

Introduction: Advances in Assessment and the Potential for Increasing the Number of Hispanics in Higher Education



Gary D. Keller

Assessment and Access: Hispanics in Higher Education is the first book of its kind. It explores for the purpose of increasing Hispanic access to higher education a variety of developments in educational assessment and test familiarization, both theoretical and practical. At the same time, the book aspires to present a balanced treatment of the relevant issues by a distinguished group of specialists, both within the higher education community and at the Educational Testing Service which is making history in this area.

I have several purposes in writing this introductory essay. First, I would like to make some observations about the value of educational aptitude and achievement tests, including standardized tests for the United States Hispanic community. My view, which is not necessarily shared by everyone, is that tests such as the Scholastic Aptitude Test (SAT), the American College Test (ACT), the Advanced Placement (AP), the Graduate Record Examination (GRE), and others have been moderately helpful in expanding access for U.S. Hispanics to college, and have the potential for being more so in the future. (I limit my affirmation to educational tests; the issues that revolve around other assessments, particularly mental tests, are more complex and the value for Hispanic Americans of these measures is more dubious.¹) Indeed, part of the rationale of this book is to explore the ways that tests can be better constructed, more effectively implemented, and better prepared for by Hispanic youths so that their aptitudes will be more accurately measured.

A second goal of this essay is to recount, for the first time, some of the historical background that has led to the creation of *Assessment and Access*, focusing particularly on the earlier interactions and subsequent formal partnership between the Hispanic Higher Education Coalition/Hispanic Association of Colleges and Universities (HHEC/HACU), the College Board, and the Educational Testing Service (ETS). I think readers will find this an interesting story with instructive implications for what can be accomplished on behalf of minority students and parents when minority organizations work in partnership with educational organizations toward common, well-defined goals. In turn this historical background should provide some additional understanding of the papers that constitute this book. Closely related to the purpose of providing background information, I will also review some of the most productive or potentially productive projects in the present or recent past that have utilized assessment instruments as an integral component of their academic initiatives. The projects or products that I have in mind include the Garfield High School initiatives in East Los Angeles; Options for Excellence in San Antonio; the Hispanic Student Success Program (HSSP) in San Antonio, Northern New Mexico, and other sites; Project PRIME (Project to Improve Minority Education) in Arizona; *TestSkills*; and *Sí Se Puede: Information on Academic Planning and Obtaining Financial Aid*.

My final goal is to introduce briefly each of the component chapters of this book, situating them within the context of the ongoing assessment challenges that confront the U.S. Hispanic community that will have been developed earlier in the background section of this essay.

Some Perspectives on the Relative Value of the Tests for Hispanic Access

Let me begin by facing straight on a critical issue: Can tests and other evaluation instruments help Hispanic Americans to enter higher education in greater numbers? Presumably if they can in general, then advances in evaluation, particularly if it is focused on the challenges that confront the accurate testing of Hispanics, will help our community even more effectively.

The answer to the question that is posed is not self-evident, certainly not for the U.S. Hispanic community and probably not for other minorities underrepresented in higher education as well. We all know, of course, that the attitudes of some segments of the U.S. Hispanic community toward published tests, particularly standardized tests, have sometimes not been positive. Many Hispanics have been and continue to be opposed to standardized tests, and most particularly to I.Q. tests, as unfair to Hispanic

children. This is not peculiar to the Hispanic community; a similar situation obtains with respect to other minority groups who are underrepresented in education, the professions, or privileged positions in society. Moreover, many teachers, parents, and school administrators find the current tests unsatisfactory as measures of a student's educational progress. In the past, groups such as the National Education Association and others have called for a moratorium of testing. Finally, because the overall record of higher education is not a compelling one for U.S. Hispanics and other underrepresented minorities, by reflection it is not possible to make a compelling argument for the efficacy of aptitude and achievement assessments on behalf of our students. In fact, over the past fifteen years, using our own past level of performance as our standard (even though this performance certainly is modest), while in absolute numbers we have seen very small increases in some academic areas, we have been going in reverse direction when we measure the *percentage* of Hispanics, Blacks, and Native Americans who graduate from high school from the total adolescent cohort eligible to do so and when we measure the *percentage* of high school graduates from these underrepresented minorities who in fact go on to college.

My view is that standardized tests are not the major cause of this retrogression that began about 1975. We should point to much broader problems, particularly the waning of commitments over the last decade among educators and our society at large to equality of educational access, to affirmative action, and to providing effective financial aid for economically disadvantaged students. Nevertheless, in contrast to the post-World War II period, where it can be argued that admissions tests went hand in hand with greatly expanded college attendance, the phenomenon of smaller percentages of underrepresented minorities entering college since the middle 1970s does diminish the argument that can be made in favor of educational admissions tests. Moreover, in various state legislatures of state capitols there have been movements over the past few years, as part of the goal of establishing more rigorous admission standards for public higher education, to overrely on admission test scores for that purpose (Breland, 1985; Goetz and Johnson, 1985). Gregory R. Anrig, president of the ETS, has been one of the first to caution against this overreliance and to call for a balanced use of multiple predictors in determining college admissions. (Anrig, 1985,6). In addition, Anrig has cited a Committee of Ability Testing of the National Academy of Sciences that observed, "a policy decision to base an admissions program strictly on ranking applicants in order of their expected success will tend to screen out minority candidates . . ." (Wigdor and Garner, 1982, 196). The committee judged that the goal of admissions decisions "should be to effect a delicate balance among the principles of selecting applicants who are likely to succeed in the program, of recognizing

excellence and of increasing the presence of identifiable underrepresented subpopulations . . . ” (Wigdor and Garner, 1982, 196).

The observations of the Committee on Ability Testing are especially important in light of Richard Durán’s subsequent study (1983) which reviewed a considerable amount of data and arrived at the conclusion that both high school grades and admission test scores were not as good predictors of U.S. Hispanics’ college grades as they were of White non-Hispanics’ college grades. Durán concludes that “admissions officers ought to rely critically on the overall profile of Hispanic students in making admissions decisions. The results of studies reviewed here suggest that admissions personnel need to be provided with a broader range of information on Hispanics’ background, language, and culture in weighing admissions decisions” (Durán, 1983, 105).

On the one hand we are faced with the skepticism about some tests and opposition to others by many U.S. Hispanics; the poor record of higher education since 1975 for underrepresented minorities, which in turn impinges against our making a strong argument for tests expanding Hispanic American access during these years; the movements such as those at the state level, to overrely on or use tests inappropriately supposedly to make educational standards more rigorous; and finally, the current conclusion based on the best data available that high school grade point averages and admission test scores do not predict academic aptitudes as accurately for Hispanics as for majority students. What arguments countervail these factors that argue against or at least mitigate the conclusions that tests can help more U.S. Hispanics enter college and that better tests and better test preparation can be even more effective?

I believe that a number of strong arguments can be brought in favor of the use of test for enhancing Hispanic education. Moreover, the potential for improving existing test construction and the technology of test administration, if realized, can further increase Hispanic access. Finally, I believe the pro-tests arguments outweigh the contrary, as substantive as the latter are.

While one sector of the U.S. Hispanic community either is afraid, skeptical about, or opposed to standardized tests, that is not necessarily the most knowledgeable sector. There now exists for the first time a significant group of U.S. Hispanic professionals who are either experts or quite knowledgeable about issues of educational assessment as they impact our community and who have a constructive, critical, but essentially positive outlook, especially in the longer term, about tests for our students. Some among the professional Hispanic community have argued that certain tests such as the SAT, GRE, LSAT, GMAT, and others do in fact significantly underestimate the aptitudes of Hispanic students. We make this argument, however, not

with the goal of doing away with tests, but improving them. Moreover, we have begun to use achievement tests such as the AP in distinctive ways both to certify the academic performance of our students and to enhance their college admission possibilities.

Part of the challenge in making standardized tests more accurate predictors of our community's aptitudes and achievements is to address the ignorance and anxiety about tests that affects our community at each level, from the students who take them through parents, teachers, and administrators who have uninformed attitudes about them. Reflexive kinds of anxiety, blanket opposition, and ignorance about the issues do not help us. There is much to be vigilant about with respect to the construction, administration, and interpretation of tests with respect to Hispanic Americans. However, a level of sophisticated knowledge about the tests is necessary to put them to optimal use or to combat their misuse or their faulty construction if they are doing our community a disservice. We have witnessed, for example, in the area of language assessment for the purpose of bilingual education, a relatively sophisticated and positive use of tests that has helped Hispanic students (see Keller, 1982).

While it is true that the percentage of U.S. Hispanics who graduate from high school and go on to college currently is in decline when compared with 1975, and that in some states there are initiatives to overvalue test scores in a way that excludes minorities, it is also documentable that in the history of U.S. higher education, tests have been used more often to open doors rather than close them. The direction—inclusion or exclusion—is determined not by the tests but by the establishment of educational policies for their use. For example, around the turn of the century, the College Board was formed to administer a uniform system of college entrance examinations to address the then chaotic situation where each college had its own subject matter examinations. This development permitted the enrollment of students from secondary schools that were unknown to the colleges which in turn enabled the latter to broaden their student bodies. During the post-World War II period, the use of admissions tests increased most dramatically precisely when quotas were eliminated or at least mitigated at private colleges, including the Ivy League; when public education greatly expanded; and when higher education opened for the first time in a genuine way to racial and ethnic minorities and to all socioeconomic levels. Although the early history of both the College Board and admissions tests was a checkered one, selective admissions to higher education was more a reflection of social class and economic status prior to the use of national admission tests than it has been since. Before the establishment of standardized tests, admissions officers at the elitist colleges of that period relied primarily on grades and recommendations for students from a small group

of well-known, elitist, college preparatory schools. The introduction of standardized tests opened opportunities, modest at first, greater in the post-World War II period, for students from high schools without reputations among the selective colleges to provide a demonstration of their academic abilities and to be admitted.²

The broadening of opportunities for minorities and lower socioeconomic status students was greatly enhanced by the College Board in its role in advancing and developing, through the College Scholarship Service, the concept of awarding financial aid based on need. This has been an important factor in expanding access to college for low-income students. In recent years both the College Board and the ETS have launched a variety of talent-search, guidance, scholarship, diagnostic, instructional, and other projects or products on behalf of minority and other deserving students. Several of these projects are highlighted later in this essay and elsewhere in this book.

The tendency in some states toward policies that overrely on tests is not an argument against tests themselves. Tests have been used productively in recent years, projects such as Options for Excellence, Hispanic Student Success Program, Project PRIME and the Garfield High School initiative have in one way or another used tests as an integral part of their educational activities with considerable success both from the perspective of performance and public recognition of the achievements of U.S. Hispanic students. Is it an exaggeration to view Jaime Escalante's and Henry Gradiillas' use of the AP as a certifier of academic achievement, especially their decision to retest Hispanic students when their scores were challenged by the ETS, as the most inspiring educational and assessment phenomenon for Hispanics in the United States in recent memory? The developing potential for misuse of standardized tests at the state level requires better preparedness on the issues and political strength on the part of minorities to call into account and change bad policy. It should be noted that past studies (Hilton and Rhett, 1973) have shown that a very large majority of students intending to go to college do so and that in recent decades tests haven't held back students who have taken them and completed their applications. A study conducted for the American Council on Education found that 75% of freshmen were attending their first choice college and nearly 95% were attending their first or second choice college (Astin, King, and Richardson, 1978, 18). The irony should not be lost on the Hispanic and other minority communities that it is at the point when underrepresented minorities have become the majority in many major school systems that state officials have increased their concern about admission standards in public higher education. Minority groups must mobilize at the state levels to resist policies that

will consciously or unwittingly reduce the commitment to minority students or reduce their presence on publicly funded college campuses.

That both high school grads and test scores do not measure Hispanics' aptitudes as accurately as they measure majority students' points also, in my mind, to necessary initiatives to improve education for Hispanic students and to improve standardized test measurement of our students. This is a problem *both* in education and its measurement. We know that our students are capable of achieving much more than what is asked for them in most of the schools that they attend. Durán's 1983 study suggests that interpretation. In addition, when we have established creative and distinctive programs for our students, such as some of those described later, the tests administered to those students have confirmed the reality of their high achievement, sometimes to the consternation of would-be nay-sayers. The tests need to be improved to assess more accurately U.S. Hispanics but even in their current state they have certified the achievements of students in East Los Angeles, in Greater San Antonio, and in Arizona.

I have reserved for last what in my mind are the stoniest arguments for this book, for optimism, for the improbability of tests, and for their enhanced effectiveness in both measuring Hispanics and helping them go to college in higher numbers. These arguments revolve around the essential fact of assessment, that evaluation abhors a vacuum, around the scientific nature of standardized tests as compared to the alternatives of the past (quota systems, subjective appraisals often conducted by persons hostile to minorities), and by virtue of what the HHEC/HACU, ETS, and the College Board have been able to achieve in partnership thus far.

It is surprising that the arguments against tests by some educators are made in a sort of idealized context, a vacuum if you will. Standardized tests are a relatively new development, but evaluation of one type or another always has been with us. In arguing against tests it is essential to be cognizant of the alternatives. I belong to a group of Hispanic professionals—bilingual educators—who have been exposed to some of the alternatives. We frequently have seen the value of standardized and other "objective" tests (perhaps "machine scorable" is a more appropriate term for what in fact is currently meant by "objective") for bilingual education and the assessment of other than English language skills. During the Vietnam War while some young people were sailing up the Mekong River, I made my journey as a young educator up the Rio Grande or Río Bravo, depending on which side you view the river. To gloss Conrad and Coppola, there is a "heart of darkness" along that river as well. In south Texas and other places, I found egregious misplacements of Spanish-speaking students into classes for the "educable mentally retarded." These students were bright

enough but not fluent in English, and teachers misplaced them wholesale on the basis of their own subjective, wrongheaded, and, given the system, unappealable appraisals. These experiences of subjective appraisals will haunt me to my last breath. They are countervailed, if such a thing is possible, by experiences I have had with the positive role of tests such as the Language Assessment Battery which measures fluency and dominance in English and Spanish and which was commissioned by the New York City Board of Education to implement the court-approved ASPIRA consent decree that permitted hundreds of thousands of children to receive bilingual education (Santiago, 1978 and 1986).

The point that needs to be reaffirmed is that in this world evaluation is constant and everpresent. People have been constantly appraising and sizing up each other from prehistory, and there is no likely end to this process. Those who wish to do away with more formal modes of testing in expectation that we will establish an idealized, test-free, and of course, less stressful society, are harboring ill-founded illusions. In a vacuum, evaluation will proceed on the basis of subjective opinions, often by people in power or privilege who harbor either negative stereotypes, or much worse, active, racially motivated hostility toward Hispanics and other minorities. It seems to me highly preferable that we rely on tests, including standardized tests, rather than subject our students to individual, subjective appraisals. Tests, no matter how flawed, at least have the positive features that they are the application of the social sciences, metrics, and education, and that they can be reviewed. Tests can be and usually are, as products of science, open to review, analysis, reasoned criticism, replication, analyses of their effectiveness through measurement of their reliability and validity, and above all, improvement through revision over the decades. In a world where assessment is both essential and unavoidable, I do not believe that the same positive features or levels of openness can be established for any alternatives to tests, including standardized tests.³ Certainly subjective appraisals by experts or would-be experts can not be monitored or controlled with respect to underlying attitudes toward various minority groups such as U.S. Hispanics. When the Garfield incident attained national recognition many people locally and elsewhere did not believe that so many Hispanic students in a lower socioeconomic status environment and undistinguished school could take AP calculus and excel. It was the genius of Escalante and Gradillas that amidst the tension they took advantage of all that the tests had to offer. The retest procedure was incontrovertible and cast the all too human doubts to the winds.

Standardized tests are a new scientific development. They emerged, particularly at the national level, with the perceived need by the military to test potential recruits during World War I. These robust juveniles in the

world of applied science have a long road to travel to maturity, but they are, in the main, improving. Moreover, the tests have built within them their own refining devices. Testmakers have the ability to measure the effective reliability and validity of the instruments that they have created. Because of this, tests improve with revision; over time each version or iteration usually becomes more accurate. Also, as some of the papers in this book suggest, tests can use developments in educational technology to their advantage. Testmakers can incorporate innovations in science, computerization, and educational delivery systems into new and possibly more effective assessment techniques.

An additional argument, as we shall see below, for the value of tests in the effort to increase Hispanic students on college campuses rests in the review of past and ongoing projects that have used them to effective purpose.

Historical Background and Review of Some Important Educational Access and Assessment Projects and Products

The issue of openness in relation to tests and the reporting of their results is not merely general or theoretical, but very germane to the developments that have led to the publishing of this book. This book is timely; actually it's a few decades overdue. It is timely because over the past ten or fifteen years there has accrued a sufficient number of researchers, a sufficient amount of research, and certainly a more than ample number of U.S. Hispanic students for the relevant issues to receive widespread attention. So: overdue, but timely, which is the way things work sometimes in science, but usually less well in less open systems. It has taken and continues to take a great deal of work, some of it adversarial, some cooperative, to convince educators generally and testmakers specifically, of the distinctiveness of the Hispanic American experience and of the reality of educational discrimination, test "error," and other forms of measurement inaccuracy. Part of the problem was that until the 1960s there were few Hispanic specialists in any field, much less psychometrics, statistics, educational assessment, and the like. We were victims with no voice to plead our oppression. Nevertheless, at first very gradually, and lately with vigor, we arrived, we documented, we drew our conclusions, and we promoted them with strength. This book is timely and assuredly not the last of its kind.

Beginning in 1979, as a result of the HHEC's involvement in test improvement issues, the Hispanic professional community began a process of interaction with testmakers, primarily the ETS and the College Board, and less so with the American College Testing Program. The relationship

between Hispanic educators and researchers and the testmaking industry can be separated into two phases. The first phase occurred between 1979 and 1983 and was of an informal nature. The second phase began in 1984 and continues through the present. In 1984, the HHEC, the College Board, and the ETS agreed to establish a formal partnership. The HHEC polled its organizational members on the priority issues which helped establish an agenda for the three organizations of both practical interventionist projects and research studies. At the same time the College Board and the ETS established a yearly budget with which to work and the three organizations have been meeting periodically and completing projects ever since. The partnership has worked well and is likely to continue productively for the foreseeable future.

Between 1979 and 1983 some of the achievements of the Hispanic community and the testmaking industry included the disclosure of test score distributions by race and ethnicity for most of the major tests; the hiring at the ETS of more U.S. Hispanic researchers and the expansion of research on educational and assessment topics important for U.S. Hispanics; the initiation in 1981 of the very successful Options for Excellence in Greater San Antonio; the publication of Richard Durán's seminal study (1983); and the sponsoring of research conferences, the most productive of which aided the completion of M. Olivas' *Latino College Students* (1986). Independently of other influences until the 1982 challenge by the ETS, it was also the period of academic achievement by increasingly larger numbers of Hispanic high school students at James A. Garfield High School in East Los Angeles.

The period from 1984 through the present has led to the establishment of Project 1000 to increase the number of Hispanic graduate students; the burgeoning influence of the Garfield phenomenon on Hispanic educators and its replication elsewhere; the production of *TestSkills* and *Si Se Puede: Information on Academic Planning and Obtaining Financial Aid*; the expansion of promising research in areas such as differential item functioning; the initiation of the Hispanic Student Success Program and Project PRIME; the beginning of the Consortium to Identify and Promote Hispanic Professionals; and the ETS- and College Board-sponsored conference and other forms of support that have led to the publication of this book.

The year 1979 is a critical one because the forces or factors that led to a productive collaboration between the Hispanic community and the testmaking industry coalesced during the heyday of what was then commonly called the "Truth-in-Testing" movement. As a result partially of earlier successes in consumer initiatives such as "truth in advertising" and "truth in lending" and spurred on by considerable dissatisfaction with the general way that such national tests as the SAT and ACT were being used, a vigorous initiative to reform tests or to eliminate them was undertaken in the

late 1970s by a number of groups, perhaps the most prominent of which was Ralph Nader and his associates.

In 1979, the HHEC participated in the ongoing debate by providing testimony, which I prepared and delivered with my able colleague Dr. Alvin Rivera, with respect to bills H.R. 3564 and H.R. 4949 (Truth in Testing Act of 1979; The Educational Testing Act of 1979). At the time these bills, which did not pass Congress, were before the Subcommittee on Elementary, Secondary, and Vocational Education of the Committee on Education and Labor of the House of Representatives. The HHEC testimony was constructive: critical of some aspects of testmaking, administering, reporting, and interpreting that were prevalent at that time, but also supportive of the tests and cognizant of their value both generally and for Hispanic Americans. The interaction between the testing industry and the professional Hispanic community dates from that time.

The testimony that the HHEC gave on October 11, 1979, related to a number of issues that are just as relevant today as they were over a decade ago. For example, we pointed out that researchers had demonstrated how "the tests use ambiguities and misleading language in order to make their questions inappropriately difficult," although researchers had not yet reviewed the possible disproportionate effects of such factors on ethnolinguistic minorities including U.S. Hispanics. We called for research "to determine the effect of such artificial test language on minority linguistic and cultural groups." We urged that the testing community or industry consider "a possible interaction effect between artificially created, misleading English test language and the linguistic parameters of the U.S. Hispanic communities." (Keller, 1979, 829). We asked that the testing industry consider our view that a proportion of what standardized test purported to measure in the scholastic aptitudes of Hispanic students was merely "a subterranean index" of purely linguistic or cultural peculiarities of various Hispanic communities. Even the ETS appeared to recognize that issue, although they might have been more inspired by marketing than equity at the time, because they had established ten years earlier a Spanish analog of their GRE, not a translation but a test independently developed for Spanish-speaking students. That test, the Prueba de Aptitud para Estudios Graduados (PAEG) as well as the Prueba de Aptitud Académica, were and still are used in Puerto Rico and parts of Latin America for measuring aptitudes. The first is used for admission for graduate study and for the study of law and the latter is analogous to the SAT.

The HHEC made another important suggestion: to control for the biases in the testing component by using additional norming procedures. Specifically, we proposed that "scoring be based on the achievement of a student using his or her language and culture groups as the norm or the

control." It's a little hard to believe from the vantage point of 1990 (this disbelief in itself is a good index of the progress that can be made in a scientific domain such as testing) but at the time, the College Board, ETS, and ACT had refused to report test distribution scores by race and ethnicity much less to report individual scores by major discernible ethnolinguistic groups. The claim (made verbally but somewhat veiled in writing⁴) was that because minorities did much more poorly on the tests than majority students, that they wanted to "protect" minority students from the allegations of innate inferiority made by some researchers at the time. Some Hispanic educators ventured the opinion that the refusal to provide the data might also have to do with the fact that on the face of it, Hispanic and other minority students were receiving scores about two standard deviations below White students on the average and that this could cast doubt about the accuracy of the tests. In any event, the testmaker's posture was completely counterproductive. Instead of shielding minorities, those who would promote Social Darwinism or other claims for innate racial superiorities or inferiorities were in fact aided by the sense that the testing community had a grisly secret that it couldn't permit to see the light. As for researchers whose work could have benefited from scientific openness, they had to use a roundabout method of correlating test scores with race and ethnicity by estimating them from the relationship of scores and reported levels of income which was published.

The testimony of the HHEC concluded with an expression of support of the testing industry and of objective tests, including standardized tests, although at the same time it fully supported the movement toward more accountability in testing. "It is because the Hispanic educator probably more than any other minority group professional has been privy to the worst, most biased aspects of the testing process, as well as its potential for accurately assessing competencies and determining placement, that these suggestions are entered in the spirit of a thoroughgoing reform of a service that needs to regain public trust in order to continue to be productive and beneficial to society" (Keller, 1979, 837).

In reviewing the testimony and the suggestions that were offered at that time, I am struck by how much in the 1979 document has come to some degree of fruition in the twelve years that have transpired, and as a result finds consideration in *Assessment and Access: Hispanics in Higher Education*. Examples of what I am alluding to follow below. I should observe, however, that much of what has been accomplished reflects a response by the admissions testmakers to the general movement toward test reform on the part of powerful social forces, particularly consumer forces. Thus the HHEC testimony may have framed some of the issues of the time in a way that best described them for the U.S. Hispanic community, but it

can not be claimed that the HHEC testimony or any one specific group initiative for that matter, actually caused certain broad test reforms. On the other hand, with the establishment of the informal collaboration beginning in the early 1980s and in 1984 became a formal partnership among the HHEC (which has recently affiliated with HACU), the ETS, and the College Board, some features of the HHEC program have been specifically implemented.

The HHEC had referred to the problem, particularly severe for Hispanic Americans, of the use of ambiguities and misleading language in tests and for the need to determine the effect of such artificial test language on minority linguistic and cultural groups. In December 1979 the ETS proposed and, after considerable review by the broad educational community, its Board of Trustees adopted, in October 1981, *Public Interest Principles for the Design and Use of Admissions Testing Programs*. These principles together with new *ETS Standards for Quality and Fairness* which were adopted at the same time, in my judgment, although some might disagree,⁵ have helped considerably in meeting concerns about the design and use of standardized tests in admission to higher education. Among the commitments that were affirmed or, since they were already in effect, reaffirmed were: (1) publication of test content to a degree limited only by reasonable safeguards of efficiency, cost, quality, and the educational impact of the programs; (2) commitment to continue to maintain and strengthen credible procedures for detecting bias and eliminating it from the content of tests, while making such procedures visible to the public; (3) commitment to formulate, maintain, and publish widely principles of appropriate use of scores and other test information derived from testing programs and to be alert to and actively discourage misuse; (4) assurance that operational forms of the test be independently reviewed before they are given and that the review include the appropriateness of the content of the test and in particular seek to detect and remove potential racial, cultural, or sex bias or other influences extrinsic to the characteristics, skills, or knowledge to be measured.

At the same time the ETS Board of Trustees formalized its support of the principle of openness in admissions testing. Gregory Anrig, who had just arrived at the ETS at the time from his former position as commissioner of education for the Commonwealth of Massachusetts, pointed out in testimony to Congress on November 4, 1981, that the real issue was not "truth" but openness in testing (Anrig, 1981, 3). As I have stated above, I believe openness to be a key advantage of standardized tests, one that when realized helps genuinely to distinguish them as scientific pursuits. The term "truth," taken from earlier consumer initiatives, is much too extravagant a descriptor for what tests currently can provide, whether in terms of correlations between scores and subsequent performances, philosophic explanations

of what the tests are actually measuring, or most anything else. The ETS board observed that openness should be construed very broadly and that therefore it involved more than mere test disclosure. Moreover, at the ETS the commitment to reviewing every test in order to eliminate references or language that might be stereotypical or objectionable to any subgroup was reaffirmed and strengthened. The methods have been formalized and currently are described as the "ETS Test Sensitivity Review Process" (Hunter and Slaughter, 1980). Finally, the ETS Board of Trustees committed to developing an industry-wide "Code of Fair Testing" for the development and administration of educational testing nationwide and recommended that this Code be based on the new standards then being developed by the American Psychological Association, the American Educational Research Association, and the National Council on Measurement in Education, and which were completed in 1985.

One strand of research that has emerged since 1979 and which has attained considerable prominence is called "differential analysis" or "differential item functioning." (In this volume, Alicia Pérez Schmitt and Neil J. Dorans review the emergence of this line of research from earlier research conducted since the 1950s.) While differential analysis research is relatively new, it is burgeoning and holds promise for addressing the specific problems of ambiguities, misleading language, cultural or linguistic bias, and the like to which HHEC has referred.

One of the most important developments for the U.S. Hispanic community to date with respect to testing error that may reflect linguistic and/or cultural bias has emerged from the work of Richard P. Durán, who at the time that he published *Hispanics' Education and Background: Predictors of College Achievement* (1983) was a research scientist at the ETS. Michael A. Olivas, who was one of the founding members of HHEC as well as the chairman of its board for several years (and who is currently a member of the board of trustee of the College Board) wrote the foreword to this book in which he observed, "Durán's review found that almost every study conducted showed noticeably lower correlations between standardized admissions tests and first-year college grades for Hispanos, relative to Anglos. Given the extraordinary high school attrition rate for Hispanic children and the likelihood that the very best survivors are taking tests, he is generous towards those who would employ these tests uncritically in admissions decisions" (vii).

In the test world just as in the art world provenance is important. Durán was writing as an ETS scientist and the book was published by the College Board.⁶ The best of both worlds because both organizations took seriously the fact that the tests when combined with high school grades were "roughly nine percent less accurate than use of the same procedure to

predict Whites' college grades'' (Durán, 1983, 102). As a result of the Durán study, the ETS has attempted to alert the educational community to this state of affairs and to admonish admissions decision-makers to take into account the findings of a less strong association for Hispanics compared to Whites (Anrig, 1985, 1986). The wheels grind very slowly in education, however, in fact much more slowly than in testmaking. It is incumbent on those who are interested in and knowledgeable of U.S. Hispanic academic and test-taking performance to continue to promote the distinctive situation of Hispanic students among the higher education community. This is precisely what Durán has done in his new role as professor and graduate officer at the University of California, Santa Barbara, where he has continued and expanded his research, particularly, as the paper in this volume bears out, to the assessment of cognitive skills.

A second suggestion made by the HHEC was to consider administration of Spanish language tests in addition to (and in a few distinctive cases instead of) their English analogues. The two Spanish language tests that are used very extensively in Puerto Rico and in other Spanish-speaking countries as well for assessment in higher education are the Prueba de Aptitud Académica (PAA), which is analogous to the SAT, and the Prueba de Aptitud para Estudios Graduados (PAEG), which is somewhat akin to the GRE as well as the LSAT. In chapter six, the paper by William H. Angoff and Linda L. Cook review the technical and statistical work that has been accomplished thus far in equating the PAA and the SAT. With respect to the PAEG, initiatives along the lines suggested by the HHEC have been established on an experimental bases. A number of graduate institutions that are participating in *Project 1000: Recruiting, Admitting, and Graduating Additional U.S. Hispanic Students*, have agreed to accept scores on the PAEG and compare them with scores earned on the GRE, and of course with subsequent performance in graduate school. This experimental study only started in earnest in 1989, but since in a considerable number of cases we have found the PAEG scores to be much higher than the GRE scores, we look forward to deriving interesting and useful conclusions.

In 1979, the HHEC had suggested controlling the inaccuracy in measuring Hispanics in the tests by reporting the scores that Hispanic students earn compared to other Hispanics (what we call the "within ethnic" score). We believed that this score could serve as an important supplement to the score that Hispanics earn with reference to the entire universe of test takers for a given administration. This is particularly the case since as Robert J. Solomon, then executive vice president at the ETS, pointed out that "while scores for whites and disadvantaged minority groups overlap, a typical result is to find that only 10–20% of disadvantaged minority groups score above a point that is average for white (i.e., exceeded by 50% of whites)''

(Solomon, 1979, 9). The HHEC suggestion faced some rough sledding for several years. Solomon's posture was indicative of the way the testmaking community explained away this phenomenon, in addition, until 1981–82, to refusing to even publish the test score distributions by race and ethnicity.

These differences will not surprise anyone familiar with the inequalities in the social, economic, and occupational spheres of American life. Although many special educational programs have been developed at the federal, state, and local levels to repair the effects of educational and social neglect, these programs and other measures cannot change so quickly what years of malpractice have perpetuated. . . . Nevertheless, the argument is made that tests of academic ability are biased because they represent middle class culture. These tests do reflect skills and knowledge considered important in the mainstream culture, in many jobs, and in higher education. But the fact that a test mirrors the common culture is a poor basis for calling the test biased. (Solomon, 1979, 9–10)

Solomon's traditional arguments asserting purely and strictly educational and other forms of "disadvantage" notwithstanding, the more recent work of Durán and others who have shown that the tests are less predictive for subsequent Hispanic academic performance in college suggests that there is more (or rather less than mere mirrors of common culture in standardized tests: more error and less accuracy. A good place to review what many Hispanic researchers believe to be a prime source of test measurement error is the work of Mestre (1986a, 1986b, 1988) and Cocking and Mestre (1988). Mestre has shown that even in such supposed relatively culture-free test items involving solving mathematical problems, there are interactions between the language and the problem that can disproportionately affect minority students. Mestre concludes that: "Bilingual Hispanic students possessing the same level of mathematical and computational sophistication as their monolingual peers often solve word problems incorrectly. The pattern of error suggests the language deficiencies lead to misinterpretations of word problems; the resulting solutions may be incorrect, yet mathematically consistent with the student's interpretation of the problem statement" (Mestre, 1988, 202). Mestre has described this as the "Sorry, you have solved the wrong problem" effect, where a student's mathematics is sound, but his or her understanding of English (e.g., a double negative in Spanish is still a negative) leads to an incorrect setting up of the mathematical variables. It should be noted to the credit of both the ETS and the College Board that, at least to my knowledge, the expertise of researchers like Mestre, who in earlier decades might have seemed threatening to testmakers, is quickly introduced to the industry. Mestre is one of a group of con-

sultants who helped design *Algebridge* and he also serves on the College Board's SAT Committee. Developed jointly by the ETS and the College Board, this is a very promising program of diagnostic testing and instruction that helps all students, but particularly minority students, prepare for high school algebra. Naturally, *Algebridge* takes into account the phenomena that José Mestre has been researching for the past ten years that can take place among Hispanic youth during the mathematics and science educational and assessment processes.

The first break in the posture of the testmaking community as the result of the HHEC's 1979 criticisms and suggestions came with the GRE Board's decision in 1981 to publish its test score distributions by race and ethnicity. The College Board and the American College Testing Program followed suit quickly with respect to the SAT and ACT, providing these data in their 1982 technical reports. It is now possible to determine the three separate "within ethnic" scores for Hispanics, specifically for Chicanos/Mexican-Americans, for Puerto Ricans, and for other Hispanics. For example, the within ethnic score for Chicanos/Mexican-Americans can be used to determine how well a Chicano student did compared to all other Chicano students who took the standardized test in a specific year, and likewise for Puerto Ricans and for other Hispanics. However, these distributions are available in technical reports and are not widely reviewed, except through Project 1000 which makes the distributions available for optional use by the participating graduate institutions. Project 1000 funded, among others, by the Carnegie Corporation, the Pew Charitable Trusts, and the Alfred P. Sloan Foundation is an attempt to recruit an additional 1,000 U.S. Hispanic students into graduate school. The Project uses and is tracking the validity of an extensive number of nontraditional predictors of aptitude for graduate study. These include the within ethnic score, performance on the PAEG or the TOEFL when students have taken these examinations, and a variety of personal qualities and skills. The ETS, the GRE Board, the Graduate and Professional School Financial Aid Service, and the Council of Graduate Schools, as well as dozens of graduate schools and close to 200 primarily undergraduate institutions and other educational organizations are supporting or participating in this project.

It is clear that using the within ethnic score as a supplement to the customary score has great potential as a corrective for test content error (as documented by researchers such as Durán and Mestre) and therefore as a mechanism to enable large increases in the number of Hispanic students who may be admitted into college. One way to appreciate the underrepresentation of U.S. Hispanic students in graduate programs is by reviewing the numbers of Chicanos and Puerto Ricans taking the GRE, which in itself is a necessary precondition in order to be considered for admission to

virtually every selective graduate school in the United States. The 1988 ETS summary report indicates that only 2,226 Mexican-Americans/Chicanos and 1,661 Puerto Ricans actually took the GRE during the period in question. This is an exceedingly small number given the large Hispanic American population with baccalaureates who theoretically could take the exam. This figure contrasts mightily with 147,466 Whites who took the test during the same period.

Moreover, when we look at the range of scores that would be competitive for admission to selective graduate schools, the disproportion is accentuated to a deplorable level. Utilizing Table 62 of the 1988 ETS summary report, let us hypothetically consider a 550 on the GRE verbal section to be a competitive, although not outstanding score. A 550 represents approximately the sixty-second percentile for White, and in absolute numbers that represents approximately 56,000 White test-takers in 1988. But for Mexican-Americans and Puerto Ricans, respectively, that score of 550 represents approximately the eighty-fifth and the ninety-second percentile. Seen in absolute numbers, approximately 334 Mexican-Americans and approximately 133 Puerto Ricans competed with approximately 56,000 Whites with test scores of 550 or better in the 1986–87 academic year for the seats available in selective graduate schools. I have used the GRE verbal as a point of comparison for the sake of simplicity. Comparable results are obtained in comparing the GRE Quantitative or analytic. The fact that only about 467 Chicanos and Puerto Ricans obtained a 550 on the verbal (which itself is not an outstanding score, if we were to evaluate 650 or higher the number is reduced to about 130 Chicanos and Puerto Ricans combined) is cause for despair for ever increasing significantly the number of Hispanics who enter selective graduate schools if we are to rely on the usual test score interpretation and if additional measures and forms of evidence for validating the graduate school potential of U.S. Hispanics cannot be identified.

Merely the validation of the within ethnic group test scores has important implications for Hispanic access. If selective institutions were to look at, say, the top 30% of U.S. Hispanics as being within admission range, the pool of 1988 GRE test-takers to be so considered would have numbered about 667 Mexican-Americans and 500 Puerto Ricans, for a total potential pool of 1,167 applicants. That is possibly double what the effective pool with genuine admissibility was in 1988. Thus, permitting the top 30% of U.S. Hispanics with respect to their within ethnic GRE score to be closely reviewed for possible admission into graduate programs could more than double the Hispanic American cohort.

However, to actually achieve this potential it will be necessary, just as in the case of Durán's finding, to achieve wider understanding among col-

lege admission officers of the value of the within ethnic score as well as concrete utilization of it as a supplement in the admissions process.

The most productive of the research conferences before the establishment of the formal partnership between the HHEC, the College Board, and the ETS, was the Latino College Students Conference, jointly sponsored in 1983 by the ETS and the Institute for Higher Education Law and Governance, University of Houston. This conference was valuable in the process of producing the book, *Latino College Students* (Olivas, 1986), another volume that is the first, but hopefully not the last, of its kind. The book, primarily inspired by Olivas' energies and prowess as a researcher, also emerged partially out of the HHEC's involvement with the testing industry. It is a seminal work, establishing the outlines of a theoretical framework for research on Latino college students as well as a detailed treatment through a series of papers on three broad areas: the transition from high school to college; Hispanic student achievement; and economics and stratification. *Latino College Students* is the model on which the present volume is based, although the former is a much broader treatment of issues affecting Hispanic students than the more circumscribed *Assessment and Access*. *Latino College Students* features research studies by several people who have worked at ETS including, Richard P. Durán, Vilma Ortiz, and María Pennock-Román as well as by other important researchers who have worked in the area of assessment: José P. Mestre and Richard R. Verdugo. *Latino College Students* has been influential over the last five years in establishing a better understanding of the Latino college student experience and it has helped inspire additional, at times broad, coverage within the higher education community of topics relevant to our community. A notable example of such coverage was the special issue of *Change: The Magazine of Higher Learning* (1988) that contained a special report on Hispanics in the academy. It is worth noting that the special issue listed in a section entitled "programs that work" several which feature the use of assessment instruments in a key way including Options for Excellence, Project 1000, Hispanic Student Success Program, and the National Hispanic Scholar Award Program.

Books like *Latino College Students* and hopefully the present volume, because they are ground-breaking, have a special historical importance for the U.S. Hispanic community. Arturo Madrid, in his foreword to Olivas' book, expresses this fact most poignantly:

Our history, or that part of it within the American community, has all too frequently been one of inability to speak up, to speak out, or to speak at all on the issues that concern us. Specifically, our exclusion from the institutional life of American society and from the forums it provides

has kept us from being heard, even when our voices have been loud and our concerns compelling. Denied access to power, especially the power to define, we have had to suffer in silence the denial, distortion, and trivialization of our historical experience. . . Nowhere have we felt the burden of institutional oppression nor the weight of responsibility more heavily than in education. This supposedly enlightened institution is seemingly also one of the most retrogressive. The educational system continues to blame the victim for its own failures, rather than to adapt its methods and policies to the differing populations and changing circumstances with which it is constantly confronted. Thus, the 1983 Conference on Latino College Students was a signal event, as is this volume of essays issuing therefrom: signal in its focus; signal in its range; signal in its depth. (Madrid, 1986, ix)

And then there was Garfield High School! Jaime Escalante, Henry Gradillas, and Benjamín Jiménez. And Options for Excellence! J. Quentin Jones, Joe Arriaga, George H. Hanford, and Henry G. Cisneros. About Garfield High School, I am reminded of what the character of the French scientist said, played by the late Francois Truffaut in *Close Encounters of the Third Kind* of a no more miraculous event, the apparition of people from around the country to meet with benevolent extraterrestrials: "It is an event sociological!" This extraordinary effort by a charismatic, Bolivian mathematics teacher, combining both advanced instruction and a high payoff for students in the form of evaluation of that instruction through the College Board's Advanced Placement Program (administered by the ETS) has captured the hearts of our society. The events sociological were evoked in a widely viewed film, *Stand and Deliver*; the achievements of Escalante were reflected in the latest presidential debates (Bush referred to Escalante as one of the nation's heroes in a famous moment during the Bush-Dukakis debates); and projects around the country have been set up which seek to replicate some of what teachers, Escalante, Jiménez, and others, together with then-principal Henry Gradillas were able to accomplish at Garfield High School. These projects include two large scale demonstrations: the Hispanic Student Success Program, currently in San Antonio and beginning replication in Northern New Mexico and other sites as well; and Project PRIME (Project to Improve Minority Education) in the State of Arizona.

More than 95% of the Garfield student body is Latino. At least 80% of the students qualify for the federal free or reduced-price lunch program; in order to qualify the annual incomes of their families fall below \$15,000 for a family of four (Mathews, 1988, 2). There are several extraordinary elements that need to be pointed out about the Garfield phenomenon in this introduction to a book on promoting Hispanic access through assessment procedures.