The Equitable Distribution of Effective Teachers

Can States Meet the Research Challenges Required for Success?

Allen Education LLC

Michael Allen, Ph.D

October, 2011

Please direct inquiries about this publication to:

CNA Education
4825 Mark Center Drive
Alexandria, VA 22311 USA

Phone: 703.824.2000
Fax: 703.824.2542
E-mail: relappalachia@cna.org
Website: www.relappalachia.org

Table of Contents

Introduction

One of the provisions of the No Child Left Behind Act that has gained prominence as a policy focus only in the last several years is the requirement that poor and minority children be served by highly qualified teachers to no less a degree than other, more affluent children. Under the Obama administration, the focus on "highly qualified" teachers has shifted to "effective" teachers, and the push for states to ensure equitable distribution is a key requirement for the State Fiscal Stabilization Fund program and Race to the Top (RTT) funding, as well as a key piece of the administration's Blueprint for the reauthorization of the Elementary and Secondary Education Act (ESEA).

Complying with the equitable distribution provision is a formidable challenge for states under the best of circumstances. It may involve the renegotiation of teacher union contracts in order to make transfer and seniority provisions less of an obstacle to teacher reassignment. It may meet the reality that in fields such as bilingual education, science, and mathematics, there are not enough qualified and effective teachers to go around. Or it may face the prospect that there are schools, either because of location or reputation, at which effective teachers simply do not want to teach and will not stay in spite of incentives they may be offered. Given the tough economic conditions that states across the country face at the present time, the difficulty of complying with the equitable distribution requirement is even greater as states and districts struggle to avoid teacher layoffs and are hard-pressed to find the extra cash required to lure successful teachers to schools in need and keep them there.

In addition to such concrete obstacles, however, states also face challenges of a more epistemological nature. Some challenges, such as conceptual clarification and understanding the implications of the relevant research base are identical for all states. The very terms "effective teachers" and "equitable distribution" are inherently fraught with ambiguity, and even definitions stipulated by the Department of Education don't completely resolve the problem; states are still required to decide on acceptable definitions of these and other concepts and must understand the implications of the definitions they choose.

Closely related to the need to define terms is the need to develop measures of progress toward providing good teachers for all students, measures that are grounded in a solid knowledge of the related research. What guidance does the research literature offer to ensure that states ultimately will settle on valid and reliable measures of teacher

effectiveness and distribution? Similarly, what is known from research about the success of compensation and other incentives that have been employed around the U.S. in an effort to recruit and retain capable teachers for schools that have been historically hard to staff?

Another epistemological challenge for states is gaining a solid knowledge of their educational picture that is sufficiently accurate, nuanced, and comprehensive that they can assess accurately their progress towards achieving equitable distribution, determine the impact of specific measures enacted toward that goal, and identify any unforeseen consequences these efforts may occasion. This requires not only sufficiently robust data, but also careful analysis of those data and various research studies to illuminate the impact of current programs, practices, and policies.

Being able to meet these epistemological challenges is a tall order for most states. The challenges are ongoing, and addressing them successfully requires consistent and long-term data reporting and analysis. It also requires continual monitoring of the success or failure of the various efforts employed to achieve equitable distribution and raise the quality of teacher workforce. Given that state agencies have limited research and analytic capacity in the best of times, states likely must find some way of augmenting their research and analytic capacity out-of-house if they are to derive maximum benefit from their efforts.

CNA has provided some of this external capacity in the four states it serves through its operation of REL Appalachia, but as of yet it has contributed little direct assistance to those states' efforts to achieve equitable teacher distribution. Indeed, through work in these client states, CNA has come to comprehend the significant research and analytical burden teacher distribution and many other reform efforts impose on states and the capacity gap in these areas that must be closed if state education reform efforts ultimately are to realize their full potential for success.

This discussion highlights some of the key research-related challenges that the REL Appalachia states and others are facing in their efforts to achieve more equitable distribution of effective teachers. It draws on the perceptions of the staff of CNA Education, many of whom serve client states as researchers directly embedded in state departments of education – a unique technical assistance model that CNA employs in all of its research areas, both military and civilian. And it also draws on interviews with state education officials and the study of relevant state documents by this paper's author.

Challenge #1: Conceptual Clarification

The very terms “effective teachers” and “equitable distribution” are inherently fraught with ambiguity, and even stipulated definitions offered by the U.S. Department of Education in its guidelines for the Race to the Top competition and its Blueprint for the Reauthorization of the Elementary and Secondary Education Act don’t completely resolve the problem (See Appendix A). The fact that states must accept these definitions to be eligible for federal funding does not mean that the definitions are wholly satisfactory. Indeed, they are imprecise and by design leave room for individual states to supplement and sharpen them.

As it is, the Department’s stipulated definition of effective teaching raises a number of questions that states need to answer – ideally informed by careful analysis and a good knowledge of the relevant research. For example:

- Can all teachers in all schools be reasonably expected to elicit at least one year of student achievement growth from all students?
- Besides the change in student achievement test scores, what other measures of student learning can a state validly employ?
- What other measures of teacher effectiveness, in addition to their students’ learning growth, are appropriate to include in the evaluation of teacher effectiveness?
- How should the various measures of both student learning and teacher effectiveness be weighted in the total assessment?

Defining the meaning of “equitable distribution” so that it is both sufficiently rigorous and actionable presents similar challenges. Here, too, the RTT application guidelines leave states with the responsibility for answering critical conceptual questions:

- How should the quality of the teaching staffs be compared at two schools if they have the same percentage of effective teachers but one school has a much larger percentage of both highly effective and highly ineffective teachers while the other school has many teachers who are almost effective but few who are at the upper or lower end of the spectrum?
- Does the definition still leave the possibility of significant inequity in teacher quality between schools with the same percentage of effective and highly effective teachers? Might teachers of high-performing students, for example,

whose achievement tests scores are already far above grade level and difficult to move higher, be far more capable than their effectiveness rating reflects?

States in the REL Appalachia region are indeed trying to address the difficult conceptual questions that attend the effort to promote the equitable distribution of effective teachers. Virginia, for example, is seeking to clarify the legitimate use of value-added measures as it further develops its teacher evaluation system and is wrestling with complex issues like the following: [1]

- Can a teacher's effectiveness be evaluated on the basis of the students' performance on state achievement tests in a single year, or does it require test scores over multiple years?
- What percentage of the overall assessment of a teacher's effectiveness should be constituted by her students' test performance, and what percentage by other measures?

Tennessee is wrestling with related issues:[2]

- How is effectiveness to be determined for teachers (such as those in art or social studies) who have no student achievement test data to use as a basis for assessment?
- How frequently during the course of a school year should teachers be evaluated?
- What assessment scores will teachers have to earn to qualify for designation as effective or highly effective?

[1] D. Jonas (Executive Director for Research and Strategic Planning, Virginia Department of Education), personal communication, August 23, 2010.
[2] K. Graham (Field Scientist, REL Appalachia), personal communication, July 28, 2010.

Challenge #2: Solid Understanding of Research

Inseparable from the need for basic definitional and conceptual groundwork in states' efforts to promote the equitable distribution of effective teachers is gaining a solid understanding of the relevant research in several different areas and its implications for the teachers' task.

Some research studies can contribute greatly to a state's ability to develop precise and effective definitions and fair and valid assessment criteria. Research on the validity or efficacy of various types of student and teacher assessments that already have been developed and employed would be particularly important in this respect. These include analytical studies of value-added and other forms of assessment and empirical studies of various teacher observation protocols. [3,4] Especially for states that face serious limitations on their ability to define teacher effectiveness on the basis of student test scores, understanding the research on the relationship between various teacher characteristics (e.g., academic ability, experience, certification status) and student performance would be extremely helpful in developing proxy indicators of likely teaching success.[5] Such indicators could not be used for purposes of teacher evaluation, but they could prove useful in estimating the relative strength of teachers in a particular school or district or throughout the state.

Other research studies can be of significant value in developing specific programs and strategies to facilitate more equitable teacher distribution. These would include studies on the impact of various kinds of teacher recruitment and retention incentives or on the impact of working conditions on the ability of schools to attract and retain capable

---

[3] See, for example, H. Braun, (2005). *Using student progress to evaluate teachers: A primer on value-added models.* Princeton, NJ: Educational Testing Service; L. Goe, C. Bell, and O. Little, (2008). *Approaches to evaluating teacher effectiveness: A research synthesis*. Washington, DC: National Comprehensive Center for Teacher Quality; and D. McCaffrey, J.R. Lockwood, D. Koretz, and L. Hamilton, (2003). *Evaluating value-added models for teacher accountability*. Santa Monica, CA: Rand Corporation.

[4] See, for example, L. Goe et al., 2008; K.M. La Paro, R.C. Pianta, and M. Stuhlman, (2004). The classroom assessment scoring system: Findings from the kindergarten year. *The Elementary School Journal, (104)*5, 409-426; and A. Milanowski, (2004). The relationship between teacher performance evaluation scores and student achievement: Evidence from Cincinnati. *Peabody Journal of Education, (79)*4, 33-53.

[5] See, for example, L. Cavalluzzo, (2004). Is National Board certification an effective signal of teacher quality? Alexandria, Virginia: The CNA Corporation; D.P. Mayer, J.E. Mullins, and M.T. Moore, (2000). *Monitoring school quality: Indicators of quality.* Washington, DC: National Center for Education Statistics; and J. B. Presley, B.R. White, B. R., and Y. Gong, (2005). *Examining the distribution and impact of teacher quality in Illinois.* Policy Research Report IERC 2005-2;

teachers. For states that have a shortage of teachers in particular fields or locations and cannot count on replacing ineffective teachers with new teachers who may be more effective, it would also be valuable to know whether the research on specific curricula indicates that some curricula are likely to promote the effectiveness of current teachers. Similarly, it would be important for state education leaders to have a grasp of the research on professional development and on pre-service teacher preparation in order to identify approaches that have proven particularly successful in increasing teacher effectiveness.

Recognizing the complexity of the conceptual issues the state is confronting in its efforts to improve its teacher assessment system and the valuable insights that research might afford in clarifying those issues, the Virginia Department of Education has commissioned a synthesis of current research on comprehensive teacher evaluation models. In addition, the state is looking at the research on differentiated and performance-based teacher compensation models in anticipation of possible district-level interest in linking the state's evaluation system to local compensation systems.[6] Kentucky, too, is studying the research on assessment systems, in particular value-added assessment.[7]

[6] J.W. Lanham (Director of Licensure and School Leadership, Virginia Department of Education), Report on the study and development of model teacher and administrator evaluation systems. Virginia Board of Education meeting item #1, July 22, 2010.
[7] Michael Dailey (Director of Educator Quality and Diversity, Kentucky Department of Education), personal communication, August 27, 2010.

Challenge #3: Solid Grasp of the State Education Picture

With the powerful education data systems that many states are developing, it might be taken for granted that those states have a precise and thorough grasp of their education picture, including the academic performance of their students, the adequacy of their teacher workforce in both quality and supply, the quality of their pre-service and in-service professional development efforts, the impact of various reform initiatives underway, etc. This is not necessarily the case, however, because even if states have the benefit of rich sets of education data they may not have the ability to undertake the kind of sophisticated data analysis and the research studies based on that data that are ultimately required to provide them with detailed and accurate understanding of the status of their education enterprise.

First and foremost, of course, states must have adequate student and teacher data. Working with the Data Quality Campaign, many states are moving toward the development of comprehensive student data systems.[8] But even states with finished systems may not have historical data that would enable them to identify long-term progress or important trends. And very few states have a comprehensive *teacher* database – a data source that is only now emerging as an important priority nationally.[9] Without solid teacher data, states cannot possibly develop an adequate assessment of the quality and supply of their teacher workforce – information that is absolutely essential to long-term efforts to ensure the equitable distribution of capable teachers.[10]

Even with the best data, there is a great deal of careful analysis that states must do before they can have confidence in their assessment of where they currently stand educationally and whether their efforts to achieve equitable teacher distribution are truly adequate and efficacious. States must ensure, for example, that their efforts to evaluate teachers' effectiveness in terms of the achievement test performance of their students

[8] The Data Quality Campaign identifies ten core elements of a longitudinal data system, and it supports and monitors the progress of states in implementing those elements. See http://www.dataqualitycampaign.org/# .
[9] The Data Quality Campaign has joined with several partners to issue a template for teacher data entitled "Leveraging State Longitudinal Data Systems to Inform Teacher Preparation and Continuous Improvement." Released in August, 2010, the template can be retrieved at http://www.dataqualitycampaign.org/resources/details/1008.
[10] See M. Allen, (2010, April). A guide to teacher data. In *Improving state need assessments of secondary science and mathematics teachers: Challenges, possibilities, and recommendations.* Washington, DC: Association of Public and Land-grant Universities. Can be retrieved at http://state-needs.teacher-imperative.com/.

are methodologically valid and reliable. On the one hand, this implies the need to ensure the methodological validity of the statistical model states develop to link the achievement of students with their teachers. In particular, states need to eliminate the possibility that the model is insufficiently sensitive to non-teacher effects that may influence students' performance but are mistakenly attributed to teachers. Some accountability models, for example, may unwittingly penalize teachers who teach high numbers of transient students or who teach in schools with exceptionally high-achieving students or with a high percentage of ineffective teachers.[11]

On, the other hand, it implies the need to ensure the validity and reliability of the achievement tests themselves. This is not a foregone conclusion. Apart from the larger controversy over the adequacy of standardized tests as a valid or sufficient indicator of student learning, states must be wary of a number of potential threats to the tests' validity and reliability that studies have identified, e.g., the timing and scaling of the tests, score inflation, and measurement error.[12]

Also important to states' efforts to gain an adequate understanding of their student and teacher performance profile is a comparison of their students' and teachers' performance with the performance of those in other states. This is not an issue of one-upmanship but of ensuring that a state's appraisal of the success of its students and teachers is not artificially inflated due to standards that are much lower than those of other states. Some studies[13] find that standards for student proficiency on state student achievement tests are not uniform from state to state. Similarly, some states have been criticized for having much lower teacher licensure standards than other states.[14]

Finally, states must undertake ongoing research and analysis in order to assess both short-term and longer-term success of the overall effort to achieve equitable teacher distribution and the specific impact of the different strategies employed. This implies the ability to determine not only changes in teacher distribution itself but also the impact of the changes in teacher distribution on student achievement – both in the schools that gained stronger teachers and in the schools from which the stronger teachers may have been recruited away. It also implies the need for states to identify potential indirect effects of the efforts to achieve equitable distribution, such as changes in the state and local

[11] See , for example, D. McCaffrey, J.R. Lockwood, D. Koretz, and L. Hamilton, (2003). *Evaluating value-added models for teacher accountability*. Rand Corporation, Santa Monica, CA.
[12] See, D. McCaffrey et al., (2003), esp. pp. 87-107.
[13] See V. Bandeira de Mello, C. Blankenship, and D. McLaughlin, (2009, October). *Mapping state proficiency standards onto the NAEP Scales: 2005-2007*. Washington, DC: National Center for Education Statistics.
[14] See *Not good enough: A content analysis of teacher licensing examinations*. (1999, Spring). Washington, DC: The Education Trust.

attrition rates of teachers or in the socioeconomic makeup of students in schools where there have been significant shifts in staffing.

Several states in the REL Appalachia region are taking steps to improve their ability to grasp the conditions of education in their state. West Virginia, for example, is in the process of implementing a longitudinal student data system and is presently focusing in particular on strengthening the link between student and teacher data and on linking its K-12 and post-secondary data. [15] The state is also investigating what kind of teacher data it needs to collect in order to gain a clearer picture of its teacher supply and demand situation, and it is developing a student growth model that will permit a value-added assessment of teacher effectiveness.

Tennessee is a Race to the Top grant recipient that already has a sophisticated and much-researched value-added teacher assessment system and has fully met the Data Quality Campaign's ten criteria for a longitudinal student data system. Nevertheless, the state is in the process of expanding its data system to provide a more comprehensive "360 degree view" of each student that will include data on students' physical and mental health, student and social services used, family circumstances, and delinquency. On the teacher side, the state plans to collect comprehensive data about working conditions. And because the state already has in place longitudinal statewide data on teacher effectiveness, it has undertaken significant analysis of the distribution of effective and ineffective teachers in each school across the state.[16]

[15] T. Geraghty (Field Scientist, REL Appalachia), personal communication, August 30, 2010.
[16] Tennessee Race to the Top application for initial funding (plus Appendix), (2010). Retrieved August 30, 2010 from http://www2.ed.gov/programs/racetothetop/phase1-applications/index.html. See especially Appendix, Sections C-1-1 and D-3-(1-9).

Conclusion

Clearly, the attempt to achieve a more equal distribution of effective teachers is a highly complex undertaking that imposes a heavy responsibility on states for research and research-related analysis. Senior state officials understand the critical importance of research to the policy development and policy evaluation process, and in the case of equitable teacher distribution there is likely not a single state that isn't bringing research and analysis to bear on the challenge. The question is, to what extent are individual states' research efforts satisfactory? This is a function of the adequacy of awareness, on the one hand, and of resources, on the other.

Several recent reports document the limited ability that state education agencies have to support statewide school improvement efforts, and for at least some states that limitation extends to an in-house capacity to do research.[17] Until recently, one state agency in the REL Appalachia region had too small a research staff to do much more than derive inferential statistics from the data available.[18] More time-intensive research studies are a luxury that few state departments of education seem able to afford, in part because state agency personnel must have the flexibility to respond to whatever front burner issue or issues require their attention at a particular moment.

The inability of many state education agencies to muster a strong in-house research program to address such priorities as the equitable distribution of teachers is not strictly a capacity issue, however. Nor, is it necessarily an issue of a lack of research expertise within the agencies – though expertise is certainly compromised when staffing is meager. Rather, even if they have reasonable capacity, state departments of education must deploy their staff to focus on the progress or support of high-profile or high-stakes state programs that are in the political spotlight or on gathering and interpreting the many kinds of data that contribute to the statewide education picture. In other words, the overarching issue is one of focus; given the ongoing priorities and political realities that face state

[17] K.C. LeFloch, A. Boyle, and S.B. Therriault. (2008, September). *Help wanted. State capacity for school improvement.* Washington, DC: American Institutes for Research. [Retrieved 11-16-10 from http://www.air.org/files/Research_Brief_I-State_Capacity_for_School_Improvement_091508_r1.pdf]; and B. Unger, B. Lane, E. Cutler, S. Lee, J. Whitney, E. Arruda, and M. Silva. (2008). How can state education agencies support district improvement? Providence, R.I.: The Education Alliance. [Retrieved 11-16-10 from http://www.alliance.brown.edu/pubs/csrqi/symposium.pdf].

[18] This has been true for the state of Virginia, for example, which only recently hired a statistician to provide more analytical capacity. [Personal communication with Deborah Jonas, Executive Director for Research and Strategic Planning, Virginia Department of Education, August 23, 2010.]

education agencies at any given time, it is simply difficult for staff researchers to broaden their perspective.

This is where external researchers can add important capacity to state agencies. If they have broad knowledge of important education research and familiarity with reform efforts in multiple states, they can bring new insights to bear on existing state priorities. With no stake in the outcome of various state reform efforts, external researchers can provide a credible, neutral perspective on those efforts that may be convincing to policy-makers and other stakeholders. Unencumbered by current state priorities, external researchers can help a state address new issues on its waiting list by providing important background research and other useful information. And, finally, external researchers can conduct sophisticated, multi-year research studies for a state that it would be virtually impossible for state agency researchers to undertake, given both resource limitations and the need for agency researchers to move quickly from one priority to another.

In the case of equitable distribution of effective teachers, few states are in the enviable position of Tennessee. Not only has the state had well over a decade of both in-house and external research dedicated to its groundbreaking value-added assessment system, as the recipient of $500 million in Race to the Top funds, the state is able to establish a state Consortium on Research, Evaluation, and Development (TN CRED) that should provide strong research support for Tennessee's reform efforts at least over the next several years.

The state's successful application for Race to the Top funding describes the consortium in this way:

> The group consists of expert researchers and practitioners from throughout Tennessee and beyond whose task is to identify the full research needs of our proposal based upon the assurance areas and assist in creating the learning agenda for our state. Relative to the items above, TN CRED will work to identify the research projects, engage with partners as necessary to accomplish the work, and link this back through distribution of the system support good practice. It also will recommend refinements to programming where beneficial, delve deeper into questions as necessary, and work with partners as they implement research programs that advance our knowledge of what education reform works and what investments should be abandoned. TN CRED will also engage with the First to

the Top Oversight Team to fully inform policy makers of research findings and engage in a dialogue that supports thoughtful policy making.[19]

Tennessee's efforts to achieve an equitable distribution of effective teachers may or may not succeed, but if they fall short, it is not likely to be because the state made an inadequate commitment to research. The same thing cannot be said with confidence about other states; many states may well fail to achieve substantial equity in the distribution of capable teachers in large part because they did not do the analysis and research necessary to ensure the efficacy of their policy choices. There are, however, several measures that may help avoid this outcome:

1. Just as it has contributed substantial funding and provided support to states for the development of comprehensive longitudinal data systems, the federal government needs to make significant funding available for state-level research and technical assistance to help states derive maximum knowledge from the data they're collecting.
2. With whatever resources are available, states need to develop a strategic research agenda that provides the biggest bang for the buck in terms of the light it is likely to shine on their efforts to achieve equitable teacher distribution. The research community may be able to provide important guidance in this task.
3. States need to learn from one another. Although every state is different, there are many structural features that are common to their education enterprise. Thus, a state should be able to draw on the related research and analysis carried out by other states, and several states working together may even be able to devise a way to develop a joint research base. The teacher supply and demand situation, for example, is unique to every state. But the methods for developing reliable supply and demand projections are universal, and states can borrow and adapt them from each other.

[19] Tennessee Race to the Top application for initial funding, (2010). Retrieved August 30, 2010, from http://www2.ed.gov/programs/racetothetop/phase1-applications/index.html, pp. 73-74

Appendix A: Equitable Distribution of Effective Teachers: Federal Definitions

In its Race to the Top Application for Phase 2 Funding, the U.S. Department of Education defines an effective teacher as one "whose students achieve acceptable rates (e.g., at least one grade level in an academic year) of student growth" (p. 7), and a highly effective teacher as one "whose students achieve high rates (e.g., one and one-half grade levels in an academic year) of student growth" (p. 8). Student growth, in turn, is defined as "the change in student achievement for an individual student between two or more points in time" (p. 11). And student achievement is defined in the application as follows (p. 11):

> (a) For tested grades and subjects: (1) a student's score on the State's assessments under the ESEA; and, as appropriate, (2) other measures of student learning . . . provided they are rigorous and comparable across classrooms.
>
> (b) For non-tested grades and subjects: alternative measures of student learning and performance such as student scores on pre-tests and end-of-course tests; student performance on English language proficiency assessments; and other measures of student achievement that are rigorous and comparable across classrooms.

Later in the same application, the Department of Education notes that student growth is but one measure of effective teaching and asks the states "to ensure that LEAs [d]esign and implement rigorous, transparent, and fair evaluation systems for teachers and principals that (a) differentiate effectiveness using multiple rating categories that take into account data on student growth . . . as a significant factor" (p. 33).

As for "equitable distribution," the Department of Education indicates in the RTT application that it is to be defined in terms of the percentage of effective and highly effective teachers serving high-poverty and/or high-minority schools (p. 38).