

Equity and Assessment

Paul Zachos

Published in *The Science Teachers Bulletin*, Fall 2005, Volume 69, Number 1, Pages 12-16

Over the past year the leadership of The Science Teachers Association of New York State has fixed its attention on addressing issues of equity in science education. The following essay is written in support of this effort. I intend to show how a great victory for equity can be attained by using examinations purely for educational purposes, instead of for ranking students.

Grading vs. Assessment, Educational vs. Non-educational use of Examinations.

It is generally believed that the issue of equity in the field of educational testing or assessment has to do with the unfairness of certain tests, i.e. that they are biased towards particular gender, cultures, races, social classes, etc. While this kind of bias certainly exists, it is the symptom of an even more serious problem associated with our testing practices. That problem is the use of tests for purposes other than to support teaching and learning.

In this essay I would like to introduce a distinction between two commonly used terms – assessment and testing. I will also distinguish educational from non-educational uses of assessments and tests. Applying these distinctions provides a basis for understanding how the issue of equity actually relates to testing and assessment.

Exams are constructed primarily to meet two broad purposes:

1. One kind is built around learning goals and gives information on the extent to which those learning goals have been attained – this I will call assessment.
2. The second is a collection of items which are aggregated to give a grade or score (e.g. number or percent correct) – this I will call testing. I will use the term examination or exam to include both.

Regarding the use of examinations there are also two broad categories:

1. The information may be used to help plan instruction and to support the attainment of learning goals. This is an educational use of information
2. Or, by contrast, scores may be used to rank students on a continuum of performance. These grades and ranks are then used in turn to make a host of decisions concerning who passes and who fails, who gets into the honors class, who gets to play on the basketball team, etc. This is essentially a non-educational use of information in that it is not directly supporting the attainment of intended learning outcomes.

A number of illuminating realizations emerge from these distinctions.

1. It is relatively easy to make educational use of assessments (again educational and assessment in the sense defined above). When teachers have information about the degree to which students have attained targeted learning goals they can design appropriate instructional interventions for individuals and groups.
2. It is very difficult to make educational use of test scores. A number or letter representing aggregate performance on a set of items does not give information about what individuals,

or the class, have attained or failed to attain that is specific enough to be educationally useful. (One might argue that test scores can be used to provide a motivation for students to learn, and are in that way educational but this argument fails for a very simple reason. Scores on tests are not a positive incentive for all students and in fact are a neutral or negative incentive for many. To be used effectively as a motivator, scores from tests would have to be used intentionally and carefully, student by student with a clear knowledge of what effect the grade would have on each student and this is clearly not done except in rare cases.

Ideally to be educationally useful an examination must tell the teacher how the student and the class are doing on the targeted learning goals or on the pre-requisites to attaining those learning goals. This information provides a sound and reasonable basis for planning and revising instruction for individuals and groups. But note the implication here -- to be educationally useful an assessment must be administered before instruction (this is the necessary feature of formative assessment that is so rarely considered). Examinations administered after instruction has been completed no longer have the possibility of being used to support learning for the students who were examined and so can not properly be considered to serve an educational purpose for those students.

It does not take a great deal of study or thinking at this point to recognize that most of the examinations that we use in schools, whether they are commercially produced, state examinations or teacher-made exams fall into the category of testing rather than assessment and that they are used primarily for non-educational purposes.

Equity, Grading, and Assessment

Both testing and assessment can provide useful information for dealing with issues of equity. The important thing is to break down student performance results in order to compare how the various targeted groups (gender, racial, cultural etc.) are performing.

If we find that a target group of interest (say a cultural minority group) does significantly less well on a reliable test or assessment, there are several possible reasons:

1. The students may be competent in the subject matter but the nature of the test or assessment does not allow this group to truly show what it can do. For instance the cultural features of a particular group may lead them to understand the meaning of questions and tasks differently from other groups. In this case, the nature of the test or assessment itself shows bias.
2. It may be that the procedure for scoring does not recognize that certain responses given by various target groups are in fact indicators of a sound grasp of the subject matter and so the answers of students from this group are marked incorrect when they should not be. This is another true bias situation that is in the proper sphere of concern for equity.
3. There could be legitimate disagreement as to whether one answer or another may be correct based on one's cultural perspective (e.g. a Filipino will have a different understanding from an American of what happened during what we in the US call the "Spanish American War".)

4. It is also possible that one or more groups of interest are simply not performing competently regarding the subject matter at hand.

In all four cases it is important for us to find out what the examinations can tell us about students and their performance. In fact examinations are the most direct and powerful way to identify symptoms and causes of inequity. We need to be vigilant, indeed proactive, in seeking out and neutralizing or counteracting the biases which appear in the first three cases. But it is in the fourth case (where a group is performing differentially because it actually lacks competence in some knowledge, dispositions or skills) that as educators we have the greatest opportunity to make a difference. Here we can most directly apply our own capabilities as teachers to correct a deficiency. But as we have seen, to do so we need to have tests that tell us how well students are performing on our learning goals and we need to have them scored in a timely fashion so that we can take action through instruction.

In summary then, both tests and assessments can provide us with information related to the possibilities of inequities in educational practice. Tests however lead only the possibility of correcting bias associated with testing. Assessments provide an opportunity and suggestions for assisting under-performing target groups because they provide information on the level of performance of students and groups on the learning goals which make up the curriculum.

The fact is that if we did not use examinations for ranking and grading, but strictly for educational assessment purposes there would actually be no equity problem regarding tests! We would be eager to find out which students were performing inadequately and to direct our attention toward helping under-performing groups to attain the targeted learning goals. We would actually welcome the discovery of deficits in any definable group of students because it would provide us with the key information needed to address the deficits. But as things currently work, students (and now schools and school districts) are granted privileges, priorities and preferences because of their test scores, and because test scores do not provide a useful information base for educational planning and improvement, then tests become causes as well as symptom of inequities

Indeed if assessment information was used strictly for educational purposes the No Child Left Behind initiative could be put back on track to serve its stated intentions. But to do so it would have to turn away totally from making students, teachers and schools accountable through the instrument of grading and ranking and instead concentrate on purely educational assessment and setting up resource allocation policies to target differential learning deficits.

Grant Wiggins' Word Problem

At a recent conference for Middle School Mathematics, Science, and Technology Teachers held on Long Island:

(http://www.hofstra.edu/Academics/SOEAHS/TEC/tec_mstp_showMe_curriculum_showcase_and_conference.cfm)

Grant Wiggins posed a problem. He stated it ingeniously as if it were a test item -- a conventional mathematical word problem. The problem took the following form:

Two trains are speeding towards each another on a single track.

One is standards based educational reform. The other is the concern for equity given increasingly diverse and heterogeneous student bodies.

When will these trains collide and what will happen at the collision?

Well, it turns out that this is not as difficult a problem to solve as it may seem, because the collisions are already occurring everywhere and we are everywhere seeing disastrous outcomes. Perhaps if we immerse ourselves in the spirit of grading and high stakes testing we might consider that we are too late to answer this question and so have failed the test. But I would rather think that it is not too late for us to learn from this situation and to change our ways to prevent further failure and disaster. Let us consider what can be done before it becomes too late.

The deeper question is --How can standards-based reform and concern for equity be reconstrued so that they are not on a collision course? In fact how can we put them on the same train and set them moving towards the same destination? The answer is two fold: First of all standards based reform must turn away from high stakes testing based on grading. It must turn away from all non-educational uses of tests based on grading. When examinations are used strictly for educational assessment purposes, to plan and modify instruction so that students can attain the learning goals, and not to rank and grade them, then we need not fear differential performance of different target groups, but actually look forward to it as an opportunity to help all students to succeed.

But teachers and students both need timely reports on progress for learning to work most effectively. So how can we live without using tests for grading purposes? This is a central question and there is a wonderful answer, indeed a very practical one, but as it is not within the purview of the present topic I will leave it for a future conversation.

Paul Zachos, Director
Association for the Cooperative Advancement of Science & Education (ACASE)
110 Spring St.
Saratoga Springs, NY 12866
phone: 518-583-4645
fax: 518-587-6467
paz@acase.org
<http://www.acase.org>