Liston, D., Whitcomb, J., Borko, H. (2007). NCLB and scientifically-based research. *Journal of Teacher Education, 58*(2), 99-107. **Posted with the permission of the publisher. Copyright 2007 by SAGE Publications.**



# NCLB and Scientifically-Based Research: Opportunities Lost and Found

Dan Liston, Jennie Whitcomb, and Hilda Borko[1]

University of Colorado at Boulder

In fall 2006, Secretary of Education Margaret Spellings and President George W. Bush, while stumping for reauthorization of No Child Left Behind (NCLB), asserted, "It is working, and it is here to stay" (The White House, 2006). Though we find claims of its effectiveness premature, NCLB and the subsequent and less-noticed Education Sciences Reform Act (ESR) are without question changing the landscape of public education. Since their adoption in 2002, many of the laws' provisions have been implemented. For example, high-stakes testing and reporting systems to demonstrate "adequate yearly progress" (AYP) are in place in all states. Indeed, as this editorial goes to press, headlines about schools not demonstrating AYP abound. Policies establishing criteria for "highly qualified teachers" (HQT) have been adopted, although wide variation exists across the states, particularly for veteran teachers. "Scientifically-based research" has become the stamp of approval for curriculum selection and instructional practices in many of the NCLB programs, notably Title I and Reading First grants. And, the Education Sciences Reform Act supports the newly created Institute for Educational Sciences (IES), which is a major source of funding for educational research and is overseeing contracts to overhaul the ERIC Clearinghouse into the What Works Clearinghouse. Some flexibility will undoubtedly be introduced in the reauthorization of NCLB. Regardless of changes enacted, however, we expect that the specifics of NCLB, and the broader theme of accountability to close the "achievement gap," will continue to dominate education conversations in the years ahead.

[1] As an editorial team, we write editorials collaboratively. To reflect the nature of this joint work, we rotate order of authors with each journal issue. We acknowledge and thank Robert Reichardt and Ed Wiley for their most helpful comments on earlier drafts of this editorial.

Initial scholarly responses to NCLB offered both critical analyses and defenses of the assumptions and research base underlying the laws (e.g., Carlson & Levin, 2005; Darling-Hammond & Youngs, 2002; Elmore, 2002; Hess & Finn, 2004, Linn, 2003; Thernstrom & Thernstrom, 2003). At the law's five-year mark, analyses describe evidence of its uneven effectiveness. For example, while district test scores are improving, NAEP scores do not show any significant improvement (Bracey, 2006; Center on Education Policy, 2006). In states that have taken a tougher stance implementing AYP goals, some schools have failed to make AYP for the second or third year, thus triggering NCLB sanctions that culminate in restructuring, despite rating in the top categories on state accountability measures. These multiple accountability systems have left the general public confused about the quality of their local schools (Dillon, 2006). Though the 2005-2006 deadline for teachers to meet HQT has been extended, most states report they are on track to meet these requirements. However, the wide variation in how states define teacher quality frequently dilutes the purpose of this provision (Henig, 2006). There is widespread skepticism among state and district officials that the teacher requirements are improving the quality of teaching (Center on Education Policy, 2006). And, much needed empirical analyses of the effects of NCLB on students and teachers are just beginning to emerge (e.g., Center for Education Policy, 2006; Fuller, 2006; Lee, 2006).

Deleted: b

In 2007, with or without strong empirical evidence of its effectiveness, NCLB is very likely to be reauthorized. It passed in 2002 with broad bipartisan support. Even with changes in the legislature after the 2006 midterm elections, NCLB's lofty and worthy aims – its commitment to assuring all children achieve at a high level and its strong focus on accountability for results – will likely prevail. However, further analysis and discussion of NCLB are needed. We view this critical juncture as a time for teacher educators to participate more actively in

current conversations about a number of important and highly contested questions regarding the way NCLB and ESR are shaping k-12 classroom practice, teachers' lives, and the work of teacher education. Here, we take stock of the laws' impact on teacher education research, practice, and policy. We focus our examination on only one of several controversial aspects of the laws—the emphasis on scientifically-based research (SBR). Frederick Hess and Michael Petrilli (2006) explain,

> The phrase 'scientifically based research' appears more than 100 times throughout the No Child Left Behind Act and is applied to policies addressing reading programs, teacher training, drug prevention and school safety, and a range of other topics. Scientifically based research has no title or program of its own, but it is woven into the fabric of virtually every program in the law. As a result, this emphasis has potentially far-reaching consequences for both daily classroom practice and academic research related to education. In addition, by making the federal government a more active partner in determining what specific instructional methods should be approved for classroom use, NCLB also sets a new precedent of federal involvement in curriculum and instruction (p. 94).

NCLB and ESR's bold efforts to transform educational research are both troubling and promising. We consider ways in which the pervasive emphasis on scientifically-based research narrows the conversation about teacher education research, practice, and policy, and then explore positive outcomes of these efforts.

## The Narrowing Effect of Scientifically-Based Research

Narrow construal of educational research. NCLB's emphasis on scientifically-based research unduly narrows what can be construed as acceptable educational research. The law's

definition of scientifically-based research promotes experimental designs modeled after medical research. Attempting to make education an "evidence-based field" more akin to medicine, the law establishes randomized field trials as the "gold standard." It further stipulates the following about scientifically-based research:

> [SBR]is evaluated using experimental or quasi-experimental designs in which individuals, entities, programs, or activities are assigned to different conditions and with appropriate controls to evaluate the effects of the condition of interest, with a preference for random-assignment experiments, or other designs to the extent that those designs contain within-condition or across-condition controls (NCLB, section 9101).[2]

In this sense, the law codifies one position in a complex, ongoing debate about the quality and rigor of educational research. While we do not question the elegance, power, and utility of experimental designs, they cannot answer all important questions the field faces; other genres of research are both necessary and relevant to maintain the vitality of educational research (Borko, Liston, & Whitcomb, 2007). The federal government's legislation of a particular method has a dampening effect on research in education writ large, and by extension, teacher education. Not only does the law limit the possibilities for future federally funded research, but also and equally important, it narrows the set of existing research available to inform policy decisions that guide k-12 and teacher education practices.

How experimental designs (including randomized field trials and quasi-experimental) play out in the culture of accountability concerns us on several fronts. First, a hallmark of these research designs is the use of psychometrically sound dependent measures that allow one to make claims about causality with sufficient statistical power. This critical design feature leads

[2]Readers will find a complete definition of scientifically-based research in section 9101 of General Provisions of the *No Child Left Behind Act.* http://www.ed.gov/policy/elsec/leg/esea02/pg107.html

researchers to overuse standardized achievement tests as proxies for learning. While high performance on standardized achievement tests is unquestionably a valued student outcome, other outcomes also matter. For example, too many achievement tests do not sufficiently tap conceptual understanding, thus narrowing the range of learning goals for which schools are held accountable. Furthermore, other important goals – such as social-emotional learning and motivation to learn – are only considered valuable to the extent that they are intermediary variables positively associated with achievement. We are concerned that the U.S. Department of Education's strategic goal to foster a culture of achievement, its emphasis on achievement in reading and mathematics, and its commitment to scientifically-based research, work together to pressure the education community to discount or ignore other potentially valuable aims of education.

A second concern with experimental designs is that too often program implementation is a black box. In many experimental studies, qualitative data about the nature and extent of implementation are not collected and analyzed. Such experimental studies do not consider how teachers shape the enactment of a particular program or intervention. Yet, good teachers will always focus more upon their students than upon faithful implementation of a model. Teachers tinker. They exercise professional judgment to adopt or modify an intervention model to meet their students' needs. They talk and share ideas about practice. These professional practices raise critical questions about whether one can or should ensure fidelity to an instructional model when conducting large scale randomized field trials. Berman and McLaughlin (1978) used the term "mutual adaptation" to account for the ways in which educational innovations and their users change in the process of implementation. Mutual adaptation is essential if a program is to be implemented effectively. At the same time, too much adaptation can mean that the intent of the

program is lost. By not taking into account the level of implementation and nature of adaptation, findings about the effectiveness of a particular intervention are difficult to interpret.

Indeed, a recent RAND study, *Evaluating Comprehensive School Reforms at Scale*, found that of the four reform models studied, none was fully implemented; schools using the same model varied in their level of implementation; and there were few differences in practices between model and non-model schools (Vernez, Karam, Mariano, & DeMartini, 2006). The RAND team also found that schools were able to adopt curriculum most easily; they experienced more difficulty implementing the model's instructional practices. Changing instructional practice is hard work, and many studies of professional development have shown that teachers need significant support and time to change their practices (Putnam & Borko, 2000; Wilson & Berne, 1999). In the RAND study, teachers' commitment up front and the quality of initial support were important determinants in whether a model was implemented at higher levels. Given the study's finding that comprehensive school reforms were not fully implemented, this study illustrates the importance of understanding how teachers enact a model when making claims about its effect on achievement.

Finally, we are concerned that as the law reshapes educational research, the only acceptable research for teacher education may be those studies that attempt to build the causal chain of evidence from its starting point in teacher education programs to the final link to pupil achievement. To build an effective chain of evidence, teacher education researchers must demonstrate that a particular teacher education practice/policy shapes teacher candidate learning, which in turn shapes teacher candidates' actual classroom practice, which in turn positively influences k-12 learning that is reflected in student achievement measures (Cochran-Smith & Fries, 2005). To build the linkages, researchers often use a proxy for each point in the chain of

evidence. Up to now, those studies seeking to develop these linkages have used rather thin proxies for each link in the chain. For example, Goldhaber and Brewer's (2000) study of the influence of teacher certification on student achievement, which is arguably one of the better examples of this kind of research, uses measures such as the type of certification (standard, probationary, emergency, private school, or no certification) as a proxy for teacher education. One limitation of this measure is that the certification categories (which were those used in NELS:88) do not correspond well to those used by policymakers; in addition, the teacher learning experience varies considerably both within and across these certification categories. Other widely used measures of teacher learning focus primarily on teachers' knowledge of content, which is essential for effective teaching but not sufficient. Furthermore, these measures – typically undergraduate major or number of courses – may not capture the content knowledge needed for teaching (Ball, Hill & Bass, 2005). Moreover, gathering reliable and valid assessments of classroom practice has proven to be remarkably difficult, time-consuming, and expensive. We worry that the proxies used to measure teacher education practice/policy, teacher learning, teacher practice, and ultimately student learning are each thin representations of the educational endeavor. When taken together, they reduce what is meaningful about teaching, learning, and learning to teach.

These concerns point to the value of supporting multiple genres of research in teacher education. Studies within the interpretive genre, for example, can provide a more nuanced, in-depth understanding of the relationships among teacher learning, teacher knowledge, and instructional practices than is possible in large-scale experimental studies (Borko, Liston & Whitcomb, 2007). Pressures to conduct experimental or quasi-experimental research, combined with inadequate funding for teacher education research (Zeichner, 2005) increase the probability

that researchers will rely upon thin proxies for complex educational processes and outcomes. In short, the legislative definitions of scientifically-based research narrow educational research. This, in turn, has implications for program design and practice in teacher education.

Narrow construal of teacher education practice. By narrowing what counts as research and by pressing for a tight alignment between scientifically-based research and teachers' classroom practice, NCLB is without question changing the k-12 landscape. For example, the Center on Education Policy (2006) reports that NCLB has led to both a narrowing of the curriculum and more direction about teaching. Specifically, they note,

> …seventy-one percent of the school districts we surveyed reported that they have reduced elementary school instructional time in at least one other subject to make more time for reading and mathematics and many … districts have become more prescriptive about what and how teachers are supposed to teach (p. 2).

These changes, combined with the pressure to connect teacher education with student achievement, may oblige teacher educators to tailor the knowledge base for teaching and the curriculum of teacher education to respond to the changing reality of k-12 classrooms.

Current conceptualizations of the knowledge base for teaching delineate three broad areas—learners and learning in social contexts, curriculum and subject matter, and teaching (Darling-Hammond & Bransford, 2005). Relying on scientifically-based research is likely to focus knowledge of curriculum and teaching on those curricula and instructional practices given the stamp of approval in experimental or quasi-experimental studies. For example, the "What Works Clearinghouse" provides an "intervention rating" that characterizes the effects of educational interventions (from "positive effects" to "negative effects"), and a scale for rating the quality of studies of interventions (from "meet evidence standards" to "does not meet evidence

screens") (http://www.whatworks.ed.gov/). Only studies using randomized controlled trial or quasi-experiment research designs can be rated as meeting evidence standards (either with or without reservations). At this point, very few studies in key areas such as beginning reading or elementary and middle school mathematics have received the most positive ratings. Yet, in the future, school districts that receive money through federal programs as well as schools failing to make adequate yearly progress will be expected to draw from this research base as one option when developing improvement plans. To better support NCLB, the logical inference is that candidates should learn the principles and practices of "proven" programs.

The vision of teaching implied is one of teacher as a technical-rational implementer of curriculum models rather than a reflective practitioner (Zeichner & Liston, 1996) or an "adaptive expert" (Darling-Hammond, 2006) who makes judgments and decisions in the face of uncertain and complex situations. This more restricted vision of teaching suggests that an appropriate role for teacher education is to ensure that teachers have the knowledge and skills necessary to implement only those instructional practices and curricular materials that have scientifically based evidence for their effectiveness. Thus far, teacher educators have resisted this narrowing of the teacher education curriculum, but it may become increasingly more difficult to do. As schools that fail to make AYP are required to develop and follow improvement plans where only scientifically-based practices are allowed, they will likely seek new teachers who are well trained to implement these specific practices.

Narrow policy construals. Recent discussions have focused on policy development that promotes teacher quality. In these discussions teacher quality is usually defined narrowly in terms of student achievement on standardized tests: Purportedly, the more students achieve, the greater their teacher's quality. Debates have centered on which aspects of teacher knowledge

matter most in defining quality teachers and whether teacher certification makes a difference. Most recently "value-added assessment" has been suggested as a way to determine quality teacher preparation.

Policymakers have found value-added assessment a helpful tool to identify whether teachers make a difference on student achievement. This statistical method has been employed in educational research to show the impact a teacher can have on student standardized test scores (Sanders, 1998). States and districts are now developing policies and data management systems that allow them to match and connect individual teachers, information about the teacher's preparation experience (whether university-based program or other route), and the teacher's students' test scores. The implication appears to be that good teacher preparation will produce quality teachers who enable students to learn and show greater gains on their standardized tests. Ineffective teacher education routes will not produce teachers who enhance their students' test scores.

Teacher educators need to be aware of this developing evaluative logic. In essence, teacher education programs will be evaluated on how well their graduates' students do on standardized tests. Indeed, Ohio's Teacher Quality Partnership is using "value-added assessment to estimate effects of new teachers, an effort that provides information about teacher preparation programs and practices" (Wiley, 2006, p. 34). Numerous issues and problems can be identified. For example, value-added assessment, as a tool to determine teacher education program quality, would seem to rely on too distal and too narrow a measure. Many intervening variables must be accounted for to make definitive causal claims and corresponding evaluative judgments. While some studies now account for student characteristics and include more measures of teacher preparation (e.g., number of courses or scores on PRAXIS exams), more complex models would

also consider substantive features of teacher preparation experiences and measures of actual teaching performances. Workplace factors such as teaching assignment, curricular materials, mentoring or coaching, and administrative leadership all have a bearing on a teacher's performance and ability to foster student achievement and need to be included. To make well-reasoned causal claims regarding the effectiveness of teacher education, an analysis to determine which teacher preparation routes "add value" would need to trace the entire chain of evidence. An informative value-added analysis would need to account for the impact at each point in the chain of evidence

Quantitative methodologists have identified critical concerns regarding the validity of estimates made in each of the value-added assessment (VAA) approaches currently in use (AERA, 2004; McCaffrey, Lockwood, Koretz, & Hamilton, 2003). Wiley cautions the field:

> Implementers are always wise to remember that because it is generally impossible to randomly group teachers and students, and VAA estimate of teacher effects may be influenced by other factors…. While a particular approach may attempt to deal with this issue by using statistical adjustments, no set of adjustments can fully compensate for the lack of randomization. Furthermore, important questions remain about whether student achievement scores are appropriate measures of teacher effects; whether achievement tests are appropriately designed; whether and how assessment errors may affect estimates; and when assessments are best administered. Although these are crucial and complex questions, no VAA approach yet takes them into effect (p. 51).

Another set of issues concerns the definition of and practices associated with a "professional" education. Are there bodies of knowledge and understandings that teachers should master that are not aimed directly at raising students' test scores? Should teachers be acquainted

with the history of US public education? Should educators explore their ethical obligations to students, parents and their profession? Should teachers be prepared to be curriculum developers or deliverers? Should some joy and satisfaction be expected in the learning and teaching process?

Certainly additional issues could be outlined. For our purposes here we want to underscore the currently developing value added means of appraising teacher education efforts. Unfortunately it appears that the early 1900's logic of and concern for social efficiency has reappeared in the logic of value added appraisals. Raymond Callahan's (1962) delineation of the public school's "cult of efficiency" has become the "cult of VAA."

Promising Possibilities of Scientifically-Based Research.

Thus far we have depicted several ways in which NCLB and the Education Science Reform Act's emphasis on scientifically-based research narrows teacher education research, practice, and policy. We also see, however, possibilities for this legislation to focus the teacher education community on the quality and rigor of our research, to examine critically the practices and knowledge conveyed in teacher education programs, and to argue more strenuously for a broadened policy agenda that will improve linkages between teacher education, induction, and teachers' ongoing development.

Improving research. Legislative definitions of scientifically-based research press educators to debate what constitutes quality and rigor in educational research and more specifically research on teacher education. Such discussions open the door for much-needed studies that build the complex evidentiary chain illustrating how powerful teacher education practices contribute to candidates' teaching performance, which in turn foster student learning, and ultimately their performance on achievement tests. The field of teacher education stands to benefit from research of this type, especially if these studies demonstrate impact on the

achievement of diverse learners. This is particularly true if the research is not limited by a narrow construal of acceptable research designs, but rather incorporates multiple methods in a single project. Two promising examples of such research are the *Teacher Pathways Project* which examines multiple outcomes of different pathways into teaching in New York City (Boyd, Grossman, Lankford, Loeb, Michelli, & Wyckoff, 2006) and the *Teacher Quality Partnership*, which explores how the preparation and development of new teachers in Ohio affects students' academic performance (Teacher Quality Partnership, 2006). By giving prominence to rigorous and quality research and by placing the role of teacher education and certification within those discussions, we may be more likely to build the chain of evidence that will link teacher education to student learning.

Examining practices and knowledge base. Though we are concerned that sound classroom practices may be discounted because they presently lack scientifically-based evidence to support them, we recognize that current debates about which practices raise student achievement challenge teacher educators to evaluate those k-12 practices introduced and modeled within teacher education programs. We hope teacher educators have hearty deliberations about the instructional practices and curriculum models candidates encounter and develop beginning mastery of in their teacher education programs.

The NCLB themes of scientifically-based research and accountability suggest one area where the knowledge base for teaching (and associated teacher education practices) ought to be more developed. Some argue that teachers need to know more about educational research design so they may be "critical consumers" of research. For example, in defending the federal government's position on scientifically-based research, Virginia Reyna asserts, "As teachers, administrators, and policy makers become educated about the scientific method, they will

become more skeptical consumers. Less snake oil will be sold and more real medicine" (2005, p. 17). In this issue, Walsh considers this position in her book review. Currently, knowledge of research design is not heavily emphasized in many discussions of the knowledge base for teaching. We agree that it is beneficial for teachers to know about research, all genres of research, though we also recognize that teachers are seldom acknowledged as decision makers when it comes to curriculum or instructional programs as adopted; thus, detailed knowledge of research design may not be of much practical value for teachers, as they are unlikely to be asked to evaluate the research evidence for a particular program.

It is perhaps more critical that teacher educators' knowledge of research design, and in particular large-scale randomized field trials and quasi-experimental designs, be enhanced. Stronger understanding of these designs will help teacher educators to evaluate instructional practices and curricular materials for possible inclusion in their courses.

Broadening policy discussion. Current policy efforts that use thin measures to determine teacher education program quality reinforce the "horse-race" approach to evaluating teacher education routes. More importantly, they deflect attention from much-needed efforts to build a more comprehensive policy approach that will strengthen both teacher education programs and the teaching force. They also deflect attention from a critical policy problem of the growing gap in the distribution of quality teachers. If you are a child growing up in a rural or poor urban community, you are more likely to have an under-qualified teacher.

This is a time for teacher educators to insert themselves more deliberately in policy discussions. We need a broader policy agenda that will address this growing teacher quality gap. For instance, the field needs a coherent policy agenda that simultaneously addresses higher standards for all preparation routes, rigorous performance assessments for teacher candidates,

improving the funding levels of teacher preparation to reflect the true costs of high quality preparation, incentives to recruit and retain candidates willing to teach in hard-to-teach areas, more rigorous evaluations of teachers in their initial years, and improvements to the working conditions, particularly for beginning teachers.

Conclusion

As this editorial goes to press, the 2006 midterm elections have just concluded. Though the election has reshaped the party balance in the federal legislature, we suspect that it will not have a strong effect on NCLB's reauthorization in 2007. Five years after its inception, NCLB's impact is beginning to be felt. As Tommy Thompson, Co-Chair of the Commission on No Child Left Behind, noted in his remarks at the Commission's hearings,

> NCLB has dramatically changed the national conversation about education by bringing a stronger focus on accountability for results and a commitment to assuring that all children – regardless of race, economic status or where they live – achieve at a high level. The law has provided a context for looking beyond how our own children are performing in our own schools. We are now much more aware and informed about the quality of education being provided to students across the country. (The Commission on NCLB, 2006).

When Congress convenes to consider the reauthorization, it will certainly consider the Commission's recommendations for ways of revising the law and improving its implementation so as to make it even more effective in closing the achievement gap. We urge the teacher education community to speak up – offering commentary about the positive and negative impact of the legislation on teacher education research, policy and practice. Such commentary is sorely needed. Certainly the nation's approach to teacher education must be examined, especially in light of NCLB and scientifically-based research. Accountability is required and so is a rich and

full sense of what it means to be humane and human in our schools, both as learners and as teachers.

References

American Educational Research Association (2004). Teachers matter: Evidence from value-added assessments. *Research* Points, *2* (2). Retrieved on November 30, 2006 from http://www.aera.net/uploadedFiles/Journals_and_Publications/Research_Points/RP_Summer04.pdf

Ball, D.L., Hill, H.C., & Bass, H. (2005, Fall). Knowing mathematics for teaching: Who knows mathematics well enough to teach third grade, and how can we decide? *American Educator,* 14-17, 20- 22, 43-46.

Berman, P., & McLaughlin, M.W. (1978). *Federal programs supporting educational change, Volume VII: Implementing and sustaining nnovations.* Santa Monica, CA: RAND Corporation.

Borko, H., Liston, D., & Whitcomb, J. (2007). Genres of empirical research in teacher education. *Journal of Teacher Education*, *58* (1), xx-xx.

Boyd, D., Grossman, P., Lankford, H., Loeb, S., Wyckoff, J., & McDonald, M. (2006). *Examining teacher preparation: Does the pathway make a difference?* Retrieved February 4, 2006, from http://www.teacherpolicyresearch.org/portals/1/pdfs/Examining_Teacher_Preparation_Full-Description.pdf

Bracey, G.W. (2006). The 16th Bracey Report on the Condition of Public Education. *Phi Delta Kappan*, *88* (2), 151-166.

Callahan, R.E. (1962). Education and the cult of efficiency. Chicago, IL: University of Chicago Press.

Deleted: ¶

Carlson, J.S. & Levin, J.R. (2005). *The* No Child Left Behind *legislation: Educational research and federal funding*. Greenwich, CT: Information Age Publishing.

Center on Education Policy (2006). From the capitol to the classroom: Year 4 of the No Child Behind Act, Summary and Recommendations. Retrieved 11/23/06 from http://www.cep-dc.org/nclb/Year4/NCLB-Year4Summary.pdf

Cochran-Smith, M. & Fries, K. (2005). Researching teacher education in changing times: Politics and paradigms. In M. Cochran-Smith & K. Zeichner (Eds.). *Studying teacher education: The report of the AERA panel on research and teacher education* (pp. 69-109). Washington, DC: American Educational Research Association.

Commission on No Child Left Behind (2006). National Education Organizations and Policy Experts Testify at Final Commission Hearing. Press release downloaded 11/23/06 from http://www.aspeninstitute.org/site/c.huLWJeMRKpH/b.1200931/k.6F95/Press_Releases Advisories.htm

Darling-Hammond, L. (2006). *Powerful teacher education: Lessons from exemplary programs*. San Francisco: Jossey-Bass.

Darling-Hammond, L. & Bransford, J. (2005). *Preparing teachers for a changing world: What teachers should learn and be able to do*. San Francisco: Jossey-Bass.

Darling-Hammond, L, & Youngs, P. (2002). Defining "Highly Qualified Teachers": What Does "Scientifically-Based Research" Actually Tell Us?. *Educational Researcher*, 31(9), 13-25.

Dillon, S. (2006). As 2 Bushes Try to Fix schools, the Tools Differ. *New York Times*. September, 28, 2006.

Elmore, R.F. (2002). Testing trap. *Harvard Magazine*, *105*(1), 35.

**Deleted:** Cochran-Smith, M. (2005). No child left behind: 3 years and counting. *Journal of Teacher Education, 56* (2), 99-103.¶

**Deleted:** xx (xx

**Deleted:** )

**Deleted:** Dillon, S. (2006a). Reports says education officials violated rules. Retrieved September 23, 2006 from http://www.nytimes.com/2006/09/23/education/23education.html?ei+5070&en=49ac85d0aaf523... ¶

**Deleted:** b

**Formatted:** Font: Italic

Fuller, B. (2006). *Is the No Child Left Behind Act Working?*. Berkeley: Policy Analysis for California Education.

Goldhaber, D. & Brewer, D. (2000). Does teacher certification matter? High school teacher certification status and student achievement. *Educational Evaluation and Policy Analysis*, 22(2), 129-145.

Henig, S. (2006). Back to School for Teachers: The No Child Left Behind Act has Changed how Colleges Serve Classroom Practitioners. Retrieved 10/06/2006 from http://chronicle.com/weekly/v53/i07/07a02001.htm

Hess, F.M. & Finn, C.E. (2004). *Leaving no child behind? Options for kids in failing schools*. New York: Palgrave.

Hess, F.M. & Petrilli, M.J. (2006). *No Child Left Behind: Primer*. New York: Peter Lang.

Lee, J. (2006). *Tracking Achievement Gaps and Assessing the Impact of NCLB on the Gaps*. Cambridge, MA: Civil Rights Project, Harvard University.

Linn, R. (2003). Accountability: responsibility and reasonable expectations. *Educational Researcher*, *32* (7), 3-13.

McCaffrey, D.F., Lockwood, J.R., Koretz, D.L., Hamilton, L.S. (2003). *Evaluating value-added models for teacher accountability*. Santa Monica, CA: RAND.

Reyna, V. F. (2005). The *No Child Left Behind Act* and Scientific Research: A View from Washington, DC. In J.S. Carlson & J.R. Levin (Eds.) *The* No Child Left Behind *Legislation: Educational research and federal funding*, (pp. 1-25). Greenwich, CT: Information Publishing.

Putnam, R. T. & Borko, H. (2000). What do new views of knowledge and thinking have to say about research on teacher learning? *Educational Researcher*, *29*(1), 4-15.

**Deleted:** Feuer, M.J., Towne, L., & Shavelson, R. (2002). Scientific culture and educational research. *Educational Researcher*, *31* (8), 4-14.¶

**Formatted:** Indent: First line: 0"

**Deleted:** http://chronicle.com/cgi-bin/printable.cgi?article=http:.....¶
Hess, F.M. (2006). The wrong kind of bipartisanship. *Teachers College Record*, Date Published: November 10, 2006 http://www.tcredord.org ID Number: 12836, Date Accessed: 11/14/2006 7:10:31 PM.¶
H

**Deleted:** Kane, T., Rockoff, J., & Staiger, D. (2007). Photo finish: Certification doesn't guarantee a winner. *Education Next*, (winter), 61-67.¶

**Deleted:** National Research Council. (2002). *Scientific research in education*. R.J. Shavelson & L. Towne (Eds.), Committee on Scientific Principles in Education Research. Washington, DC: National Academy Press.¶

**Formatted:** Font: Times New Roman, 12 pt

**Formatted:** Font: Times New Roman, 12 pt, Italic

**Formatted:** Font: Times New Roman, 12 pt

Sanders, W.L. (1998). Value-added assessment. *School Administrator*, *11*(55), pp. 24-27.

Teacher Quality Partnership (2006). http://www.teacherqualitypartnership.org/

The White House, Fact Sheet: The No Child Left Behind Act: Challenging Students through High Expectations. (2006). Retrieved on 10/10/2006 from http://www.whitehouse.gov/news/releases/2006/10/print/20061005-2.html

Thernstrom, A. & Thernstrom, S. (2003). *No excuses: Closing the racial gap in learning*. New York: Simon and Schuster.

Vernez, G., Karam, R., Mariano, L.T., DeMartini, C. (2006). *Evaluating Comprehensive School Reform Models at Scale: Focus on Implementation*. Retrieved on November 18, 2006 from http://www.rand.org/pubs/monographs/MG546/

Wiley, E.W. (2006). A practioner's guide to value added assessment. (Educational Policy Studies Laboratory Research Monograph). Tempe, AZ: Arizona State University.

Wilson, S. M., & Berne, J. (1999)." Teacher learning and the acquisition of professional knowledge: An examination of research on contemporary professional development." In A. Iran-Nejad & P. D. Pearson, (Eds.), *Review of research in education.* (pp. 173-209). Washington, DC: American Educational Research Association.

Zeichner, K. (2005). A research agenda for teacher education. In M. Cochran-Smith & K. M. Zeichner (Eds.). *Studying teacher education: The report of the AERA Panel on Research and Teacher Education* (pp 737-759). Mahwah, NJ: Erlbaum.

Zeichner, K. & Liston, D. (1996). Reflective teaching: An introduction. Mahwal, NJ, Lawrence Erlbaum.

**Deleted:** ). Insert rest of citation