



American Federation
of Teachers, AFL-CIO

AFT Teachers
AFT PSRP
AFT Higher Education
AFT Public Employees
AFT Healthcare

555 New Jersey Ave. N.W.
Washington, DC 20001
202/879-4400
www.aft.org

Randi Weingarten
PRESIDENT

Antonia Cortese
SECRETARY-TREASURER

Loretta Johnson
EXECUTIVE VICE PRESIDENT

VICE PRESIDENTS

Shelvy Y. Abrams
Mary J. Armstrong
Barbara Bowen
Linda Bridges
Kenneth Brynien
Elsie P. Burkhalter
Stacey Caruso-Sharpe
Kathy Chavez
Lee Cutler
Edward Doherty
Kathleen M. Donahue
Thomas A. Dooher
Eric Feaver
Andy Ford
Ed Geppert, Jr.
David Gray
Judy Hale
David Hecker
Richard Iannuzzi
Jerry T. Jordan
Dennis Kelly
Ted Kirsch
Francine Lawrence
Alan Lubin
Louis Malfaro
Merlene Martin
Michael Mulgrew
Maria Neira
Ruby Newbold
Candice Owley
Sharon M. Palmer
Marcia B. Reback
Laura K. Rico
Pat Santeramo
Sandra Schroeder
Phillip H. Smith
Marilyn Stewart
Ann Twomey
Adam Urbanski

August 28, 2009

The Honorable Arne Duncan
U.S. Department of Education
400 Maryland Avenue, SW
Washington, DC 20202

ATTN: Race to the Top Comments
Docket ID: ED-2009-OESE-0006

Dear Secretary Duncan:

Last month, the U.S. Department of Education (ED) released its notice of proposed priorities, requirements, definitions and selection criteria for the Race to the Top program, authorized as the Incentive Grants under the American Recovery and Reinvestment Act of 2009 (ARRA).

At that time, the American Federation of Teachers (AFT) said that we would use the following criteria to review the proposals in the draft notice: Do they help students? Are they fair to teachers? Are they transparent to the public? Do they require shared responsibility? When held to these standards, the ED proposals succeed in some areas and fall short in others.

Our section-by-section comments are attached. I write here on behalf of the more than 1.4 million members of the AFT to stress several overarching points.

The ARRA was proposed by President Obama and passed by Congress to help stabilize a teetering economy, prevent draconian cuts that would undermine education and other public services for years to come, and to save and create jobs. The approximately \$100 billion targeted for education has helped avert many cutbacks and layoffs, and has preserved proven programs. Although many districts, teachers and school staff are still struggling, this infusion of money has helped move us away from the precipice. It has been a lifeline for public education, and on behalf of the children we serve and the educators we represent, we were and are fully supportive.

The Race to the Top (RTTT) program was tucked into the ARRA; its presence seemed to suggest a broader goal—not simply to stem the tide of economic disaster, but to foster innovation in programs and policies that improve teaching and learning, including promoting better labor-management relationships. The AFT supports this goal.

If properly interpreted, implemented and administered, the Race to the Top program has the potential to spark innovation, replicate proven programs and promote promising ideas—all of which can help bolster school improvement.

However, after examining the draft notice, it appears that it is ED's intent that this program go much further, effectively creating and implementing education policy outside of a legislative process that would afford a broader, deeper and more open give-and-take among all stakeholders, including elected officials and the people they represent. We believe that bypassing the legislative process is inappropriate and not in keeping with the goals of ARRA.

For example, the Elementary and Secondary Education Act (ESEA) set out a detailed map of how states are expected to identify academically struggling schools, and then provided even more details on consequences and timelines for these schools. While the AFT, among many others, believes that adequate yearly progress (AYP), and its consequences and timeline need serious revision, the priorities proposed in the draft would simply layer another top-down accountability system on top of the current faulty one. Something as important and complex as identifying and assisting struggling schools deserves a thorough legislative review and real feedback from stakeholders and experts.

Given today's economy—in which all sectors, including school districts, are struggling—the promise of additional funding is a heady incentive to sign on the dotted line. It would be a mistake, however, to assume that funding alone will create good policy. No Child Left Behind has not failed solely because of underfunding. And the new RTTT program, which has so much potential, will not succeed simply because of an abundance of funding. What was true with NCLB is true of RTTT: True reform requires more than funding alone; it requires valid, reliable, sustainable and fair policies, thoughtful implementation and the collaborative approach necessary for success.

So that it is clear not only to ED but to all—particularly those who so often seek to define for us our views on key issues—the AFT believes:

1. There is obviously a role for student achievement in teacher evaluations. However, standardized assessments should not be the sole or predominant measure in a teacher's evaluation. We believe this for a number of reasons, not the least of which is that all research that examines equating student test scores on standardized assessments with teacher effectiveness acknowledges the imprecisions and limitations of doing so. And yet, under ED's proposal, a teacher's livelihood and career could depend upon an overreliance on an unproven idea.

It is imperative that we find common ground on teacher quality and compensation, namely, how to continuously develop, fairly compensate and accurately evaluate teachers on an ongoing basis. Unfortunately, the proposed guidelines conflate these three crucial tasks. Clearly, they are inextricably linked and must be aligned. However, failing to recognize their unique characteristics is a critical mistake.

For example, the potential consequences of a teacher's evaluation vary greatly when that information is used as a basis for determining if a teacher needs targeted professional development versus whether that teacher should be

granted tenure or dismissed. Consequently, the standard of reliability and validity imposed on high-stakes compensation, tenure and termination decisions must be different and, arguably higher, than what would be necessary when designing targeted professional development programs.

Relevant research into these areas recognizes the differences and suggests that while there continues to be optimism for the use of data to help improve programs such as student instruction and professional development, there is hesitancy about using the same results for decisions affecting compensation and tenure. A large body of scholarly research demonstrates the inability of so-called value-added methodologies to attribute student gains on tests to particular teachers with anything approaching the type of accuracy we all should want when making decisions about tenure and compensation. The AFT is, as we have repeatedly said, open to the discussion about how best to reform tenure and how teachers are compensated, but that discussion will have no legitimacy if it does not also include a discussion of the current statistical, practical and contextual limitations involved in tying student test scores to individual educators.

Before ED requires that unproven measures be imposed upon teacher evaluation systems that universally have been derided—by teachers, researchers, administrators and the Department of Education—we suggest that ED propose the following: that states first use RTTT funding to actually develop solid, reliable and transparent teacher evaluation systems. RTTT is an unprecedented opportunity to do what the recent New Teacher Project’s “The Widget Effect” recommends and which the AFT supports: develop meaningful and effective teacher evaluation systems that can measure teacher performance and inform decisions in a way that helps students and is fair to teachers.

2. The AFT strongly believes that charter schools, when appropriately authorized, operated with the input of teachers and held accountable to the same standards as non-charter public schools can be laboratories for innovation in both education and labor-management relations. For us, this is not just rhetoric: The AFT represents teachers and other staff in 70 charter schools, and the United Federation of Teachers, our largest local affiliate, founded and runs two charter schools. But, charter schools must meet the same standards for academic performance and must ensure access for all students, regardless of their test scores, language proficiency or disabilities. Because these are public, taxpayer-funded schools, parents and citizens have the right to demand appropriate oversight and regulation. When these fundamental qualities are met, charter schools become “public” in the broadest sense.

However, even when charters meet all these standards, we do not think, as the draft proposals would suggest, that they are a panacea for struggling schools. This point has been confirmed by a recent Stanford CREDO study, which indicates that while some charters are doing a good job, in most cases they perform no better and are frequently worse than non-charter public schools. The first step in helping struggling schools always should include the use of research-based proven programs. Therefore, it concerns us that the draft

proposal states that replacing staff, converting to a charter school or a privately managed school, and closing the school are the first, second and third options for districts to turn around struggling schools. ED's proposal limits use of the fourth option, which relies on comprehensive research-based programs, such as strategies for recruitment, retention and professional development, differentiating instruction, extending learning and enrichment time, and engaging families and communities. We recommend that this option also include co-locating other services children need.

Limiting use of the fourth option strikes us as counterintuitive. Innovation should encourage, not discourage, the use of research-based, proven programs. Rather than giving short shrift to it, ED should encourage and provide support for the type of collaboration and shared responsibility among teachers, school districts and other partners to turn around struggling schools that is permissible under the fourth option. We can point to districts in which the use of these strategies has proven successful. More districts should be using them, but that will not happen if the option is effectively taken away.

ED should not limit the fourth option but instead should encourage its use, require that it be comprehensive in nature, and look for opportunities for districts to create and implement model turnaround strategies that strengthen the schools while increasing capacity.

3. The AFT strongly supports the development and implementation of common standards and the inclusion in the proposal of a clear message that states should be working toward this goal.

But the standards and assessments should not stand alone; rather, they should be bookends of a program that includes a content-rich, sequenced curriculum with aligned assessments, standards-based guides and model lesson plans for teachers; professional development including mentoring and induction; time for planning and collaboration during the normal school day; and accountability measures that track whether teachers are provided what they need to make standards work in their classrooms.

ED should not simply ask that states adopt common standards; it should ask states to provide a roll-out plan for those standards and all of their supporting components so that they become not a hollow statement of intent but a catalyst for real change and improvement.

4. The AFT is pleased to see that the proposed selection criteria include the demonstration of a commitment of support from the state teachers union and a memorandum of understanding signed by the local teachers union. The reason is obvious: No reform model works over the long term without the support of the educators who must implement it.

Before I close, I want to comment on the four assurances that frame these draft regulations and which some suggest will frame the reauthorization of ESEA. As constructed, the assurances will not achieve the stated goal of helping to improve

teaching and learning, for the simple fact that they don't provide the latitude, content, context or 360-degree accountability that are needed to address the complex problems facing schools. Rather, they offer only a baseline route to gather data, again without context or content. While the AFT fully supports the four assurances as spelled out in the ARRA as a **framework** for innovation, we believe that these proposed regulations overstep the letter and intent of the law. For example, the ARRA's assurance on improving struggling schools simply says that states must comply with the corrective action and restructuring provisions of ESEA. But, the RTTT proposal interprets this assurance as prohibiting caps on charter schools. Similarly, the assurance related to teachers requires states to comply with highly qualified and equitable distribution provisions of ESEA. But the RTTT proposal inserts requirements around alternative certification, teacher evaluation, and tenure under this assurance. It is unclear how these expansive interpretations of the assurances further the goals set forth in the ARRA.

There is no doubt that \$4.35 billion is an extraordinary amount of money, and its reach is magnified by the pressures of a struggling economy. But our states and districts and unions, which rightly will be asked for statements of support, should not be presented with a Hobson's choice. They should, instead, be able to enthusiastically embrace a chance at real innovation, real collaboration and a real commitment to building programs that are branches on a growing, vibrant tree and not, as the adage warns, branches without a tree.

The AFT hopes that ED will carefully consider these comments. We look forward to working with the department to implement a program that helps create sustainable change that improves teaching and learning in our schools.

Sincerely,



Randi Weingarten
President

RW : mb opeiu#2 afl-cio
Enclosures (1)

Eligibility Requirements

Regarding proposed eligibility requirement (a): The AFT is seeking clarity as to how the Race to the Top (RTTT) awards timeline intersects with the State Fiscal Stabilization Fund application approval timeline. The U.S. Department of Education (ED) has publicly stated that Phase 1 Race to the Top applications will be available beginning in fall 2009, yet this notice sets Dec. 31, 2009, as the deadline for Stabilization Fund application approval to be eligible for Phase 1 Race to the Top awards. It appears that states may not have enough time to complete their Stabilization Fund applications in order to be cleared to receive Race to the Top funds.

Regarding proposed eligibility requirement (b), student achievement and student growth data are powerful tools that, when valid, reliable and used appropriately, can help inform instruction and improve outcomes. However, we have serious concerns about the validity and reliability of linking student achievement data with individual teachers, especially for the high-stakes uses outlined within this notice of proposed priorities.

It is imperative that any state data systems that include unique educator and student identifiers, and which link these identifiers via achievement data, protect the privacy rights of students as well as educators.

Finally, summative decisions about teacher evaluation, teacher pay and teacher benefits must be made at the local school district level, and where applicable, either in good faith cooperation with the school district's teachers, or with the exclusive bargaining representative of the school district's teachers.

Selection Criteria

Regarding the selection criteria, this notice states that sometime in the future, ED will "announce the maximum number of points assigned to each criterion." These points are essentially the basis upon which ED will determine which states receive RTTT funds and which do not. As with the actual 19 criteria, the allocation of points to each of these criteria should be available for public comment.

Standards and Assessments

(A)(1) Developing and adopting common standards

Regarding selection criterion (A)(1): For many years, the AFT has supported development of rigorous common state standards, thus we agree with the intent of (A)(1). However, if standards are to be helpful in improving teaching and learning, they cannot be adopted in a vacuum. States should be asked to demonstrate the ways that they intend to build upon the standards. For a standards-based system to achieve its goals, which include helping inform instruction, it must consist of:

- Standards that are detailed and explicit and build on knowledge and skills previously acquired as students move through the education system. They must be rooted firmly in subject matter content and specific enough to lead to a knowledge-rich curriculum that can be mastered during the school year. These standards must pay attention to both content and skills, and must be grade by grade for K-8 and by course at the high school level.
- Curriculum that provides teachers with a detailed road map for helping students reach the standards. The curriculum must focus on the content and concepts to be mastered grade by grade, and include instructional resources, instructional strategies, performance indicators, and unit and lesson plans.
- Assessments that provide information on how well the system and/or students are doing and indicate where changes in instructional strategies and resources are necessary if we are to improve learning for all children. They must be aligned to the standards and curriculum, valid, reliable and used for the purposes for which they were designed.
- Accountability must hold all parties responsible for providing the supports for student achievement. This includes assisting students who are having difficulty meeting the standards, providing professional development for teachers, and implementing standards for strong teaching and learning environments, as well as school policies that encourage students to take learning seriously by providing rewards and consequences based, in part, on state assessment results.
- Professional development that is aligned to all other components of the system and helps teachers and other instructional staff deliver the content, differentiate instruction and adjust delivery based on data analysis and best practices, as well as multiple sources of information about student learning.
- Time for collaboration and data analysis. The system must provide common planning time as well as individual planning time for teachers and instructional staff. This time is essential for educators to share and model lessons; review student achievement data; and discuss how to adapt instruction, planning and assessments to meet the needs of their students.

Both the development and implementation of such a system must be informed by teachers' collective experience and must be supported by teaching and learning conditions that foster student achievement.

The selection criteria give priority to states that adopt the standards developed by a multistate consortium. But standards alone are not enough. Adoption of the standards should be accompanied by the necessary supporting components outlined above.

Finally, the June 2010 date for adopting the new standards is unrealistic, because it does not appear to build in sufficient time to properly develop grade-by-grade common state standards and the necessary supporting components described here.

(A)(3) Supporting transition to enhanced standards and high-quality assessments

Selection criterion (A)(3) lists what state or LEA activities might include in order to successfully implement the use of enhanced standards. Given that teachers will be primarily responsible for ensuring successful implementation of the standards, the AFT recommends prioritizing the following specific strategies that translate the standards into classroom practice:

- A content-rich, sequenced curriculum and aligned assessments;
- Standards-based guides for teachers that provide essential background knowledge;
- Lesson plans that can be used with heterogeneous groups of students that new teachers could teach from and that more experienced teachers could draw from as they see fit;
- Pre-service teacher education and in-service professional development that prepare teachers to teach the specific content for which they are responsible; and
- Textbooks that, because they are based on clear standards of a reasonable length, are slim and focused.

Data Systems To Support Instruction

(B)(1) Fully implementing a statewide longitudinal data system

(B)(2) Accessing and using state data

(B)(3) Using data to improve instruction

Regarding selection criteria (B)(1), (B)(2) and (B)(3): The AFT strongly supports the use of comprehensive and reliable data to improve schools and inform instruction. To ensure that this happens, the criteria should specify that these data come from multiple sources, including a variety of measures of student learning and contextual factors, not from one state test. Additionally, the statewide longitudinal data systems should include data from these multiple sources, and include things like model lesson plans for teachers and other tools that will assist teachers in the delivery of instruction.

Likewise, to improve instruction, the criteria should include professional development for teachers and principals on how to interpret and use student achievement data gathered from multiple indicators. Further, to truly promote the use of data to improve instruction, data must be accessible to educators in a timely fashion, and educators must be provided time to analyze and interpret the data and develop future lesson plans based on these data, preferably in collaboration with other teachers.

We also recommend that criterion (B)(2) stipulate that state use and management of education data must protect the privacy of students and educators and must protect their personal information from accidental or intentional release to unauthorized persons and from intentional or accidental use for unauthorized purposes.

Great Teachers and Leaders

(C)(1) Providing alternative pathways for aspiring teachers and principals

Regarding selection criterion (C)(1): Research demonstrates that high teacher turnover is costly¹ and detrimental to student achievement.² And, not all alternative route programs are equal with regard to quality and to their ability to adequately prepare and retain teachers.³

The AFT believes that alternative certification programs should prepare teachers who want to be successful in the classroom. While the pathways to teaching may look different, the results should be the same: effective teachers who remain in the profession for more than a couple of years and who can facilitate student academic success.⁴

One evidence-based example of an alternative pathway for teachers to enter the profession is the Boston Teacher Residency program. This program recruits, prepares and retains teachers and principals for the high-need subjects in the Boston Public Schools; it does so by ensuring that for a resident's first full academic year he or she co-teach with a mentor in a school four days per week and, for one day per week, take coursework toward a master's degree. Residents are provided ongoing support for the first three years in the classroom.⁵

The "minimum proposed evidence" for criterion (C)(1) also should include data about the number of teachers rated "effective" and "highly effective," the distribution of these teachers as well as retention rates for alternatively (and traditionally) certified candidates from different programs.

The definition of alternative certification routes should ensure that all teacher preparation programs, traditional or alternative, emphasize both subject matter knowledge and pedagogy; provide a meaningful and realistic clinical experience, including mentoring and induction programs; and institute rigorous and relevant exit and licensure exams. The definition of alternative certification routes should include all components necessary to ensure that these programs provide prospective teachers with the knowledge, skills and resources needed to take on the challenges they will face in their classrooms.

(C)(2) Differentiating teacher and principal effectiveness based on performance

Regarding selection criterion (C)(2): The AFT believes there is a place for student learning in a teacher's evaluation. However, standardized assessments should not be the single or predominant factor in teacher evaluation systems.⁶ Evaluating individual teachers using their students' standardized test scores is of serious concern because current testing instruments are limited in their ability to capture the full range of learning, and because of the instability of value-added measures.⁷

Research shows us that even the best value-added models provide measures of student learning that vary enormously from year to year, especially for individual

teachers (versus whole school), and even more so for teachers in small classes and in small schools.⁸ Reasons for the annual instability include:

- Substantial statistical “noise” in both the pre- and post-test years from small sample size, test measurement error,⁹ sampling error,¹⁰ and changes in the classroom and school environment outside the teacher’s control.¹¹
- Testing periods that include two teachers—but with the results attributed to only one teacher—when tests are administered before the end of the school year.
- Non-random assignment of students to teachers.
- Bad data systems and mismatching of students to teachers.
- Failure of some value-added models to include student background and the fade out of prior teacher effects.

Although test scores may play a role, student achievement should include evidence of growth in knowledge and skills based on multiple measures. Just as no single measure can evaluate teacher performance, no single measure can or should account for student learning. Some examples of the multiple sources that can provide evidence of student learning:

- Student performances, group work or presentations scored using a rubric;
- Writing samples;
- Student progress toward targeted learning objectives;
- Portfolios;
- Grades;
- IEP goals and objectives;
- language proficiency goals for English language learners;
- Student “capstone” projects (e.g., graduation, end-of-course research or thesis paper).

Noticeably missing from the notice of proposed priorities is evidence of learning by examining student work. This would be a more meaningful way to assess student growth, but would require states and districts to invest resources in the development of a standardized approach to analyzing student work.

Criterion (C)(2) (and its attendant definitions) does not suggest workable options for evaluating the vast majority of teachers. Based on the definitions of student achievement and student growth, it is clear that the current draft focuses on, at best, 30 percent of the current teaching force (teachers for whom value-added data are available).

For non-tested grades and subjects, the examples included in this draft either relate back to “tested grades and subjects” (because for the most part, interim assessments are not available for subjects for which there are not summative assessments, i.e., those already included in “tested grades and subjects” or those with end-of-course exams) or cannot validly be tied to an individual teacher’s instruction. Criterion (C)(2) would require that for at least 70 percent¹² of teachers, decisions relating to evaluation, compensation and tenure for **individual** teachers be based on factors that in no way can be tied solely to an **individual** teacher: “rates at which students are on track to graduate from high school, percentage of students

enrolled in Advanced Placement courses who take Advanced Placement exams, rates at which students meet goals in individualized education programs.”

Given the high stakes attached to teacher evaluation in this proposal, a more thoughtful and research-based approach to including evidence of student achievement from non-tested grades and subjects is needed. As written, criterion (C)(2) encourages the development of additional standardized, high-stakes tests for the current non-tested grades and subjects, and the use of these tests for purposes other than their intended purpose. It is clear that educators and parents agree that we do not need more tests.¹³ We suggest a prohibition on the development of more standardized, high-stakes tests, especially for kindergarten through second grade, and a requirement that tests be used only for their intended purposes. Instead, the multiple sources listed above should be considered and expanded.

Teacher ratings should take into account indicators other than data on student achievement. These should be developed based on standards of practice that define what effective teachers do to facilitate student achievement. For example, there is consensus among educators that the planning and preparation of lessons, activities and classroom-based assessments are important behaviors that effective teachers engage in regularly. A comprehensive teacher evaluation system focused on increasing teacher quality and improving student achievement would ensure that the domain of planning and preparation was accounted for, measured and given weight in the system.

We agree that an improved teacher evaluation system must include the necessary teaching and learning supports as well as ongoing, job-embedded professional development, including induction and mentoring, for all teachers. Teacher evaluation should truly be about improving the quality of the workforce and student achievement. This can be done by identifying, developing and supporting effective teachers. But effective teachers do not exist within a vacuum; they exist within a school culture and environment that nurture and develop their talent. Accountability no longer can be focused on teachers alone. All members of the school community—administrators, teachers, parents and school staff—must be held accountable for creating a school environment that promotes student and teacher success. Teachers need resources including the time to collaborate with their colleagues around student achievement data if their students are to be successful. Neither students nor teachers will thrive in an environment that is not conducive to teaching and learning.¹⁴ Therefore, in order to foster shared responsibility, measures for assessing a school’s teaching and learning conditions should be developed and included in a teacher evaluation system.

Regarding (C)(2)(ii): Research on alternative compensation approaches for teachers shows that the programs that have achieved some measure of success do not pay teachers solely for increased student achievement, but also for their knowledge, skills and position on a career ladder.¹⁵ Any alternative compensation system should take these factors into account. Further, because of all of the technical reasons already outlined, and because collaboration breeds success, alternative compensation systems that provide schoolwide bonuses should be made a higher priority than those that make individual determinations.

Criterion (C)(2) has laid out four principles that should be included when designing a teacher evaluation system to determine teacher effectiveness; however, several additional factors should be included. For example, language about validity and reliability of evaluation instruments, mandatory training of evaluators, training of teachers in a new evaluation system, including teachers as evaluators, ensuring teacher buy-in, and linking principal and teacher evaluations would increase the likelihood of a rigorous and comprehensive evaluation system aimed at improving teacher quality and student achievement.¹⁶

Timeliness of the data is another problem. Standardized test score data are rarely available to schools, educators, etc., during the school year in which the instruction took place. (For example, student test scores from the 2008-09 school year were, for the most part, preliminarily released during summer 2009. Most states then allow for several weeks of review and appeals before these data are considered final.) It is therefore unclear how the decisions spelled out in criterion (C)(2)(d) can be made in a timely way. The AFT recommends that criterion (C)(2) include plans for how to compensate for the lack of available data.

The AFT recommends that criterion (C)(2)(iii) read: “Granting tenure to and dismissing teachers and principals based on rigorous and transparent procedures for awarding tenure (where applicable) and for removing tenured and untenured teachers and principals after they have had ample opportunities to improve **in full accordance with due process.**”

Finally, the standard of proof (e.g., regarding accuracy, validity, reliability) when using student achievement data to evaluate teachers will differ depending on the decisions being made. For example, the potential consequences of a teacher’s evaluation vary greatly when that information is used as a basis for determining if a teacher needs targeted professional development versus whether that teacher should be granted tenure. Consequently, the standard of reliability and validity imposed on high-stakes compensation and tenure decisions must be different and, arguably higher, than what would be necessary when designing targeted professional development programs.

(C)(3) Ensuring equitable distribution of effective teachers and principals

Regarding selection criterion (C)(3): The goal of an equitable distribution of effective teachers needs to be broadened to ensure that all students are taught by high-quality teachers. Criterion (C)(3) refers to plans for the “implementation of incentives and strategies ...” Forced transfers of teachers are not a solution. Such efforts have failed in the past. Instead, these plans should include investments in teacher support and development. This includes teacher preparation, induction and mentoring programs, and innovative teacher-led peer assistance and review. The plans also should include addressing the conditions that make attracting and retaining teachers in some schools persistently difficult. Criterion (C)(3) and the notice’s attendant definitions of “effective” and “highly effective” assume that teachers could be “effective” in any school setting without regard to the school’s

conditions. This simply turns a blind eye to what students and teachers need to be successful.

The conditions that affect teaching and learning include:

- Lack of on-site support and intervention for students experiencing learning difficulties;
- Poor administrative leadership and support;
- Unhealthy physical plant;
- Lack of faculty influence on decisions that affect student learning;
- Inadequate, ongoing, job-embedded professional development and other supports;
- Lack of student academic success;
- Student discipline and personal safety concerns;
- Inadequate time for planning, preparation and instruction; and
- Excessive classroom intrusions.

Additionally, plans should include provisions for necessary wraparound services to low-income students so that students can focus on learning and teachers can focus on teaching rather than on a student's medical, dental and nutritional needs. The solution is to make hard-to-staff schools desirable places in which students can learn and teachers can teach. School systems need to identify the strategic mix of programs, professional supports and incentives to ensure that each school offers a positive environment for students and teachers. The responsibility for securing the resources and supports to achieve this strategic mix must be shared by all stakeholders, including institutions of higher education and teacher preparation programs.

(C)(4) Reporting the effectiveness of teacher and principal preparation programs

Regarding selection criterion (C)(4): The AFT supports the collection and reporting of these data so that we can gather best practices about what successful teacher and principal preparation programs do to ensure that teachers and principals are effective once they enter the workforce. However, our serious concerns with using student test scores to determine teacher and principal quality as outlined throughout this document also apply to using these test scores to determine the quality of preparation programs.

Creating an arbitrary threshold of 20 or more graduates would have the effect of only requiring data for larger teacher and principal preparatory programs. This would likely exclude many alternative preparation programs that are localized recruitment models. Effectively, this would mean that schools of education with large programs or programs like New Leaders for New Schools would be held to a far higher standard than local teacher and principal corps programs that prepare fewer than 20 individuals a year. If we care enough to collect data on teacher and principal candidates and to link that data to their preparatory programs for a number of purposes, including accountability purposes, then we should work hard to apply that same level of care across the board. A minimum of 20 is too high in this case. We recommend that all teacher and principal preparatory programs be held accountable.

Finally, we recommend that the teacher of record—no matter his or her certification status—be included in this data collection effort around teacher preparation. Past data collection efforts of this nature have excluded individuals who are in alternative certification programs or who are otherwise not fully certified. This aligns with the goal of ensuring that all preparation programs be held accountable.

(C)(5) Providing effective support to teachers and principals

Regarding selection criterion (C)(5): The AFT agrees that data should be available to inform lesson planning and instruction, but we are concerned that setting an arbitrary time limit of 72 hours will lead to the development of low-quality assessments that can be graded quickly and returned to teachers. Not all data need to be available to teachers within 72 hours, especially data that will inform professional development opportunities. Rather, the focus should be on both the timeliness of available data as well as the quality of the data. And quality data depend on quality assessments. More important, a systematic way to include teachers' input into the effectiveness of the evaluation system must be incorporated. This way, when problems arise, such as the timeliness or availability of data to inform instruction, teachers have a means to adjust the system to make it more effective.

Additionally, data must be collected on standards of effective teaching practice. Value-added data in particular can tell you nothing about why a teacher is effective or ineffective. If the ultimate goal of an evaluation system is to improve teaching practice to foster student achievement and growth, then we must collect data on teaching practices. We need to include a variety of evaluation techniques to capture the breadth of effective teaching and professional practice incorporating classroom observations, review of lesson plans, self-assessments, teaching artifacts and portfolio assessments.

Finally, we agree that data can be a useful tool in terms of improving professional development, including mentoring and induction programs.

Turning Around Struggling Schools

(D)(1) Intervening in the lowest-performing schools and LEAs

Regarding selection criterion (D)(1): The AFT objects to this priority. States have poor track records in intervening in the management of low-performing schools, especially in high-need urban areas. Further, research suggests that districts often remain under state control for decades without significant improvement in student achievement. For the most part, state education agencies simply do not have the capacity or an understanding of local contexts to effect positive changes that will improve schools.¹⁷ State education agencies lack the funding to sustain a program and a staff with the necessary expertise to intervene in low-performing schools.¹⁸ This was true of the less populous states before the current economic recession; now it is true of nearly all the states.

Research also indicates that it is very difficult to “scale up” a strategy that has worked in one district or school.¹⁹ To be successful, school improvement must address the particular needs of the school and the community. Simply because of proximity, school districts have more capacity to tailor interventions to the particular needs of schools and their communities than states do.

(D)(2) Increasing the supply of high-quality charter schools

Regarding selection criterion (D)(2)(i): The AFT has long supported the role that charter schools play as laboratories for innovation. However, policy edicts to lift charter caps should be weighed against the research on charter school quality. The recent Stanford University CREDO report²⁰ provided a significant national snapshot of how charter schools are faring, revealing that while in a few cases charter schools do a good job, in most cases they perform no better and are frequently worse than traditional public schools. The CREDO report reinforces the AFT's position that charter schools are not a panacea, and that the focus of federal and state authorities should be on how to support and improve our non-charter public schools, which the majority of America's students attend.

The report shows that of the states doing the worst—Texas, Florida, Ohio, Arizona, Minnesota and New Mexico—the first four have 300 or more charter schools and rank only below California in the total number of charters. The states with charter schools outperforming non-charter public schools have fewer charter schools: Arkansas, Illinois (Chicago) and Missouri have fewer than 100 charters each, and Colorado has 140.

In addition, in a number of states, virtual charter schools operate without adequate oversight and accountability. For example, in Oregon, school districts can “sponsor” virtual charter schools. These charters can enroll out-of-district students, with the money following the student to the district in which the student is “enrolled.” There is an incentive for small, cash-strapped districts to sponsor virtual charter schools because the sponsoring district receives a cut of the money that flows to the virtual charter. These for-profit virtual charters have, in essence, created a back-door voucher program that often operates outside the system of state standards and curriculum; high dropout rates and many fiscal irregularities have been found in these virtual charters.

These patterns strongly suggest that students are not well-served by state or federal policies that encourage charter proliferation without having a rigorous entry process, adequate oversight or speedy closure policies.

Charter caps serve as a control on quality. In states without effective caps, quantity becomes more important than quality. Caps and the need to periodically “lift” or re-examine them are important incentives for charter authorizers to pay attention to the issues of quality.

Regarding (D)(2)(ii): We appreciate the fact that the selection criteria include considering the extent to which the states holds charter schools accountable, but

we believe this item would be improved by adding more specific language to consider the extent to which the states have statutes and guidelines that define a rigorous approval process that charter authorizers must follow, including an examination of the educational and financial track record of proposed school operators. We also recommend that when reporting on the extent to which the state has closed or not renewed charter schools, the state should report on the number of schools that have closed for each of the following reasons: for academic reasons, financial reasons, low enrollment or mismanagement.

Regarding (D)(2)(iv): The AFT recommends adding to this criterion the extent to which the public retains an ownership interest in facilities that have been acquired or improved with public funds. If public funds are used for these purposes, the public interest should be protected.

We also recommend adding a new criterion (D)(2)(v) to consider the extent to which the state collects data on the student populations served by its charter schools, including students with disabilities, English language learners and low-income students as well as the extent to which charter schools serve student populations comparable to those in non-charter public schools in the districts in which they are located. Criterion (D)(2)(v) also should gauge whether the charter schools provide their students physical education, recess and lab science courses at the same level as non-charter public schools.

(D)(3) Turning around struggling schools

Regarding selection criterion (D)(3): The AFT is committed to improving struggling schools. But, reforms must be evidence-based. For example, Baltimore city schools have made solid improvements in both reading and math for two consecutive years. They have done this by investing in lower class sizes, more school and classroom resources, and relevant professional development for teachers and other instructional staff consistent with the needs of each school and the needs of its students. Secretary Arne Duncan just visited Abbottston Elementary School within the Baltimore district to celebrate its success in implementing evidence-based reforms. Yet, it appears as though the reforms implemented by Abbottston—reforms that led directly to increased academic achievement—would not count as one of the four reform options listed in criterion (D)(3) .

Overall, we believe that the four options are too narrow, rigid and preclude the very reforms that have proven to be effective. AFT affiliates have collaborated extensively with local school districts and other partners. Based on our experience and evidence, here are common characteristics that run through many—if not most—of these successful models. We recommend that these components be included in the fourth option, and that the fourth option be a research-based alternative available to all schools.

Focus on Students' Needs

- Standards-based, common curriculum
- Smaller class sizes
- Individual and small-group tutoring before or after school

- Extended day/year
- Schoolwide behavioral expectations in a safe and orderly environment
- Same student populations in redesigned schools
- Block scheduling (secondary schools)

Focus on Teachers' Concerns

- Ongoing, embedded professional development
- Data-driven instruction and ongoing supports for both novice and experienced teachers
- Teacher voice in instructional and school decision-making
- Common planning time for staff, including the flexible schedules to allow for this
- Additional compensation for extended time (pro-rata)
- Early involvement and buy-in with turnaround planning
- Peer involvement in staffing selections

Parent and Community Roles

- Early buy-in on part of parents and community
- Ongoing parental involvement in all aspects of school life
- Appropriate wraparound services
- Increased opportunity for parent-teacher interaction
- Parent education programs regarding instructional programs and supports that reflect the diversity of families, including language, family composition and cultural differences

Additional Factors

- Leadership is educationally oriented and supportive
- Planning is collaborative and often negotiated as part of contracts
- District and union leadership play very supportive role in schools
- Schoolwide pay incentive plans in some models
- School governance is both bottom-up and top-down
- Additional trained staff (paraprofessionals, specialists, etc.) as appropriate to support classroom needs

Comments specific to the four bulleted options within Criterion (D)(3) follow.

There are many items within the four options that fall under collective bargaining provisions, including staffing, time, evaluation and compensation. This notice should require that all of these items be negotiated locally by the collective bargaining representative; and nothing in these requirements should be construed to alter or otherwise affect the rights, remedies and procedures afforded school or school district employees under federal, state or local laws (including applicable regulations or court orders) or under the terms of collective bargaining agreements, memoranda of understanding, or other agreements between such employees and their employers. (This is consistent with Section 1116(d) of ESEA.)

Also, the goal in all of these options should be to serve the entire former student population. The school should not be shuttered for a year and then reopened after a full school year has passed; this displaces children and disrupts communities.

Any planning year should happen while the students attend their current school, which may be reopened the following year under a different name, under different management and/or with turnaround elements in place.

Option 1:

We recommend that this be revised to allow all current staff to reapply for their jobs, and for the school to retain at least 50 percent of current qualified staff who meet all of the requirements of the redesigned school.

Option 2:

All schools, including schools reopened as charter schools and those managed by education management organizations (EMOs) should serve the entire former student population. As with the traditional public schools they are replacing, EMO and charter schools should be prohibited from refusing any students based on test scores, special needs status or any other factor.

In light of the recent CREDO report indicating that charter schools often have lower rates of student growth and achievement than non-charter public schools, we recommend adding that the selected EMO or charter schools must have a demonstrated track record of success.

We are opposed to any school reform initiative that would displace students and disrupt communities, as Option 3 would.

Option 4 should be available to all schools, not confined to schools for which the other strategies “are not possible.”

The “Proposed Performance Measures” section in the Appendix states that only three of the options would be allowed. This should state that all four options would be allowed.

Overall Selection Criteria

(E)(3) Enlisting statewide support and commitment

The AFT commends the inclusion in criterion (E)(3) of the demonstration of a commitment of support from the state teachers union, and if applicable, a memorandum of understanding signed by the local teachers union. When the union is an equal partner in the development and implementation of reform, it not only increases leadership and builds professionalism, it all but guarantees success. When unions are not partners and programs are imposed on teachers, the effect is opposite. Unions need to be offered a real opportunity to help shape programmatic reforms—not given a token sign-off.

Definitions

Overall, the AFT believes that the definitions of **student achievement**, **effective teacher** and **highly effective teacher** all rely too heavily on student test scores. (See

comments under criterion (C)(2) regarding appropriate evaluation systems for teachers.)

Further, we are concerned that the suggested definitions equate “effective” with advancing students “one grade level in an academic year” and “highly effective” as advancing students “more than one grade level in an academic year.” This simplistic approach ignores the fact that research has not identified a standard for how much should be learned in a given school year in a given subject.

Regarding the definition of **effective principal**, in addition to our strong objection to the notion that the effectiveness of a school be determined predominantly by test scores, we do believe that principals have a role in creating and shaping the conditions conducive to student success. Therefore, we recommend that at minimum, the definition of “effective principals” should include data about both staff turnover rates and working conditions.

Regarding the definition of **high-need LEA**, the AFT is concerned with the proposed definition of a “high-need LEA” as it is inconsistent with the actual definition used in Section 14013 (2) of P.L. 111-5, the American Recovery and Reinvestment Act of 2009. Under this overly broad definition, districts with a minimum of just one high poverty school would be considered a “high-need LEA.” We believe that setting the poverty threshold at such a level undermines the Education Department's intent to prioritize funding for truly “high-need LEAs.”

Consistent with the enabling legislation, “high-need LEA” should be defined as a district “... (A) that serves not fewer than 10,000 children from families with incomes below the poverty line; or (B) for which not less than 20 percent of the children served by the agency are from families with incomes below the poverty line.”

¹ According to Barnes, Crowe & Schaefer (2007), teacher turnover costs the nation more than \$7 billion a year. For more information, see Barnes, G., Crowe, B. & Schaefer B. (2007). *The cost of teacher turnover in five school districts: A pilot study*. Washington, DC: National Commission on Teaching and America's Future.

² Research uniformly suggests that new teachers struggle to perform their jobs (Donaldson, 2008). First-, second- and third-year teachers are consistently less effective than more experienced teachers (Hanushek, Kain, O'Brien & Rivkin, 2005; Rockoff, Kane & Staiger, 2006; and Rockoff, 2004). The positive impact of experience is the most common and strongest effect in the value-added student assessment literature (Harris, 2007). For more information on this topic, see Donaldson, M. L. (2008). *Teach for America teachers' careers: Whether, when, and why they leave low-income schools and the teaching profession*. Harvard Graduate School of Education Project on the Next Generation of Teachers. Paper prepared for the 2008 annual meeting of the American Educational Research Association: New York, NY; Hanushek, E., Kain, J. F., O'Brien, D. M. & Rivkin, S. G. (2005). *The market for*

teacher quality Working Paper 11252, Washington, DC: National Bureau of Economic Research; Rockoff, J. E., Kane, T. J. & Staiger, D. O. (2006). *What does certification tell us about teacher effectiveness? Evidence from New York City*. Working Paper 12155. Cambridge, MA: National Bureau of Economic Research; Rockoff, J. E. (2004). The impact of individual teachers on student achievement: Evidence from panel data. *American Economic Review*, 94; and Harris, D. N. (2007). *The policy uses and “policy validity” of value-added and other teacher quality measures*. University of Wisconsin at Madison. Paper prepared for the Educational Testing Service (ETS), San Francisco, Sept. 24-25, 2007.

³ Sykes and Dibner (2009) conducted a review of 50 years of federal teacher policy to help inform the conversation about teacher quality, including alternative certification. Several lessons can be gleaned from this review:

- Alternative routes have been successful in recruiting more candidates (especially candidates of color) but have failed to increase teacher retention.
- Alternative route programs are currently unregulated by the federal government (with the exception of ESEA’s HQT requirements), thus leading to an unevenness in program quality.
- Alternative route teachers are less likely to stay in teaching than are those from traditional programs.
- Scaling-up successful alternative route programs may be difficult.

Finally, Sykes and Dibner (2009) caution policymakers not to rely on alternative route programs as the solution to systemic problems. A study by Grossman & Loeb (2008) found significant variability within alternative certification programs; they recruit different candidates, include different features and place teachers in different settings. They conclude that all of these factors influence the effects of alternative route programs. Sykes & Dibner use the example of New York City to illustrate that a combination of state policy to close certification loopholes, alternative route programs and teacher salary raises all were part of the formula needed to increase teacher retention and improve student achievement.

For more information, see Sykes, G. & Dibner, K. (2009). *Fifty years of teacher policy: An appraisal*. Washington, DC: Center on Education Policy; and Grossman, P. & Loeb, S. (Eds.). (2008). *Alternative routes to teaching. Mapping the new landscape of teacher education*. Cambridge, MA: Harvard Education Press.

⁴ Humphrey, Wechsler & Hough (2008) studied seven alternative certification programs to identify characteristics of successful programs: the Teacher Education Institute (Elk Grove, California, Unified School District); New Jersey’s Provisional Teacher Program; Milwaukee’s Metropolitan Multicultural Teacher Education Program; the New York City Teaching Fellows Program; North Carolina’s NC TEACH; Teach for America; and the Texas Region XIII Education Service Center’s Educator Certification Program. The researchers found that school context (i.e., where alternatively certified teachers are placed) matters most. “Alternative certification participants working in schools with strong leadership, adequate supplies and materials, and a collegial work environment were more likely to plan to stay in teaching, had more confidence in their teaching skills, and had a stronger sense of professional growth than those working in challenging schools” (p. 35). Further, “when alternative certification programs cannot control placement, accommodating for challenging school environments is of utmost importance” (p. 36).

Effective alternative certification programs also “select well-educated individuals or take steps to strengthen candidates’ subject-matter knowledge” (p. 38). Coursework is also important, particularly coursework that is relevant to the needs of participants in alternative certification programs. However, “few alternative certification participants reported that their coursework contributed greatly to their ability to teach special education students and English-language learners” (p. 37). Finally, effective programs provided trained mentors to assist candidates with a variety of essential teaching practices such as lesson planning and delivery as well as share resources such as curricula. For more information, see Humphrey, D.C., Wechsler, M.E. & Hough, H.J. (2008). Characteristics of effective alternative teacher certification programs. *Teachers College Record*, 110 (4), 1-63.

⁵ The Boston Teacher Residency (BTR) program helps the Boston Public Schools (BPS) address high-needs hiring areas including teachers of color, math and science teachers, and teachers of students with disabilities and English language learners. According to BTR, more than 50 percent of residents and graduates are educators of color, and over 50 percent of residents and graduates teach math or science. More than 90 percent of the program's graduates are still teaching in BPS. Eighty-eight percent of principals rate BTR graduates as good as or better than other new teachers; 55 percent rate them “significantly more effective.” Sixty-four percent of principals rate BTR graduates as good as or better than the entire faculty (Freeman, H. R., Wiley, J. & Ostiguy, C. [2009, July]. *Boston Teacher Residency: Peer assistance and review for teacher leaders*. Workshop presented for the American Federation of Teachers QuEST conference, Washington, D.C.).

⁶ Relying solely on standardized test scores as the measure of student learning is problematic. According to the National Comprehensive Center for Teacher Quality, defining teacher effectiveness as a teacher’s ability to improve student gains on standardized achievement tests should be avoided because:

- Teachers are not exclusively responsible for students’ learning;
- Consensus should drive research, not measurement innovations;
- Test scores are limited in the information they can provide; and
- Learning is more than average achievement gains.

For more information, see Little, O., Goe, L. & Bell, C. (2009). *A practical guide to evaluating teacher effectiveness*. Washington, DC: National Comprehensive Center for Teacher Quality.

⁷ Value-added measures aren’t ready for primetime. Subjective evaluations (such as those done by principals and/or peers) and value-added measures that attempt to identify which teachers are effective can produce results that are very different. (For a discussion of this, see Rockoff, J. E., Jacob, B. A., Kane, T. J., & Staiger, D. O. (2008). *Can you recognize an effective teacher when you recruit one?* Working Paper #14485. Cambridge, MA: National Bureau of Economic Research.). Further, these measures don’t tell us anything about *why* teachers vary in effectiveness making it impossible to predict which teachers will be most effective (Goe, L., Bell, C. & Little, O., 2009).

⁸ See Aaronson, D., Barrow, L. & Sander, W. (2003). *Teachers and student achievement in Chicago public high schools*. Technical report, Federal Reserve Bank of Chicago; Ballou, D. (2005). Value-added assessment: Lessons from Tennessee. In R. Lissetz (Ed), *Value Added*

Models in Education: Theory and Applications. Maple Grove, MN: JAM Press; Bos, M. D. McCaffrey, T. Sass, H. Doran, D. Harris, J. Lockwood (2006). *An empirical investigation of the value-added effects of Florida*. Unpublished manuscript submitted to U.S. Department of Education, Institute for Education Sciences; and Goldhaber, D. & Hansen, M. (2008). *Assessing the potential of using value-added estimates of teacher job performance for making tenure decisions*. National Center for Analysis of Longitudinal Data in Education Research (CALDER).

⁹ Measurement error is related to such fluctuating factors as health, motivation, attention and fatigue around each student's hypothetical true score.

¹⁰ Sampling error is caused by random variations in student ability, early preparedness, family background and motivation from grade to grade in the same school.

¹¹ Changes in the environment not under the school's control are such things as changes in attendance areas, changes in housing patterns or natural disasters.

¹² For estimates on how many teachers could be evaluated with value-added measures, see Prince, C.D., Shuermann, P.J., Guthrie, J.W., Witham, P.J., Milanowski, A.T. & Thorn, C.A. (2008). *The other 69 percent: Fairly rewarding the performance of teachers of non-tested subjects and grades*. Washington, DC: Center for Educator Compensation Reform.

¹³ According to the 40th Phi Delta Kappan/Gallup poll, more than four of 10 public school parents believe there is too much emphasis on achievement tests, while only one in 10 agrees that there is not enough. For more information, see Phi Delta Kappan (2008). Highlights of the 40th PDK/Gallup poll. Retrieved Aug. 19, 2009, from http://www.pdkmembers.org/members_online/publications/e-GALLUP/kpoll_pdfs/pdkpoll40_2008.pdf.

¹⁴ Positive working conditions found to contribute to high-quality teaching, student learning and teacher retention include: adequate school facilities, a safe and orderly environment, administrative support, manageable class sizes, opportunities to collaborate regularly with peers and opportunities for teachers to be true decision-makers in their schools. For more information, see Bryk, A. S. & Schneider, B. (2002). *Trust in schools: A core resource for improvement*. New York: Russell Sage Foundation; Hanushek, E. & Rivkin, S. (2007, Spring). Pay, working conditions and teaching quality. In *The future of children*, 17(1), 69-86. Retrieved May 19, 2009, from http://www.futureofchildren.org/usr_doc/7_04.pdf; Hirsch, E. (2008). *Identifying professional contexts to support highly effective teachers*. Washington, DC: National Comprehensive Center for Teacher Quality; Hirsch, E. (2006). *Recruiting and retaining teachers in Alabama: Educators on what it will take to staff all classrooms with quality teachers*. Hillsborough, NC: Center for Teaching Quality; and Horng, E. L. (2009). Teacher tradeoffs: Disentangling teachers' preferences for working conditions and student demographics. *American Educational Research Journal*, 46(2).

¹⁵ See Schacter, J. & Thum, Y. M. (2004). Paying for high- and low-quality teaching. *Economics of Education Review*, 23, 411-430; Schacter, J., Thum, Y. M., Reifsnider, D. & Schiff, T. (2004). *The Teacher Advancement*

Program report: Year three results from Arizona and year one results from South Carolina TAP schools. Santa Monica, CA: Milken Family Foundation; and Solmon, L., White, T., Cohen, D. & Woo, D. (2007). *The effectiveness of the Teacher Advancement Program.* Santa Monica, CA: National Institute for Excellence in Teaching.

¹⁶ For example:

- Research points to the importance of systematically training classroom observers and evaluators. For a discussion of this, see Little, Goe & Bell (2009) and The Joint Committee on Standards for Educational Evaluation (1995). *The program evaluation standards* (2nd ed.). Thousand Oaks, CA: Sage.
- Communication prior to, during and after the evaluation is essential. For a discussion of this, see Joint Committee on Standards, 1995; Mathers, C., Oliva, M. & Laine, S. (2008). *Improving instruction through effective teacher evaluation: Options for states and districts.* Washington, DC: National Comprehensive Center for Teacher Quality; and Wise, A. E., Darling-Hammond, L., McLaughlin, M. W. & Bernstein, H. T. (1984). *Case studies for teacher evaluation: A study of effective practices.* Santa Monica, CA: RAND Corporation. Retrieved Jan. 26, 2009, from <http://www.rand.org/pubs/notes/2007/N2133.pdf>.
- Multiple observations are essential to a fair evaluation process as well as the improvement of practice. Current research suggests including at least four or five observations in an overall single evaluation. For a discussion of this, see Blunk, M. (2007, April). *The QMI: Results from validation and scale-building.* Paper presented at the annual meeting of the American Research Association, Chicago.
- Research shows that teachers are more receptive to and incorporate instruction and advice from those who have significant knowledge of curriculum, instruction and academic content such as expert teachers. For a discussion of this, see Stiggans, R. J. & Duke, D. L. (1988). *The case for commitment to teacher growth: Research on teacher evaluation.* Albany, NY: State University of New York Press; Wise, Darling-Hammond, McLaughlin & Bernstein, 1984.
- Research highlights the importance of including teachers in both the design and implementation of the evaluation system. For a discussion of this, see Joint Committee on Standards, 1995; and Kyriakides, L., Demetriou, D. & Charalambous, C. (2006). Generating criteria for evaluating teachers through teacher effectiveness research. *Educational Research*, 48(1), 1-20.

¹⁷ For a discussion about states' ability to improve low-performing schools, see Anderson, L.M. & Welsh, M.E. (2000). *Making progress: An update on state implementation of federal education laws enacted in 1994.* Washington, DC: U.S. Department of Education, Office of the Under Secretary; McDermott, K.A. (2006). Incentives, capacity and implementation: Evidence from Massachusetts education reform. *Journal of Public Administration Research and Theory*, 16(1), 45-65; O'Day, J. (1999). *One system or two? Title I accountability in the context of high stakes for schools in local districts and states.* Washington, DC: U.S. Department of Education; and Rennie Center for Education Research and Policy. (2005). *Reaching capacity: A blueprint for the state role in improving low-performing schools and districts.* Boston.

¹⁸ For a discussion of this, see Center on Education Policy (2007). *Educational architects: Do state education agencies have the tools necessary to implement NCLB?* Washington DC: Author.

¹⁹ For a discussion on the difficulty of scaling up improvement strategies, see Coburn, C. (2003). Rethinking scale: Moving beyond numbers to deep and lasting change. *Educational Researcher*, 32, 6, 3-12.

²⁰ Center for Research on Educational Outcomes (CREDO) (2009). *Multiple choice: Charter school performance in 16 states*. Stanford, CA: CREDO, Stanford University. Retrieved Aug. 27, 2009 from http://credo.stanford.edu/reports/MULTIPLE_CHOICE_CREDO.pdf.