

ROOT & BRANCH



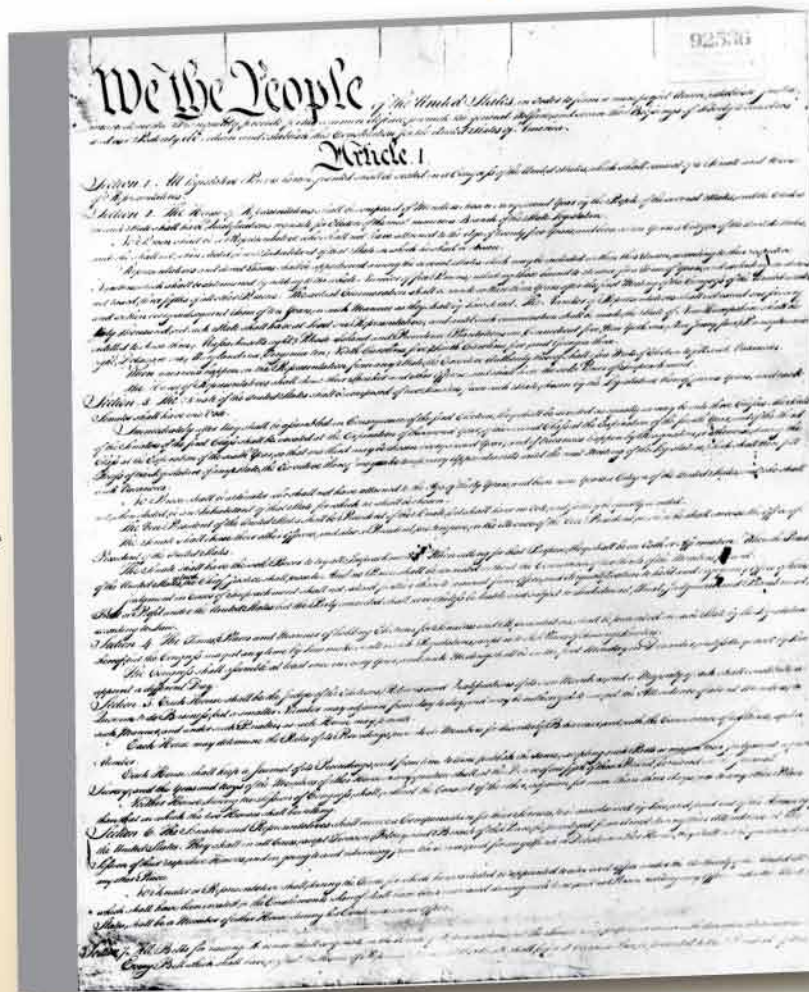
"Growing Bottom Up Movements"

Fall 2009 - Volume II Issue 2

www.advancementproject.org

The Next Phase of Election Reform

Constitutional Right to Vote



Detroit's Dropout Crisis:

A Child is **NOT** A Test Score

Race in Review

Are Whites the **NEW** Blacks?

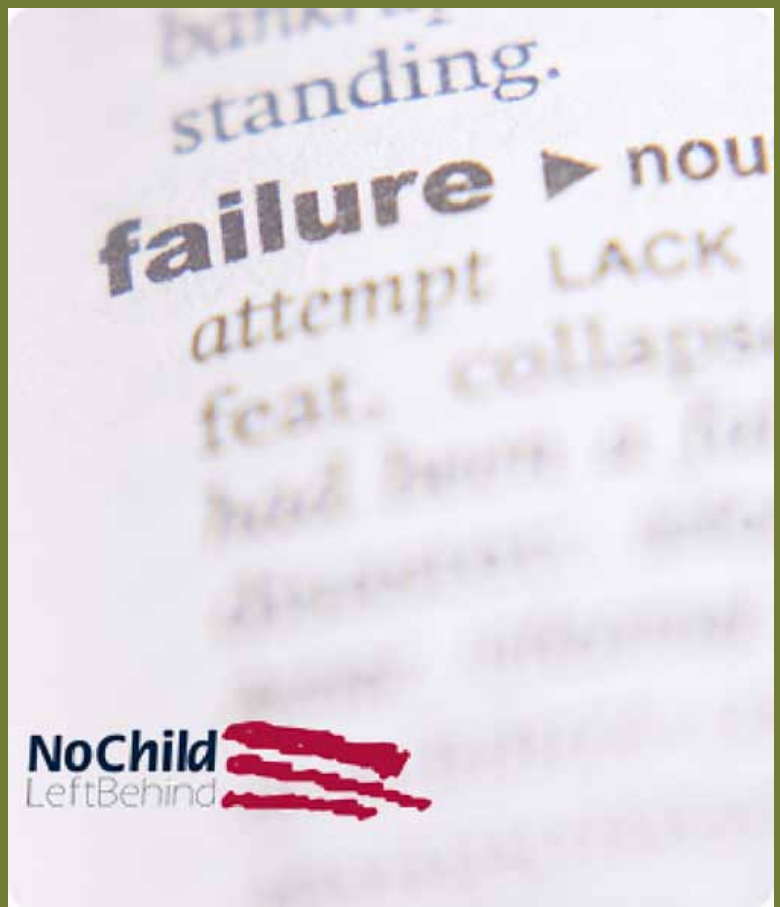
www.rootandbranchmagazine.org

www.justdemocracyblog.org

A Child Is Not a Test Score:

Assessment as a Civil Rights Issue

by: Monty Neill



“The function of education is to teach one to think intensively and to think critically. Intelligence plus character - that is the goal of true education.”

- Martin Luther King, Jr.

Introduction

On the campaign trail, President Obama declared, “We should not be forced to spend the academic year preparing students to fill in bubbles on standardized tests,” and he called for “a broader range of assessments that can evaluate higher-order skills.” The nation and its students need assessments in all the important areas, both so the public knows what is happening in schools and to avoid narrowing curriculum and instruction to fit tests that cannot indicate real success and readiness for future learning.

If the nation’s goal is a high-quality education for all, why not use assessments that can at least tell us if that goal is being met? Why not rely on multiple sources of evidence to inhibit narrowing curriculum and teaching to one test format? Why not make decisions about students and schools based on information gathered over time? Why not transform assessment and accountability

to serve the educational needs of all students? A truly healthy educational system will prioritize high-quality classroom instruction and use school-based assessment information to monitor classroom, school and district progress. Sadly, the nation’s public education system does not function that way.

Standardized tests have been prevalent for much of the latter half of 20th century. The emphasis on standardized testing has intensified in recent decades as elected officials, business leaders and others have fostered the idea that the U.S. economy will decline unless student achievement and school progress is increasingly monitored through testing.

In 2002, former President George W. Bush won passage of the No Child Left Behind Act (NCLB), pushing the emphasis on multiple-choice, paper-and-pencil tests to new heights. Under NCLB, an ever-escalating percentage of students in every public school and district is expected

to score at a proficient level on statewide standardized tests each year. Students as a whole and also specific ethnic and racial groups must meet this “Adequate Yearly Progress” (AYP) requirement.

The pressure to pass standardized tests intensified dramatically under NCLB. Schools that repeatedly fail

then to northern states with large populations of students of color in their cities, such as New York, Ohio and New Jersey. In 1995-96, 11 of 16 states in the Southern Regional Education Board had exit exams, compared with only six of the remaining 35 states (including DC). Those states also tested an average of 7.5

Teachers often lack the authority to deviate from the mandated curriculum regardless of student needs, emerging issues or the teachers' recognition that these curricula fail to prepare students for future success. School staff fear that without narrowing the curriculum and tailoring the instruction to fit the tests, their students fare poorly, putting the students themselves and their schools at risk of severe sanctions.

to make AYP face escalating sanctions, culminating in “restructuring,” which can include replacing a school’s staff or turning it over to private management. Schools that have struggled the most to make AYP are those with the highest percentages of poor students, which typically have fewer resources. In these schools, teachers are frequently expected to rigidly “deliver” a pre-programmed, often tightly scripted curriculum, each day covering a set of skills to prepare students for the tests. Teachers often lack the authority to deviate from the mandated curriculum regardless of student needs, emerging issues or the teachers’ recognition that these curricula fail to prepare students for future success. School staff fear that without narrowing the curriculum and tailoring the instruction to fit the tests, their students fare poorly, putting the students themselves and their schools at risk of severe sanctions.

One risk students face in a majority of states is the graduation test. These tests began in Florida in the late 1970s. A lawsuit delayed the use of the graduation test on the grounds that many Black students had not had a fair opportunity to learn the material on which they were tested because they had attended schools segregated by law. The courts ruled that once students that had begun school after the end of de jure segregation had graduated, the tests could be used. This ruling ignored the extensive de facto segregation and the vastly disparate resources available to Blacks and Whites.

Graduation tests quickly spread across the South and

grades, substantially higher than the national average of 5.28 grades. In effect, the worst-performing systems and those with the highest proportions of African Americans were most likely to implement high-stakes testing.

The next wave of states to enact graduation tests--after a mid-1990s halt in the growth of such tests--were disproportionately Latino. New Mexico and Texas, which imposed exit exams in the first wave, were joined by Arizona and California, for example. States with tests comprise about 70 percent of the nation’s student population, but over 80 percent of its African American and Latino students. The states without graduation tests form a belt from Illinois to Idaho, and north of Oklahoma – and in most cases have predominantly White student populations. Thus, there is clearly a racial dimension to the use of graduation exams, and youth of color, those who speak English as a second language or who have a disability or are from low-income families are disproportionately denied a diploma because of a test score.

The same is just as true for tests students must pass to move to the next grade, which are found in Florida, Louisiana, Texas and many large cities, such as Chicago and New York. As with diploma denial, the damage of grade retention falls disproportionately on youth of color. Extensive research has demonstrated that students who are held back progress more slowly than comparable students who are promoted, they suffer significant loss of self-esteem, and they are far more likely to not graduate.



The Education Consequences of High-Stakes Testing

Across the country, students have exposed the damaging educational consequences of high-stakes standardized tests. They have decried being denied a diploma because of a test score and exposed the way incessant test preparation deforms curriculum, instruction and learning.

Macario Guajardo, a 16-year-old from south Texas who for years boycotted the state's standardized test, the Texas Assessment of Knowledge and Skills (TAKS), explained to a state legislative committee on education reform the consequences for learning: "When I was in elementary, schools were basically like a TAKS factory, and students were almost like little robots. I don't remember there being any room for serious, creative and critical thinkers."

Carolyn, also aged 16, wrote in the *California Bee*, a daily newspaper, "District tests, including the high school exit exam, should be eliminated since there is no educational point to them... Too much classroom time is wasted on test preparation and taking tests. That time

should be spent on actual learning of subjects, not on the steps of how to eliminate answers" (i.e. incorrect options on multiple choice questions). She added, "The focus of our education system should not be based on tests, but on the individual needs of students."

The sharp disparities in educational opportunity are also visible to at least some students. Afrisha Lavine from Akron, Ohio, compared her school to a wealthier nearby school: "If they...put the same programs in the failing schools, then they would be good schools. The failing schools are not bad, it's just that they have a great disadvantage."

Jackie, a Boston high school student, similarly explained, "In going to other schools and finding what opportunities other students are getting made me realize what opportunities I am not getting at my school."

Districts in poorer communities, especially communities of color that have fewer qualified teachers and inadequate

books, laboratories and libraries, are expected to perform at the same levels as districts that have far more financial and educational resources. The inequity is compounded when districts gut art, music classes and sports for the rote memorization, constant quizzing and testing that limit time for creative and analytical thinking. Wealthier districts, whose students are better prepared for these tests, devote far less time to test preparation and don't suffer the impact of a narrowed of curricula.

A Californians for Justice report explained, "Any conversation with high school students from around the state reveals that students are extremely demoralized by the exit exam. It is clear that large numbers of students of color, low-income students, and immigrant students, feel that their futures are being destroyed by a test for which they have not been prepared."

As Boston student Gregory pointed out, "They are just training us for the workforce... trying to train you to sit in one place and do simple operations for eight hours." Caroline added, "You always have a lot of people saying that you kids are the future. But how can we be the future if we are not getting what we need?"

The stories of educational damage occur and recur because high-stakes standardized testing has come to dominate learning and class time. Tests are widely used as a sole hurdle for student grade promotion, graduation, or program placement, and they control opportunities, curriculum and instruction in the name of accountability.

The interaction of under-resourced schools and testing most powerfully hits students of color. They are disproportionately denied diplomas or grade promotion, and the schools they attend are the ones most likely to fare poorly on the tests and face sanctions such as restructuring.

Professor Gloria Ladson-Billings uses the term "education debt" to explain the lack of adequate educational opportunity for African-American students accumulating since slavery and segregation. She thinks that focusing on this inequality is far more meaningful than the commonly used "achievement gap," which is only refers to unequal test results. The debt includes the school-based debt in resources. It also includes the housing debt that forced people of color to suffer inferior living conditions, exemplified by the racial covenants that ensured African Americans could not move to many suburbs after World War II. Billings speaks also of the medical care debt, the pervasive historical and current unequal access to medical care by race, and the employment debt--African American families earn three-fifths of what White families earn while U.S. income inequality grows rapidly.

Test-based "school reform" such as NCLB, which passed

with support from both Democrats and Republicans, is an effort to improve results while ignoring the existence of the education debt.

The tools used to improve results -- tests and sanctions -- actually make things worse. Low-income students, who are disproportionately children of color, go to under-resourced schools that serve up a thin gruel of test preparation. So long as such a system remains in place, the pipeline to college and good jobs for low-income and minority-group youths will remain narrow, but the pipelines to prison and unemployment will remain wide.

High-stakes testing undermines school quality. What is it about the use of standardized tests as the primary, even sole arbiter of school quality that is problematic? Partly it is because, in the face of escalating sanctions, some schools and districts have taken harmful actions such as increasing suspensions and expulsions of low scorers--removing perceived problem kids from the classrooms instead of dealing with their problems. And partly is it the damage done to teaching.

Testing's control over teaching is unevenly applied. The drill-and-kill school practices that guarantee students will not be ready for college, skilled employment, lifelong learning or effective citizenship are most prevalent in schools serving low-income children of color. No one has documented this more powerfully than Jonathan Kozol in *Shame of the Nation*. Building on his earlier exposé, *Savage Inequalities*, of the vastly unequal opportunities provided in different communities across the nation, Kozol describes in painful detail the brain-deadening, emotionally stultifying consequences of scripted curricula and test preparation in what he terms "apartheid education."

Suburban middle- and upper-class schools succumb to a degree to teaching to state exams, but teaching to the test is nowhere near as prevalent or powerful in those communities. And the suburban schools certainly do not employ the tightly scripted curricula widely used in urban schools.

The learning gaps revealed by standardized tests mask worse gaps in more advanced learning skills. For example, students in well-to-do schools typically learn to write research papers, which colleges expect students to do. There are no research papers on standardized tests. If the primary goal is to boost test scores so students, teachers will not take time out to teach needed research and writing skills. As noted psychologist Robert Sternberg wrote, "The increasingly massive and far-reaching use of conventional standardized tests is one of the most effective, if unintentional, vehicles this country has created for suppressing creativity." That suppression, too, most powerfully affects students who are most subject to the tests.

The Impact of Low Graduation Rates



With tests as one key factor, African-American and Latino graduation rates barely reach 60 percent. The consequences are severe. Non-graduates have significantly lower lifetime earnings and less stable families, they are more likely to be unemployed or imprisoned. Graduation rates for low-achieving minority students and girls have fallen nearly 20 percentage points since California implemented high school exit exams, according to *Effects of the California High School Exit Exam on Student Persistence, Achievement, and Graduation* a research paper published by Stanford University's Institute for Research on Education Policy and Practice. In 2007-8, 40,000 more students failed to graduate than did so in pre-test years.

Similarly, more than 40,000 Texas students were denied diplomas in 2007 because they did not pass all four parts of the state's graduation exam. These casualties are a direct result of high-stakes accountability systems designed to maximize test scores.

Current research by John Robert Warren and his colleagues clearly demonstrates that graduation tests increase the number of dropouts, do not lead to improved test scores, and do not produce better results in college or employment. They are, in effect, wholly negative.

The Standards for Educational and Psychological Testing, produced by the American Educational Research Association, American Psychological Association, and National Council on Measurement in Education, warn against these practices: "[A] decision... that will have major impact on a student should not be made on the basis of a single test score." Similarly, the American Evaluation Association concludes, "High stakes testing leads to under-serving or mis-serving all students, especially the most needy and vulnerable, thereby violating the principle of 'do no harm.'" "Policy makers have ignored the wisdom of the very people who make, use and research tests.

Civil rights organizations have long battled these make or break tests. They point out that systems that deny diplomas or promotions based on test scores typically fail to provide an adequate or equitable opportunity to all students to learn the material on which students are tested. This places the burden of accountability on the backs of children, hitting children with the worst education systems and the fewest resources hardest.



Grade Promotion Also Hurts Disproportionately

African-American and Latino children are more frequently retained in grades than are Whites. In Chicago in 2008, 98.6% of Whites passed the grade promotion test, compared with 85.5% of African Americans. These disparities have not changed much over the years. Since 2002, 12.9% of Chicago's Black students have been held back while only 2.3% of Whites have faced the same fate. In 2008, 5.4% of Latinos were retained.

Chicago-based researchers evaluated the consequences and concluded that retention is harmful. Retained students did less well academically than comparable students who were promoted, and retention increased the likelihood of dropping out.

The Chicago studies confirmed decades of previous African-American and Latino children are more frequently retained in grades than are Whites. In Chicago in 2008, 98.6% of Whites passed the grade promotion test, compared with 85.5% of African Americans. These disparities have not changed much over the years. Since 2002, 12.9% of Chicago's Black students have been held back while only 2.3% of Whites have faced the same fate. In 2008, 5.4% of Latinos were retained.

Chicago-based researchers evaluated the consequences and concluded that retention is harmful. Retained students did less well academically than comparable students who were promoted, and retention increased the likelihood of dropping out.

The Chicago studies confirmed decades of previous research showing that flunking students diminishes their self-esteem, reduces their likelihood of graduation, and fails to increase achievement. Because grade retention is harmful and test-based policies lead to more retention overall with disproportionate increases for African Americans and Latinos, test-based retention intensifies race-based inequalities in school systems, such as Chicago's public schools.

There is a ready solution to the "social promotion" versus "retention" dispute: promote students, but provide intensive extra support to those who are not doing well, as soon as academic problems are identified. Providing such support would help schools avoid inflicting the damage of retention while being able to offer the help many students need. Helpful support would focus not on boosting test scores but on strengthening real academic knowledge and skills.



Better Assessments in an Improved System

A good assessment—understanding what has been learned and student learning processes --is essential for

teaching and learning. It is also a core component of holding students, teachers, schools and districts accountable for their work, class time and resources. While the processes of assessment and accountability often overlap, they do not go hand in hand. Many assessments used in education have nothing to do with accountability. And demonstrating a student's, teacher's or school's success, in short, accountability, should involve far more than academic assessments.

In *Grading Education: Getting Accountability Right*, authors Richard Rothstein, Rebecca Jacobsen, Tamara Wilder wrote that their research showed that the general public, legislators and school board members view the purposes of education broadly and think it should serve many purposes, including the teaching of: academics, critical thinking, arts and literature, preparation for skilled work, social skills, work ethics, citizenship, physical health, and emotional health. Other researchers have come up with similar results.

In *Empowering Schools and Improving Learning*, the Forum on Educational Accountability proposes accountability structures that would look at inputs, what students are getting on the front end, including the quality of health care and housing in addition to teacher quality and school resources. FEA proposes that schools and districts collaborate with families and communities to meet the needs of the whole child -- cognitive/intellectual, social, civic, emotional, psychological, ethical, and physical -- while preparing them for successful citizenship in a multicultural world.

Assessments should include multiple kinds of evidence, from multiple-choice questions to essays and projects, teacher observations and student self-evaluations. Good teachers know how to use a broad range of assessments and that one can use many different tools to assess knowledge. Unfortunately, pressure to boost scores on standardized tests has reduced the range of assessments teachers use. For example, one teacher, in a FairTest report on NCLB, described how she had to reduce the number of book reports she assigned because of the time required for test prep. These kinds of stories have been told thousands of times across the nation.

Teachers use high quality assessment results to adjust their teaching ("formative" assessing) and to evaluate student success ("summative" assessing). This means that good teachers use a variety of measures to gather a great deal of evidence about student learning. Most of the time, this evidence stops with the teacher. It may show up as a grade, or in a discussion with parents or next year's teacher, but it rarely informs efforts to improve schools or shape policy or provide public accountability.

In short, a much richer sampling of learning is ignored in favor of a narrow set of data called test scores.

Good and meaningful learning often involves more extended work. Assignments and projects produce a great deal of information about the learning process and student achievement. But it is not easy to create meaningful tasks, and no teacher can be expected to create all the tasks she might need. Therefore, districts, states or consortia of districts, or even test companies, should assemble banks of high-quality tasks. The tasks would be available for teachers to use during their courses, as they deem appropriate. The completed tasks become part of the record of student achievement. Some tasks might be administered statewide, but research suggests great caution in trying to make sound inferences based on the results of one or two tasks. One study found it would take between nine and 10 one-hour tasks to be able to make a fair judgment about what a student learned in a high school biology class. No state can or should administer 10 tasks, but a good teacher can do so during the course of the year. However, when regarded as one part of the overall evidence, centrally required tasks can be a useful component of assessing learning.

Gathering evidence over time from the many kinds of work students complete and assembling the pieces in a useful format is not simple. It requires a strong evaluation structure. An example of such an evaluation structure is the Learning Record (LR), first developed in London, England, for use with multilingual, multicultural immigrants. Without dictating the specific content, the LR provides a structure for gathering samples that can illuminate the teacher's evaluation of a student. It provides a means for scoring; in the case of the reading record, by placing students on progress (developmental) scales. If anyone else looks at the LR, its structure enables rapid verification of the teacher's evaluation. If teachers had low expectations or did not have students read much of consequence, it would be revealed. If the students read challenging novels and plays and wrote thoughtful papers, that would be revealed.

"Sampling" procedures, such as those employed by the National Assessment of Educational Progress, also can be used. NAEP is the only nationally representative and continuing assessment of what America's students know and can do in various subject areas. These assessments are conducted periodically in mathematics, reading, science, writing, the arts, civics, economics, geography, and U.S. history. It reports at state and national levels, as well as a few large cities. A state sampling procedure will not produce individual student scores, but rather school-level scores. These standardized tools, which can include

performance tasks, can serve primarily as a check on the system.

New assessment approaches require teachers, administrators and other professionals to learn more. Unfortunately, teachers are poorly prepared in college or in their early years of teaching to do assessments in beneficial ways, even though assessments are a primary component of teaching. Collaborative professional development, primarily at the school level, can include a focus on developing and using better, broader, richer assessments. Educators can then use the resulting assessment data for further staff learning.

Standardized tests can be part of the assessment and accountability mix. However, there is no reason for the federal government to mandate that states test all children with a standardized instrument every year. (No other country comes close to mandating as much testing as the U.S., and many are doing better on international comparisons of student learning than does the US.) If states continue to choose to test annually, those tests should be used in accountability the same way sampling would: Where there are significant shortfalls in a school district, the state would investigate to determine the cause, and if warranted, direct any needed changes to help teachers, students and school administrators.

Using a variety of assessments is the best means to obtain evidence about school quality and student learning. This shift is necessary to get out of the dangerous, educationally destructive trap of high-stakes testing. Escaping that trap does not mean no one will be watching or that schools and districts won't be held accountable. The assessment alternatives outlined here not only provide a richer set of accountability tools, they provide far richer information for improving education. The information provided in a Learning Record or a set of complex tasks can be used to guide improvement efforts with much greater accuracy and effectiveness than the sparse data from 40-to-50-item standardized tests.

U.S. Secretary of Education Arne Duncan has announced his intention to reform assessment as part of the Congressional reauthorization of NCLB in 2010. Although he has said that a testing system should be reformed to fairly and accurately measure student growth, it remains to be seen whether he will push for a real overhaul in the actual assessments as well as their use. Signs such as his pushing for the use of current tests to evaluate teachers are discouraging, and must be opposed along with other high-stakes misuses.

Conclusion

High-quality performance assessments with strong local components within a supportive accountability and improvement system are a vital component of educational opportunity and equity. Such assessments can contribute to improving education where an over-emphasis on standardized tests coupled with misguided accountability procedures does not.

Schools alone cannot overcome poverty, but they can make a powerful difference. A focus on improvement must also address the reality that the nation cannot wish away inequity and inadequacy with the magic wand of testing. Rational efforts must use input, process and outcome information to guide improvement and each school's capacity to serve all children well, as the Forum on Educational Accountability has proposed. This means holding government accountable for providing adequate and equitable resources for all children. Schools like those in the New York Performance Standards Consortium show that while students in the under-resourced schools might not catch up on standardized tests, they can be prepared for the more important goal of succeeding in college because they can learn to think and use knowledge well.

What do we as a nation want to prioritize? Spurious and illusory steps toward equity through standardization, or real improvement efforts in which high-quality assessment is one essential part.

More significantly, how much less should low-income communities and communities of color and their advocates settle for? We might agree that it is pie in the sky to think that all kids will be in schools that spend \$25,000 per pupil each year, as many elite private schools do, or spend in the upper teens per pupil as many wealthy suburban schools do. But, are no art classes or science labs acceptable? Is no consideration of the whole child and her or his relationship to actual communities okay? We might indeed prioritize reading, writing and math, but all of those skills can and should be integrated into richer opportunities - and assessments and accountability need to take those broader needs into account.

To settle for less is not only to settle for obvious inequality, it is to consign the children of the poor to perpetually less--to not give them the educational opportunities they need to succeed in higher education, at work and as effective citizens. Settling for less means leaving them behind while pretending to enable them to catch up. Dr. King's epigram that opens this article clearly does not support the emphasis on rote learning of "basics" or drills for filling in the bubbles on multiple-choice tests.

The answers to these questions will depend on activism by parents, students, educators, communities, and organizations. Without a concerted push for change, our nation is all too likely to continue undermining education for our most vulnerable youth.

-Monty Neill is Interim Executive Director of the National Center for Fair & Open Testing (FairTest) in Boston, MA. <http://www.fairtest.org>

Resources

FairTest's website, <http://www.fairtest.org>, contains fact sheets, articles, the FairText Examiner newsletter, bibliographies on testing and positive assessment, materials for reform activists. It includes information and links on graduation and grade retention, the impact of testing on curriculum and instruction, and on authentic assessment and accountability, including the Learning Record and other materials referenced in this article. It also hosts a list of 830 colleges that do not require test scores from some or all of their applicants.

The Forum on Educational Accountability's website is <http://www.edaccountability.org>.

What Kids Can Do has a wealth of student voices, in texts and on video, including materials exposing the consequences of high stakes testing. <http://www.whatkidscando.org>. See also <http://www.teenempowerment.org>.

The New York Performance Standards Consortium has a waiver from most New York high school exams, and use locally developed performance assessments instead. See more at <http://performanceassessment.org/>.

Good videos include Ondine Rarey's Testing Mrs. Grube and Louis Kruger's Children Left Behind.