



The Unraveling of No Child Left Behind:
How Negotiated Changes Transform the Law

By

Gail L. Sunderman

Foreword by Gary Orfield

February 2006

Copyright © 2006 by President and Fellows of Harvard College

All rights reserved. No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopy, recording, or any information storage and retrieval systems, without permission in writing from The Civil Rights Project.

This publication should be cited as:

Sunderman, G. L. (2006). *The Unraveling of No Child Left Behind: How Negotiated Changes Transform the Law*. Cambridge, MA: The Civil Rights Project at Harvard University.

Additional copies of this report may be obtained from our website at:

<http://www.civilrightsproject.harvard.edu>

Produced with generous support from the Bill and Melinda Gates Foundation, the Charles Stewart Mott Foundation, and the National Education Association.

TABLE OF CONTENTS

| | |
|--|----|
| LIST OF TABLES | 3 |
| LIST OF FIGURES | 4 |
| ACKNOWLEDGEMENTS | 5 |
| FOREWORD | 6 |
| EXECUTIVE SUMMARY | 9 |
| Two-Fold Approach to Amending State Accountability Plans | 9 |
| Implications of Changes | 10 |
| INTRODUCTION | 11 |
| METHOD | 12 |
| EXAMINING THE FEDERAL ROLE IN EDUCATION..... | 13 |
| FEDERAL “FLEXIBILITY” TO IMPLEMENT NCLB | 15 |
| A ‘New Path’ for NCLB | 21 |
| State Initiated Policy Changes | 27 |
| TRACKING CHANGES TO STATE ACCOUNTABILITY PLANS..... | 31 |
| IMPACT OF RULE AND STATE INITIATED POLICY CHANGES | 39 |
| Effect on AYP and Accountability | 40 |
| Who Benefits/Who Loses? | 45 |
| Added Complexity | 48 |
| DISCUSSION | 52 |

LIST OF TABLES

| | |
|---|----|
| Table 1: Summary of Rule Changes to NCLB Allowed by US Department of Education. | 15 |
| Table 2: Percentage of Core Classes Taught by Highly Qualified Teachers in all Schools in Six States, and in High- and Low-Poverty Schools, 2002-03 and 2003-04. | 26 |
| Table 3: State Initiated Policy Changes: State Accountability Plan Amendments Approved by U.S. Department of Education. | 27 |
| Table 4: Summary of District Waivers Granted by U.S. Department of Education. | 31 |
| Table 5: States Taking Advantage of Policy Changes Adopted by U.S. Department of Education through December 2005. | 32 |
| Table 6: States Taking Advantage of State Initiated Changes: Adopted Same Subject and Grade Span for Accountability, Changed Subgroup Size, and Made Statistical Changes. | 34 |
| Table 7: States Taking Advantage of State Initiated Changes: Changed Definition of AYP, Annual Measurable Objectives, Additional Indicator. | 36 |
| Table 8: States Taking Advantage of State Initiated Policy Changes: Adopted Changes in Assessment System, How Schools are Included in Accountability System, How States Intervene in Schools, and How States Report NCLB Data. | 38 |
| Table 9: Range and Average Number of Changes Adopted by States by Type of Change (Rule Change, State Initiated, Total Changes), December 2005. | 39 |
| Table 10: Number and Percent of Schools by Enrollment Category, Georgia, 2005-06. | 41 |
| Table 11: Number and Percent of Districts by Enrollment Category, Virginia, 2004-05. | 41 |
| Table 12: Number of Districts Reporting Subgroups at Time 1 (if subgroup >50) and Time 2 (if subgroup >50 or 1% of total enrollment), Virginia, 2004-05. | 42 |
| Table 13: School Level AYP: Percentage of Schools Making AYP, Identified for Improvement, and “Provisional AYP,” Florida, 2003-04 and 2004-05. | 44 |
| Table 14: The Number and Percentage of Districts Making Adequate Progress and Identified for Improvement Using the AYP Method and the Grade Span Method, California and Georgia, 2004-05. | 45 |
| Table 15: Demographic Characteristics (Percentages of Minority, Low-Income, and English Language Learners) by District Accountability Status under the AYP and Grade Span Methods, California, 2004-05. | 46 |
| Table 16: Number of Districts and Schools Participating in the SRSA and RLIS Programs, Enrollment and Demographic Characteristics of Students, 2004-05. | 48 |
| Table 17: Number and Percentage of Students Who Must Meet an AYP Goal of 50% Proficient Using Different Confidence Intervals. | 50 |

LIST OF FIGURES

Figure 1: Plot of confidence intervals around different percent proficiency levels..... 50

ACKNOWLEDGEMENTS

The authors are grateful to numerous individuals for their assistance with this report. Christopher Tracey contributed greatly by developing the matrix used to track the amendments and changes made to state accountability plans. We are very grateful to our research assistants, Damon Clark, Jongsun Kim, and Ann Owens who worked tirelessly to help prepare this report and assisted with the research. We would also like to thank members of the Civil Rights Project's administrative team, including Jennifer Blatz, Lori Kelley, Pamela Lomax, Jerry Monde, and Christina Safiya Tobias-Nahi who helped with the production of the report. We received very useful comments from our external reviewer, Richard Elmore. Finally, we would like to acknowledge the contributions of state and district officials who have contributed knowledge, time, and data to our study of NCLB.

FOREWORD

A fundamental problem with the No Child Left Behind Act (NCLB) as reauthorization approaches is that what once seemed a clear if highly controversial policy has now become a set of bargains and treaties with various states. Federal aid programs rarely produce raging controversy, but NCLB, the major domestic policy accomplishment of the Bush Administration, has stirred passions across the country. Some of the attacks, coming from conservatives as well as liberals claim that the act itself or the policies being enforced under the act are illegal or even unconstitutional. You can see these claims in the report of the National Conference of State Legislatures and from the leader of the legislative battle to withdraw Utah from the program. This report is not about claims regarding the constitutionality of the law, but it does consider the idea of lawfulness in another important sense—the creation of uniform, neutral, clear standards that can be readily understood by those who are duty bound to obey them. Needless to say, if serious sanctions are connected to far-reaching and substantive requirements, people on the front line want clear and consistent policies so that they can know what is expected and what will trigger loss of funds or imposition of drastic educational changes such as dissolving a school and removing its faculty.

Inevitably when a sweeping, one-dimensional policy is imposed on a vast and sprawling nation with a tradition of striking decentralization and massive differences in many conditions affecting schooling, there will be problems and adjustments to make. In this sense, it is certainly not surprising that the most expansive assertion of federal power over schools in American history should encounter unexpected and sometimes very deep problems that require changes and modifications. These problems were greatly intensified by requirements that schools show improvements in performance every year, and in many cases, achieve above average performance for all students, something that had never been done in any high poverty district. These unforeseen circumstances are exactly what congressional oversight and good evaluation by administrators and researchers are supposed to discover and correct.

It is not surprising that controversy has erupted over NCLB; it was wholly predictable. A major problem has been the failure of Congress and the Administration to acknowledge and correct the problems even as opposition has grown across the nation. Congress has not provided serious oversight of the working of the law and has not adopted timely amendments. The Bush Administration has not commissioned independent research on the implementation of the policy and refused to admit rather obvious mistakes until virtual rebellion took hold in the field. For years the Administration engaged in political attacks on those who pointed out the problems and then insisted that no substantial changes in the law were needed. Any federal program, however, needs continuing support from a Congress elected at the state and local level. Recognizing that something had to be done to respond to critical problems with the law and the large number of schools being identified as failing, the Administration negotiated changes individually with states. While these ad hoc deals to respond to the massive political and professional criticism of the policy have led to something that may be useful

for short-term political support, they have created massive confusion and resentment across the country and cynicism about the real meaning of the law and its administration.

Among the claims of supporters of the law is that it provides invaluable information to parents of children of every race, those who do not speak English or are handicapped. The law sets out what seem to be very clear goals and consequences. However, as researchers predicted before the law was enacted, huge numbers of schools would be branded as failures, including many that are seen as successes and often have rising achievement levels. In response, the Administration is permitting a wide variety of changes that lower the failure rate. In the process, though, interpreting reports on how well schools are doing is becoming nearly impossible. How many parents or citizens can understand statistical confidence levels, multi-year averaging to count for a single year's progress, elimination of some students from various subgroups when calculating progress, the expanded "safe harbor" provision and many other changes that lower the number of apparent failures. I doubt very much that members of Congress or even most researchers would have any idea how to interpret the resulting changes in failure levels. It is becoming as complex as the adjustments made to various schedules in federal income taxes that are now impossible to understand or calculate without a computer and the right software. Because the Administration wants to pretend that it has not changed the law while actually changing it in very consequential ways, it has reduced the reports of school success and failure to unintelligible gobbledygook.

What is a law? A political scientist or lawyer could spend a semester discussing that question but a rough definition might be something like this: a rule of general application established by and enforced by governmental power. A lawful policy is one that operates within the principles established by the law. Is NCLB being administered in a lawful manner? One vital part of enforcing a law is that it is intelligible and predictable. In other words, those subject to enforcement that is backed by public sanctions need to know what the law requires, that the requirements are consistent with the law itself, that they have reliable information about what they must or must not do, and what the consequences may be.

By this standard, some of the most important policies currently being enforced under NCLB are not lawful. They are not rules of general application and they are not consistent from state to state. Sometimes even individual districts have been allowed to ignore the policy requirements. Exceptions available in some places are not publicized or made available in others. There is no certainty in the law and no way to know what it will be in the future. This means that the policy is essentially a product of negotiation, of power and discretion, not law.

Since NCLB concentrates serious power and sanctions around demands that turned out to be unattainable and, sometimes, irrational, school officials and the public face the contradiction between explicit and very harsh policies and a patchwork of policies that have been compromised in an inconsistent and unpredictable pattern. These compromises constitute the real policy, and, since they are applied differently to different regions of the country, there is no single, national policy. If Texas and Florida are

permitted particularly sweeping exceptions, there is suspicion of favoritism. This risks the perception that federal education policy is both extremely intrusive and has become nothing more than another arena politics where what you must do depends on who you know or which official you ask. For a policy that is the central education law in the nation and has forced many states and school districts to do things they did not want to do, imposed negative ratings on many schools, and threatened very serious sanctions against schools that did not meet federal requirements for annual progress, the inconsistency and unpredictability of the process for interpreting the law introduces confusion and suspicion and diverts energy from the goals to the process of winning exceptions and special treatment.

From a policy standpoint, allowing some states to ignore or change various elements without changing the law or communicating a clear alternative policy that school officials can reliably plan around may solve current political problems but undermines the relationships and the legitimacy of both the policy and the Department of Education. It certainly weakens the position of officials wishing to enforce standards against clear and willful violations of the law. If the perception is that there is no law, any enforcement action will be seen as arbitrary and unfair.

An Administration so strongly committed to the principles of a law it has consistently hailed as positive is now backing away from it by making compromises on many dimensions. The dangerously high levels of opposition at the state and local level will inevitably affect support for the law in Congress. In this situation, since Congress controls reauthorization, appropriations, and oversight, the law cannot be sustained as originally written. The Administration's only choice is between denying that it has changed the law while continuing to make compromises with states or making open changes of general application and admitting that mistakes were made in the enactment and enforcement of a very ambitious reform. I believe that the Administration has chosen the wrong alternative and that its political success will be short-lived. Congress has abdicated its responsibility. The better alternative would be for Congress and the Administration to call in state and local educators and experts, community leaders and civil rights groups to listen to what has been learned, to reaffirm the viable and workable parts of the law and to make the needed changes, in a clear and open way, by eliminating or changing policies that can no longer be sustained. As part of that process, a review of what has been demanded and granted to the various states would be a valuable first step. This report displays the inconsistencies across states in how the law is being interpreted and could be a valuable part of thinking through what needs to be changed by law, by regulation, and by the announcement of clear new interpretations available to all states and school districts uniformly.

Gary Orfield

EXECUTIVE SUMMARY

Over the past two years, the U.S. Department of Education's (ED) has made such extensive compromises in implementing the No Child Left Behind Act of 2001 (NCLB) that the law's legitimacy is in serious question. In response to growing state and local opposition to the law, political and professional criticisms of its requirements, and the increasing number of schools and districts identified for improvement, the administration has allowed a wide variety of changes in state accountability plans. These changes reflect a political strategy by the administration to respond to the growing state opposition to the law by providing relief from some of the law's provisions and reducing, at least temporarily, the number of schools and districts identified for improvement. But they are also a concession by ED officials that NCLB is not working and have created a policy that has no consistent meaning across states.

This report documents the changes states have made to their accountability plans and examines how these policy shifts affect the meaning of accountability and who benefits (and loses) from the changes. We reviewed decision letters sent to all 50 states that outlined the changes approved by ED through December 2005. The intent of this report is to provide policymakers with information they can use to develop a systemic approach to correcting the flaws in NCLB by documenting the requirements that are difficult for states to implement and identifying areas where the law may not be working as intended. The report provides an easy to understand synopsis of the changes allowed by ED and state-by-state summaries of the amendments each state adopted.

Two-Fold Approach to Amending State Accountability Plans

During the first two years of NCLB, the administration strictly interpreted and enforced the NCLB requirements, rebuffing any attempt to introduce policies that would respond to the concerns raised about the law. This changed beginning in late 2003 and early 2004 when ED announced a series of new policies aimed at aspects of the law that were a source of dissatisfaction. This series of policy changes affected how the students with disabilities and Limited English Proficient (LEP) subgroups were counted for accountability purposes, changed how states could calculate participation rates, and relaxed the highly qualified teacher requirements.

The next major set of changes came in April 2005 when Secretary Spellings announced the "Raising Achievement: A New Path for *No Child Left Behind*" initiative. By this time, hopes that the political opposition to the law would subside with minor adjustments in policy were quelled. States continued to introduce legislation or resolutions during the 2005 legislative session that placed restrictions on the implementation of NCLB. This opposition cut across the political and ideological spectrum, with some of the strongest opposition coming from Republican states. This new initiative took a different approach than the initial rule changes. It required states to adhere to a set of core principles when requesting changes to their accountability plans. It was unclear, however, what would count as evidence or how ED would evaluate whether states met these principles.

At the same time, another process was taking place that produced “state initiated” policy changes. These changes to state accountability plans were negotiated on a state-by-state basis. There were no guidelines on the types of changes states could request, no information on how the requests would be judged, and no guarantees that changes approved in one state would be approved in another. The state initiated amendments were extensive and varied, and included amendments that changed how states determined adequate yearly progress (AYP), allowed states to set different targets for different subgroups, and increased the kinds of statistical techniques states could use when calculating AYP, among others. These behind-the-scene agreements further eroded consistency in how the law was being applied across states.

Implications of Changes

Since the number and kinds of changes that states have adopted are not uniform across states, with each state requesting its own configuration of amendments, accountability no longer has a common meaning across states or even within states. Accountability now depends on which subgroups are included in the system, how each state calculates adequate yearly progress, and which district, school, or subgroup benefit from the various changes states adopted. The allowable statistical techniques states have adopted add complexity to the NCLB accountability system by complicating the meaning of AYP and obscure the ability of states, districts, and schools to show improvements in student performance. Many of the changes simply reduced the number of schools and districts identified for improvement, but without requiring any educational improvement.

Additionally, there were clear winners and losers from the changes. Some of the changes, such as the change in the method used to identify districts for improvement, have compounded the flaws in the NCLB accountability provisions by making it harder for some districts, primarily those serving minorities, to make AYP. Others, such as the changes in the highly qualified teacher requirements benefit some regions of the country over others.

Because this process was politically motivated to respond to growing state opposition to the law, ED has not systemically addressed the underlying flaws in the law. These include the double counting of students in some subgroups for accountability purposes, the reliance on mean proficiency to determine AYP, arbitrary timelines for improving achievement and unrealistic achievement goals that have no connection to what can actually be achieved, a reliance on testing and sanctions to improve schools without corresponding attention to the resources or expertise schools need, and insufficient attention in the law to state capacity to turn around huge numbers of failing schools and districts. If these issues are debated and addressed, it will mean a major overhaul of NCLB and not just a tinkering at the edges in response to political pressure. To improve NCLB, policymakers need to reexamine the core assumptions that underlie NCLB and reevaluate the mechanisms used by NCLB to improve schools and student achievement. To restore legitimacy to the process, policymakers need to include educators, experts, community leaders, and civil rights groups in an open and honest debate about what is needed to reform schools and improve student achievement.

INTRODUCTION

In 2001, Congress passed the No Child Left Behind Act of 2001 (NCLB) in a bipartisan effort to improve education in the United States, particularly for low-income and minority students. Frustrated with the slow pace of change that had characterized previous educational reform, this act was fundamentally different from its predecessor, the Improving American Schools Act (IASA), in the specificity of its requirements and in the seriousness of its effort to expand federal power over education. It required states to implement a single, statewide accountability system based on performance on state reading and mathematics tests. Unlike previous legislation, it specified timelines states must follow to insure all students are proficient, told states how often to test students and which subjects to emphasize, and prescribed a series of sanctions for low-performing schools that failed to improve scores on standardized tests. At the core of the system was the idea that all schools would be held to the same high standards and be accountable for the performance of all of their students. In meeting these goals, little consideration was given to differences in the resources between schools or to the type of students they served.¹ Nor did Congress consult with state and local policymakers or educators when it developed the legislation. Instead, the law represents a political compromise between the political parties and factions within those parties.

Now, four years later, there is a major political upheaval going on in the states that crosses political and ideological lines. States are voicing their dissatisfaction with the law by passing or introducing resolutions and legislation that attempts to limit or contain the role of the federal government in what they view as a state function. As more and more schools and districts are identified as “needing improvement” under the law and subject to sanctions, including many considered as the best in a state, states are rebelling against the strategy for reforming schools established in the law. They are also increasingly seeking waivers to the NCLB provisions. In response, after two years of strictly interpreting the law, the U.S. Department of Education (ED) has negotiated changes in state accountability plans with individual states and announced a series of policy changes designed to ease the stress on the system politically. Since these changes often mean different things in different states and within states have different consequences for different types of districts and schools, they change the meaning of accountability as originally conceived in NCLB. This approach reflects a political strategy that will decrease, at least temporarily, the number of schools and districts identified for improvement. It reduces the stress on the system by taking away accountability for some subgroups and strategically controlling the number of schools subject to sanctions without addressing the fundamental flaws in the law. At the same time, it acknowledges that ED could not sustain a strategy of strictly enforcing the NCLB provisions.

A central component of NCLB is the requirement that all states develop an accountability plan that outlines how the state will implement the law’s provisions and

¹ One exception to the lack of attention to resource differences between schools were the provisions requiring a “highly qualified teacher” in every classroom and that low-income and minority students were not taught at higher rates than other children by inexperienced, unqualified, or out-of-field teachers.

improve student achievement. The first of these plans was due by January 31, 2003, one year after the act was signed into law. Since then, states have amended their plans or changed them to comply with demands from the ED. Both the processes that states followed to amend these plans as well as the substantive changes they have made has not been closely scrutinized. Since changes to state accountability plans may signal either implicit or explicit policy shifts in how NCLB is actually implemented, it is important to understand the nature of these changes, how they affect the meaning of accountability, and who benefits (and loses) from the changes.

Accountability is at the heart of No Child Left Behind. As the largest source of federal funding for K-12 schools, NCLB incorporates tough accountability provisions that require all students, including low-income and minority students, students with disabilities, and students with limited English proficiency, to meet the same achievement goals. Accountability and the toughest sanctions are aimed at schools, but the law also extends accountability to the district, and to a limited extent, to the state. The stakes attached to this accountability are high, which create incentives for states to seek waivers or changes in how the law is implemented.

In this report we summarize the changes states have made to their accountability plans in an effort to inform policymakers and the public about the current status of NCLB. These revisions provide information about requirements that are difficult for states to implement and identify areas where the law may not be working as intended. Rather than continue an ad hoc process of amending state accountability plans, this information can be used to develop a systematic approach to correcting the flaws in NCLB.

METHOD

To understand how states have amended their accountability plans we reviewed the decision letters sent by the U.S. Department of Education to all 50 states. The decision letters outline the changes to state accountability plans approved by the ED.² This review included all decision letters posted by ED on their website through December 2005 and includes amendments approved in 2004 and 2005. To put the amendments in perspective, we reviewed NCLB policy letters to states and key policy letters signed by the Education Secretary or Deputy Secretary that established new parameters within which states could amend their accountability plans.³ We also reviewed press releases and speeches made when announcing policy changes and the non-regulatory guidance that incorporated the new policies. In analyzing the changes made to the state accountability plans, we determined if these were uniform policy changes (i.e., resulting in changes in policy or new rules governing the implementation of NCLB and available to all states) or state-initiated policy changes (i.e., changes or waivers granted to some states through negotiations with ED). Finally, we examined

² The decision letters are available at <http://www.ed.gov/admins/lead/account/letters/index.html>.

³ The NCLB policy letters to states are available at <http://www.ed.gov/print/policy/elsec/guid/stateletters/index.html> and the key policy letters at <http://www.ed.gov/print/policy/gen/guid/secletter/index.html>.

whether these changes have a differential impact on schools and districts depending on the demographic make-up of their students and how the changes affected the meaning of accountability and the goals of NCLB.

EXAMINING THE FEDERAL ROLE IN EDUCATION

Federalism is often viewed from the perspective of the relationship between the national government and the states, with a focus on programmatic and fiscal relationships between federal and local authorities in various domains (Nathan, Gais, & Fossett, 2003). This perspective focuses on which level of government has policy, financial, and administrative responsibility over various governmental functions. Education is typically viewed as a state and local responsibility, particularly in areas related to core educational functions. NCLB has moved this relationship in the direction of expanding the federal role over states, while strengthening the role of states over local districts in education policy (Sunderman & Kim, 2004).

Much of the research on federalism has been on the relationship between levels of government in education without differentiating between the types of issues the federal government can most effectively address. On civil rights issues and access to education, the federal government played a critical and important role by challenging local practices of exclusion and discrimination, issues that states ignored or actively resisted. The federal government was instrumental in the desegregation of schools, for example, following the 1954 Supreme Court decision, *Brown v. Board of Education* outlawing state-mandated separate schools for black and white students. Title IX and the Individuals with Disabilities Act (IDEA) expanded access to educational opportunities for women and students with disabilities. With these laws, Congress defined the educational rights of students and expanded access to public education to previously excluded groups or students.

The federal government has also played an important role in fostering innovation through the use of grants-in-aid to states, providing resources for research and development, and in collecting and disseminating data. Title I of the Elementary and Secondary Education Act of 1965 (ESEA), Head Start, and other federal early education programs provided additional resources to states to develop programs that assist educationally and economically disadvantaged students. This legislation provided the impetus for states and private and public organizations to develop programs aimed at addressing the educational needs of disadvantaged students. While Head Start and Title I are often criticized for not meeting their lofty goals, they were a source of new programs and spurred research on curriculum and effective instructional practices. Congress has also funded educational research and development, although federal guidance in this area has been disappointing and lacked the commitment needed to develop a systematic and rigorous approach to supporting innovative research (Vinovskis, 1999). To provide the nation with data, the National Center for Education Statistics (NCES) collects, analyzes and disseminates statistics and information on education.

Beginning in the 1980s with the publication of *The Nation at Risk*, the federal role in education began to change as the federal government moved into regulating substantive educational issues. *The Nation at Risk* shifted the debate from a focus on educational access and equity to a concern with educational quality (Sunderman, 1995). Policies emphasized raising educational standards, increased student assessment, and implementing a core curriculum rather than access or programs to promote educational equity and innovation.⁴ America 2000 adopted under President Bush in 1989, Goals 2000 adopted during the Clinton administration in 1993 and the Improving America's School Act of 1994 all reflected this new agenda. Under NCLB, the federal government moved even further into regulating core educational functions by adopting policies that affect the development and implementation of curriculum, the delivery of instruction, how students are assessed, and the regulation and supervision of the teaching force.

Research on the implementation of federal policies suggests that a high degree of cooperation between the state and federal government would be needed to successfully implement the NCLB requirements. Earlier attempts to expand the federal role into controlling substantive educational issues met with limited success because of state resistance to an expanded federal role and inadequate funding to meet the federal goals (Superfine, 2005). Like NCLB, both Goals 2000 and the Improving America's Schools Act of 1994, the predecessor to NCLB, were built on bipartisan compromises that quickly dissipated once the laws were implemented. In contrast to these earlier federal laws that were loosely regulated in order to quell political opposition, the administration has strictly enforced the implementation of the NCLB requirements.

The success of the federal government in expanding civil rights and access to education has led some to believe that federal policy could be applied to improving the performance of schools. Supporters of the law argue that federal policy could provide the leverage to change how states allocate resources and create professional and political incentives for states to provide the support and technical assistance necessary to help low-performing schools and districts. However, changing how students learn and transforming educational practice are a fundamentally different order of complexity than enforcing civil rights. With civil rights issues, a court or administrative agency can order a school to accept a student. Research on school reform suggests that the active participation and consent of those involved (and required to change their behavior) is necessary before reform ideas take root or substantive changes in teaching practice are made (Elmore, 2004). Regulating core educational functions also raises questions about the feasibility of managing educational activities in 50 states, each with its own unique context. In the following sections, we show how the federal approach to implementing NCLB has evolved in response to political push back from the states and from errors in the design of a law that attempts to regulate core educational functions.

⁴ The Excellence Movement also promoted the application of market principles to education and school choice as remedies for educational problems.

FEDERAL “FLEXIBILITY” TO IMPLEMENT NCLB

During the first two years of NCLB, ED strictly interpreted and enforced the NCLB requirements, rebuffing any attempt to introduce policies that would respond to the concerns raised about parts of the law. Many of the early policy letters sent to states or chief state school officers reinforced a strict interpretation of the law and sought to clarify state responsibilities under the law, remind them of implementation deadlines, or provide additional guidance on various aspects of the law. For example, in a letter to the superintendent of public instruction in Indiana, Raymond Simon (then Assistant Secretary, Office of Elementary and Secondary Education) turned down the state’s attempt to identify schools districts for improvement on the basis of students in Title I schools. In declining to approve this request, Simon said, “We believe NCLB requires a State to consider the progress of students in all schools in a district” (U.S. Department of Education, 2004, October 4). Other letters responded to inquiries from members of Congress.

This changed beginning in late 2003 and early 2004 when ED announced new policies for the inclusion of students with disabilities and English language learners into the state accountability system and modified the highly qualified teachers provisions. Table 1 summarizes these and other rule changes announced by ED through December 2005. The first of these changes affected how students with disabilities were counted when determining adequate yearly progress (AYP). In response to objections to holding all students with disabilities to grade-level standards, ED announced a new regulation that allowed states to develop alternative achievement standards and use them to measure the progress of students with the most significant cognitive disabilities (U.S. Department of Education, 2003, December 11).⁵ The change allowed states to include disabled students who scored “proficient” or “advanced” on the alternative assessments as making AYP. While the law left it to states to define which students could be classified as having “significant cognitive disabilities,” it limited the number of scores based on alternative rather than grade-level standards that could be included in AYP calculations to 1% of all assessed students.

Table 1: Summary of Rule Changes to NCLB Allowed by U.S. Department of Education.

| Date of Change | Summary of Rule Change |
|----------------|--|
| 12-11-03 | <p><i>Inclusion of students with disabilities in assessment programs—1% rule.</i> Permits states to use alternate achievement standards to measure the progress of students with the most significant cognitive disabilities. In calculating AYP, states may count the “proficient” or “advanced” scores of students who take assessments based on alternate achievement standards so long as the number of these proficient and advanced scores does not exceed 1% of all students in the grades</p> |

⁵ The final regulations adopted on December 2, 2002 did not allow states to use alternative achievement standards for students with the most significant cognitive disabilities when determining AYP (U.S. Department of Education, 2002).

| | |
|----------------------------------|--|
| <p>2-20-04</p> | <p>tested at the state and district level.</p> <p><i>Flexibility regarding English language learners.</i></p> <p>a. <i>Allow LEP students a one-year exemption from the reading content assessment:</i> States are allowed to exempt LEP students from taking the reading content assessment during the first year they are enrolled in U.S. schools. LEP students must still take the English language proficiency assessment as well as the mathematics assessment, with accommodations as appropriate. States have the flexibility to decide whether or not to include results from the mathematics assessments and the reading content assessments (if given) when calculating AYP. LEP students are included in the 95% participation requirement needed to achieve AYP.</p> <p>b. <i>Allow greater flexibility in LEP subgroup definition:</i> States are allowed to include former LEP students in the LEP subgroup for AYP calculations up to two years after they have achieved English proficiency (U.S. Department of Education, 2004, February 19).</p> |
| <p>3-15-04 & 3-31-04</p> | <p><i>Highly Qualified Teacher requirements.</i></p> <p>a. <i>Flexibility for Rural Schools:</i> Gives local educational agencies that are eligible to participate in the Small Rural School Achievement Program (SRSA) additional time for teachers who teach multiple subjects and are highly qualified in one subject to become highly qualified in the additional subjects. Veteran teachers have an extra year (until 2006-07) and new teachers have three years from the date of hire to gain certification in the additional subjects.</p> <p>b. <i>Flexibility for Veteran Teachers of Multiple Subjects:</i> Allows states to streamline their HOUSSE procedures by developing a process for current, multi-disciplinary teachers to demonstrate subject matter mastery in each of their subjects through a single set of procedures.</p> <p>c. <i>Flexibility for Science Teachers:</i> Allows states to rely on their own teacher certification requirements for science to determine areas in which teachers must have subject matter knowledge in order to be considered highly qualified.</p> |
| <p>3-29-04</p> | <p><i>Calculating participation rates for schools and/or subgroups.</i></p> <p>a. <i>Significant Medical Emergency:</i> States do not have to include a student with a documented significant medical emergency in the participation rate (states define what constitutes a “significant medical emergency”).</p> |

| | |
|---------------------------|--|
| <p>4-7-05 5-10-05</p> | <p><i>b. Uniform Averaging:</i> States can use test data from the previous one or two years to average the participation rate for a school and/or subgroup. If the two- or three-year average meets or exceeds 95%, the school will have met the AYP participation requirement.</p> <p><i>Raising Achievement: A New Path for No Child Left Behind.</i> Provides states “new flexibility” to amend their accountability plans or consolidated applications if they meet certain requirements. Specific areas where this will apply have not been determined. To be eligible for this new flexibility, states must:</p> <ul style="list-style-type: none"> <i>a. Provide evidence of raising achievement and closing the achievement gap.</i> <i>b. Demonstrate rigorous implementation of strong standards, assessment, and accountability systems.</i> <i>c. Demonstrate evidence that information is accessible and that school districts are implementing supplemental services in good faith and expanding the capacity for public school choice.</i> <i>d. Implement a rigorous system to ensure teachers are highly qualified.</i> |
| <p>5-10-05</p> | <p><i>Accountability for students with academic disabilities—2% rule.</i></p> <ul style="list-style-type: none"> <i>a. Accountability for Students with Academic Disabilities:</i> Allows states to develop modified assessments that are based on modified achievement standards for a limited group of students with academic disabilities. Up to 2% of the proficient scores of students with academic disabilities who took the modified assessments can be included when determining AYP. <i>b. Interim Flexibility: Accountability Plan Amendments for 2004-05.</i> In determining accountability for 2005-06 (based on assessments administered in 2004-05) only, states were given two interim policy options for calculating AYP for students with disabilities. <p><i>Transition Option #1:</i> Under this option, eligible states can calculate a proxy to determine the percentage of students with disabilities that is equivalent to 2% of all students assessed and add this group of students to the percent of students with disabilities that are proficient. This option applies only to schools and districts that did not make AYP solely based on the scores of the students with disabilities subgroup.</p> <p><i>Transition Option #2:</i> States that meet certain requirements may count the proficient scores of students with disabilities that are based on modified achievement standards when calculating AYP, subject to a 2% cap. This option is only available to those states with modified achievement standards and assessments.</p> |

10-21-05

Meeting the Highly Qualified Teacher requirements.

States that do not quite reach the 100% goal by the end of the 2005-06 school year will not lose federal funds if they are implementing the law and making a good-faith effort to reach the 100% goal, subject to approval of a revised plan. States must show a “good faith effort” by meeting four requirements.

- a. *Have a definition of a highly qualified teacher that is consistent with the law.*
- b. *Provide accurate and complete reports on the number and percentages of highly qualified teachers.*
- c. *Highly qualified teacher data reported to ED must be complete and accurate.*
- d. *Take action to ensure that experienced and qualified teachers are equitably distributed.*

11-21-05

Growth Model Pilot Project.

States may propose a growth-model accountability model for making AYP determinations. ED will approve models for up to 10 states to participate in a pilot program for 2005-06. To be approved, the growth model proposal must show how it will meet a set of core principles outlined by ED that include:

- a. *The accountability model must insure all students are proficient by 2013-14 and include annual goals.*
- b. *The accountability model must establish high expectations for low-achieving students but cannot set expectations for annual achievement based on student demographic or school characteristics.*
- c. *The accountability model must set separate achievement goals for reading/language arts and mathematics.*
- d. *The accountability model must include all students in tested grades; the performance of student subgroups; and all schools and districts.*
- e. *The assessment system must include annual assessments in grades 3-8 and high school in reading/language arts and mathematics; been operational for more than one year; and produce comparable results from grade to grade and year to year.*
- f. *The accountability model and state data system must track student progress.*
- g. *The accountability model must include student participation rates and student achievement on an additional academic indicator.*

Source: US Department of Education. Policy letters to states are available at <http://www.ed.gov/print/policy/elsec/guid/stateletters/index.html> and the key policy letters at <http://www.ed.gov/print/policy/gen/guid/secletter/index.html>.

This change was followed by a second announcement in February 2004 that modified accountability requirements for students with limited English proficiency (LEP). Researchers had identified a number of challenges to implementing the NCLB requirements for LEP students, including the instability of the LEP subgroup, the failure of standardized test scores to accurately reflect what LEP students understand, and the lack of proven accommodations that would make these scores more reliable, among others (Abedi, 2004; Batt, Kim, & Sunderman, 2005; Coltrane, 2002, November). State and local education officials questioned the fairness of the provisions because students who achieved English proficiency are generally moved out of the subgroup while new students with very low levels of English proficiency are continually added to the subgroup. This greatly diminishes the chances that schools serving large numbers of LEP students will be able to improve the performance of this subgroup and make AYP. In addition, states were finding that schools reporting an LEP subgroup were more likely to be identified as needing improvement than those without this subgroup, an issue with the students with disabilities subgroup as well (Sunderman, Kim, & Orfield, 2005).

The new policies allowed states to exempt LEP students from taking the reading content assessment during the first year they enrolled in U.S. schools and gave states the flexibility to decide whether or not to include results from the mathematics assessments and the reading content assessments (if given) when calculating AYP. To accommodate for the changing composition of the LEP subgroup, states were allowed to continue to count LEP students for up to two years after they exited the program when calculating AYP.

The third set of changes was in response to the difficulties states and districts encountered in implementing the highly qualified teacher provisions. These provisions require states and district receiving Title I funds to develop a plan to “ensure that all teachers teaching in core academic subjects . . . are highly qualified not later than the end of the 2005-06 school year (NCLB, 2002, §6319 (a)(2)).⁶ The statute defines highly qualified as a teacher who has obtained a full state certification, has a bachelor’s degree, and has demonstrated subject matter competency in core academic subjects (NCLB, 2002, §7801(23)).

States encountered a number of technical challenges in implementing the highly qualified teacher requirements, including the lack of information and data systems needed to determine whether teachers were highly qualified, difficulties attracting teachers to high poverty, urban and rural areas, and budget constraints that limited their ability to invest more in teacher salaries or incentives (General Accounting Office, 2002, 2003; Sunderman & Kim, 2005). State studies on teacher supply and demand reported shortages in mathematics, special education, science, and foreign languages, with rural and central city areas experiencing the most severe shortages (Georgia Professional

⁶ The law defines core academic subjects as English, reading or language arts, mathematics, science, foreign languages, civics and government, economics, arts, history, and geography (NCLB, 2002, §7801(11)). A teacher who teaches a subject or grade level for which they are not highly qualified is considered an “out-of-field” teacher (NCLB, 2002, §6602(5)).

Standards Commission, 2002; Illinois State Board of Education, 2002; Virginia Department of Education, 2000, November).

Among the requirements, demonstrating subject matter competency presented the most challenges, particularly in regions of the country or schools where teachers taught more than one subject or were in short supply. Rural states were concerned about their ability to ensure a highly qualified teacher in every core subject, particularly in smaller schools where a teacher may teach multiple courses and multiple grades (Jimerson, 2004, January; Richard, 2003a, 2003b). Meeting the highly qualified requirements in middle and high schools was complicated because many teachers would be required to demonstrate subject matter competency in several subjects. Since some states allowed middle school teachers to teach with elementary certification, which requires less evidence of subject matter knowledge than secondary certification, there was concern that teachers would leave middle schools rather than gain the additional requirements. In addition, middle school teachers often are responsible for teaching more than one subject, something that the organization of middle schools, which encourages team teaching, promotes. At the high school level, teachers may teach subjects where they do not have subject matter expertise, in part because of assignment practices which create out-of-field teaching (Ingersoll, 2001) as well as the shortage of teachers with subject matter expertise. Special education teachers, who are generally certified in special education or their specialty (for example, speech therapy), are also required to meet the subject matter requirements, a challenge when they teach multiple subjects. The combination of teacher shortages in some subject areas with the increased requirements threatened to worsen the situation in rural and central city districts, particularly at the middle and high school levels.

ED initially responded to these issues by encouraging states to expand alternative routes to teacher licensure, enhance teacher preparation and recruitment programs, and emphasize professional development to improve teacher knowledge and instructional practices (U.S. Department of Education, 2004, January 16). In March 2004 in response to continuing public concern about some of the teacher provisions, the administration relaxed the regulations governing the teacher quality provisions of the law in three ways. First, veteran teachers in rural districts who were certified in one subject were allotted an additional year (until 2006-07 rather than 2005-06) to demonstrate subject matter competency in the additional subjects, and new teachers that were certified in one subject were given three years from the date of hire to gain certification in the other subjects that they taught.⁷ Second, science teachers could be certified across science-related subjects (i.e., biology, chemistry, etc.) instead of needing to be certified in each science subject area. Third, teachers who taught multiple subjects would be allowed to prove mastery by passing a test that covered multiple subjects. In the past, they had been required to pass a test in each subject. Though rural areas were targeted for increased flexibility in implementing the teacher quality provisions by such changes, no such relief was directly

⁷ New teachers in other districts must meet the highly qualified requirements when they are hired. Only teachers that are participating in an alternative route to certification program can be hired without being fully certified and still be considered as meeting the certification requirements (U.S. Department of Education, 2005, August 3)

targeted at the challenges facing urban school districts as they attempted to comply with the law.

Another area where ED made a policy change was in the procedures for determining the participation rate. Under NCLB, 95% of students in each subgroup and at each grade level must take the state's assessments in order to make AYP. Since a school could meet its performance targets and still fail to make AYP if it did not meet the participation requirements for any of the subgroups, many schools, including those that were recognized locally as successful schools, failed to make AYP. In response, ED allowed states to average participation rates for schools or a subgroup over three years and to excuse those students with a "significant medical emergency" from being counted when calculating AYP (U.S. Department of Education, 2005, March 29). While this change eased the burden for a particular year, it still meant schools that fell below the 95% participation rate would have to exceed 95% in other years, thus providing only temporary relief.

A 'New Path' for NCLB

The next major set of changes came in April 2005 when Secretary Spellings announced the "Raising Achievement: A New Path for *No Child Left Behind*" initiative. By the time of this announcement, opposition to the law in the states had reached a crescendo. Hopes were quelled that the state opposition that existed during the 2004 state legislative session would die down. Instead, states continued to introduce legislation or resolutions that placed restrictions on the implementation of NCLB. This opposition cut across the political and ideological spectrum, with some of the strongest opposition coming from Republican states. For example, the Utah legislature passed a bill in April 2005, which the governor signed, that required state and local education officials to give first priority to meeting Utah's educational goals when they conflict with NCLB and to minimize the amount of additional state resources used to meet the NCLB requirements ("Implementing federal educational programs," 2005; Lynn, 2005). This followed 15 months of negotiations between ED and Utah officials and threats from ED to withhold the \$107 million that Utah received in federal education funds (Davis, 2005; Dillon, 2005). While the state backed away from earlier legislation to opt out of the law altogether, it remained among the strongest anti-NCLB legislation in the country.

Other states continued to oppose the federal law as well, with one source identifying 47 states that had taken action to amend or fix the law or limit its impact in their state.⁸ The Texas Education Agency (TEA) continued to follow state law when determining AYP by allowing districts to follow state rather than federal rules for counting the test scores of students with disabilities (Hoff, 2005). TEA took this action after ED had changed the regulations on testing special education students, which allow the state to exempt 9% of its special education population from grade level standards compared to 3% under the federal rules (Marks, 2005). By following the Texas rule

⁸ This information is available from Communities for Quality Education, a national advocacy organization, at <http://www.qualityednow.org/reports/revolt/index.php>. The website covers actions states had taken through August 2005 (site last visited on October 3, 2005).

rather than the federal rule on counting special education students for AYP purposes, the state reduced the number of schools and districts identified for improvement.⁹ When Connecticut could not resolve a dispute with ED officials over testing every student every year, the state followed through with its threat to sue the U.S. Department of Education. In a bipartisan review of the law, the National Conference of State Legislatures (NCSL) issued a consensus report on NCLB in February 2005 that was highly critical of the federal law. Further highlighting the limited support from state legislators, this report was critical of the expanded role of the federal government in education and what the report called “a coercive relationship between states and the federal government” (National Conference of State Legislatures, 2005).

The administration’s approach to working with state officials failed to generate the cooperation needed to fully implement the law. Continuing the negotiating style and rhetoric of her predecessor, Secretary Spellings threatened to withhold federal money, accused state officials of dodging the law’s requirements, or questioned the integrity of state officials when they disagreed with federal officials. For example, Secretary Spellings criticized Connecticut officials’ position on testing as “un-American” and accused Connecticut of shirking its responsibility to low-income and minority students on national television (Frahm, 2005). Connecticut Commissioner of Education Betty J. Sternberg responded in a letter to Spellings, calling her accusations inaccurate and seeking an apology (Sternberg, 2005). Utah legislators were also offended by Spellings’ response to the Utah law. In a letter to Senator Orin Hatch, Spellings wrote, “Several of the principles in the bill are fundamentally troublesome, and appear to be designed to provoke noncompliance with federal law and needless confrontation” (Sack, 2005; Spellings, 2005, April 18).

Amid the continuing criticism of the law, Spellings announced a new plan for giving states additional flexibility in implementing the law. At the April 2005 event announcing the policy, Spellings stated that “these past three years have helped us be smarter about how this law is working in the schools” and outlined a structure where “states that show results and follow the principles of *No Child Left Behind* will be eligible for new tools to help them meet the law’s goals” (Spellings, 2005, April 7). Under these provisions, ED will give states additional flexibility in meeting the federal mandates if they meet certain requirements. In her remarks, Spellings laid out the principles of the new plan – saying that states with “sound education policies” and those that closely followed the tenants of NCLB would be provided these ‘new tools.’ ED would develop these principles further in later public announcements.

This new plan was two-fold. First, it outlined a set of guiding principles that ED would take into account, along with states’ own unique situation, when considering amendments to state accountability plans or consolidated applications. These principles included the following:

⁹ Under the state rule, 86 districts did not make adequate progress whereas 517 would have been identified for improvement using the stricter definition and 402 schools did not make AYP compared with 1,718 that would not have made AYP without the rule change (Hoff, 2005).

1. ***Ensuring students are learning.*** States must provide evidence of progress in raising overall achievement and closing the achievement gap, including improvements in AYP results or significant improvements in overall achievement trends on state tests or on NAEP.
2. ***Making the school system accountable.*** This principle includes providing the “tools” for demonstrating improvements in student achievement. To meet this principle, states must demonstrate evidence of rigorous implementation of strong standards, assessment, and accountability systems.
3. ***Ensuring information is accessible and options are available.*** This principle highlights access to information and the implementation of public school choice and supplemental educational services. States seeking flexibility must demonstrate evidence of a coherent and easy-to-understand explanation of accountability and provide report cards that are readily available. In addition, school districts must be implementing supplemental services in good faith, and the state must be working with districts to expand capacity for public school choice.
4. ***Improving the quality of teachers.*** States must demonstrate that they are implementing a rigorous system for ensuring teachers are highly qualified (U.S. Department of Education, 2005, April 7).

In determining whether to approve an amendment request, “Department staff will consider existing State data to determine how the State meets the guiding principles under its current system of accountability and how the allowance of additional flexibility will contribute to the State’s reaching its goals and increasing student achievement” (U.S. Department of Education, 2005, May 10a). Absent from the announcement were the areas where ED might consider granting additional flexibility or how officials would judge the data presented by a state. Apparently, ED intended to announce at a later date the areas where states can apply for “additional flexibility.”

Second, the Secretary announced intentions for another change in the rule for counting students with disabilities for AYP purposes, this time for “students with academic difficulties” (U.S. Department of Education, 2005, May 10b).¹⁰ This rule would allow states to develop modified achievement standards and assessments aligned to those standards for students with academic difficulties. The state could include the proficient scores from those alternative assessments when determining AYP for up to 2% of all students tested. Combined with the earlier change for students with “significant cognitive disabilities,” this rule brought to 3% the proportion of students that could take alternative assessments and have their scores counted when calculating AYP.

Until ED completed the process for developing the new rule, states were given “interim flexibility” for counting students with disabilities when determining AYP for the 2005-06 school year (based on 2004-05 assessments). There were two options. The first

¹⁰Students with academic difficulties are defined in the Individuals with Disabilities Education Act (IDEA).

option—called the proxy method—applied “only to schools or districts that did not make AYP based solely on the scores of students with disabilities” (U.S. Department of Education, 2005, May 10b). It allowed states to make a mathematical adjustment in the way the achievement scores for the students with disabilities subgroup were considered for AYP determinations. Following the approach introduced in the earlier “New Path” announcement, only states that met certain “eligibility guidelines” could apply for this option. The second option applied to states with modified achievement standards and assessments. These states could count the proficient scores of students with disabilities that were based on modified achievement standards when calculating AYP, subject to a 2% cap.

Highly Qualified Teacher Requirements. Following the approach outlined in the New Path initiative, Secretary Spellings announced a policy change in October 2005 to allow states additional time to meet the highly qualified teacher requirements if they can demonstrate they are “implementing the law and making a good-faith effort to reach the HQT goal in NCLB as soon as possible” (Spellings, 2005, October 21). In the letter, Spelling said states that do not have highly qualified teachers in every classroom by the end of the 2005-06 school year can submit a revised plan for meeting these requirements by the end of the 2006-07 school year. ED’s approval of the revised plan is contingent on states demonstrating that they are making “a good-faith effort” to meet the law’s requirements. To do that, they must:

1. Have a definition of a highly qualified teacher that is consistent with the law.
2. Provide accurate and complete reports on the number and percentages of highly qualified teachers.
3. Report data to ED on the number of highly qualified teachers that is complete and accurate.
4. Take action to ensure that experienced and qualified teachers are equitably distributed.

In approving the revised plan, ED officials indicated that they would also consider states’ efforts to recruit, retain, and improve the quality of the teaching force.

In contrast to the earlier policy change, which relaxed some of the rules governing the highly qualified teacher provisions, this rule change acknowledged the broader issue of the difficulties states have in meeting the requirements and was a way for the Secretary to avoid withholding Title I funds from states not meeting them. By delaying the implementation deadline, this change also recognized the complexity of the issues surrounding the implementation of the teacher quality provisions without specifically addressing them or helping states figure out how to meet these challenges. Challenges ranged from having adequate data systems to track teacher qualifications to how to determine if veteran teachers are highly qualified. While states report that the majority of their teachers are highly qualified, there is wide variation among states and few that have reached the 100% mark. Self reported baseline data (for 2003) from the states indicate that the percentage of classes taught by highly qualified teachers ranged from 16% in

Alaska to 98.6% in Wisconsin (The Education Trust, 2003).¹¹ However, a number of analysts questioned the reliability of the state data, particularly when it is compared with national data (Center on Education Policy, 2004; The Education Trust, 2003). For example, the Secretary's second annual report on teacher quality found that only 54% of the nation's secondary teachers were highly qualified during the 1999-2000 school year (U.S. Department of Education, 2003).¹² States also have difficulty applying the requirements to special categories of teachers, including special education, and middle and high school teachers, ensuring that districts serving large percentages of low-income and minority students have highly qualified teachers, and meeting the teacher requirements in rural districts. Finally, there are questions about whether the NCLB requirements are the most important ones and concern that they ignore issues specific to a particular state and district (Sunderman & Kim, 2005).

As the deadline for having all teachers highly qualified approached (2005-06), it was becoming clear that many states would not reach the 100% goal. As shown in Table 2, variation among six states (representing the states participating in this study) in the percentage of classes taught by highly qualified teachers remained, especially in insuring that low-income students had access to highly qualified teachers. While states have reported improvements in the percentage of highly qualified teachers, because states lacked adequate data collection systems, it is hard to know what the improved numbers mean (Keller, 2005). The Government Accountability Office (GAO) reported that states were making progress in their ability to track and report the percentage of core academic classes taught by teachers meeting the NCLB requirements, but that there were several issues that limited the quality and precision of this data (Government Accountability Office, 2005). Among these were a reliance on district reported data, which varied by district in both the data collection processes used and quality of the data. Some states did not include all teachers in their calculations and the rigor of the teacher requirements varied across states, making it difficult to compare states or to be sure how meeting these requirements would affect the quality of instruction or student performance.

¹¹ States were required to submit baseline data on the percentage of classrooms taught by highly qualified teachers by September 1, 2003.

¹² This analysis was based on data from the Schools and Staffing Survey.

Table 2: Percentage of Core Classes Taught by Highly Qualified Teachers in all Schools in Six States, and in High- and Low-Poverty Schools, 2002-03 and 2003-04.

| State | Statewide | | High-Poverty Schools | | Low-Poverty Schools | |
|------------|-----------|---------|----------------------|---------|---------------------|---------|
| | 2002-03 | 2003-04 | 2002-03 | 2003-04 | 2002-03 | 2003-04 |
| Arizona | 84.0* | 96.0 | N/A | 96.0 | N/A | 96.0 |
| California | 48.0 | 52.0 | 35.0 | 40.0 | 53.0 | 60.0 |
| Georgia | 94.0* | 97.0 | N/A | 97.5 | N/A | 97.2 |
| Illinois | 97.9 | 98.2 | 94.6 | 93.4 | 99.5 | 99.7 |
| New York | N/A | 93.8 | N/A | 88.5 | N/A | 97.7 |
| Virginia | 85.5 | 94.5 | 77.1 | 92.2 | 87.4 | 96.5 |

Source: 2002-03: AZ: (The Education Trust, 2003).

CA: 2002-03 California's Consolidated State Performance Report,

<http://goldmine.cde.ca.gov/nclb/sr/rt/documents/cspr.pdf>;

<http://www.cde.ca.gov/nclb/sr/sa/documents/vr04csa0901.pdf>

GA: http://www.gapsc.com/nclb/Admin/admin_fedreport.html

IL: 2003 Illinois State Report Card and 2003 Illinois District (Chicago) Report Card;

http://206.230.157.60/publicsite/reports/2003/state/English/2003_StateReport_E.pdf

NY: (The Education Trust, 2003).

VA: Professional Qualifications of Teachers, 2002-03, <http://www.pen.k12.va.us/VDOE/src/vasrc-pqt.pdf>

2003-04: AZ: (Government Accountability Office, 2005)

CA: 2003-2004 California's Consolidated State Performance Report: Part I <http://www.cde.ca.gov/nclb/sr/rt/index.asp>

GA: 2003-2004 Georgia's Annual Report Card:

<http://reportcard.gaosa.org/vr2004/k12/accountability.aspX?ID=ALL:ALL&TestKey=HQ&TestType=acct>

IL: 2004 Illinois State Report Card: <http://webprod1.isbe.net/ereportcard/publicsite/getSearchCriteria.aspx>, State Report Card FY 2004.

NY: Pat O'Brien, New York State Education Department.

VA: The Virginia School Report Card 2005 <http://www.pen.k12.va.us/VDOE/src/index.shtml>

* Reflects the percentage of teachers statewide that met state and federal criteria required to be considered highly qualified rather than the percentage of classes taught by highly qualified teachers.

Note: High poverty schools are those in the bottom quartile; low poverty schools are those in the top quartile.

Growth Model Pilot Program: In a November 2005 announcement, Secretary Spellings said that states “that follow the bright-line principles of *No Child Left Behind*” can submit proposals to implement a growth model of accountability for the 2005-06 school year (Spellings, 2005, November 18). The Spelling’s announcement called it a pilot program, which ED will evaluate, and limited participation to 10 states. This announcement followed months of speculation about the possibility of allowing states to adopt growth models and considerable, long-term support among education advocates to incorporate value-added or growth models into state accountability systems (Olson, 2004, 2005b). States wanting to participate in this program were required to submit an application by February 2006. To be considered, states must show how their proposal meets a set of core principles laid out by ED. Similar to earlier policy changes, these principles reinforced the importance of accountability as a model for improving student achievement, the use of a timeline for showing that all students are proficient and included an emphasis on assessing all students in tested grades and all subgroups.

State Initiated Policy Changes

In addition to the policy changes outlined above, another process was taking place that produced “state initiated” policy changes. State accountability plans have undergone numerous revisions since they were first submitted in January 2003. States revised their plans in response to reviews by ED that found the plans out of compliance with the law or made adjustments to reflect changes in regulations. In addition, beginning in 2004 states could submit a request to amend their plans for consideration by ED. However, there were no guidelines on the kind of changes states could request or how these requests would be judged and no guarantees that changes approved in one state would be approved in another. Since these amendments were the result of negotiations between individual states and ED, learning about them in other states depended on informal communication among states. The first time ED officially addressed any of these changes was in a draft document released in November 2005, a full two years after the amendment process began (U.S. Department of Education, 2005, November 10).¹³ It is worth noting that many of the early requests for changes from states that were consistently rejected were later approved or the guidance governing them was changed (Erpenbach & Forte, 2005, October). The state initiated amendments, approved by ED in 2004 and 2005, are summarized in Table 3.

Table 3: State Initiated Policy Changes: State Accountability Plan Amendments Approved by U.S. Department of Education.

Summary of Amendment

School level identification; same subject. To be identified for improvement, schools must fail to make AYP for two consecutive years in the same subject—i.e., two years in math or two years in reading.

Grade span method of calculating district accountability. Allows states to aggregate district level test score data by grade span (i.e., grades 3-5, 6-8, 9-12) rather than by individual grades for accountability purposes. To be identified for improvement, districts must fail to make AYP in the same subject and for all grade spans for two years.

Grade span method of calculating state accountability. Allows states to aggregate state level test score data by grade span (i.e., grades 3-5, 6-8, 9-12) rather than by individual grades for accountability purposes. The state is identified for improvement if it fails to make AYP in the same subject and for all grade spans for two years.

Minimum group size. Allows states to change the minimum group size required for reporting subgroup scores for accountability purposes.

- a. ***Minimum group size for reporting on subgroups.*** States can change the minimum group size required for reporting subgroup scores for accountability purposes. Changes include increasing the minimum group size or using proportional minimum ‘n’ models (i.e., adopting a group size that is proportional to the size of
-

¹³ This document was labeled “pre-publication copy.” It is unclear how widely it was disseminated.

-
- the student population).
- b. *Minimum group size for Limited English Proficient (LEP).* States can use a different group size for reporting on the LEP subgroup.
 - c. *Minimum group size for students with disabilities (SWD).* States can use a different group size for reporting on the SWD subgroup.

Statistical changes in how data is used.

- a. *Use of confidence intervals when calculating AYP.* States can use a “confidence interval” when making AYP determinations, which allows states to create a band around the test scores of any subgroup or group of students. Confidence intervals were included in the original accountability workbooks of some states. Most states sought to add the use of a confidence interval to their plan or to increase the size of the confidence interval from 95% to 99%. Confidence intervals can be applied to all AYP indicators.
- b. *Use of confidence intervals when calculating safe harbor.* States can apply a confidence interval when making safe harbor determinations. States may use a 75% interval.
- c. *Use of standard error of measurement when calculating student test scores.* States can add a standard error of measurement when calculating individual student tests scores.
- d. *Uniform averaging applied to AYP.* Allows states to apply uniform averaging to one or more AYP indicators (i.e., test scores, graduation rates, attendance rates). States can average the two or three most recent years of data (current year plus one or two previous years) and compare that to the current year data. States can use the higher of the two to determine AYP status for 2006-07. For 2005-06, states can average two years of data.
- e. *Define how rounding rules will be used.* Use standard rounding rules when making AYP decisions.

Revised/implemented an Index system.

Allows states to use an index to determine AYP in lieu of calculating the percent of students scoring proficient on state assessments. Some models weight performance by achievement levels (for example, extra weight is given for performance at higher achievement levels than at lower achievement levels) as long as the performance at the higher levels does not compensate for performance at the lower levels. Other models adjust for differences in the number of students at each grade level.

Definition of AYP.

- a. *Changed definition of AYP.* Florida added a new designation to the definition of AYP, “Provisional AYP” for those schools that received an A or B school letter grade under the state accountability plan but did not make AYP.
 - b. *Changed intermediate goals.* States were allowed to change the starting point or intermediate goals along the timeline towards 100% proficiency. Some states did this because they adopted new assessment systems while others adopted equal increments (AMOs that increase every year) or equal stair-step increments.
 - c. *Changed graduation rate/attendance rate/other indicator.* Allows states to
-

change or revise goals for graduation rate, attendance rate, and/or the additional indicator required for determining AYP. These changes allow schools/districts to make progress toward a specific target rather than meeting or exceeding the target.

- i. *Goals/standards for LEP students.* Allows states to establish separate goals for the LEP subgroup for graduation rate, attendance, and/or the additional indicator when determining AYP. This includes allowing LEP students extra time to graduate.
- ii. *Goals/standards for students with disabilities.* Allows states to establish separate goals for the SWD subgroup for graduation rate, attendance, and/or the additional indicator when determining AYP. This may include, for example, using a graduation rate established by a student's IEP or allowing SWD extra time to graduate.

Changes in assessment system.

- a. *State level changes.* Made changes to state assessment system used for NCLB accountability or adopted a new state test.
- b. *Delayed implementing 3-8 assessment system.*
- c. *Test retakes.* Allowed students to take NCLB assessments multiple times (with highest achievement levels used for accountability purposes).
- d. *Testing out of grade level.* Clarified how states use test results for students in advanced courses taking out of grade level assessments. For example, middle school students who take a middle school math course and the high school math test can bank their test results until they are enrolled in high school.

Including all schools in accountability system.

- a. *Determining accountability for small schools and schools without tested grades.* States revised the process for determining AYP decisions for small schools and/or schools without tested grades.
- b. *Defined/clarified how schools not subject to NCLB will be included in accountability system.* States defined or clarified the process for including schools not subject to NCLB (for example, alternative schools, psycho-educational centers, kindergarten only schools, new schools or recently reconstituted schools) in the accountability system.
- c. *Define/clarified how states ensure all schools are included in accountability decisions.*
- d. *Changed/clarified AYP for targeted assistance schools (TAS).*

Interventions in schools/districts.

- a. *Established additional interventions for SWD and LEP students.* Developed a system to offer additional monitoring and/or support services for SWD and LEP students.
 - b. *Clarified plans for intervening in under-performing schools/districts.* States clarified how they would intervene in poorly performing schools and districts, including prioritizing interventions to districts based on the extent to which the district did not make AYP or providing direct services or intervention to schools
-

on the bases of the elements that were involved in identifying a school for improvement.

- c. *Implemented or revised award (or sanction) system for schools/districts.*

Reporting related changes.

- a. *Incorporated/aligned aspects of state system into AYP plan.*
b. *Changed accountability plan to accommodate SWD.*
c. *Changed/revised process for appealing AYP determinations.*
d. *Defined new timeline for identifying schools/districts.*
e. *Changed/clarified reporting requirements and timeline for releasing report cards.*
f. *Defined “full academic year” for use in AYP determinations.*
g. *Defined subgroups that are required to report for AYP determinations.*
h. *Redefined/clarified how participation rate is calculated.*
-

State-Level Waivers: Unlike the implicit policy changes discussed above, which theoretically all states could apply if they knew about them, there was another process going on where deals were negotiated with a very few individual states and districts. As mentioned earlier, Texas education officials decided to use their own rule, rather than follow the federal guidelines, for counting special education students for AYP purposes. These rules allowed the state to exempt 9% of its special education population from grade level standards compared to 3% under the federal rules (Marks, 2005). In adopting this rule, Texas educational officials said the state has the flexibility under NCLB to bypass the federal rules and tied this flexibility to state authority to consider school district appeals over their AYP status (Stutz, 2005). This is the only state to apply their own rules to counting special education students, and they did so with impunity for two years (2003-04 and 2004-05). In July 2005, the Texas Education Agency negotiated a deal with ED to limit the percent of special education students who take alternative assessment and are counted for accountability purpose to 5% for the 2004-05 school year. For 2005-06, Texas must meet the 3% cap and may be fined for not using the lower cap in 2003-04 (Grant, 2005).

District-Level Waivers: In August, ED extended waivers to a handful of districts that allowed them to change how they implemented the supplemental educational services program (Table 4). These waivers were not made available to all districts. The first beneficiaries of these waivers were four school districts in Virginia—Alexandria, Henry County, Newport News, and Stafford County. This waiver, granted in August 2005, allowed these four districts to offer students supplemental educational services before requiring them to offer students the option to transfer to another school. Under NCLB, districts are required to offer transfers to all students in any school that is in its first year of school improvement and to provide supplemental educational services to students in a school that is in its second year of school improvement. District officials had long advocated reversing the order of these sanctions, arguing that it makes more sense to attempt to improve a school by offering the tutoring program before implementing the more disruptive transfer option.

Table 4: Summary of District Waivers Granted by U.S. Department of Education.

| Date of Waiver | Summary of Waiver |
|-----------------------|---|
| 8-27-05 | Reverse order of sanctions: Four Virginia school districts were given a waiver to offer supplemental educational services before requiring transfers (Alexandria, Henry County, Newport News, and Stafford County). |
| 9-1-05 | Supplemental educational service providers: Allowed the Chicago Public Schools to be a provider under the NCLB supplemental educational services program even though the district had been identified as needing improvement. The waiver was for one year. This waiver was later extended to include New York and Boston in separate agreements. |

The second waiver was first granted to the Chicago Public Schools. Following months of conflict between the Chicago Public Schools and ED officials over the provision of supplemental services by the district, Secretary Spelling announced a waiver that would allow the district to continue to be a supplemental service provider even though it had been identified for improvement (Banchemo, 2005; Gewertz, 2005c). The waiver was for one year. This dispute began in the 2004-05 school year when the district was first identified as needing improvement. ED officials told the district they could no longer be a service provider under NCLB, a stipulation in the law. The district, which had provided services to 79% of the students participating in the program in 2003-04 (Sunderman et al., 2005),¹⁴ continued to offer services, arguing that stopping services would disrupt tutoring for tens of thousands of students (Gewertz, 2005a). The district reached an agreement with ED in February 2005 that would allow them to continue tutoring students for the remainder of the school year, but they would have to use money other than the NCLB set-aside for the supplemental services program (Gewertz, 2005b). The status of the program for the 2005-06 school year remained unclear until the one-year waiver was granted in September 2005. In November 2005, the Boston Public Schools and the New York City Public Schools received a similar waiver—called flexibility agreements—in exchange for greater oversight by the Education Department (Lindsay, 2005, November 3; Saulny, 2005).

TRACKING CHANGES TO STATE ACCOUNTABILITY PLANS

Table 5 shows the number of states taking advantage of the rule or policy changes adopted by ED through December 2005. Many, but not all states have amended their accountability plans to incorporate these changes. Table 5 shows that 31 states changed their method of counting students with the most significant cognitive disabilities for

¹⁴ Of the 64,500 students served in the 2003-04 school year, 49,607 (76.9%) received services from the district and 14,893 (23.1%) received services from private providers. In the 2004-05 school year, of the 61,466 students served, 30,803 (50.1%) received services from the district and 30,663 (49.9%) from private providers.

purposes of calculating AYP by adopting the 1% rule. The most recent change, which allowed states to mathematically adjust the percentage of students in the students with disability subgroup for the 2004-05 school year (“interim flexibility”), was approved in 28 states. There were 43 states that took advantage of the additional flexibility offered for testing LEP students and 41 states that changed how they calculated participation rates for subgroups by adopting uniform averaging.

Table 5: States Taking Advantage of Policy Changes Adopted by U.S. Department of Education through December 2005.

| State | Alt. Assessments/ SWD 1% Rule | Alt. Assessments/ SWD 2% Rule | Assessing LEP's | Participation Rate-Subgr. |
|----------------|----------------------------------|----------------------------------|--------------------|------------------------------|
| Alabama | x | | x | x |
| Alaska | x | | x | x |
| Arizona | x | x | x | x |
| Arkansas | x | | x | x |
| California | x | | x | x |
| Colorado | x | | | |
| Connecticut | | | x | x |
| Delaware | x | x | x | x |
| Florida | | x | | |
| Georgia | x | x | x | x |
| Hawaii | | | x | x |
| Idaho | | x | x | x |
| Illinois | x | x | x | x |
| Indiana | x | x | x | x |
| Iowa | | x | | |
| Kansas | x | | x | x |
| Kentucky | | | x | x |
| Louisiana | | x | x | x |
| Maine | x | x | x | x |
| Maryland | x | x | x | x |
| Massachusetts | x | x | x | x |
| Michigan | x | x | x | x |
| Minnesota | x | | x | x |
| Mississippi | x | x | x | |
| Missouri | | | | x |
| Montana | x | x | x | x |
| Nebraska | x | | x | x |
| Nevada | x | x | x | |
| New Hampshire | | x | | |
| New Jersey | x | | x | x |
| New Mexico | x | x | x | x |
| New York | x | x | | x |
| North Carolina | x | x | x | x |
| North Dakota | | x | x | x |
| Ohio | | | x | x |
| Oklahoma | x | x | x | x |
| Oregon | | x | x | |
| Pennsylvania | | x | x | x |

| State | Alt. Assessments/ SWD 1% Rule | Alt. Assessments/ SWD 3% Rule | Assessing LEP's | Participation Rate- Subgr. |
|----------------|----------------------------------|----------------------------------|--------------------|-------------------------------|
| Rhode Island | | | x | x |
| South Carolina | x | | x | x |
| South Dakota | | x | x | |
| Tennessee | x | x | x | x |
| Texas | | | x | x |
| Utah | x | | x | x |
| Vermont | | | | |
| Virginia | | x | x | x |
| Washington | x | | x | x |
| West Virginia | | x | x | x |
| Wisconsin | x | | x | x |
| Wyoming | x | | x | x |
| Total | 31 | 28 | 43 | 41 |

States are not required to include how they plan to meet the highly qualified teacher requirements in their accountability plans. Instead, ED monitors states for compliance with meeting these requirements (U.S. Department of Education, 2005, August 3). Thus, all states were eligible for the changes in the highly qualified teacher requirements that pertain to multi-subject and science teachers. The flexibility for rural districts was extended to districts that are eligible for the Small Rural Schools Achievement (SRSA) program. ED maintains a list of eligible districts and for the 2004-05 school year there were 4,906 eligible districts.¹⁵ States seeking an extension in the timeline for meeting the requirements must seek approval from ED.

The number of states where specific state initiated amendments were approved is summarized in Tables 6 through 8. These tables summarize the amendments approved by ED and do not include states that may have adopted the change in their initial accountability plan. Seven states adopted a rule that says a school must fail to make AYP in the same subject for two consecutive years before it is identified for improvement. There were 33 states that were allowed to use the grade span method when calculating district accountability and 1 state (Maryland) where the state could apply this method to the state when determining AYP (Table 6). Seventeen states changed the minimum group size required for reporting subgroup scores. These changes raised the minimum group size, either by adopting a higher minimum number or using a method that is proportional to the overall student population. For example, Georgia will use either a minimum group size of 40 or 10% of the overall student population, whichever is greater (with a 75 student cap) and Virginia will use 50 or 1% of the student population, with a 200-student cap. There were seven states (Alaska, Kansas, Missouri, New Jersey, New York, South Carolina, and Washington) that were permitted to establish a minimum group size for the students with disabilities subgroup that was different from that for other subgroups. Five states (Alaska, Minnesota, Missouri, South Carolina, and Washington) were allowed to have a different subgroup size for LEP students.

¹⁵ The list of eligible districts is available at www.ed.gov/programs/reapsrsa/eligible04/index.html.

Table 6: States Taking Advantage of State Initiated Changes: Adopted Same Subject and Grade Spans for Accountability, Changed Subgroup Size, and Made Statistical Changes.

| State | Same Subject | Grade Span for Accountability | | Minimum Subgroup Size | | | Statistical Changes In Use of Data | | | | |
|----------------|--------------|-------------------------------|-------------|-----------------------|-----|-----|------------------------------------|-----------------------------------|----------------|------------------------|----------------|
| | School Level | District Level | State Level | All Subgrps | SWD | ELL | Confidence Intervals/ AYP | Confidence Intervals/ Safe Harbor | Standard Error | Uniform Averaging/ AYP | Rounding Rules |
| Alabama | | x | | | | | x | | | x | |
| Alaska | | x | | | x | x | | x | | | |
| Arizona | | x | | x | | | | | | | |
| Arkansas | | x | | x | | | x | | | | |
| California | | x | | | | | x | x | | x | x |
| Colorado | | | | | | | | | | | |
| Connecticut | | x | | | | | | | | | |
| Delaware | | x | | | | | x | x | | | |
| Florida | | | | x | | | | | | | |
| Georgia | | x | | x | | | | | | | |
| Hawaii | | | | x | | | | x | x | | |
| Idaho | | | | | | | | x | | | |
| Illinois | x | x | | x | | | | | | | |
| Indiana | | x | | | | | x | | | | |
| Iowa | | | | | | | | | | | |
| Kansas | | | | | x | | x | x | | | |
| Kentucky | | | | x | | | | | | x | |
| Louisiana | x | x | | | | | | x | | | |
| Maine | | x | | | | | | x | | | |
| Maryland | | x | x | | | | | | | | |
| Massachusetts | | | | x | | | | | | | |
| Michigan | | | | x | | | | | x | | |
| Minnesota | | | | | | x | | | | x | |
| Mississippi | | x | | | | | x | | | | |
| Missouri | | x | | | x | x | x | x | | | |
| Montana | | | | x | | | x | x | | | |
| Nebraska | | x | | x | | | x | x | | | |
| Nevada | | x | | | | | | | | | |
| New Hampshire | | x | | | | | | | | | |
| New Jersey | | x | | x | x | | | | | | x |
| New Mexico | | x | | | | | x | | | | |
| New York | x | x | | x | x | | | | x | | |
| North Carolina | x | x | | x | | | x | | | | |
| North Dakota | | | | | | | | x | | | |
| Ohio | | x | | | | | | | | | |
| Oklahoma | | x | | | | | x | x | | | |
| Oregon | x | x | | | | | | | | | |
| Pennsylvania | | x | | | | | x | | | | |
| Rhode Island | | | | | | | | | | x | |

| State | Same Subject | Grade Span For Accountability | | Minimum Subgroup Size | | | Statistical Changes In Use of Data | | | | |
|----------------|--------------|-------------------------------|----------------|-----------------------|-------------|----------|------------------------------------|---------------------------|-----------------------------------|----------------|------------------------|
| | | School Level | District Level | State Level | All Subgrps | SWD | ELL | Confidence Intervals/ AYP | Confidence Intervals/ Safe Harbor | Standard Error | Uniform Averaging/ AYP |
| South Carolina | x | x | | | x | x | | | x | x | |
| South Dakota | | x | | | | | | x | | | |
| Tennessee | | x | | x | | | x | | | | x |
| Texas | | | | | | | | | | | x |
| Utah | | | | | | | | x | | | |
| Vermont | | | | | | | | | | | |
| Virginia | | x | | x | | | | | | | |
| Washington | | | | x | x | x | x | | | x | x |
| West Virginia | x | x | | | | | | | | x | |
| Wisconsin | | x | | | | | x | x | | | |
| Wyoming | | x | | | | | | x | | x | |
| Total | 7 | 33 | 1 | 17 | 7 | 5 | 17 | 17 | 4 | 9 | 5 |

There were several amendments states requested that changed statistically how they calculate AYP. Seventeen states added or changed the way they use confidence intervals when making AYP determinations (Table 7). A similar process, the use of standard error of measurement to calculate AYP, was adopted by four states. Since 31 states already included confidence intervals in their original accountability plans (Center on Education Policy, 2005, November 16), most states now use them to calculate AYP. In addition to adding confidence intervals, some states adopted larger confidence intervals. While most states use a 95% interval, a number of states increased the confidence interval from 95% to 99%. The other trend was to adopt confidence intervals for calculating safe harbor. There were 17 states that added the use of a 75% confidence interval to safe harbor determinations, the limit allowed by ED. Other statistical changes in how states calculated AYP included adopting uniform averaging procedures (9 states) or rounding rules (5 states), which allow states to round AYP calculations to the nearest whole number.

Other amendments affected how states defined AYP (Table 7). There were 8 states that adopted or revised a system that uses an index to determine the AYP status of districts and schools. Alabama, New Mexico, Pennsylvania, and South Carolina added an index in 2005 and Connecticut announced plans to investigate using an index to compute proficiency levels after the spring 2006 test administration. Pennsylvania received permission to incorporate the state's Pennsylvania Performance Index (PPI) into its AYP determinations. Under the Pennsylvania amendment, if a school or district does not meet the state's proficiency targets or make AYP through the safe harbor provisions, they may still make AYP if it reaches the PPI performance targets. Massachusetts, Mississippi, and Oklahoma revised an index that was approved in their original plan. Other states already using an index to calculate proficiency levels that was approved in their original state accountability plan include Louisiana, Minnesota, New York, Rhode Island, Vermont, bringing to 13 the number of states using an index.

In April 2005, ED approved an amendment that allows Florida to add a new designation to the definition of AYP, “provisional AYP” for those schools that received an A or B school letter grade under the state accountability plan but did not make AYP. Florida was the first state, and at this writing, the only state with a “provisional AYP” rating. Six states changed their annual measurable objectives (AMOs) by adopting intermediate goals that increased in equal increments, the approach taken by Florida, or equal stair step increments. Another eight states adopted new AMOs because they had adopted new assessments.

States also made changes in the additional indicator required by NCLB when calculating AYP (Table 7). States were allowed to set new goals or change existing goals for both the graduation rate requirement (19 states) and the additional indicator (16 states), which were usually attendance rate targets. Finally, states could set different graduation rate targets and/or attendance targets for students learning English and for students with disabilities.

Table 7: States Taking Advantage of State Initiated Changes: Changed Definition of AYP, Annual Measurable Objectives, and Additional Indicator.

| AYP | | Annual Measurable Objectives | | | Additional Indicator: New/ Existing Goals Set | | Additional Indicator: New/Existing Goals Set: LEP | | Additional Indicator: New Existing Goals Set: SWD | |
|---------------|-------|------------------------------|--------------------|--------------------------------|---|-----------------|---|-----------------|---|-----------------|
| State | Index | Def of AYP | Intermediate Goals | Intermediate Goals/New Assess. | Graduation Rate | Attendance Rate | Graduation Rate | Attendance Rate | Graduation Rate | Attendance Rate |
| Alabama | x | | | X | | | | | | |
| Alaska | | | | X | | | | | | |
| Arizona | | | | | | x | | | | |
| Arkansas | | | | | x | x | | | | |
| California | | | | | x | | | x | | |
| Colorado | | | | | | | | | | |
| Connecticut | x | | | | x | | | | | |
| Delaware | | | | | | x | | | | |
| Florida | | x | x | | | | | | | |
| Georgia | | | | X | x | | x | | x | |
| Hawaii | | | | | | x | | | | |
| Idaho | | | | | | x | | | x | |
| Illinois | | | | | | | | | | |
| Indiana | | | | | | | | | | |
| Iowa | | | | | | | | | x | |
| Kansas | | | | | | x | | | x | |
| Kentucky | | | x | | x | x | | | x | |
| Louisiana | | | | | | | | | | |
| Maine | | | | | x | | | | | |
| Maryland | | | | | x | x | | | | |
| Massachusetts | X* | | x | | x | | | | | |
| Michigan | | | | | x | | | | | |
| Minnesota | | | | X | | | | | | |
| Mississippi | X* | | | | | | | | | |
| Missouri | | | x | | | | | | | |
| Montana | | | | X | | | | | | |

| State | AYP | | Annual Measurable Objectives | | Additional Indicator: New/Existing Goals Set | | Additional Indicator: New/Existing Goals Set: LEP | | Additional Indicator: New Existing Goals Set: SWD | |
|----------------|----------|------------|------------------------------|--------------------------------|--|-----------------|---|-----------------|---|-----------------|
| | Index | Def of AYP | Intermediate Goals | Intermediate Goals/New Assess. | Graduation Rate | Attendance Rate | Graduation Rate | Attendance Rate | Graduation Rate | Attendance Rate |
| Nebraska | | | | | | | x | | x | |
| Nevada | | | | X | x | x | | | | |
| New Hampshire | | | | | | | | | | |
| New Jersey | | | | | x | | x | | x | |
| New Mexico | x | | | X | x | x | | | x | |
| New York | | | | | x | | | | | |
| North Carolina | | | | | | | | | | |
| North Dakota | | | | | x | x | x | | x | |
| Ohio | | | | | | | | | | |
| Oklahoma | X* | | | | | | | | | |
| Oregon | | | | | | | | | | |
| Pennsylvania | x | | | | x | x | | | | |
| Rhode Island | | | | | | x | | | | |
| South Carolina | x | | | | x | | | | x | |
| South Dakota | | | | X | x | | | | | |
| Tennessee | | | | | x | x | | | | |
| Texas | | | | | | | | | | |
| Utah | | | | | | | | | | |
| Vermont | | | | | | | | | | |
| Virginia | | | x | | | x | x | | x | |
| Washington | | | x | | x | x | x | x | x | |
| West Virginia | | | | | | | | | | |
| Wisconsin | | | | | | | | | | |
| Wyoming | | | | | | | | | x | |
| Total | 8 | 1 | 6 | 8 | 19 | 16 | 6 | 2 | 13 | 0 |

Note: *Index approved in original plan; revised in subsequent amendments. Louisiana, Minnesota, New York, Rhode Island, and Vermont also included an Index in their original state accountability plan.

Table 8 summarizes the number of states making changes in their assessment systems, and in how they include schools in the accountability system, intervene in schools and districts needing improvement, and report data and information related to NCLB. The changes states made in their assessment systems included allowing high school students to take exams that are used for AYP purposes multiple times. The administration first adopted a rule requiring states to count the score a student earned the first time they took the test. This change allowed states to align NCLB testing requirements with state priorities. Since many states use these same tests as high school exit exams, making them high stakes, they allow students to re-take the exams multiple times.

Table 8: States Taking Advantage of State Initiated Policy Changes: Adopted Changes in Assessment System, How Schools are Included in Accountability System, How States Intervene in Schools, and How States Report NCLB Data.

| State | Changes in Assessment System | | | | Inclusion/Interventions/Reporting | | |
|----------------|------------------------------|-------------------------|--------------------|----------------------------|--|---------------|-----------|
| | State Level | Delayed Testing Program | Allow Test Retakes | Testing Out of Grade Level | Including All Schools in Accountability System | Interventions | Reporting |
| Alabama | x | | | | | x | x |
| Alaska | | | | | | x | |
| Arizona | | | | | | | x |
| Arkansas | | | | | | | |
| California | | | | | x | | x |
| Colorado | | | | | | | |
| Connecticut | | | | | | x | x |
| Delaware | | | | | | x | x |
| Florida | | | | | | | |
| Georgia | x | | | | x | x | x |
| Hawaii | | | | | x | | x |
| Idaho | | | | | | | x |
| Illinois | | | | | | x | x |
| Indiana | | | | | | | |
| Iowa | | | | | | | |
| Kansas | | | | | x | | |
| Kentucky | | | | | | | x |
| Louisiana | x | | | | | | |
| Maine | x | | | | x | | x |
| Maryland | x | | | x | | | |
| Massachusetts | | | | | x | | |
| Michigan | | x | | | x | | x |
| Minnesota | x | | | | x | | |
| Mississippi | | | | | | | |
| Missouri | | | | | | | |
| Montana | | | | | x | | |
| Nebraska | | | | | x | | |
| Nevada | | | | | x | | |
| New Hampshire | | | | | | | |
| New Jersey | x | | x | | | x | x |
| New Mexico | | | | | x | | x |
| New York | | | | | | | |
| North Carolina | x | | | | x | x | x |
| North Dakota | | | | | x | | x |
| Ohio | | | | | | | |
| Oklahoma | | | | | | | |
| Oregon | | | | | | | x |
| Pennsylvania | | | | | | | |
| Rhode Island | | | | | | | x |
| South Carolina | x | | | | x | x | x |
| South Dakota | | | | | x | x | x |
| Tennessee | x | x | | x | | | |
| Texas | | | x | | x | | x |

| State | Changes in Assessment System | | | | Inclusion/Interventions/Reporting | | |
|---------------|------------------------------|-------------------------|--------------------|----------------------------|--|---------------|-----------|
| | State Level | Delayed Testing Program | Allow Test Retakes | Testing Out of Grade Level | Including All Schools in Accountability System | Interventions | Reporting |
| Utah | | | | | | | |
| Vermont | | | | | | | |
| Virginia | | | x | | | | x |
| Washington | | x | x | | x | | |
| West Virginia | | | | | | | x |
| Wisconsin | | | | | x | | |
| Wyoming | x | | x | | x | x | x |
| Total | 11 | 3 | 5 | 2 | 20 | 11 | 24 |

Table 9 summarizes the data on state accountability plan amendments. The number of amendments adopted by each state ranged from zero in Vermont to 15 in Washington for the state initiated changes and from zero to 18 for all changes (Table 9). States adopted an average of 2.84 of the rule changes, 5.56 of the state initiated changes, for a total average of 8.4 changes. The various combinations of amendments within any one state and the range in the number of changes states made show that widely different compromises have been made between ED and the states on what NCLB means. The result is that NCLB has no common meaning across states and accountability has little relationship to student achievement.

Table 9: Range and Average Number of Changes Adopted by States by Type of Change (Rule Change, State Initiated, Total Changes), December 2005.

| Type of Change | Range of Changes | Average # of Changes |
|-----------------------|------------------|----------------------|
| Rule, N=4 | 0-4 | 2.84 |
| State Initiated, N=27 | 0-15 | 5.56 |
| Total, N=31 | 0-18 | 8.40 |

IMPACT OF RULE AND STATE INITIATED POLICY CHANGES

Since the kinds and number of changes states have adopted are not uniform across states, with each state requesting its own configuration of amendments, it is difficult to unravel the impact these changes are likely to have. The combination of changes as much as any one will determine their impact. In general, both the policy changes and state requested amendments affect which students are included under the law's accountability provisions and which districts and schools are deemed successful and which are labeled as needing improvement. Because of the array of changes states can adopt, accountability will mean different things in different states as well as between districts and schools within a state, as we show in this section. In addition, the allowable statistical techniques add complexity to the NCLB accountability system, complicate understanding what AYP and accountability mean, and obscure the ability of states, districts, and schools to show improvements in student performance.

In this section we first examine the impact of the NCLB changes on AYP and accountability. We provide data on how changing the minimum group size has a differential impact, depending on the size of schools or districts, examine amendments approved for Florida that allow the state to change intermediate goals and add a “provisional AYP” category, and show how the graduation rate amendments further weaken graduation rate accountability. The second section looks at the question of who benefits (and loses) from the grade span method of identifying districts for improvement and the changes in the highly qualified teacher requirements. Finally, we examine how the statistical techniques that states can now use add complexity to the accountability system. In commenting on the impact of these changes, we are cognizant of both the political and practical implications of the changes, while at the same time recognizing that they introduce incoherence into the accountability system because they do not address the underlying flaws of NCLB.

Effect on AYP and Accountability

Some changes, such as those allowing states to use alternative achievement standards to measure the progress of students with disabilities, will likely reduce the number of schools identified for improvement solely because of the scores of students in this subgroup. For example, the 1% and 2% rules that allow states to use alternative achievement standards for students with disabilities when determining AYP were designed to mathematically reduce the number of students who must meet grade level standards. These changes should allow states to use assessments that more accurately reflect what these students learn but since they apply a uniform rule to different state systems for assessing special education students, they provide no guidelines for which students should be exempt from regular assessments. At the same time, these policies do not resolve the issue of how to hold schools accountable for the performance of this subgroup. At the school level, the effect of these changes is likely to depend on the overall demographic makeup of the school. Since schools with more subgroups have a higher probability of being identified for improvement, those schools with fewer subgroups are likely to benefit more than those with more subgroups (Kim & Sunderman, 2005).

Other changes, such as the rule allowing schools and districts to average participation rates over three years, will provide temporary relief since a low participation rate in one year will have to be compensated for in other years. Changing the minimum group size is an easy change to make but is likely to have minimal effect on the number of schools and districts identified for improvement. The use of confidence intervals to determine AYP is likely to have a greater impact on the number of schools and districts identified for improvement, but adds complexity to understanding the meaning of AYP and complicates efforts to show improvements in student performance.

Minimum group size: States adopted different approaches to establishing a minimum group size. Some adopted different group sizes for students with disabilities and for students with limited English proficiency. Other states raised the minimum group size or made it proportional to the school or district enrollment. Overall, the larger group

sizes are likely to have minimal impact on whether a school or district is identified for improvement, in part because it does not address the issue of counting students in multiple categories (Kim & Sunderman, 2005). A more significant impact will be on which subgroups schools and districts are required to report.

The affect of the proportional group size will depend on the size of the school or district. For example, Florida adopted a minimum group size of 30 or more as long as that group represents 15% of the school’s total enrollment, with a cap of 100 students per group. This change means that the subgroup size will be proportional to those schools that enroll between 200 and 666 students. In a school with fewer than 200 students, the group size is 30 students, and in schools with more than 666 students, it is 100 students regardless of enrollment. Georgia also adopted a proportional subgroup size of 40 students or 10% of the student enrollment, with a cap of 75 students. For small schools (those with less than 40 students), Georgia uses a confidence interval methodology to determine whether the school’s total group passing rate is significantly below the state’s AYP passing rate. This rule makes the subgroup size proportional to schools that enroll between 400 and 750 students. There are 993 schools, or 47.3%, in the 400-750 student range (Table 10).

Table 10: Number and Percent of Schools by Enrollment Category, Georgia, 2005-06.

| Less than 40 Students | | Between 40 and 400 Students | | Between 400 and 750 Students | | Greater than 750 Students | |
|-----------------------|-----|-----------------------------|------|------------------------------|------|---------------------------|------|
| # | % | # | % | # | % | # | % |
| 7 | 0.3 | 294 | 14.0 | 993 | 47.3 | 806 | 38.4 |

Source: Georgia Department of Education. Calculations are our own.

Virginia adopted a proportional group size that affects primarily districts. The minimum group size in Virginia is 50 or 1% of the student population with a cap of 200 students, which is proportional to enrollment in districts that enroll between 5,000 and 20,000 students. Among the 133 districts in Virginia, there are 36 districts or 27.1% that fall within the 5,000 to 20,000 range (Table 11).

Table 11: Number and Percent of Districts by Enrollment Category, Virginia, 2004-05.

| Less than 5,000 Students | | Between 5,000 and 20,000 Students | | Greater than 20,000 Students | |
|--------------------------|------|-----------------------------------|------|------------------------------|-----|
| # | % | # | % | # | % |
| 84 | 63.2 | 36 | 27.1 | 13 | 9.8 |

Source: Virginia Department of Education. Calculations are our own.

Changing the group size will impact the reporting on those subgroups with fewer students. We compared the number of school districts in Virginia that would be required to report the major racial subgroups under two conditions: using the same subgroup size for all districts compared to using the proportional subgroup rule. The results are presented in Table 12. There is no difference in the number of districts that would be

required to report a white or black subgroup, the largest subgroups in Virginia.¹⁶ There is a slight change in the number of districts reporting a Latino subgroup (2.3% change), an Asian subgroup (3.8% change), or a Native Hawaiian subgroup (2.3% change). The largest change was in the Native American subgroup, with only 2 districts reporting this subgroup under the proportional rule versus 17 districts under the one size subgroup rule (11.3% change). Still, the effect is small.

Table 12: Number of Districts Reporting Subgroups at Time 1 (if subgroup >50) and Time 2 (if subgroup >50 or 1% of total enrollment), Virginia, 2004-05.

| Subgroup | # of districts at T1 | % of districts at T1 | # of districts at T2 | % of districts at T2 | # of districts w/ change | Change in % of districts reporting | % change in districts reporting |
|---------------------|----------------------|----------------------|----------------------|----------------------|--------------------------|------------------------------------|---------------------------------|
| White | 133 | 100.0 | 133 | 100.0 | 0 | 0.0 | 0.0 |
| Black | 120 | 90.2 | 120 | 90.2 | 0 | 0.0 | 0.0 |
| Hispanic | 75 | 56.4 | 72 | 54.1 | 3 | 2.3 | 4.0 |
| Asian | 44 | 33.1 | 39 | 29.3 | 5 | 3.8 | 11.4 |
| Native American | 17 | 12.8 | 2 | 1.5 | 15 | 11.3 | 88.2 |
| Native Hawaiian | 3 | 2.3 | 0 | 0.0 | 3 | 2.3 | 100.0 |
| Unspecified* | 20 | 15.0 | 10 | 7.5 | 10 | 7.5 | 50.0 |
| Total N (districts) | 133 | | | | | | |

Source: Virginia Department of Education. Calculations are our own.

Changing Intermediate Goals—Reducing the Number of Schools Identified for Improvement: Secretary Spellings approved several amendments to the Florida accountability plan that illustrate how changing the intermediate goals can reduce the number of schools identified for improvement, at least temporarily. In a letter to Secretary Spellings (dated April 1, 2005), the Florida Commissioner of Education requested three changes:

1. Revise the annual proficiency targets for reading and mathematics to ten intermediate goals that increase in equal increments.
2. Maintain a subgroup size of 30, but it must represent 15% of the total school population before the performance of the subgroup will impact the total school rating.
3. Replace the current Safe Harbor measure with individual student improvement (learning gains) to proficiency (Florida Department of Education, 2005, April 1).

Florida based its request on the argument that the low percentage of schools making AYP was because of the size and diversity of Florida's schools, the growth of special populations, and the exposure to one criterion affecting the AYP status of schools (Florida Department of Education, 2005, April 1). In 2004-05, only 23% of Florida

¹⁶ Percentage of enrollment by subgroup for Virginia, 2004-05: 59.7% white, 26.8% black, 7.0% Latino, 4.8% Asian, 0.3% Native American, 0.1% Native Hawaiian, and 1.4% unspecified. Source: Virginia Department of Education.

schools made AYP. To increase this number, Florida argued that the number of schools making AYP had less to do with student achievement than it did with other issues. The state has seen its school age population increase by 23.5% between 1990 and 2000, with particularly large increases in the number of LEP, low-income, and students with disabilities subgroups. Florida argued that the state was “more exposed to the impact of the NCLB AYP calculations than other states because of our cell size, school size and diversity of schools, as well as our high academic standards” (Florida Department of Education, 2005, April 1). In other words, large numbers of Florida schools were likely to be identified for improvement because the subgroup accountability rules mean large numbers of schools must meet multiple performance targets and the use of mean proficiency disadvantages schools that start further behind the proficiency targets, a situation that exists in other states as well (Kim & Sunderman, 2005; R. L. Linn, 2003; R. L. Linn, Baker, & Betebenner, 2002).

Revising the annual proficiency targets changed the percentage of students needing to demonstrate proficiency for 2004-05 from 48% to 37% in reading and from 53% to 44% in mathematics (Florida Department of Education, 2005, April 1). Thereafter, Florida proposed annual increases in proficiency rather than stair-step increases at three-year intervals. The change in the subgroup size would remove some subgroups from calculating the school’s AYP status. The last requested change would allow Florida to replace the Safe Harbor measure with a value-added or growth method for determining AYP based on individual student scores rather than mean proficiency. Taken together, these changes would reduce the number of schools identified for improvement.

In May 2005, federal officials approved two of the requested amendments: revisions in the timetable for meeting the proficiency targets and the change in the subgroup size (Matus, 2005; Olson, 2005a). The third request, the use of learning gains to measure progress, remained under consideration. Since other states were asking for a similar change, Spellings stated that she would form a panel to look at the idea of measuring proficiency through individual learning gains.

In addition to the changes requested in the April 2005 letter, ED approved an amendment that allows Florida to add a new designation to the definition of AYP, “provisional AYP” for those schools that received an A or B school letter grade under the state accountability plan but did not make AYP. Florida was the first state, and at this writing, the only state with a “provisional AYP” rating. This creates a new category of schools that are exempt from the NCLB sanctions as long as they maintain an A or B grade under the state accountability system. The effect of “provisional AYP” together with the other changes reduces the number of schools identified for improvement by shifting many of them to the new category. In 2003-04, 73.4% of schools were identified for improvement; in 2004-05, 37.4% of schools were identified for improvement and 26.6% were identified as “provisional AYP” (Table 13). The increase in schools making AYP is likely related to lowering the proficiency target for 2004-05.

Table 13: School Level AYP: Percentage of Schools Making AYP, Identified for Improvement, and “Provisional AYP,” Florida, 2003-04 and 2004-05.

| % Schools Making AYP | | % Schools Identified for Improvement | | % Schools Identified as Provisional AYP | |
|----------------------|---------|--------------------------------------|---------|---|---------|
| 2003-04 | 2004-05 | 2003-04 | 2004-05 | 2003-04 | 2004-05 |
| 22.4 | 36.0 | 73.4 | 37.4 | 0 | 26.6 |

Source: Florida Department of Education. Note: In 2003-04, 4.2% of the schools were not rated.

Graduation Rate Requirements: NCLB includes a provision for graduation rate accountability designed to counter incentives to exclude low-performing students from schooling in their efforts to improve achievement test scores. States must include graduation rates as an additional indicator of academic achievement when they determine if a school or district is making adequate progress. However, regulations developed by ED that allow states considerable flexibility in how they meet this requirement have substantially weakened graduation rate accountability. The regulations do not include subgroup accountability (states do not have to disaggregate graduation rates by subgroups) and allow states to make progress towards meeting graduation rates rather than require states to meet a set target (Losen, 2005).

Many states cited a lack of a student record system that would allow the state to track individual students longitudinally from grade nine through graduation as limitations in establishing graduation rate accountability. In addition, states opposed stricter graduation rate accountability unless there was some relief from the strict AYP provisions. They fear that stricter graduation rate accountability will add another way for schools and districts to fail AYP. This could be balanced if gains in graduation rates were counted as part of an indicator system rather than as an add-on to the NCLB test score accountability. Although states are required to set a goal, there are no requirements for minimum gains in graduation rates or for raising the graduation rate of subgroups.

Many of the amendments approved by ED further weaken graduation rate accountability. Amendments that incorporated “making progress” towards the graduation rate target as sufficient for making AYP or adjusted the graduation rate target furthered this trend towards weak graduation rate accountability. For example, California, which required schools and districts to demonstrate an increase of 1/10th of a percentage point in graduation rate (from any starting point) or to meet an annual status target that begins at 82.8%, added a definition of progress for schools averaging two years of data (schools must demonstrate an increase of 2/10th of a percentage point when averaging data) (California Department of Education, 2005, August 23). Georgia removed from its graduation rate calculations students receiving certificates of attendance and Special Education diplomas. This lowered the percentage of students graduating from high school in 2002 from 72.7% to 61.8%. Based on this revised graduation rate, Georgia set its graduation rate target at 60% and allowed schools below this rate to still make AYP if they showed progress over the preceding year (Georgia Department of Education, 2005, June 28). South Dakota changed its graduation rate target from 90% to 80% and Maryland incorporated a measure of progress (1/10th of a percentage point improvement over the previous year) into its AYP determinations.

Among the amendments are provisions that allow states to set separate graduation rate targets for LEP students and students with disabilities. For example, Georgia extended the standard number of years for LEP students to graduate from four to five years and a summer. For students with disabilities, the amendment allows a student’s IEP team to determine the “standard” number of years needed for a student with disabilities to graduate. New Jersey allows up to 6 years for LEP students entering the U.S. during high school and students with disabilities to earn a high school diploma. These changes could decrease incentives for schools to push out LEP or special education students if they are held accountable to a more realistic standard and may make sense from the perspective of a student’s academic background. However, allowing states to set different graduation rate standards or reporting procedures for LEP students and students with disabilities creates a different accountability system for these students, something advocates for these students oppose.

Who Benefits/Who Loses?

Grade Span Method: Research on the grade span method shows that using this method both reduced the number of districts identified for improvement and benefited some districts more than others (Tracey, Sunderman, & Orfield, 2005). States that have been allowed to use the grade span method can aggregate test scores within a grade span (i.e., grades 3-5, grades 6-8, and grades 9-12) and a district is identified for improvement only if it fails to make adequate progress in the *same* content area for *all* grade spans. In states without this agreement, districts must show adequate progress at *each* grade level and for *each* content area (AYP method). The effect of this change in California and Georgia is shown in Table 14. In California, the grade span method reduced the number of districts identified for improvement by 60.3%, from 378 districts (36.4% of districts statewide) to 150 districts (14.4% of districts statewide). In Georgia, the grade span method reduced the number of districts needing improvement by 89.6%. Using this method to determine district accountability in Georgia reduced the number of districts identified for improvement from 115 (representing 63.5% of total districts) to 12 (6.6% of total districts).

Table 14: The Number and Percentage of Districts Making Adequate Progress and Identified for Improvement Using the AYP Method and the Grade Span Method, California and Georgia, 2004-05.

| State | AYP Method | | | | Grade Span Method | | | |
|----------------------|----------------|------|--------------|------|-------------------|------|--------------|------|
| | Adequate Prog. | | Need Improv. | | Adequate Prog. | | Need Improv. | |
| | N | % | N | % | N | % | N | % |
| California Districts | 661 | 63.6 | 378 | 36.4 | 889 | 85.6 | 150 | 14.4 |
| Georgia Districts | 66 | 36.5 | 115 | 63.5 | 169 | 93.4 | 12 | 6.6 |

Source: California Department of Education and Georgia Department of Education (calculations are our own).

There are a number of practical and political reasons for wanting to reduce the number of districts identified for improvement, including the lack of state capacity to intervene in large numbers of districts labeled as needing improvement and the political expediency of removing highly performing districts from the failing list. However, our

research shows that the grade span method benefited some districts more than others and has a disparate impact on districts serving minorities. Using data from California, those districts serving White students were more likely than their counterparts serving minority students to be removed from the needs improvement category. Table 15 compares the demographic characteristics of students enrolled in districts identified for improvement and districts making adequate progress using both the AYP method and the grade span method. Using the grade span method to identify districts for improvement reduced the percentage of White students enrolled in districts needing improvement from 21.4% to 17.7% and increased the percentage of Latino students from 58.5% to 62.1%. There was a slight change in the percentage of low-income students in improvement districts (60.1% under the AYP method and 62.9% under the grade span method). The percentage of LEP students also increased, from 31.5% under the AYP method to 34.2% under the grade span method. The grade span method also disadvantaged elementary and high school districts, which contain few grade spans, while favoring unified districts that include all or multiple grade spans (see Tracey, et al., 2005 for more information on how the grade span method benefits some districts more than others).

Table 15: Demographic Characteristics (Percentages of Minority, Low-Income, and English Language Learners) by District Accountability Status under the AYP and Grade Span Methods, California, 2004-05.

| Subgroup | AYP Method | | | | Grade Span Method | | | |
|-----------|------------------------------|------|-----------------------------|------|------------------------------|------|-----------------------------|------|
| | Adeq Prog. Districts (N=661) | | Need Imp. Districts (N=378) | | Adeq Prog. Districts (N=889) | | Need Imp. Districts (N=150) | |
| | Enrollment | | Enrollment | | Enrollment | | Enrollment | |
| | N | % | N | % | N | % | N | % |
| White | 1,190,763 | 45.2 | 782,213 | 21.4 | 1,503,160 | 41.5 | 469,816 | 17.7 |
| Latino | 809,906 | 30.8 | 2,134,523 | 58.5 | 1,291,061 | 35.7 | 1,653,368 | 62.1 |
| Black | 149,162 | 5.7 | 354,336 | 9.7 | 234,813 | 6.5 | 268,685 | 10.1 |
| Asian | 292,452 | 11.1 | 207,086 | 5.7 | 342,523 | 9.5 | 157,015 | 5.9 |
| Soc. Dis. | 908,125 | 34.5 | 2,194,984 | 60.1 | 1,429,599 | 39.5 | 1,673,510 | 62.9 |
| LEP | 446,806 | 16.7 | 1,150,456 | 31.5 | 686,805 | 19.0 | 910,457 | 34.2 |

Source: California Department of Education (calculations are our own)

Rather than addressing the flaws inherent in the NCLB accountability provisions, the grade span method of calculating AYP compounds them. Both the original method of calculating AYP and the grade span method make it harder for some districts to make AYP than others for reasons unrelated to academic performance. Two core components of NCLB—the subgroup rules and the reliance on mean proficiency—pose particular challenges for high-poverty districts that enroll large numbers of minority students (Kim & Sunderman, 2005; R. L. Linn, 2003; Tracey et al., 2005). The subgroup rules disadvantage both large and diverse districts because they must meet more performance targets than smaller, more homogeneous districts. The reliance on mean proficiency does not take into account academic growth and disadvantages districts that enroll students who begin their schooling with lower skill levels. The change to the grade span method of determining district accountability left these two mechanisms in place while, at the same time, altering the definition of AYP in ways that was unrelated to academic performance.

Highly Qualified Teacher Requirements: As with changing the way districts are identified for improvement, changing the highly qualified teacher requirements reflect both political and practical considerations. Politically, by making the changes, ED acknowledged the difficulties states have in meeting the requirements, particularly in rural parts of the country, and in applying the subject matter requirements to special categories of teachers. Extending the deadline for states to meet them allows states more time to develop the necessary data systems to track teacher qualifications, an issue that has raised concerns about the reliability of the data. Nonetheless, the flexibility given rural schools in meeting the highly qualified teacher requirements benefits some regions of the country more than others because of the definition ED used to define a rural district. It also favors districts that are less poor and serve fewer minorities. Allowing states more time to meet the 100% goal lacks transparency since states must apply for this provision and ED has not clearly stated how it will evaluate state efforts to show how they meet a set of core principles.

The rule granting flexibility to rural districts applies only to those districts that are eligible to participate in the Small Rural School Achievement (SRSA) program and was not extended to districts participating in another federal program, the Rural Low Income Schools (RLIS) program. The criteria for participating in these two programs differ. To qualify for the SRSA program, a district must either have fewer than 600 students or be located in a county with fewer than 10 people per square mile. They also must be located in communities with fewer than 2,500 residents. The RLIS program is designed for rural districts in larger communities. There is a poverty level requirement (20% of students served by the district must be low-income) and the district must be located in a town no larger than 25,000 people.

Extending the highly qualified teacher flexibility only to those districts eligible for SRSA benefits some regions of the country while disadvantaging others. The SRSA districts are concentrated in the Midwest and Great Plains whereas RLIS districts are located in the Southeast or Southwest. A report by the Rural School and Community Trust found that 44% of the school districts participating in the SRSA program and 9% of the RLIS districts were located in 12 Midwestern and Great Plains States (The Rural School and Community Trust, 2004, May 13).¹⁷ In contrast, only 1% of the SRSA districts and 42% of the RLIS districts were located in 11 Southeastern states.¹⁸ This regional bias is related to how districts are organized in different regions of the country. Because districts in the Southeast organize on the county level, they tend to be large schools systems located in rural areas whereas the Midwest and Great Plains districts are decentralized and more likely to organize around a local community located in sparsely populated areas.

In addition to the regional bias, the flexibility favors districts that are less poor and have fewer minorities. In 2004-05, there were 4,609 districts participating in SRSA

¹⁷ These states include Illinois, Indiana, Iowa, Kansas, Michigan, Minnesota, Montana, Nebraska, North Dakota, Ohio, South Dakota, and Wisconsin.

¹⁸ These states include Alabama, Florida, Georgia, Kentucky, Louisiana, Mississippi, North Carolina, South Carolina, Tennessee, Virginia, and West Virginia.

and 1,303 districts eligible for the RLIS program (Table 16). The districts participating in RLIS enrolled 42% more students than SRSA, and these students were more likely to be minority and low-income (38% are minority in RLIS compared to 22% in SRSA; 57% are low-income compared to 38%).

Table 16: Number of Districts and Schools Participating in the SRSA and RLIS Programs, Enrollment and Demographic Characteristics of Students, 2004-05.

| Program | # Districts | # Schools | Enrollment | % Minority | % Low Income* |
|---------|-------------|-----------|------------|------------|---------------|
| SRSA | 4,906 | 10,106 | 1,677,116 | 22% | 38% |
| RLIS | 1,303 | 7,244 | 2,889,520 | 38% | 57% |

Source: The Rural School and Community Trust (2004, May 13).

Note: SRSA is the Small Rural School Achievement program and RLIS is the Rural Low Income Schools program.

*Low-income is defined as the percentage of students participating in the free/reduced priced lunch program.

Added Complexity

The combination of statistical techniques states can now use to calculate AYP adds complexity to the system and complicates understanding what AYP and accountability mean. States can now apply confidence intervals to determine AYP and safe harbor, use standard error of measurement in calculating test scores, average data across two or more school years, or use an Index to determine AYP. Ten states may receive approval to adopt a growth model of accountability. Combine these changes with changes in which students are included in the accountability system (i.e., the 1% and 2% rules governing the inclusion of students with disabilities and the exemptions for LEP students) and changes in the minimum subgroup size and the grade span method of calculating AYP, and the complexity of the accountability system increases further.

Use of Confidence Intervals: Using confidence intervals to determine AYP makes sense from a statistical perspective because it takes into account natural fluctuations in test scores. It is beneficial politically since it will reduce the number of schools and districts identified for improvement. On the negative side, it obscures the ability of schools and districts to show improvement and complicates public understanding of AYP and accountability. Below we explain how confidence intervals work and show how it increases the probability of making AYP for schools of different sizes and at different levels of confidence.

Confidence intervals are a statistical technique that creates an interval around the percentage of students who have to score at the proficient level or above. It accounts for natural fluctuations in test scores that arise from sampling errors (i.e., small sample sizes) and factors unrelated to student learning (fluctuations in student scores). It is used to report on the certainty of a measured value. The breadth of the confidence interval (or range within which acceptable scores fall) depends on the number of measurements (number of students tested), the level of confidence desired (i.e., 60%, 95%, 99%), and

the distribution of the data, or test scores.¹⁹ The more certainty desired, the wider the interval will be.

Determining whether a school makes AYP is much like tossing a coin numerous times, where each coin toss represents each student taking the test. Without a confidence interval, there is no margin to account for uncertainty in measurements or random fluctuations. For example, tossing a “fair” coin 10 times should yield a result of 5 heads and 5 tails. Yet this outcome is not certain, even if the coin is “fair.” An outcome of 6 or 7 heads and 4 or 3 tails may also occur. Similarly, if one tosses a coin 100 times, if it is fair, it is unlikely to come up heads 90 times, but it is also unlikely to come up heads exactly 50 times. If the coin toss is heads 55% of the time, one cannot be certain that the coin is not fair and that random fluctuations caused the measurement to be off.

Confidence intervals around the measurement show the range within which the actual mean is likely to fall, and how likely it is to be within that range. In the case of a coin toss, if one tosses a fair coin 100 times, the 95% confidence interval predicts that the real mean for the coin would be in the range between 45 and 65. With a 99% confidence interval, the range is broader and includes outcomes between 42 and 68. Similarly, the more times one flips the coin, the more certain the measurement and the narrower the range of acceptable outcomes. The model that states use to report on the percentage of students scoring at or above a particular target is much like this coin toss example. It assumes that the students in a particular grade/subgroup/school have some actual probability of being measured as proficient and the purpose of the test is to measure all the student’s proficiencies and deduce that probability.

Using a confidence interval to report on the percentage of students scoring at proficient or above improves the probability of making AYP. Table 17 shows the minimum number and percentage of students that must score proficient for schools of different sizes (25 students, 100 students, and 500 students) to meet an AYP goal of 50% proficient. To meet the 50% target using confidence intervals changes the number and percentage of students needing to score at proficient or above. Without confidence intervals, a school with 25 students would need to have half or 13 students score proficient. With a 95% confidence interval, that same school would need to have a minimum of 8 students score at the proficiency level and 7 students using a 99% confidence interval. This represents 32% and 28% of the students in the school who must meet the cut-off score. As the size of the school increases, the minimum percentage of proficient students increases, depending on the size of the confidence interval. In a school with 100 students, 41% of students need to demonstrate proficiency in order to make AYP using a 95% confidence interval and 38% using a 99% confidence.

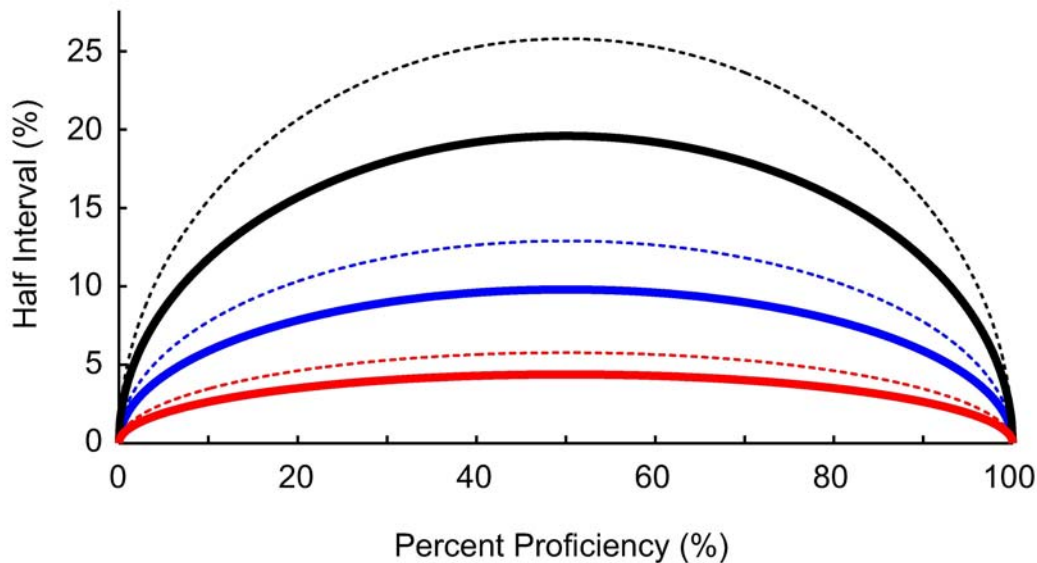
¹⁹ The distribution of test scores does not matter when determining AYP since NCLB requires students to score at or above a set target, which means a student score equals 1 or 0.

Table 17: Number and Percentage of Students Who Must Meet an AYP Goal of 50% Proficient Using Different Confidence Intervals.

| Number of Students | Confidence | Minimum Number of Proficient Students | Minimum Percentage of Proficient Students | Confidence Interval |
|--------------------|-------------|---------------------------------------|---|---------------------|
| School 1: 25 | No Interval | 13 | 52% | N/A |
| | 95% | 8 | 32% | 14%-50% |
| | 99% | 7 | 28% | 4%-52% |
| School 2: 100 | No Interval | 50 | 50% | N/A |
| | 95% | 41 | 41% | 31%-52% |
| | 99% | 38 | 38% | 25%-51% |
| School 3: 500 | No Interval | 250 | 50.0% | N/A |
| | 95% | 229 | 45.8% | 41.4%-50.2% |
| | 99% | 222 | 44.4% | 38.7%-50.1% |

Figure 1 shows the confidence intervals around different measured percent proficiencies. The different lines show the confidence intervals for tests taken by different numbers of students and for different levels of confidence.

Figure 1: Plot of confidence intervals around different percent proficiency levels.



Note: 95% (thick lines) and 99% (dotted lines) confidence intervals are plotted for different measured percent proficiencies. Lines are shown for sample sizes of 25 students (black; top two lines), 100 students (blue; middle two lines), and 500 students (red; bottom two lines). To calculate the confidence interval for a particular measurement, find the measured percentage of students who are proficient (percent proficiency) and its corresponding half interval; the confidence interval ranges from the measured percentage minus the half interval to the measured percentage plus the half interval. The approximation shown here works best for large numbers of students and for proficiencies that fall between 20% and 80%.

While confidence intervals may be a more accurate way to report the percentage of students who are proficient, there are limitations in its use. For one, they obscure the ability of states, districts, and schools to show improvements in student performance. For example, if a school with 100 students has 47% of its students score at the proficient level in year 1 and 53% score proficient in year 2, both are within the range of acceptable outcomes when a 95% confidence interval is used (and the goal is 50% proficient). The school will have made AYP in both cases but it is unclear whether the school is improving, staying the same, or even getting worse. In addition, incentives to improve the test scores of students who score close to the cut-off score remain. Much of the variance comes from scores that fluctuate around the required cut off score and improving those scores will increase the chances a school makes AYP. Likewise, using confidence intervals benefits schools with large numbers of students scoring just below the threshold level.

Index: Using an Index to calculate AYP has the potential to address some of the flaws in NCLB. Under NCLB, all schools must meet a single mean proficiency level in reading and mathematics (the same cut score on standardized achievement tests). With this system, students either make the cut score level or they do not. No credit is given for improving student performance. Since mean proficiency does not correct for selection bias (i.e. it does not reflect differences in student's cognitive skills or background characteristics before they enter school), it does not take into account the contribution of schools to school learning and disadvantages schools serving minority and low-income students (Kane & Staiger, 2002; Kim & Sunderman, 2005; R. Linn, 2000; Raudenbush, 2004). Students with lower test scores must reach the same proficiency target as students with higher test scores. Research suggests that schools serving large numbers of low-income and minority students do generate achievement gains, but since they are likely to start further behind, they have a higher probability of failing to meet the AYP requirements.

An Index that measures student progress allows schools to show movement from one performance level to the next and credits improvement in student achievement across different performance levels. For example, if there are four proficiency levels—below basic, basic, proficient, advanced—schools that move students from below basic to basic would receive credit for this improvement that could be applied to meeting AYP even though those students did not meet the cut score for proficient performance on the standardized test. An Index also creates incentives for schools to focus on students within performance categories rather than just those students who score close to the cut score level. It is likely to benefit schools serving minority and low-income students since improvements in performance are taken into account and to be perceived as a fairer way of holding schools accountable than a system that relies on meeting a fixed target. What an Index does not account for is movement within a proficiency range.

DISCUSSION

The rule changes and amendments were a concession by ED officials that NCLB is not working. Because of the difficulties of applying a one-size-fits-all approach to states with very different educational and political contexts, to districts facing assorted educational challenges, and to students with a range of academic abilities, ED allowed amendments and changed some rules governing the implementation of NCLB in an effort to subdue the growing political opposition to the law. Making these changes were clearly in response to the increasing state and local opposition to the law. As the level of political opposition from the states intensified, ED loosened some of the key accountability requirements and modified many of the regulations governing the law. It has promised more “flexibility” to states meeting certain core principles and granted waivers to some states and districts. The approach the federal government has taken with NCLB is reminiscent of its response to state implementation of earlier federal legislation that was loosely regulated in response to political opposition.

While some of these changes were in the right direction, the administration has resisted addressing many of the underlying flaws in the law, including the double counting of students in some subgroups for accountability purposes, the reliance on mean proficiency to determine AYP, arbitrary timelines for improving achievement and unrealistic achievement goals that have no connection to what can actually be achieved, a reliance on testing and sanctions to improve schools without corresponding attention to the resources or expertise schools need, and insufficient attention in the law to state capacity to turn around huge numbers of failing schools and districts. The result has been a process that lacks transparency and consistency across states designed to decrease, at least temporarily, the number of schools and districts identified for improvement. Some of the changes, such as the change to using the grade span method to identify districts for improvement, have compounded the flaws in the NCLB accountability provisions by making it harder for some districts, primarily those serving minorities, to make AYP. Others, such as the changes in the highly qualified teacher requirements benefit some regions of the country over others. The combination of statistical techniques that can now be used to calculate AYP add complexity to the states’ accountability systems and complicates understanding what AYP and accountability means. Finally, the accumulation of rule changes and amendments create accountability systems that have no common meaning across states and little relationship to actual student achievement.

Instead of correcting the policy errors in the design of the law, the administration has called them the “bright lines” or core principles of NCLB (U.S. Department of Education, 2005, November 10). These include improving achievement by focusing on accountability and assessments, attaining 100% proficiency by 2014, the use of market mechanisms (supplemental educational services and choice) to improve school performance, an emphasis on the collection and reporting of data, and maintaining a goal of having 100% of teachers “highly qualified.” It portends to use these core principles in deciding which states will be given additional flexibility with NCLB, yet it is far from clear what will count as evidence in meeting these core principles or how this evidence will be used. Additionally, many of the core principles articulated by ED reflect an

ideological approach to school improvement rather than something based on sound, scientific research. Supplemental educational services are a case in point. There is no body of research providing clear and consistent evidence documenting the effect of supplemental educational services on learning outcomes for low-income and minority students, the students most affected by this program.

The scope of the changes allowed by ED also raises questions about the ability of the federal government to intervene in core educational functions traditionally governed by states and local districts. While the politics of implementing NCLB have centered on the level of funding provided by the federal government to meet the costs of implementing NCLB and opposition to the expansion of federal power over state authority to govern education, there has been less debate about the broader question of which functions the federal government can most effectively perform. Our research supports a federal role in civil rights enforcement, and in providing resources for data collection and dissemination, and for fostering innovation and research on new programs, curriculum, and instructional practices. It also supports disaggregating data by subgroups and accountability that reflects a realistic research based set of goals, is based on more than just test scores in two subjects, and informs instructional practice. Other research by the Civil Rights Project has strongly supported graduation rate accountability for all groups of students. However, the process of amending state accountability plans described here shows the unraveling of NCLB and raises questions about the feasibility of regulating core educational processes from Washington, DC.

Some of the rule changes introduced by ED begin to address some of the flaws in the law. Allowing states to use an Index and the tentative move towards a growth model to calculate AYP have potential to address some of the flaws in the law and the negative incentives created by NCLB. However, these positive changes will encounter similar challenges as long as the timeline for reaching 100% proficiency and a reliance on a test-based accountability remain in place. Many other changes, particularly those designed to temporarily reduce or delay the number of schools and districts identified for improvement, begin to dismantle the law without addressing its underlying flaws. To do that will require confronting the very principles the administration has identified as its “bright lines.” To improve NCLB, policymakers need to:

- Reexamine the core assumptions that underlie NCLB, including the assumption that market mechanisms alone can improve schools, the idea that external accountability and the imposition of sanctions will force schools to improve and motivate teachers to change their instructional practices, the premise that schools can make large and rapid improvements in student achievement, and the view that standardized tests are valid, related to the school curriculum, and fair to all students, among others.
- Reexamine the mechanisms used by NCLB to improve schools and student achievement, including the reliance on test-based accountability, arbitrary timelines and unrealistic goals for improving student achievement and teacher quality, negative (and unproven) sanctions to improve schools, and an

accountability measure (AYP) that lacks validity as a measure of school effectiveness and has a disparate impact on diverse schools and schools serving large numbers of minority and low-income students.

- Include educators in an open and honest debate about what is needed to reform schools and improve student achievement, about what is productive and what is counterproductive, and what resources are required to reach NCLB's goals. This should include a realistic assessment of state capacity to turn around huge numbers of failing schools and districts.

If these issues are debated and addressed, it will mean a major overhaul of NCLB and not just a tinkering at the edges in response to political pressure.

REFERENCES

- Abedi, J. (2004). The No Child Left Behind Act and English language learners: Assessment and accountability issues. *Educational Researcher*, 33(1), 4-14.
- Banchero, S. (2005, September 2). Tutoring waiver for Chicago a 1-time deal: Other Illinois districts must follow U.S. rules. *Chicago Tribune*.
- Batt, L., Kim, J., & Sunderman, G. L. (2005). *Limited English proficiency students: Increased accountability under NCLB*. Cambridge, MA: The Civil Rights Project at Harvard University.
- California Department of Education. (2005, August 23). *State of California consolidated state application accountability workbook*. Sacramento, CA: Author.
- Center on Education Policy. (2004). *From the capital to the classroom: Year 2 of the No Child Left Behind Act*. Washington, DC: Author.
- Center on Education Policy. (2005, November 16). *States test limits of federal AYP flexibility*. Washington, DC: Author.
- Coltrane, B. (2002, November). *English language learners and high-stakes tests: An overview of the issues*. Retrieved May 13, 2004, from <http://www.cal.org/ericell/DIGEST>
- Davis, M. R. (2005, February 9). Utah is unlikely fly in Bush's school ointment. *Education Week*, pp. 1, 21.
- Dillon, S. (2005, April 20). Utah vote rejects parts of education law. *New York Times*, p. 14.
- Elmore, R. F. (2004). *School reform from the inside out: Policy, practice and performance*. Cambridge, MA: Harvard Education Press.
- Erpenbach, W. J., & Forte, E. (2005, October). *Statewide educational accountability under the No Child Left Behind Act: A report on 2005 amendments to state plans*. Washington, DC: Council of Chief State School Officers.
- Florida Department of Education. (2005, April 1). *Letter to the Honorable Margaret Spellings*. Retrieved October 13, 2005, from http://www.fldoe.org/NCLB/AYP_Amend_Letter_2005_04_01.pdf
- Frahm, R. A. (2005, April 9). State takes on feds: Officials deny they are 'un-American' in education policy. *Hartford Courant*.
- General Accounting Office. (2002). *Title I: Education needs to monitor state's scoring of assessments* (No. GAO-02-393). Washington, DC: Author.
- General Accounting Office. (2003). *No Child Left Behind Act: More information would help states determine which teachers are highly qualified*. (No. GAO-03-631). Washington D.C.: Author.
- Georgia Department of Education. (2005, June 28). *State of Georgia: Consolidated state application accountability workbook, revised for 2004-05*. Atlanta, GA: Author.
- Georgia Professional Standards Commission. (2002). *The Georgia educator workforce 2002: A report of the supply, demand, and utilization of teachers, administrative, and student services personnel in Georgia public schools*. Atlanta: Author.
- Gewertz, C. (2005a, January 5). Chicago resisting federal directive on NCLB tutoring. *Education Week*, pp. 1, 15.
- Gewertz, C. (2005b, February 9). Chicago, Ed. Dept. settle tutoring dispute. *Education Week*, pp. 3, 11.

- Gewertz, C. (2005c, September 1). Ed. Dept. allows Chicago to continue NCLB tutoring. *Education Week*.
- Government Accountability Office. (2005). *No Child Left Behind Act: Improved accessibility to Education's information could help states further implement teacher qualification requirements*. (No. GAO-06-25). Washington, DC: Author.
- Grant, A. (2005, July 11). TAKS exemptions limited to 3 percent. *Houston Chronicle*, p. B 3.
- Hoff, D. J. (2005, March 9). Texas stands behind own testing rule. *Education Week*, pp. 1, 23.
- Illinois State Board of Education. (2002). *Educator supply and demand in Illinois: 2002 annual report*. Springfield, IL: Author.
- Implementing federal educational programs, Utah State Legislature, 2005 First Special Sess.(2005).
- Ingersoll, R. M. (2001). *Teacher turnover, teacher shortages, and the organization of schools*. Seattle, WA: Center for the Study of Teaching and Policy, University of Washington.
- Jimerson, L. (2004, January). *Special challenges of the "No Child Left Behind" act for rural schools and districts*. Retrieved September 29, 2005, from <http://www.ruraledu.org/docs/nclb/jimerson.pdf>
- Kane, T. J., & Staiger, D. O. (2002). Volatility in school test scores: Implications for test-based accountability systems. In D. Ravitch (Ed.), *Brookings papers on educational policy, 2002* (pp. 235-283). Washington, DC: Brookings Institution Press.
- Keller, B. (2005, December 14). Actual measure of 'highly qualified' teachers just beginning to come to light across nation: First and second rounds of reported data based largely on guess work. *Education Week*, pp. S6, S8, S9.
- Kim, J., & Sunderman, G. L. (2005). Measuring academic performance under NCLB: Implications for educational equity. *Educational Researcher*, 34(8), 3-13.
- Lindsay, J. (2005, November 3). *Boston schools allowed to continue tutoring struggling students*. Retrieved November 7, 2005, from http://www.boston.com/news/education/k_12/articles/2005/11/03/boston_schools_allowed..
- Linn, R. (2000). Assessments and accountability. *Educational Researcher*, 29(2), 4-16.
- Linn, R. L. (2003). Accountability: Responsibility and reasonable expectations. *Educational Researcher*, 32, 3-13.
- Linn, R. L., Baker, E. L., & Betebenner, D. W. (2002). Accountability systems: Implications of requirements of the No Child Left Behind Act of 2001. *Educational Researcher*, 31, 3-16.
- Losen, D. (2005). Graduation rate accountability under the No Child Left Behind Act. In G. L. Sunderman, Kim, J. S., & Orfield, G. (Ed.), *NCLB meets school realities: Lessons from the field*. Thousand Oaks, CA: Corwin Press.
- Lynn, R. (2005, April 20). Utah bucks feds on schools. *The Salt Lake Tribune*.
- Marks, A. (2005, April 19). States hit back on school reform law. *The Christian Science Monitor*, p. 1.
- Matus, R. (2005, May 17). Education official lets Florida loosen rules. *St. Petersburg Times*, pp. 5, B.

- Nathan, R. P., Gais, T. L., & Fossett, J. W. (2003, November 7). *Bush federalism: Is there one, what is it, and how does it differ?* Paper presented at the Association for Public Policy Analysis and Management, Washington, DC.
- National Conference of State Legislatures. (2005). *Delivering the promise: State recommendations for improving No Child Left Behind*. Washington, DC: Author.
- Olson, L. (2004, August 11). Critics float 'No Child' revisions. *Education Week*, 23, 1, 33.
- Olson, L. (2005a, May 25). Florida gains flexibility on NCLB provisions. *Education Week*, 24, 18, 20.
- Olson, L. (2005b, October 26). 'Value added' models for gauging gains called promising. *Education Week*, 25, 11.
- Raudenbush, S. W. (2004). *Schooling, statistics, and poverty: Can we measure school improvement?* Princeton, NJ: Educational Testing Service.
- Richard, A. (2003a, April 2). Montana leads choir of rural concerns over 'No Child' law. *Education Week*, p. 26.
- Richard, A. (2003b, October 22). Rural states concerned about federal law. *Education Week*, pp. 25, 27.
- Sack, J. L. (2005, April 27). Utah passes bill to trump 'No Child' law. *Education Week*, 24, 22, 25.
- Saulny, S. (2005, November 8). Restriction is lifted on funds for tutors in failing schools. *The New York Times*.
- Spellings, M. (2005, April 7). *Raising achievement: A new path for No Child Left Behind; prepared remarks for Secretary Spellings at Mount Vernon, VA*. Retrieved April 7, 2005, from <http://www.ed.gov/print/news/speeches/2005/04/04072005.html>
- Spellings, M. (2005, April 18). Letter to Senator Orrin G. Hatch. Washington, DC.
- Spellings, M. (2005, November 18). *Secretary Spellings announces growth model pilot, addresses chief state school officers' annual policy forum in Richmond*. Retrieved November 11, 2005, 2005, from <http://www.ed.gov/print/news/pressreleases/2005/11/11182005.html>
- Spellings, M. (2005, October 21). *Letter to chief state school officers regarding the "highly qualified teacher" provisions of the No Child Left Behind Act*. Retrieved November 2, 2005, from <http://www.ed.gov/policy/elsec/guid/secletter/051021.html>
- Sternberg, B. J. (2005). *Commissioner Sternberg's letter to Secretary Margaret Spellings*. Retrieved May 6, 2005, from <http://www.state.ct.us/sde/nclb/correspondence/spellings.pdf>
- Stutz, T. (2005, February 26). Texas defies No Child law on special ed. *Dallas Morning News*.
- Sunderman, G. L. (1995). *The politics of school reform: The educational excellence movement and state policymaking*. Unpublished Dissertation, The University of Chicago, Chicago, IL.
- Sunderman, G. L., & Kim, J. (2004). *Expansion of federal power in American education: Federal-state relationships under the No Child Left Behind Act, year one*. Cambridge, MA: The Civil Rights Project at Harvard University.

- Sunderman, G. L., & Kim, J. (2005). *Teacher quality: Equalizing educational opportunities and outcomes*. Cambridge, MA: The Civil Rights Project at Harvard University.
- Sunderman, G. L., Kim, J. S., & Orfield, G. (2005). *NCLB meets school realities: Lessons from the field*. Thousand Oaks, CA: Corwin Press.
- Superfine, B. M. (2005). The politics of accountability: The rise and fall of Goals 2000. *American Journal of Education*, 112(1), 10-43.
- The Education Trust. (2003). *Telling the whole truth (or not) about highly qualified teachers*. Washington, DC: Author.
- The Rural School and Community Trust. (2004, May 13). *Most rural students left behind under Department of Education's new "flexibility" rules regarding highly qualified teachers*. Retrieved September 29, 2005, from http://www.ruraledu.org/issues/nclb/HQT_Flex_factsheet.pdf
- Tracey, C. A., Sunderman, G. L., & Orfield, G. (2005). *Changing NCLB district accountability standards: Implications for racial equity*. Cambridge, MA: The Civil Rights Project at Harvard University.
- U.S. Department of Education. (2002). Title I--Improving the academic achievement of the disadvantaged; Final rule. *Federal Register*, 67(231).
- U.S. Department of Education. (2003). *Meeting the highly qualified teachers challenge: The secretary's second annual report on teacher quality*. Washington, DC: Author.
- U.S. Department of Education. (2003, December 11). *Title I regulation on alternate achievement standards*. Retrieved September 20, 2005, from <http://www.ed.gov/print/policy/elsec/guid/secletter/031211.html>
- U.S. Department of Education. (2004, February 19). *Fact sheet: NCLB provisions ensure flexibility and accountability for limited English proficient students*. Retrieved March 4, 2004, from <http://www.ed.gov/print/nclb/accountability/schools/factsheet-english.html>
- U.S. Department of Education. (2004, January 16). *Improving teacher quality state grants Title II, part A*. Washington, DC: Author.
- U.S. Department of Education. (2004, October 4). *Identification of districts for improvement*. Retrieved June 7, 2005, from <http://www.ed.gov/print/policy/elsec/guid/stateletters/indi.html>
- U.S. Department of Education. (2005, April 7). *Raising achievement: A new path for No Child Left Behind*. Retrieved April 7, 2005, from <http://www.ed.gov/print/policy/elsec/guid/raising/new-path-long.html>
- U.S. Department of Education. (2005, August 3). *Highly qualified teachers and improving teacher quality state grants non-regulatory guidance*. Washington, DC: Author.
- U.S. Department of Education. (2005, March 29). *Secretary Paige issues new policy for calculating participation rates under No Child Left Behind*. Retrieved June 7, 2005, from <http://www.ed.gov/print/news/pressreleases/2004/03/03292004.html>
- U.S. Department of Education. (2005, May 10a). *The pursuit of raising achievement: A new path for No Child Left Behind, May 10, 2005*. Retrieved May 11, 2005, from <http://ed.gov/print/policy/elsec/guid/raising/disab-amendment.html>

- U.S. Department of Education. (2005, May 10b). *Spellings announces new special education guidelines, details workable, "common-sense" policy to help states implement No Child Left Behind*. Retrieved May 11, 2005, from <http://ed.gov/print/news/pressreleases/2005/05/05102005.html>
- U.S. Department of Education. (2005, November 10). *No Child Left Behind: A road map for state implementation, pre-publication copy*. Retrieved November 23, 2005, from <http://www.ed.gov/admins/lead/account/roadmap/index.html>
- Vinovskis, M. A. (1999). Do federal compensatory education programs really work? A brief historical analysis of Title I and Head Start. *American Journal of Education*, 107, 187-209.
- Virginia Department of Education. (2000, November). *Report on supply and demand of instructional personnel in Virginia: 1999-2000*. Richmond, VA: Author.