

CARROTS, STICKS, AND THE BULLY PULPIT

*Lessons from a Half-Century of Federal
Efforts to Improve America's Schools*

Edited by

FREDERICK M. HESS

ANDREW P. KELLY

2011

Harvard Education Press
Cambridge, Massachusetts

Stepping-Stones to Success or a Bridge Too Far?

The Federal Role in Educational Accountability

Paul Manna and Jennifer Wallner

I ncreasing accountability for educational performance was one of the main objectives of the No Child Left Behind Act (NCLB). Now essentially a decade into the law's implementation, NCLB's track record raises fundamental questions about the federal government's ability to improve the nation's schools and the lives of its students. In which ways can the federal government play a constructive role in helping to realize those goals? Where does federal assertiveness become federal overreach that produces more harm than good? In this chapter, we focus on three aspects of accountability that NCLB has embraced: ensuring that subnational governments (for example, states and localities) produce policy outputs consistent with federal mandates; fostering educational outcomes that create better opportunities for the nation's students; and, ultimately, promoting substantive outcomes that result in high-quality educational experiences for all. We explore those three areas, what we call the crucial stepping-stones of accountability, in the context of NCLB's implementation.

Overall, we find that federal policy makers have had some success with NCLB in holding subnational governments accountable for adopting certain policies that hew to federal law. But several barriers have prevented NCLB from producing the desirable outcomes to improve educational opportunity and quality that its advocates had envisioned. As our analysis makes clear, this is not simply a tale of NCLB loyalists doing battle with the law's skeptics. The story is more nuanced than that. Laying down what we call the stepping-stones of accountability is challenging work, especially

when done to produce substantively important outcomes that have eluded previous generations of policy makers.

By analogy, consider an intrepid hiker trying to cross a body of water. The hiker would need little or perhaps no assistance to cross a small stream or shallow tributary—a sturdy set of boots would do just fine. Wider rivers require dependable stepping-stones that provide a safe path above the rushing current. Similarly, the nation's experience with NCLB has shown that it becomes more difficult to use federal policy to promote educational accountability as the ultimate goals of policy become more distant. Encouraging state or local policy changes, what we call *outputs*, is a more proximate result that requires less effort than enhancing educational opportunities or, more demanding still, producing substantively valuable educational *outcomes*. As one moves across these three stepping-stones—from outputs, to opportunity, to results—the challenges become more difficult, and federal policy makers struggle to promote more consistent levels of success.

In practice, wielding federal power to improve educational quality, at least in the form that NCLB has taken, may be a bridge too far. We develop that argument by describing our three stepping-stones of accountability, situating them alongside accountability efforts in other countries, and relating them to NCLB's ambitious implementation goals. In the process, we derive several lessons about the federal government's ability to advance policies that foster educational accountability.

CONCEPTUALIZING ACCOUNTABILITY

It can be difficult to conceptualize accountability because the term carries many meanings. It is, as one author notes, "one of those evocative political words that can be used to patch up a rambling argument, to evoke an image of trustworthiness, fidelity, and justice, or to hold critics at bay."¹ With humble origins in bookkeeping, accountability has evolved into a powerful symbol for good governance to advance, among other things, fair management of public affairs, effective evaluation of processes and outcomes, and a widespread improvement of public services. It is also indelibly connected to other politically desirable features such as responsibility, efficiency, integrity, and equality. Collectively, these concepts represent a formidable collection of goals and objectives nested beneath a single conceptual umbrella.

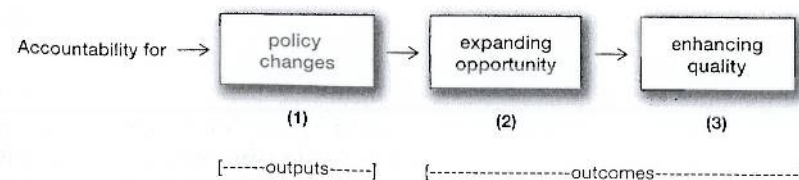
At its core, accountability requires people to answer to others for their actions and the ensuing results they produce. It ensures an adherence to policies and practices to see that goals are effectively met and that procedures are sufficiently followed. Furthermore, accountability can be a tool

to improve performance and elevate the quality of programs and practices through systematic evaluation. In general, three conditions help promote accountability. First, expectations must be clear to those who are required to act. Second, the expectations should be reasonable and accompanied with a means of evaluating them. Effective information is critical here so that overseers can accurately determine whether their agents are following procedures and accomplishing the desired goals. Finally, clear consequences (positive or negative or both) must follow in light of these evaluations.

In the political realm, elections are the most obvious manifestation of accountability, with the electorate empowered to "throw the rascals out" when political representatives fail to meet expectations or fulfill campaign promises. In the policy world, however, identifying accountability relations and the mechanisms for answerability are less clear. Especially in social policy, the delivery of a single program can involve numerous agents and organizations, meaning that responsibilities are diffused and program goals may be either misinterpreted or contested by the various participants involved.² Furthermore, traditional accountability provisions have tended to focus on either the supportive policy outputs or the procedures that implementers must follow as they use those outputs to accomplish goals. In education, for example, the traditional emphasis is accountability focused on the answerability of schools and districts for state and federal funds and on complying with various regulations for things such as directing dollars to needy students or maintaining certain class sizes. This focus on compliance consequently prioritized obeying bureaucratic rules and producing certain policy outputs over a more intense focus on substantive outcomes.

The distinction between policy outputs and outcomes, and the accountability challenges associated with each, is at the core of our argument.³ Figure 7.1 begins to illustrate these key concepts, which form the stepping-stones of accountability. In the context of NCLB, the figure considers three different areas in which the federal government might be interested in holding subnational governments accountable. Moving from the first to third

FIGURE 7.1 Three stepping-stones of accountability



stepping-stone presents an increasingly difficult set of challenges for federal overseers.

The first stepping-stone focuses on policy outputs, which are the concrete actions that subnational governments might take as they implement federal policy. Here federal officials focus on whether states or localities have established policies, rules, regulations, and budgets to be consistent with federal expectations. The next two stepping-stones focus on different substantive policy outcomes, which reveal whether federal, state, and local policies have actually provided program participants with something valuable. Put differently, outputs focus on how policies have changed, while outcomes consider how the world has changed due to those outputs.⁴ Thus, the second stepping-stone of accountability focuses on the substantive outcome of guaranteeing opportunities to individuals eligible for federal programs or other benefits. Federal officials holding others accountable for such opportunities would focus on the availability and scope of benefits offered across subnational governments and rates of program participation. The third stepping-stone looks beyond the outputs realized on the first step and the opportunities of the second step to consider whether, ultimately, these policies and opportunities have improved the lives of program beneficiaries.

As prior research on performance management has shown, it is relatively easier for overseers to hold others accountable for producing required policy outputs, relatively more difficult to make sure that those outputs produce meaningful opportunities, and more difficult still to ensure that those opportunities improve the quality of individuals' lives.⁵ The increasing array of challenges associated with constructing such a multilevel accountability regime—akin to strategically laying a series of stepping-stones across a wide and fast-moving river—is especially apt for schooling where the ability to measure performance is difficult and where deep disagreements persist over the elements of a quality education.⁶ The following section, which connects the framework of figure 7.1 with accountability in education more generally and the specific American experience with NCLB's design, begins to explain why. As our discussion will reveal, federal policy makers who established and implemented NCLB were interested in promoting accountability for all of the steps along the path that figure 7.1 outlines.

Regardless of the stepping-stone, prior research on public management and organizational performance has identified three empirical realities that influence how accountability regimes, including those in education, are likely to operate across levels of government.⁷ First, unmeasured activities tend to receive low priority in accountability arrangements. The systematic collection of standardized information that is effectively aligned to the

identified goals and priorities is thus a crucial component of accountability. However, authorities must also attend to and use the collected information. As one report from the Organization for Economic Cooperation and Development (OECD) stated, "The collection of information is perhaps of little use if nothing is done with that information."⁸ Seeing that information motivates effective actions must therefore be integrated into any meaningful accountability regime.

Second, enforcing accountability is challenging when a higher level of government relies heavily on lower levels to administer and implement various programs. Federal education policy in the United States is a perfect example. Although leaders in Washington have become increasingly interested and active in K-12 education during the last two decades, federal education policies continue to be administered by states and local school districts.⁹ The lengthy delegation chains that result, which themselves are embedded in complex networks of governmental and nongovernmental organizations that attempt to influence education policy, make it difficult for leaders at the top to hold their agents accountable. From the bottom looking up, ground-level school administrators and teachers can find themselves cross-pressured by directives of authorities at state and federal levels that attempt to hold them accountable for obeying certain policy rules, offering students opportunities, and producing high-quality teaching and learning.

Finally, all actors involved in implementing accountability systems must believe these systems are credible. Political leaders and agency executives who use them to inform their policy making and distribution of public resources, as well as public managers and street-level employees who are subject to their requirements and consequences, all must have confidence in the results that accountability systems produce. Desirable outputs, outcomes, and continuous organizational improvements are unlikely to occur when that confidence is lacking. Accountability mechanisms developed with little understanding of how policies are implemented on the ground run the risk of producing results that are substantively vacuous, which in turn can prompt opposition and undercut the goals that advocates of accountability hoped to accomplish in the first place.

EDUCATIONAL ACCOUNTABILITY AROUND THE GLOBE

During the last twenty years, worldwide, politicians, other public officials, and advocates outside government have tried to improve the quality of information about education available to wide audiences. Driven by the belief that economic competitiveness rests on citizens' skills, governments have redoubled their efforts at improving schooling outcomes and elevating the

overall quality of their education systems. This trend, combined with the fact that public education is one of the largest expenditure items on government balance sheets, has prompted increasing concerns about educational quality. As a result, the tradition of maintaining accountability systems for policy outputs (stepping-stone 1 in figure 7.1) and promoting greater opportunities (stepping-stone 2) has continued with an added emphasis on producing quality academic outcomes (stepping-stone 3).¹⁰ Considering NCLB's embrace of accountability in light of these broader global changes is instructive because it illustrates how governments around the world have approached accountability across the three stepping-stones.

Developments north of the forty-ninth parallel are especially instructive given that the United States and Canada are both federal systems that empower midlevel (state or provincial) governments with much responsibility for education in the intergovernmental system. Canadian provinces are entirely responsible for elementary and secondary schooling, but, unlike in the United States, local board authority in Canada is much more heavily circumscribed under the auspices of provincial leadership.¹¹ Canadian provinces have all but eliminated the role of local boards in education funding, installed centrally developed curriculum, instituted centrally designed tests, and developed a pan-Canadian assessment program designed to evaluate the respective educational achievements of the provinces and territories. Additionally, the provinces provide representative samples for international tests, such as the OECD's Programme for International Student Assessment (PISA), further elevating their ability to measure and compare results across the country. Each of the provinces has instituted policies that require disabled students to be included in regular classrooms, compel local boards to hire provincially certified teachers, and assess the relative achievements of schools within each jurisdiction. These programmatic features share much in common with many American initiatives, even some elements of NCLB. However, and importantly, none involves the federal government. Federal leaders in Ottawa play no role in K-12 education in Canada. Rather, it is the provinces, acting sporadically in concert through the Council of Ministers of Education Canada (CMEC), that have independently fashioned this accountability architecture in education.

Casting a wider net, accountability systems exist in other countries across the OECD, but with markedly different configurations and national responsibilities. In contrast to the Canadian structure, the French national government has installed a powerful board that plays a major role in defining educational accountability expectations, drafting assessment regimes and ranking schools according to their performance on a national curriculum. England has a national curriculum accompanied by assessments and

measures for rating schools. Hong Kong, China, and Japan use national assessments to measure students and school progress and to inform their policy decisions. To receive federal funding, Australian schools must conduct standardized tests written by Canberra officials. This brief snapshot demonstrates that diverse intergovernmental arrangements can promote accountability in education. However, the common characteristic across all these systems is the explicit definition of what students must know and how well they must know it.¹² A further key is that one administrative unit, either at the national or subnational level, is responsible for setting standards and ensuring compliance with the expectations for certain policy outputs or the promotion of desired outcomes.

However, not all countries have followed the route of centralization, either at the subnational or national level. Finland, a country renowned for its educational achievements, empowers local schools and individual teachers.¹³ The Finnish approach to education reform has not embraced test-based accountability and externally determined learning standards. Instead, the country has opted for school-based curriculum development, individualized learning practices, and the adoption of "intelligent accountability policies and gradual building of a culture of trust within the education system that values teachers' and headmasters' professionalism."¹⁴ These arrangements are viable in part because teachers in Finland are highly trained and must hold a master's degree to gain permanent employment. Furthermore, the country has a fairly homogeneous population with little economic inequality, which eases the complexity of schooling the citizenry.

As in the United States with NCLB, across these diverse nations student achievement results (a measure of educational quality) are not simply examined as broad aggregates. National systems also assess educational opportunities. Globally, no longer are nations content with simply holding educational providers accountable for allowing educational access or equal treatment, a focus of the second stepping-stone. Those opportunities and policy outputs that created them must also produce meaningful academic gains for students, the focus of the third stepping-stone. Holding schools or school districts accountable for the dual outcomes of providing opportunities and enhancing quality requires the disaggregation and study of student achievement results across diverse populations including students from families with low incomes, those with disabilities, and minority groups based on language or race and ethnicity. The effectiveness of these initiatives, however, turns critically on the data collection and analytical capacities of the respective countries.

Recognizing the importance of consistent monitoring, many countries also have instituted nationally or regionally regulated school inspections

and evaluation programs. Eighteen countries in the OECD, for example, have a specific national or regional school inspectorate tasked with evaluating individual schools.¹⁵ In most cases, these accountability regimes are supplemented by requirements for schools to conduct periodic self-evaluations. These inspectors help to systematize and institutionalize the assessment of school performance. Certainly, the quality of these evaluations and their ability to inspire confidence depend on the capabilities and diligence of the people doing the work. Regardless of the organization and execution, they are based on the idea that central authorities should take stock of and regularly assess work of subnational actors. To date, the United States has not adopted a similar arrangement of uniform school inspections that produces comparable results across the country.

ACCOUNTABILITY OBJECTIVES OF NCLB

The broad theories of action and specific requirements that U.S. federal officials embraced with NCLB reveal parallels and contrasts with developments in other nations. Consistent with global trends is NCLB's effort to incorporate concerns about accountability across all three stepping-stones. Under NCLB's provisions, Washington aimed to bring some uniformity to the uneven approach that previously persisted in K-12 accountability while paying more attention to student outcomes (both opportunity and quality) and still preserving prior concerns about policy outputs. The law did encourage a transformation in accountability measures from a primary emphasis on outputs to new approaches that embraced outputs and outcomes simultaneously.

Before NCLB, some local districts had developed policies to hold their individual schools accountable, some states had promulgated laws to hold both districts and schools accountable, and other federal policies had been enacted to hold governments at all levels accountable. Collectively, this culminated in a morass of policies ostensibly designed to provide answers about education to people inside and outside government. What is more, the bulk of these early provisions prioritized adherence to bureaucratic rules and audits associated with receiving government grants rather than drawing attention to the performance of the education system itself. In crafting NCLB, the law's authors and advocates stated their desire to move past accountability dominated by a focus on the production of policy outputs to focus more squarely on outcomes that generated real opportunities and higher academic quality.¹⁶

Given that NCLB was an expansive law with several dozen program authorizations and numerous moving parts, we delimit our discussion to

a few key elements to better understand the federal government's role in promoting educational accountability.¹⁷ The areas we consider appear in table 7.1 along with their relationship to the three stepping-stones of accountability. In this section we discuss each area of concern and its particular manifestation in NCLB.

Outputs. Although NCLB's various elements required states and local governments to take several specific actions, the law's substantive core focused on two main policy outputs. First, states were to develop comprehensive plans to hold schools accountable for academic performance. Among other things, this meant that states had to define reading and math standards and then test students against those standards each year in grades 3 through 8 and once in high school. Those results were to be disaggregated by student subgroups so that the data would reveal whether students of different racial and ethnic backgrounds, in addition to subgroups based on disability status and knowledge of the English language, were making academic progress. States then were to use those test scores, along with other indicators, to rate schools for Adequate Yearly Progress (AYP). Schools failing to make AYP for a specified number of consecutive years were subject to an escalating cascade of federally defined remedies. The states would work with their local districts and schools to ensure that the remedies were implemented.

Second, NCLB required states to develop a menu of policies addressing teachers. Specifically, states needed to create requirements to guarantee that all teachers in core subjects were "highly qualified," meaning they all possessed at least a bachelor's degree and full state certification, and

TABLE 7.1 Educational accountability and NCLB

Area of federal concern	Examples from NCLB
Outputs	<ul style="list-style-type: none"> • States develop systems to hold schools accountable • States develop systems to define high quality teachers • States establish plans to ensure that high quality teachers are evenly distributed across a state's schools
Outcomes—Opportunity	<ul style="list-style-type: none"> • Students attending schools not making yearly progress can <ul style="list-style-type: none"> – transfer to other schools – receive supplemental educational services (e.g., tutoring) • Students' teachers meet the highly qualified standard • Highly qualified teachers are equitably distributed
Outcomes—Quality	<ul style="list-style-type: none"> • Students learn rigorous academic content • Student achievement increases • Gaps in student achievement narrow or even close • Highly qualified teachers have deep knowledge and skills

that the state maintained a system to ensure that they knew their subject matter through, for example, teacher testing. Additionally, states needed to develop plans to ensure fair distribution of highly qualified and experienced teachers across the state so that certain student groups, typically disadvantaged ones, had access to these teachers as NCLB, via state policy, defined them.

Outcomes—Opportunity. The authors and advocates of NCLB also were interested in holding states, districts, and schools accountable for promoting educational outcomes to enhance educational opportunities for the nation's students, especially those at risk of falling behind. We focus on three specific opportunities, two of which were tied to the list of remedies for schools failing to make AYP. A school struggling to meet yearly performance goals for two consecutive years, and the school district in which the school resided, was to provide opportunities for students to transfer to other schools that were making AYP. Further, if the school missed AYP for another year in a row, then students were to receive additional academic help in the form of free out-of-school tutoring. Finally, faithful implementation of the teacher quality plans was to ensure that all students enjoyed opportunities to learn from teachers who were experienced and highly qualified.

Outcomes—Quality. Flowing from the policy outputs and educational opportunities that NCLB was designed to create, the law's authors also expected educational quality to improve. The most important quality outcomes focused on student achievement with the expectation that as students learned more rigorous content, overall achievement would increase and gaps between student groups would narrow or even disappear. The law set a deadline: by 2014, all students would be proficient in reading and math. Further, the law's authors intended for NCLB's teacher quality requirements to ensure that the nation's teachers did indeed possess deep subject-matter knowledge and were prepared to do their jobs well.

In sum, NCLB's approach to accountability built on prior federal policies while simultaneously extending the federal government's reach in education. Requiring states to develop detailed plans for accountability and teacher quality, a focus of the first stepping-stone, was consistent with several prior decades of federal education policy, which had typically devoted much attention to policy outputs and whether states had met specific legal and regulatory requirements. The concerns in NCLB regarding student opportunities were also consistent with prior federal efforts, such as those to outlaw racial segregation and to increase access to the classroom for stu-

dents with disabilities. Federal concerns about outcomes to promote educational quality, at least on the scale contemplated in NCLB, represented a more recent federal ambition, building on and dramatically extending expectations from the prior Elementary and Secondary Education Act (ESEA) reauthorization of 1994.

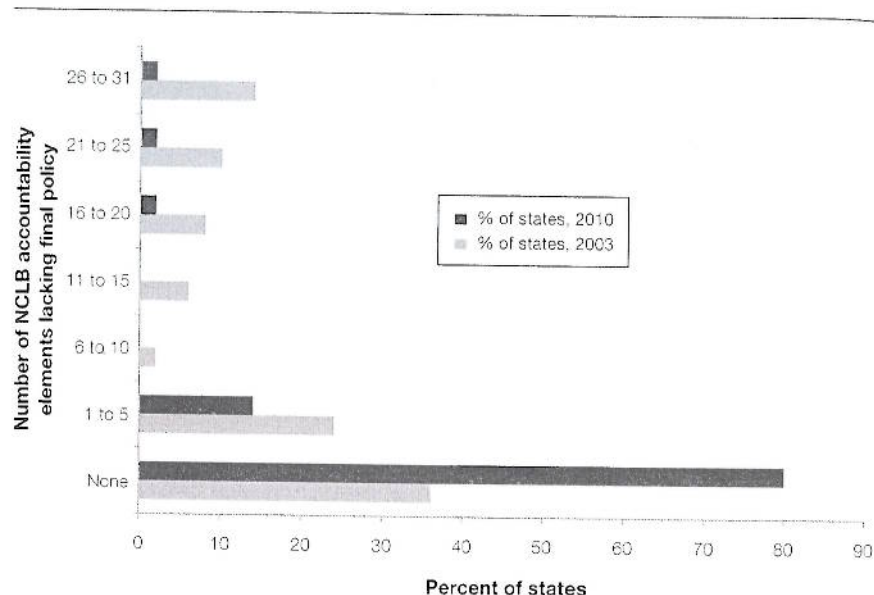
ACCOUNTABILITY AND NCLB'S IMPLEMENTATION

What can the nation's experience with NCLB reveal about the federal government's ability to promote accountability in elementary and secondary education? We offer four specific lessons that draw on evidence from the law's implementation and the three stepping-stones of accountability.

Federal education funds and the requirements that accompany them can prompt changes in state policy outputs. Authors of federal grant programs have long embraced the theory of action that grant funds can entice lower levels of government to embrace federal priorities. This view has been especially prevalent in federal social policies—such as the provision of public assistance, health care, and education—in which Washington relies heavily on subnational governments to develop the policy infrastructure and make the implementation choices that ultimately will help accomplish federal goals.¹⁸ In education, much historical evidence suggests that state and local policy outputs frequently respond to the demands that accompany federal dollars. For example, although the U.S. Supreme Court receives much credit for ending legal segregation of schools in its famous *Brown v. Board of Education* decision of 1954, measurable changes in policies to enhance desegregation only began to gain momentum once federal education assistance became available and tied to such policy changes. Thus, while *Brown v. Board* is properly considered a historic decision, federal grant funds were crucial for producing state and local policy outputs consistent with the court's view that separate was inherently unequal.¹⁹

Evidence from NCLB's implementation also suggests that states respond to federal grant programs by passing policies designed to meet federal program requirements. Figure 7.2 presents one illustration that examines policy development in all fifty states. The figure focuses on the thirty-one policy elements that NCLB required states to include in their district and school accountability plans. Using the most recently amended plans available, the figure illustrates how many of these elements states still had to complete as of 2003, when these plans were originally due, and in 2010.²⁰ If states were generating policies consistent with NCLB's expectations, then we would expect to see most states clustering in the bottom bars in the

FIGURE 7.2 State Adoption of NCLB's Required Accountability Elements, 2003 and 2010



Sources: (1) 2003 data are from Manna (2006) Educational Policy article; (2) 2010 data are from www2.ed.gov/admins/lead/account/stateplans03/index.html, accessed in February 2011.

table, because that would show they have few or no elements needing additional work.

Two main results stand out from figure 7.2. First, as states were attempting to meet the law's early deadlines, some had much difficulty completing final policies for its required accountability elements. Although a plurality of states completed their work by the 2003 deadline, the others demonstrated varying levels of progress, with some having much work still remaining at that time. The spread of the gray bars across the figure illustrates this result. Second, by 2010, nearly all states had either produced completed plans or had plans that were nearly complete. The combined total of the bottom two black bars comprises more than 90 percent of the states. One reason for that progress was the amount of administrative attention and frequent vetting of state plans (and revisions of those plans) that the federal Department of Education has offered since NCLB became law. Holding the states accountable for producing these policy outputs was a top priority for federal officials, who also allocated vital human resources to ensure that federal oversight actually occurred.

Changes in state policy outputs can be sensitive to federal administrative priorities and state capacities. Although the carrot of federal dollars can influence the content of state policy outputs, money alone is not sufficient to prompt policy change. An additional important factor is the federal government's willingness to enforce the requirements that accompany federal dollars. Sometimes resources devoted to enforcement, such as administrative time and money, are more concentrated in some areas and spread thinly in others. An interesting contrast in NCLB that illustrates this difference is the development of state policy outputs to meet the law's testing and accountability provisions versus those required to advance the law's teacher quality objectives.

Since NCLB became law, federal education officials have devoted more enforcement and administrative effort to testing and accountability matters than to teacher quality. In contrast to the results in figure 7.2 that show state progress in developing plans to hold schools and districts accountable for academic performance, states made less progress in developing policy outputs to meet the law's Highly Qualified Teacher (HQT) provisions. In particular, states have been much less responsive to NCLB's requirement that they design plans to ensure that quality and experienced teachers are equitably distributed. Those teacher equity plans were due at the end of the 2005–2006 school year, but at the time twenty-four states failed even to submit them.²¹ Among the twenty-six states that did submit an equity plan, only seven had plans that met the federal requirements, twelve offered plans that partially met the requirements, and the remaining seven provided plans that failed to meet the requirements. Put differently, more than four years after NCLB had become law, thirty-one states still had not submitted a teacher equity plan that even partially met the law's requirements. It is important to note that the 2005–2006 deadline for these plans did not require states to have eliminated inequities in teacher distribution (a focus of the second accountability stepping-stone), only that they had produced a plan to do so (the first stepping-stone).

Stricter federal enforcement of the law's HQT provisions perhaps could have produced policy outputs more consistent with NCLB's expectations. Yet those areas of the law, relative to the issue of state testing and accountability systems, received limited attention from the George W. Bush administration. In part that was because the law's HQT provisions became incorporated into the law during the legislative process, at the urging of Democrats (especially Representative George Miller, D-CA), and were not part of the administration's original NCLB proposal. The policy guidance and adjustments that the administration transmitted to states on this issue sent strong signals that few, if any, consequences would ensue if the states

let slide the development of the policy infrastructure needed to meet the law's HQT requirements. By 2007, the administration was simply expecting states to demonstrate a good-faith effort to do so.²² That approach was in marked contrast to the administration's more assertive efforts with the law's testing and school accountability elements.

Federal policy is better able to promote outcomes that enhance educational opportunity when those outcomes depend primarily on individual subnational governments rather than on the behavior of multiple governments and nongovernmental actors. Moving from the first stepping-stone of accountability in figure 7.1, accountability for policy outputs, to the second, accountability for educational outcomes that enhance opportunity, creates added complexity. The reason is that producing valued outcomes in education and other social policy areas increasingly depends on the ability of federal policy to help mobilize networks of actors. These networks include diverse agencies of subnational governments, extragovernmental organizations in the private and nonprofit sectors, and individual people who are potential program beneficiaries. Simply holding a lower level of government accountable for producing policies consistent with federal requirements does not guarantee that such networks will be effectively cultivated and then mobilized. When the opportunities that federal laws aim to create come to rely heavily on such cultivation and mobilization, they become more difficult to provide.²³

One way to see the increasing difficulty of promoting accountability for outcomes is to contrast two opportunities that federal education policy has attempted to promote. The first area is the federal goal, established in the Individuals with Disabilities Education Act (IDEA), that students with disabilities should have the same opportunities to attend public schools as their nondisabled peers. Despite criticisms of IDEA, one of its undeniable successes has been to dramatically increase in the number of students with disabilities who attend traditional public schools and learn in typical classroom environments.²⁴ Opportunities for traditional classroom instruction have increased dramatically for these students during the four decades that IDEA has existed.

One can compare IDEA's success with NCLB's attempt to increase educational opportunity for students attending schools that receive Title I dollars but consistently miss AYP. NCLB requires that students attending these schools should be able to transfer to other schools making AYP and that they should be able to receive free extra tutoring, known as supplemental educational services. Unlike the opportunities that IDEA created, NCLB has

struggled to produce the opportunities that its authors intended. Few students have taken advantage of these opportunities relative to the numbers that have been eligible. In the 2008–2009 school year, for example, 2.7 percent of eligible students exercised school choice under the law, while 15.6 percent of eligible students received supplemental educational services.²⁵

An important reason for the difference is the network of actors required to guarantee the envisioned opportunities and the avenues for redress when the actors fail to comply. The principal actors needed to achieve IDEA's objectives of fuller participation have been school districts, which define local enrollment and attendance policies, and parents of students with disabilities, who have long-expressed an interest in having their children afforded the same opportunities as the nondisabled. Furthermore, when districts failed to comply with IDEA, the law empowered parents to take them to court to secure a legal remedy to gain the desired opportunity. This is a relatively simple network where responsibilities and authority are clear.

Providing opportunities via NCLB's public school choice and supplemental services provisions has required the operation of more complicated networks. In large urban districts, where in some cases nearly all schools have missed AYP, the ability to transfer to a more successful school depends on several factors. These include the degree to which district leaders can strike transfer agreements with neighboring districts that have open seats in successful schools; determinations about whether student transfers will conflict with court orders or state education board or legislated policies that dictate a certain racial balance or class size in schools; and the willingness of parents to pick the NCLB choice option rather than exercising some other form of choice, which often exists in urban districts.

Similarly, broad student access to supplemental educational services only occurs if providers choose to make themselves widely available in communities across a state. Some such providers have found it more profitable or easy to operate in densely populated areas or to serve only particular student groups (e.g., economically disadvantaged but not English language learners or the disabled). Further, state and local government coordination is required, given that state education agencies are responsible for keeping a list of certified providers while local districts are charged with administering funds for the program on the ground level. Finally, in contrast to IDEA, NCLB does not provide parents with access to the courts to guarantee access to choice or supplemental services. In some communities, that has led other advocacy groups and information providers, such as local nonprofit organizations, to stand by parents to help them take advantage of these opportunities. Absent more powerful influence from the courts, state and

local officials have felt less pressure to work creatively to implement these two remedies under the law.

Traditions of American federalism limit the federal government's ability to hold subnational governments accountable for producing high-quality educational outcomes. Promoting outcomes that enhance opportunity has been a mainstay of federal education policy since the 1960s when the original ESEA became law. Increasing opportunities for students to participate in public education to help them accomplish their personal goals has an appealing quality that is consistent with American ideals about individual initiative and perseverance. It is also consistent with prior studies of federalism, which tend to agree that higher levels of government are better positioned to advance outcomes supporting equality of opportunity than are lower levels of government.²⁶ Unequal opportunities frequently result from disparities in wealth across states or local communities, which make the federal government's redistributive role potentially powerful.

Crafting and defending a strong federal role becomes more difficult when concerns over accountability for outcomes shift from advancing opportunity to promoting quality. Approaching the stepping-stone of quality has always been a challenge for federal officials given that state and local governments remain the primary developers and implementers of education policy in the United States. And citizens generally like the idea that control over the content of what is taught in school should reside in subnational governments. This division of labor has moved federal officials to avoid passing federal laws or writing regulations that define quality outcomes in specific ways. Otherwise, they open themselves to criticism that they are trying to seize control of the nation's schools from the state and local leaders who are better positioned to act.

Two outcomes from NCLB's implementation illustrate the challenge facing federal officials who hoped to promote accountability for educational quality: student achievement in core subjects and the quality of the nation's teachers.²⁷ On the surface, NCLB's requirement that all students be proficient in reading and math by 2014 suggested a federal willingness to more aggressively promote a certain vision of quality. In reality, though, federal policy makers deferred to state governments to define the all-important substantive meaning of *proficient*. States retained the power to define content standards, what students should know and be able to do in reading and math. They also retained control over the definition of performance standards, which amounted to the cut scores that students needed to achieve on state-defined tests in order to earn proficient marks. Thus, proficiency

meant something different in every state. Consequently, changes in student achievement on state tests that occurred after NCLB's enactment would have reflected different levels of rigor depending on the states in which students attended school.

A similar pattern emerged from the law's HQT requirements. Again, federal policy did articulate a specific definition of *quality*, which required teachers to possess a bachelor's degree, full state certification, and expertise in their subjects. As with content and performance standards for students, though, the law left the states to tackle substantive matters. States continued to define what requirements teachers needed to fulfill to earn a teaching certificate (number of education credits, nature of the student teaching experience, and other requirements). They also possessed the power to decide when teachers knew their subjects. Some states allowed completion of a college major to suggest subject-matter knowledge, but there was much diversity across the states in defining how much college course work constituted a major. Similarly, some states allowed teachers to take a series of standardized tests, often from the Praxis series, to demonstrate their knowledge. But the cut scores states used to determine who was an expert and who was not also varied, with many cut scores set relatively low. So even though many states could produce data saying that more than 90 percent of their teachers were highly qualified, expectations regarding substantive quality were by no means consistent or uniform across the nation.

As we see with both defining student proficiency and teacher qualifications, the federal government is limited in its ability to define and promote substantive educational outcomes. The obstacles are two-fold. First, offering assertive federal definitions of what substantive quality means would require a complete reversal of the hallowed principle of state and local control over education in the United States. Although today federal officials more confidently assert their interests regarding adherence to federal requirements (as in the first stepping-stone) or the promotion of educational opportunities (the second stepping-stone), they become much more cautious when, for example, discussions turn toward considering more specific requirements that would define the substance of quality educational outcomes. Second, promoting accountability for quality outcomes also requires the presence of reliable funding and capable staff to monitor achievements, provide administrative support, and enforce meaningful sanctions when states, districts, or schools resist federal objectives. In a political climate that produces sloganeering about sending every possible dollar to the classroom, it is difficult to argue for greater investments in federal and state bureaucratic capabilities even when such efforts could enhance the ability

of both levels of government to improve substantive results. These two factors cast a long shadow on the ability of the federal government to reach the third accountability stepping-stone.

EDUCATIONAL ACCOUNTABILITY AND THE FEDERAL ROLE

In essence, federal efforts that attempt to promote accountability strive to lay a series of stepping-stones to build a bridge from federal ambitions to eventual student success. Washington policy makers face increasing obstacles when their goals move from fostering the creation of policy outputs at subnational levels to producing outcomes that enhance educational opportunity and, ultimately, to generating subsequent outcomes that improve educational quality. The nation's experience with NCLB illustrates the challenges inherent in moving from approaches to accountability that depend primarily on subnational policy production to improving the quality of teaching and the quality of student academic experiences. Considering this experience in the larger global context, and in light of other federal efforts in the United States, we close with some broader insights about the federal government's ability to promote educational accountability for the production of desirable policy outputs and outcomes in education.

First, federal leaders should never forget that their efforts to promote educational accountability do not operate in isolation from other levels of government. Other countries tend to avoid this problem by designating a single administrative unit as the principal definer of accountability protocols. In the United States, however, the responsibility for defining and determining accountability is shared across federal, state, and local levels. The large number of federal education programs, including the few dozen or so embedded in NCLB, along with state and local education initiatives, all embrace, to varying degrees, elements of accountability distributed across the three stepping-stones outlined in figure 7.1. Federal policy designed to enhance accountability that fails to recognize this reality is likely to foster duplicated effort, confusion, and challenges on the ground in districts and schools. The existence of dual accountability systems, in which states rate schools with their own parallel approaches that may or may not be consistent with NCLB's approach using AYP, has meant that schools can receive high marks from one but low ones from the other. Similarly, the Obama administration's Race to the Top Fund (RTTT) aimed to alter the meaning of quality teaching in NCLB's requirements, which focused on credentials and demonstrated teacher knowledge, to also incorporate elements capturing gains in student achievement. It is worth noting that in adding that approach, RTTT did not dismantle the HQT requirements present in NCLB,

and in fact some of them may be in tension.²⁸ As a result, state and local implementers that won RTTT grants will have to attend to both sets of accountability expectations.

Second, federal education policies to promote educational accountability have been, and likely will continue to be, administered by subnational governments. Although the federal government has become a more active player in elementary and secondary education, the bulk of the funding and effort in this policy area still resides in state and local governments, and likely will continue to for several years to come. Furthermore, contrasted with the powerful authority enjoyed by provinces and the marginalization of local boards in Canada, implementation of federal policies in the United States occurs in an environment characterized by fragmented authority. That has major implications for the ability of the federal government to lay a series of stepping-stones for accountability that move from overseeing the formation of policy outputs to the realization of substantively valuable student outcomes.

No matter how compelling or well-intentioned federal policy may be, getting to the results that really matter cannot occur without states and local governments mustering the political will and administrative capacity to do so. That division of labor for education in the American federal system strongly suggests that federal efforts to influence outcomes should expect to encounter many hurdles at subnational levels. If the nation's experience with NCLB is any guide, conceptualizing an accountability system of federal carrots and sticks to improve student outcomes may be a bridge too far. That is not to say that improving student achievement and eliminating achievement gaps are not worthy goals. They are. Rather, the point here is that in light of the political logrolls that help produce federal policy, and given the diverse states and communities in which policy plays out, it is incredibly difficult to craft federal policy to promote accountability for substantive outcomes that focus on educational quality.

Finally, if indeed it is a bridge too far to expect federal policy makers to capably take a leading role in promoting accountability for quality outcomes, they certainly can help create conditions that enable state and local leaders—those closest to the ground level—to take up the accountability charge. Looking abroad, other national governments enjoy considerably more authority over their respective education systems, compromising the principle of local control. In lieu of abandoning this hallowed principle of the American education system, ironically, then, the federal government may best leverage its comparative advantages if it focuses less on trying to pass policies that lay down a line of stepping-stones that directly connect policy outputs to student opportunities to academic results.

A better approach would be to think carefully about how to wield federal power to foster the development of knowledge, administrative infrastructure, and transparency that is required for a vibrant accountability system to flourish, a theme that other chapters in this volume echo. Federal efforts to seed the development of state longitudinal data systems is one example, as is NCLB's requirement that schools and school districts make information about their teaching staffs and student performance across subgroups transparent to the public. Focusing on those sorts of efforts would keep federal attention targeted on the first stepping-stone of accountability that we discussed earlier while at the same time minimizing regulatory burdens. It would then be up to state and local leaders, teachers, parents, and interested groups to mobilize that capacity and information to lay down the subsequent stepping-stones that could produce better outcomes. Such an approach may seem frustrating to federal leaders who see themselves as occupying a more assertive role. In the long run, that approach would better play to the federal government's comparative advantages.

We thank Andrew Kelly, Frederick Hess, and Daniel Lautzenbeiser for their helpful comments on earlier drafts. We also thank Claire Delcourt for her superb research assistance.

59. For more on these networks, see: www.pie-network.org/, <http://nestcentral.org/> and www.aspeninstitute.org/sites/default/files/content/docs/edinno/index.html.
60. For more on how the federal DOE could be restructured to better support innovation, see Sara Mead and Andrew Rotherham, *Changing the Game: The Federal Role in Supporting 21st Century Educational Innovation* (Washington, DC: Brookings Institution Press, 2008).
61. Jeffrey Mervis, "Obama Proposes Education Technology Agency Modeled after DARPA," *Science Insider*, February 4, 2011.
62. See <http://ccsr.uchicago.edu/content/index.php> and www.serp.institute.org/ for more information on these two promising research programs. See also A. S. Bryk and L. M. Gomez, "Reinventing a Research and Development Capacity," in *The Future of Educational Entrepreneurship*, ed. Frederick M. Hess (Cambridge, MA: Harvard Education Press, 2008), 81–206.

Chapter 7

1. Mark Bovens, "Public Accountability," in *The Oxford Handbook of Public Management*, ed. Ewan Ferlie, Laurence E. Lynn Jr., and Christopher Pollitt (New York: Oxford University Press, 2005), 182.
2. Kenneth J. Meier and Laurence J. O'Toole Jr., *Bureaucracy in a Democratic State: A Governance Perspective* (Baltimore: Johns Hopkins University Press, 2006).
3. For the distinction between policy outputs and substantive outcomes, see James Q. Wilson, *Bureaucracy: What Government Agencies Do and Why They Do It* (New York: Basic Books, 1989).
4. Actually attributing outcomes to specific policy outputs can be challenging, especially when policies are implemented in complex networks involving several actors with varying degrees of responsibility. For discussion of this idea, see Meier and O'Toole, *Bureaucracy in a Democratic State*; and Wilson, *Bureaucracy*.
5. Donald P. Moynihan, *The Dynamics of Performance Management: Constructing Information and Reform* (Washington, DC: Georgetown University Press, 2008); William T. Gormley and David L. Weimer, *Organizational Report Cards* (Cambridge, MA: Harvard University Press, 1999); Beryl A. Radin, *Challenging the Performance Movement: Accountability, Complexity, and Democratic Values* (Washington, DC: Georgetown University Press, 2006); Wilson, *Bureaucracy*.
6. Joanne Anderson, *Accountability in Education*, Education Policy Series (Paris: International Academy of Education; Brussels: International Institute for Education Planning, 2005); Richard Rothstein, *Grading Education: Getting Accountability Right* (Washington, DC: Economic Policy Institute; New York: Teachers College Press, 2008).
7. Gormley and Weimer, *Organizational Report Cards*; Wilson, *Bureaucracy*.
8. *Education at a Glance* (Paris: OECD, 2007), 415.
9. Paul Manna, *School's In: Federalism and the National Education Agenda* (Washington, DC: Georgetown University Press, 2006).
10. *Education at a Glance*, 413. See also Anderson's *Accountability in Education*, which discusses enhanced efforts to promote accountability in Great Britain's Office for Standards in Education, in Australia's National Education Monitoring Performance Task Force, and in the United States with NCLB.
11. Jennifer Wallner, "Beyond National Standards: Reconciling Tension Between Federalism and the Welfare State," *Publius: The Journal of Federalism* 40, no. 4 (2010): 646–671.
12. Finland is a notable exception to the trend focusing on externally determined learning standards. Rather than prioritizing a results-based system rooted on test-based accountability and standards articulated by outside forces, Finland has emphasized teacher professionalism and mutual trust: Schools have been encouraged to craft positive learning environments and implement content that best helps students reach gen-

- eral goals. See Pasi Sahlberg, "Education Policies for Raising Student Learning: The Finnish Approach," *Journal of Education Policy* 22, no. 2 (2009): 147–171.
13. Ibid.
14. Ibid., 152.
15. *Education at a Glance*, 414.
16. Paul Manna, *Collision Course: Federal Education Policy Meets State and Local Realities* (Washington, DC: CQ Press, 2011).
17. For summaries of the law's content and theories of action, see Frederick M. Hess and Michael J. Petrilli, *No Child Left Behind Primer* (New York: Peter Lang, 2006); and Manna, *Collision Course*.
18. Martha Derthick, *The Influence of Federal Grants* (Cambridge, MA: Harvard University Press, 1970); David R. Beam and Timothy J. Conlan, "Grants," in *The Tools of Government: A Guide to the New Governance*, ed. Lester M. Salamon (New York: Oxford University Press, 2002).
19. Gerald N. Rosenberg, *The Hollow Hope: Can Courts Bring About Social Change?* 2nd ed. (Chicago: University of Chicago Press, 2008).
20. Data come from the authors' analysis of the states' NCLB accountability workbooks. Original documents from 2003 are available from the authors. The latest workbooks are available at <http://www2.ed.gov/admins/lead/account/stateplans03/index.html>.
21. Statistics cited in this paragraph come from the authors' analyses of the state HQT plans, which are available at <http://www2.ed.gov/programs/teacherqual/hqtplans/index.html>.
22. The administration's evolving position on HQT is discussed in Manna, *Collision Course*, 55–58.
23. Meier and O'Toole, *Bureaucracy in a Democratic State*.
24. Chester E. Finn Jr., Andrew J. Rotherham, and Charles R. Hokanson Jr., eds., *Rethinking Special Education for a New Century* (Washington, DC: Thomas B. Fordham Foundation and Progressive Policy Institute, 2001).
25. Statistics are from the authors' analysis of data available from the U.S. Department of Education's ED Data Express Web site, www.eddataexpress.ed.gov/.
26. Paul E. Peterson, Barry G. Rabe, and Kenneth K. Wong, *When Federalism Works* (Washington, DC: Brookings Institution Press, 1986); Paul E. Peterson, *The Price of Federalism* (Washington, DC: Brookings Institution Press, 1995).
27. Manna, *Collision Course*.
28. An example of this tension is the NCLB requirement that students not be tested on material outside of their current grade level. In contrast, RTTT aimed to push states to develop accountability systems in which teachers could be rewarded if their students made more than one year of academic gain in a single academic year. For teachers whose students all began the school year either at or above grade level, it would be impossible to demonstrate such gains under NCLB.

Chapter 8

1. It was also in 1955 that the AFL merged with the CIO to form the nation's largest labor union. And conservative economist Milton Friedman fired the first salvo in the battle over school choice that year when he proposed a method of improving public education by opening it to marketplace competition. Levittown-type housing in the nation's suburbs represented 75 percent of all new residential construction starts in 1955, as well, marking the point when the majority of the nation's largest cities began losing population and a profound shift in the demography of urban public education commenced.
2. As part of the process to incorporate the six national education goals into Goals 2000, the new Clinton administration's version of George H. W. Bush's America 2000 initiative, Congress added two more goals of their own in an attempt to put its imprimatur