



UNITED STATES DEPARTMENT OF EDUCATION
OFFICE OF THE ASSISTANT SECRETARY FOR CIVIL RIGHTS

June 21, 1999

Dear Colleague:

I am writing to thank you for your interest and input regarding our draft of *Non-discrimination in High-Stakes Testing: A Resource Guide*, and to confirm some of the foundations and objectives for this work. In addition, I want to let you know of our plans for further work on the Guide. We appreciate your willingness to assist in our effort to refine the existing draft of the Guide.

Overview

As you know, our objective is to provide educators and policymakers with a useful, practical tool that will assist in their planning and implementation of policies relating to the use of tests as conditions of conferring educational opportunities to students. As Secretary Riley said in his commemoration of the 45th anniversary of the *Brown v. Board of Education* decision, the courtroom is not the optimal place to make education policy. If, therefore, we can provide meaningful tools to our constituents in an area that is too frequently the subject of confusion—and, therefore, sometimes unnecessary controversy—we will have helped provide a stronger foundation for better educational decisions that serve all of our students. That is our goal.

Background

The U.S. Department of Education Office for Civil Rights enforces laws that prohibit discrimination on the basis of race, national origin, sex, disability and age by educational institutions that receive federal funds. Our work on the Guide is just one part of our effort to prevent problems of discrimination in the area of test use. Our educational stakeholders at all levels have come to us requesting advice and technical assistance in a variety of test use contexts, particularly as states and districts use tests as part of their standards based reforms. And, increasingly, we are addressing testing issues in a broader and more extensive array of complaints of discrimination filed with our office. These corresponding developments confirm the need to provide a useful resource that will capture legal, test measurement, and educational principles, references, and resources to assist educators and policymakers.

In our effort to develop such a resource, we have consulted with dozens of teacher, administrator, policymaker, business, advocacy, and test publisher organizations to solicit input and advice regarding the scope, framing and kinds of resources to include in the guide. Notably, we have contracted with the National Academy of Sciences Board on Testing and Assessment, which has independently reviewed and which will again assess this document to ensure that it comports with professional standards.

400 MARYLAND AVE., S.W. WASHINGTON, D.C. 20202

Our Mission is to Ensure Equal Access to Education and to Promote Educational Excellence throughout the Nation.

Dear Colleague
June 21, 1999
Page 2

Central principles

The draft of the Guide reflects and we believe that tests can be critical components of educational strategies designed to promote educational excellence for all students. Our aim is to ensure that tests are used appropriately. Indeed, our resolutions of claims of discrimination recognize most clearly that the solution to concerns regarding discrimination is *not* to eliminate the very tools that help provide a meaningful picture of the educational opportunities provided to students. In short, this Guide is not about eliminating tests; rather, the Guide should help promote decisions ensuring that tests are used in ways that are consistent with their design and purpose, and that all students are afforded the opportunity to achieve to high standards.

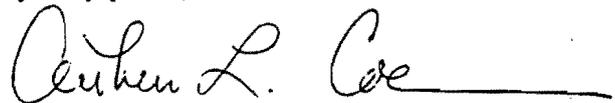
The Guide does not articulate new legal principles or test measurement standards. Citing thirty federal court decisions and over thirty test measurement standards, it provides a synthesis of often very dense information in one place, so that readers who are not as familiar with the area can have a readable, understandable resource.

Next steps

We have already received many useful and constructive suggestions regarding the Guide and are incorporating many of those ideas in our revisions. We have asked all with whom we have consulted to provide us with comments by the end of this month so that we may work during July and August to further refine and develop the Guide. Given the interest in this work, we are pleased to extend this deadline until July 16 for those who need additional time to comment. Our plan is to re-circulate the document in draft to the groups and individuals with whom we have consulted, just prior to our submission of the document to the National Academy of Sciences Board on Testing and Assessment, for its final review. Thereafter, we anticipate making a draft available to the public for review. We will publish a federal register notice to that effect and will have the revised draft posted on our web site. Once final comments are in, we will issue the Guide in final form.

Thank you again for your interest in this Guide, and do not hesitate to call me if you have questions or additional suggestions. You may also contact David Berkowitz at (202) 205-5557 or David_Berkowitz@ed.gov.

Very truly yours,



Arthur L. Coleman
Deputy Assistant Secretary

United States Commission on Civil Rights
Public Briefing
Prepared Statement of Arthur L. Coleman
Deputy Assistant Secretary
U.S. Department of Education Office for Civil Rights
June 18, 1999

Introduction

Thank you for the opportunity to represent the U.S. Department of Education and to discuss the existing legal, educational and test measurement principles that guide the work of the Department's Office for Civil Rights ["OCR"]. All individuals making important high-stakes decisions affecting the lives of students should understand the central principles on which so many in the education community agree. These points of agreement provide a very fertile common ground that should be the basis of our efforts to ensure that tests are used appropriately and that, as a consequence, accurate educational decisions are made—permitting all students achieve to their full potential.

Our goal is to provide some critical foundations for fulfilling the promise of the new civil right identified by U.S. Secretary of Education Richard Riley in his commemoration of the 45th anniversary of the *Brown v. Board of Education* decision. He said: "A quality education must be considered a key civil right for the 21st century." That point of consensus, along with the common ground that exists regarding good testing policies and practices, provides an important context for this discussion and affirms the need to move from the polarizing, either-or rhetoric that too frequently surrounds this issue. Secretary Riley has noted that too often in education, people are "choosing sides, not solutions." We seek to promote educationally sound solutions through our work related to the use of tests as foundations for high-stakes decisions affecting students.

When we talk about promoting the goal of achieving high standards education for all students, we mean it. The issue of testing in education should not be about favoring either standards or equity. Neither should the issue of testing in education be about blanketly favoring or opposing tests.

We believe that high standards for all means high standards for all. We believe that good test use practices advance high standards learning and equal opportunity—just as educationally inappropriate uses of tests do not. Tests are, as Secretary Riley has said, “important tools for educators to assess and assist students as they strive to meet high standards.” And, as they often provide a meaningful picture of educational opportunities provided to students, our goal is to preserve these critical measures of student performance just as we work to ensure that they are used appropriately.

We believe that the use of tests in education is an issue that should be the subject of informed and constructive dialogue. We welcome the opportunity today to continue our effort to advance a constructive discourse that can result in educational excellence for all students.

OCR's Work Regarding High-Stakes Decisions and the Use of Tests

The recently published *High Stakes: Testing for Tracking, Promotion, and Graduation* (National Research Council, Heubert and Hauser, eds., 1999) observed that the controversy surrounding the use of tests for high stakes decisions affecting students is often based upon misinformation and misperceptions about what tests are designed to do and, correspondingly, about good (and bad) test use practices. The application of federal non-discrimination laws to testing practices is subject to the same fate of misinformation—and sometimes, ill-informed commentary. To promote a better understanding and better practices regarding the use of tests for high stakes purposes, the

OCR is developing a resource guide for educators and policymakers. The guide will describe the existing non-discrimination, educational and test measurement foundations relating to the use of standardized tests that confer educational benefits to students.

Educational stakeholders at all levels have come to us requesting advice and technical assistance in a variety of test use contexts, particularly as states and districts increasingly use tests as part of their standards based reforms. And, we are addressing testing issues in a broader and more extensive array of complaints of discrimination filed with our office. These corresponding developments confirm the need to provide a useful resource that will capture legal and educational principles, references and resources to assist educators and policymakers.

We have worked with literally dozens of educator, parent, teacher, business, policymaker and testing groups and individuals to solicit input and advice regarding the scope, framing and kinds of resources to include in the guide. Notably, we have contracted with the National Academy of Sciences Board on Testing and Assessment, which has independently reviewed and which will again assess this resource to ensure that it comports with professional standards.

Perhaps the controversy that has surfaced in the wake of our extensive outreach was unavoidable. Nonetheless, we are perplexed that some of the central principles were considered by a few individuals to be novel and that the document was read by certain individuals to be something that it is not. Our effort should be understood in the clearest of terms. It is not our aim to establish standards or definitions of merit for educational institutions, as some have claimed. Nor is our effort to advocate for the elimination of standardized tests, such as the SAT. The draft of the guide reaffirms this point, in one instance stating: “[h]igh quality assessments can make high standards meaningful.” Instead, our effort is straightforward: to

explain existing legal and test measurement principles and to provide a collection of related resources—all in an effort to promote accurate decision making affecting the educational opportunities for all of our students.

Identifying and Debunking Some Myths

The misperceptions and erroneous understandings of test use principles that led to some of the controversy that has surfaced in recent weeks calls to mind the admonition by Nancy Cole, the president of the Educational Testing Service. Presciently, she has reminded us with regard to testing policies and practices that “we must acknowledge the myths that seem[] to make the issues simpler...even though acknowledging these myths makes finding solutions even more difficult.” (Cole, *Merit and Opportunity: Testing and Higher Education at the Vortex*, 1997). Given some of the inaccuracies regarding our ongoing work, it is an appropriate time to acknowledge—and rebut—some of the myths regarding federal non-discrimination standards and principles of sound test use.

Myth One: The goals of excellence and equity are irreconcilable.

FALSE, as a matter of law and policy.

The view that the goals of establishing standards (as in, for instance, establishing a standard of merit in college admissions) and complying with federal laws designed to ensure non-discrimination are inconsistent is, simply, erroneous. Indeed, if the federal courts teach us anything it is this: compliance with federal non-discrimination standards rests, in the first instance upon the school’s educational judgments, to which deference is appropriately given. Correspondingly, the ultimate question upon which the federal legal analysis affecting the use of high-stakes tests depends is one of educational sufficiency: is the test valid for the purposes used? Are the inferences derived from test scores, and the educational judgments based on those inferences, accurate and fair?

The educational foundations that guide any federal legal analysis suggest that policies promoting excellence can be and should be fully aligned with the promotion of equal opportunity for all students. For the hope of a high standards education for all students to become a reality for this generation of test-taking students, we must insist on high standards for tests that have consequences for students—just as we do for schools, teachers, and the students that they teach. As foundations for the judgments that shape the lives—and lifetimes—of students, these tests must be used in ways that accurately reflect educational standards and that do not inappropriately deny opportunities to students based on their race, national origin or sex.

Myth Two: Significant disparities in the test performance by subgroups of students indicate that the test discriminates illegally.

FALSE, as a matter of law.

Test results indicating that groups of students perform differently should be a cause for further inquiry and examination, with a focus upon the relevant educational programs and testing practices at issue. The existence of significant disparities does not mean, however, that the test illegally discriminates. Differences in test scores may result from a range of factors, including: lack of preparation; poor skills or knowledge; inadequate exposure to the material tested; poor motivation; or problems with the test itself.

The guarantee under federal law is for equal opportunity—not equal results. The legal non-discrimination inquiry regarding neutral practices (referred to by the courts as the “disparate impact” standard) illustrates this point: If the educational decisions based upon test scores reflect significant disparities in the kinds of educational benefits afforded to students based on race, national origin or gender, then ask more probing questions about what’s going on to ensure non-discriminatory,

educationally sound practices. This common sense framework is paralleled in the Code of Fair Testing Practices in Education. The Code provides, in relevant part:

“Test users should...[r]eview the performance of test takers of different races, gender, and ethnic backgrounds when samples of sufficient size are available [and e]valuate the extent to which performance differences may have been caused by inappropriate characteristics of the test.” [Joint Committee on Testing Practices, Code of Fair Testing Practices in Education (1988)]

The alignment of testing principles and legal standards could not, therefore, be clearer.

Myth Three: Test scores, alone, tell the whole story.
FALSE, as a matter of good educational practice.

Tests provide very valuable guidance in making educational judgments affecting students. Decisions such as college admissions decisions frequently—and appropriately—include consideration of test scores. The value that test results can provide when making educational decisions about students does not mean, however, that test scores should as a matter of good educational practice trump the need for thoughtful educational decision making. (Note here that federal non-discrimination laws do not preclude the prospect of the permissible use of a standardized test as a sole criterion where that test has been validated for such use.)

Moreover, a test's value as an educational tool is dependent upon its design, the context in which the test is administered, and the ultimate uses of the test. For example, the SAT may be valid as a tool to be used in a university's admissions decisions. At the same time, that same test is clearly inappropriate as a basis for making decisions about whether to promote a student from eleventh to twelfth grade in high school or whether to confer a passing grade in chemistry for the year.

Even when a test is used for the purposes consistent with its design, a test is one tool among many. Just as tests are not perfect barometers of learning, conclusions based on those test results are not always error free. Many variables can affect a student's test performance, including: the quality of the student's education; the student's skill, ability, or knowledge about a particular topic; preparation for the test; or what the student ate for breakfast on the day the test was administered. Does this mean that we should do away with tests? Absolutely not. What it does suggest is precisely what test measurement standards affirm: the importance of considering multiple and educationally appropriate measures when making life-defining decisions about students. The 1985 American Psychological Association Standards for Educational and Psychological Testing state, for instance: "In elementary and secondary education, a decision ...that will have a major impact on a test taker should not automatically be made on the basis of a single test score" (APA Standard 8.12). About this point, the guidance from test developers in higher education is instructive. Consider, for instance:

- Test uses "that should be avoided" include "using test scores as the sole basis for important decisions affecting the lives of individuals, when other information of equal or greater relevance and the resources for using such information are available." [The College Board, Guidelines on the Uses of College Board Test Scores and Related Data (1988)].
- The SAT works "very well in many different circumstances...[but] there are differences in how it works for different groups of students, for different types of educational programs, and for different institutions." [The College Board, Research Notes, RN-01 (June 1997)].

Conclusion

Ultimately, good educational practices—frequently reflected in test measurement standards—and federal case law highlight the importance of considering objective measures such as tests in appropriate ways when making decisions about students. In short, they affirm that not all tests are created equal and that tests should be used in ways that are valid for the particular purpose for which they are used.

This is the driving force behind the U.S. Department of Education Office for Civil Rights' continuing effort to provide assistance to policymakers and educators as we continue to enforce federal laws that prohibit discrimination against students. Rather than creating false and polarizing "win-lose" choices on this all-important set of issues, we need to, as Secretary Riley admonishes, "step back, lower our voices, truly listen to each other and search for common ground." That is our objective as we work to fulfill the promise of longstanding education goals and non-discrimination protections: high standards learning for all students.

Bradley, Bridget

From: Coleman, Arthur
Sent: Friday, June 11, 1999 10:12 AM
To: 'Peter_Rundlet@oa.eop.gov'; 'Martin_T@a1.eop.gov'; 'kmitchell@nas.edu'; 'mfeuer@nas.edu'; Bowers, Susan; Lewis, Cathy H; Lim, Jeanette; Slayton, Lester; Cantu, Norma V; Pierce, Raymond; Bradley, Bridget; Jovicich, Catherine; Wohl, Alexander
Subject: FW: NYT response to Thernstrom op ed going today...

FYI, we're submitting this to NYT today. Call if questions. (Bridget, Mike read a longer version yesterday and liked it; we had to cut to get to NYT designated length.) Art

-----Original Message-----

From: Coleman, Arthur
Sent: Friday, June 11, 1999 10:02 AM
To: Green, Julie; Frank, David
Cc: Murphey, Rodger; Lyon, Tom; Saunders, Kelly; Berkowitz, David; Fitch, Rebecca; Winnick, Steve; Craig, Susan; Kole, Adina
Subject: NYT response

Julie and David,

Here is the revised letter cut to about 290 words. Julie, since we were well under 300, I left in two segments that you can still choose to cut if you think advisable...[1] parenthetical in second line of third para ("and conclusion...") and [2] last line of third para ("The guarantee...)--both of which I think are impt.

Hate losing the SAT para, but no way to pare it down much shorter than it was.

It's yours...! Pls send a copy of whatever the final is so that we can provide to staff. Thanks for your help. Art



NYTThernstromltr3.doc

In her commentary, "Testing, the Easy Target," Abigail Thernstrom makes a number of erroneous conclusions regarding the U.S. Department of Education Office for Civil Rights (OCR) draft resource guide regarding test use.

Our effort is to help policymakers and educators understand the foundations for legal and educationally sound testing policies-to avoid controversy and litigation. With citations to over sixty federal court opinions and test measurement standards of the American Psychological Association, this is nothing new.

Ms. Thernstrom's equation of racial gaps in test performance with discrimination (and conclusion that tests are, therefore, "guilty until proved innocent") reflects a misreading of the draft, just as it reflects a gross misunderstanding of the settled legal principles upon which it is based. Differences in students' test scores may result from a range of factors, including, lack of preparation for or motivation to do well; poor skills or knowledge related to what is being measured; or problems with the use of the test itself. The guarantee under federal law is for equal opportunity-not equal results.

Thernstrom also accuses the U.S. Department of Education's Office for Civil Rights of seeking to ban all standardized testing. That outrageous assertion is squarely refuted by the draft, as well as by our numerous policy statements and resolutions of discrimination claims. Our guiding principle is that anti-discrimination standards give substantial deference to sound educational judgments. As a result, tests that are used in educationally appropriate ways and that are valid for the purposes used are very important instruments to help educators do their job. Our resolutions of discrimination claims recognize clearly that the solution to many such problems is *not* to eliminate the very tools that help provide a meaningful picture of the educational opportunities provided to students.

Arthur L. Coleman
Deputy Assistant Secretary for Civil Rights
U.S. Department of Education

In her commentary, "Testing, the Easy Target," Abigail Thernstrom makes a number of erroneous conclusions regarding the U.S. Department of Education Office for Civil Rights (OCR) draft resource guide regarding test use.

Our effort is to help policymakers and educators understand the foundations for legal and educationally sound testing policies—to avoid controversy and litigation. With citations to over sixty federal court opinions and test measurement standards of the American Psychological Association, this is nothing new.

Ms. Thernstrom's equation of racial gaps in test performance with discrimination (and conclusion that tests are, therefore, "guilty until proved innocent") reflects a misreading of the draft, just as it reflects a gross misunderstanding of the settled legal principles upon which it is based. Differences in students' test scores may result from a range of factors, including, lack of preparation for or motivation to do well; poor skills or knowledge related to what is being measured; or problems with the use of the test itself. The guarantee under federal law is for equal opportunity—not equal results.

Thernstrom also accuses the U.S. Department of Education's Office for Civil Rights of seeking to ban all standardized testing. That outrageous assertion is squarely refuted by the draft, as well as by our numerous policy statements and resolutions of discrimination claims. Our guiding principle is that anti-discrimination standards give substantial deference to sound educational judgments. As a result, tests that are used in educationally appropriate ways and that are valid for the purposes used are very important instruments to help educators do their job. Our resolutions of discrimination claims recognize clearly that the solution to many such problems is *not* to eliminate the very tools that help provide a meaningful picture of the educational opportunities provided to students.

Arthur L. Coleman
Deputy Assistant Secretary for Civil Rights
U.S. Department of Education



"Coleman, Arthur" <Arthur_Coleman@ed.gov>
10/21/99 03:38:55 PM

Record Type: Record

To: Peter Rundlet/WHO/EOP, Andy Rotherham/OPD/EOP

cc:

Subject: FW: Testing resource guide...strategies for rollout, next steps and schedule

I meant to copy you on this when I sent it on Tuesday. Happy to discuss if you want to talk. Art

> -----Original Message-----

> From: Coleman, Arthur

> Sent: Tuesday, October 19, 1999 12:10 PM

> To: Frost, Susan; Smith, Mike; Liu, Goodwin; Johnson, Judith; Cohen,

> Mike; Jovicich, Catherine; Rairdin, Kae; Fleming, Scott; Frank, David;

> Murphey, Rodger; Winston, Judith; Winnick, Steve; Kole, Adina; McLaughlin,

> Maureen

> Cc: Cantu, Norma V; Thornton, Leslie; Wohl, Alexander

> Subject: Testing resource guide...strategies for rollout, next steps

> and schedule

> Importance: High

>

> Friends,

>

> This is a recap of where we are regarding the revisions to the resource

> guide re testing, and plans for moving forward. (A timetable is located at

> the bottom of this message.) Please feel free to email or call with

> reactions and input. Next week, we'll begin extending invitations for

> some of the early November conversations that are discussed below.

>

> We've received extensive input regarding the resource guide from the 40+

> (education, business, parent, testing, advocacy) groups with whom we've

> conferred. We've almost completed our efforts to incorporate much of the

> feedback in terms of scope, clarity, etc. We plan to circulate within the

> Department a draft of the revised guide next week, allowing one week for

> review and comment. We'll then incorporate comments as appropriate and be

> in a position to circulate among our stakeholders toward the end of

> November, prior to forwarding the draft to the National Academy of

> Sciences Board on Testing and Assessment for its final review.

>

> In the meantime, we plan to schedule a series of meetings with groups

> during the first week in November to provide an overview of the comments

> received, and to discuss the changes that can be expected in the draft.

> The new draft will NOT be circulated during the meetings during the first

> week of November; rather, these efforts will be designed to provide a

> 'sneak preview' -- to gauge reaction, conceptually, and to prepare the

> groups for what they'll soon be seeing. (We've had offers of assistance

- > by ACE and K-12 test publishers to facilitate convening of such meetings.)
- > We expect to provide a letter to summarize the presentations to minimize
- > the risk of mischaracterization of the meetings. Through this process we
- > can identify likely trouble spots and be in a position to strategize
- > before distribution of the guide.
- >
- > The NAS Board on Testing and Assessment's Forum on Educational Excellence
- > and Testing Equity (funded by the Dept., with Dept. representation by
- > Judith Johnson and me), with broad representation from the K-12, academic,
- > and testing world, has agreed to meet on November 9 in Washington to serve
- > as a focus group for the guide. We will meet all day, receive feedback
- > from this group, and refine the document to address issues prior to the
- > re-release to educational stakeholders later in the month.
- >
- > I plan to follow up with Susan re group meetings, and Kae re possible
- > meetings with Congressional staffers.

> Thanks.
> Art

> Week of	Event
> October 25	Circulation within ED and WH contacts
> November 1	Meeting with external groups pre-release of the
> guide	Editing, based on ED comments
> November 8	NOVEMBER 9: Focus group of NAS Forum members in DC
> November 15	Editing
> November 22 and 29	Release of Guide to stakeholders
> December 15	Guide to BoTA, following revisions based on external
> feedback	
>	
>	

External Stakeholder Groups OCR Met With on Testing Guide

- post sec
 - K-12
 - civil rights
 - test pub/APA
-
- business/parent

Group Type	Organization Name	Contact Person	Address	Phone #	E-mail	Fax #
Board on Testing and Assessment	Attorney At Law	Taylor, William L.	2000 M Street, NW Suite 400 Washington, DC 20036	202-659-5565		202-223-5302
Board on Testing and Assessment	Board on Testing and Assessment	Feuer, Ph.D., Michael J.	National Research Council 2101 Constitution Avenue, NW Washington, DC 20418	202-334-3087		202-334-3584
Board on Testing and Assessment	Board on Testing and Assessment	Mitchell, Karen	2101 Constitution Avenue, N.W. Washington, D.C. 20418	202-334-3407		202-334-3584
Board on Testing and Assessment	Cornell University	Neisser, Ulric	Ithaca, New York 14853	607-255-6355		607-255-8433
Board on Testing and Assessment	Harvard Law School	Edley, Jr., Christopher F.	Giswold Hall 405 Cambridge, MA 02138	617-495-4614		617-496-5156
Board on Testing and Assessment	Harvard University	Ferguson, Ronald	John F. Kennedy Street Cambridge, MA 02138	617-495-1104		508-230-7291

Group Type	Organization Name	Contact Person	Address	Phone #	E-mail	Fax #
Board on Testing and Assessment	Texas Education Agency	Smisko, Ph.D., Ann	1701 North Congress Ave Austin, TX 78701	512-463-9087		512-475-3667
Board on Testing and Assessment	University of Illinois, Urbana-Champaign	Trent, William T.	1310 South Sixth Street Champaign, Illinois 61820	217-333-6153		217-244-4121
Business	Council on Competitiveness	Berman, Ph.D., Marshall	1401 H Street, NW Suite 650 Washington, DC 20005	2026824292		2026825150
Business	National Alliance of Business	Goldberg, Milton	1201 New York Ave, NW Suite 700 Washington, DC 20005-6143	202-289-2944		202-289-2908
Business	National Alliance of Business	Guidera, Aimee Rogstad	1201 New York Ave, NW Suite 700 Washington, DC 20005-6143	202-289-2901		202-289-2908
Business	National Alliance of Business	Lindsey, Thomas A.	1201 New York Ave, NW Suite 700 Washington, DC 200056143	202-289-2932		202-289-1303
Business	National Center on Education and the Economy	Clark, Charles S.	Suite 750 700 11th Street, NW Washington, DC 20001	2027833668, ext. 127		202-783-3672

Group Type	Organization Name	Contact Person	Address	Phone #	E-mail	Fax #
Business	Partners in Education	St. Amand, Barbara E.	901 North Pitt Street, Suite 320 Alexandria, VA 22314- 1536	703-836- 4880		703-836- 6941
Business	Siemens Corporation	Tobin, John P.	1301 Avenue of the Americas New York, NY 10019	212-258- 4046		212-258- 4019
Business	U.S. Chamber of Commerce	Bober, Chris	1516 Street, NW Washington, DC 20062- 2000	202-463- 5548		
Civil Rights	ASPIRA Association, Inc.	Crespo, Hilda	1444 I Street, NW Suite 800 Washington, DC 20005	202-835- 3600		202-835- 3613
Civil Rights	Center for Law and Education	Stoneman, Christine	1875 Connecticut Ave, NW Suite 510 Washington, DC 20009	202-986- 3000		202-986- 6648
Civil Rights	Center for Law and Education	Weckstein, Paul	1875 Connecticut Ave, NW Suite 510 Washington, DC 20009	202-986- 3000		
Civil Rights	Leadership Conference on Civil Rights	Henderson, Wade J.	1629 K Street, NW Suite 1010 Washington, DC 20006	202-466- 3311		
Civil Rights	MCA Enterprises, Inc.	Lenehan-Razzuri, Moira	1156 15th Street, NW Suite 220 Washington, DC 20005	202-862- 8514		202-862- 9814

Group Type	Organization Name	Contact Person	Address	Phone #	E-mail	Fax #
Civil Rights	MCA Enterprises, Inc.	McAdam, Marilyn	1156 15th Street, NW Suite 220 Washington, DC 20005	202-862-8514		202-862-9814
Civil Rights	Mexican American Legal Defense & Educational Fund	Hernandez, Antonia	1518 K Street, NW Suite 410 Washington, DC 20005	202-628-4074		202-393-4206
Civil Rights	Multicultural Education, Training & Advocacy	Rice, Roger L.	240A Elm Street, Suite 22 Somerville, Massachusetts 02144			
Civil Rights	NAACP - Legal Defense and Educational Fund, Inc.	Shaw, Theodore M.	99 Hudson Street, Suite 1600 New York, NY 10013	212-219-1900		212-226-7592
Civil Rights	NAACP Legal Defense and Education Fund	Byrd, Janelle	1444 I Street, NW 10th Floor Washington, DC 20005	202-682-1300		202-682-1312
Civil Rights	National Association for Bilingual Education	Loera, Patricia E.	1220 L Street, NW Suite 605 Washington, DC 20005-4018	202-898-1829		202-789-2866
Civil Rights	National Council of La Raza	Rodriguez, Roberto	1111 19th Street, NW Suite 1000 Washington, DC 20036	202-776-1787		202-776-1792

Group Type	Organization Name	Contact Person	Address	Phone #	E-mail	Fax #
Civil Rights	LULAC	Collantes, Cesar	Nassif Building, Northwest Plaza Room PI 100 400 7th Street, SW Washington, DC 20590	202-267-3159		
ESE	American Association of School Administrators	Hunter, Dr. Bruce	1801 N. Moore Street Arlington, VA	703-528-7000		
ESE	Council of Chief State School Officers	Martin, Wayne	1 Massachusetts Avenue, NW Suite 700 Washington, DC 20001-1431	202-336-7010		202-789-1792
ESE	Council of the Great City Schools	Lewis, Sharon	1301 Pennsylvania Ave, NW Suite 702 Washington, DC 20004	202-393-2427		
ESE	Council of the Great City Schools	Simmering, Jeff	1301 Pennsylvania Ave, NW Suite 702 Washington, DC 20004	202-393-2427		
ESE	Council of the Great City Schools	Uro, Gabriela	1301 Pennsylvania Ave, NW Suite 702 Washington, DC 20004	202-393-2427		

Group Type	Organization Name	Contact Person	Address	Phone #	E-mail	Fax #
ESE	National Association of Elementary School Principals	Sava, Samuel	1615 Duke Street Alexandria, VA 22314	703-684-3345		
ESE	National Association of Secondary School Principals	Tirozzi, Gerald	1904 Association Drive Reston, VA 20191	703-860-0200		
ESE	National School Boards Association	Underwood, Julie	1680 Duke Street Alexandria, VA 22314	703-838-6722		703-683-7590
ESE	National School Boards Association	Villani, Joseph S.	1680 Duke Street Alexandria, VA 22314-3493	703-838-6799		703-548-5516
ESE	San Antonio Schools Develop. Foundation	Lam, Diana	4002 Shady Oak Drive San Antonio, TX 78229	210-377-2902		210-377-2981
ESE Test Publisher	MCA Enterprises, Inc.	Snowwhite, Larry S.	1156 15th Street, NW Suite 20 Washington, DC 20005	202-862-8514		202-862-9814
ESE Test Publisher (Law Firm)	Hale and Dorr LLP	Urwitz, Jay P.	1455 Pennsylvania Ave, NW Washington, DC 20004	202-942-8400		202-942-8484
ESE Test Publisher (law firm)	Taylor Thiemann & Aitken	Thiemann, Alan	908 King Street Alexandria, VA 22314	7038369400		7038369409

Group Type	Organization Name	Contact Person	Address	Phone #	E-mail	Fax #
Judiciary	U.S. Court of Appeals for the DC Circuit	Goodwin, Liu	Chambers of the Honorable David S. Tatel 3818 U.S. Courthouse Washington, DC 20001			
Law Firm	Jorden Burt	Dillon, Alanna M.	Suite 400 East 1025 Thomas Jefferson St, NW Washington, DC 200070805	202-965-8115		202-965-8100
Parent/Teacher	National PTA	Oakes, Mary Beth	1090 Vermont Ave, NW Suite 1200 Washington, DC 20005	202-289-6790		
Parent/Teacher	National PTA	Oakes, Maribeth	1090 Vermont Avenue, NW Suite 1200 Washington, DC 20005-4905	202-289-6790, ext. 203		202-289-6791
Parent/Teacher	The National PTA	Igo, Shirley	713 El Paso Street Plainview, TX 79072-6207	806-296-9202		806-296-2549
Policy/Legislative	National Conference of State Legislatures	Shreve, David	444 N. Capitol Street, NW Suite 515 Washington, DC 20001	202-624-8187		
Policy/Legislative	National Governors' Association	Sullivan, Patricia	Hall of the States 444 North Capitol Street, Suite 267 Washington, DC 20001-1512	202-624-7723		202-624-5313

Group Type	Organization Name	Contact Person	Address	Phone #	E-mail	Fax #
PostSec	American Assoc. of Community Colleges	Kee, Arnold M.	1 Dupont Circle, NW Suite 410 Washington, DC 20036	202-728-0200 ext. 26		202-833-2467
PostSec	American Council on Education	Baer, Michael	1 Dupont Circle Suite 800 Washington, DC 20036	202-939-9551		
PostSec	American Council on Education	Barr, Vickie	1 Dupont Circle Suite 800 Washington, DC 20036	202-939-9327		
PostSec	American Council on Education	Garza, Hector	1 Dupont Circle Suite 800 Washington, DC 20036	202-939-9390		
PostSec	American Council on Education	Ikenberry, Stanley O.	1 Dupont Circle, NW Suite 800 Washington, DC 200361193	202-939-9310		202-659-2212
PostSec	American Council on Education	King, Jaci	One Dupont Circle, NW Suite 800 Washington, DC 20036-1193	202-939-9310		202-659-2212
PostSec	American Council on Education	Steinback, Shelley	One Dupont Circle, NW Suite 800 Washington, DC 20005	202-939-9310		202-659-2212
PostSec	American Council on Education	Wilds, Deborah	1 Dupont Circle Suite 800 Washington, DC 20036	202-939-9451		

Group Type	Organization Name	Contact Person	Address	Phone #	E-mail	Fax #
PostSec	American Council on Education/Cent er for Adult Learning	Robinson, Susan	1 Dupont Circle Suite 250 Washington, DC 20036	202-939-9486		
PostSec	American Council on Education/GED TS	Schaefer, Lynn	1 Dupont Circle Suite 250 Washington, DC 20036	202-939-9701		
PostSec	American Council on Education/GED TS	Skaggs, Gary	1 Dupont Circle Suite 250 Washington, DC 20036	540-890-0952		
PostSec	American Council on Education/GED TS/	Auchter, Joan E.	1 Dupont Circle Suite 250 Washington, DC 20036	202-939-9435		
PostSec	Boston College	Pullin, Diana C.	Campion Hall 100 Chestnut Hill, MA 02167	617-552-8407		617-552-0812
PostSec	Boston University	Loury, Glenn C.	704 Commonwealth Ave Boston, MA 02212	617-353-5852		617-353-4072
PostSec	Brandeis University	Regunberg, Michael	415 South Street - MS136 Waltham, MA 02454-9110	781-736-4212		781-736-8699
PostSec	Brandeis University	Reis, Arthur H.	Mail Services MS-134 PO Box 9110 415 South Street Waltham, MA 02454-9110	781-736-2106		781-736-3457

Group Type	Organization Name	Contact Person	Address	Phone #	E-mail	Fax #
PostSec	Brown University	Quinn, Marisa	Box 1920 Providence, RI 02912-1920	401-863-1632		401-863-3189
PostSec	California Institute of Technology	Burke, Esq., April	Lewis-Burke Associates 1233 20th Street, NW, Suite 610 Washington, DC 20036	202-466-4111	April@lewis-burke.org	
PostSec	Case Western Reserve University	Dziak, Adrienne L.	14 Adelbert Hall 10900 Euclid Avenue Cleveland, OH 44106-7024		Ald8@po.cwru.edu	
PostSec	Columbia University	Smith, Ellen S.	535 West 116th Street 301 Low Memorial Library Mail Code 4319 New York, NY 10027	212-854-3394		212-666-1952
PostSec	Cornell University	Johnson, Stephen Philip	Office of Government Affairs 114 Day Hall Ithaca, NY 14853-2801	607-255-4347		607-255-5396
PostSec	Duke University	Vick, Paul	PO Box 90433 Durham, NC 27706		Pvick@duke.edu	
PostSec	Harvard University	Casey, Kevin	2 Garden Street Cambridge, MA 02138		Kevin_casey@harvard.edu	
PostSec	Indiana University	Walker, George E.	Bryan Hall 104 Bloomington, IN 47405-1220	812-855-6153	Walkerg@indiana.edu	812-855-6396

Group Type	Organization Name	Contact Person	Address	Phone #	E-mail	Fax #
PostSec	Information for Life's Transitions	Minchew, Daniel	One Dupont Circle, NW Suite 800 Washington, DC 20005	202-939-9310		202-659-2212
PostSec	Massachusetts Institute of Technology	Crowley, John C.	820 First Street, NW Suite 410 Washington, DC 20002	202-789-1828	Jcrowley@mit.edu	
PostSec	Massachusetts Institute of Technology	Parravano, Paul	77 Massachusetts Avenue Room 5-208 Cambridge, MA 02139-4307	617-253-3873		617-258-6096
PostSec	Massachusetts Institute of Technology	Smith, Tobin L.	820 First Street, NE Suite 410 Washington, DC 20002	202-789-1828	Toby@mit.edu	
PostSec	National Collegiate Athletic Association	Dixon, Doris	Federal Relations Office 1 Dupont Circle, NW Suite 310 Washington, DC 20036	202-293-3050		
PostSec	New York State Dept of Education	Kadamus, James A.	89 Washington Ave Rm 875EBA Albany, NY 12234	518-474-5915		518-486-2233
PostSec	New York University	Brown, Lynne P.	Elmer Holmes Bobst Library 70 Washington Square South, Room 1228 New York, NY 10012-1091		Lynne.brown@nyu.edu	

Group Type	Organization Name	Contact Person	Address	Phone #	E-mail	Fax #
PostSec	New York University	Thompson, Marilyn Berry	Jorden Burt Boros Chichetti Berenson & Johnson, LLP 1025 Thomas Jefferson Street, NW Suite 400 East Washington, DC 20007	202-965-8100	Mbt@wdc.jordenusa.com	
PostSec	Northwestern University	Bridges, Steven	2-640 Rebecca Crown Center 633 Clark Street Evanston, IL 60208-1117		sbridges@nwu.edu	
PostSec	Princeton University	Wells, Nan Shelby	444 North Capitol Street, NW Suite 728 Washington, DC 20001	202-639-8420		202-639-8423
PostSec	Purdue University	Greenkorn, Robert A.	1283 CHME Building, Room 101 West Lafayette, IN 47907-1283	765-494-4051		765-494-0805
PostSec	Stanford University	Hakuta, Kenji	CERAS 207S Stanford, CA 94305	650-725-7454		650-723-7578
PostSec	Stanford University	Levin, Henry M.	CERAS Buildin, Rm 109 Stanford, CA 94305-3084	650-723-0840		650-725-6140
PostSec	Teachers College, Columbia University	Heubert, Jay P.	525 W. 120th Street, Box 157 New York, NY 10027	212-678-3775		718-796-2783

Group Type	Organization Name	Contact Person	Address	Phone #	E-mail	Fax #
PostSec	The Johns Hopkins University	Etten, Thomas A.	3400 North Charles Street Baltimore, MD 21218		Tetten@jhuadig.admin.jhu.edu	
PostSec	The Ohio State University	Stoddard, Richard S.	205 Bricker Hall 190 North Oval Mall Columbus, OH 43210-1321	614-292-6245		614-292-1231
PostSec	The University of Arizona	McGonagill, Margaret A.	Administration Building, Room 116 PO Box 210066 Tucson, AZ 85721	520-621-3108		520-626-3390
PostSec	The University of North Carolina at Chapel Hill	Samors, Robert J.	1300 Pennsylvania Avenue, NW North Tower, Suite 370 Washington, DC 20004	202-289-8433		202-331-4119
PostSec	Tulane University	Courtney, Sharon	327 Gibson Hall New Orleans, LA 70118-5698		Sharonc@mailhost.tcs.tulane.edu	
PostSec	University at Buffalo - State Univ. of New York	Penksa, Janet	520 Capen Hall Buffalo, NY 14260	716-645-2922		716-645-2933

Group Type	Organization Name	Contact Person	Address	Phone #	E-mail	Fax #
PostSec	University of California	Gourlay, Diana	Federal Governmental Relations 1523 New Hampshire Avenue, NW Washington, DC 20036	202-588-0002	Dgourlay@popserver.ucop.edu	
PostSec	University of California	Peterson, Matt	Federal Governmental Relations 1523 New Hampshire Avenue, NW Washington, DC 20036	202-588-0002	Matt.peterson.ucop.edu	
PostSec	University of California, Berkeley	Suduth, Esq., A. Scott	Federal Government Relations 1523 New Hampshire Avenue, NW Washington, DC 20036	202-588-0066		202-785-2669
PostSec	University of California, LA	Marthias, Sarah Avellar	1523 New Hampshire Avenue, NW Washington, DC 20036	202-588-0002	Sarah.marthias@ucop.edu	
PostSec	University of California, Los Angeles	Kallick, Deborah B.	10920 Wilshire Boulevard Suite 1500 Los Angeles, CA 90024-6517	310-794-6806		310-794-6827
PostSec	University of California, San Diego	Furtek, Edward	9500 Gilman Drive Mailcode 0993 La Jolla, CA 92093-0993		Efurtek@ucsd.edu	

Group Type	Organization Name	Contact Person	Address	Phone #	E-mail	Fax #
PostSec	University of Chicago	Hedges, Larry V.	Department of Education 5835 South Kimbark Avenue Chicago, IL 60637	773-702-1589		773-702-0248
PostSec	University of Illinois, Chicago	Walberg, Herbert J.	3123 ECSW Chicago, IL 60612	312-951-6044		312-951-8547
PostSec	University of Iowa	Willard, Derek H.	Office of Vice President for Research 201 Gilmore Hall Iowa City, IA 52242	319-335-3714		319-335-2104
PostSec	University of Michigan	Bank, Cindy A.	499 South Capitol Street, SW Suite 501 Washington, DC 20003	202-554-0578	Bankc@u mich.edu	
PostSec	University of Michigan	Butts, Thomas A.	499 South Capitol Street, SW Suite 501 Washington, DC 20003	202-554-0578	Tombutts @umich. edu	
PostSec	University of Minnesota, Twin Cities	Peterson, Donna	6 Morrill Hall 100 Church Street, SE Minneapolis, MN 55455-0110	612-624-9022		612-624-6369
PostSec	University of North Carolina, Greensboro	Wightman, Linda F.	217 Curry Guilding Spring Garden road Greensboro, NC 27412-5001	336-334-3045		336-334-4120

Group Type	Organization Name	Contact Person	Address	Phone #	E-mail	Fax #
PostSec	University of Oregon	Martin, Patricia A.	202 Johnson Hall Eugene, OR 94703-1260		Pmartin@oregon.uoregon.edu	
PostSec	University of Pennsylvania	Murphy, Micheline	133 South 36th Street Suite 501 Philadelphia, PA 19104		Emurphy@pobox.upenn.edu	
PostSec	University of Rochester	Pryor, Karen C.	University of Rochester Medical Center PO Box 601 Rochester, NY 14627		Karen_pryor@urmc.rochester.edu	
PostSec	University of Washington	Perry, Barbara F.	444 North Capitol Street Suite 418 Washington, DC 20001	202-624-1420		202-624-1429
PostSec	University of Wisconsin System	Boatright, Kevin J.	1762 Van Hise Hall 1220 Linden Drive Madison, WI 53706		Kboatright@ccmail.uwsa.edu	
PostSec	Vanderbilt University	Gardner, Midge	1112 16th Street, NW Suite 500 Washington, DC 20036	202-785-2980		202-785-3948
PostSec	Washington University in St. Louis	Lokken, Pamela	Campus Box 1191 One Brookings Drive Saint Louis, MO 63130	314-935-5616		314-935-5188
PostSec	Yale University	Jacob, Richard	PO Box 208255 2 Whitney Avenue New Haven, CT 06520-8255	203-432-4949		203-432-7960

Group Type	Organization Name	Contact Person	Address	Phone #	E-mail	Fax #
PSE Test Publisher	ACT	Ferguson, Richard L.	2201 North Dodge Street Iowa City, IA 52243-0168	319-337-1079		319-337-1059
PSE Test Publisher	ACT	Petersen, Nancy	2201 North Dodge Street Iowa City, IA 52243-0168	319-337-1067		319-339-3000
PSE Test Publisher	Educational Testing Service	Carnevale, Tony	Rosedale Road Princeton, NJ 08541	202-659-8000		
PSE Test Publisher	Educational Testing Service	Cole, Nancy	Rosedale Road Princeton, NJ 08541	609-734-1010		609-921-0235
PSE Test Publisher	Educational Testing Service	Engel, Penny	Rosedale Road Princeton, NJ 08541	609-734-1010		609-921-0235
PSE Test Publisher	Educational Testing Service	Johnson, Eugene G.	Statistics and Psychometric Research Rosedale Road Princeton, NJ 08541	609-734-5305		609-734-5420
PSE Test Publisher	Educational Testing Service	Von Mayrhauser, Stanford	Rosedale Road Princeton, NJ 08541	609-734-5070		
PSE Test Publisher	Law School Admission Council	Schrieber, Stephen T.	661 Penn Street Box 40 Newton, PA 18940-0040	215-986-1184		215-986-1169
PSE Test Publisher	Law School Admission Council	Van Tol, Joan	PO Box 40 Newtown, PA 18940	215-968-1137		
PSE Test Publisher	The College Board	Camara, Wayne J.	45 Columbus Avenue New York, NY 10023-6992	212-713-8069		212-649-8427
PSE Test Publisher	The College Board	Childers, John	45 Columbus Avenue New York, NY 10023-6992	202-822-5907		202-822-5920

Group Type	Organization Name	Contact Person	Address	Phone #	E-mail	Fax #
PSE Test Publisher	The College Board	Clark, Constance R.	45 Columbus Ave New York, NY 10023-6992	202-822-5900		
PSE Test Publisher	The College Board	Stewart, Donald M.	45 Columbus Avenue New York, NY 10023-6992	212-713-8000		
PSE Test Publisher (Law Firm)	Wilmer, Cutler & Pickering (For ETS)	Berman, Bruce	2445 M Street, NW Washington, DC 20037-1420	202-663-6173		202-663-6363
PSE Test Publisher (Law Firm)	Fulbright & Jaworski (for ACT & GED)	Burgoyne, Bob	801 Penn Ave, NW Washington, DC 20004	202-662-4513		202-662-4643
PSE Test Publisher (Law Firm)	Sullivan & Cromwell (for College Board)	Graham, Jr., Phillip L.	125 Broad Street New York, NY 100042498	212-558-3842		212-558-3588
Teachers	American Federation of Teachers	Rosenberg, Bella H.	555 New Jersey Ave, NW Washington, DC 20001	202-879-4562		202-393-7481
Teachers	National Education Association	Teasley, Mary Elizabeth	1201 16th Street, NW Washington, DC 20036-3290	202-822-7321		202-832-7741
Test Measurement	American Educational Research Association	Russell, William J.	750 First Street, NE Washington, DC 20002-4242	202-336-6000		
Test Measurement	American Psychological Association	Roberts-Fox, Heather	750 1st Street, NE Washington, DC 20002	202-336-6000		202-336-5953

TEST USE AND CIVIL RIGHTS

Why Is Test Use Important?

Tests continue to be used in many ways to measure competence in today's competitive world. From elementary through graduate school, test results often serve as a basis for decision-making that affects our youth. The progress of young schoolchildren increasingly is measured by test scores. Efforts to promote excellence in education are frequently associated with the use of high-stakes tests — tests whose results are used to make placement, promotion and graduation decisions, for instance. To best understand tests, all of us — parents, students, teachers, school administrators and policymakers — should understand the kinds of tests used and why and how they are used.

The issue of nondiscrimination in testing and assessment is properly viewed as consistent with standards-based reforms -- the cornerstone of many of the U.S. Department of Education's initiatives. The U.S. Department of Education is committed to the support of high standards and challenging assessments for all students. Nondiscrimination in testing and assessment is essential to ensuring that equal opportunities for educational excellence are provided regardless of race, national origin, or sex. All students need an educational system which both expects high performance and offers real and meaningful educational opportunities. It is critical that high standards for academic achievement be coupled with the necessary instruction and support that help students reach those standards - as determined by valid and reliable assessments.

The improper use of high-stakes tests can violate civil rights laws that prohibit discrimination against students on the basis of their race, national origin or sex. Any use of a high-stakes test must be considered in the context of the educational interests at issue, constitutional guarantees and civil rights laws.

This pamphlet describes civil rights requirements that apply to high-stakes tests. Specific examples of test use and civil rights requirements are discussed throughout this pamphlet and are also included in the questions and answers section at its end. The constitutional requirements are outlined in one question and answer below. For a more complete discussion of the legal principles related to test use and civil rights, please see OCR's *Nondiscrimination In High-Stakes Testing: A Resource Guide*.

What Tests Have Consequences For An Individual Student?

Most students take tests that are prepared by their teachers and designed to measure knowledge of topics covered in the classroom. For younger students, a weekly spelling test would be an example. For older students, a final exam on American government given at the end of the semester is an example. In addition, there are some tests that are given on a large scale to measure the performance of an entire school or school district. In many such cases, individual student scores are not reported to the school, student or parent. Instead, only group scores are reported for the school or for the school district.

Both classroom tests and broad school assessments are important. The focus of this pamphlet is on another type of test: those that are generally given on a state-wide or district-

wide basis and that are used to make educational decisions that have very important consequences for an individual student. Examples of these decisions are:

- whether or not students will be placed in gifted and talented programs;
- whether or not students will be promoted to the next grade or permitted to graduate; and
- whether or not students will be offered such benefits or opportunities as admissions or scholarships to specific colleges and universities, or to vocational education programs.

In cases like these, tests are used by schools to make major decisions about a student's educational future. Because these tests have important consequences for students, they commonly are called high-stakes tests.

Federal laws prohibit discrimination against students on the basis of race, national origin or sex in testing. This pamphlet outlines the relevant civil rights laws and the legal standards, along with frequently asked questions and answers regarding test use. (In addition, although beyond the general scope of this pamphlet, Federal law also prohibits discrimination on the basis of disability in testing and some of the basic requirements in this area are outlined in the question and answer section.)

Federal Civil Rights Legal Standards Regarding High-Stakes Test Use

Different Treatment

Discrimination against students on the basis of race, national origin or sex can occur in the use of high-stakes tests. One form of discrimination is called different treatment. This occurs when students are treated differently solely because of their race, national origin, or sex in terms of how a test is given or how its results are used, absent an appropriate legal justification (i.e., to remedy past illegal discrimination). One example of different treatment would be if a school district uses test scores to place girls in advanced placement math class only if they achieve higher test scores than boys placed in the same class. Another example would be a school district that puts minority group students in a special education program for mentally retarded students based on their scores on an intelligence test but does not assign white students with the same scores to that program, even though all other placement factors are equivalent.

Disparate Impact

A test may be discriminatory even if it is used in the same manner for all students and even if it is given under policies or practices that are the same for all students: it may result in a disproportionate denial of education benefits or opportunities to a particular group of students. Sometimes test scores result in students of a particular race, national origin or sex being denied — in numbers that are very different from their representation in the general student population — such education benefits or opportunities as promotion, graduation or placement. For example, a test would likely have a

disparate impact if it results in placement of 10 percent of the school's total number of minority group students and two percent of the school's total number of white students in special education classes for mentally retarded students. (Whether or not any particular set of numbers results in the type of disparate impact that triggers concern depends on the outcome of a statistical analysis, a topic discussed in more depth in *OCR's Nondiscrimination In High-Stakes Testing: A Resource Guide*.)

It is important to underscore that such disparate impact, by itself, does not mean that discrimination has taken place. Instead, it is merely a red flag — an indication of possible discrimination — that suggests that additional questions regarding test use should be answered. Before deciding if disparate impact discrimination has taken place, a complete set of questions must be asked and answered in a process involving several steps, outlined below.

Disproportionate Numbers of Students

- ❖ First, using our example, has the test resulted in markedly disproportionate numbers of students of a certain race, national origin or sex being placed in a special education class, as compared to the proportion of students of another race, national origin or sex? If the answer is yes, the next step is to determine the educational necessity of the test.

Validity and Reliability

- ❖ Where the test has a disparate impact, the school district must show that the test is educationally necessary. In determining whether a test is

educationally necessary, it must be shown that the test use is valid and reliable. Professionals in the field of testing use professionally accepted standards to assess the validity and reliability of a test in an educational decision-making context. Inferences from a test are valid if research demonstrates that the test measures what it is designed to measure when used appropriately, if the test is being used by the school in a manner consistent with its designed purpose, and if the test results are relevant to the educational decision in question. For a test to be considered reliable, there should be evidence that the same students, taking the test multiple times with no change in preparation, receive corresponding scores. (Additional information about test validity and reliability is set out in the box accompanying this text.)

If a test has a disparate impact and a school district cannot show that the test is both valid and reliable for its particular use, the test will not be found to be educationally necessary. Where a test with a disparate impact is not shown to be both valid and reliable for its particular use and the school district continues to use it in the same way, the district is in violation of Federal civil rights laws in education. If the test has a disparate impact and a school district can show that the test is both valid and reliable for its particular use, the next step is to determine whether there are any practical alternatives to the test in question that meet the school's educational needs.

More Information on Test Validity and Reliability

Professionals in the field of testing assess the use of a test to make educational decisions according to professionally accepted standards. The following set of inquiries introduces the complex process of determining the appropriateness of a test for use in a particular situation. In general, testing professionals ask the following types of questions.

Central inquiries - The central inquiries are:

- Does research demonstrate that the test measures what it is designed to measure for all students who are taking the test?
- Are the results reliable indicators of what the test is designed to measure?
- Is the test being used by the school in a manner that is consistent with its designed purpose?
- Is this measure relevant to the educational decision in question?

Where the answer to any of these basic inquiries is "no," the results are not appropriate for use in the particular situation. For example, let's use the example of a math achievement test designed for use state-wide in making decisions regarding whether a student is adequately prepared to move to the next grade level in math. A school might use this math achievement test in two ways - one way being a valid test use and one way being an invalid use of the test. First, it might use the test results in making decisions about whether to promote students to the next math grade. This use is consistent with the design of the

test. But what if the school uses the test to place students in a gifted and talented program in language arts? This use is invalid because it is inconsistent with the design of the test; the information from the test is not relevant to placement in a language arts gifted and talented program. In the first instance, test use is valid; in the second, it is invalid.

Where a test is being used as the sole criterion to make a high-stakes decision, the test must be designed for this use and there must be evidence indicating that it is appropriate to use the test as a sole criterion.¹ For example, a test designed to measure general intelligence would not be a good test to use as the sole measure in selecting students for a gifted and talented program. Because information provided by the test publisher indicates that the test was not designed for this purpose, this use of the test is not valid.

Achievement exams - Tests can be used in making decisions about whether students have acquired a certain degree of knowledge and skills. For example, a school district might require that

¹ The Standards for Educational and Psychological Testing, which are generally accepted professional standards that guide testing in schools, state that, "[I]n elementary and secondary education, a decision ... that will have a major impact on a test taker should not automatically be made on the basis of a single test score. Other relevant information for the decision should also be taken into account by the professionals making the decision." See American Psychological Association Standards for Educational and Psychological Testing (1985) at 8.12.

students pass a statewide test created to measure knowledge and skills in mathematics and language arts in order to receive a high school diploma. Schools have the obligation to ensure that all students are provided the opportunity to receive instruction that is fully consistent with and aligned to curriculum goals or standards set by the state or district for all students. Additionally, the state or school district that is testing the students has the obligation to ensure that the assessment is aligned with curriculum goals or standards. States or school districts must also be able to demonstrate that students have had enough time and opportunity to learn the material tested. If the instruction is not consistent with the goals or standards and if students have not had a fair opportunity to learn the material tested, schools have the responsibility to correct these problems.

Professionals in the field of testing use professionally accepted standards to assess the validity and reliability of a test in an educational decision-making context. It is important to note that a test is not necessarily valid or reliable merely because it is widely used by other school systems or because the company that developed it has a sound reputation.

For a more comprehensive discussion about technical considerations regarding tests, readers may consult OCR's Nondiscrimination in High-Stakes Testing: A Resource Guide, which includes a list of references on this topic

Alternatives With Less Disparate Impact

- ❖ Even where a test is valid and reliable, there still may be another test, or another way of measuring student achievement — or of measuring a characteristic such as level of proficiency in English or whatever characteristic the school wants to measure — that would serve the school's purpose as well as the test in question, and that would have a less negative impact on students of a particular race, national origin, or sex. If the use of such tests or measures is a practical alternative to the test in question and meets the school's educational needs, these alternatives must be used.

It is a good educational practice for school administrators to review the results of any testing program. If school administrators find that a test results in a disparate impact based on race, national origin or sex, the best educational practice is to inquire about other testing instruments or measures that would serve the school's purpose, be valid and reliable for that purpose, and that would provide all students with equal access to the school's programs and benefits.

Ways To Improve Test Use

Even if use of a certain test is found to be discriminatory, frequently there are steps a school district can take without eliminating the use of the test. For example, a school district can enhance student learning opportunities to help students master the skills and knowledge measured by the test. Or the district can add to its decision-making process such other evaluation standards as grades, teacher evaluations, portfolios containing student work, or even a second and different test. Lastly, the school can revise the test to make it valid and reliable for the purpose for which it is to be used.

QUESTIONS AND ANSWERS

Q. What information about testing may be useful to parents and students?

A. When parents or students talk to school staff about programs or individual progress, they may want to ask about any tests with high-stakes consequences that will be given during the school year. Parents and students may find it helpful to ask the name of each test; what knowledge, skill or ability each test is supposed to measure (for example, math achievement or general intelligence); and when each test will be given. Parents and students should know the school's overall decision - making process, including how the school judges the test's importance, what factors may be used along with the test, how each of these other factors may be weighted, and what the consequences are for students who fail the test. Parents and students may want to find out what remedial work will be offered to the student who performs poorly on the test and determine what additional opportunities, such as optional after-school classes to improve test performance, will be offered. A group of parents or an organization like the P.T.A. may want to meet with school counselors and administrators to learn about school tests. Many school districts find that the best educational practice is for the district to take the initiative in providing information about testing to parents and students as early as possible.

When high-stakes decisions are made, including decisions based, at least in part, on tests, parents and students may need information that will enable them to understand how the decisions were made and to assess whether a student or group of students is being treated fairly, regardless of race, sex, or national origin. In terms of facilitating understanding on the part of parents and students, the best educational practice is for school administrators to be able to explain how the overall decision-making process worked.

Q. What are proficiency and achievement tests?

A. Proficiency tests evaluate the mastery of knowledge and skills. They can include such tests as those which evaluate students' reading and writing skills in English. They might also include tests which evaluate academic mastery in subject areas taught in school, such as mathematics or science. These tests are often called achievement tests.

In connection with the national trend toward increasing accountability and encouraging high standards, many states and school districts are using the results of achievement tests to help determine graduation or grade promotion. There are many reasons for this, including ensuring that high school graduates are prepared to either enter college or compete in the job market. States and districts may also want to motivate students to work toward greater academic achievement, or to ensure that high school diplomas meaningfully represent a particular level of achievement.

Q. Are there any Federal requirements affecting public elementary and secondary school students with disabilities in connection with the use of high-stakes tests?

A. As two of the Federal civil rights laws listed on the inside pamphlet cover note, public elementary and secondary schools are required to provide a free appropriate public education to all students with disabilities in their jurisdiction. To comply with the requirements of these two Federal civil rights laws, a school must provide regular or special education and related aids and services necessary to meet the student's educational needs so that the student can participate in and benefit from the school's education program -- including participating in the public school's testing program.

When students with disabilities are tested, the civil rights requirements discussed in this pamphlet apply to them. In addition, schools must, among other requirements, select and administer the tests so that the results accurately reflect what the student knows or is able to do, rather than the student's disability. This means that students must be given appropriate accommodations and adaptations in the administration of the tests. Examples include oral testing, large print tests, Braille versions of tests, individual testing and separate group testing.

One high-stakes decision that affects some students with disabilities and that may involve tests, as well as other types of information, is the decision as to whether a student should be provided with special education. This decision

involves other decisions including: whether the student is an individual with a disability, covered by Section 504 and Title II; whether the student should be provided regular education with related aids and services or special education; and whether the student would be eligible under the Individuals with Disabilities Education Act (IDEA) (discussed below). Under both the civil rights laws and the IDEA, any determination of whether a student should be provided special education must be made on an individual basis in accordance with specific statutory and regulatory requirements, including requirements regarding the use of tests for that purpose.

The IDEA provides funds to states, and through them to local school districts, to assist in providing a free appropriate public education to students residing within the state in mandatory age ranges, and it establishes conditions for receipt of such funds. Under IDEA, the determination of whether a student needs special education must be made on an individual basis through a process which involves the use of tests or other evaluation materials and procedures. Under IDEA, states also must have nondiscriminatory procedures for purposes of evaluation and placement of students in special education, as well as procedures regarding the participation of students with disabilities in general state and districtwide assessment programs (discussed in the next question and answer). State and districtwide assessments of student achievement cannot be used alone for determining whether a student has an impairment and needs special education and related services under the IDEA. However, an individual student's performance on such an assessment could be considered by parents and school districts as a part of an evaluation for the

purposes of determining the student's need for special education under IDEA. For more information on how IDEA requirements affect high-stakes testing, please telephone the Department's Office of Special Education Programs at 202-205-5507.

Q. Should public school students with disabilities be included in proficiency tests with high-stakes consequences that are given throughout a district or state?

A. Under IDEA, this decision must be made on an individual basis by the student's individualized education program (IEP) team, and must be reflected in the student's IEP. For students who are not covered by the IDEA, but who are covered by Section 504, this decision must be made on an individual basis through other applicable evaluation and placement processes. It would be a violation of the civil rights laws prohibiting discrimination on the basis of disability, if a student with a disability who, based upon his or her IEP or Section 504 plan, should be preparing for and taking a state- or district-wide proficiency test, is excluded from these opportunities on the basis of disability. It is generally expected that students with disabilities should be included in these assessments. As described immediately above, where necessary, appropriate accommodations and adaptations in the administration of the test must be provided to students with disabilities who take these tests and should be specified in the student's IEP or Section 504 plan.

The newly enacted Individuals with Disabilities Education Act Amendments of 1997 (IDEA '97) requires States, as a condition of receiving IDEA funds, to include students with disabilities in State and districtwide assessment programs, with appropriate accommodations, where necessary. IDEA '97 also requires that the student's IEP specify any individual modifications in the administration of State or districtwide assessments of student achievement that are needed in order for the student to participate in such assessment. Similarly, if the IEP team determines that the student will not participate in a particular State or districtwide assessment of student achievement (or part of such an assessment), the student's IEP must include statements of why that assessment is not appropriate for the student and how the student will be assessed. IDEA '97 also requires state or local educational agencies to develop guidelines for disabled students who cannot take part in state and districtwide tests to participate in alternate assessments. These alternate assessments must be developed and conducted not later than July 1, 2000. For more information about the IDEA, please call the Department's Office of Special Education Programs at 202-205-5507.

Q. How are limited English proficiency (LEP) students affected by discrimination laws in the use of high-stakes tests?

A. LEP students ordinarily must be included in school or district assessment programs. When LEP students are included in these assessment programs, the inferences and interpretations drawn from the students' responses to the test or assessment procedure must be valid and reliable. Further, the students must be provided appropriate

accommodations in order to ensure valid and reliable results. Accommodations might occur in the test format (including editing accommodations) and/or in the administration, response or scoring conditions. Depending upon the nature and purpose of the test and the particular needs of a LEP student, providing a valid and reliable version of the test in the student's native language might be an appropriate accommodation. Other accommodations may include extended time or the use of bilingual dictionaries. Title VI requires the inclusion of LEP students in assessment programs, absent an educational or psychometric justification for their exclusion. If students are excluded from assessment programs, comparable information about their academic progress must be collected for these students.

Q. In addition to the civil right laws discussed in this pamphlet, are there any other Federal rights or requirements related to the use of proficiency tests for high-stakes decisions about which parents, students and school staff should be aware?

A. Yes, there are rights and requirements that arise from the Constitution of the United States. These apply to students and their public schools and they somewhat overlap with statutory civil rights requirements. An overview of these rights follows. However, it is important to point out that OCR does not enforce constitutional rights, unless there are claims of discrimination which are covered by the

federal civil rights statutes. For this reason, private legal counsel should be consulted for more information or to seek enforcement of these rights in Federal Court.

The Constitution requires that fundamental fairness be present in situations in which a government institution — such as a public school — creates an understanding on the part of students that they are entitled to something important: for example, a high school diploma. What if a student enters high school and the rules provide only that successful completion of coursework is necessary to receive a diploma? And then, when the student enters the senior year, the rules change to require that students also pass a proficiency test to earn a diploma? In this case, there are key issues related to whether a high-stakes test complies with constitutional standards for due process. Among the key issues addressed by Federal courts in making this determination are:

- whether there is a reasonable educational justification for the test;
- whether the test measures what it is designed to measure;
- whether the test represents a fundamental change in the rules of the school related to an important expectation, such as high school graduation, and, if so, whether the students had an adequate opportunity to prepare for, take, and pass the test.

Q. When should a parent or student file a complaint with OCR regarding test use?

A. School districts may take voluntary action to correct discrimination when it is brought to the attention of school decision-makers. However, parents or students may file a complaint with OCR immediately if they believe that discrimination has occurred.

Q. What are some of the steps OCR takes when investigating a discrimination complaint about the disparate impact of a test used to make a high-stakes decision?

A. OCR seeks information regarding whether the use of the test in question has resulted in a markedly disproportionate number of students of a certain race, national origin or sex being placed in or denied access to a particular program. Next, OCR determines if the school has shown that the test is educationally necessary. In determining educational necessity, OCR examines evidence of the test's validity and reliability, as described above. OCR then determines if there are practical alternatives to the test. Specifically, are there other tests or assessment procedures that would have a less negative effect on the basis of race, national origin or sex; or is there another reasonable way to achieve the school's objectives that would have a less negative impact on groups of students, while accomplishing the same educational objective as the test in question? Where appropriate, OCR also would be interested in the ways in which the school or school district would enhance learning opportunities so that students are prepared for the high-stakes test in question.

Q. How does a student or parent file a complaint with OCR?

A. If a student or a parent, or another person, decides to file a complaint with OCR, the complaint should be filed with the OCR enforcement office responsible for the state in which the school is located. The offices are listed on the last page of this pamphlet. Generally, the complaint should be filed within 180 days of the last act of alleged discrimination.

The complainant should give OCR his or her name, address, and ~~daytime phone number~~, and provide the date(s) and enough information about the alleged discriminatory act(s) so that OCR can understand the nature of the complaint.

OCR may extend the time for filing a complaint in certain circumstances. For example, when a student has filed a grievance under school procedures within 180 days of the last act of alleged discrimination, OCR will generally accept a complaint raising the same allegations up to 60 days after the end of the school procedures because it encourages students to file grievances with their own school first. How OCR resolves the complaint usually then would be limited to the allegations raised in the grievance.

[END]

4/99 DRAFT: For Internal Governmental Handling Only

Nondiscrimination in High-Stakes Testing: A Resource Guide



I disagree with the proposition that there are inherent racially based differences in the capacity of the American people to reach their full potential.

President Bill Clinton, October 21, 1994

An invalid test cannot measure merit.

Walls v. Mississippi State Dept. of Public Welfare,
542 F. Supp. 281, 311 (N.D. Miss. 1982), aff'd in
relevant part, 730 F. 2d 306 (5th Cir. 1984).

Table of Contents

NONDISCRIMINATION IN HIGH-STAKES TESTING: AN OVERVIEW	i
I. Introduction	i
II. Scope of the Resource Guide	iii
III. Foundations of the Resource Guide.....	iii
RESOURCE GUIDE	1
I. Introduction	1
II. Basic Federal Standards	1
III. Disparate Impact Analysis	2
IV. Different Treatment Analysis.....	8
V. Equal Opportunity for Limited-English Proficient Students.....	9
VI. Analysis Where Prior Dual System.....	9
VII. Remedies.....	10
APPENDIX: COMPENDIUM OF LEGAL AND TECHNICAL RESOURCES.....	1
I. Basic Federal Standards	1
II. Disparate Impact Analysis.....	3
III. Different Treatment Analysis.....	7
IV. Analysis Where Prior Dual System.....	8
TAB A	
SAMPLE QUESTIONS FOR EVALUATING EVIDENCE OF EDUCATIONAL NECESSITY AND DETERMINING WHETHER THERE ARE PRACTICABLE ALTERNATIVES WITH LESS DISPARATE IMPACT.....	A - 1
TAB B	
LIST OF POLICY AND TECHNICAL ASSISTANCE DOCUMENTS ADDRESSING THE CLASSIFICATION OF STUDENTS AND THE NONDISCRIMINATORY PROVISION OF SERVICES	B - 1
TAB C	
VALIDITY, A GLOSSARY.....	C - 1
TAB D	
BIBLIOGRAPHY.....	D - 1
Attachment - A	
Attachment - B	

NONDISCRIMINATION IN HIGH-STAKES TESTING: AN OVERVIEW

I. Introduction

The issue of nondiscrimination in high-stakes testing is, at its core, a critical issue concerning access to education. When tests are used to make educational decisions, they should be used to measure students' abilities, knowledge, or qualifications, regardless of race, national origin, or sex. The U.S. Department of Education's Office for Civil Rights (OCR) has developed this resource guide in order to provide our staff and members of the educational community that we serve with practical guidance on testing and assessment principles that lie at the core of Title VI of the Civil Rights Act of 1964 (Title VI) and Title IX of the Education Amendments of 1972 (Title IX) case law.

This Resource Guide provides an overview of the federal standards and related educational principles that should guide the use of tests for making high-stakes educational decisions, such as those that involve: student placement in gifted and talented programs or programs serving students with limited English proficiency; referral of students for special education services; student promotion from one grade to another grade level; diploma awards; and higher education admissions decisions and scholarship awards. This Guide applies to norm-referenced and criterion-referenced tests¹ as well as to professionally designed alternative forms of assessment, which are used for making high-stakes educational decisions. The Resource Guide is not intended to apply to tests that are used to measure the performance of schools but have no high-stakes consequences for individual students nor does the Resource Guide address teacher-created classroom tests, even when such tests are being used for high-stakes educational decisions.

The issue of nondiscrimination in testing and assessment is properly viewed as consistent with standards-based reforms. Education leaders and the general public agree that there must be challenging standards for all students. In recent years, States and communities across the nation have embarked on far-reaching systemic efforts to reform their schools. Uniting their efforts has been an emphasis on high academic standards and high-quality assessments geared to those standards.

By defining what students should know and be able to do, standards keep schools focused on the desired results for students and can stimulate the development of appropriate curricula and the application of effective teaching strategies to make these results possible. Standards also indicate what assessments must measure in order to show achievement.

¹ Norm-referenced tests are tests used to identify an individual's performance in relation to the performance of other people in a specified group on the same test. American Psychological Association Standards for Educational and Psychological Testing (1985) (APA Standards) at p. 92. Criterion-referenced tests allow users to make score interpretations in relation to a functional performance level. APA Standards at p. 90. In other words, criterion-referenced tests are designed to measure to what degree a learner has mastered a certain skill.

High-quality assessments can make high standards meaningful by providing communities with a mechanism by which to hold schools accountable for achievement. It is critical that high standards for academic achievement be coupled with the necessary instruction and support that help students reach those standards - as determined by valid and reliable assessments.

The U.S. Department of Education is committed to the support of high standards and challenging assessments for all students. By outlining the relevant requirements of federal civil rights law, this Guide should assist states and local educational agencies instituting high-stakes assessments for all students. The Guide is intended to help states and local educational agencies avoid potential pitfalls in their implementation of high standards when using large scale assessments with educational consequences for individual students.

Federal civil rights laws ensure that all students have equal educational opportunities. Although many of the federal legal standards that should guide sound educational decisions are importable from the federal cases addressing employment discrimination, there are, nonetheless, critical differences. The educational institution's obligation to a student does not ordinarily end once a decision is reached to, for example, place the student in a particular educational program. The educational institution is responsible for ensuring that the student has appropriate educational opportunities throughout his or her educational career to improve and develop needed academic skills. Indeed, observing the differences between the employment and education settings, a federal court recognized:

If tests predict that a person is going to be a poor employee, the employer can legitimately deny the person the job, but if tests suggest that a young child is probably going to be a poor student, a school cannot on that basis alone deny that child the opportunity to improve and develop the academic skills necessary to success in our society.

Larry P. v. Riles, 793 F.2d 969, 980 (9th Cir. 1984).²

Similarly, the question of test use cannot be examined in a vacuum. While the Resource Guide focuses specifically on the discriminatory use of tests which are used for high-stakes educational decisions, this issue must be considered in the context of the educational objectives involved and the effect of the particular testing practice in question upon students, particularly where classification of students and the provision of services is at issue. (Tab B of this Resource Guide lists policy and technical assistance documents that provide resource information and legal guidance relating to the nondiscriminatory classification of students and the provision of services to students.)

² See also National Research Council, High Stakes Testing for Tracking, Promotion, and Graduation, at pp. 61 - 62, 76 - 77, 97 (National Academy Press, Washington, D.C. 1999).

II. Scope of the Resource Guide

The Resource Guide does not apply to modifications of tests and/or testing conditions required for the purpose of accommodating individuals with disabilities under Section 504 of the Rehabilitation Act, the Individuals with Disabilities Education Act (IDEA) or the Americans with Disabilities Act (ADA).³ Although the legal theories of discrimination discussed in the Resource Guide are generally applicable to disability issues that arise under Section 504, the IDEA and the ADA, an additional analysis regarding testing accommodations provided to individuals with disabilities is also required. *See, e.g.*, Attachment A: Dear Colleague Letter (September 29, 1997) (addressing the inclusion of students with disabilities in statewide assessment systems). This analysis is beyond the scope of the Resource Guide.

The Resource Guide, along with the attached Compendium of Legal and Technical Resources (Appendix), should be read as an explanation of the legal and conceptual framework needed for understanding the issues raised by challenges to high-stakes testing. The model (and pragmatic) questions set out in Tab A should be viewed as a starting point for addressing questions of great complexity regarding challenges to testing and assessment practices. These model questions do not define the "floor" of what must be asked any more than they define the "ceiling" of what may be asked. Those decisions are inherently case-specific. Tab C provides a glossary of terms relating to test validity.

III. Foundations of the Resource Guide

A. Professional Standards

Generally-accepted professional standards for evaluating standardized tests provide a significant foundation for this guide. They include those described in the Standards for Educational and Psychological Tests prepared by a joint committee of the American Psychological Association, the American Educational Research Association, and the National Council on Measurement in Education; the Code of Fair Testing Practices in Education prepared by the Joint Committee on Testing Practices; and the Uniform Guidelines on Employee Selection Procedures. At OCR's request, the National Academy of Sciences' Board on Testing and Assessment (BOTA) reviewed earlier drafts of this guide and provided comments, which have helped to ensure that the Resource Guide is consistent with existing professional standards.

³ Section 504 of the Rehabilitation Act and Title II of the Americans with Disabilities Act are enforced by OCR; the Individuals with Disabilities Education Act is administered by the U.S. Department of Education's Office of Special Education Programs.

B. Legal Standards

This guide outlines two separate legal theories of discrimination: disparate treatment and disparate impact. Each theory is based on settled federal legal principles under Title VI of the Civil Rights Act of 1964 and Title IX of the Education Amendments of 1972 and neither breaks any new legal ground.

A disparate treatment analysis is used to determine whether a policy or practice regarding testing is being applied differently to an individual student or group of students because of their race, national origin, or gender, without legal justification for doing so, e.g., as a remedy for past de jure discrimination. This analysis would be used to determine, for example, whether black students and white students are being tested under different conditions or whether students with the same test scores are being treated differently by an educational institution.

Under a disparate impact analysis, the focus is on the "effects" of the application of a facially neutral policy or practice, regardless of whether the adverse consequences for a particular race, national origin, or gender were intended. The use of a disparate impact analysis is appropriate when the use of a test pursuant to a race-neutral policy or practice creates a significant difference in the granting or denial of benefits or opportunities on the basis of race, national origin or sex. Tests that have a disparate impact on the basis of race, national origin, or sex must be educationally necessary; otherwise, they are not permissible under Title VI or Title IX. Educational necessity involves a showing that the test is valid and reliable for the purpose for which it is being used⁴. The use of the test is still not permissible under Title VI or Title IX if the test is not the least discriminatory practical alternative that can serve the education institution's educational purpose. See Attachment B: Memorandum from the Attorney General for Heads of Departments and Agencies that Provide Federal Financial Assistance, "Use of the Disparate Impact Standard in Administrative Regulations under Title VI of the Civil Rights Act," July 14, 1994.

⁴ Section III. B. of the Resource Guide contains a discussion of test validity and reliability. Tab C provides a glossary of terms relating to test validity.

RESOURCE GUIDE

I. Introduction

When tests are used to make educational decisions, they should be used to measure students' abilities, knowledge, or qualifications, regardless of race, national origin, or sex. Civil rights concerns arise when test uses do not satisfy federal antidiscrimination standards. This Resource Guide outlines the requirements of federal law prohibiting misuse of tests and other assessment procedures that result in discrimination based on race, national origin, or sex. It is designed to provide a general analytical framework under Title VI and Title IX for determining the proper use of tests and other assessment procedures in the educational context.

In evaluating a test or other assessment procedure, it is important to consider how the test is being used. In some cases, it may be used to make a certification or selection decision (*e.g.*, admission to a school, awarding of a scholarship, or teacher certification). In other cases, it may be used to classify students (*e.g.*, to identify students as needing special education or special language services or to identify students as gifted and talented).

When high-stakes educational decisions are made, tests may be used in conjunction with other criteria, such as teachers' recommendations. Ordinarily, if there are allegations or evidence regarding possible discrimination on the basis of race, national origin, or sex with respect to the use of a test or other criteria as part of a high-stakes decision making process, there should be an inquiry into the operation of the entire assessment process. There should be an inquiry into what criteria are being utilized as part of the entire process and the weight being given to each of the criteria in the process. This Resource Guide focuses on cases where the test or assessment contributes significantly to the high-stakes decision. However, if other criteria are contributing to a disparate impact on the basis of race, national origin, or sex, they should be evaluated to ensure that they are educationally appropriate and necessary, as well.

II. Basic Federal Standards

The requirements of Title VI and Title IX apply to all educational institutions that receive federal funds. These laws apply to all of the academic, athletic, and extracurricular programs of the institution, whether conducted in facilities of the recipient or elsewhere. Title VI prohibits race and national origin discrimination in programs and activities that receive Federal financial assistance. Title IX prohibits sex discrimination in education programs that receive Federal financial assistance. Title VI and Title IX cover the uses of property that the recipient owns and the activities that the recipient sponsors. Title VI and Title IX cover these operations, whether the individuals involved in a given activity are students, faculty, employees, applicants, or other participants. See Compendium at pp. 1 - 3.

Some federal courts have addressed challenges to the use of tests for high-stakes purposes under the equal protection and due process clauses of the Fourteenth Amendment to the U.S. Constitution. Although OCR enforces statutory rights under Title VI and Title IX rather than constitutional rights, to the extent the claim is that a school district's use of tests is discriminatory, those actions may violate both the statutes and the Constitution. OCR normally would not be involved, however, in cases in which there were no allegations of discrimination. Thus, those cases challenging the use of tests for constitutional reasons unrelated to discrimination would not fall within OCR's jurisdiction. Some federal cases in which discrimination claims have been raised have also involved equal protection challenges to a jurisdiction's use of tests in which the claim is based not on discriminatory intent but on the jurisdiction's use of tests to separate out those students who should not be allowed to graduate.⁵ Under these circumstances, since there is no claim of discrimination based on membership in a suspect class, the equal protection claim is reviewed under the rational basis standard. The jurisdiction thus need show only that the use of the tests has a rational relation to a valid state interest. See Debra P. v. Turlington, 644 F.2d 397, 406 (5th Cir. 1981); Erik V. v. Causby, 977 F. Supp. 384, 389 (E.D.N.C. 1997).⁶

Due process challenges to the use of tests fall into two categories, substantive and procedural due process. Analyses under the due process clause address whether students have been denied, based on test scores, educational benefits or opportunities to which they had a legitimate claim of entitlement. Such cases typically involve a procedural due process claim that student were not given sufficient notice of the test and its requirements, or a substantive due process claim that the students were not taught the material on which the tests were based. Debra P., 644 F.2d at 404-405; Crump v. Gilmer Independent School District, 797 F. Supp. 552, 555-556 (E.D.Tex. 1992).

III. Disparate Impact Analysis

A disparate impact analysis may be applied to allegations involving discriminatory test use by educational institutions. Under this analysis, the use of any educational test which has a significant disparate impact on members of any particular race, national origin, or sex is discriminatory, and a violation of Title VI and/or Title IX, respectively, unless it is

⁵ As a general matter, courts express reluctance to second guess a state's educational policy choices when faced with such challenges, although recognizing that a state cannot "exercise that [plenary] power without reasons and without regard to the United States' Constitution." Debra P. v. Turlington, 644 F.2d 397, 403 (5th Cir. 1981).

⁶ Where, however, the use of a facially race-neutral test perpetuated the effects of the prior dual school system in which students were intentionally segregated on the basis of race, such a test could violate the equal protection clause even absent direct evidence of discriminatory intent. Debra P., 644 F.2d at 407, citing Arlington Heights v. Metropolitan Housing Development Corp., 429 U.S. 252, 266-268 (1977); see also Anderson v. Banks, 520 F. Supp. 472, 500 (S.D.Ga. 1981) (discriminatory impact of the test cannot be considered separately from the de jure discrimination that preceded it).

educationally necessary and there is no practicable alternative form of assessment which meets the educational institution's educational needs and would have less of a disparate impact on the basis of race, national origin, or sex.

In applying a disparate impact analysis, the following questions should be addressed:

- A. Does the educational institution's use of an educational test result in the significantly disproportionate denial of an educational benefit or opportunity to members of a particular race, national origin, or sex?
- B. If so, is the use of the test educationally necessary?
- C. If so, do there exist practicable alternative forms of assessment which would substantially serve the school's stated purpose and are valid and reliable for that purpose, but which have less of a disparate impact on the basis of race, national origin, or sex?

Each question is discussed in more detail below. Where, based on evidence, there is a finding that the use of a test or assessment procedure caused or contributed to a disparate impact on members of a particular race, national origin, or sex (the first question), and the test or procedure does not meet the legal standard of educational necessity (the second question) or there is a practicable alternative form of assessment which would meet the educational institution's educational needs and would have less of a disparate impact on the basis of race, national origin, or sex (the third question), there is a violation of Title VI or Title IX under this disparate impact analysis.

A. Establishing Disparate Impact

Under a disparate impact analysis, a school's use of an educational test that causes or contributes to a disproportionate denial of an educational benefit or opportunity to members of a particular race, national origin, or sex is sufficient information to indicate a possible failure of compliance with Title VI or Title IX which should be investigated further. It is important to note that disparate impact by itself does not necessarily mean that discrimination has taken place. Disparate impact may lead to a finding of discrimination only when the use of the test in question is not educationally necessary or when there is no practicable alternative form of assessment which would meet the educational institution's educational needs and have less of a disparate impact on the basis of race, national origin, or sex.

B. Establishing Educational Necessity

Once it has been determined that a disparate impact exists, it must then be determined whether the use of the test or assessment procedure is educationally necessary.⁷ To meet

⁷ Where a test is being used as the sole or principal criterion for making educational decisions and where it was clearly not designed to be used as such, there is no basis upon which to conclude that the test is educationally necessary.

the educational necessity standard, the test or assessment procedure must be valid and reliable for the purpose for which it is being used.

In evaluating the validity and reliability of a test or assessment procedure, generally accepted professional standards should be the foundation for such decision making. These standards include the Standards for Educational and Psychological Testing prepared by a joint committee of the American Psychological Association, the American Educational Research Association, and the National Council on Measurement in Education; the Code of Fair Testing Practices in Education prepared by the Joint Committee on Testing Practices; and the Uniform Guidelines on Employee Selection Procedures⁸. All decisions as to whether a test or procedure has met professionally accepted standards should be made in consultation with experts.

As discussed below, in determining whether a test or assessment procedure is educationally necessary, it must be shown that the test or procedure is valid and reliable for the purpose for which it is being used.

1. Technical Considerations

Validity

Establishing validity is the process of evaluating the degree to which a test measures what it claims to measure and leads to legitimate inferences that are appropriate or meaningful.⁹ The demonstration of validity is multifaceted and depends on the type of assessment and the purposes for which the test was designed to be used.

Often, validity demonstrations will require careful analysis of data according to existing professional standards. This is a complex and specialized endeavor, and professionally accepted validation standards and techniques are evolving (for example, the 1985 Standards for Educational and Psychological Testing are currently being revised). Tab C contains a glossary of terms related to test validity.

⁸ Although there are many principles in the Uniform Guidelines that apply to educational testing in general terms, the Uniform Guidelines do not address educational testing issues. There are critical, contextual differences between employment and educational testing that should not be overlooked when using the Uniform Guidelines as a resource in the educational setting. The Uniform Guidelines were adopted by and are currently used by the U.S. Equal Employment Opportunity Commission, the U.S. Department of Labor, and the U.S. Department of Justice.

⁹ Indeed, it may not be technically correct to refer to a test or assessment procedure as being valid. Rather, it is the inferences and interpretation drawn from the responses to the test or procedure that must be valid. However, for simplicity's sake, this guidance will often use the more common approach of referring to the test or procedure as being valid for the purpose for which it is being used.

In some cases, a test will clearly not be valid for the purpose for which it is being used. For instance, where a test manufacturer states that a test is not valid for use as a sole criterion in educational decision making, it is a clear misuse of that test if the school, in fact, uses only the test results in making a high-stakes educational decision.

Construct validity is relevant when an assessment is used to measure a particular characteristic, property, skill, ability, capacity, academic achievement, or behavior. The construct validation of a test usually involves a series of studies, using a variety of research methodologies.

The validation of constructs of academic content are relevant when a recipient is using a test to measure the acquisition of specific knowledge or academic skills. For example, a statewide proficiency test designed to measure whether students have learned specific skills or gained specific knowledge in order to determine whether they should receive a diploma would be subject to an assessment of the validity of the constructs of its content.

Criterion-related validity is relevant when scores on a test or assessment procedure are related to the examinee's performance on some other measure, which is known as a criterion. For example, when a recipient is using test scores to accept or reject applicants to a particular program, school, or curriculum, it should have evidence that the test scores correlate significantly with success in the program, school, or curriculum.

Reliability

Along with evidence of a test's validity, evidence of a test's reliability over time and over students should be considered and must conform to accepted professional standards.¹⁰ Reliability is the degree to which test scores are consistent, dependable, or repeatable. For a test to be considered reliable, there should be evidence that the same students, taking the test multiple times with no change in preparation, receive corresponding scores. No test is perfectly reliable and differing amounts of error or unreliability are tolerated, depending upon the purposes for which the test or procedure is designed to be used. Reliability may be affected by the type of assessment procedure at issue, e.g., a standardized test versus a performance-based assessment.¹¹

¹⁰ The 1985 Standards for Educational and Psychological Testing discuss reliability. See APA Standards at pp. 19 - 23.

¹¹ Performance-based assessment requires students to generate rather than choose a response. Students are required to actively accomplish complex and significant tasks, while bringing to bear prior knowledge, recent learning, and relevant skills to solve problems. Demonstrations, written or oral responses, journals and portfolios are examples of performance-based assessment. Herman, J.L., Aschbacher, P.R., & Winters, L. (1992). A Practical Guide to Alternative Assessment. Alexandria, VA: Association for Supervision and Curriculum Development.

Fairness

Within the constraints of the defined purposes of a test or procedure, it is expected that the assessment will be valid and reliable for all students taking the assessment.¹² That is, there must be adequate evidence that the test is measuring the same academic constructs for all students, and that the results are sufficiently precise for all students.

Use

Assessment results can be used appropriately or inappropriately. Misuse can stem from two test-related considerations, as well as other problems in the decision-making process. That is, users may suggest a test or procedure is measuring what it is not, thereby producing invalid inferences. They may attempt to use results in making decisions which require a higher level of precision or reliability than the assessment is designed to produce. An example of this type of misuse is a school district using results from a test as a sole criterion in making a high-stakes decision when the test publisher has stated that the test is not to be used as a sole criterion.¹³ The processes which users engage in to make decisions about individuals or groups may themselves be flawed, so that the results of tests with reasonably valid and reliable inferences are used inappropriately.

Invalid inferences can stem from misalignment between what is described as being assessed and what is actually being measured. It might also stem from a misalignment between curriculum goals or standards and what the high-stakes test or procedure is measuring, or between what is being assessed and what is being taught in classrooms. In each situation, the source of the misalignment must be established so that it can be determined where changes are needed. For instance, in determining whether a high-stakes test is being used appropriately, it may be appropriate to determine the degree to which schools provide instruction in the knowledge and

¹² The 1985 Standards for Educational and Psychological Testing address technical issues of fairness in testing. See e.g., APA Standards at standard 1.2, 1.5, 1.8, 1.10, 1.13, 3.5, and 3.10. See also Paul W. Holland & Howard Walner, Differential Item Functioning (Lawrence Erlbaum Associates, Publishers 1993); National Research Council, High Stakes Testing for Tracking, Promotion, and Graduation, at pp. 78 - 82.

¹³ See also APA Standards at standard 8.12 ("[i]n elementary and secondary education, a decision ... that will have a major impact on a test taker should not automatically be made on the basis of a single test score. Other relevant information for the decision should also be taken into account by the professionals making the decision."); National Research Council, High Stakes Testing for Tracking, Promotion, and Graduation, at p. 3.

skills measured by the test.¹⁴ A statewide proficiency test could be subject to an analysis of the degree to which the schools in the State provide adequate instruction in the content areas measured by the test. Often, it is necessary to determine whether curriculum goals or standards have been clearly identified. If they have been clearly identified, then the alignment of instruction and assessment should flow from these standards. Misalignment would occur if either the instruction or assessment is not consistent with the standards.

2. Establishing Technical Merit

Tab A includes guidance on the types of questions to ask and information to obtain regarding the technical merit of assessments. These sample questions should be considered as starting points for appropriate inquiry. In most cases, these questions should be refined, modified, and supplemented based on the facts of the case and the advice of testing and/or other education experts.

The following guidelines should be considered when evaluating evidence of technical merit:

a. **No assumption of technical merit.** The general reputation of a test, its author, or its publisher, or casual reports of its validity are not evidence of a test's technical merit. A test is not considered technically viable under federal law based on a test's name or descriptive labels; promotional literature about the test; data regarding the frequency of a test's use; or testimonial statements and credentials of test publishers, consultants, or schools which have previously used the test. A publisher's test manual may provide technical evidence; this alone is not sufficient to determine technical merit.

b. **Acceptable types of evidence.** The use of a test should be supported by studies of the same test conducted by test publishers or professional researchers which demonstrate adequate validity and reliability for the particular use. Such studies must show that the use of the test by the school is the professionally accepted equivalent to the use for which the test was validated. The use of the test by the school should be within the technical parameters defined by the publisher and demonstrated by the evidence.

¹⁴ Several federal court decisions have addressed the degree to which schools have provided adequate instruction in the knowledge and skills measured by a test. See Debra P. v. Turlington, 644 F. 2d 397, 405 (5th Cir. 1981); Crump v. Gilmer Independent School District, 797 F. Supp. 552, 555-6 (E.D. Tex. 1992). The inquiry regarding whether there is an alignment between knowledge and skills that are being tested and the curriculum and instruction that are being provided to students is critical when the test use in question involves an assessment of learning or achievement in school.

As one part of the process of showing that a test or procedure is technically sound, it may be appropriate to assess the degree of relationship between test scores and performance criteria. This may be done by researchers using professionally accepted research and statistical procedures.

3. Cutoff Scores

In determining whether a test or procedure with a disparate impact is educationally necessary, it is necessary to look to how the test or procedure is actually used by the recipient. In some cases, a test or assessment procedure may be used without a specific passing or cutoff score.

In other cases, a score may be set, either by the test developer or the test user. Standard 6.9 of the 1985 Standards for Educational and Psychological Tests states that "[w]hen a specific cut score is used to select, classify or certify test takers, the method and rationale for setting that cut score, including any technical analyses, should be presented in a manual or report." This information must be considered in determining whether the cutoff score used by a recipient was set by some systematic process that reflects the good faith exercise of professional judgment.

C. Alternatives With Less Disparate Impact

Even if a school can show that a test or assessment procedure is valid and reliable, the school's continued use of the test or procedure may be in violation of federal law if one or more instruments, criteria, or procedures are available as a practicable alternative to the challenged test or procedure, and if any such alternative 1) substantially serves the educational purposes for which the test or procedure is used, 2) is valid and reliable for those purposes, and 3) would have a lesser disparate impact.

It is a good educational practice for school administrators to review the results of any testing program. If school administrators find that a test results in a disparate impact based on race, national origin or sex, the best educational practice is to inquire about other testing instruments or measures that would serve the school's educational purpose, be valid and reliable for that purpose, and have a less negative impact on students of a particular race, national origin, or sex.

IV. Different Treatment Analysis

If warranted by the nature and scope of the allegations or evidence, a different treatment analysis may be utilized, as described below, to determine whether the educational institution administered a test or assessment procedure differently or used scores differently because of the students' race, national origin, or sex, without a legitimate, nondiscriminatory reason. If the reason for the different treatment was, e.g.: 1) the provision of testing accommodations or auxiliary aids to qualified individuals with disabilities as required by Section 504 of the Rehabilitation Act of 1973 or Title II of the Americans with Disabilities Act of 1991; or 2)

voluntary or remedial affirmative action undertaken in accordance with federal law, the educational institution may have a legal justification for the different treatment and there may be no violation of federal law.

Otherwise, tests and assessment procedures must be administered and scores used in the same manner regardless of race, national origin, or sex. Even if a test or procedure is supported by sufficient evidence of educational necessity, an educational institution may still be in violation of Title VI and/or Title IX if the test or procedure is administered differently or the scores are used differently for students because of their race, national origin, or sex.

V. Equal Opportunity for Limited-English Proficient Students

The requirements of Title VI discussed above regarding the use of tests for making high-stakes educational decisions are applicable when tests are being used to make high-stakes educational decisions concerning students with limited English proficiency. Under Title VI and other federal laws, State educational agencies and school districts are required to ensure that students are not denied equal educational opportunities because a student has limited proficiency in English.

Limited English proficient (LEP) students must ordinarily be included in assessment programs. When LEP students are included in assessment programs, the inferences and interpretations drawn from the students' responses to the test or assessment procedure must be valid and reliable. Further, the students must be provided appropriate accommodations in order to ensure valid and reliable results. Accommodations might occur in the test format (including editing accommodations) and/or in the administration, response or scoring conditions. Depending upon the nature and purpose of the test and the particular needs of a LEP student, if students are literate in their native language, and if the instruction has been in that language, providing a valid and reliable version of the test in the student's native language might be an appropriate accommodation. Other accommodations may include extended time or the use of bilingual dictionaries. If students are excluded from assessment programs, based on legitimate educational or psychometric justifications for their exclusion, comparable information about their academic progress must be collected for these students.

VI. Analysis Where Prior Dual System

School districts that have operated dual systems and have not been declared unitary have an obligation to dismantle their prior de jure segregated systems. The use of any educational test or assessment procedure may be a violation of Title VI if it had been used to achieve the segregation or if it perpetuates the segregation. Where such tests or assessment procedures are being used, school districts have an obligation to identify, consider and implement less discriminatory criteria consistent with sound educational policy, to the extent practicable.

VII. Remedies

Where an educational institution is in violation of Title VI or Title IX, there are a range of remedies that may be used to come into compliance with federal law. Remedies should be designed to ensure that educational institutions comply with civil rights statutes when meeting their educational goals. Depending on the facts of a given case, there are many permissible responses to correcting a violation. If the administration or design of a test is discriminatory on the basis of race, national origin, or sex, appropriate remedies might include: supplementing the use of the test with other assessment measures; revising the test instrument within a reasonable period of time to address compliance concerns; or substituting the test with another available instrument that more appropriately measures what is intended to be measured. If the test or assessment procedure reflects discriminatory educational practices with respect to the adequacy of instruction provided to students to prepare them to take the test, an appropriate remedy might include enhancing learning opportunities for students to perform well on the test.

APPENDIX: COMPENDIUM OF LEGAL AND TECHNICAL RESOURCES

This compendium provides an outline of key legal and technical resources to serve as a reference for inquiries regarding potential discrimination in the use of an educational test or assessment procedures.

The investigation and analysis of disparate impact cases under Title VI of the Civil Rights Act of 1964 (Title VI), 42 U.S.C. § 2000d, and Title IX of the Education Amendments of 1972 (Title IX), 20 U.S.C. § 1681, rely, to a large extent, on case law developed under Title VII of the Civil Rights Act of 1964, 42 U.S.C. § 2000e, which prohibits discrimination on the basis of race, color, national origin, sex, and religion in employment. See United States v. LULAC, 793 F.2d 636, 648-49 (5th Cir. 1986); Georgia State Conference of Branches of NAACP v. Georgia, 775 F.2d 1403, 1417 (11th Cir. 1985); NAACP v. Medical Center, Inc., 657 F.2d 1322 (3rd Cir. 1981); Dillon County District No. 1 and South Carolina State Department of Education, No. 84-VI-16 (Civil Rights Reviewing Authority 1987).

I. Basic Federal Standards

A. Title VI and Title IX Prohibit Discrimination in Federally Funded Programs and Activities

Title VI prohibits race and national origin discrimination in programs and activities that receive Federal financial assistance. Title IX prohibits sex discrimination in education programs that receive Federal financial assistance. See also 34 C.F.R. Part 100 (regulations implementing provisions of Title VI) and 34 C.F.R. Part 106 (regulations implementing provisions of Title IX). Under the Civil Rights Restoration Act of 1987, OCR generally has institution-wide jurisdiction over a recipient of Federal funds. See 42 U.S.C. § 2000d-4 (1989).

B. Specific Discriminatory Actions Prohibited By Title VI and Title IX¹⁵

The regulations implementing Title VI do not specifically address the use of tests and assessment procedures, but do include a general provision prohibiting discrimination based on race or national origin. 34 C.F.R. § 100.3(b)(5).

The regulations implementing Title IX specifically prohibit the discriminatory use of tests or assessment procedures in admissions, 34 C.F.R. § 106.21, employment, 34 C.F.R. § 106.52, and counseling, 34 C.F.R. § 106.36. Title IX further prohibits discrimination in areas in which test or assessment procedure results are often used to allocate benefits and opportunities. See 34 C.F.R. § 106.37(a) (prohibition against discrimination in financial aid awards and against assisting any entity which provides financial aid to students in a manner which discriminates based on sex); 34 C.F.R. § 106.31(b)(6) (prohibition against providing "significant assistance" to entities which discriminate on the basis of sex in providing any aid, benefit or service to students or employees).

See also 34 C.F.R. § 100, Appendix B, part K (Guidelines for Eliminating Discrimination and Denial of Services on the Basis of Race, Color, National Origin, Sex, and Handicap in Vocational Education Programs) ("if a recipient can demonstrate that . . . criteria [that disproportionately exclude persons of a particular race, color, national origin, sex, or disability] have been validated as essential to participation in a given program and that

¹⁵ Some federal courts have addressed challenges to the use of tests for high-stakes purposes under the equal protection and due process clauses of the Fourteenth Amendment to the U.S. Constitution. Although OCR enforces statutory rights under Title VI and Title IX rather than constitutional rights, to the extent the claim is that a school district's use of tests is discriminatory, those actions may violate both the statutes and the Constitution. Some federal cases in which discrimination claims have been raised have also involved equal protection challenges to a jurisdiction's use of tests in which the claim is based not on race or sex discrimination, but on the jurisdiction's use of tests to determine, for example, those students who should be allowed to graduate. Under these circumstances, the claim is reviewed under the rational basis standard and the jurisdiction need show only that the use of the tests has a rational relation to a valid state interest. See Debra P. v. Turlington, 644 F.2d 397, 406 (5th Cir. 1981); Erik V. v. Causby, 977 F. Supp. 384, 389 (E.D.N.C. 1997).

Due process challenges to the use of tests fall into two categories, substantive and procedural due process. Analyses under the due process clause address whether students have been denied, based on test scores, educational benefits or opportunities to which they had a legitimate claim of entitlement. Such cases typically involve a procedural due process claim that students were not given sufficient notice of the test and its requirements, or a substantive due process claim that the students were not taught the material on which the tests were based. Debra P., 644 F.2d at 404-405; Crump v. Gilmer Independent School District, 797 F. Supp. 552, 555-556 (E.D.Tex. 1992); cf. Williams v. Austin Independent School District, 796 F. Supp. 251 (W.D.Tex. 1992).

alternative equally valid criteria that do not have such a disproportionate adverse effect are unavailable, the criteria will be judged nondiscriminatory. Examples of admission criteria that must meet this test or assessment procedure are ... interest inventories ... and standardized test or assessment procedures").

II. Disparate Impact Analysis

Because the regulations that implement Title VI and Title IX incorporate an effects standard, a recipient's use of facially neutral policies that have a disparate impact on the basis of race, national origin, or sex may constitute a violation of Title VI or Title IX. See 34 C.F.R. § 100.3(b)(2); 34 C.F.R. § 106.21(b)(2); 34 C.F.R. § 106.36(b); 34 C.F.R. § 106.52. See also Guardians Assn. v. City Service Commission of City of N.Y., 463 U.S. 582 (1983) Lau v. Nichols, 414 U.S. 563 (1974). See also Attachment B: Memorandum from the Attorney General for Heads of Department and Agencies that Provide Federal Financial Assistance, "Use of the Disparate Impact Standard in Administrative Regulations under Title VI of the Civil Rights Act," July 14, 1994.

A. Establishing Disparate Impact

There is no rigid mathematical threshold that must be met to demonstrate a disparate impact. Watson v. Fort Worth Bank and Trust, 487 U.S. 977, 994-95 (1988)("statistical disparities must be sufficiently substantial to raise ... an inference of causation," i.e., "show that the practice in question caused the exclusion of applicants for jobs or promotions because of their membership in a protected group").

Groves v. Alabama State Board of Education, 776 F.Supp. 1518, 1523-1529 (M.D. Ala. 1991)(discussion on establishing a statistical prima facie case of disparate impact).

Georgia State Conference of Branches of NAACP, supra at 1421 ("Generally, to establish a prima facie case of disparate impact based on race the plaintiffs must show that the defendants' racially neutral practice detrimentally affects persons of a particular race to a greater extent than other races. . .").

B. Establishing Educational Necessity

The use of an educational test or assessment procedure which has a disparate impact on members of any race, national origin, or sex group is discriminatory, and a violation of Title VI or Title IX, unless the recipient justifies the use as educationally necessary. See Board of Education v. Harris, 444 U.S. 130, 151 (1979)(in disparate impact cases in the education context defendants are required to show an educational necessity instead of a business necessity); Griggs v. Duke Power Company, 401 U.S. 424 (1971)(sets similar standard for disparate impact of an employment test or assessment procedure); Branches of NAACP v. State of Georgia, 775 F.2d 1403, 1417 (11th Cir.

1985); and Sharif v. New York State Education Department, 709 F.Supp. 345 (S.D. N.Y. 1989)(standard for disparate impact of an educational test or assessment procedure is educational necessity).

See also Memorandum from then Acting Assistant Secretary for Civil Rights, Alicia Coro, to then Acting Regional Civil Rights Director, Region V, Linda A. McGovern (PCD # 70 October 22, 1986)(standard in case involving alleged disparate impact of LSAT scores as an admissions criterion at DePaul University College of Law and Illinois Institute of Technology/Chicago/Kent College of Law, is educational necessity).

Whether a test or assessment procedure is educationally necessary depends on whether the test or assessment procedure is valid for the purpose for which it is being used. Sharif, supra; State of Georgia, supra; cf. Final Order of the Civil Rights Reviewing Authority, Dillon County School District No. 1, Docket No.84-IV-16.

The Guidance is consistent with professional standards - See APA Standards and the Code of Fair Testing Practices in Education.

1. Validity

Memorandum from Harry M. Singleton to John E. Palomino (PCD # 57 April 4, 1985)(valid test or assessment procedures "successfully measure what they claim to measure; are used only for the specific purpose(s) for which they were developed; and, are administered in conformance with the instructions provided by the publisher").

See APA Standards at p. 11 (defining criterion-related evidence) and standard 1.11, 1.12, 1.18 (describing criterion-related validation studies).

See APA Standards at p. 10 (defining content-related evidence) and standard 1.6 and 1.7 (describing content-related validation studies).

See APA Standards at p. 9 (defining construct-related evidence) and standard 1.8, 1.9, and 1.10 (describing construct-related validation studies).

Larry P. v. Riles, 495 F. Supp. 926 (N.D. Cal. 1979), aff'd, 793 F.2d 969 (9th Cir. 1984)(State of California, requesting approval from the court to use standardized IQ test or assessment procedures for the purpose of placing black children in EMR classes, required to, among other things, provide statistics showing the mean scores of blacks and whites on the test or assessment procedure and information supporting the validity of the test or assessment procedure for the purpose of identifying and placing students in EMR classes).

Debra P. v. Turlington, 730 F.2d 1405 (11th Cir. 1984)(court approved validity study which consisted of a number of surveys and site visits that analyzed whether the students had received the instruction necessary for them to have mastered the skills that were being tested).

See also Code of Fair Testing Practices in Education at A-1 and A-7 (encouraging test developers to describe the population for which the test is appropriate and encouraging test users to select tests appropriate for the testing purpose and population of test takers).

American Psychological Association Standards for Educational and Psychological Testing (1985) (APA Standards) at standard 1.16 (permitting criterion-related validation inferences to be drawn from a set of prior studies, where "local validation evidence" is not available, depending on the degree of similarity between the test or assessment procedure use and validation sample); and at pp. 12-13 (describing the concept of differential prediction).

2. Reliability

APA Standards at pp. 19-20 (discussing reliability and error of measurement).

3. Cutoff scores

Evans v. City of Evanston, 881 F.2d 382 (7th Cir. 1989) (while test was valid for the job, cutoff score was set one standard deviation above the mean; the court rejected this because there was no attempt to connect the score to level of performance: "...the ability to perform firefighting tasks adequately depends not on relative but on absolute test performance.").

Richardson v. Lamar County Bd. of Education, 729 F. Supp. 806 (M.D. Ala. 1989) (passing score rejected because of the lack of any relationship to actually measuring competence; instead, based on what was "politically acceptable").

APA Standards at standard 6.9 (when a specific cut-off score is used to select, classify or certify test takers, the method and rationale for setting that cut score, including any technical analyses, should be presented in a manual or report).

C. Alternatives With Less Disparate Impact

Albemarle Paper Co. v. Moody, 422 U.S. 405, 425 (1975)(in Title VII case challenging use of employment test that had a disparate impact, court stated that

employee can still prevail, even if test is valid, if other tests or selection devices with less disparate impact would serve the employer's interests).

NAACP v. State of Georgia, 775 F.2d 1403 (11th Cir. 1985) (considering less discriminatory alternatives in Title VI education context).

Sharif v. New York State Education Department, 709 F.Supp. 345 (S.D.N.Y. 1989) (where use of the SAT had a disparate impact based on sex in awarding state merit scholarships, court approved awards being based on a combination system - using both grade point averages and SAT scores - as a legally sufficient alternative to sole reliance on the SAT; court found that, compared with sole reliance on the SAT, combination system would better advance the state's goal of awarding high school performance and would better provide all students with an equal opportunity to compete for prestigious state scholarships; court found that feasibility argument about the combination system advanced by the state education agency lacked merit).

Bridgeport Guardians, Inc. v. City of Bridgeport, 735 F. Supp. 1126, 1136-1137 (D. Conn. 1990), aff'd, 933 F.2d 1140, 1148 (2nd Cir. 1991), cert. denied, 502 U.S. 924 (1991) (where making promotion decisions for police department on a strict rank-order basis based on examination scores had a disparate impact on racial minority candidates, district court rejected the use of video simulations as an alternate selection criteria on the basis that they have generally not increased the relative standing of minority candidates and because it substantially adds to the cost of the promotion process; district court approved use of banding as an alternate selection criteria as it found that there is no evidence that any added burdens that a banding analysis would impose are more than minimal; appeals court upheld the use of banding noting, based on testimony from the city's industrial psychologist, that small variances in the examination scores did not indicate that there were real differences in the qualifications of the candidates).

Brunet v. City of Columbus, 1 F.3d 390, 411-412 (6th Cir. 1993), cert. denied, 510 U.S. 1164 (1994) (where making hiring decisions for fire fighter positions on a strict rank-order basis based on the results of a physical capability test (PCT) and a cognitive ability test (CAT) had a disparate impact on the basis of sex, although the appeals court rejected both alternatives presented by the plaintiffs, the appeals court found error because there was no indication in the record that the district court required the city, pursuant to the Uniform Guidelines on Employee Selection Procedures, to conduct its own investigation of viable alternatives with lesser or no impact on female applicants before implementing the strict rank-order process; the appeals court also found nothing in the record that requires the CAT and the PCT to be weighted equally; the appeals court also

indicated that the city should be required to demonstrate why the CAT, which arguably is more predictive than the PCT, should not be weighted more than the PCT, noting that the change might result in a lesser disparate impact on women).

FairTest v. College Entrance Examination Board and Educational Testing Service, OCR Case No. 02-94-2048 (where use of the PSAT had disparate impact based on sex in selecting National Merit Scholarship semi-finalists, recipients agreed to modify the test to include a writing skills component and to study whether academic records could also be considered).

III. Different Treatment Analysis

As with other claims of race, national origin, or sex discrimination under Title VI and Title IX, a different treatment analysis may apply when a policy or practice regarding testing or assessment is being applied differently by an educational institution to different groups of students because of their race, national origin, or sex. This is the touchstone of what is a classic violation of Title VI and Title IX and their implementing regulations.

Where there is direct evidence that an educational decision was made based on race, national origin, or sex, a prima facie violation of Title VI or Title IX has been established. The recipient then has the burden of establishing a legitimate reason (an affirmative defense) for the different treatment, such as showing that the disparate treatment was the result of a valid affirmative action plan. See Regents of the University of California v. Bakke, 438 U.S. 265, 320 (1978) (Although the U.S. Supreme Court found that an applicant to medical school had been discriminated against on the basis of race under an unlawful admissions process, five justices agreed that the portion of the lower court decision that enjoined the university from ever considering the race of any applicant should be reversed. Justice Powell recognized that the "State has a substantial interest that legitimately may be served by a properly devised admissions program involving the competitive consideration of race or ethnic origin.") But see Hopwood v. Texas, 78 F.3d 932 (5th Cir. 1996), cert. denied, 518 U.S. 1033 (1996) (The Fifth Circuit held that a university's interest in achieving a diverse student body can never constitute a compelling governmental interest justifying the use of race in university admissions selections.)

Note that there need not be direct proof of intentional discrimination in order to make a disparate treatment case. "In most disparate treatment cases, intent to discriminate is established inferentially, through circumstantial evidence." Equal Employment Opportunity Commission's "Revised Enforcement Guidance on Recent Developments in Disparate Treatment Theory," July 14, 1992, Number N 915.002. at 2. The basic elements of a different treatment case in which there is no direct evidence of discrimination were set out by the U.S. Supreme Court in McDonnell Douglas Corp. v. Green, 411 U.S. 792 (1973), a Title VII employment case. See also United States Postal Service Board of Governors v. Aikens, 460 U.S. 711 (1983); Texas Department of Community Affairs v. Burdine, 450 U.S. 248 (1981).

IV. Analysis Where Prior Dual System

United States v. Fordice, 112 S.Ct. 2727 (1992) (Mississippi's admission policy which required higher ACT scores for historically white public universities than for historically black public universities was constitutionally suspect as it was originally enacted by historically white universities to maintain prior dual system of higher education. States and schools districts that have operated a dual system have an obligation to dismantle the prior de jure segregated system.).

Knight v. Alabama, 14 F.3d 1534, 1540-42 (11th Cir.1994) (A recipient's "burden of proving that [less discriminatory] alternatives are impractical or educationally unsound is a heavy one.").

Debra P. v. Turlington, 644 F.2d 397, 407 (5th Cir. 1981) (In an equal protection challenge to the requirement that students pass a functional literacy test or assessment procedure to receive a high school diploma, trial court instructed to consider whether disproportionate failures of black students could be attributed, in part, to unequal education received during period of dual school system).

TAB A

**SAMPLE QUESTIONS FOR
EVALUATING EVIDENCE OF EDUCATIONAL NECESSITY AND
DETERMINING WHETHER THERE ARE PRACTICABLE ALTERNATIVES WITH
LESS DISPARATE IMPACT**

Once it has been determined that a test or assessment procedure has a disparate impact on students of a particular race, national origin, or sex, there should be an inquiry into the test or assessment procedure's educational necessity. Educational necessity means that the challenged test or assessment procedure is valid and reliable for the purpose for which it is being used. There should then be an inquiry to determine whether there are any practicable alternatives to the test or assessment procedure that are available, which would (i) substantially serve the educational purpose identified by the educational institution, (ii) be valid and reliable for that purpose, and (iii) have a lesser disparate impact.

As stated in the resource guide, evidence of the general reputation of a test or assessment procedure is not sufficient to establish validity. Also, a manual developed by the test or assessment procedure publisher is not presumptive evidence of validity. Rather, the use of a test or assessment procedure may be supported by validity studies of the same test conducted by the school, other schools, test publishers or distributors, or professional researchers.

The following questions are designed to assist OCR staff and members of the educational community in evaluating evidence of educational necessity and in determining whether there are any practicable alternatives to the test or assessment procedure which would meet the educational institution's educational needs and have a lesser disparate impact:

1. What test or assessment procedure is the educational institution administering?
2. For what purpose is the test or assessment procedure being administered (purpose should be stated in specific terms, i.e., predicting grades in algebra, rather than in general terms, i.e., measuring intelligence of seventh-graders)?
3. What is the educational institution's justification for the purpose for which the test or assessment procedure is being used? If the school cannot show that the test is educationally necessary, the use of the test or assessment procedure will be in violation of Title VI or Title IX, as appropriate.
4. For what purpose was the test or assessment procedure developed (if it is clear from preliminary evidence)? Is the school using the test or assessment procedure for this purpose? If not, the use of the test or assessment procedure cannot be justified as educationally necessary and will be in violation of Title VI or Title IX, as appropriate.

5. Is the test or assessment procedure being used as the sole criterion for making an educational decision? If so, was the test or assessment procedure designed to be used this way? If not, the test or assessment procedure used cannot be justified as educationally necessary and there is a violation of federal law.

6. What is the racial, ethnic, and/or sex composition of the test or assessment procedure-taking population?

7. Does the school have evidence that it has developed or that has been developed by the test or assessment procedure publisher that the test or assessment procedure is valid and reliable for the purpose for which it is being used?

a. What is the form of the evidence, i.e., study or report? Raw data or very general information is not acceptable evidence of the validity of a test or assessment procedure.

b. Who conducted the study, i.e., the school, another school, the test or assessment procedure developer, an independent researcher?

c. How long ago was the study conducted?

d. Is there evidence under professionally accepted standards that the test or assessment procedure is valid and reliable?

e. For what use was the test or assessment procedure validated? Is this use identical to, or the professionally accepted equivalent of, the purpose for which the school is using the test or assessment procedure?

f. Was the test or assessment procedure administered and scored properly?

i. Were all the students treated the same way as to how the test or procedure was administered or scored?

ii. Did school officials clearly articulate to school personnel who administered, scored, or interpreted the results the construct(s) or variable(s) the procedure was designed to measure?

iii. Did those who administered, scored, or interpreted the results have the appropriate skills to perform these functions adequately? For example, are those who interpret the scores able to understand and interpret commonly reported scores, such as percentile ranks, standard scores, stanines, normed curve equivalents, and grade equivalents (as appropriate to the particular test)? Are

they able to understand and interpret commonly reported summary indexes such as central tendency measures, estimates of reliabilities, and standard errors of measurement (again, as appropriate to the particular test)?

8. What type of statistical test was used in the validation study upon which the school relies?

9. If the school has empirical evidence that the test or assessment procedure is valid and reliable for the purpose for which it is being used, has the school looked at whether there are practicable alternatives to the test or assessment procedure with less impact that would substantially serve the school's stated purpose and that are valid and reliable for that purpose?

a. How great was the disparate impact caused by the challenged test or assessment procedure?

b. What alternatives has the school looked at?

c. Did the criteria or criterion have less impact than the challenged test or assessment procedure?

d. If the school did not look at alternatives, what is the reason for this decision?

e. If the school did look at alternatives, but chose not to use them, what is the reason for this decision?

10. Are there one or more criteria which, either alone, or in combination with other criteria, would have less impact, serve the school's educational purpose, and be valid and reliable for that purpose?

11. Does the school use a "cutoff score" on the test or assessment procedure which determines whether a student receives an educational benefit or opportunity? If so:

a. Does the cutoff score have a disparate impact on students of a particular race, national origin, or sex?

b. Is the cutoff score being used as the sole criterion for making an educational decision? If so, was the cutoff score designed to be used in this way? If not, the use of the cutoff score cannot be justified as educationally necessary and there is a violation of federal law.

c. Does the school have evidence regarding the method and rationale for setting the cutoff score?

d. Does the evidence provided under c. reflect a systematic process that evidences the good faith exercise of professional judgment?

TAB B

**LIST OF POLICY AND TECHNICAL ASSISTANCE DOCUMENTS
ADDRESSING THE CLASSIFICATION OF STUDENTS AND THE
NONDISCRIMINATORY PROVISION OF SERVICES¹⁶**

I. Provision of Educational Services to Limited-English Proficient Students

U.S. Department of Education, Office for Civil Rights. (1991). "Policy Update on Schools' Obligations Toward National Origin Minority Students With Limited-English Proficiency."

This policy update is primarily designed for use by OCR staff to determine whether schools are complying with their obligation under Title VI to provide any alternative language programs necessary to ensure that national origin minority students with limited-English proficiency (LEP) have meaningful access to schools' programs. The policy update provides additional guidance for applying the May 1970 and December 1985 memoranda, which are described below.

U.S. Department of Education, Office for Civil Rights. (1985). "The Office for Civil Rights' Title VI Language Minority Compliance Procedures."

This memorandum provides a description of the procedures followed by OCR in making determinations of compliance with Title VI as regards the treatment of LEP students in educational programs that receive federal financial assistance from the Department of Education.

U.S. Department of Education, Office for Civil Rights. (1970). "Identification of Discrimination and Denial of Services on the Basis of National Origin," 35 Fed. Reg. 11595 (May 1970 Memorandum)

This memorandum was designed to clarify the policy of the Department of Health, Education and Welfare, now the Department of Education, on issues concerning the responsibility of school districts to provide equal educational opportunity to LEP students. This memorandum was a foundation for the

¹⁶ The policy and technical assistance documents that are listed were either developed by the U.S. Department of Education's Office for Civil Rights or were developed by other organizations, where the U.S. Department of Education was a contributing party.

U.S. Supreme Court's decision in Lau v. Nichols, 414 U.S. 653 (1974). The memorandum was affirmed by the U.S. Supreme Court in the Lau decision.

August D. & Hakuta, K. (Eds.). (1997). Improving America's Schools for Language-Minority Children: A Research Agenda. Committee on Developing a Research Agenda on the Education of Limited-English-Proficient and Bilingual Students, National Research Council. Washington, D.C.: National Academy Press.

The National Research Council (NRC) completed an extensive study of LEP students, which is summarized in this report. The report provides a review of the state of knowledge regarding the education of LEP students and identifies a research agenda that will address key gaps in present knowledge on the topic. Among the topics covered in the report are student assessment and program evaluation.

U.S. Department of Education, Office for Civil Rights. (1991). "Technical Assistance Resource Package on the Provision of Equal Educational Opportunity to National Origin Minority and Native American Students Who Are Limited-English-Proficient."

This resource package focuses on the provision of equal educational opportunities to national origin minority and native American students who are limited-English-proficient. The package contains: 1) information about the history and importance of issues pertaining to the education of LEP students; 2) technical definitions and explanations of prominent educational approaches to teaching LEP students; 3) a summary of case law regarding civil rights requirements for educating LEP students; 4) a summary of compliance review letters of findings; 5) summaries and listings of major research publications, studies, and reports addressing LEP issues; and 6) a listing of major professional and beneficiary organizations involved in gathering information about LEP issues.

II. Minority Students and Special Education

U.S. Department of Education, Office for Civil Rights. (1995). "Minority Students and Special Education - Legal Approaches for Investigations."

This memorandum provides an overview of the legal theories and approaches to be employed in OCR investigations regarding the disproportionate representation of minority students in special education.

Markowitz, J., García, S., & Eichelberger, J.H. (1997). Addressing the Disproportionate Representation of Students from Racial and Ethnic Groups in Special Education: A Resource Document. Alexandria, VA: Project FORUM, National Association of State Directors of Special Education.

This document is intended to enhance the knowledge base of technical assistance providers to enable them to provide more effective technical assistance and guidance to state and local education personnel who are addressing the problem of disproportionate representation. The document includes a compilation of approaches that have the potential for effectively preventing and correcting disproportionate representation, an annotated bibliography of print resources, and a list of individuals who are knowledgeable about one or more of the main topic areas presented in this document.

Morison, P., White, S.H., & Feuer, M.J. (Eds.) (1996). The Use of IQ Tests in Special Education Decision Making and Planning: Summary of Two Workshops. Board on Testing and Assessment, National Research Council. Washington, D.C.: National Academy Press.

To assist the U.S. Department of Education, the Board on Testing and Assessment convened two workshops to facilitate an examination and discussion of research evidence regarding the uses of intelligence tests (IQ) tests in special education placement decisions, with particular focus on mental retardation and learning disabilities. The workshops had the following objectives: 1) to provide an overview of legal, policy, and measurement issues regarding the use of IQ tests in special education; 2) to examine issues related to the validity and fairness of IQ testing for classification and placement of students in special programs, with emphasis on potential adverse effects on minority students; and 3) to explore some possible alternative assessment methods that could be used in combination - or as substitutes for - traditional IQ tests. The summary report provides a synthesis of the key themes and ideas discussed at the workshops.

TAB C

VALIDITY, A GLOSSARY

Introduction

"Validation is the most important consideration in test evaluation. ... Test validation is the process of accumulating evidence to support ... inferences (made from test scores)." (American Psychological Association Standards for Educational and Psychological Testing (1985). (APA Standards at p. 9).

In general the inference that a test or assessment procedure is valid is justified when the research evidence indicates the following is true:

1. the test or assessment procedure measures the construct (characteristic, property, skill, ability, capacity, or behavior) it was intended to measure;
2. the test or assessment procedure is used in a correct and appropriate manner, with regard to testing setting, testing procedure (including the qualifications of the test-giver and the manner in which the test is given), tested sample of people (e.g., using a test validated for adults to assess children would be improper); and
3. the inferences drawn from the resulting test or assessment procedure data are appropriate and correct.

"Traditionally, the various means of accumulating validity evidence have been grouped into categories called *content-related*, *criterion-related*, and *construct-related evidence of validity*. ... These categories are convenient ... but the use of the category labels does not imply there are distinct types of validity or that a specific validation strategy is best for each specific inference or test use ... An ideal validation includes several types of evidence, which span all three of the traditional categories ... Professional judgment should guide the decisions regarding the forms of evidence that are most necessary and feasible in light of the intended uses of the test and any likely alternatives to testing." (APA Standards at p. 9).

Evidence can be gathered by use of such particular statistical techniques as correlation and regression analyses with test items or scores and other test or non-test variables, factor analysis, item response theory (IRT) and other level of difficulty techniques, and differential item functioning (DIF) analyses. It might also be gathered by the systematic judgment evaluation of individual responses, or a formal evaluation of one or a number of test construction, implementation, or data analytic processes. (Aiken, 1994; Holland and Wainer, 1993; Wainer and Braun, 1988).

Regardless of technique, evidence is obtained which demonstrates that information derived from the assessment is accurately reflective of what is supposed to be measured. This is done by designing investigations which focus on convergence, or high positive relationships, and discriminatory analyses, which seek to demonstrate divergence between the assessment information and related, but distinct, variables.

Construct-Related Validity¹⁷

Construct-related evidence of validity is "evidence that supports a proposed construct interpretation of scores on a test based on theoretical implications associated with the construct label." (APA Standards at p. 90). The construct-related validity of a test or assessment procedure is the extent to which the assessment may be said to measure a theoretical construct or trait. (Aiken, 1994; Anastasi, 1988; Groth-Marnat, 1990). "Reasoning ability, spatial visualization, and reading comprehension are constructs ...The construct of interest for a particular test should be embedded in a conceptual framework ... The conceptual framework specifies the meaning of the construct, distinguishes it from other constructs, and indicates how measures of the construct should relate to other variables." (APA Standards at pp. 9 - 10).

Any data throwing light on the nature of the trait under consideration and the conditions affecting its development and manifestations represents appropriate evidence for this validation. (Anastasi, 1988). The process of compiling construct-related evidence for test validity starts with test development and continues until the pattern of empirical relationships between test scores and other variables clearly indicates the meaning of the test score. "[V]alidating inferences about a construct also requires paying careful attention to aspects of measurement such as test format, administration conditions, or language level, that may affect test meaning and interpretation materially." (APA Standards at p. 10).

Construct validity, which is the most general type of validity, is not determined by a single way or by one investigation. Rather it involves a network of investigations and other procedures designed to determine whether an assessment instrument that purportedly measures a certain variable is actually doing its job. (Aiken, 1994; Groth- Marnat, 1990).

¹⁷ As indicated throughout the Resource Guide, OCR relies upon generally accepted existing professional standards when evaluating the validity and reliability of a test or assessment procedure. However, it should be noted that there is a trend among measurement theorists to consider construct validity to be the fundamental, unifying framework for conceptualizing validity evaluations (see, e.g., Shepard, 1993, and Wainer and Braun, 1988). Under this framework, since all validation is subsumed under construct validation, there are not different types of validity. Also, as part of this framework, various sources of evidence, including, but not necessarily limited to, content-related evidence, criterion-related evidence, and prediction-related evidence, can be, and usually are, used to evaluate the degree to which score-based inferences and actions are supported. Some testing and assessment experts include such additional evidence as the consequences of test use on individuals and groups in society as part of the construct validity framework (Messick, 1989).

Content-Related Evidence of Validity

Content-related evidence of validity is "evidence that shows the extent to which the content domain of a test is appropriate relative to its intended purpose." (APA Standards at p. 90). "In general, content-related evidence demonstrates the degree to which the sample of items, tasks, or questions on a test are representative of some defined universe or domain of content. ... For some educational decisions, it is important to determine the agreement between the test and the curricular or instructional domains it is meant to cover. ... [I]nferences about content are linked to test construction as well as to establishing evidence of validity after a test has been developed and chosen for use." (APA Standards at pp. 9 - 10).

Content validity is concerned with whether the content of the test or assessment procedure elicit the range of responses representing the entire domain or universe of skills, understandings, or other behaviors that the test or assessment procedure was supposed to measure (Gregory, 1992; Aiken, 1994; Anastasi, 1988). "Methods classed in the content-related category should often be concerned with the ... construct underlying the test as well as the character of test content. There is often no sharp distinction between test content and test construct." (APA Standards at p. 11.).

Criterion Validity¹⁸

Criterion-related evidence of validity "demonstrates that test scores are systematically related to one or more outcome criteria." (APA Standards at p. 11). This type of validity evidence is produced by relating scores on the test or assessment procedure to performance criterion measures, standards, or variables (Aiken, 1994). According to Anastasi (1988), a criterion is a direct and independent measure of that which the test is designed to predict. Criterion-related procedures indicate the effectiveness of the assessment where performance on the test is checked against a criterion. "The choice of the criterion and the measurement procedures used to obtain criterion scores are of central importance. Logically, the value of a criterion-related study depends on the relevance of the criterion measure that is used." (APA Standards at p. 11).

Two types of criterion-related evidence are those obtained from investigations which focus on prediction and those which focus on concurrent relationships. "Two designs for obtaining criterion-related evidence - predictive and concurrent - can be distinguished. A predictive study obtains information about the accuracy with which early test data can be used to estimate criterion scores that will be obtained in the future. A concurrent study serves the same purpose, but it obtains prediction and criterion information simultaneously. Predictive

¹⁸ Correlations between a test and a criterion are validity coefficients. A study of predictive or concurrent validity is nearly always reported in terms of a correlation coefficient. Cronbach, L.J. (1990). *Essentials of Psychological Testing* (5th ed.). New York:Harper Collins Publishers, Inc.

studies are frequently, but not always, preferable to concurrent studies of selection tests for education or employment, whereas concurrent evidence is usually preferable for achievement tests, tests used for certification, diagnostic clinical tests, or for tests used as measures of a specified construct." (APA Standards at p. 11).

TAB D

BIBLIOGRAPHY

Selected Validity References

Aiken, L.R. (1994). Psychological Testing and Assessment, (8th ed.). Needham Heights, MA: Allyn and Bacon.

American Educational Research Association, American Psychological Association, National Council on Measurement in Education. (1985). Standards for Educational and Psychological Testing.

Anastasi, A. (1988). Psychological Testing (6th ed.). New York: Macmillan Publishing Company.

Anastasi, A. (1986). "Evolving Concepts of Test Validation." Annual Review of Psychology, 37, 1-15.

Carver, C.S., & Schier, M.F. (1996). Perspectives on Personality (3rd ed.). Needham Heights, MA: Allyn & Bacon.

Cattell, R.B. (1964). "Validity and reliability: A proposed more basic set of concepts." Journal of Educational Psychology, 55(1), 1-22.

Corsini, R.J., & Marsella, A.J. (1983). Personality Theories, Research, and Assessment. Itasca, IL: F.E. Peacock Publishers.

Cronbach, L.J. (1990). Essentials of Psychological Testing, (5th ed.). New York: Harper Collins Publishers, Inc.

Cureton, E.E. (1968). International Encyclopedia of the Social Sciences. U.S.A.: Crowell, Collier and Macmillan.

Ebel, R.L., & Frisbie, D.A. (1991). Essentials of Educational Measurement (5th ed.). Englewood, NJ: Prentice Hall.

English, H.B. & English, A.C. (1962). A Comprehensive Dictionary of Psychological and Psychoanalytic Terms. U.S.A.: David McKay Company, Inc.

- Funder, D.C. (1997). The Personality Puzzle. New York: W.W. Norton & Company.
- Gatewood, R., & Field, H. (1994). Human Resource Selection. Fort Worth: Dryden Press.
- Gregory, R.J. (1992). Psychological Testing: History, Principles, and Applications. Needham Heights, MA: Allyn and Bacon.
- Groth-Marnat, G. (1990). Handbook of Psychological Assessment (2nd ed.). New York: John Wiley and Sons.
- Hartigan, J., & Wigdor, A. (1989). Fairness in Employment Testing. Committee on the General Aptitude Test Battery, National Research Council. Washington, DC: National Academy Press.
- Holland, P.W. & Wainer, H. (Eds.). (1993). Differential Item Functioning. Hillsdale, NJ: Lawrence Erlbaum Associates.
- McCloskey, D., & Ziliak, S. (1996). "The standard error of regressions." Journal of Economic Literature, 37, 97-114.
- Meir, R. (1994). The Chronic Crisis in Psychological Measurement and Assessment: A Historical Survey. New York: Academic Press.
- Messick, S. (1989). "Validity." In R.E. Linn (Ed.), Educational Measurement (3rd ed.). New York: Macmillan.
- National Research Council. (1998). High Stakes Testing for Tracking, Promotion, and Graduation. Washington, D.C.: National Academy Press.
- Office of Technology Assessment. (1992). Testing in American Schools: Asking the Right Questions. OTA-SET-519. Washington, DC: U.S. Government Printing Office.
- Reilly, R., & Warech, M. (1994). "The validity and fairness of alternative to cognitive tests." In H. Wing & B. Gifford (Eds.), Policy Issues in Employment Testing. Boston: Kluwer.
- Rust, J., & Golombok, S. (1989). Modern Psychometrics: The Science of Psychological Assessment. New York: Routledge.
- Shepard, L.A. (1993). "Evaluating Test Validity." Review of Research in Education, 19, 405-450.
- Travers, B.E. (Ed.). (1996). The Gale Encyclopedia of Science. Detroit: Gale Research.
- Wainer, H., & Braun, H. (Eds.). (1988). Test Validity. Hillsdale, NJ: Hove and London.

**Selected References Regarding Appropriate Test Use In The Context of
Standards-Based Reform**

Coleman, A.L., "Excellence and Equity in Education: High Standards for High-Stakes Tests." 6 Virginia Journal of Social Policy & the Law 81(1998).

Federation for Children with Special Needs. (1998). All Kids Count.

Hansche, L.N. (1998). Handbook for the Development of Performance Standards: Meeting the Requirements of Title I. Washington, D.C.: Council of Chief State School Officers and the U.S. Department of Education.

Jaeger, R.M. & Tucker, C.G. (1998). Analyzing, Disaggregating, Reporting, and Interpreting Students' Achievement Test Results: A Guide to Practice for Title I and Beyond. Washington, D.C.: Council of Chief State School Officers.

Kopriva, R.J. & Saez, S.M. (1997). Guide for Scoring Open-Ended Mathematics Responses of LEP Students. Washington, D.C.: Council of Chief State School Officers.

Kopriva, R.J. (in press). Ensuring the Accurate Assessment of LEP Students: A Practical Guide for Test Development. Washington D.C.: Council of Chief State School Officers.

Linn, R.L. & Herman, J.L. (1997). A Policymaker's Guide to Standards-Led Assessment. Denver, CO: Education Commission of the States.

National Research Council. (1999). High Stakes Testing for Tracking, Promotion, and Graduation. Washington, D.C.: National Academy Press.

Tucker, M. & Coddling, J. (1998). The State of Standards: Powerful Tool or Symbolic Gesture? Washington, D.C.: National Center on Education and the Economy.
U.S. Department of Education, Office of Educational Research and Improvement. (1997). Assessment of Student Performance.

U.S. Department of Education, Office of Elementary and Secondary Education. (in progress). Toolkit for the Assessment of Language Minority Students.

U.S. Department of Education, Standards, Assessments, and Accountability. (1997).
<http://www.ed.gov/pubs/legsregs.html#policy>. <http://www.ed.gov/offices/OESE/policy.html>

Attachment - A



UNITED STATES DEPARTMENT OF EDUCATION
WASHINGTON, D.C. 20202

SEP 29 1997

Dear Colleague:

We are writing to you today to highlight the importance of including students with disabilities in all educational reform activities and, in particular, in statewide assessment systems. As you know, President Clinton has announced a bold, national education initiative which includes the goal of learning to challenging and clear standards of achievement for all students, including students with disabilities. In his 1997 State of the Union address, the President announced a ten-point call to action including rigorous, voluntary national tests in reading and math embodying national standards, teaching every student to read independently by the end of the third grade, and increased accountability in public education.

Assessment is an integral aspect of accountability. Assessment systems have varied purposes. Whatever the focus of the particular assessment system - program evaluation, school and staff accountability or measuring student progress - assessments provide valuable information which benefits individual students, either directly, such as in the measurement of individual progress against standards, or indirectly, such as in evaluating programs. Given the emphasis on assessment in recent educational reform efforts, including State and Federal legislation linking assessment and school accountability, it is of utmost importance that students with disabilities be included in the development and implementation of assessment activities. Too often, in the past, students with disabilities have not fully participated in State and district assessments only to be short-changed by the low expectations and less challenging curriculum that may result from exclusion.

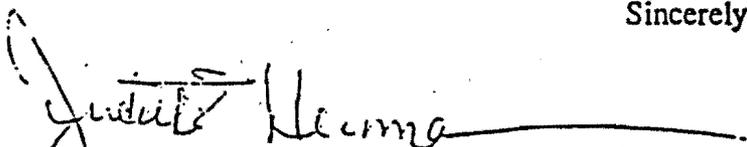
Given the benefits that accrue as a result of assessment, exclusion from assessments based on disability generally would not only undermine the value of the assessment but also violate Section 504 of the Rehabilitation Act of 1973 (Section 504), which prohibits exclusion from participation of, denial of benefits to, or discrimination against, individuals with disabilities on the basis of their disability in Federally-assisted programs or activities. 29 U.S.C. 794. Similarly Title II of the Americans with Disabilities Act (ADA) of 1990 provides that no qualified individual with a disability shall, by reason of such disability, be excluded from participation in or be denied the benefits of the services, programs, or activities of a public entity, or be subjected to discrimination by such entity. 42 U.S.C. 12132.

alternate assessments, if it can be reported in a statistically sound manner and would not result in disclosure of performance results identifiable to individual children. The reports must be provided with the same frequency and in the same level of detail as the State's reports on the assessment of nondisabled children. For assessments conducted after July 1, 1998, data relating to the performance of children with disabilities in regular assessments is required to be disaggregated. For those assessments conducted prior to July 1, 1998, the data for children with disabilities participating in regular assessments, is only required by IDEA to be disaggregated if the State requires disaggregation. Section 612(a)(17)(B); 111 Stat. 67-68.

The Office of Special Education Programs within OSERS has a cooperative agreement with the National Center on Educational Outcomes (NCEO) at the University of Minnesota to study and provide information on including students with disabilities in statewide and other assessments. We have enclosed a brochure on the NCEO, which may be contacted for more information.

As we work together to reform our educational system, we must ensure that all children, including students with disabilities, are part of that reform. Including students with disabilities in the development and implementation of assessments is a vital step towards providing access to the general curriculum and learning to challenging standards.

Sincerely,



Judith E. Heumann
Assistant Secretary for
Special Education and
Rehabilitative Services



Norma V. Cantú
Assistant Secretary for
Civil Rights

Attachment - B



Office of the Attorney General

Washington, D. C. 20530

July 14, 1994

MEMORANDUM FOR HEADS OF DEPARTMENTS AND AGENCIES
THAT PROVIDE FEDERAL FINANCIAL ASSISTANCE

FROM:

THE ATTORNEY GENERAL

SUBJECT:

Use of the Disparate Impact Standard in
Administrative Regulations Under Title VI
of the Civil Rights Act of 1964

This month marks the 30th anniversary of the passage of Title VI of the Civil Rights Act of 1964 (42 U.S.C. §§2000d to 2000d-6), which prohibits discrimination on the basis of race, color, or national origin in programs and activities that receive Federal financial assistance. The anniversary of this landmark legislation is a fitting time to remind agencies that administrative regulations implementing Title VI apply not only to intentional discrimination but also to policies and practices that have a discriminatory effect. In Guardians Association v. Civil Service Commission, 463 U.S. 582 (1983), the Supreme Court held that while Title VI itself requires proof of discriminatory intent, agencies may validly adopt regulations implementing Title VI that also prohibit discriminatory effects. Nearly all agencies have adopted such regulations. In Alexander v. Choate, 469 U.S. 287 (1985) (construing Section 504 of the Rehabilitation Act of 1973), a unanimous Supreme Court restated the holding in Guardians that disparate impact violations could be addressed through regulations implementing Title VI.

This Administration will vigorously enforce Title VI. As part of this effort, and to make certain that Title VI is not violated, each of you should ensure that the disparate impact provisions in your regulations are fully utilized so that all persons may enjoy equally the benefits of federally financed programs.

Enforcement of the disparate impact provisions is an essential component of an effective civil rights compliance program. Individuals continue to be denied, on the basis of their race, color, or national origin, the full and equal opportunity to participate in or receive the benefits of programs assisted by Federal funds. Frequently discrimination results from policies and practices that are neutral on their face but

have the effect of discriminating. Those policies and practices must be eliminated unless they are shown to be necessary to the program's operation and there is no less discriminatory alternative.

Under Executive Order 12250, the Department of Justice is responsible for ensuring that funding agencies meet their responsibilities under Title VI. This Department is committed to productive and effective enforcement of the civil rights laws by each agency that extends Federal financial assistance. Facially neutral policies and practices that act as arbitrary and unnecessary barriers to equal opportunity must end. This was the goal of Title VI when it became law and it remains one of the highest priorities of this Administration.