative weighting of the individual measures) that carries with it accountability-related consequences. Low-stakes states regard scores on individual indicators as signals that either provide reassurance to state officials of satisfactory performance or trigger further review.

Assigning composite scores to programs using either weighted or unweighted indicators may satisfy policymakers, the public, and some higher education members who require some indication of readiness of the state’s teacher preparation programs. Implementation of this effort, however, comes with a number of cautions.

First, if the composite score is assigned at the EPP level and not the individual program level, it may mask significant differences in the performance of individual certification programs at the EPP and is thus misleading. It is common for institutions to operate multiple programs, and states should be cautious not to inadvertently imply that all programs fall under a singular rating. Furthermore, ranking could mean grading programs on a curve. A provider ranked high in a state where program performance is generally low may be far weaker than its ranking would connote. Conversely, a provider ranked in the lower half in a state with a significant number of high-performing programs may be much stronger than its ranking would indicate.

Weighting indicators can also be problematic. It seems reasonable that some indicators should be considered to be more important than others—especially indicators of completer outcomes, such as persistence or impact on the P-12 classroom. What precise weight to assign the various indicators is ultimately an arbitrary decision, and a composite program score masks actual program strengths and weaknesses on the various individual indicators much as an EPP aggregate score masks strong and weak individual programs.

The core question for all states is: which policy actions are justified given the reliability of the measures used to assess program performance?

The Present and the Future

The challenges and limitations described above are not insurmountable. If anything, this report shows a story of progress, despite the fact that the snapshot of the data presented here captures only the near present. State officials, teacher educators, school leaders, policymakers, and the public must be patient in the face of those limitations. We may never have the perfect accountability system for

teacher (and educator) preparation; we are far from it at present. Moving forward, this effort should be guided by confidence in what can be accomplished, and a realistic understanding of what cannot.

Many states have already implemented thoughtful and effective program accountability systems, and many others are making clear progress toward that goal. The best systems remain incomplete, with many measures that are serviceable but not quite adequate. TPA has found, however—not just from this 50-state review but also by monitoring the progress made by the NTEP states and several others over the past three years—that states are beginning to use these systems to provide some of the most powerful and effective state-sponsored program accountability and improvement systems we’ve ever seen.

States must be cautious about the application of these measures and the power tied to them because these systems and the measures they employ are often underdeveloped. The entire education enterprise needs to see itself as engaged in the process of continuous improvement with respect to educator preparation program accountability across the nation. We’re still testing the utility and applicability of indicators, trying to improve the validity and reliability of the assessments we use, and improve the quality of the measures we construct to make use of those assessments.

None of this means that states and providers should refrain from using the indicators and measures currently developed and described in this report. In the fields of nutrition or medicine, we follow dietary guidelines and accept diagnoses that we trust are based upon the best science available. Over time, we will inevitably receive new, better guidance and more exacting diagnostic tests, because science is always progressing. The same holds true for the field of education.

No matter how advanced our preparation program indicators and measures become, they will never tell the entire story. Indicators and measures should continue to improve in their ability to indicate how a program is performing and identify the general locus of problems. These indicators remain a signal, not a cause, and more, often qualitative information will be required to identify the true cause of a problem, and to devise a strategy for addressing it. It will take state officials working *with* teacher educators and other experts—not in opposition to them—in order for preparation providers and programs, and their accountability systems, to reach their full potential.

ACKNOWLEDGMENTS AND AUTHOR NOTES

This document, *Accountability in Teacher Preparation: Policies and Data in the 50 States & DC*, was authored by Dr. Michael B. Allen and Dr. Charles R. Coble, Partners, Teacher Preparation Analytics. The effort has been steadfastly encouraged and supported by Mary-Dean Barringer and Katrina Miller of the Council of Chief State School Officers (CCSSO). In addition to the support provided by funders of CCSSO's Network for Transforming Educator Preparation (NTEP), Stephanie Banchemo, Program Director for Education with The Joyce Foundation, provided significant funding to CCSSO for TPA to accomplish this first-of-a kind national scan of teacher preparation accountability data policies. Saroja Barnes, Director of NTEP, has continued to extend to TPA the opportunity to engage the 14 NTEP states to "test" the utility of our policy work in advancing teacher preparation licensure and program approval at the state level. The authors thank these individuals and their organizations for their vision, generosity, and trust in TPA to conduct the analysis and produce this report.

We trust this report will be of interest and useful to teacher educators, state education officials (specifically those dealing with educator preparation program accountability and approval), and education policymakers. The report is an attempt to serve as a first look at the kinds of data and measures that each state requires their teacher preparation programs to report out as part of the program review process.

Although the authors, in working with CCSSO, have attempted to make the report as reader-friendly as possible, the report deals with technical issues related to state accountability policies and practices, which are not the grist of everyday language. To ease readability, we avoided the use of in-text references, choosing instead to list our sources at the end of the narrative section of the report. The Appendices contain much of the essential and more technical information that provided the basis of the report narrative and which the reader is encouraged to read.

TPA engaged the very capable services of APA Consultants, based in Denver, who conducted the bulk of the state policy analysis. TPA is grateful to Bob Palaich, CEO of APA, for supporting the collaboration with TPA and to the very skilled and knowledgeable Robert Reichardt, who lead the APA team of researchers in this work: Eric Eagon, Jack Hill, Jennifer Kramer-Wine, Amber Minogue, Kathryn Rooney, and Yilan Shen. Cathy Walker, administrative assistant to TPA, was helpful in so many ways from the beginning to completion of the project. This was truly a team effort. Finally, it would have been impossible to construct many of the state profiles without the enormously patient and helpful state contacts who helped the research team and authors better understand the specifics of their states' current efforts and who verified the summaries contained in the report.

In spite of all of the assistance received from others, the authors know that there may be weaknesses and errors in the report that remain uncorrected. The authors take full responsibility for those and any other shortcomings, as well as for any opinions and points of view expressed in the report narrative.

SOURCES

- Allen, M., Coble, C., & Crowe, E. *Building an Evidenced-Based System for Teacher Preparation*. Council for the Accreditation of Educator Preparation (CAEP). Website. Accessed 5.1.16 at <http://caepnet.org/knowledge-center?page=2>.
- Cantrell, S. & Kane, T. J. (2013). *Ensuring Fair and Reliable Measures of Effective Teaching: Culminating Findings from the MET Project's Three-Year Study*. Seattle, WA: Bill and Melinda Gates Foundation. Retrieved 4-20-14 from http://www.metproject.org/downloads/MET_Ensuring_Fair_and_Reliable_Measures_Practitioner_Brief.pdf.
- Coble, C., Crowe, E., & Allen, M. (2016). *CAEP Standard 3.2 Research, Study and Analysis A Report to the Council for the Accreditation of Educator Preparation*. Retrieved 5-10-16 from <http://www.caepnet.org/standards/standard-3> [See "CAEP STANDARD 3.2 Research, Study and Analysis"].
- Council for the Accreditation of Educator Preparation (CAEP). (2013). *CAEP Accreditation Standards*. Washington, DC: CAEP. Retrieved 4-20-14 from <http://caepnet.org/accreditation/standards/>.
- Council of Chief State School Officers. (2012). *Our Responsibility, Our Promise: Transforming Educator Preparation and Entry into the Profession*. Washington, DC: CCSSO.
- Crowe, E. (2010). *Measuring What Matters: A Stronger Accountability Model for Teacher Education*. Washington, DC: Center for American Progress. Retrieved 4-20-14 from http://www.americanprogress.org/issues/2010/07/pdf/teacher_accountability.pdf.
- Duncan, A. (2009). *The Secretary's Sixth Annual Report on A Highly Qualified Teacher in Every Classroom*. Washington, DC: US Department of Education.
- edTPA. Website. Accessed 6-7-14 at <http://edtpa.aacte.org/>.
- Ehlert, M. Koedel, C. Parsons, E., & Podgursky, M. (2013). "The Sensitivity of Value-Added Estimates to Specification Adjustments: Evidence from School- and Teacher-Level Models in Missouri." *Statistics and Public Policy*, 1, 19-27.
- ETS National Observational Teaching Exam (NOTE). Website. Accessed 5.1.16 at <http://www.ets.org/note/>.
- ETS Practice Performance Assessments Website. Accessed 5.1.16 at <http://www.ets.org/ppa/test-takers/teachers/about>.
- Greenberg, J., McKee, A., & Walsh, K. (2013). *Teacher Prep Review*. Revised December 2013. Washington, DC: National Council on Teacher Quality. Website. Accessed 5-2-14 at http://www.nctq.org/dmsView/Teacher_Prep_Review_2013_Report.
- Henry, G. T., Fortner, C. K., & Bastian, K. C. (2012). "The Effects of Experience and Attrition for Novice High School Science and Mathematics Teachers." *Science*: 335, 1118-21.

- Henry, G.T., Thompson, C.L., Bastian, K.C., et al. (2011, July). UNC Teacher Preparation Program Effectiveness Report. Chapel Hill, NC: Carolina Institute for Public Policy.
- Kane, T. J., & Staiger, D.O. (2012). *Gathering Feedback for Teaching*. Seattle, WA: Bill & Melinda Gates Foundation.
- Levine, A. (2006). *Educating School Teachers*. Washington, DC: The Education Schools Project. Website. Accessed 4-13-14 at http://www.edschools.org/teacher_report.htm.
- Missouri Educator Profile – MEP. Website. Accessed 5.1.16 at http://www.mo.nesinc.com/TestView.aspx?f=HTML_FRAG/MO006_TestPage.html.
- Mitchell, K.A., Robinson, D.Z., Plake, B.S., & Knowles, K.T. (eds.) (2001). *Testing Teacher Candidates: The Role of Licensure Tests in Improving Teacher Quality*. Washington, DC: National Academies Press.
- National Association of State Directors of Teacher Education and Certification. Project: Interstate Data Sharing. Website. Accessed 4-20-14 at: <http://www.nasdtec.net/?page=InterstateDataShare>.
- National Council on Teacher Quality. 2014 Teacher Prep Review. Website. Accessed 5.1.16 at: http://www.nctq.org/dmsStage/Teacher_Prep_Review_2014_Report.
- National Research Council. (2010). *Preparing Teachers: Building Evidence for Sound Policy*. Washington, DC: The National Academies Press. Website. Accessed 4-20-14 at: http://www.nap.edu/catalog.php?record_id=12882.
- Pearson. Teacher Licensure Testing and Performance Assessment. Website. Accessed 6-7-14 at: <http://www.pearsonassessments.com/teacherlicensure.html>.
- TRIPOD. Website. Accessed 5.1.16 at: <http://tripoded.com/teachers/>.

APPENDIX A

TPA KEY EFFECTIVENESS INDICATORS

Assessment Categories	Key Indicators	Measures
I Candidate Selection Profile	Academic Strength	PRIOR ACHIEVEMENT —(1) <i>For Undergraduate Programs</i> : Non-education course GPA required for program admission. Mean and range of high school GPA percentile (or class rank) for-candidates admitted as freshmen. Mean and tercile distribution of candidates' SAT/ACT scores. GPA in major and overall required for program completion. Average percentile rank of completers' GPA in their major at the university, by cohort. —(2) <i>For Post-Baccalaureate Programs</i> : Mean and range of candidates' college GPA percentile and mean and tercile distribution of GRE scores TEST PERFORMANCE — <i>For All Programs</i> : Mean and tercile distribution of admitted candidate scores on rigorous national test of college sophomore-level general knowledge and reasoning skills
	Teaching Promise	ATTITUDES, VALUES, AND BEHAVIORS SCREEN —Percent of accepted program candidates whose score on a rigorous and validated "fitness for teaching" assessment demonstrates a strong promise for teaching
	Candidate/Completer Diversity	DISAGGREGATED COMPLETIONS COMPARED TO ADMISSIONS —Number & percent of completers in newest graduating cohort AND number and percent of candidates originally admitted in that same cohort: overall and by race/ethnicity, age, and gender
II Knowledge and Skills for Teaching	Content Knowledge	CONTENT KNOWLEDGE TEST —Program completer mean score, tercile distribution, and pass rate on rigorous and validated nationally normed assessment of college-level content knowledge used for initial licensure
	Pedagogical Content Knowledge	PEDAGOGICAL CONTENT KNOWLEDGE TEST —Program completer mean score, tercile distribution, and pass rate on rigorous and validated nationally normed assessment of comprehensive pedagogical content knowledge used for initial licensure
	Teaching Skill	TEACHING SKILL PERFORMANCE TEST —Program completer mean score, tercile distribution, and pass rate on rigorous and validated nationally normed assessment of demonstrated teaching skill used for initial licensure
	Completer Rating of Program	EXIT AND FIRST YEAR COMPLETER SURVEY ON PREPARATION —State- or nationally-developed program completer survey of teaching preparedness and program quality, by cohort, upon program (including alternate route) completion and at end of first year of full-time teaching
III Performance as Classroom Teachers	Impact on K-12 Student Learning	TEACHER ASSESSMENTS BASED ON STUDENT LEARNING —Assessment of program completers or alternate route candidates during their first three years of full-time teaching using valid and rigorous student-learning driven measures, including value-added and other statewide comparative evidence of K-12 student growth overall and in low-income and low-performing schools
	Demonstrated Teaching Skill	ASSESSMENTS OF TEACHING SKILL —Annual assessment based on observations of program completers' or alternate route candidates' first three years of full-time classroom teaching, using valid, reliable, and rigorous statewide instruments and protocols
	K-12 Student Perceptions	STUDENT SURVEYS ON TEACHING PRACTICE —K-12 student surveys about completers' or alternate route candidates' teaching practice during first three years of full-time teaching, using valid and reliable statewide instruments
IV Contribution to State Needs	Entry and Persistence in Teaching	TEACHING EMPLOYMENT AND PERSISTENCE —(1) Percent of completers or alternate route candidates, by cohort and gender – race-ethnicity, employed and persisting in teaching years 1-5 after program completion or initial alternate route placement, in-state and out-of-state —(2) Percent of completers attaining a second stage teaching license in states with multi-tiered licensure
	Placement/Persistence in High-Need Subjects/Schools	HIGH-NEED EMPLOYMENT AND PERSISTENCE —Number & percent of completers or alternate route candidates, by cohort, employed and persisting in teaching in low-performing, low-income, or remote rural schools or in high need subjects years 1-5 after program completion or initial alternate route placement, in-state and out-of-state

APPENDIX B

TPA 50-STATE POLICY PROTOCOL

TPA 50-State Policy Protocol

11-3-15

1. Who conducted review?
 - 1a. State
 - 1b. Date
2. Source(s) of data

Enter the names of online or written sources, and URLs where possible, of the information entered on this data form (e.g., CAEP Program Impact Data presentation, ECS Teacher Preparation Data base, state DOE website, state code, etc.).
- 2a. Respondent #1 Information

Enter name, agency, title, email, phone number
- 2b. Respondent #2 Information
- 2c. Respondent #3 Information
3. Scope of state preparation program review and approval (select all that apply)?
 - a. State university-affiliated programs
 - b. County/city university-affiliated programs (example: City University of New York)
 - c. Private university-affiliated programs
 - d. State non-university-affiliated programs
 - e. County/city/district non-university affiliated programs
 - f. Private non-university-affiliated programs
 - g. Online programs run by out-of-state providers
 - h. Other:
 - i. Notes on Question 3
4. Frequency of state's normal review of established programs (i.e., for programs approved without conditions) (select one)?
 - a. Annually
 - b. Every 2 – 3 years
 - c. Every 4 – 6 years
 - d. Every 7 years
 - e. Less frequent than every 7 years
 - f. Other:
 - g. Notes on Question 4
5. More frequent review for "watch list" programs?
 - a. Yes
 - b. No
 - c. Other:
 - d. Notes on Question 5

6. State-required national (CAEP) accreditation (select one)?
 - a. Not required by state of either public or private programs
 - b. Not required, but all programs must meet CAEP standards
 - c. Required of state programs only
 - d. Required of all university-based programs in the state (public and private)
 - e. Required of all university-based and non-university-based programs other than out-of-state providers
 - f. Required of all university-based and non-university based programs including out-of-state providers
 - g. Other:
 - h. Notes on Question 6
7. Specialized Professional Association (SPA) approval (example: there is one for Elementary Education, Sciences (NSTA)) (select one)?
 - a. Required of all nationally (CAEP) accredited teacher preparation programs (TPPs) as a condition of state program approval
 - b. Required of all of SPA-eligible programs as a condition of state program approval, even if education program provider (EPP) not nationally accredited
 - c. All SPA-eligible programs must meet applicable SPA standards but don't require formal SPA approval
 - d. Not required of any programs
- 8.1. Annual HEA Title II Report (select all that apply)?
 - a. State posts annual federally required HEA Title II report on teacher preparation on state website
 - b. State posts link on state website to USDE annual Title II report
 - c. State uses Title II report or parts of report for state program approval or accountability purposes
- 8.2. Annual or biannual Teacher Preparation Program Productivity Report (select one)? (other than annual Title II report—if yes provide more detail on 9)
 - a. Yes (continue to 8.2a)
 - b. No (go to question 9)
 - c. Notes on use of the Title II report
9. Regularly required TPP performance report (select all that apply)?

This is an annually or biannually updated report/dashboard/accessible database provided to programs and/or the public that tracks indicators other than Title II, example value added scores in TN, or # of teachers in fields):

 - a. State requires annual or biannual TPP performance report other than Title II report and report other than annual or biannual preparation program productivity report.
 - b. Annual or biannual TPP performance report provided for public information
 - c. Annual or biannual TPP performance report intended for program improvement purposes
 - d. Annual or biannual TPP performance report used by state officials to screen annually for possible program deficiencies
 - e. Annual or biannual TPP performance report used as evidence in multi-year state program approval decisions
 - f. Other:
 - g. Notes on Question 9

10. Regular performance report on school leader/administrative preparation programs (select one)? Enter descriptions of planned policy overhauls in the notes section for this question. This should include links to any documents describing the policy changes.
 - a. State has implemented (at least partially) an annual or biannual LPP performance report along the lines of its TPP performance report
 - b. State is currently developing an annual or biannual LPP report
 - c. Other:
 - d. Notes on Question 10
11. Is candidate academic strength currently used or slated to be used as a measure of program performance on annual or biannual reports?
 - a. If NO Go to question 12
 - b. Yes (continue to 11a)
 - c. No (go to 12)
 - d. Other:
 - e. Notes on use of Candidate Academic Strength: Question 11
- 11a. Implementation status of the use of academic strength of program candidates as part of TPP performance report (select one)?
 - a. Fully implemented (all intended programs and subjects)
 - b. Partially implemented (or in pilot)—Put full implementation date below
 - c. Adopted in policy but not yet implemented
 - d. Other:
 - e. Notes about reporting candidate strength. If partially implemented (ii above) what is the implementation target date
- 11b. Candidate population reported on for strength of program candidates (select all that apply)?
 - a. Most recent entering cohort
 - b. Most recent completer cohort
 - c. Current group of candidates enrolled in program
 - d. All enrolled students during reporting period
 - e. Other:
 - f. Notes about candidate population reported for strength of program candidates
- 11c. Reporting level on strength of candidates (select all that apply)?
 - a. Individual certification/endorsement programs where data sufficient
 - b. Education program provider as a whole (data on all individual programs combined)
 - c. Endorsement area/subject
 - d. Other:
 - e. Notes on Reporting level on strength of candidates
- 11d. Data use (select all that apply)?
 - a. High School GPA
 - b. High school class rank
 - c. College GPA in all courses
 - d. College GPA in general distribution courses (required general education courses)
 - e. College GPA in content major
 - f. College GPA for last year or last 2 years of coursework
 - g. College class rank

- h. SAT scores
- i. ACT scores
- j. GRE scores
- k. MAT scores
- l. Praxis Core (Praxis 1) scores
- m. Other:
- n. Notes about data used to measure strength of candidates (11d)

11e. Measures used to evaluate program and/or provider performance (select all that apply)?

- a. Program cohort average score(s)
- b. National percentile average(s) for cohort
- c. Cohort score distribution(s) (e.g., percentage of cohort in each tercile of national percentile distribution or average score of each program tercile)
- d. Comparison with statewide average scores
- e. Minimum state-established program admissions score benchmarks for candidates-- enter these in following text box
- f. Other:
- g. Minimum state-established benchmarks (if checked in 11.e.v. above)
- h. Additional notes on measures of strength of program candidates

12. Is teaching promise or non-academic competencies of program candidates currently used or slated to be used as a measure of program performance in annual or biannual performance reports? (Expect very few yes answers.)

- a. Yes (continue to 12a)
- b. No (go to 12f)
- c. Other:
- d. Notes on teaching promise and other non-academic measures question 12

12a. Implementation status of use of teaching promise or non-academic competencies of program candidates as part of TPP performance report (select one)?

- a. Fully implemented (all intended programs and subjects)
- b. Partially implemented (or in pilot) --Put full implementation date below
- c. Adopted in policy but not yet implemented
- d. Other:
- e. If partially implemented (ii above) what is the implementation target date

12b. Candidate population teaching promise or non-academic competencies of program candidates (select all that apply)?

- a. Most recent entering cohort
- b. Most recent completer cohort
- c. All enrolled students during reporting period
- d. Other:
- e. Notes about candidate population for teaching promise or non-academic competencies of program candidates (12b above)

- 12c. Reporting level on candidate teaching promise or non-academic competencies (select all that apply)?
- a. Individual certification/endorsement programs where data is sufficient
 - b. Education program provider as a whole (data on all individual programs combined)
 - c. Endorsement area/subject
 - d. Other:
 - e. Notes about reporting level on teaching promise or non-academic competencies of program candidates (12c above).
- 12d. Sources of data (select all that apply)?
- a. State-or national-normed assessment
 - b. Locally developed assessment (e.g. district)
 - c. Candidate interview
 - d. Other:
 - e. Notes about sources of data strength of program candidates (12d above).
- 12e. Measures used to evaluate program and/or provider performance (select all that apply)?
- a. Program cohort average score(s)
 - b. Cohort score distribution(s) (e.g., percentage of cohort in each tercile of state or national percentile distribution, or average score of each program tercile)-
 - c. Minimum state-established program admissions score benchmark for candidates--list benchmark in notes below
 - d. Other:
 - e. Notes about measures of candidate teaching promise or other non-academic skill including benchmarks used (12e above)
- 12f. Other state-required use of measures of candidate teaching promise or other non-academic competencies besides or in addition to assessment of preparation program performance (select all that apply)?
- a. None
 - b. Program admission criterion for individual candidates
 - c. Basis for individual candidate individual development plan (IDP)
 - d. Other:
- 12g. Any additional notes about use of measures of teaching promise or other non-academic skills.
13. Are candidate completion rates currently used or slated to be used as a measure of program performance in annual or biannual reports?
- If collected but used for a purpose other than for program evaluation check "Other" and explain the use of these data in the "Notes" box below. If the data are not collected, at all, check NO and go to question 14."
- a. Yes (continue to 13a)
 - b. No (go to question 14)
 - c. Other:
 - d. Notes on Candidate Completion Rates: Question 13

- 13a. Implementation status of use of completion rates of program candidates as part of annual or biannual performance report (select one)?
- Fully implemented (all intended programs and subjects)
 - Partially implemented (or in pilot) --Put full implementation date below
 - Adopted in policy but not yet implemented
 - Other:
- 13b. Candidate population program completion rates are reported (select all that apply)?
- Most recent entering cohort
 - Most recent completer cohort
 - All enrolled students during reporting period
 - Other:
 - Notes on candidate population used for completion rate data (13b)
- 13c. Reporting level on candidate completion rate (SELECT all that apply)?
- Individual certification/endorsement programs where data is sufficient
 - Education program provider as a whole (data on all individual programs combined)
 - Endorsement area/subject
 - Other:
 - Notes on reporting level for completion rate data (13c)
- 13d. Definition of program completion rate includes (select all that apply)
- Program completion rate of entire entering cohort
 - Program completion rate of candidates recommended for student teaching
 - Licensure examination pass rate of completing cohort
 - Program completion rate (however defined) disaggregated by gender and race/ethnicity
 - Minimum state-established program completion percentage benchmark(s) (list benchmarks in following textbox)
 - Other:
 - Notes on completion rate measures used to evaluate programs and if appropriate benchmarks used (13d)
- 13e. Is the measure disaggregated by demographic factors?
- No
 - Gender
 - Race and ethnicity
 - Other:
- 13f. Additional notes on candidate completion rate.
14. Is candidate content mastery currently used or slated to be used as a measure of program performance in annual or biannual reports? This is expected to be the Praxis subject assessment (formerly Praxis II), a version of the Pearson NES (which may be renamed by the state), or a state developed assessment. Based on state-required licensure examination of teaching content.
- Yes (continue to 14a)
 - No (go to 15)
 - Other:
 - Notes on Content Mastery. Question 14

- 14a. Implementation status of use of program candidate content mastery as part of annual or biannual performance report (select one)?
- a. Fully implemented (all intended programs and subjects)
 - a. Partially implemented (or in pilot) --Put full implementation date below
 - b. Adopted in policy but not yet implemented
 - c. Other:
- 14b. Candidate population program content mastery is reported on (select all that apply)?
- a. Most recent entering cohort
 - a. Most recent completer cohort
 - b. Other:
 - c. Notes on candidate population reported on for content mastery (14b)
- 14c. Reporting level on candidate content mastery (select all that apply)?
- a. Individual certification/endorsement programs where data sufficient
 - b. Education program provider as a whole (data on all individual programs combined)
 - c. Endorsement area/subject
 - d. Other:
 - e. Notes on reporting level for content mastery (14c)
- 14d. Assessment(s) required to measure content mastery (select all that apply)?
- a. Praxis Subject Assessments (formerly Praxis II)
 - b. Pearson NES--states might have their own names
 - c. Other:
 - d. Notes on assessments used to measure content mastery (14d)
- 14e. Measures used to evaluate program or provider performance (select all that apply)?
- a. Pass rate percentage of candidates--passing level is a benchmark (list required pass rates in notes text box below)
 - b. Program cohort average exam score(s)
 - c. Distribution: state or national percentile average(s) for cohort
 - d. Cohort score distribution(s) (e.g., percentage of cohort in each tercile of state or national percentile distribution on assessment(s) used, or average score of each program cohort tercile)
 - e. Other:
 - f. Notes on measures used to evaluate program performance AND required pass rates if appropriate (14e)
- 14f. Additional notes on content mastery information used in annual or biannual reports.
15. Is candidate pedagogical skill and general professional knowledge used or slated to be used as a measure of program performance in annual or biannual reports? Based on state licensure examination or state-required culminating program assessment. Could be paper and pencil test, performance assessment and/or portfolio assessment. We expect these assessments to be the edTPA, Praxis Performance Assessment for Teachers (PPAT), Principles of Learning and Teaching (PLT), Praxis Teaching Foundation Test (PLT), Assessment of Professional Knowledge (NES) or a state developed assessment.
- a. Yes (continue to 15a)
 - b. No (go to question 16)
 - c. Other:
 - d. Notes on candidate pedagogical skill and general professional knowledge: question 15

- 15a. Implementation status of use of program candidate pedagogical skill or general professional knowledge as part of annual or biannual performance report (select one)?
- a. Fully implemented (all intended programs and subjects)
 - b. Partially implemented (or in pilot) --Put full implementation date below
 - c. Adopted in policy but not yet implemented
 - d. Other:
- 15b. Candidate population program that pedagogical skill and general professional knowledge is reported on (select all that apply)?
- a. Most recent entering cohort
 - b. Most recent completer cohort
 - c. All enrolled students during reporting period
 - d. Other:
 - e. Notes on candidate population for reporting pedagogical skill and general professional knowledge (15b)
- 15c. Reporting level on candidate pedagogical skill and general professional knowledge (select all that apply)?
- a. Individual certification/endorsement programs where data is sufficient
 - b. Education program provider as a whole (data on all individual programs combined)
 - c. Endorsement area/subject
 - d. Other:
 - e. Notes about reporting level for candidate pedagogical skill and general professional knowledge (15c)
- 15d. State assessment(s) required (select all that apply)?
- a. edTPA
 - b. Praxis Performance Assessment for Teachers (PPAT)
 - c. Principles of Learning and Teaching (PLT)
 - d. Praxis Teaching Foundations Tests
 - e. Assessment of Professional Knowledge (NES)
 - f. Other:
 - g. Notes on assessments used to measure candidate pedagogical skill and general professional knowledge (15d)
- 15e. Measures used to evaluate program and/or provider performance (select all that apply)?
- a. Pass rate percentage of candidates--passing level is a benchmark
 - b. Program cohort average exam score(s)
 - c. State or national percentile average(s) for cohort
 - d. Cohort score distribution(s) (e.g., percentage of cohort in each tercile of state or national percentile distribution on assessment(s) used, or average score of each program cohort tercile)
 - e. Other:
 - f. Notes on measures of candidate pedagogical skill or general professional knowledge used in annual or biannual reports (15e)
- 15f. Additional notes on candidate pedagogical skill or general professional knowledge used in annual or biannual reports.

16. Is candidate pedagogical content knowledge (PKC) for teaching used or slated to be used as measure of program performance in annual or biannual reports? Based on state licensure examination or state-required culminating program assessment specifically focused on depth and breadth of content knowledge for teaching. There is no dedicated national assessment for this—only if they have a locally developed assessment (Praxis, edTPA and other assessments mentioned above do not count).
 - a. Yes (continue to 16a)
 - b. No (go to question 17)
 - c. Other:
 - d. Notes on Candidate Pedagogical Content Knowledge: Question 16
- 16a. Implementation status of use of program candidate pedagogical content knowledge as part of annual or biannual performance report (select one)?
 - a. Fully implemented (all intended programs and subjects)
 - b. Partially implemented (or in pilot) --Put full implementation date below
 - c. Adopted in policy but not yet implemented
 - d. Other:
- 16b. Candidate population program that candidate pedagogical content knowledge is reported on (select all that apply)?
 - a. Most recent entering cohort
 - b. Most recent completer cohort
 - c. All enrolled students during reporting period
 - d. Other:
 - e. Notes on candidate population for reporting pedagogical content knowledge (16b)
- 16c. Reporting level on candidate pedagogical content knowledge (select all that apply)?
 - a. Individual certification/endorsement programs where data sufficient
 - b. Education program provider as a whole (data on all individual programs combined)
 - c. Endorsement area/subject
 - d. Other:
 - e. Notes on reporting level of candidate pedagogical content knowledge (16c)
- 16d. Assessments required for measuring candidate pedagogical content knowledge (please identify)?
- 16e. Measures used to evaluate program and/or provider performance (select all that apply)?
 - a. Pass rate percentage of candidates
 - b. Program cohort average exam score(s)
 - c. State or national percentile average(s) for cohort
 - d. Cohort score distribution(s) (e.g., percentage of cohort in each tercile of state or national percentile distribution on assessment(s) used, or average score of each program cohort tercile)
 - e. Other:
 - f. Notes on pedagogical content knowledge measures used in annual or biannual reports (16e)
- 16f. Additional notes on pedagogical content knowledge used in annual or biannual reports?

17. Is program candidate and/or completer (as employed teacher) perceptions of preparation program effectiveness included as a measure of program performance in annual or biannual reports?
 - a. Yes (continue to 17a)
 - b. No (go to 18)
 - c. Other:
 - d. Notes on Perceptions of Preparation Program Effectiveness: Question 17
- 17a. Implementation status of use of candidate/completer perceptions of preparation program effectiveness as part of annual or biannual performance report (select one)?
 - a. Fully implemented (all intended programs and subjects)
 - b. Partially implemented (or in pilot)—Put full implementation date below
 - c. Adopted in policy but not yet implemented
 - d. Other:
- 17b. Reporting level (select all that apply)?
 - a. Individual certification/endorsement programs where data is sufficient
 - b. Education program provider as a whole (data on individual programs combined)
 - c. Endorsement area/subject
 - d. Other:
- 17c. Source(s) of data (select all that apply)?
 - a. Written or on-line survey
 - b. Formal oral interviews
 - c. Other:
- 17d. When is information on the perceptions collected (select one)?
 - a. Administered to candidates at or near program completion
 - b. Administered during completers' first year of full-time teaching
 - c. Administered during completers' second year of full-time teaching
 - d. Other:
- 17e. Measures used to evaluate program and/or provider performance (select all that apply)?
 - a. Overall program rating by respondents
 - b. Ratings on individual program features
 - c. Program rating(s) compared to benchmark or performance requirement (e.g., meets, exceeds, or fails to meet benchmark score(s))
 - d. Other:
- 17f. How is the data used (select all that apply)?
 - a. Candidate/completer program ratings used internally for program improvement
 - b. Candidate/completer program ratings used for state program performance assessment
 - c. Other:
- 17g. Is the data limited to completers who are working in-state (select one)?
 - a. Yes
 - b. No
 - c. Other:
- 17h. Additional notes on candidate/completer perceptions of program effectiveness

18. Is program completer (as employed teacher) composite scores on a required annual teacher performance assessment—used as a measure of program performance separate from single teacher assessment components articulated previously? For this particular program performance measure to be used, the state would need (1) to administer an annual performance assessment for all teachers and (2) to combine two or more of the measures it uses to evaluate teachers into a composite score, which then provides the basis for evaluating preparation programs. If any of the measures used in this teacher performance assessment are used individually for the evaluation of preparation programs, then questions about those individual measures should be answered in the appropriate questions below.
 - a. Yes (continue to 18a)
 - b. No (go to 19)
 - c. Other:
 - d. Notes on composite scores on a required teacher evaluation, question 18
- 18a. Implementation status of use of aggregate performance evaluation as part of annual or biannual performance report (select one)?
 - a. Fully implemented (all intended programs and subjects)
 - b. Partially implemented (or in pilot) --Put full implementation date below
 - c. Adopted in policy but not yet implemented
 - d. Other:
- 18b. Completer population program that aggregate performance evaluation is reported on (select all that apply)?
 - a. Completers in 1st year of employment as teacher
 - b. Completers in 2nd year of employment
 - c. Completers in 3rd year of employment
 - d. iv. Completers in 4th year of employment
 - e. Other:
- 18c. Limitations based on where completers work (select all that apply)?
 - a. Completers teaching in-state only
 - b. Completers working in public schools only
 - c. Other:
- 18d. Reporting level for aggregate performance evaluation (select all that apply)?
 - a. Individual certification/endorsement programs where data is sufficient
 - b. Education program provider as a whole (data on all individual programs combined)
 - c. Endorsement area/subject
 - d. Other:
 - e. Notes on reporting level for aggregate performance evaluation (18d)
- 18e. Sources of aggregate performance evaluation data (select all that apply)?
 - a. Observation of teacher's classroom performance that IS also reported individually (see measure 20)
 - b. Observation of teacher's classroom performance that IS NOT reported individually (in measure 20)
 - c. Teacher work portfolio
 - d. Teachers' measured impact on P-12 students' learning that IS also reported as an individual program measure (see measure 19)
 - e. Teachers' measured impact on P-12 students' learning that IS NOT also reported as an individual program measure (on measure 19)
 - f. Report from teachers' principal or supervisor that IS also reported as an individual program

- measure (see measure 21)
 - g. Report from teachers' principal or supervisor that IS NOT also reported as an individual program measure (in measure 21)
 - h. Teacher self-evaluation
 - i. P-12 pupil perceptions of their teachers that IS also reported as an individual program measure (see measure 22)
 - j. P-12 pupil perceptions of their teachers that IS also reported as an individual program measure (in measure 22)
 - k. Other:
- 18f. Scope of subjects for aggregate performance evaluation data used in annual or biannual reports (select all that apply)?
 - a. Teachers in all subjects
 - b. Teachers in limited subjects (list subjects in notes below)
 - c. Detail on limited subjects (18f)
- 18g. Scope of grades used for aggregate performance evaluation data used in annual or biannual reports (select all that apply)?
 - a. Pre-K
 - b. Kindergarten
 - c. 1
 - d. 2
 - e. 3
 - f. 4
 - g. 5
 - h. 6
 - i. 7
 - j. 8
 - k. 9
 - l. 10
 - m. 11
 - n. 12
 - o. Other:
- 18h. Aggregate performance evaluation measures reported in annual or biannual reports (select all that apply)?
 - a. Average teacher or program score, rating, or percentile rank by PROGRAM
 - b. Average or composite teacher or program score, rating, or percentile rank by INSTITUTION OR EPP (all programs in aggregate)
 - c. Distribution by program: e.g. percentage of teachers by PROGRAM meeting state-established performance benchmark (list benchmark in notes below)
 - d. Distribution by institution or EPP: e.g. percentage of teachers by INSTITUTION OR EPP (for all teachers or programs in aggregate) meeting state-established performance benchmark
 - e. Data are not used for program evaluation but given to program for improvement purposes only
 - f. Other:
 - g. Benchmarks used for aggregate performance evaluation data (18h)

- 18i. How are the performance evaluation measures used (select all that apply)?
- a. Data are used for program evaluation
 - b. Data are NOT used for program evaluation but given to program for improvement purposes only
 - c. Other:
- 18j. Additional notes on use of aggregate performance evaluation measures reported in annual or biannual reports.
19. Is program completer (as employed teacher) impact on P-12 student learning used as a measure of program performance in annual or biannual reports? This is usually a value-added or growth score.
- a. Yes
 - b. No
 - c. Other:
 - d. Notes on Program Completer Impact on P-12 Student Learning: Question 19
- 19a. Implementation status of use of completer impact on student learning as part of annual or biannual performance report (select one)?
- a. Fully implemented (all intended programs and subjects)
 - b. Partially implemented (or in pilot) --Put full implementation date below
 - c. Adopted in policy but not yet implemented
 - d. Other:
- 19b. Completer population program that impact on k-12 learning is reported on (select all that apply)?
- a. Completers in 1st year of employment as teacher
 - b. Completers in 2nd year of employment
 - c. Completers in 3rd year of employment
 - d. Completers in 4th year of employment
 - e. Completers in 5th year of employment
 - f. Other:
- 19c. Limitations based on where completers work (select all that apply)?
- a. Completers teaching in-state only
 - b. Completers working in public schools only
 - c. Other:
- 19d. Reporting levels for this measure (select all that apply)?
- a. Individual certification/endorsement programs where data is sufficient
 - b. Education program provider (EPP), with scores disaggregated by tested subjects (i.e., reflecting scores of completers from any program whose P-12 pupils (in all grades) were tested in one or more of those subjects)
 - c. EPP as a whole (aggregated data on all relevant programs)
 - d. Other:
- 19e. Source(s) of student data for this measure (select all that apply)/
- a. P-12 student scores on statewide or nationally normed assessment(s)
 - b. P-12 student scores on district-level assessment(s)
 - c. P-12 student scores on school -level or classroom teacher-developed Student Learning Objectives assessments
 - d. Other:

- 19f. Source(s) of teacher data for this measure (select all that apply)?
- a. Teacher effectiveness calculation derived from Student Growth Model
 - b. Teacher Value-Added Model
 - c. Teacher effectiveness calculation derived from Student Learning Objectives assessment
 - d. Other:
- 19g. Number of teacher cohorts used for this measure (select one)?
- a. 1
 - b. 2
 - c. 3
 - d. 4
 - e. 5+
 - f. Years
 - g. Notes on years of impact data used in annual or biannual reports (19g)
- 19h. Grade 3 assessment data used (select all that apply)?
- a. Not used
 - b. Reading/language arts
 - c. Mathematics
 - d. Other:
- 19i. Grades 4-8 assessment data used (select all that apply)?
- a. Not used
 - b. Reading/language arts
 - c. Mathematics
 - d. Science
 - e. Social Studies
 - f. Other:
- 19j. High school assessment data used (select all that apply)?
- a. Not used
 - b. English/Language arts
 - c. Mathematics
 - d. Science(s)
 - e. Social Studies
 - f. Other:
- 19k. Are measures of teacher impact (e.g. value-added) reported as averages (or other measure of central tendency e.g. median) (select all that apply)?
- a. Not reported as an average
 - b. Reported by program and tested subject
 - c. Reported by program, multiple tested subjects together
 - d. Reported by EPP and tested subject
 - e. Reported by EPP, multiple tested subjects together
 - f. Reported by school type (e.g. low performing) either by program or EPP and by tested subject
 - g. Reported by school type (e.g. low performing) either by program or EPP and for ALL subjects averaged together
 - h. Other:

- 19l. Are measures of teacher impact (e.g. value-added) reported as a distribution such as percent of teachers in each quartile) (select all that apply)?
- a. Not reported as a distribution
 - b. Reported by program and tested subject
 - c. Reported by program, multiple tested subjects together
 - d. Reported by EPP and tested subject
 - e. Reported by EPP, multiple tested subjects together
 - f. Reported by school type (e.g. low performing) either by program or EPP and by tested subject
 - g. Reported by school type (e.g. low performing) either by program or EPP and for ALL subjects averaged together
 - h. Other:
- 19m. Are measures of teacher impact (e.g. value-added) reported relative to a benchmark (e.g., percent below a state average) (select all that apply)?
- a. Not reported relative to a benchmark
 - b. Benchmarks are based on state averages
 - c. Other:
 - d. Notes on benchmarks used to measure teacher impact (19m)
- 19n. Additional notes on reporting of teacher impact on student learning.
20. Is observation-based performance a measure in annual or biannual reports?
- a. Yes (continue to 20a)
 - b. No (go to 21)
 - c. Other:
 - d. Notes on Program Completer Observation-based Performance: Question 20
- 20a. Implementation status of use of completer observation-based performance as part of annual or biannual performance report (select one)?
- a. Fully implemented (all intended programs and subjects)
 - b. Partially implemented (or in pilot)—Put full implementation date below
 - c. Adopted in policy but not yet implemented
 - d. Other:
- 20b. Completer population program that observation is reported on (select all that apply)?
- a. Completers in 1st year of employment as teacher
 - b. Completers in 2nd year of employment
 - c. Completers in 3rd year of employment
 - d. Completers in 4th year of employment
 - e. Completers in 5th year of employment
 - f. Other:
- 20c. Limitations based on where completers work (select all that apply)?
- a. Completers teaching in-state only
 - b. Completers working in public schools only
 - c. Other:

- 20d. Reporting level (select all that apply)?
- a. Individual TPP (e.g. certification or endorsement area) where data is sufficient
 - b. Institution or EPP as a whole with data on all individual programs combined
 - c. Other:
- 20e. Scope of grades included for observation data used in annual or biannual reports (select all that apply)?
- a. All
 - b. Pre-K
 - c. Kindergarten
 - d. 1
 - e. 2
 - f. 3
 - g. 4
 - h. 5
 - i. 6
 - j. 7
 - k. 8
 - l. 9
 - m. 10
 - n. 11
 - o. 12
 - p. Other:
- 20f. Aggregate observations measures reported in annual or biannual reports (select all that apply)?
- a. Average teacher or program score, rating, or percentile rank
 - b. Distributional by program: e.g. percentage of teachers meeting state-established performance benchmark (list benchmark in notes below)
 - c. Other:
 - d. Benchmarks used for aggregate observation data (20f)
- 20g. Is data reported by program or institution (EPP) (select one)?
- a. Program
 - b. Institution or EPP
 - c. Other:
- 20h. How is the data used? (select one)
- a. Program accountability
 - b. Program improvement only
 - c. Other:
- 20i. Additional notes on observational data reported in annual or biannual reports.
21. Is data from a school principal and/or supervisor perceptions of the strength of new teachers' pre-service preparation based on their teaching performance included as a measure of program performance in annual or biannual reports (not based on performance evaluation)?
- a. Yes (continue to 21a)
 - b. No (go to 22)
 - c. Other:
 - d. Notes on Supervisor Perceptions of Teachers Pre-Service Preparation: Question 21

- 21a. Implementation status of use of supervisor perception of preparation as part of annual or biannual performance report (select one)?
- a. Fully implemented (all intended programs and subjects)
 - b. Partially implemented (or in pilot) – Put full implementation date below
 - c. Adopted in policy but not yet implemented
 - d. Other:
- 21b. Completer population reported on (select all that apply)?
- a. First year of employment
 - b. Second year of employment
 - c. Third year of employment
 - d. Fourth year of employment
 - e. Other:
- 21c. Limitations based on where completers work (select all that apply)?
- a. Completers teaching in-state only
 - b. Completers working in public schools only
 - c. Other:
- 21d. Reporting level (select all that apply)?
- a. Individual TPP (e.g. certification or endorsement area) where data is sufficient
 - b. Institution or EPP as a whole with data on all individual programs combined
 - c. Other:
- 21e. Survey instruments used (select all that apply)? Do not expect a national survey.
- a. Nationally developed, state-administered survey
 - b. State-developed and administered survey
 - c. Locally-developed and administered survey
 - d. Other:
 - e. 21e cont. Name or describe the survey instrument used
- 21f. Main focus of survey (select one)?
- a. Strengths and weaknesses of individual teachers
 - b. Strengths and weaknesses in new teachers' preparation (and, by implication, in their preparation program)
 - c. Other:
- 21g. Scope of data (select one)?
- a. All Teachers
 - b. Other:
- 21h. Measures used to evaluate program and/or provider performance (select all that apply)?
- a. Average by program
 - b. Average by institution or EPP
 - c. Benchmark based on state average
 - d. Other benchmark (describe in text box below)
 - e. Data are not reported publicly but provided to programs
 - f. Other:
 - g. Additional information on supervisor perceptions measures reported including benchmarks (21h)

- 21i. Are P-12 supervisor perceptions of their teachers reported as an independent measure of program performance or used only in combination with other measures (select one)?
 - a. Reported independently
 - b. Combined with other measures
 - c. Other:
- 21j. Additional notes on supervisory perceptions measures used in annual or biannual reports.
- 22. Pupil perceptions of their teacher's effectiveness is reported.
- 22. Are measures of pupil perceptions of their teacher's effectiveness used in annual or biannual reports on EPP (P-12 students—do not expect many)?
 - a. Yes (continue to 22a)
 - b. No (go to 23)
 - c. Other:
 - d. Notes on pupil perceptions of their teachers' effectiveness: question 22
- 22a. Implementation status of use of pupil perception as part of annual or biannual performance report (select one)?
 - a. Fully implemented (all intended programs and subjects)
 - b. Partially implemented (or in pilot)—Put full implementation date below
 - c. Adopted in policy but not yet implemented
 - d. Other:
- 22b. Completer population reported on (select all that apply)?
 - a. First year of employment
 - b. Second year of employment
 - c. Third year of employment
 - d. fourth year of employment
 - e. Other:
- 22c. Limitations based on where completers work (select all that apply)?
 - a. Completers teaching in-state only
 - b. Completers working in public schools only
 - c. Other:
- 22d. Reporting level (select all that apply)?
 - a. Individual TPP (e.g. certification or endorsement area) where data is sufficient
 - b. Institution or EPP as a whole with data on all individual programs combined
 - c. Other:
- 22d. Reporting level (select all that apply)?
 - a. Individual TPP (e.g. certification or endorsement area) where data is sufficient
 - b. Institution or EPP as a whole with data on all individual programs combined
 - c. Other:
- 22e. Survey instrument used (select all that apply)?
 - a. Nationally developed, state-administered survey
 - b. State-developed and administered survey
 - c. Locally-developed and administered survey
 - d. Other:
 - e. Name or describe the survey instrument used.

- 22f. Scope of subjects taught for pupil perception measures (select one)?
- a. Teachers in all subjects
 - b. Other:
- 22g. Scope of grades used for observation data used in annual or biannual reports (select all that apply)?
- a. All
 - b. Pre-K
 - c. Kindergarten
 - d. 1
 - e. 2
 - f. 3
 - g. 4
 - h. 5
 - i. 6
 - j. 7
 - k. 8
 - l. 9
 - m. 10
 - n. 11
 - o. 12
 - p. Other:
- 22h. Measures used to evaluate program and/or provider performance (select all that apply)?
- a. Average by program
 - b. Average by EPP
 - c. Distribution at each level of the perceptions scale
 - d. Benchmark based on state average
 - e. Other benchmark (describe in text box below)
 - f. Data are not reported publicly but provided to programs
 - g. Other:
- 22i. Are P-12 pupils' perceptions of their teachers reported as an independent measure of program performance or used only in combination with other measures (select one)?
- a. Reported independent
 - b. Combined with other measures
 - c. Other:
- 22j. Additional notes on use of pupil perceptions of teachers in annual or biannual reports.
23. Is the placement of completers used in annual or biannual reports?
- a. Yes (continue to 23a)
 - b. No (go to 24)
 - c. Other:
 - d. Notes on Placement Data: Question 23

- 23a. Is placement combined with persistence as one measure?
An example of this combination of placement and persistence would be the percent of 2013 completers working in 2015. Note that a state can report both a combined measure (23a) and a measure of placement only (23i)
- Yes (continue to 23b)
 - No (go to 23i)
 - Other:
- 23a. Implementation status of use of completer placement/persistence as part of annual or biannual performance report (select one)?
- Fully implemented (all intended programs and subjects)
 - Partially implemented (or in pilot) --Put full implementation date below
 - Adopted in policy but not yet implemented
 - Other:
- 23b Time period reported for placement/persistence (select all that apply)?
- 1st year after completion
 - 2nd year after completion
 - 3rd year after completion
 - 4th year after completion
 - 5th year after completion
 - Other:
- 23c. Measure used to report placement/persistence (select all that apply)?
- Proportion who are working at a given time
 - Proportion that has been placed over a period of time (e.g. percent who worked 3 out of 4 years)
 - Other:
- 23d. Additional information on measures used (select all that apply)?
- Use 1-year average over multiple cohorts (describe in notes below)
 - Use two-year placement over multiple cohorts
 - Other:
 - 23b-d Additional information on measures used
- 23e. Limitations based on where completers work (select all that apply)?
- Completers teaching in-state only
 - Completers working in public schools only
 - Other:
- 23f. Reporting level (select all that apply)?
- Individual TPP (e.g. certification or endorsement area) where data is sufficient
 - Institution or EPP as a whole with data on all individual programs combined
 - Other:
- 23g. Measures used to evaluate programs?
- Benchmark based on state average
 - Other state established benchmark
 - Dis-aggregation by race/ethnicity
 - Other:

- 23h. Additional information on how measures are evaluated including benchmarks.
- 23i. Implementation status of use of completer placement as part of annual or biannual performance report? If placement is combined with persistence and you completed 23a-h, states might also report this information. This is for measures of placement rates only. An example of this measure is proportion of a completer cohort that worked any of 3 years after completion.
- a. Fully implemented (all intended programs and subjects)
 - b. Partially implemented (or in pilot)—Put full implementation date below
 - c. Adopted in policy but not yet implemented
 - d. Other:
- 23j Time period reported for placement (select all that apply)?
- a. 1st year after completion
 - b. 2nd year after completion
 - c. 3rd year after completion
 - d. 4th year after completion
 - e. 5th year after completion
 - f. Other:
- 23k. Measure used to report placement (select all that apply)?
- a. Proportion who are placed over a given time: e.g. Percent who were placed within 3 years of completion.
 - b. Other:
- 23l. Measure used to report placement (select all that apply)?
- a. Proportion who are placed over a given time: e.g. Percent who were placed within 3 years of completion.
 - b. Other:
- 23m. Additional information on measures used (select all that apply)?
- a. Use 1-year average over multiple cohorts (describe in notes below)
 - b. Use two-year placement over multiple cohorts
 - c. Other:
 - d. 23i-m, additional information on measures used
- 23n. Limitations based on where completers work (select all that apply)?
- a. Completers teaching in-state only
 - b. Completers working in public schools only
 - c. Other:
- 23o. Reporting level (select all that apply)?
- a. Individual TPP (e.g. certification or endorsement area) where data is sufficient
 - b. Institution or EPP as a whole with data on all individual programs combined
 - c. Other:
- 23p. Measures used to evaluate programs (select one)?
- a. Benchmark based on state average
 - b. Other state established benchmark
 - c. Dis-aggregation by race/ethnicity
 - d. Other:

- 23q. Additional information on how measures are evaluated including benchmarks.
24. Is persistence of completers in teaching reported in annual or biannual reports?
- a. Yes (continue to 24a)
 - b. No (go to 25)
 - c. Other:
 - d. Notes on Persistence of Program Completers: Question 24
- 24a. Implementation status of use of completer placement as part of annual or biannual performance report (select one)?
- a. Fully implemented (all intended programs and subjects)
 - b. Partially implemented (or in pilot)—Put full implementation date below
 - c. Adopted in policy but not yet implemented
 - d. Other:
- 24bi. Measures reported on persistence: 1 year after completion (select all that apply)?
- a. Retained in a school
 - b. Retained in a district
 - c. Retained in the state
 - d. Retained in or out of state in education
 - e. Other:
- 24bii. Measures reported on persistence: 2 years after completion (select all that apply)?
- a. Retained in a school
 - b. Retained in a district
 - c. Retained in the state
 - d. Retained in or out of state in education
 - e. Other:
- 24biii. Measures reported on persistence: 3 years after completion (select all that apply)?
- a. Retained in a school
 - b. Retained in a district
 - c. Retained in the state
 - d. Retained in or out of state in education
 - e. Other:
- 24biv. Measures reported on persistence: 4 years after completion (select all that apply)?
- a. Retained in a school
 - b. Retained in a district
 - c. Retained in the state
 - d. Retained in or out of state in education
 - e. Other:
- 24bv. Measures reported on persistence: 5 years after completion (select all that apply)?
- a. Retained in a school
 - b. Retained in a district
 - c. Retained in the state
 - d. Retained in or out of state in education
 - e. Other:

- 24bvi. Measures reported on persistence: More than 5 years after completion (select all that apply)?
- a. Retained in a school
 - b. Retained in a district
 - c. Retained in the state
 - d. Retained in or out of state in education
 - e. Other:
 - f. Additional notes on persistence measures reported (24b)
- 24c. Limitations based on where completers work (select all that apply)?
- a. Completers teaching in-state only
 - b. Completers working in public schools only
 - c. Other:
- 24d. Reporting level (select all that apply)?
- a. Individual TPP (e.g. certification or endorsement area) where data is sufficient
 - b. Institution or EPP as a whole with data on all individual programs combined
 - c. Other:
- 24e. Measures used to evaluate programs (select all that apply)?
- a. Percentage of completers still working in education
 - b. Percentage of completers working in education after X years (describe X in notes below)
 - c. Average years employed for cohorts reported
 - d. Percent who have achieved a second stage license
 - e. Other:
- 24f. Measures used in program evaluation (select all that apply)?
- a. Comparative performance of subgroups (gender, race/ethnicity)
 - b. Comparison to state average benchmark
 - c. Comparison to some other benchmark
 - d. Other:
 - e. Additional information on criteria used to describe measures including benchmarks (24f)
- 24g. Additional notes on persistence measures used to report on programs.
25. Is placement of completers in high needs schools a reported measure? High needs schools are often defined by proportion of free and reduced lunch eligible students, possibly Title 1.
- a. Yes (continue to 25a)
 - b. No (go to 26)
 - c. Other:
 - d. Notes on placement of completers in high needs schools: question 25
- 25a. Implementation status of use placement in high-needs school as part of annual or biannual performance report (select one)?
- a. Fully implemented (all intended programs and subjects)
 - b. Partially implemented (or in pilot) --Put full implementation date below
 - c. Adopted in policy but not yet implemented
 - d. Other:

- 25b. Completer population reported on (select all that apply)?
- a. First year after program completion
 - b. Second year after program completion
 - c. Third year after program completion
 - d. Fourth year after program completion
 - e. Fifth year after program completion
 - f. Sixth year after program completion
 - g. Other:
- 25c. Limitations based on where completers work (select all that apply)?
- a. Completers teaching in-state only
 - b. Completers working in public schools only
 - c. Other:
- 25d. Reporting level (select all that apply)?
- a. Individual TPP (e.g. certification or endorsement area) where data is sufficient
 - b. Institution or EPP as a whole with data on all individual programs combined
 - c. Other:
- 25e. Additional information on measures used (select all that apply)?
- a. 1-year average over multiple cohorts
 - b. 2-year average over multiple cohorts
 - c. Other:
- 25f. Measure used to report placement (select all that apply)
- a. Proportion who are working at a given time:
 - b. Total proportion placed over a given time, e.g. percent who were placed at a high needs school within 3 years of graduation
 - c. Disaggregation by race, ethnicity, and/or gender
 - d. Compared to state average benchmark
 - e. Compared to some other benchmark (describe below)
 - f. Other:
 - g. Additional information on criteria used to describe measures including benchmarks (25f)
- 25g. Additional information on measures based on placement in high needs schools or subjects.
26. Is persistence in high needs schools a measure reported?
- a. Yes (continue to 26a)
 - b. No (go to 27)
 - c. Other:
 - d. Notes on Persistence in High Needs Schools: Question 26
- 26a. Implementation status of use placement in high-needs school/subject as part of annual or biannual performance report (select one)?
- a. Fully implemented (all intended programs and subjects)
 - b. Partially implemented (or in pilot) --Put full implementation date below
 - c. Adopted in policy but not yet implemented
 - d. Other:

- 26b. Completer population reported on (select all that apply)?
- a. First year after program completion
 - b. Second year after program completion
 - c. Third year after program completion
 - d. Fourth year after program completion
 - e. Fifth year after program completion
 - f. Sixth year after program completion
 - g. Other:
- 26c. Limitations based on where completers work (select all that apply)?
- a. Completers teaching in-state only
 - b. Completers working in public schools only
 - c. Other:
- 26d. Reporting level (select all that apply)?
- a. Individual TPP (e.g. certification or endorsement area) where data is sufficient
 - b. Institution or EPP as a whole with data on all individual programs combined
 - c. Other:
- 26e. Definition of "persistence" used?
- a. Retention in teaching and/or school leadership in one high-need school
 - b. Retention in teaching and/or school leadership in one high-need district
 - c. Retention in teaching and/or school leadership in high-need schools in the state
 - d. Retention in teaching and/or school leadership in high-need schools in-state or out-of-state
 - e. Other:
- 26f. Measures used to evaluate program performance (select all that apply)?
- a. Percentage of members of completer cohorts reported on still in teaching or school administrative positions in a high-need school, district, or subject --NOTE number of years this is measured in notes below
 - b. Average number of years that completers in the cohorts reported on were employed in teaching or school administration in a high-need school, district, or subject
 - c. Percentage of completers teaching in high-need subjects who achieved a second-stage license to teach in that schools
 - d. Other:
 - e. Notes on persistence measures (26f)
- 26g. Measures used in program evaluation (select all that apply)?
- a. Comparison persistence by subgroup (gender, race/ethnicity)
 - b. Comparison to state average persistence benchmark
 - c. Comparison to some other persistence benchmark
 - d. Other:
 - e. Additional information on criteria used to describe measures including benchmarks (26g)
- 26h. Additional notes on use of persistence in high needs schools as a measure.

27. Is placement of completers in high needs subjects reported measure? High needs subjects are defined by the state as part of teacher shortage reporting.
 - a. Yes (continue to 27a)
 - b. No (go to 28)
 - c. Other:
 - d. Notes on Placement of Completers in High Needs Subjects: Question 27
- 27a. Implementation status of use placement in high-needs subject as part of annual or biannual performance report (select one)?
 - a. Fully implemented (all intended programs and subjects)
 - b. Partially implemented (or in pilot) --Put full implementation date below
 - c. Adopted in policy but not yet implemented
 - d. Other:
- 27b. Completer population reported on (select all that apply)?
 - a. First year after program completion
 - b. Second year after program completion
 - c. Third year after program completion
 - d. Fourth year after program completion
 - e. Fifth year after program completion
 - f. Sixth year after program completion
 - g. Other:
- 27c. Limitations based on where completers work (select all that apply)?
 - a. Completers teaching in-state only
 - b. Completers working in public schools only
 - c. Other:
- 27d. Reporting level (select all that apply)?
 - a. Individual TPP (e.g. certification or endorsement area) where data is sufficient
 - b. Institution or EPP as a whole with data on all individual programs combined
 - c. Other:
- 27e. Additional information on measures used (select all that apply)?
 - a. 1-year average over multiple cohorts
 - b. 2-year average over multiple cohorts
 - c. Other:
- 27f. Measure used to report placement (select all that apply)?
 - a. Proportion who are working at a given time: Average placement, above this is described as a measure of both placement and persistence
 - b. Total proportion placed from at a given time: Cumulative placement
 - c. Other:
- 27g. Measures used in program evaluation (select all that apply)?
 - a. Comparative performance of subgroups (gender, race/ethnicity)
 - b. Comparison to state average benchmark
 - c. Comparison to some other benchmark
 - d. Other:
 - e. Additional information on criteria used to describe measures including benchmarks (27g)

- 27h. Additional information on measures based on placement in high needs schools or subjects.
28. Is persistence in high needs SUBJECTS a measure reported?
- Yes (continue to 28a)
 - No (go to 29)
 - Other:
 - Notes on Persistence in High Needs Subjects: Question 28
- 28a. Implementation status of use placement in high-needs school/subject as part of annual or biannual performance report (select one)?
- Fully implemented (all intended programs and subjects)
 - Partially implemented (or in pilot) --Put full implementation date below
 - Adopted in policy but not yet implemented
 - Other:
- 28b. Completer population reported on (select all that apply)?
- First year after program completion
 - Second year after program completion
 - Third year after program completion
 - Fourth year after program completion
 - Fifth year after program completion
 - Sixth year after program completion
 - Other:
- 28c. Limitations based on where completers work (select all that apply) ?
- Completers teaching in-state only
 - Completers working in public schools only
 - Other:
- 28d. Reporting level (select all that apply)?
- Individual TPP (e.g. certification or endorsement area) where data is sufficient
 - Institution or EPP as a whole with data on all individual programs combined
 - Other:
- 28e. Definition of "persistence" used (select one)?
- Retention in teaching and/or school leadership in one high-need school
 - Retention in teaching and/or school leadership in one high-need district
 - Retention in teaching and/or school leadership in high-need schools in the state
 - Retention in teaching and/or school leadership in high-need schools in-state or out-of-state
 - Other:
- 28f. Measures used to evaluate program performance (select all that apply)
- Percentage of members of completer cohorts reported on still in teaching or school administrative positions in a high-need school, district, or subject --NOTE number of years this is measured in notes below
 - Average number of years that completers in the cohorts reported on were employed in teaching or school administration in a high-need school, district, or subject
 - Percentage of completers teaching in high-need subjects who achieved a second-stage license to teach in that subject
 - Other:
 - Notes on number of years that persistence percentage is measured over (28f)?

- 28g. Additional notes on use of persistence in high needs schools as a measure.
29. Are you using or planning to use any other measures of teacher preparation programs, in addition to those we've asked about?
- Yes (continue to 29a)
 - No (go to 30)
 - Other:
- 29A. If Yes, please name the additional measure 1 (Question 29).
- 29Aa. Implementation status of additional measure 1 as part of annual or biannual performance report (select one)?
- Fully implemented (all intended programs and subjects)
 - Partially implemented (or in pilot) --Put full implementation date below
 - Adopted in policy but not yet implemented
 - Other:
- 29Ab. Reporting level for additional measure #1 (select all that apply)?
- Individual TPP (e.g. certification or endorsement area) where data is sufficient
 - Institution or EPP as a whole with data on all individual programs combined
 - Other:
- 29Ac. Detail on additional measure #1?
- 29B. If there is a 2nd additional measure, please name the additional measure #2
- 29Ba. Implementation status of additional measure #2 as part of annual or biannual performance (select one)?
- Fully implemented (all intended programs and subjects)
 - Partially implemented (or in pilot) --Put full implementation date below
 - Adopted in policy but not yet implemented
 - Other:
- 29Bb. Reporting level for additional measure #2 (select all that apply)?
- Individual TPP (e.g. certification or endorsement area) where data is sufficient
 - Institution or EPP as a whole with data on all individual programs combined
 - Other:
- 29Bc. Detail on additional measure #2?
- 29C. If there is a 3rd additional measure, please name the additional measure #3
- 29Ca. Implementation status of additional measure #3 as part of annual or biannual performance (select one)?
- Fully implemented (all intended programs and subjects)
 - Partially implemented (or in pilot) --Put full implementation date below
 - Adopted in policy but not yet implemented
 - Other:
- 29Cb. Reporting level for additional measure #3 (select all that apply)?
- Individual TPP (e.g. certification or endorsement area) where data is sufficient
 - Institution or EPP as a whole with data on all individual programs combined
 - Other:

29Cc. Detail on additional measure #3?

Final data system questions

30. Does the state use the program performance measures to identify satisfactory and unsatisfactory program performance?
- Yes (continue to 30a)
 - No (go to 31)
 - Notes on Question 30
- 30a. If Yes on satisfaction level of program performance (question 30), how does the state use them (select all that apply)?
- The state sends a full performance report on all indicators to each EPP or TPP
 - The state notifies programs of unsatisfactory performance
 - The state intervenes in programs that persistently under-perform on one or more measures
 - Other:
31. How does the state use the data from the performance measures it employs to identify program strengths and weaknesses and/or to decide to intervene (select all that apply)?
- For each individual measure or indicator, the state compares all programs and registers a concern programs that score particularly low or high in comparison with other programs
 - For each individual measure or indicator, the state registers a concern about programs that fail to meet or greatly exceed state-established performance benchmarks
 - The state registers a concern about program performance only for programs that consistently score high or low on individual benchmarks over a period of several years
 - The state registers a concern about program performance only for programs that score high or low on multiple measures
 - Other:
32. Does the state aggregate program (or institutional) scores on individual measures or indicators into a single composite performance score for the EPP or its individual programs?
- Yes (continue to 32a)
 - No (go to 33)
 - Other:
- 32a. If the state creates aggregate scores, what is the reporting level for this score (select one)?
- Individual TPP (e.g. certification or endorsement area) where data is sufficient
 - Institution or EPP as a whole with data on all individual programs combined
 - Other:
- 32b. If the state develops a composite score, how is it developed (select all that apply)?
- State subtracts the number of indicators or measures on which a program performs significantly below average or fails to meet state established benchmarks from the number of measures on which a program performs significantly above average or exceeds established benchmarks
 - State constructs a composite average for each program based on scores for all indicators or measures, if yes, describe weighting in 32c
 - Other:

- 32c. Describe the weighting, list each measure and its overall weight (32bii).
- 32d. If the state ranks institutions, what, if any, additional information is provided at the program level?
33. Are teacher preparation providers and/or programs required to follow state-approved protocols in reviewing the data on program performance for purposes of self-assessment and improvement?
- Yes (continue to 33a)
 - No (go to 34)
 - Other:
- 33a. Are they required to report on that process to state officials (select one)?
- Yes
 - No
 - Other:
 - Provide more detailed information on protocols for reviewing state data (33a)
34. Are teacher preparation providers and/or programs required to submit to state officials a program improvement plan based on program performance data and/or state-developed protocols for reviewing that data?
- Yes --provide detail below
 - No
 - Other:
 - Is more detailed information available on these plans (34)
35. Please describe any additional kinds of actions employed by the state to address poor performance by programs on the various performance measures? Data system characteristics.
36. Please indicate which of the following data sets you have in place in your state that support your ability to gather evidence on the outcomes you have identified (select all that apply)?
- Longitudinal student assessment data that can be linked to teachers
 - Longitudinal teacher data that can be linked to the schools where they work
 - Longitudinal teacher data that can be linked to TPP or EPP
 - Longitudinal school level data that describes student characteristics in the school (student count, race/ethnicity)
 - Longitudinal school level data that describes special program participation at the school level (free and reduced lunch participation, special education)
 - School level data that describes student behavior in schools (absenteeism, transience, suspensions and violent crime).
 - Notes on Question 36
37. Does the state provide or impose data guidelines, definitions, and protocols for local school districts or TPPs/EPPs in order to ensure accuracy and uniformity of data collected (select all that apply)?
- On school districts (please complete 37a)
 - On EPPs (please complete 37b)
 - No (go to 38)
- 37a. If state provides or imposes reporting guidelines for local districts on their students, teachers, or school characteristics (question 37), please provide sources of guidance if possible?
- 37b. If state provides or imposes reporting guidelines for TPPs/EPPs on their programs, teacher candidates, and program completers, please provide sources of guidance if possible?

38. Please indicate which of the following state and local agencies are involved in maintaining the databases the state uses to assess its teacher preparation programs (select all that apply):
 - a. State department of education
 - b. State educator professional standards board
 - c. State higher education coordinating agency
 - d. Independent higher education coordinating agency
 - e. State department of labor and employment
 - f. State board of teacher licensure
 - g. Office of the governor
 - h. Other:
39. What is the agency primarily responsible for coordinating the preparation program data used (answer "none" if appropriate)?
40. Is the state able to gather data on the placement and/or performance of program completers who take teaching jobs out-of-state? (select one)
 - a. Yes-detail in text box below
 - b. No
 - c. Provide detail on data on placement/performance of completers in other states (Question 40)
41. Are there additional features of your teacher preparation performance assessment system – either currently implemented or in the actual process of implementation – that you think are important, but we’ve not discussed? If in the process of implementation, what are the actual status and timeline?
42. Are there other significant changes underway or planned by your state related to the evaluation of teacher preparation programs? If so, where can we find more information about these changes?
43. What are the greatest challenges you face to the sustainability and further development of your program performance assessment system?
44. Is there anything else you’d like to add in order to help us understand how your state is evaluating teacher preparation programs or the data system that supports it?
45. May we contact you if we have any clarifying questions and to have you verify an eventual draft report we’ll put together on your state’s preparation program reporting and accountability requirements?
46. Are there other people in your agency or in another agency we might talk to in order to learn information we may not have been able to obtain in this interview?

APPENDIX C

BASIC STATE PROGRAM REVIEW, APPROVAL, AND REPORTING FEATURES

Assignment Criteria

The Assignment Key: ○ = No; ● = Yes; **na** = not reported, or not discernible

States were assigned ● on indicators in the summary tables, if the feature: (a) involved data that are comparable between programs [using the same scoring rubric would be sufficient to satisfy this criterion]; and (b) used the data for accountability purposes, not just for information - i.e., so that an indicator score at some lower level would be considered problematic and a potential justification for state inquiry or action – either in relation to an annual or biennial accountability review or to a longer-term review of program performance.

Features not satisfying these two criteria were assigned ○.

These assignments were applied to states that currently implemented the feature or provided evidence that the item would be in place by fall 2017.

If the reviewers could not make a confident assessment whether a state did or did not have (or would not have by fall 2017) a certain feature at least partially implemented, the indicator was assigned **na**.

TABLE 1 (A) Basic State Program Review, Approval, and Reporting Features

Assignment Key: ○ = No; ● = Yes; na = not reported, or not discernible

State/Rating	AL	AK*	AZ	AR	CA	CO	CT	DE	DC*	FL	GA	HI	ID	IL	IN	IA	KS
1. Continuing program approval is required of all EPPs in state (including alternative routes)	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
2. Continuing state program approval is based on a comprehensive review every 5-7 years	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
3. CAEP accreditation is required of public college and university EPPs	○	●	○	●	○	○	○	●	●	○	●	●	●	○	●	○	●
4. CAEP accreditation is required of all EPPs approved to operate in the state	○	○	○	○	○	○	○	●	○	○	○	●	○	○	●	○	○
5. All individual certification programs must meet applicable SPA or other national discipline association standards	●	○	○	○	○	○	○	●	○	○	●	●	na	○	●	○	○
6. State has an accountability-related annual or biennial program performance review	●	○	●	●	●	●	○	●	●	●	●	●	○	○	●	●	○
7. State reports annual or biennial program performance review data to EPPs and/or the public	●	○	●	●	na	●	○	●	●	●	○	●	○	○	○	●	●
7a. Report covers all state and private college/university-based EPPs	●	○	●	●	○	●	○	●	●	●	●	●	○	○	●	●	●
7b. Report covers all non-traditional certification routes operating in the state	●	○	●	●	○	●	○	●	○	●	●	●	○	○	●	na	●
7c. Report data are aggregated to the EPP level	na	○	●	●	○	●	○	●	na	●	●	●	○	○	●	na	●
7d. Some data are reported out on individual certification programs or teaching subjects	na	○	●	●	○	●	○	●	na	●	●	○	○	○	●	na	●
8. Public has access to EPP-level accountability data	○	○	○	●	○	●	○	●	●	●	●	○	○	○	○	●	○
9. Public has access to accountability data on at least some individual certification programs or teaching subjects	○	○	○	●	○	●	○	○	na	●	○	○	○	○	●	na	○

There were a few states that did not respond to repeated requests for official verification of the information we collected about their accountability systems. These states are marked with an asterisk (*) next to their initials in the table.

TABLE 1 (B) Basic State Program Review, Approval, and Reporting Features

Assignment Key: ○ = No; ● = Yes; na = not reported, or not discernible

State/Rating	KY	LA	ME	MD	MA	MI	MN	MS	MO	MT*	NE	NV	NH	NJ	NM*	NY	NC
1. Continuing program approval is required of all EPPs in state (including alternative routes)	●	-	●	●	●	●	●	●	●	○	●	●	○	●	●	●	●
2. Continuing state program approval is based on a comprehensive review every 5-7 years	●	-	●	●	●	●	●	●	○	●	●	●	●	●	●	●	●
3. CAEP accreditation is required of public college and university EPPs	○	-	○	○	○	●	○	●	○	○	○	○	○	●	○	●	●
4. CAEP accreditation is required of all EPPs approved to operate in the state	○	-	○	○	○	●	○	○	○	○	○	○	○	●	○	●	○
5. All individual certification programs must meet applicable SPA or other national discipline association standards	●	-	○	○	●	○	○	○	○	○	○	○	○	○	○	○	○
6. State has an accountability-related annual or biennial program performance review	●	-	○	○	●	●	●	○	●	○	●	●	○	○	na*	na	●
7. State reports annual or biennial program performance review data to EPPs and/or the public	●	-	●	○	●	●	●	●	●	○	●	●	●	●	●	na	●
7a. Report covers all state and private college/university-based EPPs	●	-	●	○	●	●	●	●	●	○	●	●	●	●	●	○	●
7b. Report covers all non-traditional certification routes operating in the state	●	-	●	○	●	○	●	●	●	○	●	●	○	●	na	○	○
7c. Report data are aggregated to the EPP level	●	-	●	○	●	●	●	●	○	○	●	●	○	○	na	○	●
7d. Some data are reported out on individual certification programs or teaching subjects	●	-	na	○	●	○	○	○	○	○	●	●	○	●	na	○	●
8. Public has access to EPP-level accountability data	●	-	○	na	●	●	●	●	●	○	●	○	○	●	●	na	●
9. Public has access to accountability data on at least some individual certification programs or teaching subjects	●	-	○	na	○	●	●	○	○	○	●	○	○	●	na	na	○

Note: The Louisiana Department of Education requested not to be included in this analysis.

TABLE 1 (C) Basic State Program Review, Approval, and Reporting Features

Assignment Key: ○ = No; ● = Yes; na = not reported, or not discernible

State/Rating	ND*	OH	OK	OR	PA	RI	SC	SD	TN	TX	UT	VA	VT	WA	WV	WI	WY
1. Continuing program approval is required of all EPPs in state (including alternative routes)	○	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
2. Continuing state program approval is based on a comprehensive review every 5-7 years	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
3. CAEP accreditation is required of public college and university EPPs	○	●	○	●	○	○	●	○	○	○	●	○	○	○	●	○	●
4. CAEP accreditation is required of all EPPs approved to operate in the state	○	na	○	○	○	○	○?	○	○	○	●	○	○	○	○	○	●
5. All individual certification programs must meet applicable SPA or other national discipline association standards	○	○	●	○	○	●	●	●	●	○	○	○	○	●	○	○	●
6. State has an accountability-related annual or biennial program performance review	●	●	●	●	●	●	○	○	●	●	○	○	○	●	●	●	○
7. State reports annual or biennial program performance review data to EPPs and/or the public	●	●	●	●	●	●	●	○	●	●	●	●	●	●	●	●	○
7a. Report covers all state and private college/university-based EPPs	●	●	●	●	●	●	●	○	●	●	●	●	●	●	●	●	○
7b. Report covers all non-traditional certification routes operating in the state	○	na	○	●	●	●	○	○	●	●	●	○	●	●	●	●	○
7c. Report data are aggregated to the EPP level	na	●	●	●	●	●	○	○	●	●	○	●	na	●	●	●	○
7d. Some data are reported out on individual certification programs or teaching subjects	na	na	○	●	●	●	○	○	●	○	○	●	na	●	na	●	○
8. Public has access to EPP-level accountability data	●	●	●	○	●	●	○	○	●	●	○	●	na	●	○	●	○
9. Public has access to accountability data on at least some individual certification programs or teaching subjects	○	●	●	○	●	●	○	○	●	○	○	●	na	●	○	●	○

APPENDIX D

STATE PROGRAM REVIEW AND/OR REPORT INDICATORS

Assignment Criteria

The Assignment Key: ○ = No; ● = Yes; **na** = not reported, or not discernible

States were assigned ● on indicators in the summary tables, if an indicator: (a) involved data that are comparable between programs [using the same scoring rubric would be sufficient to satisfy this criterion]; and (b) used the data for accountability purposes, not just for information - i.e., so that an indicator score at some lower level would be considered problematic and a potential justification for state inquiry or action – either in relation to an annual or biennial accountability review or to a longer-term review of program performance.

Indicators not satisfying these two criteria were assigned ○.

These assignments were applied to states that currently implemented the indicator or provided evidence that the item would be in place by fall 2017.

If the reviewers could not make a confident assessment whether a state did or did not have (or would not have by fall 2017) a certain indicator at least partially implemented, the indicator was assigned **na**.

TABLE 2 (A) **State Program Review and/or Report Indicators**
Assignment Key: ○ = No; ● = Yes; na = not reported, or not discernible

State/Rating	AL	AK*	AZ	AR	CA	CO	CT	DE	DC*	FL	GA	HI	ID	IL	IN	IA	KS
1. Candidate academic proficiency – at program Entry	●	○	○	○	na	○	○	●	●	na	●	○	na	●	●	●	○
2. Completer academic proficiency – at program Completion	○	○	○	○	na	○	○	●	●	na	●	○	na	○	na	●	○
3. Candidate teaching promise – at program entry	○	○	○	○	○	○	○	●	○	○	○	○	na	○	○	●	○
4. Demographic composition of entering candidate cohort	○	○	○	○	○	○	○	●	○	○	○	○	na	○	○	○	○
5. Completion rate of entering candidates	○	○	●	○	na	○	○	●	●	●	●	●	○	○	●	●	○
6. Completion rate by demographic sub-group	○	○	○	○	○	○	○	●	○	○	○	○	○	○	na	○	○
7. Completers' mastery of content for certification field	●	○	●	●	●	○	○	●	●	●	●	○	○	●	●	●	○
8. Completers' mastery of content-specific pedagogical knowledge	●	○	●	○	●	○	○	○	●	●	●	○	○	●	●	○	○
9. Completers' general teaching skill and proficiency	●	○	●	○	●	●	○	○	●	●	●	●	○	●	●	●	○
10. Completers' evaluation of program at completion	●	○	na	●	○	○	○	●	○	○	●	○	○	○	○	○	○
11. Completers' evaluation of program in first year teaching	●	○	na	●	○	○	○	●	○	○	●	○	○	○	●	○	○
12. Supervisors' evaluations of new teachers' pre-service preparation	●	○	●	○	○	○	○	●	○	○	●	○	○	○	●	●	○
13. Completers' impact on P-12 pupils as teachers of record	○	○	○	○	○	○	○	●	○	●	○	○	○	○	●	○	○
14. Completers' observed performance as teachers of record	○	○	●	○	○	○	○	●	○	○	●	○	○	○	na	○	○
15. P-12 students' perceptions of new teachers' effectiveness	●	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○
16. Placement rate of completers in teaching	●	○	●	●	○	○	○	●	○	●	●	○	○	○	○	●	○
17. Persistence of completers in teaching (minimum 3-5 years)	○	○	●	○	○	○	○	●	○	●	●	○	○	○	○	●	○
18. Completer placement/persistence in high-need schools	○	○	●	○	○	○	○	●	○	○	○	○	○	○	○	○	○
19. Indicators (some or all) are combined into overall program score, rating, and/or ranking	○	○	○	○	○	○	○	●	○	●	●	○	○	○	○	○	○
20. Indicators are assigned various weightings	○	○	○	○	○	○	○	●	○	●	●	○	○	○	○	○	○

There were a few states that did not respond to repeated requests for official verification of the information we collected about their accountability systems. These states are marked with an asterisk (*) next to their initials in the table.

TABLE 2 (B) **State Program Review and/or Report Indicators**
Assignment Key: ○ = No; ● = Yes; na = not reported, or not discernible

State/Rating	KY	LA*	ME	MD	MA	MI	MN	MS	MO	MT*	NE	NV	NH	NJ	NM*	NY	NC
1. Candidate academic proficiency – at program Entry	●	-	○	●	●	○	na	●	●	○	○	○	●	○	○	●	●
2. Completer academic proficiency – at program Completion	○	-	○	●	na	●	●	●	●	○	○	○	●	●	○	●	●
3. Candidate teaching promise – at program entry	○	-	○	○	○	○	○	○	○	○	○	○	●	○	○	○	○
4. Demographic composition of entering candidate cohort	●	-	○	○	○	○	○	●	○	○	○	○	○	○	○	○	●
5. Completion rate of entering candidates	○	-	○	●	●	○	○	●	○	○	○	○	○	○	○	●	○
6. Completion rate by demographic sub-group	○	-	○	○	●	○	●	●	○	○	○	○	○	○	○	○	●
7. Completers' mastery of content for certification field	●	-	○	●	●	●	●	●	○	○	○	●	●	●	●	●	●
8. Completers' mastery of content-specific pedagogical knowledge	○	-	○	○	●	●	●	○	○	○	○	●	●	○	○	●	○
9. Completers' general teaching skill and proficiency	●	-	○	○	●	○	○	○	○	○	○	●	●	●	●	●	●
10. Completers' evaluation of program at completion	●	-	○	○	●	●	●	○	●	○	○	○	○	○	○	●	●
11. Completers' evaluation of program in first year teaching	●	-	○	○	●	●	●	○	●	○	○	●	○	○	○	na	●
12. Supervisors' evaluations of new teachers' pre-service preparation	●	-	○	○	○	○	●	○	●	○	○	○	○	○	○	●	na
13. Completers' impact on P-12 pupils as teachers of record	○	-	○	●	●	●	○	○	○	○	○	○	○	○	○	●	●
14. Completers' observed performance as teachers of record	○	-	○	○	●	●	○	○	○	○	○	○	○	○	○	●	○
15. P-12 students' perceptions of new teachers' effectiveness	○	-	○	○	○	○	○	○	○	○	○	○	○	○	○	●	●
16. Placement rate of completers in teaching	●	-	○	○	●	○	○	○	○	○	○	●	○	●	○	●	●
17. Persistence of completers in teaching (minimum 3-5 years)	●	-	○	○	○	○	○	○	○	○	○	●	○	○	○	●	○
18. Completer placement/persistence in high-need schools	○	-	○	○	○	○	○	○	○	○	○	●	○	●	○	○	○
19. Indicators (some or all) are combined into overall program score, rating, and/or ranking	○	-	○	○	○	●	○	●	○	○	○	○	○	○	○	○	○
20. Indicators are assigned various weightings	○	-	○	○	○	●	○	na	○	○	○	○	○	○	○	○	○

Note: The Louisiana Department of Education requested not to be included in this analysis.

TABLE 2 (C) **State Program Review and/or Report Indicators**
Assignment Key: ○ = No; ● = Yes; na = not reported, or not discernible

State/Rating	ND*	OH	OK	OR	PA	RI	SC	SD	TN	TX	UT	VA	VT	WA	WV	WI	WY
1. Candidate academic proficiency – at program Entry	○	●	○	●	●	●	●	na	●	●	○	○	●	●	○	●	na
2. Completer academic proficiency – at program Completion	○	●	●	○	●	●	○	○	●	●	○	○	●	●	○	●	na
3. Candidate teaching promise – at program entry	○	○	○	●	●	na	○	○	○	○	○	○	●	○	○	○	na
4. Demographic composition of entering candidate cohort	○	○	○	●	●	●	○	○	○	na	○	○	○	○	○	○	na
5. Completion rate of entering candidates	○	○	○	●	●	●	○	○	○	●	○	○	○	●	●	○	na
6. Completion rate by demographic sub-group	○	○	○	●	●	na	○	○	○	○	○	○	○	●	○	●	na
7. Completers' mastery of content for certification field	○	●	●	●	●	●	●	○	●	○	○	●	●	●	●	●	na
8. Completers' mastery of content-specific pedagogical knowledge	○	○	●	○	●	●	●	○	○	○	○	○	○	●	○	●	na
9. Completers' general teaching skill and proficiency	○	●	●	○	●	●	●	○	●	○	○	●	○	●	●	na	na
10. Completers' evaluation of program at completion	○	●	○	○	na	●	○	○	○	●	○	○	●	○	○	○	na
11. Completers' evaluation of program in first year teaching	○	●	●	○	●	●	○	○	●	○	○	○	○	●	○	na	na
12. Supervisors' evaluations of new teachers' pre-service preparation	○	○	○	○	●	●	○	○	●	●	○	●	○	○	○	○	na
13. Completers' impact on P-12 pupils as teachers of record	○	●	○	○	○	●	○	○	●	●	○	○	○	●	○	○	na
14. Completers' observed performance as teachers of record	○	○	○	○	●	●	○	○	●	○	○	○	na	●	○	○	○
15. P-12 students' perceptions of new teachers' effectiveness	○	○	○	○	○	○	○	○	○	○	○	○	○	●	○	●	○
16. Placement rate of completers in teaching	○	○	○	●	●	●	○	○	●	●	○	○	○	●	●	○	○
17. Persistence of completers in teaching (minimum 3-5 years)	○	●	○	●	●	●	○	○	●	●	○	○	○	●	○	na	○
18. Completer placement/persistence in high-need schools	○	○	○	●	○	●	○	○	○	○	○	○	○	○	○	○	○
19. Indicators (some or all) are combined into overall program score, rating, and/or ranking	○*	○	○	○	○*	○	●	○	●	○	○	○	○	○	○	○	○
20. Indicators are assigned various weightings	○	○	○	○*	○	○	●	○	○	●	○	○	○	○	○*	○	○



One Massachusetts Avenue, NW, Suite 700
Washington, DC 20001-1431
voice: 202.336.7000 | fax: 202.408.8072