Inadequate
As almost everyone knows by now, the central aim of the No Child Left Behind (NCLB) law is to make every public-school student proficient in reading and math by the year 2014. It is a laudable goal, as the overwhelmingly bipartisan congressional support for the legislation in 2001 proved. The law’s drafters even had the foresight to know that fixing the deficits in student proficiency would be accomplished within the allotted time only if “each State” established “a timeline for adequate yearly progress,” or AYP, that would steadily close the gap between current levels of performance and the ideal proficiency level each state established. Accordingly, adequate yearly progress toward proficiency on the part of every student became the heart and soul of NCLB.

Unfortunately, however, four years into the life of the law—and fewer than ten years from 2014—there are signs of an irregular heartbeat. Though NCLB is absolutely correct in insisting that schools make measurable improvements on the way to the 2014 goal—and is right also to demand that they do so with subgroups of students who have lagged behind others in the past—those responsible for implementing the legislation have yet to find the best way of giving concrete meaning to each of the key words: proficiency, adequate, progress, every, yearly.

The spirit of NCLB—and even the language included in the actual provisions of the legislation—is right. The implementing regulations, however, too often thwart the estimable intent of the law. When measuring progress, the legislation repeatedly states that AYP assessment should be based on statistics that are scientifically valid, but the regulations that help states abide by NCLB sometimes guide them toward statistically invalid calculations. Thus a school may be identified as failing when in fact it is not. Such mistakes undermine the legitimacy of the law, not just among parents and teachers, but even among administrators and legislators who are proponents of accountability.

By Caroline M. Hoxby
Because so many of the key issues are administrative, a legislative fix is not necessary to address the basic problems. In this essay, I want to suggest five relatively easy fixes—improved ways of giving operational meaning to “proficiency,” “adequate,” “yearly,” “progress,” and “every”—that are designed to fill the worst potholes on the current road. The fixes are relatively easy because they can be implemented without new legislation, are respectful of the autonomy of the states, and are fairly easy for schools to understand. Just as important, they will neither penalize schools unfairly nor dilute NCLB goals and objectives.

The Spirit of AYP
Before turning to these fixes, though, I wish to take a few moments to emphasize some of the basic principles of NCLB. A core principle of NCLB is that every student must reach the desired level of performance: no group of students—minority, disabled, poor, limited English proficient, mobile—should be left behind.

Another core principle of NCLB is that every child is capable of attaining proficiency, defined in an appropriate way. Thus, while progress is important, NCLB deliberately emphasizes reaching proficiency, not making gains each year, regardless of past performance. NCLB provides no special recognition to students or schools that exceed the minimum. This is not a good thing or a bad thing, but it clearly demonstrates that the focus of NCLB is on bringing low-achieving students to a sound level of academic achievement.

A third principle of NCLB is that it works through the states, long the workhorses of the country’s education system. States and localities provide more than 90 percent of funding for schools, so it makes sense for them to exercise control. Furthermore, with fewer schools to watch, states are in a much better position than the federal government to monitor multiple targets. Thus, even though NCLB monitors only proficiency, it encourages states, in their own accountability systems, to reward schools that make gains along the entire spectrum of achievement.

These three principles—attention to every child, a focus on reaching proficiency, and respect for state autonomy—lie at the core of NCLB. The five relatively easy fixes that I propose—benchmarking state definitions of proficiency, measuring progress scientifically, reporting clearly whether the school is making adequate yearly progress, encouraging participation by every student in test taking, and taking seriously the requirement that schools track yearly progress—can be implemented immediately.

1. DEFINING PROFICIENCY:
Use a National Benchmark to Clarify State Standards
NCLB allows each state to determine the level of proficiency that a student needs to achieve in reading and math. While this provision of the law respects state autonomy, it has given rise to a wide variation in state definitions of this key concept. Ironically, the states that took accountability the most seriously before NCLB set high standards for their students and
The spirit of NCLB—and even the language included in the actual provisions of the legislation—is right. The implementing regulations, however, too often thwart the estimable intent of the law.

tough proficiency standards, one needs to find a common benchmark that allows for objective comparisons among the states. Fortunately, a quite good benchmark, the National Assessment of Educational Progress (NAEP), is readily available. It is administered in every state to a representative sample of children in grades 4 and 8. Although it is not a perfect bridge between states’ tests—it is not, for instance, administered to students in all the grades that are required to be tested under NCLB—it is by far the best available. (How each state currently performs on this benchmark is shown in this issue of Education Next, page 52.)

Were the federal government to give its official imprimatur to an interstate benchmark, it would greatly enhance the transparency of state proficiency standards. It would also encourage states with low standards to raise them and to discourage others from reducing them as a strategy for helping local schools achieve compliance with AYP provisions.

Benchmarking will also give the U.S. Department of Education a scientific basis for deciding whether some states deserve extensions of the time to reach proficiency because they have set their proficiency standards significantly higher than those of the typical or average state. In sum, it will encourage states with ambitious proficiency standards to keep them, expose proficiency standards that are too modest, and inform political debates within the states themselves.

2. MEASURING PROGRESS: A Statistically Valid Approach

The purpose of NCLB is to make sure that every child reaches proficiency, and measuring AYP is meant to help schools figure out whether they are on track to do so. Unfortunately, current methods of measuring AYP sometimes incorrectly evaluate the progress of schools. This is troubling because misclassifying a school undermines the credibility of AYP itself.

Measuring AYP is in fact a fairly straightforward statistical activity. To use a school’s existing record of performance
similar to the method used to project the arrival time of airplanes, the trajectory of a hurricane, and so on. Conventional statistics give us the most likely forecast and also give us high-end and low-end forecasts. The same principles can be applied to forecasting progress toward proficiency under NCLB.

A major benefit of using conventional forecasting methods is that they automatically determine, with a specific level of confidence, whether ethnic, low-income, or other subgroups are making adequate yearly progress. Currently, the crude manner in which the progress of these subgroups is measured has stirred controversy, because the number of such students varies from one school to the next—and is often too small to permit a confident forecast. The statistical technique I have described will standardize and clarify the degree of accuracy that can be achieved.

My proposed forecasting method is quite different from the technique now used, which includes or excludes subgroups based on the number of students falling into a subgroup. A subgroup may be shown as failing even though it is actually making AYP according to conventional statistics. Moreover, the current technique treats schools differently if they fall slightly above or below the subgroup threshold, often without good statistical reason for doing so. Such anomalies matter because a whole school will fail to make AYP if a single subgroup fails. Thus, if a statistically invalid test is applied to a particular subgroup, AYP for the whole school can end up being wrong. When there is an error in measuring AYP, it undermines the credibility of NCLB itself. A standard forecasting technique that can be applied consistently to every school is more consistent with NCLB principles and will enhance the legitimacy of the legislation among administrators, teachers, and parents.

A second major benefit of the method described above is that it automatically takes account of progress that students make toward proficiency, even if they do not cross the proficiency threshold. There is no need for the current unsatisfactory, difficult-to-explain “safe harbor” provision, which is the partial substitute for what I am proposing. Even those students now performing far below proficiency will help a school make AYP if those students are improving quickly enough.

3. ADEQUACY: Conveying Its Meaning Readily

The adequacy of a school’s progress ought not to be mysterious. It ought to convey what it means to make adequate yearly progress—in other words, what it means for a school to be on track or off track toward the goal of universal proficiency. Thus an AYP report should contain figures (with graphics) showing where the school is currently and where the school needs to be in the year 2014. The figures should show the best forecast of where a school will be in each year between the current year and 2014. They should indicate low-end and high-end forecasts so that readers understand whether the forecast is precise or noisy. A figure ought to be created both for the school overall and for subgroups. Such figures can automatically be generated using the forecasts described above (see Figure 1).

4. EVERY CHILD: Participation in Proficiency Testing

Under current legislation, a school fails to make AYP if less than 95 percent of its students participate in state assessments of achievement or if less than 95 percent of the students in any of its subgroups participate in assessment. (If a subgroup is too small for the AYP calculation, it is currently exempt from the participation requirement.) The goal of the participation provision is excellent. Nevertheless, the rule can be improved by placing it on a sound statistical basis.

Instead of requiring 95 percent participation, regulations should simply assign the minimum score to any student who does not participate. Since all students will score at or above the minimum score, every school has a strong incentive to...
ensure that every child take the test during the weeks set aside for this purpose. And a school will be penalized only to the extent that the minimum scores recorded for nonparticipants drag down the overall score for the school or the relevant subgroup.

The calculation I propose is statistically valid, amply rewards schools that maximize participation, and penalizes schools that fail to get full participation. In contrast, a 95 percent participation cutoff is an arbitrary way to determine whether a school is failing to make AYP.

5. YEARLY PROGRESS: Keep It Simple

Finally, inadequate yearly progress should be taken seriously: a student’s achievement should not be fully factored into a school’s AYP unless the student has been enrolled for at least 90 percent of the year since the last test was administered. If students switch schools within a local education agency, their achievement should be factored into AYP with a weight equal to the share of the year that they spent in each school.

Back to Basics

AYP is the heart of NCLB. Many consequences of NCLB depend on it. Moreover, if the goal of NCLB—making every child proficient—is a correct one, then it is crucial to track whether we are on the path to achieving that goal.

Fortunately, reasonable administrative action can correct deficiencies in the way in which AYP is measured and reported today. AYP can be refined simply by paying closer attention to the operational definitions of key words in the law. We need to benchmark state definitions of proficiency, measure progress by forecasting how well each school is moving toward the 2014 goal, publicize adequacy by means of simple figures that show where each school stands, encourage schools to test every child by assigning minimum scores to those who are not tested, and hold schools accountable for only that portion of the year the child spent in the school.

Caroline M. Hoxby is professor of economics at Harvard University, the director of the Economics of Education Program at the National Bureau of Economic Research, and a member of the Koret Task Force at the Hoover Institution, Stanford University.

“Highly Recommended” — Library Journal

Ask for The Common Review at your favorite library, newsstand, or bookstore. Better yet, subscribe to The Common Review, the magazine of the Great Books Foundation ($12 US/$20 Int.), and become part of America’s most thoughtful reading community. Visit us online at www.thecommonreview.org for selected content and more, or call 1-800-222-5870 to subscribe today.