



## The Case for Being Mean

By Frederick M. Hess

*Performance-based accountability promises to ensure that every student, regardless of background, masters crucial knowledge and skills. But to realize that promise, accountability needs to be coercive, that is, it must confront failure with real consequences for both educators and students.*

The enactment of the federal No Child Left Behind Act in January 2002 made performance-based education accountability a federal mandate. The legislation followed a decade of concerted state activity across the states that produced an array of high-stakes accountability systems. Those state systems have already come under fire. In such places as Nevada, Florida, and Massachusetts, where thousands of high school seniors are at risk of being denied diplomas in 2004, angry parents are protesting, civil rights groups have threatened boycotts over the high rates of failing minority students, and educators worry that their schools will be targeted by state education agencies as low-performing or inadequate.

The allure of performance-based accountability is its promise to ensure that *all* students, even the most disadvantaged, will master crucial knowledge and skills. An overwhelming percentage of adults, often 90 percent or higher, support accountability in the abstract, recognizing the appropriateness of holding public educators responsible for teaching essential material instead of permitting them to use public classrooms as personal forums. Aside from a few ideological critics, even most educators are sympathetic to the goals of performance-based accountability. The important split is not between

ideological proponents and opponents of accountability, but between those who support tough-minded accountability, despite all its warts, and those who like the ideal of accountability but shrink from its reality.

### Nice versus Mean Accountability

Simply put, there are two kinds of accountability: suggestive and coercive, or, more plainly, “nice” and “mean.”

Advocates of nice accountability presume that the key to school improvement is to provide educators with more resources, expertise, training, support, and “capacity.” They view accountability as a helpful tool that seeks to improve schooling by developing standards, applying informal social pressures, using tests as a diagnostic device, increasing coordination across schools and classrooms, and making more efficient use of school resources through standardization. The educational benefits produced by nice accountability depend on individual volition.

Mean accountability, on the other hand, uses coercive measures—incentives and sanctions—to ensure that educators teach and students master specific content. Students must demonstrate their mastery of essential knowledge and skills in the areas of math, writing, reading, and perhaps core disciplines at certain key points and before graduating from high school. Educators are expected to do

---

Frederick M. Hess is a resident scholar at AEI. His next book, *Common Sense School Reform* (Palgrave), is due out in 2004. A version of this article appeared in the November 2003 issue of *Educational Leadership*.

what is necessary to ensure that they no longer pass on students unequipped for the most fundamental requirements of further education, work, or good citizenship.

In such a system, school performance no longer rests on fond wishes and good intentions. Instead, such levers as diplomas and job security are used to compel students and teachers to cooperate. Mean accountability seeks to harness the self-interest of students and educators to refocus schools and redefine the expectations of teachers and learners.

For educators, mean accountability offers many benefits that nice accountability does not. Unlike its nicer variant, mean accountability gives the school and district leadership personal incentives to seek out and cultivate excellence. It enables policymakers to roll back regulations designed to control quality by means of micromanaging procedures. It builds popular support for education by providing state officials and voters with hard evidence on school performance. And, in well-run schools and districts, mean accountability gives effective teachers new freedom to teach as they see fit and with the materials they deem appropriate, as long as their students master essential skills.

Advocates of mean accountability agree that nice accountability yields real benefits, but they point out that these benefits have been only modest and uneven. The 2002 National Assessment of Educational Progress reported that just 33 percent of U.S. fourth graders and 36 percent of twelfth graders scored at least at the “proficient” level in reading; 36 percent of fourth graders and 26 percent of seniors scored “below basic.” The results are far worse in urban communities, where two-thirds of fourth graders are routinely reading at “below basic” level.

The split between those who insist on mean accountability and the gentler souls comes down to whether one agrees with nice-accountability proponents that educators are doing all they can, that student failure is caused largely by factors outside the control of teachers or administrators, and that incentives will not productively alter educators’ behavior.

Proponents of coercive accountability reject such claims. Common sense tells us that people work more effectively when we hold them accountable for performance, reward them for excellence, and give them opportunities to devise new paths to success. In any line of work, most employees will resist changes that require them to take on more responsibility, disrupt their routines, or threaten their jobs or wages. To overcome such

resistance, we need to make inaction more painful than the proposed action. In education, this means making a lack of improvement so unpleasant for local officials and educators that they are willing to reconsider work rules, require teachers to change routines, assign teachers to classes and schools in more effective ways, increase required homework, fire ineffective teachers, and otherwise take those painful steps that are regarded as “unrealistic” most of the time.

The idea is not to simply lay more weight on the shoulders of teachers or principals. The challenge is more fundamental. In any line of work, decision-makers want to avoid unpopular decisions. But sometimes school officials have to make painful choices: to drop a popular reading program that is not working; to cut elective choices if students have not mastered the basics; to fire a well-liked principal who is not achieving results. In each case, the easier course is to not act. The way we force people to make unpleasant choices is by pressing them to do so—even if it angers employees or constituents. Coercive accountability provides the best, most straightforward way to bring that pressure to bear in support of core academic subjects.

## Rethinking Systems and Practices

For decades, U.S. schools have been constantly reforming without ever really changing. As long as we give veto power over change to those who will endure its costs, we will continue to shy away from reinventing schools as more efficient and effective organizations. We will not force painful improvement by convincing those who bear the costs of change that it really is for the best. We must leave them no choice in the matter.

It is not just a question of making people work harder, but of forcing managers and leaders to rethink systems and practices. Take the Detroit automakers who fell upon hard times in the late 1970s. They were producing oversized and poorly designed cars, had gotten lazy about quality control, had permitted costs and union contracts to spiral out of hand, and had added layer upon layer of middle management. The emergence of fierce foreign competition and a dramatic loss of market share shocked these firms into action. Energetic new leadership rethought the product line, redesigned quality control, slashed middle management, renegotiated contracts, and cut costs. The aim of transformation was not to berate workers; it was to force those in charge to focus on high performance and make painful decisions to achieve it.

Today, district and school leaders spend their time pleading with their subordinates to cooperate because they can imagine no other ways to drive change, but they are mistaken. We can drive change by requiring educators to meet clear performance goals and attaching consequences to success or failure.

## Ambivalence about Being Tough

Although public officials and educators are sympathetic to the notion of accountability, they are often squeamish about the demands of coercive accountability. The benefits of accountability—a more rigorous and focused school system—are broad and widely dispersed and often hard to isolate, whereas the costs are borne by visible students and teachers, many of whom can inevitably point to various extenuating circumstances. A Texas principal, after affirming that she believed in rigorous standards for student learning, expressed the ambivalence felt by many:

Last year I had to tell a student that she didn't pass the "last chance" TAAS exam administered in May of her senior year. I do not even want to imagine the heartbreak that she and her family felt. I've only had to do this once, but it was one time too many, and I don't know that I have it in me to do it again.

Accountability requires education officials to make five politically sensitive sets of decisions. First, they must designate a prescribed body of content and objectives to be tested. Such a course necessarily marginalizes some other goals, objectives, content, and skills. Second, officials must impose assessments that accurately measure whether or not students have mastered the requisite skills and content. Third, they must specify what constitutes mastery. Fourth, they need to decide what to do with students who fail to demonstrate mastery. Finally, for accountability to significantly alter education programs and practices, the system must reward or sanction educators on the basis of student performance.

Each decision tends to produce passionate opposition among those who bear the costs. Opponents of coercive accountability seize upon the arbitrary nature of many of these decisions, demanding modifications that will increase test validity and reduce any inequities or pernicious effects produced by misuse of assessments.

Proponents of coercive accountability often have trouble holding the line against the appeals of aggrieved constituencies. In the face of heated opposition, proponents often agree to a series of compromises on program design and implementation that eventually undercut the coercive promise of accountability.

For example, although most states have adopted mandatory graduation exams and about half offer school incentives linked to test scores, phase-in periods and implementation delays mean that graduation requirements and performance-based incentives for educators have taken effect in only a few states. Delays and adjustments may provide time to refine tests and curriculums and potential penalties, but they also conveniently push substantive challenges into the future.

To date, most states that have actually started to approach deadlines have blinked and delayed the implementation of sanctions. An analysis published in 2000 found that roughly one-third of the states that had adopted high-stakes accountability systems had slowed or scaled back their original efforts. In Arizona, for instance, when more than 80 percent of tenth graders failed the state math test in 1999 and 2000, the state board of education and the legislature scrambled to push back the graduation requirement from 2002 to 2006. In recent years, other states—including Alabama, Alaska, California, Delaware, Maryland, North Carolina, and Wyoming—scaled back testing programs or postponed their effects.

If policymakers don't delay implementation, they often soften accountability in various other ways. Although each accommodation can be justified on educational grounds, each of these common compromises also serves to dull the mean edge of accountability:

- *Lowering the stakes of the tests for students, for educators, or for both.* Weak or nonexistent sanctions offer little incentive for teachers, low-performing students, or anyone else to worry much about test results.
- *Making tests easier.* This can be done by lowering content standards, adopting easier questions, or reducing the cut-off scores for satisfactory performance.
- *Offering lots of second chances.* Giving students a number of retests or a teacher several years to boost his or her performance means that the law of averages will help a number of moderately low performers to clear the bar.
- *Permitting some students or educators to sidestep the required assessment.* Some schools issue a "basic"

diploma in lieu of a standard diploma, and some administrators exempt teachers who teach specialized classes from evaluation.

## The Temptations of Compromise

From the inception of high-stakes testing, proponents have tended to laud the requisite tests and accompanying systems as clear, scientifically defensible, manageable, and concise. Critics typically attack the tests and systems as unreliable, simplistic, overly focused on trivia, or lacking the necessary curricular and pedagogical support. They argue that linking teacher incentives or student advancement to anything so crude will pose inevitable perils. In truth, both sides are correct.

The details of accountability—the content to be tested, the assessments to use, the definition of minimum competency, and how to address the performance of educators or students—are inherently arbitrary. The closer one gets to crafting and enforcing standards, the less defensible specific program elements can appear.

Determining what students need to know, when they need to know it, and how well they need to know it is an ambiguous and value-laden exercise. Neither developmental psychologists nor psychometricians can “prove” the necessity to teach specified content at a particular grade level. Such decisions are imperfect judgments about students’ needs and capacities.

Proponents have difficulty standing firm on program details precisely because decisions regarding what students need to know, when they need to know it, and how well they need to know it are only reasonable approximations. No amount of tweaking will yield a perfect instrument.

Loath to concede that graduation testing is inevitably flawed, proponents try to placate critics with one “refinement” after another. They soften sanctions, adjust passing scores, offer exemptions, fiddle with school performance targets, delay implementation, and take other similar steps as they seek to discover just the right balance. Unfortunately, the painless, happy medium is fundamentally at odds with the purpose of coercive accountability. This series of compromises may preserve

the façade of accountability but will eventually strip accountability of its power.

## The Importance of Being Mean

The challenge for proponents of coercive accountability is to acknowledge the localized pain and dislocation that they intend to visit upon some educators and students as the price of a system that will ensure that educators are serving all of our nation’s students. The challenge for those enamored of nice accountability is to explain how they plan to ensure that schools prepare all students for their adult lives. Although their caveats about inequalities in home environments and natural student abilities have merit, surely it is not overly ambitious to demand that educators find a way to teach all students the essentials of reading, writing, math, and the rest of the key disciplines before sending them into the world.

Most accountability programs begin with at least a rhetorical commitment to the transformative ideal. Over time, critics weaken such systems, often while espousing their support for the principle of accountability. These critics trace their opposition to the specifics of existing arrangements, stating that they will support transformative accountability if only . . . it is stripped of its motivating power.

The choice is between an imperfect accountability system and none at all. In the absence of coercive accountability, we have seen how easy it is to graduate ill-equipped students and excuse inadequate school performance—especially among the most disadvantaged students.

In the end, standards are a useful and essential artifice. They and the accountability systems they support must be defended as such.

If accountability finally becomes part of the “grammar of schooling” for parents, voters, and educators, then its performance benchmarks for ensuring that students are learning, teachers are teaching, and schools are serving their public purpose will become accepted practice. State and federal officials now face the question: will accountability fulfill this potential or become another hollow rite of spring?