

## Carrots, Sticks, and Unbroken Windows

### Making NCLB Live Up to Its Promises

If we are to make use of what we knew in Dewey's day (and know even better today) about how the human species best learns, we will have to start by throwing away the dystopia of the ant colony, the smoothly functioning (and quietly humming) factory where everything goes according to plan, and replace it with a messy, often rambunctious, community, with its multiple demands and complicated trade-offs.

*Deborah Meier, Educating a Democracy*

The measurement of educational quality is as messy and complicated as education itself. But it is not impossible, and it would be a mistake to wave our hands and assume that No Child Left Behind cannot be fixed. The law undoubtedly constitutes a flawed vehicle, but it makes the kinds of promises to our most disadvantaged citizens and their children worthy of a great liberal democracy. Making NCLB work is worth the effort. As Gary Orfield of the Harvard Civil Rights Project has argued, "What is sorely needed now is an acknowledgment that the too-hasty compromises and contradictions [of NCLB] need to be sorted out, that experts in implementing deep educational change

and people who know what the reasonable expectations for progress are and how to measure progress in a more sophisticated way be brought into the process.”<sup>1</sup>

In chapter 2, I argued that experience is everything in education. It shapes success, failure, assessment, and compliance. Those closest to the production process know its quality better than anyone else, though even their assessments will never be perfectly accurate. Incentives constitute an important part of the story as well. Given the limitations on principals’ and teachers’ time resources, any high-stakes accountability plan must be formulated with careful attention to the bureaucratic responses to and unintended consequences of its evaluations and menu of consequences.

Chapter 3 introduced the empirical data for the analyses: a 2003 survey of Minnesota’s public and charter school principals on their leadership, their influence, and No Child Left Behind. Combining these results with data on student characteristics, test score results, and adequate yearly progress (AYP) status, I found a troubling though unsurprising relationship between a school’s success and failure under No Child Left Behind and the characteristics of the student population. However, principals’ leadership decisions can matter under NCLB’s test-based regime, offering encouragement that educational outcomes are not determined by socioeconomic status alone.

Chapter 4 turned the analysis around, looking at the effects of being labeled as a failing school on principals’ behaviors under and attitudes toward No Child Left Behind. Troublingly, AYP identification does not appear to be associated with a reenergized and re-focused principalship but with risk aversion and a loss of influence. I also examined principalship in charter schools, whose dependence on customer preferences—according to the theory laid out in chapter 2—offers inducements toward a more proactive and customer-focused principalship. Though Minnesota’s charter schools have a startlingly high AYP failure rate, it may well result from the effects of NCLB’s status model of quality assessment rather than from deficiencies in the charter school principalship. I concluded by questioning the logic of denying traditional public schools the benefits of an incentive structure that combines bureaucratic autonomy with a more stakeholder-focused leadership.

In chapter 5, I discussed three alternatives to NCLB's status approach to quality assessment: growth, value-added, and production models of measurement. Though more data are needed, patterns of success under growth models appear to differ from and be preferable to what we see under the current status approach. Charter schools would likely do better under either alternative, as would many traditional public schools with large percentages of minority students and of students eligible for free and reduced price-lunch that were doing the kinds of things that excellent schools do.

I then turned to data from the Minneapolis Public Schools, incorporating the results of parent and student surveys into my larger data set of AYP success and failure along with student and school characteristics. These investigations produced striking results. Though the Minneapolis Public Schools displayed a very high rate of failure—not unlike many urban school systems around the country—schools whose students and teachers felt them to be safe and orderly, focused on curricular standards, and characterized by relatively high levels of parental involvement were much more likely to make AYP than those whose students and teachers reported lower levels of these conditions. These results were even stronger when Minneapolis's relatively advantaged schools were deleted from the analysis, confirming that getting things right in education matters most to those who need it the most.

This chapter moves from questions to answers, taking these lessons as a basis to argue for specific policy proposals as we move forward under No Child Left Behind or whatever state and national policies survive it. I take it as a given that the world of high-stakes top-down accountability is here to stay and offers a series of progressively substantial modifications to the law, from small tweaks to bigger changes in thinking about the prospect of effective and thoughtful educational accountability in the United States.

Having explored issues involved with top-down and bottom-up reforms in education, I would like to propose something more along the lines of a "middle-out" model for fixing No Child Left Behind.<sup>2</sup> The central premise behind middle-out reform is to work from the inside out, basing assessments and incentives on the experiences and insight of those who know what is good and not good and what needs

to be done to improve the situation. A plan for incorporating middle-out reform into an accountability system under No Child Left Behind would be based on the following four assumptions and would be undertaken according to the following dictums:

1. *Proximity matters; therefore, go inside.* Middle-out reform begins and ends with those closest to the point of production. This is critical in experiential organizations. Assessments and consequences should be based on the wisdom of those who best know quality, including producers and consumers of the services. Both quantitative and qualitative data will be necessary, as will careful attention to any consequences—intended and unintended—of introducing or modifying the incentive structures under which principals and teachers work.

2. *Context matters; therefore, be flexible.* One size does not fit all, and those closest to production have much to offer when designing assessments and consequences. In addition, any one measure of quality will probably be insufficient in an undertaking as complex as producing high-quality education.

3. *Incentives matter; therefore, be theoretically grounded.* Incentives will be responded to, gamed, and possibly undermined if those involved in the enterprise reject the foundations of the logic behind the incentives. The primary danger with incentives is not that some individuals will game the system (which they will) but that incentive structures may unintentionally reallocate principals' and teachers' time away from aspects of their jobs that we do not wish them to ignore. Multitasking is one of the defining characteristics of principalship. Any incentive structures must be designed with an understanding that time is not an infinite resource.

4. *Resources matter; therefore, be bold.* Money does matter, but not uniformly and without careful thought into where it goes and on what basis it is handed out. Resources applied or withheld without an adequate understanding of agency, incentives, leadership, and context will probably not produce the desired educational benefits. The challenges involved in translating these resources into positive outcomes should not lead policymakers to give up on their possibility. Stan-

standardized tests have existed for more than a century; however, we as a society have never committed to providing the same level of resources to our poorest students as wealthy individuals have been willing and able to spend on the education of their children. Resources need to be provided based on evidence of success and tailored to sustaining increases in quality, and these rewards should be both financial and bureaucratic.

The effective schools movement may have failed to achieve all of its goals partly because of a lack of a theoretical paradigm that could compete with the simple (and perhaps simplistic) assumptions of those who argued that given the determinacy of student and peer characteristics, resources never matter in education. The insights of scholars of bureaucratic agency offer powerful theoretical tools for developing a policy that allows for the judicious application of rewards and positive incentives. However, we should not be overly optimistic about the instantaneous prospects for any initiative, as many educational reforms have foundered on the shoals of unrealistic expectations.<sup>3</sup> Rather, we need to create a dynamic that feeds on itself and builds toward lasting and sustainable change.<sup>4</sup> Such a system might involve four modifications (in order of increasing significance) of No Child Left Behind.

1. *We need to continue to explore the incorporation of growth and value-added models of assessment into AYP calculations, but we should not expect either to be a quick fix to NCLB's problems.*

The incorporation of either growth or value-added models of assessment into NCLB is the low-hanging fruit of educational policy in this country, and we are already seeing considerable movement on these fronts. Neither approach, however, is as easy as it sounds. Policymakers should continue to explore the use of these alternate testing formats but should place much more emphasis on implementation than on methodological advantages. Though much discussion has occurred on the merits of either a growth model or a value-added model, scholars have focused much less on what to do with this information or on the important policy issues that need to be addressed

prior to large-scale implementation of either a growth or a value-added assessment model.

The main issue is whether one of these alternatives will be used alone or in combination with NCLB's current status model for determining AYP. I would guess that the result will be a combination. If we were to adopt only a growth or value-added model of assessment, then wealthy schools might be sanctioned in roughly the same proportion as poorer and high minority schools under the status model given that students at more advantaged schools would probably be starting from a relatively high level of achievement. The political implications of such a scenario are easy to imagine and make it easy to dismiss this prospect as unlikely ever to see the light of day.

I start with the assumption that policymakers incorporate a status model and a simple growth model in their evaluations of schools, that (for the sake of simplicity) only academic proficiency (and not attendance and graduation are included), and that all schools succeed in all test participation categories. Four categories of schools would result:

1. those that failed to show both academic proficiency and sufficient gains in proficiency;
2. those that failed to show proficiency but showed yearly gains in proficiency;
3. those that displayed proficiency but failed to show gains; and
4. those that showed both proficiency and gains in proficiency.

Schools in categories 1 and 4 are straightforward: those in category 1 would fail, while those in category 4 would pass. Schools in category 3 would also likely avoid sanction, as most of them would be in wealthier areas with politically active parent communities that would act to shape the policy regulations of such a system.

Therefore, policymakers would be left to decide what to do with those schools that failed to make academic proficiency but displayed strong gains for all students. These schools probably are also failing specific subgroups of students or else the schools would have made

safe harbor. One possibility would be to extend safe harbor to the top gaining schools for one year but not the next. Given how much test score gains bounce around from year to year, however, it is not clear that such exemptions would be based on meaningful data. If we exempt from sanction only those schools that make gains for all of their failing subgroups, then nothing will really change, since most of these schools would be making safe harbor. If, conversely, we exempt schools that show gains in proficiency overall but not necessarily in every subgroup, then we run the risk of giving too much leeway to schools that are failing their students in specific subgroups of race, ethnicity, wealth, and ability—that is, exactly the students that No Child Left Behind was designed to help. This is quite a conundrum indeed.

If we were to adopt a value-added system in combination with a status model, the situation would differ slightly but remain complicated. One could imagine exempting a school from identification if it were adding value to the students overall or especially within underperforming subgroups. Again, schools adding value to students in all subgroups would probably not need any additional help. Therefore, both growth models and value-added models sound good in theory, but neither appears to be a magic bullet in practice. At present, only educational methodologists are providing in-depth analyses of these two options. Their ranks need to be supplemented by scholars who pay attention to the important details and consequences of these good ideas.

2. *Choice options, including charter schools, need to be dealt with on a fundamental level.*

Charter schools are not thriving under No Child Left Behind. However, many of these schools serve high-need student populations and display many of the production characteristics that we are trying to encourage with No Child Left Behind. It is more than a little ironic that we feel the need to apply top-down sanctions to a group of schools designed to be free from excessive regulation. Policymakers need to address this contradiction. If we believe in the power of consumer choice, then we should place our faith in it. If not, then we

should admit that we are just fumbling around as best we can. Given the complexity of the task, a bit of humility might be useful.

One researcher has advised that a “two-year grace period may be desirable before starting the clock on sanctioning new charters,”<sup>5</sup> given the dual challenges of getting schools started and avoiding sanction under NCLB. While such a plan has the benefit of giving charters a chance to succeed, it runs the risk of ignoring the achievement of the students in these charters. It is also a deeper issue than when or even if we should apply NCLB’s assessments and consequences to charter schools. It is not at all clear that more than a fraction of the parents presented with choice options (including charter schools) under NCLB choose any of these options. And if all eligible parents decided to use these options, it is not at all clear that successful schools would have enough seats for all these children.

The real question is whether we are going to treat all parents in the same way under No Child Left Behind, not whether we will provide choices to a lucky few. Parental involvement is beneficial, as is bureaucratic autonomy. We now need to figure out how to bring those benefits into NCLB’s framework. If we believe in the simultaneous power of parental evaluations of school quality and bureaucratic autonomy as a reward for this parent-centeredness, then we should consider extending the bureaucratic advantages of charter schools to all public schools.

3. *We need to incorporate a production model of quality assessment into No Child Left Behind.*

I recommend that we adopt a production model of quality assessment while refining the test-based models currently in use or under consideration. A production model of assessment would rely on surveys of students, teachers, and parents to determine the quality of the processes and conditions of education. There is nothing revolutionary about incorporating parent and teacher satisfaction surveys into an evaluation system.<sup>6</sup> However, such input usually offers little more than window dressing for the real test-based accountability provisions. The difference with what I propose here is that parent, teacher, and student surveys would be directly incorporated into the

system and that the results of these surveys would have meaningful consequences for schools.

An important task in considering alternative assessment systems under NCLB is to take a look at what school principals think about measuring the quality of their leadership and therefore the quality of their schools. Because I have suggested that we incorporate surveys of teachers, parents, and students into such a system, it makes sense to ask whether principals think that these actors would constitute effective evaluators of school quality. The comments in the follow-up interviews discussed already suggest that such may be the case. Table 9 presents principals' evaluations of which measures of performance most accurately gauge how principals are fulfilling their obligations as leaders and as local school managers. As table 9 clearly demonstrates, Minnesota's traditional public school principals felt that evaluations by the teachers at their schools represented the most effective measures of their performance. These results held true for principals whose schools made AYP in 2003 but were even stronger for those whose schools did not make AYP. For principals of charter

**TABLE 9. Most Effective Measures of Leadership: Minnesota's Principals' Views**

	Regular Public (made AYP in 2003) (1)	Regular Public (failed to make AYP in 2003) (2)	Charter School (3)
Test scores of students at this school	13%	14%	8%
Evidence of parent outreach (mailings, events, telephone contacts)	6%	6%	5%
Results of parent satisfaction surveys	16%	16%	28%
Evaluations by teachers at this school	29%	35%	25%
Evaluations by administrative superiors in this district	13%	12%	n.a.
A combination of factors	12%	10%	18%
Other	11%	8%	18%
Number of responses	870	51	40

Source: Minnesota Schools Survey 2003. AYP data from Minnesota Department of Education 2003b. Totals may not add to 100% due to rounding.

Note: The survey question was worded as follows: "Which of the following measures of school performance do you feel *most effectively measures your leadership* as principal of this school?"

n.a. = not applicable.

schools, parent evaluations were the most commonly cited best measure; however, teacher evaluations came in a close second.

This is an important finding, as it points to a divergence in focus between the traditional public and charter school principalships. Principals in traditional public schools may be more inwardly focused, looking for confirmations of their successes and failures within their teaching staffs, while charter school principals look outward to their parent communities for these verdicts.

These results are also consistent with the findings from chapter 4: principals in traditional public schools felt that their power and influence were most effective when they were exercised in concert with rather than in opposition to their teachers. These findings do not presuppose a perpetually happy principal-teacher relationship but assume only that principals and teachers are most effective when their interests are aligned. I am not aware of proposals for modifying No Child Left Behind to incorporate teachers' evaluations of principals into assessments of school performance. It is difficult, however, to imagine a private-sector system that fails to incorporate subordinates' assessments of their managers in assessing whether those managers are performing adequately. For some reason, this is typically the case in public education.

Roughly 10 percent of the principals in the Minnesota survey took the option of writing in their own responses to the question of most effective measure of leadership rather than circling an item from the menu that I provided (the "other" category in table 9). Table 10 presents the results of these write-in responses, with charter and non-charter schools combined, given the small number of respondents.

The second-most-common response was something that should have been included in the menu and a factor that showed up consistently in the follow-up interviews: evaluations by students. Students are the primary stakeholders in education, yet their feedback is excluded entirely from No Child Left Behind's evaluation and their potential role in creating a comprehensive system of incentives for schools is completely ignored. I, like many of my fellow researchers, failed to listen carefully enough to what the theory was telling me, unwittingly contributing to the omission of students from the

accountability process even though they are causes as much as examples of achievement.<sup>7</sup>

No one is closer to the production of educational quality than the students. Even elementary school children have useful feedback to offer to assess the conditions of educational production, as the coherence and consistency of student evaluations and academic performance in the Minneapolis data attest. Students know if they feel safe, if they are allowed to concentrate in class, if their teachers care about them, and if their teachers care about the material. The proponents of the “breaking ranks” model of educational accountability reject the current situation whereby “the student’s role in the educational process is passive and subordinate,”<sup>8</sup> replacing it with a model of “student-centered accountability”<sup>9</sup> that recognizes the critical role that school climate and culture play in educational excellence.

The most commonly offered measure of principals’ leadership in the responses in table 10 was a combination of teacher, parent, and student evaluations—that is, precisely the types of measures I advocate for use in modifying No Child Left Behind’s assessment regimen. Many of the Minnesota principals who chose to offer their opinions about effective measures of their leadership felt that student evaluations should be a component of a more comprehensive system

**TABLE 10. Other Effective Measures of Leadership: Minnesota’s Principals’ Write-in Responses**

	All Principals
Student evaluations	17%
Combination of parent, teacher, and student evaluations	20%
School climate and student engagement	10%
Goal attainment and mission orientation	11%
Gain models, value-added models of test scores	9%
Attendance, graduation, and postgraduation success	13%
Reputation, word of mouth	5%
Other	15%
Number of responses	109

*Source:* Minnesota Schools Survey 2003. AYP data from Minnesota Department of Education 2003b. Totals may not add to 100% due to rounding.

*Note:* The survey question was worded as follows: “Which of the following measures of school performance do you feel *most effectively measures your leadership* as principal of this school?”

of evaluation. Several also placed the utility of these evaluations within an understanding of the school as a social system producing whole citizens. Nonprompted measures of superior leadership included the following:

If students like school, feel safe, feel as though they have input into decision-making, are challenged academically, and are good citizens to each other.

School climate/school success on scores *and* student interest/belief in education. [It] has to be well-rounded—nothing [in the list] related to student perceptions.

Student success factors such as participation in leadership activities, student satisfaction, student self-efficacy, student personal growth, student confidence, belief in self.

Building relationships and trust with staff, students, and parents.

We know what principals think. They agree with the core prediction of the theories that I have discussed in this book: those closest to the production of education are best able to assess its quality. The question, of course, is how to incorporate these evaluations and assessments into No Child Left Behind's test-based world.

*4. The implementation of No Child Left Behind must incorporate meaningful rewards into its accountability system.*

At its core, No Child Left Behind is a punitive system despite the fact that its passage was associated with a promise of a significant boost in federal education funding and that its implementation has resulted in more monies for Title I (though somewhat less than was promised). Nevertheless, there are no federal rewards for schools that succeed under the law other than keeping its wolves at bay.<sup>10</sup> In addition, relatively little discussion has examined the role that incentives, positive or negative, play in evaluating No Child Left Behind's implementation and the likelihood of its success or failure:<sup>11</sup>

“Unfortunately, the prevailing mind-set among too many politicians and policy makers is something right out of *Field of Dreams*: ‘If we set high standards, students will magically achieve.’”<sup>12</sup>

One superintendent interviewed as part of the 2004 study by Minnesota’s Office of the Legislative Auditor summed up the irony of NCLB’s punishment-driven philosophy: “If research drives this law, then those who promulgated it know that punishment is the least likely way to get improvement. Yet the only form of motivation for teachers and schools is the threat of loss of revenue, prestige, and the school itself.”<sup>13</sup>

This aspect of No Child Left Behind is very difficult to reconcile with a true desire on the part of its authors to make public schools succeed. Any reasonable exploration of bureaucratic agency raises serious questions about an incentive structure built only on avoiding punishment. Given what James Q. Wilson has observed about coping organizations and what I have found about the experiential subset of these organizations, NCLB seems almost designed to produce mediocrity, risk aversion, and doing time. The empirical results of this book support this baffling conclusion. Therefore, in addition to incorporating alternative measures of educational excellence and thinking more carefully about choice, policymakers must incorporate meaningful rewards into their plans.

Table 11 presents the responses by public school principals in Minnesota to a question about the most effective incentives for improving academic achievement within their particular schools. What is striking about these results is how similar they are among different groups of principals. These individuals do not report that they want more money for themselves; rather, they seek the resources necessary to achieve their vision for excellence within their schools. Of course, it is possible that these principals are merely providing what they think is the “correct” answer to the question. Nevertheless, all the groups agree, and their observations concur with the findings of this study. Public school principals are strategic actors, and they see the potential utility or necessity of making their institutions more responsive to the needs and wishes of their parents and students; however, they desire the resources to allow them to do so. Perhaps a

theoretically and empirically sound accountability program could help make this a reality.<sup>14</sup>

Table 12 summarizes the responses of principals who chose, in a separate question, to offer their views on what other kinds of incentives might be offered to principals who demonstrated superior performance. Resources for the school (not for the principal) again dominated. Increased funding for smaller class sizes as well as resources for training were commonly suggested. However, not all principals felt that incentives were necessary, pointing to the limitations on any incentive system put into place in an endeavor where much of the motivation is intrinsic, but even these individuals might benefit from more and better leadership tools.

Several of the principals who offered their own opinions saw bureaucratic rewards as effective additions to resource rewards. Such bureaucratic enticements might include more autonomy—not necessarily power over teachers, just the power to shape the teaching team. The principal of one of Minnesota’s wealthiest metro-area high schools wanted authority not to fire his teachers but to recruit and keep them. In response to the question, “What other potential incentives that were not mentioned in the above list do you think might be

**TABLE 11. Incentives for Superior Leadership: Minnesota’s Principals’ Views**

	Regular Public (made AYP in 2003) (1)	Regular Public (failed to make AYP in 2003) (2)	Charter School (3)
Cash bonuses	6%	6%	6%
Recurring salary increases	12%	2%	6%
Sabbatical leaves	2%	6%	9%
Decreased oversight from state and local administration	14%	20%	21%
Additional support staff	66%	65%	56%
Number of responses	810	48	34

*Source:* Minnesota Schools Survey 2003. AYP data from Minnesota Department of Education 2003b. Totals may not add to 100% due to rounding.

*Note:* The survey question was worded as follows: “Of the following incentives that might be offered to public school principals for superior performance, which do you think is the *most likely* to improve academic achievement within the school?”

offered to public school principals who exhibit superior performance?” he requested,

Choosing a certain percentage of teachers to hire or keep employed even if they do not meet licensing requirements. I have had to let high quality people go in order to satisfy union and licensing demands. Without a doubt, these people had an outstanding effect on student achievement.

The desire for flexibility in hiring and retention—politically a much easier task than flexibility in firing—concur with the results of a 2004 study for the Minneapolis Public Schools that advised the district to “increase the latitude of principals to interview, select, and hire their staff.” The authors of a 2003 national study of school and district accountability also concluded that “high performance can be rewarded with additional flexibility in school management, additional resources, or group awards or gifts to school employees.”<sup>15</sup>

#### Reclaiming No Child Left Behind’s Liberal Agenda

Student, teacher, and parent surveys can be useful measures of the quality of leadership and therefore the quality of education in our

**TABLE 12. Other Effective Incentives for Superior Leadership: Minnesota’s Principals’ Write-in Responses**

	Percentage of Write-in Responses
Public recognition for principal, teachers, and/or students	18%
Leaves, vacations, etc.	3%
Fewer regulations in terms of running the school	9%
Paid training and workshops for principals and staff	14%
Other financial benefits (retirement, fringe, etc.)	3%
Increased funding for smaller class sizes, better equipment, etc.	23%
Incentives not necessary	15%
Other	15%
Number of responses	421

*Source:* Minnesota Schools Survey 2003. Totals may not add to 100% due to rounding.

*Note:* The survey question was worded as follows: “What other potential incentives that were not mentioned in the above list do you think might be offered to public school principals who exhibit superior performance?”

public schools. Incentives—both resource-based and autonomy-based—can be more thoughtfully incorporated into such a system. I conclude this book by offering some ideas on how policymakers might incorporate production-based incentives into No Child Left Behind's ambitious agenda.

The plan would focus on schools in America's urban centers and disadvantaged rural communities. Though the theoretical and empirical investigations undertaken in this book apply equally to all public schools, NCLB focuses on closing the achievement gap between advantaged and disadvantaged students. The plan would be based on combining the results of systematic and comprehensive annual surveys of teachers, students, and parents with the results of the variant of test-based accountability state officials chose to adopt. Results of these surveys would be used in two ways. First, they would be publicized. For schools operating under a system of choice, these results would align the top-down assessment with the bottom-up provision of choice in a way that makes much more theoretical sense than what we are now doing. Second, being labeled as successful would have significant consequences for a school on the NCLB sanction track as well as beyond NCLB's punitive focus.

The heart of the evaluation would be a system that awarded one point to schools serving high-need student populations for each of eight criteria, including having made AYP in the previous year as well as being in the top 25 percent of schools in the comparison group based on

- one-year gains in reading proficiency;
- one-year gains in math proficiency;
- safety and order (as measured by teacher, student, and parent surveys);
- a focus on curriculum and high standards (as measured by teacher, student, and parent surveys);
- creating a climate of respect and trust (measured by surveys of students and teachers);
- a focus on high standards (as measured by surveys of parents and teachers); and
- parental involvement (parental attendance at conferences and rate of volunteerism).

Schools would be compared to other schools in the state with similar sociodemographic characteristics. The top 20 percent of these schools, based on their point scores, would be identified as “schools of promise” and would get placed on a timeline of rewards that resembled No Child Left Behind’s current timeline of sanctions. These rewards would be cumulative, just like the sanctions for AYP failure. Rewards under this system, however, would be both financial and bureaucratic and would be funded by the federal government, but not out of previously allocated Title I monies. In other words, the plan would entail a commitment of new and dedicated funds, not another reallocation of what has already been appropriated. Table 13 lays out the increasingly significant proposed rewards for repeated identification as a school of promise under this plan.

In the first year of identification, the school’s status—along with lists of top-performing schools in each of the point categories—would be made public as part of No Child Left Behind’s reporting requirements. In addition, schools of promise that did not make AYP in the same year would not advance on the sanctions schedule but

**TABLE 13. Increasingly Significant Rewards for Schools of Promise**

Cumulative Years of Identification	Rewards for Schools
1 year	Identification as a “school of promise” with public notification Progression along AYP sanction track frozen for 1 year (if applicable)
2 years	Progression along AYP sanction track reset to year 1 (if applicable) <sup>a</sup> School receives one-year grant for paid training and leaves for teachers. <sup>b</sup>
3 years	School given control over hiring and retaining teachers. <sup>a</sup> School receives a 10% bonus (based on total staff salaries) to hire additional support staff. <sup>b</sup>
4 years	School receives a 10% bonus (based on total expenditures for teacher salaries) to hire additional teachers to reduce class sizes. <sup>a</sup> School exempted from forced layoffs and reassignment of teachers. <sup>b</sup>
5 years	Paid leave time for teachers and principal to conduct workshops at other schools Teachers receive a 10% salary bonus. <sup>b</sup> Principal eligible for leadership program

<sup>a</sup>Each subsequent year continues rewards from all previous years.

<sup>b</sup>Previous financial rewards continue in future years but are not cumulative. (Amounts are set at first year of award.)

would stay frozen at the previous year's level. The second year of identification would be rewarded with the reset of progress along the AYP sanction track to the first year, if applicable. The school would also receive a one-year grant to allow for paid training and leave for teachers and staff. This grant would recur in subsequent years of identification. The plan would not require consecutive identification to progress along the timetable, though perhaps it would include a reset button if a school failed to achieve promising status again after a number of years.

The third cumulative year of identification would reward schools with control of hiring and retention of teachers. I use the word *school* intentionally here, as I think more thought needs to be given, perhaps at the district or building level, to the relative power of principals and teachers in this reward. Principals need more autonomy in this area; however, one can imagine principal/teacher collaboration as well. In addition, a 10 percent bonus, based on current expenditures on staff, would be granted to the school to hire additional support staff. These funds would be allocated by the principal.

In the fourth year of cumulative identification, the school would receive a 10 percent bonus based on current expenditures on teacher salaries to hire additional teachers to reduce class sizes. Funding for these positions would be guaranteed for a specified number of years, since the uncertainty created by hiring new teachers without some such guarantee would likely nullify any benefits from the initiative. The school would also be exempted from forced layoffs and reassignment for that year, though teachers could still opt to be reassigned under standard district procedures.<sup>16</sup>

In the fifth year of identification, all teachers in the school would receive a 10 percent salary bonus, recurring in each subsequent year that the school was identified as promising. These salary bonuses would continue but would not be cumulative. Principals and teachers in these schools would be compensated, including the hiring of substitute staff, for conducting workshops and training of others in the district or state. In addition, principals would have the option of participating in a leadership program. New funds would be allocated to provide these principals with semester or one-year leaves to mentor other principals in the district or state. Perhaps the selection of indi-

viduals to fulfill administrative requirements during the leave could be used as an apprenticeship program for principals and assistant principals from other schools willing to undertake the challenge of leading a public school in a high-need community.

#### Issues and Concerns

This plan for modifying No Child Left Behind would require a non-trivial amount of data collection; however, this task does not seem overwhelming in comparison to the demands placed on schools and districts to collect and evaluate student achievement test results. In addition, these new data would be useful in their own right. Parents choosing schools, for example, would be very interested in finding out which schools were demonstrating quality in all or some aspects of their production. Principals in promising schools (and even those that just missed identification) could use the data to retain active parents in their school communities in the face of competition from charter schools, public schools in other districts, and possibly even private schools.

As has been discussed, any system of accountability must be carefully analyzed for intended and unintended consequences of the incentive system that it creates, and this caution applies equally well to the production model proposed here. Any approach brings the very real possibility that those subject to its reach will try to game the system to their advantage, perhaps in conflict with the system's ultimate objectives. However, the possibility that principals will reallocate their time to ensure positive survey responses from their parents, students, and teachers might lead to a beneficial skewing of the allocation of principals' time. The goal would not be to get principals to ignore performance on standardized tests but rather to facilitate a greater balance between time and energy devoted to the tests and time and energy devoted to reaching out to parents and fostering a vibrant social community.

Any incentive structure under this system would also ideally be compatible with and complimentary to both top-down and bottom-up reforms, using both carrots and sticks to align principals' and teachers' behaviors with customers' preferences. There is no theoretical basis for denying traditional public schools the benefit—now

given only to charter schools—of having an incentive and rewards system geared to becoming more aware of and responsive to parents' wishes and quality evaluations. This lack of attention to customers and autonomy, measurement and sanction represents one of the most critical gaps in current discussions of NCLB's prospects.

Teachers might have incentives to overstate their principal's leadership ability or the status of the school climate, since they would be the recipients of many of these benefits, particularly in later years. The idea, however, is to reward schools with more resources (teachers included) and more autonomy for the principal. Teachers, therefore, would be providing their survey responses with the knowledge that higher scores might contribute to more resources but would also result in the principal having more authority—including over hiring, retention, and the disbursement of grants and leaves. Schools whose teachers were performing well, partly because of the quality of the principalship, would benefit the most. Again, the goal is to strike a balance in the incentive structure.

Any consideration of financial and bureaucratic incentives must be made with dual consideration of feasibility and the fact that many of these rewards may continue beyond the fifth year. Perhaps various reset buttons would force schools to start over on the identification track; however, many of the positive incentives need to be provided with some assurance that they will continue so that school and district officials can prepare accordingly.

This plan is in no way meant to be definitive, and all of the financial incentives would need to be subject to a thorough cost-assessment process. The percentages required for identification have been set with the dual goals of making success possible for schools but not so high as to overwhelm the federal government's ability to fund the program or to make the AYP provisions meaningless. Such an outcome would very likely undercut political support for No Child Left Behind. The key is to strike a careful balance between improving NCLB and killing it. Any system that proposes to identify and reward educational quality should also be subject to intense review, debate, and modification. This plan is no different. My objective is only to help steer the debate away from complaisance or futility and toward action and hope. At the very least, policymakers should seriously consider a pilot project involving a few select cities—with a similarly

matched set of control cities—to explore this plan’s potential to improve achievement and reenergize and refocus our public schools in the communities that need it the most.

### Conclusions

The modern and not-so-modern history of reforming public schooling in the United States is littered with policy proposals that have failed because of a lack of the political will necessary to sustain them over the long haul, a failure to commit sufficient resources to implement them in full, poor planning, or a combination of these factors. No Child Left Behind is an imperfect vehicle, to be sure, but it is noble in its flaws, promising in its failures. No inherent incompatibility exists between high-stakes accountability and educational excellence. Properly devised, high-stakes accountability can have the benefits that we seek.

To make NCLB work, we need to hold schools—or more accurately, those who run and operate them—accountable for factors that they can control and that matter to student achievement yet not hold them accountable for factors beyond their control. We need to distinguish between schools that are failing their students and governments that are failing their schools, sanctioning and starving schools of promise in the communities that most need educational excellence. We must reward those who succeed with the tools, conditions, and resources needed to continue to do their jobs at a high level. This, in my thinking, would be a “fixed windows” approach to educational reform.

The promise to close the achievement gap between advantaged and disadvantaged students in this country has been codified into federal law. The apparatus—though hastily devised and wobbling under the weight of our expectations—is here. The knowledge of where we want to go has been here for a long time. We now need leaders who are unafraid of the inevitable opposition from entrenched interests on either side of the educational reform debate. No Child Left Behind’s liberal promises demand and deserve nothing less.