

Practical Assessment, Research & Evaluation

A peer-reviewed electronic journal.

Copyright is retained by the first or sole author, who grants right of first publication to the *Practical Assessment, Research & Evaluation*. Permission is granted to distribute this article for nonprofit, educational purposes if it is copied in its entirety and the journal is credited.

Volume 11 Number 1, January 2006

ISSN 1531-7714

Reconsidering The Minimum Competency Test Strategy In No Child Left Behind: An Agenda For Reform

Danielle Beth Kuper
Kuper Consulting Group

The minimum competency test (MCT) strategy used in the *No Child Left Behind Act of 2001* (NCLB) requires schools, in return for federal funding, to bring their students up to the level of proficiency (the minimum standard) in mathematics and reading/language arts by school year 2013-2014. This strategy involves both students (who have to reach the minimum level of achievement) and schools (who suffer the high stakes if students don't achieve the minimum). Reconsidering NCLB's MCT strategy suggests comparing it with other MCTs on the basic issues all MCTs must address: what achievement, what performance standards, and what consequences. Doing so suggests possible reforms to improve NCLB.

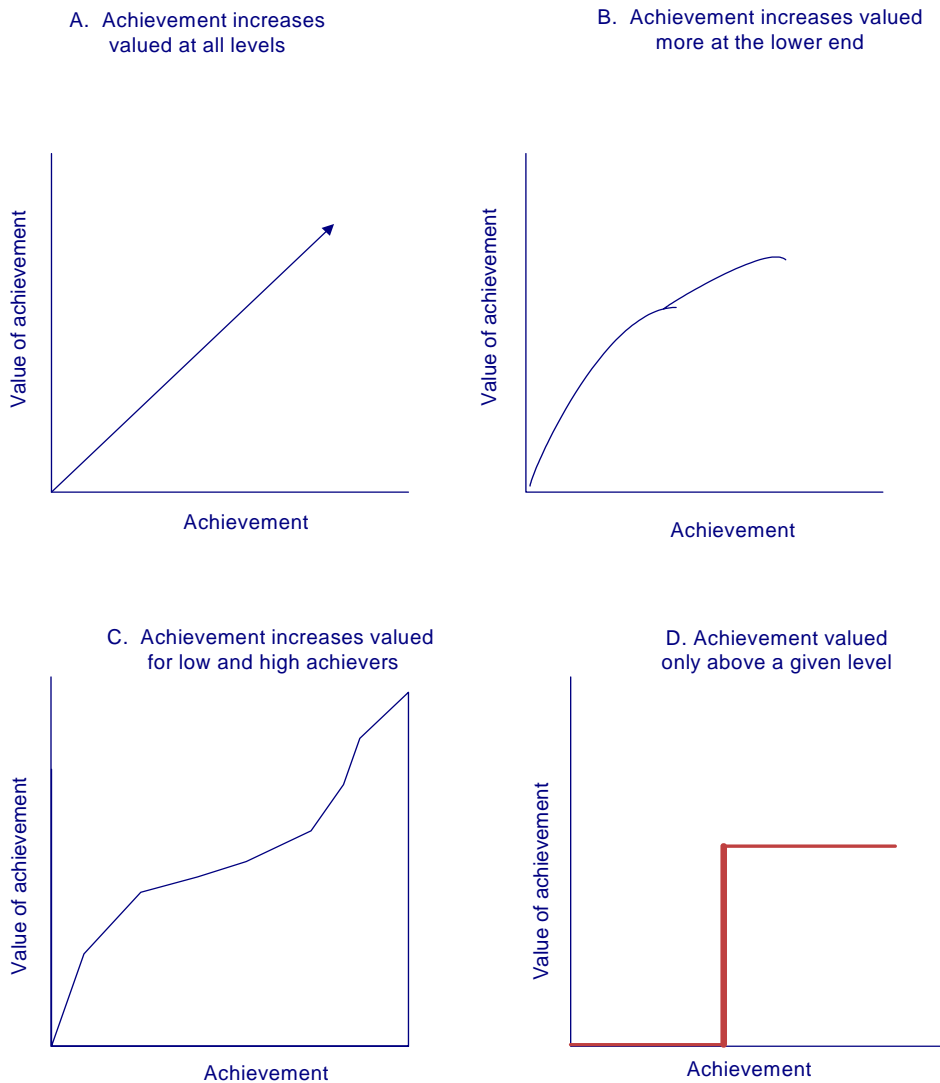
The *No Child Left Behind Act of 2001* (NCLB) establishes a new system of minimum competency tests (MCTs) as a means to increase the academic achievement of students.

Although many aspects of NCLB have been analyzed, its strategy as a system of MCTs has remained, surprisingly, relatively unexplored. Not explicitly analyzing how NCLB operates as a MCT strategy limits understanding of the broad effects of this innovative legislation. It also hinders thinking about how NCLB could be modified to make it more effective and constructive in improving American education. Evidence exists that minimum competency exams can increase achievement, and the lessons learned in earlier versions can be used in reforming NCLB (Bishop, 2005; Center on Education Policy, 2004 and 2005; Dee, 2003).

Why is NCLB an MCT Strategy?

The defining characteristic of a minimum competency test is the way it values academic achievement. Consider four alternative ways that student achievement can be valued (Klitgaard, 1974). The first way that achievement can be valued (see panel A in Chart 1) is essentially the higher the better. It doesn't matter whether achievement increases at the lower or upper end of the scale: higher achievement anywhere is better. If achievement is valued this way for every child, each student's achievement can be summed up with that of others (assuming an appropriate measurement scale), divided by the number of students in the grade (or school), and an average calculated. This has been a common practice in American education where mean scores for grades (or schools or cohorts) have been reported for virtually every test,

Chart 1. Alternative Ways to Value Academic Achievement



ranging from commercial tests and the ACT and SAT to the National Assessment of Educational Progress (NAEP).

One limitation to this way of valuing achievement is that if demographically-defined groups of students cluster at different segments of the achievement scale, just looking at an overall mean for a grade or school would not indicate that.

It would not, for example, indicate that economically disadvantaged, black, Hispanic, English-language learners, or disabled students were the ones disproportionately lower down on the scale. Not knowing that, students in those groups may not get additional services that could help them improve their performance.

More generally, valuing all achievement equally overlooks strong arguments for believing that

achievement increases at some levels of the scale are more important than at other levels. Achievement gains among low achievers (in any demographic group), for example, are often seen as more important than increases at other levels because students realizing those gains are more likely to lead productive lives, enriching themselves personally and being more productive citizens (including being less likely to engage in crime and other anti-social behavior). It is this kind of thinking that lay behind the creation of Title I in the *Elementary and Secondary Education Act of 1965*. Panel B in Chart 1 shows the way achievement is assessed when gains at the lower end of the scale are valued more than those on other parts of the scale.

While many agree that achievement gains by students at the lower end of the scale should be valued more than those at perhaps some other parts of the scale, some people argue that achievement gains among our very best students should also be valued more highly. They argue that progress is a product of intellectual elites. Elites are the people who come up with the innovations and scientific discoveries, lead large institutions and businesses, and ensure that our economy remains dynamic and growing. In an age of intense international economic competition, increasing achievement among our very highest performers is seen as the way to keep our country strong globally. People who see elite achievement as important celebrate merit scholarships, science fair winners, and valedictorians. Their perception of the value of achievement at the upper end, complementing the additional value given to achievement gains at the lower end, is seen in panel C in Chart 1. As this panel suggests, student achievement gains in the middle are seen as relatively less important, which can cause some parents to feel that their “average” kids do not get the attention they deserve.

Finally, consider panel D in Chart 1. Here achievement is valued when it reaches a certain level, a minimum level if you will. The belief is that every student needs to achieve at a certain level for their achievement to count. A common example of this way of valuing achievement is the high school exit examination. To graduate, a student must achieve a minimum score (or above) on a test. Scoring below it means no graduation (in practice,

of course, other avenues to graduation are often available); scoring above the cut score earns no additional value.

NCLB values achievement this way. It sets as the minimum valued achievement what it calls proficiency, and it requires states to set that minimum level on mandated tests (in mathematics and reading/language arts in grades 3-8 and in one grade between 10-12) in return for federal funding. Although the minimum level in NCLB is called “proficiency”—as opposed to the “minimum” (or basic level of) achievement—the way proficiency is actually valued to secure federal funding is as a minimum.¹

How Does NCLB Differ from Earlier MCT Systems?

One of the reasons NCLB has not been widely analyzed as a minimum competency system is that it differs from earlier systems in several key ways. First the NCLB strategy uses minimum competency standards on two levels, student and school (actually, district and state could be counted as different levels also). At the student level, by 2005-2006 tests are to be given to all students every year in mathematics and reading/language arts in grades 3-8 and in one grade between 10-12 (additional science tests will also soon be required). The proportions of students meeting the minimum competency standard of proficient set by the state are then determined. Minimum competency proportions (MCPs)—that is, the percent who score at the proficient level or above—are calculated over all students and for up to 9 major subgroups (American Indians, Asian/Pacific Islanders, blacks, Hispanics, multi-racial, white, disabled students, English-language learners, and economically disadvantaged), or up to 10 groups in all. If a school (or district) spanning K-12 grades has every type of major subgroup, NCLB requires that MCPs potentially be calculated for all 10 groups of students on two subjects for seven grades (3-8 and once in 10-12), or a total of 140 proportions (see Title I, *No Child Left Behind Act of 2001*).

Those student MCPs are then converted into what is called a measure of a school’s Adequate Yearly Progress (AYP). MCPs are compared to the

annual AYP targets set by a state (and approved by the federal government) called annual measurable objectives. These AYP targets represent the increases over time in the proportions of proficient students that, if realized, will result in 100 percent of students reaching the (minimum) level of proficiency set by the state by school year 2013-2014. This is the ultimate goal of NCLB.

For each school, AYP is calculated along the lines of what might be called a truth table. A truth table determines the truth of a proposition by breaking it down into a set of questions that can be answered yes or no. If all questions are answered yes, the proposition is true; otherwise it is false. In AYP, the proposition concerns whether the school is making adequate yearly progress in ensuring that all of its students collectively and as subgroups will reach proficiency by 2014. The questions asked are essentially did students as a whole and in each subgroup meet the AYP targets that would keep them on the path to universal proficiency by 2014? In other words, did each group's achievement fit the pattern of graph D in Chart 1? If any group fails to meet the required target, the school fails the test (unless the "safe harbor" provision applies, as discussed below) and is graded as not making its AYP target. All schools are held to the same AYP targets for proportion proficient (or above) whether or not achievement gains are needed from the prior year's proportions to meet the target. Schools with many high achieving students may not have to increase the proportions of students reaching proficiency to make AYP targets for a number of years, while schools with many low achieving students will need to increase the proportions of students reaching proficiency every year to make AYP targets.

When a Title I school does not make AYP for two or more consecutive years it gets classified as in need of improvement. Then the high stakes nature of that classification looms. If a Title I school goes two consecutive years not making AYP, it has to offer students the option of transferring to another public school in the district and develop a plan for improvement. If the school does not make AYP for three consecutive years, it must continue to offer public school choice to all students plus offer low-income students supplemental educational

services from a state-approved provider. If AYP is not met for four consecutive years, the district must implement corrective actions to improve the school, including replacing certain teachers or other staff, or fully implementing a new curriculum, while continuing to offer choice and supplemental services. After five years of not making AYP, the school district must restructure the Title I school, replacing all or most of the school teachers and staff, or turning over the operation of the school either to the state or a private company (NCLB; U.S. Department of Education, 2003).

The important point to realize here is that AYP is also an MCT system for schools (and ultimately districts and states). A school's accomplishments in raising achievement among all students (again, see panel A in Chart 1), or among economically disadvantaged students (see panel B), or among gifted students (see panel C), count for nothing if each and every of the specified subgroups do not pass their minimum competency standards as depicted in panel D.

Thus NCLB has a two-tiered MCT system. At the student level, students are assessed annually in terms of their meeting the minimum competency level (the state-defined proficiency level) or not. At the school level, schools are assessed annually in terms of their progress (as determined by AYP targets) in getting increasing proportions of students collectively and in subgroups to meet the minimum competency level that will ensure universal proficiency among students by 2014 (in some cases the proportionate increases required do not increase every year, but must every three years).

Here it is worth pointing out some of the similarities and differences between earlier approaches to MCTs and the MCT system used by NCLB. A core purpose of MCT systems is to establish standards for accountability and quality (Jaeger and Tittle, 1980). But earlier types of MCTs--for example, the high school exit exam now found in 19 states (CPE, 2005)--held students accountable for their own learning: they suffered the high stakes of not graduating or being promoted to the next grade if they did not pass the test. When MCTs were first established, the unfairness of putting the burden of proof exclusively on

students for the quality of their education and level of achievement generated widespread opposition. Critics said that students were being stigmatized for a failure of the schools and court cases were filed, leading to the decisions reached in a pivotal court case, *Debra P. vs. Turlington* (1981; 1983). The result was to shift substantial burden onto the schools to ensure that they had provided students with adequate opportunity to learn (curricular validity) and achieve at the required minimum level (Linn, 1983).

The striking change instituted in NCLB is that students do not necessarily face high stakes for not meeting the required minimum competency level on annual assessments: schools--that is to say, teachers and principals--do. Since everyone recognizes that both teachers and students bear some responsibility for student learning, many people perceive NCLB as unfairly allocating responsibility for achievement gains solely to teachers--and not assigning any responsibility or high stakes to students.

This use of testing may violate validity standards for assessment (see especially chapter 15 in *Standards for Educational and Psychological Testing* issued by the American Educational Research Association, American Psychological Association, and the National Council on Measurement in Education, 1999). DeMars (2000), for example, found in researching changes in the potential consequences of Michigan's high school proficiency test that students performed better on the high-stakes administration than on the low-stakes one. This finding implies that the validity of the interpretation of test results could be jeopardized depending upon the stakes of the test. NCLB allows low-stakes tests for students to be used for high-stakes consequences for schools, which may not be a valid use of those tests.

One similarity with earlier MCTs, however, is that the party assigned responsibility by NCLB for not meeting the minimum level of competency--Title I schools not meeting AYP--is stigmatized. Being classified as in need of improvement is a mark of failure to meet the requirements of NCLB, whatever their reasonableness, and this failure must be publicized in the form of report cards and, under certain conditions, in notification to parents that

their children can choose to go to another school or receive additional educational services (U.S. Department of Education, 2003).

One difference in the way stigmatizing may be working under NCLB, however, is that the perception may be growing that a Title I school could be classified as in need of improvement only because of one subgroup. This can cause special attention being paid to that group, and in some cases resentment toward the group could increase. This would be a worrisome development if public support for educating all groups should erode because of perceived injustice in classifying schools as in need of improvement.

How Can NCLB Be Improved?

Reconsidering NCLB as a MCT system involving both students and schools suggests looking at the way it currently resolves basic issues all MCTs must address so that possible modifications can be examined. The first issue concerns what achievement is valued. NCLB values achievement only in terms of state-defined proficiency in mathematics and reading/language arts (and, soon, science) as determined through AYP targets of the proportion of students above proficiency. This approach applies the kind of standard depicted in panel D to all students collectively and each and every subgroup. (The one exception is the "safe harbor" provision which allows a subgroup to be considered as meeting its AYP if its proportion not making proficiency is reduced 10 percent from the previous year and it makes progress on the other required academic indicator [e.g., graduation at the high school level or, if chosen, perhaps attendance at the elementary level].) Substantial evidence exists, however, that the family and social backgrounds of students have large and lasting effects on what children achieve academically (Rothstein, 2004). Schools having a disproportionately high number of children with backgrounds putting them at risk of not achieving proficiency are accordingly more likely to fail to meet AYP targets even though they raise student achievement substantially.

The rigidity of AYP in measuring a school's progress in improving the achievement of its

students is obvious: it awards no value to any achievement gains by any group in any subject above the proficiency minimums in mathematics and reading/language arts. AYP thus creates an incentive to narrow the curriculum, or at least increase the amount of time spent on mathematics and reading/language arts, and to focus attention on the students near the minimum who are believed able to reach the cut score with additional help (Reback, 2005).

NCLB is a radical change from previous federal legislation in large part because of AYP (Popham, 2005). While states remain responsible for establishing content standards, curricula, and performance level standards, NCLB specifies the subjects that Title I schools, districts, and states must value up to the level of proficiency. The 1994 amendments to Title I required states to create performance-based accountability systems for schools, but many states had not been able to put systems meeting federal standards in place by the target date of 2000. Perhaps in part due to frustration that the earlier strategy did not appear to be working in helping disadvantaged students, NCLB created requirements that Title I schools not just try but actually increase achievement to the level of proficiency--and not only among disadvantaged students but among all students and specific additional subgroups as well (many of which had other federal programs directed at them). While alignment of state and federal standards of accountability was expected to occur in 2000, it was not widely anticipated as involving how the achievement of all students and the various subgroups must be valued.

One possible reform to consider for NCLB is to change how AYP is defined so that the achievement of students collectively and as subgroups can be valued in different ways (as depicted in Chart 1). One option is to amend the way AYP is calculated. This could maintain federal control over how achievement is valued in what subjects. Linn (2005a and 2005b), for example, has mentioned several ways that NCLB could be amended, including calculating progress as achievement growth and expanding the safe harbor provision.²

An alternative to modifying but keeping a federal standard is to let states propose their own ways of calculating academic progress, an option that perhaps reflects an earlier tradition of federalism in American education policy. One possibility would have states specify which subjects they want to consider in calculating progress and the ways in which achievement would be valued (as indicated in Chart 1, this might include achievement above the level of proficiency or gains in achievement even below that level). States might also want to consider what kinds of assessments to use. Some evidence, for example, suggests that MCTs in the form of end-of-course exams are especially effective in raising achievement (Bishop, 2005). In any case, NCLB could be modified to allow a wider appreciation of schools' efforts and successes in improving the achievements of their students whether the federal government or the states control how those achievements are valued and measured for purposes of federal funding.

A second basic issue every MCT must address is where to set the cut score, or minimum, on its achievement scale. Setting sensible performance level standards has been a challenge for states on other MCTs (Center on Education Policy, 2004 and 2005). Standard setting is always subjective and judgmental, regardless of the sophistication with which it is done and the reasonableness (according to validity and other standards) of its results. Many states have needed a number of attempts to get a high school graduation standard that improves educational achievement and has public support.

The cut scores for NCLB are currently widely seen as unworkable (Linn, 2005a and 2005b). Although the states determined those performance level standards, NCLB specified that three levels of proficiency--basic, proficient, and advanced--had to be set. That approach probably contributed to the relatively high--and hence unworkable--standards that resulted. Even though state standards of proficiency are almost never as high as the national standard of proficiency set for the National Assessment of Educational Progress (McLaughlin and Bandeira de Mello, 2002), they are still too high for all students to reach by 2014.

Several strategies are available to amend performance standards. One approach would establish a national performance level standard to replace state standards. Linn (2005a), for example, has suggested adopting the basic standard set for NAEP. Panel A in chart 2 depicts the basic, proficient, and advanced levels of proficiency on the distribution of grade 4 reading achievement in NAEP in 2003. About 38 percent of 4th graders were below basic in reading; about 70 percent were below proficient; and about 93 percent below advanced. If all those below basic were brought up to that standard by 2014, the distribution of achievement might look like the one found in panel B. Shifting the distribution of achievement among students to that extent would be unprecedented. Even setting a national standard of basic on NAEP is probably not workable.

But the performance standards on another national assessment could be considered, namely on the Armed Forces Qualification Test (AFQT) that is part of the ASVAB, or Armed Services Vocational Aptitude Battery. AFQT may be the most widely used test battery in the world, and its tests assess word knowledge, paragraph comprehension, arithmetic reasoning, and mathematics knowledge. This assessment was developed by the Department of Defense in the 1960s to help select personnel for the military and is widely credited with helping to build the most effective military in history. Although it is considered a test of aptitude (though not of IQ), aptitude at the high school level reflects achievement (Cascio and Lewis, 2005). The standards used to determine fitness to serve in the military may also be appropriate targets for all high school graduates. They certainly provide an understandable goal for many students. Using AFQT standards as a base, standards scaled proportionately could be developed for elementary and middle school grades.

A second approach in rethinking performance standards is to have the states align them with other extant state achievement standards. One area to consider is the high school exit exams many states already use. In fact, among the 19 states with high school exit exams (used to withhold diplomas based on test results), 15 are using them to meet NCLB

requirements and 10 are using the same cut-score for high school graduation and NCLB (Center for Education Policy, 2005). Starting from the level of proficiency they already expect of their high school graduates, states could scale those expectations down to an appropriate level for the assessments given in grades 3-8. While some of those tests are explicitly aligned with state content standards and curricula, others rely on implicit standards. Nevertheless, this is a path some states have already taken and others could consider.

A second direction states could take is to revisit the performance level standards they have set for passing the General Educational Development test (GED), which is sponsored by the American Council on Education. This assessment has an illustrious history of use since its development for returning GIs after World War II to certify that persons without a high school diploma have the ability to benefit from a college education. The areas tested over about eight hours include: mathematics, language arts (reading and writing), science, and social studies--a broad swath of the curriculum. Although the GED is perceived by many to be relatively undemanding since high school dropouts gain credentials through it, over 1000 institutions of higher education accept the GED as evidence that a student is eligible to be considered for admission. It also indicates a level of cognitive achievement comparable to high school graduates who do not go on to college (Heckman and Rubinstein, 2001).

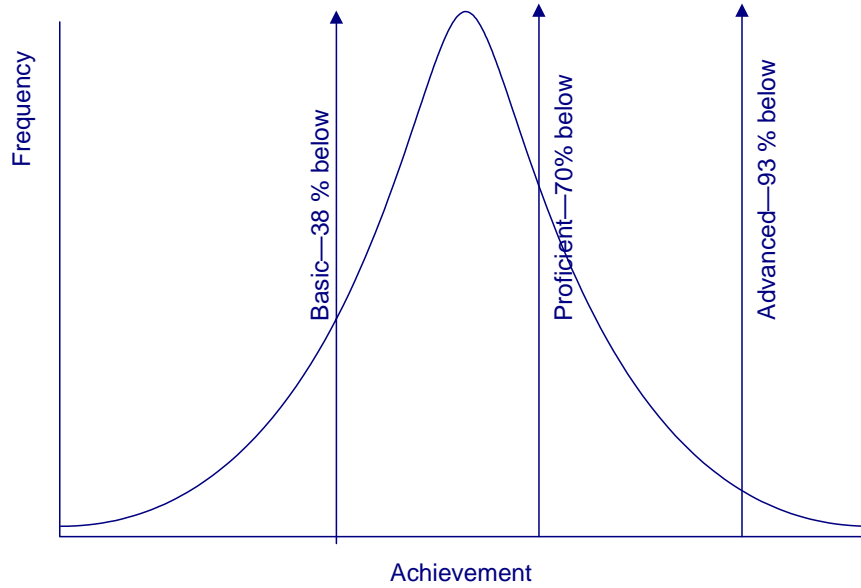
The issue of whether there should be national or state performance standards is a contentious one. But regardless of which is chosen, standards of proficiency for specific purposes--high school graduation, enlistment in the military, and enrollment in institutions of higher education--already exist and could guide federal and state policymakers in reassessing where to set appropriate levels of proficiency for NCLB that are workable.

A third issue that MCTs must address is who is to be held responsible for failing to meet the performance standards. As mentioned, the first MCTs put the burden of failure on students. But the need to have schools accept responsibility for preparing the student, ensuring exposure to tested

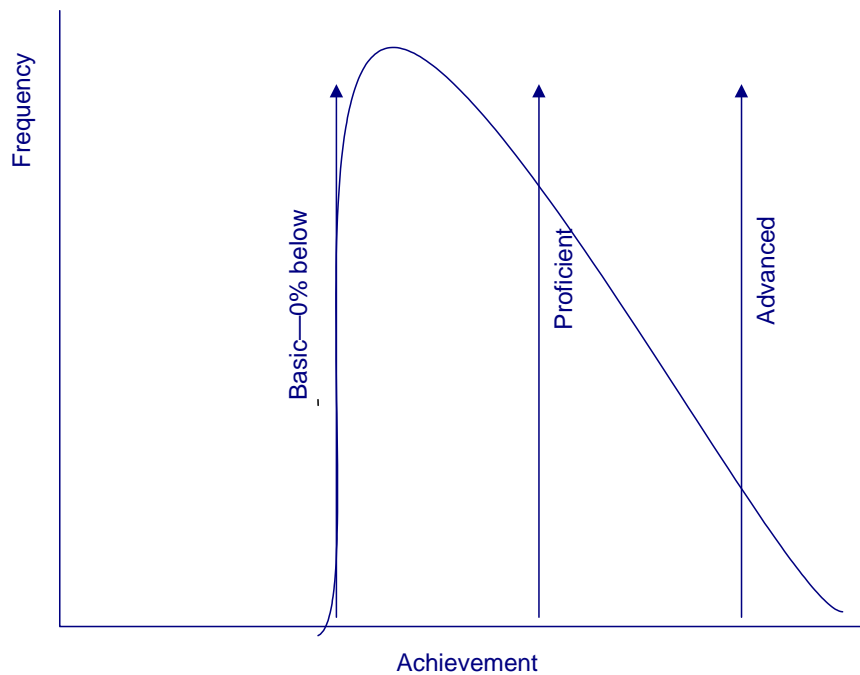
materials, and so on, was established in the Debra P. case (Linn, 1983). Under NCLB's MCT strategy,

however, students generally have no mandated

Chart 2. A. Estimated Distribution of Student Achievement on NAEP Grade 4 2003 Reading Test by Basic, Proficient and Advanced Levels



B. Hypothetical Distribution of Student Achievement on NAEP Grade 4 Reading Test If Minimum Achievement Was At or Above Basic Level



consequences: only schools face high stakes from NCLB, which is perceived by many as unfair and appears to violate standards of assessment validity, as noted above.

One possibility to consider is adding high stakes for students. States have already started doing this, as noted above, in using high school exit exams as part of their NCLB accountability systems. High stakes could also be extended to other grades in the form of grade promotion or mandatory summer school or Saturday attendance (or some other requirement), and some jurisdictions have done so (e.g. Chicago). Doing so ensures that the joint responsibility of students and teachers for learning is recognized.

A final issue that MCTs must address is the consequences of passing or not passing. In most MCTs, the consequences are negative: a student doesn't graduate from high school, get a GED, or become eligible to enlist in the military. NCLB also specifies a range of negative consequences for schools, from having to develop a plan for improvement, to offering choice, providing supplemental services, and to state or private control of the school. NCLB does, however, allow Title I funds to be awarded to schools that close achievement gaps or exceed academic achievement goals and to teachers that receive academic achievement rewards (U.S. Department of Education, 2003). But unlike other jurisdictions that require such rewards (for example, nine states made awards to school districts on the basis of performance—absolute, improved, or both—in 2002 [ECS, 2002]), NCLB only creates the possibility of reward. If students should also come to have some responsibility for their achievement under NCLB, positive consequences for them could also be considered.

Can NCLB Work?

Although NCLB may be unworkable in its current form, it has generated much positive energy concerning how to improve education, not least because it has insisted that achievement test scores be made available for many different groups of

students and then used in evaluating student and school performance. But two more general issues need to be considered.

The first concerns why accountability systems might be expected to improve achievement. The production of learning has long been carried out essentially the same way: through a dialogue between teacher and student. Possible technological improvements, from television to the computer, have born little fruit, and other possible technological innovations for improving learning seem remote. So what is the theory behind the perceived effectiveness of accountability? Two elements seem central: motivation and time on task (Jacob, 2003).

Holding students and teachers accountable could increase the motivation of teachers and students—and supportive parents—to achieve more. It could also lead to a shift in the amount of time devoted to the subject areas assessed: time on task does improve achievement. But those types of shifts may well result in a one-time gain. They do not seem promising as a basis for making continuous improvement, which is of special concern when setting ambitious long-term goals for academic growth. Experience with different types of accountability systems—in terms of coverage, types of exam, consequences, etc.—will be needed to determine what is possible in continuing to improve those systems.

In this context it will be interesting to observe how performance on NAEP unfolds. The 2005 results from NAEP assessments suggest significant (but small) progress in mathematics, but little in reading. Some have suggested that NCLB was the cause of the gains, while others have argued that earlier state efforts at establishing accountability systems may be responsible. The larger question is whether these recent NAEP results are “real gains” and will continue or reflect “inflation” perhaps resulting from a shift in time devoted to, in particular, mathematics at the expense of other subjects because of accountability requirements. Because NAEP does not cover the entire curriculum in a systematic way, it may not be possible to provide an answer to this question

definitively, but observing future changes in its achievement scores may suggest one.

The second general concern is what expectations are appropriate for improvement in achievement. As Cohen and Haney (1980) noted 25 years ago, there is no question that schools educate. Algebra, Spanish, proficiency in writing essays--those are not things that people learn without help. But "there is a very weak understanding why some students, teachers, classrooms, or schools are more productive than others in the same subjects. Thus far, all the evidence suggests that there are important differences in effectiveness, but that they have no uniform causes." Those uniform causes remain to be determined.

This suggests that perhaps the principal shortcoming of NCLB is that it embodies an engineering mentality (Nelson, 2001). In creating an automobile engine, for example, managers are able to specify, say, the horsepower wanted, and then turn the task of designing an engine meeting that specification over to engineers. The engineers are able to specify the characteristics of the parts they want for the engine, and then machinists can produce them (to a reasonable tolerance level). NCLB takes a similar approach. Policymakers decided that they wanted students who are proficient (the horsepower). The task was then turned over to psychometricians (engineers) to design the measurements and to teachers (machinists) to produce students who meet them. But as has often been said, schools are not factories--not for lack of trying but because students are not inert metals, they are human beings who have differing qualities and motivations.³ Teachers can help students realize their potential, but they are not able to produce student achievement to any given specification.⁴

Using an MCT strategy, NCLB could be a turning point in improving the quality of American education—but it needs appropriate changes to realize its potential. And those changes need to be more attuned to what is possible within the current capacity of our schools (Elmore, 2002).

References

- American Educational Research Association, American Psychological Association, and National Council on Measurement in Education (1999). *Standards for Educational and Psychological Testing*. Washington, DC.
- Bishop, J. H. (2005). High School Exit Examinations: When Do Learning Effects Generalize? In J. Herman & E. Haertel (Eds.), *Uses and Misuses of Data in Accountability Testing*. National Society for the Study of Education Yearbook, Part I. Chicago: University of Chicago Press.
- Cascio, E.U. & E.G. Lewis (2005). *Schooling and the AFQT: Evidence from School Entry Laws*. Working Paper 11113. Cambridge, MA: National Bureau of Economic Research.
- Center on Education Policy. (2004). *State High School Exit Exams: A Maturing Reform*. Washington, DC.
- Center on Education Policy. (2005). *State High School Exit Exams: States Try Harder But Gaps Persist*. Washington, DC.
- Cohen, D. & Haney, W. (1980). Minimums, Competency Testing, and Social Policy. In R. M. Jaeger & C.K. Tittle (Eds.), *Minimum Competency Achievement Testing*. Berkeley, CA: McCutchan Publishing Corporation.
- Debra P. v Turlington*, 474 F. Supp. 244 (M.D. Fla.1979); 644 F. 2d 397 (5th Cir. 1981); 564 F. Supp. 177 (M.D. Fla 1983).
- Dee, T. S. (2003, Spring). Learning to Earn. *Education Next*, 3, 64-75.
- DeMars, C.E. (2000). Test Stakes and Item Format Interactions. *Applied Measurement in Education*, 13, 55-78.
- Education Commission of the States (2002). *State Notes: Accountability—Rewards and Sanctions for School Districts and Schools*. Denver, CO. Retrieved from: <http://www.ecs.org/clearinghouse/18/24/1824.htm>.
- Elmore, R. (2002, Spring). Unwarranted Intrusion. *Education Next*, 2, 31-35.
- Goldstein, J. & P. Behuniak (2005). Growth Models in Action: Selected Case Studies. *Practical Assessment, Research & Evaluation*, 10 (11). Retrieved from: <http://pareonline.net/pdf/v10n11.pdf>
- Heckman, J.J. & Y. Rubinstein (2001). The Importance of Noncognitive Skills: Lessons from the GED Testing Program. *American Economic Review*, 91, 145-149.
- Jacob, B. (2003). Getting Inside Accountability: Lessons from Chicago. In W. G. Gale & J. R. Pack (Eds.),

- Brookings-Wharton Papers on Urban Affairs*. Washington DC: Brookings Institution Press.
- Jaeger, R. M. & C. K. Tittle (Eds.) (1980). *Minimum Competency Achievement Testing*. Berkeley, CA: McCutchan Publishing Corporation.
- Klitgaard, R. E. (1974). *Achievement Scores and Educational Objectives*. Santa Monica, CA: RAND Corporation.
- Linn, R.L. (1983). Curricular Validity: Convincing the Court It Was Taught Without Precluding the Possibility of Measuring It. In G. Madaus (Ed.), *The Courts, Validity, and Minimum Competency Testing*. Boston: Kluwer-Nijhoff Publishing.
- Linn, R. L. (2005a) *Test-based Educational Accountability in the Era of No Child Left Behind*. SE Report 651. CRESST/University of Colorado, Boulder.
- Linn, R. L. (2005b) *Fixing the NCLB Accountability System*. CRESST Policy Brief CRESST/University of Colorado, Boulder.
- McLaughlin, D. and Bandeira de Mello, V. (2002, April) *Comparison of State Elementary School Mathematics Achievement Standards Using NAEP 2000*. New Orleans: American Education Research Association Meetings.
- No Child Left Behind Act of 2001* (2002), Pub. L. No. 107-110.
- Nelson, R. R. (2001). *On the Uneven Evolution of Human Know-how*. ISERP Working Paper 01-05. New York: Columbia University.
- Popham, W. J. (2005). *America's "Failing" Schools*. New York and London: Routledge, Taylor and Francis Group.
- Reback, R. (2005). *Teaching to the Rating: School Accountability and the Distribution of Student Assessment*. Working Paper (July draft). New York: Barnard College, Columbia University
- Rothstein, R. (2004). *Class and Schools: Using Social, Economic, and Educational Reform to Close the Black-White Achievement Gap*. Washington, DC: Economic Policy Institute.
- U.S. Department of Education (2003). *No Child Left Behind: A Parents Guide*. Washington: D.C.

Notes

- ¹ Some reports have categorized high school exit tests as either “minimum competency” or “standards-based” tests, but both are technically minimum competency tests (e.g., see Center on Education Policy, 2005). The difference between the two is in how and where the minimum required score is set.
- ² The Department of Education recently announced that as many as 10 states could meet their AYP requirements by using growth in achievement as opposed to meeting the proficiency proportions in their current annual measurement objective. If the goal of NCLB remains the same—all students achieve proficiency by 2014, a minimum competency standard—the challenges of determining adequate annual achievement gains are not less than those of setting annual levels of achievement. Technical problems, for example, increase because the reliability of gain scores is less than that of a level score. Missing data becomes more of a problem because two level scores are needed to calculate one gain (which means more missing data for students who move, often those who need special attention). States choosing to measure gains would also have to worry about currently high-scoring students whose achievement growth has not been an issue under current AYP standards since growth is often found negatively correlated with achievement level. That said, there is much to be said in favor of moving to a minimum growth assessment (MGA) system from a minimum competency test (MCT) system—using achievement gains instead of proficiency levels—because the validity and fairness of results are likely to be greater. For a survey of current growth systems in 9 states and 2 cities, see Goldstein and Behuniak (2005).
- ³ The more commonly used engineering analogy rests on U.S. achievements in rocket technology. It is often said—as heard by the author from both proponents and opponents of NCLB—if the U.S. can send a man to the moon, why can't we educate every child to be proficient?

- ⁴ Former Secretary of Education Rod Paige once asked, why can't every third grader read at a third-grade level? This question reflects a common lack of understanding of how grade-level norms are set. He might as well have asked "Why isn't every third grader at least 52 inches tall?"

Acknowledgement

The author thanks Leon Nora and the anonymous reviewers for comments on an earlier draft. This paper does not reflect the opinions or positions of any group, organization, or agency the author has ever worked with.

Citation

Kuper, Danielle Beth (2006). Reconsidering the minimum competency test strategy in no child left behind: an agenda for reform. *Practical Assessment Research & Evaluation*, 11(1). Available online: <http://pareonline.net/getvn.asp?v=11&n=1>

Author

Danielle Beth Kuper
Kuper Consulting Group
Minneapolis, MN
danykuper@yahoo.com