Improving State Evaluation of Principal Preparation Programs

FEBRUARY 2016

New Leaders

COMMISSIONED BY:

The Wallace Foundation®

Recommended citation: UCEA and New Leaders (2016). *Improving state evaluation of principal preparation programs*. Retrieved from: www.sepkit.org.

Improving State Evaluation of Principal Preparation Programs

Far too many people completing state-approved principal preparation programs are not ready to assume assistant principal or principal positions.[1] This concerning fact has motivated a number of programs to pioneer effective, evidence-based practices, such as increasing the rigor of their admissions or developing partnerships with schools and districts to provide candidates with authentic opportunities to practice leadership and receive feedback on their performance.[2] Unfortunately, there is little evidence that these and other research-based practices are broadly in use across programs. In fact, there has been a proliferation of programs, particularly less-rigorous online programs, in response to consumer demand for cheaper and more flexible options.[3] Ensuring programs reflect best practices, such as providing opportunities for authentic practice, is a central opportunity for state policy-makers looking to improve the quality of principal leadership.

State policy-makers are uniquely positioned to ensure that the quality of principal preparation improves because most states have statutory authority to do so. Specifically, states grant initial and ongoing approval for principal preparation programs to operate, and they issue licenses for individuals to serve as principals. Because of this authority, a recent comprehensive report on state policy identified the approval and oversight of principal preparation programs as one of six potentially powerful areas for state policy action to improve the effectiveness of school principals.[4]

At the heart of a state's authority to approve programs is the opportunity to evaluate those programs. What is more, high-quality program evaluation is a means to improve programs. Data collected through program evaluation provides critical evidence for identifying areas where programs could improve their design or execution. Thus, in addition to providing insight into the quality of programs through a set of agreed-upon metrics, the data collected and analyzed during the evaluation process can be used to inform program changes.[5]

1. Hull, J. (2012). *The principal perspective: Full report*. Alexandria, VA: Center for Public Education; Young, M. D., & Brewer, C. (2008). Fear and the preparation of school leaders: The role of ambiguity, anxiety, and power in meaning making. *Educational Policy*, 22(1), 106–129.
2. Darling-Hammond, L., LaPointe, M., Meyerson, D., & Orr, M. T. (2009). *Preparing principals for a changing world: Lessons from effective school leadership programs*. San Francisco, CA: Jossey-Bass; Davis, S. H., & Darling-Hammond, L. (2012). Innovative principal preparation programs: What works and how we know. *Planning and Changing*, 43(1/2), 25–45.
3. Anderson, E., & Reynolds, A. L. (2015). *A policymaker's guide: Research-based policy for principal preparation program approval and licensure*. Charlottesville, VA: The University Council for Educational Administration.
4. Manna, P. (2015). *Developing excellent school principals to advance teaching and learning: Considerations for state policy*. New York, NY: The Wallace Foundation.
5. Patton, M. Q. (1997). *Utilization-focused evaluation: The new century text*. Thousand Oaks, CA: Sage Publications; Orr, M. T., Young, M. D., & Rorrer, A. K. (2010). *Developing evaluation evidence: A formative and summative evaluation planner for educational leadership preparation programs*. Charlottesville, VA: UCEA National Center for the Evaluation of Educational Leadership Preparation and Practice.

Unfortunately, most states have not yet developed robust program evaluation systems.[6] In 2015, the University Council for Educational Administration (UCEA) and New Leaders completed a project, with support from The Wallace Foundation, that aimed to produce ideas and resources that would enable states to design and conduct feasible, fair, and useful evaluation. The project resulted in a detailed model and a related set of tools for effective state evaluation of principal preparation programs.

This publication draws on findings from the UCEA/New Leaders project, highlighting key design principles that states should consider as they begin the work of improving the evaluation of their principal preparation programs. It also includes examples from two states that have done substantial work in this arena, as well as one of the tools from the UCEA/New Leaders project—a readiness assessment rubric for states.

Why design principles? This paper outlines a set of design principles for state leaders to consider as they work to improve the quality of principal preparation program evaluation. While states will undoubtedly want and need to develop systems unique to their context, they could benefit from having a set of guideposts to organize what can be complex work. The design principles are based on program evaluation generally and on the evaluation of principal preparation programs specifically, as well as in-depth conversations with a diverse group of academics, policy-makers, and practitioners.

The UCEA/New Leaders Partnership

The UCEA/New Leaders project combined the expertise of two organizations with deep collective knowledge of research and practice related to the preparation of school leaders. UCEA is a consortium of 99 higher education institutions that has a 60-year track record of building the knowledge base on effective leadership preparation, designing and utilizing preparation program standards, and developing evaluation tools and practices designed to improve the preparation and professional development of educational leaders and professors. New Leaders operates cutting-edge principal preparation programs that are producing highly effective leaders and uses the knowledge gained from rigorously evaluating its own programs to inform federal, state, and local policy and practice, as well as training other preparation programs on how to design and conduct program evaluation.[8]

Engaging Experts in Design

The New Leaders/UCEA project also involved a deep and iterative collaboration with state leaders who have experience designing or implementing preparation evaluation systems, methodologists with experience evaluating principal preparation programs, representatives of national organizations focused on issues of leadership preparation, principal preparation program leaders with experience evaluating their own or other programs, district leaders with experience evaluating internal or external programs, and principals. Twenty-five academics, policy-makers, and practitioners participated as advisors.[9] Their participation included a series of webinars addressing specific and challenging issues (i.e., state authority and leadership, data considerations, the rigor of outcomes and processes, and consumer needs and priorities); a two-day, design-focused convening; and a review of the tools developed. Five experts also conducted in-depth reviews of all of the documents before they were finalized.

6. Anderson, E., & Reynolds, A. L. (2015).
7. Orr, M. T., Young, M. D., & Rorrer, A. K. (2010).
8. Neuman-Sheldon, B., Ikemoto, G. S., Bailey, M., Erdfarb, T., Nerenberg, L., Patterson, N., & Valdez, M. (2014). *Principal preparation program self-evaluation: Lessons learned by New Leaders*. New York, NY: New Leaders.
9. Mónica Byrne-Jiménez, Hofstra University (New York); Mary Canole, Council of Chief State School Officers; Stevie Chepko, Council for the Accreditation of Educator Preparation; Matthew Clifford, American Institutes for Research; Shelby Cosner, University of Illinois at Chicago; Brian Dassler, Florida Department of Education; Jacquelyn Davis, George W. Bush Institute; Benjamin Fenton, New Leaders; Susan Gates, RAND Corporation; Mark Gooden, University of Texas at Austin; Jackie Gran, New Leaders; Steven Gross, Temple University (Pennsylvania); Sara Heyburn, Tennessee State Board of Education; Susan Korach, Ritchie Program for School Leaders, University of Denver (Colorado); Paul Manna, College of William and Mary; Tricia McManus, Hillsborough County Public Schools (Florida); Glenn Pethel, Gwinnett County Public Schools (Georgia); Diana Pounder, University of Utah Education Policy Center; Frances Rabinowitz, Bridgeport Public Schools (Connecticut); Carol Riley, National Association of Elementary School Principals; Cortney Rowland, National Governors Association; Christopher Ruszkowski, Delaware Department of Education; Erin Swanson, Martin Millennium Academy (North Carolina); Brenda Turnbull, Policy Studies Associates; David Volrath, Maryland State Department of Education.

Why examples? A small number of states in recent years have made efforts to use the evaluation of principal preparation programs as a strategy for improvement. Illinois, for example, has spent a decade developing and refining a new set of expectations for principal preparation programs, most notably requiring deeper partnerships between programs and school districts, and has required all programs in the state to redesign their programs based on the new criteria. Other states, such as Delaware, Florida, and Tennessee, are at earlier stages of development.

Such examples demonstrate some of the possibilities of high-quality program evaluation, and they surface several tensions for states. Some of these are political in nature. For example, attention to principal leadership as a statewide focus necessarily competes with other priorities, especially a focus on teacher quality and support. In addition, while some programs are increasing selectivity in the admission of students in order to improve overall quality, others depend on open admissions to generate revenue for schools of education. Other tensions are more technical in nature. For example, evaluations are ideally driven by data about program quality and outcomes, but state systems for collecting and interpreting data are often too limited to support such an approach (i.e., they lack direct measures of program quality and outcomes, and the available data for indirect measures are insufficient or of low quality). Moreover, evaluations are ideally diagnostic in nature, using program quality and outcome data to drive inquiries into the sources of successes, needed improvements, and concerns that warrant further investigation. However, state systems are not always organized to use data in this way and often lack the necessary capacity for diagnosis and support. Managing these tensions is essential in developing systems of evaluation that contribute to better outcomes.

Why tools? An exhaustive review of the existing literature on the evaluation of principal preparation programs makes plain that states lack good models, tools, and resources. To address this problem, UCEA and New Leaders worked with a diverse group of academics, policy-makers, and practitioners to develop a model approach to program evaluation and a suite of related tools and resources. A list of all the tools and resources developed by New Leaders and UCEA can be found on pages 11-12. To find the New Leaders/UCEA model, tools, and resources, see www.sepkit.org.

DESIGN PRINCIPLES FOR IMPROVING THE EVALUATION OF PRINCIPAL PREPARATION PROGRAMS

State leaders interested in improving their systems for evaluating principal preparation programs will ultimately need solutions tailored to their specific contexts. But conversations with state leaders and others committed to improving principal preparation surfaced a set of five core design principles on which to ground their efforts. While states may face capacity constraints (e.g., with regard to their data collection and analysis capabilities) that push against these design principles, the principles nevertheless represent a set of goals to which states can aspire and toward which they can work.

A. Structure the review process in a way that is conducive to continuous program improvement
Effective program review encourages ongoing improvement and innovation in program design and implementation in two ways. First, it provides programs with specific and actionable feedback about their practices and graduate outcomes. This feedback requires that the reviewers possess relevant expertise for making appropriate judgments, including content expertise in leadership, understanding of adult learning theory and practices, knowledge of current research about effective leadership preparation, and the ability to analyze curriculum and pedagogy. Second, an effective review system allows adequate time for improvement. To be truly focused on improvement, review cycles and processes provide programs with adequate time to make changes and assess their impact.

B. Create appropriate systems to hold programs accountable for effective practices and outcomes
An evaluation system is one of the key ways states can hold preparation programs accountable for their role in delivering high-quality preparation for aspiring principals. With approximately 700 programs currently in operation and new ones emerging on a regular basis, states need to be able to confidently make consequential decisions such as whether to approve a program; when to put a program on an improvement plan; and, in the most serious circumstances, when to rescind program approval. To generate that confidence, states can consider the following characteristics of system design: (1) understand the limitations of the indicators being tracked as measures of quality, and ensure that there is sufficient and valid information for making consequential decisions; (2) develop a clear and transparent rating system that has enough levels to meaningfully differentiate performance across programs and that captures performance and improvement over time; and (3) develop a clear and transparent process and timeline for intervening in the event of unacceptable performance.

C. Provide key stakeholders with accurate and useful information
When key consumers and partners—especially aspiring school leaders and school districts—have good information about key indicators of program quality, they can use that information to make more informed choices. For aspirants, a state evaluation system can provide concrete information about program features and outcomes (e.g., candidate learning and career outcomes) to inform enrollment choices. Ideally, systems would provide side-by-side, apples-to-apples comparisons of programs to help inform decision making. For districts, evaluation systems can provide specific information about program characteristics and candidate outcomes to guide decisions concerning formal partnerships with programs and the hiring of graduates. To meet these goals, effective evaluation systems provide high-quality, publicly available, reliable, and understandable data about programs.

D. Take a sophisticated and nuanced approach to data collection and use
Collecting and using data is central to program improvement, yet it presents significant challenges. Too often, data points can be misleading or misused. Taking a sophisticated and nuanced approach to data collection and use encompasses five related ideas.

1. **Evaluate what matters.** Strong data systems include the indicators that are most germane to principal preparation. These include inputs (especially the rigor of selection into a program and the diversity of candidates), processes (especially the ways in which a program increases aspirants' leadership knowledge and skills), outputs (especially aspirants' successful placement in roles as principals), and graduate outcomes (especially contributions to student academic achievement measures, student attainment measures such as graduation, and noncognitive measures such as student engagement and social/emotional growth).

2. **Evaluate accurately.** Strong data systems use the most accurate data available, and interpretations are made cautiously and with awareness of data limitations. Special caution

is needed in establishing confidence in the accuracy of measures of leadership effectiveness, which are in the early stages of development. Limitations related to the reliability and validity of data from particular sources can help determine whether to use and how much to weight those data in an evaluation system.

3. **Include data that can be realistically gathered and shared.** In strong systems, data are feasible to gather, efficient to report, and can be used in conjunction with other sources of information to strengthen evidence. Further, data collection is ongoing and conducted according to an established schedule. In many states, new investments in data collection and reporting will be essential to creating stable, consistent data.

4. **Consider contextual factors.** Data are means, not ends. In order to make appropriate judgments based on accurate results, states will need additional contextual information. For example, a 100% admissions rate for a program could signal that the program does not have rigorous admissions—or it could be the result of targeted recruiting and effective prescreening, such that only strong applicants apply. Therefore, it is important to provide programs with opportunities to explain their results so that states can better interpret results and draw conclusions. These information exchanges about root causes can be the basis of productive conversations about program quality and improvement.

5. **Clearly and transparently communicate how results will be used.** Ensure that program leaders understand which data will be made public, including how and when that will occur. Program leaders also need to understand how component parts of the program evaluation will be used to make substantive judgments and decisions about program status.

At heart, this design principle is about triangulating multiple data sources to arrive at more accurate judgments. Research has consistently demonstrated the limitations of specific data indicators as measures of quality (including, for example, a recent statement from the American Education Research Association on the use of value-added models of student achievement[10]). Put simply, single types of data on their own are often imperfect. But imperfect data can provide important and useful information when used appropriately. For example, it is well documented in the medical field that mammograms have a 50–60% false-positive rate and a 20% false-negative rate. Given the costs and risks of mammograms (such as exposure to radiation and stress from false-positive tests), leading medical organizations differ in their recommendations about the age at which women should start receiving mammograms and how often they should be administered. However, no one in the medical field recommends discontinuing mammogram testing altogether. To the contrary, mammograms are a critical diagnostic tool that generates data that doctors examine in conjunction with other data to make decisions about further testing and treatment. No doctor would ever use mammogram data alone to recommend surgery or make other make high stakes decisions.

What is true for physicians reading a mammogram result is also true for state education leaders looking at the graduation rate of a principal preparation program or the growth in student achievement in schools led by a program's graduates. By themselves, these indicators offer only limited insights, but combined with a deeper professional review, they can help state officials arrive at a full and accurate picture of program quality. That fuller picture can be the basis for states to make consequential decisions about program approval and can be the impetus for continuous improvement of all programs.

E: Adhere to characteristics of high-quality program evaluation

Effective state systems of program evaluation reflect best practices in program evaluation in education. The Standards for Educational Evaluation, issued by the Joint Committee on Standards for Educational Evaluation, serve as a basis for judging best practices. These standards focus on utility (i.e., the extent to which stakeholders find processes and results valuable), feasibility (i.e., the effectiveness and efficiency of evaluation processes), propriety (i.e., the fairness and appropriateness of evaluation processes and results), accuracy (i.e., the dependability of evaluation results, especially judgments of quality), and accountability (i.e., having adequate documentation to justify results).[11] These standards may often be in tension with one another; for example, data gathered from first-hand observations of program processes may be of high utility but may also be restrictively expensive to gather, thus making them less feasible to include.

---

10. AERA Council. (2015). AERA statement on use of value-added models (VAM) for the evaluation of educators and educator preparation programs. *Educational Researcher*, 44(8), 448–452.

New Leaders

University Council for Educational Administration
Quality Leadership Matters   since 1954

EXAMPLES TO CONSIDER

Implementing a quality system for evaluating principal preparation programs is complex work. For many states, this work will be new and difficult and will require a significant commitment of time, expertise, and resources. A few states have paved the way, and their experience can provide a road map for others.

Illinois

By 2015, Illinois had been working for more than a decade to develop and implement policies focused on improving principal preparation. Formal conversations among stakeholders culminated in the passage of comprehensive legislation that replaced a general administrative credential with one specifically focused on preparation for the principalship, increased the rigor of selection into programs and program content, required programs to collaborate with school districts, and required programs to include an internship that gave candidates authentic leadership experiences. The law mandated that all preparation programs in operation be approved under these new requirements; once regulations were finalized in 2011, programs had three years to fully meet new requirements. The state convened and trained a review panel consisting of teachers, principals, superintendents, university representatives, and members of the business community to provide constructive feedback in advance of consequential decisions about program approval by the Illinois State Educator Preparation and Licensure Board (ISEPLB). Twenty-six of 31 programs which previously offered a general administrative credential received approval from ISEPLB to prepare principals, and the policy requirements caused them to view districts, rather than individuals, as their primary consumers. Further, "[t]hese changes require[d] programs to move beyond the focus on a single program outcome—graduates securing administrative positions—to the actual impact the principal candidate ultimately has on school improvement and student outcomes."[12]

Illinois' approach is notable not only for a substantial increase in the rigor of expectations for programs, but also for a consistent effort to engage a wide array of stakeholders. Over the last 15 years, formal committees have had a hand in developing policy ideas, monitoring the quality of implementation on an ongoing basis, and suggesting tweaks to the rules and regulations. According to a case study on Illinois' principal preparation legislation, the involvement of stakeholders from the outset and their continued collaboration allowed the group to "capitalize on specific windows of opportunity" to advance their collective agenda.[13]

Delaware

Delaware also has prioritized program evaluation as a means for improving principal preparation. Regulatory changes adopted by the Delaware State Board of Education in 2014 require the Delaware Department of Education to develop scorecards for teacher and leader preparation programs.[14] The scorecards are based on data submitted by each program to the state and calculated from state data systems, and they supplement each program's accreditation by the Council for the Accreditation of Educator Preparation (CAEP). Taken together, the scorecards and the CAEP accreditation process will allow Delaware to assess the four critical dimensions of program evaluation described in design principle D1 above—inputs, processes, outputs, and graduate outcomes. Delaware is currently working with stakeholders to finalize the indicators that will be used for the first year of assessing leader preparation programs. The table below outlines how Delaware's draft indicators correspond to the four key dimensions of program evaluation.

---

11. See http://www.jcsee.org/program-evaluation-standards for more details on the standards.
12. See p. 4 of Baron, D., & Haller. A. (2014). *Redesigning principal preparation and development for the next generation: Lessons from Illinois*. Normal, IL: Illinois State University, Center for the Study of Education Policy.
13. Baron & Haller (2014), p. 21.
14. The regulations can be found at http://regulations.delaware.gov/AdminCode/title14/200/290.shtml

Table 1. Crosswalk of Delaware Draft Indicators and Recommended Indicator Categories

Recommended Indicator Categories[15]	Delaware Draft Indicators
Program Inputs: Indicators that reflect the program's ability to recruit and select high-potential aspirants and to diversify the pool of aspiring principals	• Selectivity in admissions • Candidates' prior teaching performance • Diversity of candidates
Program Processes: Indicators that reflect the quality of learning experiences for aspiring principals	• Accreditation by CAEP
Program Outputs: Indicators that reflect the success of aspirants in completing a rigorous program and being hired into principal and assistant principal roles	• Placement in administrative roles within one and three years • Placement in administrative roles in high-need schools • Retention in administrative roles
Graduate Outcomes: Indicators that reflect the impact that program graduates have, both on practices in the schools they lead and on student learning	• Improvements in culture and climate in schools led by graduates • Student growth in schools led by graduates • Percentage of graduates deemed highly effective on the state's administrator evaluation instrument • Perceptions of graduates' performance as measured by perceptual surveys of districts and program participants

The scorecards are designed to provide districts and candidates with comparable information about programs. Because making such information public is new and because presenting individual program and cross-program comparative data is challenging, a key aspect of the state's approach is to publish the information in year one (expected to be 2016) without any expectation of using it for decisions about program status. Once state leaders and stakeholders have had an opportunity to shape the content and formatting of the scorecards through use in the first year, data from the scorecards will be used to determine whether programs remain in good standing until their next CAEP accreditation (which occurs on a seven-year cycle) or whether they are given a probationary status.[16]

---

15. For more detail on these indicator categories, see the companion guide developed by New Leaders and UCEA.
16. This is the approach that Delaware is already taking with teacher preparation. Scorecards have already been published, along with a detailed description of the method used for choosing indicators (see http://www.doe.k12.de.us/Page/2573 for details). Delaware is currently working with stakeholders to finalize the indicators that will be used for the first year of assessing leader preparation programs.

New Leaders

University Council for Educational Administration
Quality Leadership Matters   *since 1954*

Lessons for Other States

Both Delaware and Illinois offer useful models and lessons for other states looking to improve their evaluation of principal preparation programs. One key lesson is the importance of state context. It is clear that each state's focus and pace necessarily will be influenced by current conditions in at least the following two areas.

A. Focus, alignment, and positioning of state leadership: The extent to which state leaders have prioritized school leadership—and specifically school leader preparation—in the state's educational improvement agenda, and the extent to which the state education agency (SEA) is positioned to be an effective resource for local education agencies and leadership preparation programs.

State education leaders (i.e., governors, legislators, state board of education members, chief state school officers, deans of schools of education, associations, and others) may be focused on a wide array of issues, ranging from the adequacy of state public education funding, to the content of student learning standards and assessments, to the quality of teacher evaluation systems. They also may differ on the relative priority of issues, let alone particular solutions to those issues. Ideally, these leaders and stakeholders share an understanding of how improved principal leadership would contribute to improved educational outcomes. If state political leaders and relevant agencies have a shared commitment to improving principal preparation programs, there are a number of steps that can be taken, such as the modification of existing policies that support program review processes and the targeting of funds to support implementation of a program evaluation system informed by the guidance offered in this document. Finally, meaningful improvement is more likely if the SEA is seen as a supportive partner interested in the improvement and innovation of preparation programs, and not just concerned with compliance.

B. Technical capabilities of the state education agency: The extent to which the SEA has crucial capabilities needed to support a new evaluation system, especially those related to data collection and the analysis and substantive review of programs.

In order to implement a strong evaluation system, a state needs a robust system of current data that includes important data on individual educators (e.g., their role, licensure status, evaluation ratings, etc.), enables tracking over time, and allows for connections between school-level data on leaders and preparation programs. Without these capabilities, a state might start small—for example, limiting the evaluation to available data (e.g., program input data collected and submitted by programs) and incentivizing programs to collect and report on their own output and impact data. However, data collected in this way should be interpreted with caution and not made public due to concerns about verification. Meanwhile, the state could invest in building a more robust data system.

In order to implement a strong evaluation system, a state also needs substantial capacity to compile, clean, and analyze data. This capacity is both a resource consideration, in that the state needs to fund the analytical capability, and an expertise consideration. Ideally, those conducting the analyses have experience in evaluating preparation programs, and particularly principal preparation programs. If the state does not have these resources, it might consider partnering with research institutions or consortia with data analysis capabilities.

Finally, effective implementation requires investment in program review. In particular, the state may need to train and maintain a cadre of reviewers if it intends to conduct periodic in-depth reviews of individual programs. If resources are limited, the state might consider limiting the number of programs requiring in-depth review on an annual basis or partnering with approved professional associations to conduct the in-depth reviews.

WHERE SHOULD STATES BEGIN? TOOLS AND RESOURCES FOR CONSIDERATION

With funding from The Wallace Foundation, UCEA and New Leaders have developed the *Guide to State Evaluation of Principal Preparation Programs* and a related set of tools and resources for states as they begin the process of developing a quality evaluation system for principal preparation programs. The guide recommends that states begin by diagnosing existing conditions related to the two areas discussed in the above section: the prioritization of school leadership preparation and the technical capabilities needed to conduct effective reviews. The readiness assessment rubric found in Appendix A is one of the tools developed by UCEA and New Leaders; it is designed to help states distinguish whether conditions are underdeveloped, workable, or ideal for the enhanced evaluation of principal preparation programs.

States might want to consult additional tools and resources in the toolkit. The resources provide background knowledge to support states in determining what factors to consider in designing a system that fits their context. The tools include model examples that can be modified and adapted according to specific state contexts.

Summary of the Tools and Resources Included in the New Leaders and UCEA Toolkit		
Tool/Resource	**Purpose**	**Description**
Guide to State Evaluation of Principal Preparation Programs	Provides states with detailed recommendations for the design and implementation of evaluation systems, as well as a road map for designing and implementing such systems based on an examination of their starting point.	Narrative, including design principles for systems of evaluation, a model for states to use and adapt, an overview of all the tools developed by UCEA and New Leaders, and recommendations for how to get started. Includes a two-stage model for evaluation. Stage 1 involves the collection of information about all programs in the state and the publication of results in an annual report; stage 2 involves an in-depth review of program practices and outcomes.
State Readiness Diagnostic Rubric	Helps states assess the extent to which they are ready to implement the recommendations provided in the guide.	Rubric detailing when conditions for this work are underdeveloped, workable, and ideal. Conditions addressed include the focus, alignment, and positioning of state leadership, as well as technical capabilities of the SEA.
Program Indicators, Rubric, and Report	Provides states with specifications for an annual report that would contain consistent information to help states decide when to initiate a targeted review. The annual report would also provide candidates and districts basic data on programs.	Suggested design for an online platform with data and information for each preparation program in the state, including the status of the program and when it is due for review, and annual data points for multiple measures in each area (inputs, processes, outputs, and program graduate outcomes).
Handbook for an In-Depth Review Process	Provides states and programs with clear processes and tools to support periodic in-depth reviews of programs.	Detailed process guide for undertaking reviews, including sections on data, process, the review team, and rubrics to assess programs.

New Leaders

University Council for Educational Administration
Quality Leadership Matters　*since 1954*

Summary of the Tools and Resources Included in the New Leaders and UCEA Toolkit		
Handbook for a Targeted Review Process	Provides states and programs with processes and tools to support targeted reviews in response to concerning data.	Detailed process guide for undertaking reviews, including sections on purposes, measures, process, and reviewer credentialing.
Overview of How Preparation Programs Are Currently Reviewed by States (as of 2015)	Provides states with information to compare their current principal preparation program evaluation system to recommendations proposed in the guide.	State-by-state summary tables of what each state requires for program approval and oversight, including the data that each state requires and an analysis of gaps between what data are currently collected versus what data are recommended for annual review.
Description of Other In-Depth Program Review Processes (as of 2015)	Provides states with background on the four types of reviews that higher education leadership preparation programs are likely to experience, so they can design their system with alignment to these other review systems in mind.	Narrative description of what each type of review entails, how often it typically occurs, and common challenges. The final section describes factors preparation providers consider to be beneficial sources of change.
List of Other Tools and Resources	Provides links and reference citations for tools, research, and resources created outside of the UCEA/New Leaders project.	Bibliographic list of tools and resources, including descriptions and directions for accessing them.

CONCLUSION

In an era of increased accountability for results and an ever-improving understanding of the role that school leaders play in improving student outcomes, states face significant choices in the design and implementation of strategies for improving school leadership. Some states, such as those highlighted in this document, have included the evaluation of principal preparation programs among their core strategies for improving school leadership. Others are considering doing the same. Florida, for example, has considered legislation to overhaul the process for evaluating principal preparation programs. Tennessee is considering a new approach to evaluation as well. As these states and others proceed along these lines, it is our hope that the design principles and examples provided here—as well as the tools and resources developed collaboratively by UCEA and New Leaders, with the substantial involvement of expert academics and practitioners—provide a solid foundation for their deliberations. This is challenging work, to be sure, but developing sophisticated approaches to assessing the quality of preparation programs and acting wisely on those assessments holds real potential to boost the quality of leaders in our schools.

Appendix A:
Readiness Assessment Rubric and
Process Recommendations

I. INTRODUCTION

Implementing a better system of evaluating principal preparation programs is complex work, and it requires that states have certain conditions and capabilities already in place. Before undertaking the work of designing and implementing a new evaluation system, we recommend that states assess their capacity to implement the recommendations in the guide. This tool is designed to help with that assessment of readiness. It has two parts: (1) a readiness assessment rubric and (2) process recommendations for completing the assessment.

The readiness assessment rubric includes information in two broad areas:

A. Focus, alignment, and positioning of state leadership: The extent to which state leadership has prioritized school leadership—and specifically school leader preparation—in the state's educational improvement agenda, and the extent to which the state education agency (SEA) is positioned to be an effective resource for local education agencies and leadership preparation programs.

B. Technical capabilities of the state education agency: The extent to which the SEA has crucial capabilities needed to support a new evaluation system, especially those related to data collection and the analysis and substantive review of programs.

The process recommendations outline how states might use information in these two areas to arrive at conclusions about their readiness to restructure or refine their assessment of leadership preparation programs. Completing this rubric will enable states to determine whether current conditions are ideal, workable or underdeveloped for implementing the recommendations in the guide.

When conditions are ideal, states may move forward with confidence. When conditions are workable in most areas, states may decide to move forward and work on improving conditions at the same time. When conditions are underdeveloped, states would benefit from developing supportive conditions before adopting our relevant recommendations. To move forward when conditions are underdeveloped would invite low-quality implementation and could unintentionally result in poor and potentially negative outcomes.

It is important to note that this is not a scientifically validated instrument. They do not lend themselves well to absolute determinations. Rather, it is a heuristic, allowing states to make sensible judgments about where to start and how fast to proceed.

A. Focus, Alignment, and Positioning of State Leadership			
	Underdeveloped	Workable	Ideal
A1. Commitment to improving school leadership			
State leadership prioritization: Public commitment by state leaders and key stakeholders to improving school leadership	State political leaders (governor, state chief, state board, legislative leaders) rarely discuss school leadership as a way to improve schools. Stakeholders (e.g., associations, prominent local education agency [LEA] leaders, university leaders) have major disagreements on the importance of leadership.	State political leaders communicate about school leaders as one among many issues of concern. Stakeholders have a broad array of perspectives on the importance of leadership.	School leadership is a top agenda item for state political leaders. Stakeholders' perspectives are highly aligned around the importance of school leadership
Investment in leadership: Visibility of school leadership in state strategic plan and in allocation of resources	The state's strategic plan says very little about strategies to improve school leadership. No discretionary dollars are allocated to improving school leadership, and no effort is made to encourage local investments in leadership.	School leadership is in the strategic plan but is a secondary priority or one on a long list of stated priorities. Investments in leadership are regularly communicated as allowable expenditures in state grant programs (as appropriate).	School leadership is a major focus of the state's strategic plan; the state has a clear understanding of how improved school leadership will contribute to improved educational outcomes. The state has targeted funds (including public and privately sourced funds) to specific efforts to improve school leadership.
A2. Commitment to improving principal preparation			
State leadership alignment: Unified stance of state leaders to improve principal preparation	Lines of authority for improving principal preparation programs are divided or ambiguous. Agencies with responsibility for principal preparation and licensure (e.g., SEA, professional licensing boards) have limited communication and differing priorities.	Lines of authority for improving principal preparation programs are clear. Agencies with responsibility for principal preparation and licensure have goals and strategies that do not conflict, and they communicate regularly.	State political leaders and relevant agencies are unified in a commitment to improving principal preparation programs and agree on the need to rigorously assess the quality of programs, help programs improve, and take action to address underperformance. Agencies with responsibility for principal preparation and licensure have shared goals and are committed to collaboration with each other and with programs (especially when it comes to sharing data).
Policy framework: Policies in place to foster innovation	There is little effort by state political leaders to influence the practices of principal preparation programs.	The state's policy framework allows for innovative principal preparation program design.	The state pursues new partners and encourages universities and other providers to create innovative principal preparation programs.

NewLeaders

	Underdeveloped	Workable	Ideal
A3. Commitment to and capacity for continuous improvement[17]			
Collaboration: Perceptions of the SEA as a collaborative partner	LEAs and preparation programs have little interaction with the SEA, viewing the agency as primarily concerned about compliance with statutes and regulations.	LEAs and preparation programs have mixed interaction with the SEA: Some interactions are overly driven by compliance concerns, while other interactions are focused on genuine improvement and making rules work for them.	LEAs and preparation programs view the SEA as a trusted partner committed to continuous improvement; compliance still matters, but the SEA works to make it as seamless as possible.
Communication: SEA systems for communication with partners	Information coming from the SEA to LEAs and preparation programs is either nonexistent or perceived by programs as excessive and disjointed, often sending mixed messages.	Information coming from the SEA to LEAs and preparation programs is perceived by programs as organized, predictable and reasonably clear.	The SEA convenes local partners in ways that foster two-way communication.
Innovation: Perceptions of the SEA as a source of ideas	LEAs and preparation programs do not look to the SEA for new ideas to improve schools and universities.	The SEA serves as an effective information clearinghouse, making innovations in the field visible to LEAs and preparation programs.	The SEA shares data; engages LEAs and preparation programs in conversations about improvement; and offers new learning opportunities, including creative strategies for implementing federal and state policy.
Decision making: Use of evidence in SEA decisions	The SEA offers little explanation or unclear justification for policy changes.	The SEA reports on data used in the design of new policies and articulates the reasons for policy changes.	The SEA transparently shares data, data analysis, and operating theories that underlie policy design and implementation decisions.
Expertise: Knowledge and skills to manage change process for leadership work	LEAs and preparation programs view SEA leadership as having limited understanding of core leadership issues and as being unresponsive or unhelpful in managing the process of large-scale change.	SEA leadership communicates a solid understanding of the connections between leadership and student outcomes, as well as the adaptive challenges associated with large-scale change.	SEA leadership communicates a strong understanding of—and solutions for—the adaptive challenges associated with large-scale change. SEA leadership is deeply involved in national and statewide conversations about the practice and impact of school leaders.

17. Determined through anonymous surveys of program leaders.

New Leaders

B. Technical Capabilities of the State Education Agency

B1. Data and data system requirements

	Underdeveloped	Workable	Ideal
Program data system: System that collects program data (e.g., number of applicants, clinical hours required, 100-word description) from preparation programs	Data are available in isolated locations without an overarching system for integrating the different sources or linking the data points.	A program data system is in place, but it may not include all data points needed for the SEA's annual report; some data may be missing, inaccurate, or lack comparability. Systems support might be needed to design new tools or interfaces to collect needed information from multiple sources and/or agencies. Substantial budgeting would be required for staff time to request missing data, monitor data completion, and build necessary data systems. Time is allocated to ensure data integrity.	A program data system is in place and includes all fields/variables needed for the SEA's annual report. The system enables consistent reporting and data aggregation. Data are complete and accurate. Programs use common definitions of indicators, making the data comparable across programs. The system is not overly burdensome for programs, districts, or school partners.
Placement data systems: Systems that track individual educators and their annual placement role (teacher, principal, assistant principal, other school leader, district leader)	Data are available in isolated locations without an overarching system for integrating the different sources or linking the data points.	Placement data systems exist and are coordinated but have lots of inaccuracies and missing data. Budgeting would be required for staff time to request missing data and monitor data. Time is allocated to clean data.	Placement data systems are complete and accurate.
Unique identifiers for program participants: Identifiers that link data from preparation programs, licensure status, placement data systems, and effectiveness ratings from educator evaluation system	It is not possible to link individuals across data systems (for programs, licensure, placements, school outcomes).	Unique identifiers do not exist, but it is possible to link two or more data systems, and the SEA has the capacity to do this. Budgeting would be required for junior analyst time to link data systems.	Unique state-level identifiers are in place to link individuals to all of the data required by the evaluation system.
Comparable survey data: Common survey administered to program graduates that gathers their perceptions of program process indicators	Graduates of most programs are not surveyed, or the response rates are too low to make results meaningful.	Surveys of program graduates exist, and response rates are reasonable, but the surveys differ, preventing comparison of data across programs.	A common survey is administered to all graduates in the state with reasonable response rates, enabling comparison of data across programs.

NewLeaders

	Underdeveloped	Workable	Ideal
Measures of teacher and leader effectiveness: Ratings of individual teachers and principals on the state performance evaluation system	Measures of teacher and leader effectiveness: Ratings of individual teachers and principals on the state performance evaluation system.	Measures exist and have some variability but lack validity and reliability. The SEA does not share results publicly and does not provide caveats that caution users on interpretation.	Measures exist, have variability, and have been found to be both reliable and valid. The SEA has the capacity to use measures in contextually appropriate ways. The SEA ensures that any public release of data meets federal and state privacy guidelines.
Measures of student learning gains: Student achievement scores across grade levels in core subject areas	Measures are not based on individual student growth from year to year.	Consistent and methodologically sound measures of individual student growth, including proper controls for student- and school-level variables, exist, but they are not comprehensive across grade levels and subject areas. Measures exist, but ns are small (less than 10 individuals) for most programs.	Consistent and methodologically sound measures of individual student growth, including proper controls for student- and school-level variables, exist. These measures allow for assessment of school leaders' influence on student learning after three years at a school site. Adequate consideration is given to bias against high-needs schools.
B2. Data compilation and analysis capacity			
Monitoring data reporting completion and accuracy: Requires staffing to ensure the submission and accuracy of data from preparation programs and other data sources	No staff or resources exist.	Staff assignments and/or resources could be prioritized for data monitoring.	Staff and/or resources are already assigned to data monitoring.
Creating and publishing annual reports: Requires technical skill for website/report design and senior analytical skill to make methodological decisions	No staff or resources exist.	Staff assignments and/or resources could be prioritized for data reporting.	Staff and/or resources are already assigned to data reporting.
Creating and implementing methodology for summative rating: Requires specialized assessment and statistical skill	No staff or resources exist.	Staff assignments and/or resources could be prioritized for data analysis/methodology.	Staff and/or resources are already assigned to data analysis/methodology.

New Leaders

	Underdeveloped	Workable	Ideal
B3. Review process capabilities			
Staffing: Requires specialized leadership experience and skills	There is no SEA staff committed to leadership preparation, or those responsible have multiple other roles.	There are staff members at the SEA focused on school leadership, including preparation, but they have limited backgrounds in school leadership or adult leadership.	There are staff members at the SEA focused on school leadership, including preparation, and they are deeply credible with leaders and preparation providers in the state.
Management and training of reviewers: Requires specialized review process capabilities	No staff or resources exist.	The state has a reasonably adequate pool of high-quality, credible reviewers but does not have a track record of systematically vetting them for leadership expertise or training them for inter-rater reliability. The state does not have a strong track record of outsourcing functions and maintaining quality.	The state has a robust pool of high-quality, credible reviewers who have been (or could be) trained for inter-rater reliability and normed to provide useful feedback to programs. or The state has a strong track record of out-sourcing functions and maintaining quality. This allows for bringing in an established process (e.g., review by the Educational Leadership Constituent Council).
Implementation of reviews: Requires financial and human resources	No staff or resources exist.	A review process exists, but it is not sufficient for quality, in-depth review of all programs flagged.	Sufficient resources exist to carry out in-depth reviews for all programs flagged as needing it, and for conducting periodic reviews of all programs.

NewLeaders

III. PROCESS RECOMMENDATIONS

The readiness assessment rubric can be used in more than one way. A state working to build political support for an evidence-based approach to assessing the quality of principal preparation programs may want a formal process to engage stakeholders in completing the rubric and agreeing on next steps for the work. Meanwhile, a state already committed to an evidence-based approach to assessing the quality of principal preparation programs may want the SEA to simply undertake an internal diagnosis of conditions in order to surface critical gaps and needed resources.

For a more extensive process, these general steps are recommended:

1. Create a vision for the work. In order to demonstrate executive-level commitment to an open and honest process of assessing the state's readiness for implementing a better system of evaluating principal preparation programs, it can be helpful to write a purpose statement outlining why the work is important and how it connects to the state's broader vision of leadership. The state's strategic plan for education is an important resource for this step.

2. Create a project plan. In order to ensure that the right people will be engaged and will have access to authentic information, it can be helpful to craft a project plan that includes roles and responsibilities and to assemble available data to conduct the readiness assessment.

3. Convene stakeholders. In order to build trust in and commitment to the process, it can be helpful to convene leaders from universities, preparation programs, administrator associations, districts and schools. The purpose of such a convening is to share the goals and work plan, ask for authentic feedback, and ask for a commitment to participating in the process.

4. Conduct the assessment. This is the heart of the work: gathering data, making sense of it, surfacing and discussing important substantive issues and agreeing on rubric ratings.

5. Set action steps. With the assessment complete, state leaders and stakeholders need to make decisions about their readiness and identify areas of focus that are consistent with the conclusions from the readiness assessment. This is also an opportunity to establish a new work plan for the implementation phase, including strategies for addressing any areas of weakness that need to be remedied in the short term.

For a more targeted approach within an SEA, the critical steps are numbers 4 and 5 above, as well as some amount of stakeholder engagement (step 3). Note, however, that some categories of the rubric require information from sources outside of the SEA (e.g., perceptions held by LEA leaders and program leaders), so some level of external engagement is helpful regardless of the scope of the analysis.