

Assessment and Educational Equity

*Based on interviews with
Dennie Palmer Wolf*

In many schools, right up through graduate school or law school, standardized achievement tests often create situations in which equity exists only in minimal ways....Students from poorer economic groups are often in tracks that fail to give them the information they need to prepare for college or to do well on standardized college aptitude tests. Until we change this underlying inequality in education, tests will reflect it.

Weaknesses In Current Definitions of Educational Equity

I am deeply concerned about the effect of testing on equity in education. In many schools, right up through graduate school or law school, standardized achievement tests often create situations in which equity exists only in minimal ways. For instance, achievement tests often result in the segregation of populations *within* a school instead of the segregation which formerly existed *among* schools.

Many districts use externally developed tests designed to assess students' basic skills; for example, Boston uses the Metropolitan Standardized Test. Performance on such tests, and educational tracks based on test performance, are very often correlated to first language and to socioeconomic status, which are, in turn, correlated to ethnicity because of uneven distribution of social and economic opportunity. Nevertheless, these tests are often the basis for entry into special programs, tracking and retention. The result is many schools that are only technically unsegregated; in reality there is segregation within individual schools rather than full educational equity.

Historically, educational equity has been defined in three separate but inadequate ways: as input, "easily countable equity data," and "poorly-framed achievement data."

Unfortunately, the technical legacy of cases like *Brown v. Board of Education* has requir-

ed equity that is defined in terms of input. For example, all schools, just as they must be racially balanced, are often judged or equated on teacher/pupil ratio, number of school days, or equal physical facilities. Yet that doesn't create equity in terms of educational opportunity. Even if it is technically true that the overall school population is racially balanced, within a given school students are tracked in "general education," "advanced," or "college bound" programs, which provide vastly different educational opportunities. Equitable educational opportunity and equity of *output* is lacking, especially since poor and minority children are disproportionately in educationally inferior tracks. Equity of output would require greater efforts to provide equalized opportunities to all children. Standardized tests make little contribution to fulfilling this requirement.

"Easily countable equity data" means making sure the population of a program--e.g., a gifted and talented program--is racially balanced *at inception*. But this is ineffective in producing true

This article was drafted by LTV on the basis of several telephone interviews with, and was reviewed by, **Dennie Palmer Wolf**. Dr. Wolf is the director of PACE, Performance Assessment Collaborative for Education, at the Harvard Graduate School of Education. Funded by the Rockefeller Foundation, PACE does independent research about the consequences of new forms of assessment, particularly in urban settings, and seeks out new forms of assessment fair to a diverse population. Dr. Wolf is currently also working with the College Board (CEEB) on more diversified approaches to assessment.

equity because programs that begin balanced often end up imbalanced. Rather than focusing only on the inception, we should instead be paying attention to who stays in a program, who graduates from school, who is ready upon graduation to go on to substantial post-secondary education. Those are much harder questions because they are about quality, not just numbers. And, once again, standardized tests do not help provide equity here.

"Poorly-framed achievement data" equity relates to the data available at a single point in time, e.g., how students are doing in April of a given year. While standardized tests do give us a sense of student performance at a single moment in time, the data are not organized to tell us students' progress *over* time, even when schools give a standardized test every year. Therefore, we don't know the growth curve for different populations of students in a given district, but only the achievement of an "average" student. Moreover, when we publish data, we link it directly to ethnicity, not to level of family income, first language, the degree of social stress in students' lives or other factors that are more causal.

Additional Problems of Current Tests

Currently, testing and evaluation is a secret science. It is technically driven, with standards and procedures so internal to itself that it is often difficult for there to be public debate about it. Thus, the minimal data which is made public is difficult to question. Consumers of education do not know, when scores are said to increase or decrease, whether all segments of the student population are doing better or whether schools have excluded the test scores of certain groups such as students in Special Education or bilingual classes.

Nor do teachers do more than distribute and monitor standardized tests; they do not determine what items are on the tests, nor what items *should* be on a test because they concern important educational matters, nor how items are

determined to be right or wrong. The bulk of such determinations are made by non-teacher technicians, who historically have concentrated on creating a distribution in performance (the bell-curve) rather than criterion-referenced reports of performance. All this should be questioned.

Furthermore, test scores are treated as if *they*, not the underlying education, were the problem. Take, for example, tests sponsored by the College Board, which invented the notion of standardized college entrance examinations. Originally those examinations were intended to increase equity by means of anonymity: any-one, independent of ethnicity or class or ability to pay tuition, could earn a score that would make him or her a viable candidate for college. But whatever the original intention, these scores are now taken as indicating how able different populations are or how good certain schools or districts are. Yet we now know that test scores often correlate with the kind of high school that students attend and the track students are in. Students from poorer economic groups are often in tracks that fail to give them the information they need to prepare for college or to do well on the College Board's standardized college aptitude tests. Until we change this underlying inequality in education, tests will reflect it. And changing the scores *per se* (through practice courses or other means) is only a cosmetic answer.

Proposed New Forms Of Assessment

The underlying problems may be exacerbated rather than ameliorated by current proposals for national tests put forth by former President Bush and now being taken up by the Clinton administration. The plan is to create national achievement tests for students in the fourth grade, the eighth grade, and the tenth grade. The initiative has been designed in hopes that these achievement tests will drive up the standards of public education. But it is not a foregone conclusion that national testing will cause the level of education to rise, unless

such testing were to be tied to equalized access to educational opportunity. Moreover, such high stakes testing, which will certify students as having met or not met standards at early points in their lives, could drive large numbers of students out of school at the tenth grade or even earlier. Very undesirably, such testing, which could have crucial consequences for students, teachers and entire school systems, may be another set of tests for which people will want coaching both in school and, for the more affluent who can afford it, outside of school as well.

These new proposals *are* nonetheless different from current testing in that most versions of them seek performance based testing. Instead of consisting of dozens of multiple choice questions, newer tests could consist of two or three day tasks that require skills and numerous "types" of knowledge. For example, in a math test for fourth graders developed by one team working on a national examination system called the New Standards Project, students would be given a certain amount of hypothetical money to spend to create an aquarium in which the fish have adequate resources and can live compatibly. The students will be given an entire data base about the fish and about accessories that are available. The project would take several days; students would have to create a relevant graph, and would have to argue in favor of creating the aquarium as they did.

Though it would be different from current standardized tests because it would use performance based assessment, the very serious risk of NSP's proposal is that, like current tests, it may or may not directly and deliberately affect curricula. Consequently, at Performance Assessment Collaborative for Education, we are collaborating with partners as diverse as four urban and two rural school districts, as well as the College Board, to create assessments that will cause and be based on a strong curriculum and good teaching. In these collaborations we *are* seeking to directly and intentionally affect curricula in order to increase true educational equity.

In the College Board's PACESETTER pro-

gram, the necessary bodies of knowledge will be determined by having national committees of teachers and scholars design courses. A shared template will help all teachers to impart the same "large ideas" and "worthwhile skills," but will not require each teacher to use the same subject matter when doing so. For example, one teacher of literature might use a significant amount of African American literature, while another might use American writers of the 1920s and 1930s. But both would require all students to meet the same high levels of academic achievement--by doing independent reading, writing essays, etc. The essays along with other projects would become part of a portfolio of the student's work. The portfolio itself would be graded by the student's teacher, and then portfolios from that teacher's students would be "cross read," or "cross rated," by other teachers.

At the end of a semester or a year, each student would undertake a "culminating project," which would be done within a specified time period. In a literature course, for example, students might be given one week to investigate and examine the literature of their particular geographic region. They might interview writers and put together an anthology of regional literature, with an introduction explaining the special characteristics of that literature, why it is interesting, and what it contributes to American literature as a whole.

As long as the materials and tasks are demanding, it would not matter whether, in these projects, students are dealing with Huckleberry Finn, or with the oral and written traditions of the Spanish speaking community.

What we are suggesting for *all* students is currently done only for the best and brightest—for those in advanced placement. We believe that, if the country wants high standards for all of its students, there must be public syllabi and knowledgeable public discussions that describe the curricular features needed to attain such standards. We further believe, as said, that assessment must be based on, and not be independent of, the needed curriculum.

The suggestion being made here not only replicates what is already done for the nation's best students, but is also similar to what is required for the International Baccalaureate, which is based on a set of examinations given in international high schools. In the pertinent foreign language exam, for instance, a student must give an oral report in the foreign language, and must take and answer questions from the floor in the foreign language. The exams themselves are created by committees of teachers, who change the specific details of the examinations annually while continuing to require mastery of the same "large ideas" and "large skills." Many European nations, such as Holland, have made and continue to make large investments in creating and giving these kinds of examinations.

The Feasibility Of The Approach Suggested By The Collaborative And The College Board

The immediate question, of course, is whether the suggestions of PACE and the College Board are feasible on a national scale for all American students. Implementation would cost more—perhaps 200 or 300 percent more—than standardized multiple choice tests. However, this is not the way to think of it. The better way to think of it is as industry thinks of the steep but important expense of quality control, which many companies consider to be worth the investment. After all, we are talking about serious educational reform and the use of assessment to help achieve it.

Moreover, we may soon have little choice. In Kentucky, for example, after the state's school system was declared unconstitutional, the legislature voted substantial sums of money to create a new statewide assessment system based on portfolios and performance tests. Or again, in a pending Connecticut case called *Sheff v. O'Neill*, a 14 year old African American is claiming that, because of the inequality between

the schools of inner city Hartford and of the surrounding suburban areas, he is not getting an education which comports with the state constitutional provision guaranteeing equality of educational opportunity. Such inequalities are common throughout the country because states and politicians have been far more concerned with maintaining the decision making autonomy of local school districts than with questions of true educational equity. Interestingly, the governor of Connecticut has taken an independent stand in response to the *Sheff* case, saying that the inequalities of education should be rectified and proposing that the state should be divided into six regions, each of which is economically mixed, and schools within each region should be responsible for achieving educational equity within the region. This would be a real contribution to the system of education—to providing high quality education to children of poverty who today are often stultified in lower tracks and poorer schools.

Conclusion

The need to provide high quality education to students who are currently marginalized in schools is pressing. It is a very troublesome fact of our culture that today about one in five children grows up in poverty. What happens to these students in school has life-long consequences: students who are not placed in tracks that include serious academic subjects like mathematics, sciences and foreign languages may be hindered for the rest of their lives and may indeed never earn much above the minimum wage. The nation can no longer afford this. Events like the Los Angeles riots or the case in Hartford show that the victims may no longer suffer it in silence. But changing tests or creating national assessment systems is not, *per se*, the answer; changing fundamental, daily educational opportunity is.