16 HUMAN DIVERSITY AND EQUITABLE ASSESSMENT

Edmund Gordon
~~City University of New York~~
Yale University (Emeritus) and
The College Board

The topic of human diversity and its implications for assessment is extremely complex. In this chapter, I will address just four of the major issues arising in this contentious domain. The first issue concerns the relationship between assessment and pedagogy itself. The second issue has to do with the implications of assessment that grow out of an assertion that the problems of diversity are probably not primarily problems for assessment but problems for education. Third, I will discuss the complexities of diversity as distinct from pluralism, both of which have implications for education and assessment. Finally, I explore some notions of how educational assessment can be made more sensitive and more appropriate for persons from diverse cultural backgrounds.

PROBLEMS OF ASSESSMENT OR OF PEDAGOGY

With respect to the first issue, I argue that the most fundamental problem concerning human diversity and equity in educational assessment has to do with the effectiveness, sufficiency, and adequacy of teaching and learning. When teaching and learning are sufficient, when they are truly effective, most of the problems posed for equitable assessment as a function of diverse human characteristics become manageable. The problems are not eliminated but they at least become manageable when the educational work on the front end is appropriate. Unfortunately, it often is not. It is when teaching and learning are insufficiently effective for the universe of students served that problems arise in the pursuit of equity for diverse human populations. It can be argued that the problems of equity in educational assessment are largely secondary to the failure to achieve equity through educational treatments. However, the fact that equitable educational assessment is only secondarily a problem of assessment does not mean that those of us in the assessment enterprise have no responsibility for doing something constructive.

It is not by accident that existing approaches to standardized assessment of educational achievement are insufficiently sensitive to the diversity of the stu-

dent population served as well as to pluralism and the social demands placed upon students—shortly, I will clarify what I mean by pluralism. Prevailing standards by which academic competence is judged are calibrated in large measure against either what most people at a specific level of development are considered able to do or what we agree is necessary in order for students to engage effectively in the demands of the next level of work. The fact that some persons have greater difficulty than others and seem unable to achieve these standards is usually thought to be a problem in the individual or reflective of group differences in abilities. It is not often thought to be a problem of the appropriateness of the assessment instruments or of educational practices.

In our efforts to be responsive to diverse learning characteristics and plural social standards, prevailing wisdom suggests that there may be limits to what can be done to design and develop sensitive assessment technology and procedures. We may be able to make the assessment processes more instructive; we may make them more supportive of diverse learning experiences; we may find more varied contexts and vehicles through which students can demonstrate their competencies; the items can be made more process sensitive and could give less emphasis to narrowly defined products. But in the final analysis, any assessment procedure is most likely to reveal the effectiveness of the teaching and learning to which students have been exposed. Thus the facts of diversity and pluralism may have more serious implications for teaching and learning than for equitable educational assessment technology and practice.

Previously, I asserted that this does not mean that those of us in the assessment community are off the hook. What I think it does mean is that both the teaching and learning end and the assessment end need to give greater attention to better understanding of the complexities of human diversity.

COMPLEXITIES OF DIVERSITY AND PLURALISM

When I speak of diversity of human characteristics, I am generally referring to the different positions people hold in this society, differences in status as well as differences in function. In both education and assessment if we do not distinguish between status characteristics and functional characteristics, we are likely not to make much progress.

Status defines one's position in the social hierarchy and that status, that position, often determines one's access to the resources of the society, to the economic and political power structures. Status influences access to opportunity and access to rewards. It influences how other people treat you, what other people may expect of you, and often what one expects from one's self. Traditionally, differential status has been assigned on the basis of social class or caste, of ethnicity or race, of gender, even of language and national origin. There is a host of so-called social dividers by which we define status.

In contrast, diversity in functional characteristics refer to the ***hows*** of behavior, to the ways in which behavior is manifested, the ways in which people act. These functional characteristics may be locally associated with certain status groups, but the manner of behavior is not associated with status. We include as functional characteristics such traits as cognitive styles, interests, and identity.

With each of these functional characteristics—remembering now that they define the how of behavior, the way people behave—we think more and more that they reflect the way persons engage learning experiences and the way people engage the environments to which they are exposed.

In the past year or so, I have been thinking that this may be too narrow a view of either functional or status characteristics. It may be that the most crucial aspect of the categories into which one falls is its contribution to the way one feels about oneself. Cultural identity may be more important than culture itself. In the modern world there is so much overlap in the manifestations of culture that if we were to focus only on cultural practices, only on the external facts of culture, we would have a hard time explaining how it is that Ed Gordon, who is exposed to so much of what is the mainstream culture of the United States, still acts like an African American, functions like an African American, and most important thinks of himself as an African American. What is important here is Ed Gordon's identification of self with that symbolic culture we call African American. Even when one deconstructs the culture itself, it has elements of many other cultures. Now this point may be made for all of the functional characteristics, that the differences between groups with respect to these functional characteristics are less important than the differences within groups. What may be most important is the contribution that the point of reference—the culture or the group—makes to the way the person identifies the self.

In order to clarify some of their implications for assessment and education, diversity of status and function needs to be distinguished from pluralism. In many of our writings we tend to use pluralism and diversity interchangeably as if they were synonymous. However, pluralism refers to the social demand of demonstrating multiple concurrent competencies in situationally relevant contexts. We recognize pluralistic demands most readily with respect to different cultures and languages. Those of us who are bilingual or multilingual have clear advantages over those of us who are monolingual. Similarly, those of us who can make ourselves at home in more than a single culture have advantages over those who can function in only one culture. I take great pride in being able to go back to North Carolina and be at home with my colleagues there or to come to Princeton or to City College and be at home with my colleagues there. If the context of the settings in which I am called upon to express my competencies qualitatively affects their expression, then this is tantamount to the development of pluralistic competencies. The requirement that all of us are being called upon to develop pluralistic competencies in a diverse world and a diverse society presents problems for us in what to teach and what to assess.

Let us turn next to the questions of equity and equality, which I contend often get confused. Much of our legal approach to democracy has been based upon ensuring that people are treated equally. However, when we come to human services, particularly education and health, equal treatment may actually be dysfunctional for some folks. What one needs is equitable treatment, that is, treatment that is appropriate to the characteristic and sufficient to the need.

When we begin to talk about the issues of diversity and of pluralism in the context of equity, I am simply contending that equity in these terms requires considerably more attention, first, so we can better understand it and, second, so that this understanding can inform what we do about inequity.

Because the issues concerning diversity and pluralism are far more complex than is often reflected in public debates, it may be useful to identify some of the possible ways in which a concern for population diversity and pluralistic outcomes impacts upon teaching, learning, and assessment. It is becoming more and more obvious that these sources of variance influence student motivation to engage in academic learning and to master its content. Recent research shows that some African American students are unwilling to engage in appropriate academic activity because they view those activities as identified with another group (**reference** ?). Furthermore, it is dysfunctional in their primary social interactions to exhibit behaviors that set them apart from their group rather than place them within the group. In my own work with Thomas and Allen where we are trying to help young citizens improve their cognitive skills and strategies, we found that many of these students actually have relevant strategies (**reference**). The students used them appropriately in a different context but were unwilling to use them in the academic context. Hence, the primary problem is not one of cognitive deficiency but the absence of a disposition to use particular kinds of cognitive skills in particular kinds of situations.

My oldest son, who is an anthropologist, talks about the possibility that much of the behavior we see in young Black males that we call dysfunctional or antisocial is actually resistant behavior. If the society is going to do something productive with these young people, we have to find ways to turn their resistant behavior into socially productive activities rather than try to control or contain the resistant behavior itself.

As a consequence, the ways in which I identify myself, the ways in which I perceive my status and my functional relationships with society, influence the way in which I engage in the learning opportunities of that society. Parenthetically, they probably also influence the ways in which I engage in assessment experiences. Not only do cultural identities influence the ways in which things are acquired but also the ways in which things are played out in day-to-day behavior as well as in assessment functions.

TOWARD MORE EQUITABLE ASSESSMENT

Many years ago, actually it was in the 1950s, I had the good fortune to work with a German woman who had come to this country just about the time of the rise of Hitler. Her name was Elsa. Elsa was a gentile but she did not want to be part of a society gearing itself up to treat Jews and Gypsies and other so-called undesirable folk the way Germany was treating them at that particular moment. And if I can digress formally at my age, I take advantage of this opportunity to express my politics. It is troubling to me that we see beyond the horizon in this country some of the same mean spirit that one saw in Germany in 1929, 1930, and 1931, namely, efforts to divide people in terms of their group characteristics, some as desirable versus others as undesirable and, more seriously, some as

superior versus others as inferior. Worse still, there is a strong tendency to view these group differences as God-given and genetically determined. These beliefs led to the destruction of German society, and many of us have been troubled by Herrnstein and Murray's (1995) treatment of these issues for similar reasons. But, many of us are almost as troubled by the concerted rush to reject these ideas because in the process of doing so we also gave undue prominence to that point of view.

Unfortunately, the nation seems to be ready for this kind of minority-group baiting because we seem to be increasingly fearful of the sufficiency of our material resources to serve all of us adequately, so some of us are to be set apart. Elsa saw this happening in Germany, but she was not politically assertive enough to fight it there. However, she did say that she would not stay in Germany and be a part of it. But Elsa's great contribution was not political but rather her scientific work in trying to assess educational potential in neurologically impaired children. She went to great lengths to construct alternative procedures for doing so. But the piece of her message that I want to stress as we move into a discussion of what can be done to make assessment more responsive to diverse human characteristics is her argument that any serious changes one might wish to make in assessment procedures need to be preceded by a change in our perception of the purpose of assessment.

Elsa saw the traditional purposes of assessment as being focussed primarily, almost exclusively, on classification and prediction. She argued that what we really need in assessment is better description and understanding as a basis for prescribing and intervening. It seems clear that Elsa's message is more appropriate for lower levels of education than for higher education and more appropriate for youngsters with serious disabilities than for those in the normal range. However, there are messages in her work that can be applied to higher education and to a wide range of diverse individuals because, in her view, the assessment experience should result in the examinees having more and better understanding of how they function and of how they can use the assessment information for subsequent learning and development.

What can be done to improve assessment? It must be 20 or 25 years ago now that Sam Messick and I put together a little proposal (that was never funded) to unbundle standardized test data in an effort to identify task demands and associated cognitive processes contributing to the total score. We wanted to isolate clusters of items with common task requirements so that subscores could be constructed at the process level. We hoped to better understand the construct-relevant sources of task difficulty in relation to the ways in which examinees engage or fail to engage relevant cognitive processes, so that we could provide richer score interpretations for informing instruction. Although our study died aborning, recent work by Snow and his colleagues has pursued similar aims in a much more sophisticated fashion, using rich data sets and detailed multivariate analyses (Hamilton, Nussbaum, Kupermintz, Kerkhoven, & Snow, 1995; Kupermintz, Ennis, Hamilton, Talbert, & Snow,1995).

In the past few years we have seen the emergence of what has been called portfolio assessment, which aims to collect a wide array of qualitative and quantitative information about students in an effort to improve understanding of the person being assessed. Portfolios afford an opportunity for students to present themselves in a variety of ways, even in ways that offer the best picture of the self. Another vehicle is to embed assessment in the curriculum itself or to embed teaching and learning experiences within assessment procedures. However, some of us worry that the assessment tail could wag the instructional dog. We become cautious when assessment is used for instructional as well as for evaluative purposes.

To dispel some of these concerns and to provide some concrete suggestions as to how to make assessment more responsive to human diversity, I next present seven assessment/educational interventions. These seven suggestions are not offered as firm recommendations but rather are in the spirit of needed research and development. The first suggestion is to radically increase the teaching, learning, and assessment experiences, to increase the diversity of the kinds of learning and assessment tasks available as well as diversity in the contexts and demands of learning and assessment. This particular suggestion stems from considerations outlined in earlier sections of this chapter that what we have to build into the assessment experience is a much wider opportunity for choice in ways in which students can demonstrate what they know and do not know.

Incidentally, I recognize that there are problems in systematizing evaluative judgments based on individualized choice of tasks, whether by students or by teachers. This issue of systematization or standardization takes me back to one of Elsa's abiding concerns. Elsa went into retirement and ultimately to her death complaining that she was never able to standardize her assessment procedures. Some of us thought that if her procedures had been standardized, this would have defeated her purpose. It may very well be, as we think about our other research and development suggestions, that if we give too much attention to how we standardize scores now, to how we calibrate them systematically to make comparable judgments across persons and circumstances, it may be counterproductive to the purposes behind assessment responsive to human diversity.

The first of these suggestions was for far greater diversity in the opportunities to learn and to express one's learning in test situations. The second has to do with increased flexibility in the timing of assessments, which was examined by Bennett in Chapter 14.

The third suggestion has to do with recognizing and enhancing the multiplicity of perspectives to which students are exposed in their learning as well as the perspectives from which they are encouraged to express themselves in their assessments. What we see is influenced by the frame of reference or perspective we bring to the stimulus field. When I am exposed to a phenomenon or come to any experience, I see through eyes that have been shaped by a particular background and by particular experiences. If what is being looked for by the person who is assessing or what is being offered by the person who is teaching does not take my perspective into account, my responses and constructions may be misinterpreted. Insensitivity to student perspectives presents problems for us because it is not just that multiple perspectives represent alternative viewpoints, some of

them right and some of them wrong. The question of how one selects from a variety of perspectives and makes judgments is one of the research issues that needs to be pursued. I am convinced that at both the instructional level and the assessment level if one is insensitive to student perspectives the interpretation of assessment data will be off-target, distorted, or otherwise invalid. Furthermore, a way of seeing is also a way of not seeing in our assessment and teaching. We have to be much more sensitive to differences in perspective and recognize that these differences are influenced by the cultural backgrounds from which people come.

The fourth suggestion is a proposal to sample from noncanonical as well as canonical voices, knowledge, and techniques. The key question is what is appropriate to the canon in the context of cultural diversity? As we hear new voices in different cultures, their automatic inclusion in the canon is almost as bad as their exclusion. If one is confronted with bad literature from Asian culture or African culture, say, it does not do justice to students from those cultures by accommodating them with the inclusion of bad literature in teaching or assessment. Rather, it is patronizing them. One of the practical problems here is not being sufficiently familiar with these new and diverse voices in literature in general so as to select appropriately a representative teaching or assessment package. Selections from noncanonical sources need to be made with the same care and attention that we give to the more familiar canonical selections.

The fifth suggestion has to do with allowing student self-selected choice as opposed to teacher- or examiner-determined options for demonstrating what is known. The introduction of widespread student choice into assessment presents problems because there is some evidence that students, the people being tested, are not often the best judges of what they ought to be tested on or of the best way of presenting themselves. Hence, we need to do more research not only on the effects of self-selected choice in learning and assessment, but also on how to help students make better decisions about what they need as well as what they want and about how to put their best foot forward. If we go back to my work with Thomas and Allen **(reference)** on cognitive modifiability, what we began to recognize was that when the target students were able to identify what we were talking about in a context familiar to them, when they were able to use their own experience base, they understood the concepts. But when we presented material in the academic language and academic context, they did not. This finding implies that we should provide choices in learning and assessment that are compatible with the students' own experiences.

As a sixth point, we need to provide opportunities for both individual and cooperative learning as well as cooperative performance opportunities. In Lauren Resnick's **(reference)** article on learning in and out of school, she reminds us of how in school settings and formal assessment settings, we do not want people to cooperate with each other. Yet on entering the world of work, if you cannot cooperate with others you are lost.

The final suggestion is for student design of tests to cover learner- or examinee-generated knowledge. If you are trying to learn what I know about chemistry, biology, English, or the social sciences, why not give me the opportunity to define the domain of knowledge I think I am an expert in? Some years ago out at the Rand Corporation they were experimenting with approaches to computer-based instruction. One of the lines of investigation being pursued was self-designed tests whereby an examinee constructed his or her own assessment from a pool of items and assessment situations. Letting students define the boundaries of their knowledge and understanding in a subject area may be the best way of assessing their knowledge structures.

Having laid out this program of research-and-development to make assessment more responsive to human diversity, I must remind us of the point at which we began: Unless persons have had adequate educational and social opportunities to learn and develop, whatever we do in assessment is not going to help very much.

REFERENCES

Hamilton, L. S., Nussbaum, E. M., Kupermintz, H., Kerkhoven, J. I. M., & Snow, R. E. (1995). Enhancing the validity and usefulness of large-scale educational assessments: II. NELS:88 science achievement. *American Educational Research Journal*, *32*, 555-581.

Ed - we have this reference!

Herrnstein,

Kupermintz, H., Ennis, M. E., Hamilton, L. S., Talbert, J. E., & Snow, R. E. (1995). Enhancing the validity and usefulness of large-scale educational assessments: I. NELS:88 mathematics achievement. *American Educational Research Journal*, *32*, 525-554.

pg 4 1st (ref)

Fordham, S., & Ogbu, J. U. (1986). Black students' school success: Coping with the "burden of 'acting white'." Urban Review, 18, 176-206.

pg 4 & 7 1st 2nd (ref)

Gordon, E.W., Allen, B.A., Armour-Thomas, E. (1988) "The Development & Enhancement of Cognitive Competence of Educationally Disadvantaged High School Students." A Report to the Exxon Foundation.

pg 7 2nd ref

Resnick, Lauren B. (1987). Learning in School and Out. Educational Researcher, 16,9 pp13-20.