

11-1-1980

Truth-in-Testing Legislation: A Brief for the Status Quo

Alan B. Asay

Follow this and additional works at: <https://digitalcommons.law.byu.edu/lawreview>



Part of the [Consumer Protection Law Commons](#), and the [Legislation Commons](#)

Recommended Citation

Alan B. Asay, *Truth-in-Testing Legislation: A Brief for the Status Quo*, 1980 BYU L. Rev. 902 (1980).

Available at: <https://digitalcommons.law.byu.edu/lawreview/vol1980/iss4/10>

This Comment is brought to you for free and open access by the Brigham Young University Law Review at BYU Law Digital Commons. It has been accepted for inclusion in BYU Law Review by an authorized editor of BYU Law Digital Commons. For more information, please contact hunterlawlibrary@byu.edu.

Truth-in-Testing Legislation: A Brief for the Status Quo

"[E]mpirical psychology," wrote Kant in 1786, "does not rise to the rank of a natural science because mathematics cannot be applied to the phenomena of inner sense" ¹ Later philosophers have taken a kinder view of the status of psychology, ² but the difficulty of arriving at numerical descriptors of the mind persists.

Measuring the mind is the job of the psychometrician, who uses instruments commonly called tests. His testing is evaluated according to its "validity," which in the jargon of the profession refers to the accuracy of the tests as measuring instruments; a test that measures what it purports to measure is said to be valid. ³ The scores of a perfectly valid test correlate exactly with the levels of whatever psychological phenomenon is being measured. ⁴

However, as Kant intimated, perfectly valid tests are an ideal rather than a reality. Validity rarely exceeds 0.60 (where one equals perfect validity), ⁵ and in practice, levels of validity are considerably lower. Among college admission tests, ⁶ for example, median validity coefficients range from 0.41 for the Scho-

1. I. KANT, *METAPHYSISCHE ANFANGSGRÜNDE DER NATURWISSENSCHAFT* (Riga 1786).

2. See, e.g., P. DIESING, *PATTERNS OF DISCOVERY IN THE SOCIAL SCIENCES* (1971); A. KAPLAN, *THE CONDUCT OF INQUIRY* (1964).

3. E. BURNS, *THE DEVELOPMENT, USE, AND ABUSE OF EDUCATIONAL TESTS* 131 (1979); R. EBEL, *ESSENTIALS OF EDUCATIONAL MEASUREMENT* 567 (1972).

4. A number of common criticisms of testing are included within the concept of validity. For example, the allegation that test results correlate with the race or family income of the test taker is really a validity question. A test that actually measures race or family income when it is designed to measure other factors is invalid unless there is an external correlation, independent of the test, between race or family income and any qualities validly measured by the test.

Similarly, the recent findings of the Federal Trade Commission, *FED. TRADE COMM., EFFECTS OF COACHING ON STANDARDIZED ADMISSION EXAMINATIONS* (1979), to the effect that pre-test "cramming" improves performance on the tests is also a validity problem. A test is supposed to measure certain long-term learning abilities not susceptible to cramming. Thus, if tests are vulnerable to cramming, they are less valid measures of the long-term abilities they seek to assess.

5. L. CRONBACH, *ESSENTIALS OF PSYCHOLOGICAL TESTING* 135 (3d ed. 1970).

6. As used in this Comment, the word "college" is to be understood in a broad sense and includes graduate, law, medical, and dental schools as well as undergraduate institutions.

lastic Aptitude Test to 0.29 for the Graduate Management Admission Test.⁷

Despite their imperfect validity, however, tests are used as a basis for making important decisions. Tests largely determine which students are placed in programs for the mentally retarded,⁸ which are denied diplomas due to their "functional illiteracy,"⁹ and which are to be admitted to college, graduate study, or professional school. Tests are also used in deciding who may enter a certain profession or be employed at a certain job. Thus, to a great degree, testing charts one's course through life. At every decisive point in our educational system, a test appears either to direct a student toward further education and a more promising profession or to shunt him into a career requiring less education.

Using partially invalid tests to determine whether educational opportunities will be granted causes these opportunities to be unjustifiably withheld from some.¹⁰ For example, there are undoubtedly students who are refused admission to the colleges they prefer because their test scores are inaccurately low, even though the same students would have been welcomed at the same institutions had their scores more accurately reflected their abilities.

This is a problem that is rightfully the subject of concern. It is manifestly unfair to deny an individual an opportunity because of a test's misrepresentation of his abilities. Legislatures have shared this concern which has been reflected in the "truth-

7. The median validity coefficient of the Scholastic Aptitude Test is 0.41; that of the Graduate Record Examination is 0.33; those of the Law School Admission Test and Graduate Management Admission Test are 0.36 and 0.29 respectively. These figures show the degree of correlation between test scores and the abilities to be measured. 1.0 represents a perfect, positive correlation, 0.0 no correlation, and -1.0 a perfect negative correlation. EDUCATIONAL TESTING SERVICE, TEST USE AND VALIDITY 16 (1980).

8. See *Parents in Action on Special Education v. Hannon*, No. 74 C 3586 (N.D. Ill. July 7, 1980); *Larry P. v. Riles*, 495 F. Supp. 926 (N.D. Cal. 1979); *Hobson v. Hansen*, 269 F. Supp. 401, 475 (D.D.C. 1967), *aff'd sub nom. Smuck v. Hobson*, 408 F.2d 175 (D.C. Cir. 1969); cf. *Lora v. Board of Educ.*, 456 F. Supp. 1211 (E.D.N.Y. 1978) (methods of classifying students as emotionally disturbed violated federally guaranteed rights).

9. See, e.g., *Debra P. v. Turlington*, 474 F. Supp. 244 (M.D. Fla. 1979). See generally Clague, *Competency Testing and Potential Constitutional Challenges of "Every Student"*, 28 CATH. U.L. REV. 469 (1979); Lewis, *Certifying Functional Literacy: Competency Testing and Implications for Due Process and Equal Educational Opportunity*, 8 J.L. & EDUC. 145 (1979).

10. Employment testing that discriminates against minorities has been the subject of considerable litigation. See, e.g., *Washington v. Davis*, 426 U.S. 229 (1976); *Griggs v. Duke Power Co.*, 401 U.S. 424 (1971).

in-testing" legislation recently enacted in California and New York.¹¹ Similar statutes are under consideration in a number of other states.¹² Nevertheless, even though a problem undeniably exists, there is good reason to doubt that legislation will be more effective than the status quo in solving it. This Comment will explore the various legislative possibilities in the testing area and compare them with present efforts to improve testing.

I. THE NEW YORK STATUTE

Like much other consumer-oriented legislation, the New York statute uses disclosure as a means of effecting reform.¹³ It contains two separate disclosure provisions. The first requires that a testing company send a requesting test taker a copy of his raw score, his answer sheet, and a list of correct answers.¹⁴ The

11. CAL. EDUC. CODE §§ 99150-99160 (Deering Supp. 1980); N.Y. EDUC. LAW §§ 340-347 (McKinney Supp. 1980-1981). Although test scores are used as criteria in making a number of crucial decisions in the course of a test taker's life, the New York and California statutes deal with only one of these decisions: college admission. Query whether this limitation is defensible in light of the fact that the validity of all testing is imperfect, regardless of the use to which it is put. Has college admission testing been singled out for legislative treatment because it is more important than testing which can classify an individual as being mentally retarded or functionally illiterate?

12. Letter from Allan Nairn to Alan Asay, (Sept. 16, 1980). The states listed by Nairn as considering truth-in-testing legislation are Connecticut, Illinois, Indiana, Louisiana, Maine, Maryland, Massachusetts, Michigan, Minnesota, Missouri, New Jersey, Ohio, and Pennsylvania. The bills before the U.S. House of Representatives are H.R. 3564, 96th Cong., 1st Sess. (1979), and H.R. 4949, 96th Cong., 1st Sess. (1979). As of this writing, both bills are still pending before the Committee on Education and Labor. The last action taken by the committee was a hearing commenced on July 31, 1979. [1979-1980] 2 CONG. INDEX (CCH) ¶ 35,017, 35,030.

13. The New York statute is the most imitated truth-in-testing statute, and will therefore be discussed in this Comment as typical. See Robertson, *Examining the Examiners: The Trend Toward Truth in Testing*, 9 J.L. & EDUC. 167, 180-84 (1980). However, the California statute differs from its New York counterpart in that it attempts to make testing companies more financially responsible. CAL. EDUC. CODE § 99154 (Deering Supp. 1980).

14. The relevant portions of the New York law provide:

2. Within ninety days after filing a standardized test pursuant to subdivision one of this section and for a period of not less than ninety days after the offer is made, the test agency shall provide to the test subject the opportunity to secure:

a. a copy of the test questions used to calculate the test subject's raw score;

b. a copy of the test subject's answer sheet, or answer record where there is no answer sheet, together with a copy of the correct answer sheet to the same test with questions used to calculate the test subject's raw score so marked; and

c. a statement of the raw score used to calculate the scores reported to the

second provision requires that a test company file with a government agency a list of questions used in scoring, the accompanying correct answers, and "any study, evaluation, or statistical report pertaining to a test."¹⁵ The documents thus filed become public records accessible under freedom of information laws.¹⁶ The statute neither provides aggrieved test takers a cause of action nor establishes any regulatory authority over the testing industry; rather, it relies solely on disclosure to accomplish its objective.

A. *Disclosure to the Test Taker: Identifying the Consumer*

The rationale underlying disclosure laws is that an informed consumer can better his lot in the marketplace because a competitive market offers several alternatives, which, when combined with accurate information, enable a consumer to select the option most advantageous to him. Certain information is difficult to obtain, however, and so it is sometimes necessary to resort to a governmental mandate to force disclosure in order to give the consumer an informed choice and thereby restore a healthy degree of competition to the market.¹⁷ Accordingly, lenders who are required to disclose actual interest rates and other hidden costs in lending agreements theoretically put the borrower in the position to choose the most favorable terms he can find, which in turn gives lenders an incentive to make their

test subject.

N.Y. EDUC. LAW § 342(2) (McKinney Supp. 1980-1981).

15. The pertinent provisions of the New York law state:

1. Within thirty days after the results of any standardized test are released, the test agency shall file or cause to be filed in the office of the commissioner;

a. a copy of all test questions used in calculating the test subject's raw score;

b. the corresponding acceptable answers to those questions; and

c. all rules for converting raw scores into those scores reported to the test subject together with an explanation of such rules.

....

7. Documents submitted to the commissioner pursuant to this section shall be public records

N.Y. EDUC. LAW § 342(1), (7) (McKinney Supp. 1980-1981).

16. The New York freedom of information law is N.Y. PUB. OFF. LAW §§ 84-90 (McKinney Supp. 1980-1981).

17. R. POSNER, *ECONOMIC ANALYSIS OF THE LAW* 273-76 (2d ed. 1977); Rhoades, *Reducing Consumer Ignorance: An Approach and its Effect*, 20 ANTITRUST BULL. 309 (1975).

terms more favorable.¹⁸

The New York law apparently assumes that the test taker is the consumer in the admissions testing situation since disclosure is directed to him. It is more accurate, however, to view the college using the test in its admission process as the consumer. It is true that the test taker pays for the test,¹⁹ but he has little use for the test scores. Although he is naturally interested in any assessment of himself, he does not take the test to gain greater self-understanding but rather to gain entry into college. The college exercises consumer decisionmaking power inasmuch as it is the college that decides which test, if any, it will use in its admission program.²⁰ The college chooses in the market without permitting the test taker to question the choice made.²¹ Thus, the college is the consumer, the testing company is the seller, and the test taker is merely a third party whose measurement is the commodity bought and sold.

Since the test taker has little opportunity to comment about what goes on between the college and the testing company, disclosure of information to the test taker has little effect on the test's validity. In directing disclosure to the test taker, the New York statute serves to inform someone who cannot act on the information provided. Although the test taker may learn by disclosure that he has been invalidly tested, he remains powerless to improve the validity of the test.

18. See 15 U.S.C. §§ 1601-1681 (1976). However, the practical effect of the truth-in-lending law has been disappointing, especially in connection with low-income consumers. R. POSNER, *supra* note 17, at 275; Kripke, *Gesture and Reality in Consumer Credit Reform*, 44 N.Y.U. L. REV. 1 (1969); Whitford, *The Functions of Disclosure Regulation in Consumer Transactions*, 1973 WIS. L. REV. 400.

19. Someone must bear the cost of an admissions program, and the college has no alternative but to charge the applicant. It would not be fair to expect students who have already been admitted to bankroll the applications of prospective students, especially those who are never admitted. Alumni and state governments also have little interest in financing the comparatively mundane admissions process.

20. This market choice could not be made by the test taker because allowing him to choose from among several tests would force the college to accept the results of any of the test he might select. Such a practice would make it impossible to compare different candidates since the tests not only employ different scoring scales but may also measure different abilities.

21. See *DePina v. Educational Testing Serv.*, 31 A.D.2d 744, 297 N.Y.S.2d 472 (1969); *K.D. v. Educational Testing Serv.*, 87 Misc. 2d 657, 386 N.Y.S.2d 747 (Sup. Ct. 1976). Here again, one would not have it any other way. Since the college, not the test taker, is the primary user of the measurement, the test taker should not be the one to determine how he is to be measured. To give the test taker control over the measurement procedure would amount to giving him control over one of the criteria of his own admission, a matter in which he has considerable personal interest.

On the other hand, colleges are in a position to improve validity. They are already organized into groups for testing purposes, such as the College Board or the Law School Admissions Council. The typical contract between a college group and the testing company provides for development and administration of the test by the company but leaves ownership of the test in the college group, which also retains the right to set general testing policy and determine what the testing company will be paid.²² Therefore, by acting collectively as owners of the test, they contract for the administration, scoring, and development of their test according to their specifications. The testing company is induced to cooperate because a contract for testing services could be awarded to any of several companies.²³

An example of the control a college group wields over its test is found in the creation of the new Medical College Admission Test (MCAT). Medical colleges decided to revise the old MCAT, and in consultation with practitioners, other interested parties, and testing experts appointed by the colleges, they developed an entirely new test. The colleges initiated and carried out the revision, retaining the testing companies in an advisory role only.²⁴

There is no evidence that the testing experts hoodwinked the learned doctors or that a governmental disclosure mandate would have been of any assistance to the medical colleges. Disclosure law theory presumes an ignorant or gullible consumer at the mercy of a seller with superior knowledge, but in the case of college admission testing, the theory does not apply. The college

22. The property rights of colleges in their tests are currently the subject of litigation arising out of the New York statute. In January, 1980, the Association of American Medical Colleges, which administers the Medical College Admission Test, was granted a preliminary injunction against application of the statute because of possible infringement or preemption by the Federal Copyright Act. *Association of Am. Medical Colleges v. Carey*, 482 F. Supp. 1358 (N.D.N.Y. 1980).

Colleges are already organized as groups of test users and bargain collectively with test companies. See, e.g., *Agreement Between the Law School Admission Council and Educational Testing Service (July 1, 1974)* [hereinafter cited as *Agreement*] reprinted in *Truth in Testing Act of 1979: The Educational Testing Act of 1979: Hearings on H.R. 3564 and H.R. 4949 Before the Sumcomm. on Elementary, Secondary, and Vocational Education of the House Comm. on Education and Labor, 96th Cong., 1st Sess. 88 (1979)* [hereinafter cited as *Hearings*].

23. The nation's largest publishers of standardized educational tests are Harcourt, Brace, Jovanovich Co.; McGraw-Hill Co.; IBM; Educational Testing Service (ETS); Houghton Mifflin Co. In addition, there are many smaller companies in the testing business. M. HOLMEN & R. DOCTER, *EDUCATIONAL AND PSYCHOLOGICAL TESTING* 33-69 (1972).

24. *Hearings*, supra note 22, at 70-77 (statement of John Cooper).

groups are not gullible, and if they lack information, they possess the resources to obtain it since they control the spending of the fees obtained from test takers.²⁵ Thus, the colleges, as consumers of testing, do not require the forced disclosure that might be needed in the case of a poorly educated applicant for a revolving charge account. Moreover, test companies have no incentive to deceive the colleges. The test company stands only to gain from efforts to develop a more valid test, and there is no reason to believe that they would not cooperate fully with the colleges' efforts to do so. Consequently, since colleges are the consumers of admission testing, it does not follow that disclosure by the testing company to the colleges needs to be required by statute.

B. Disclosure of Test Research

The second disclosure provision of the New York statute requires testing companies to make public their research on testing. However, this provision is of little value to the colleges, whose strong bargaining position already gives them ample access to the companies' research. Their contracts with the testing companies require timely reporting to the colleges of all relevant studies.²⁶ However, the New York statute's second disclosure provision does not direct disclosure exclusively to the college-consumer; rather, the entire public is made the object of disclosure. Apparently this provision dispenses altogether with the usual rationale for disclosure laws. Since disclosure is not directed to any identified consumer group but rather to the public as a whole, the rationale cannot be a matter of rectifying defects in the market structure. Instead, the purpose of the second provision seems to be to subject the testing establishment to whatever public criticism may be aroused through exposure of its research, with the hope that such criticism will give testers an incentive to test better.²⁷

Any incentive that might thus be gained, however, is damp-

25. Agreement, *supra* note 22.

26. *Id.*

27. Robertson, *supra* note 13, at 198-99. However, it is doubtful that the statute, as it stands, has enough teeth in it to bring about the disclosure it seeks. The statute, which calls for disclosure of "any study . . . pertaining to a test," is vague, and no regulations have been promulgated to specify more adequately what must be disclosed. Testers are therefore free to launder their reports, delete unfavorable findings, and even to delete the data and submit only bald conclusions.

ened by the fact that the tests' lack of validity is an open secret. It is common knowledge that the tests are partially invalid, and testing companies readily admit as much.²⁸ Since the disclosure of research about test invalidity will not reveal anything really new, it cannot create a significantly greater incentive than already exists.²⁹

Moreover, the colleges already have a strong incentive to improve testing. All major colleges make substantial efforts to recruit the best young scholars. It is in their interest to identify the most desirable prospective students by the most accurate means. In turning away good students because of invalid scores, both the college and the students suffer loss. In fact, the college's loss is compounded by the fact that poor students with invalid high scores are admitted where they would otherwise be screened out. The need to use valid criteria to assemble the most capable student body possible is a powerful incentive to improve testing. Adding the threat of public embarrassment as an incentive where a strong incentive already exists amounts to using two whips where one would do.

II. ALTERNATIVES TO DISCLOSURE

Because the colleges have both the power and the incentive to improve test validity, they can be relied upon to produce the desired improvement. The legislatures in New York and California, however, deemed their own involvement to be necessary as well. Although their disclosure enactments will not be effective in solving the validity problem, the question remains whether a different legislative approach might be effective.

A. *Prohibition of Admission Testing*

Since the partial invalidity of college admission tests causes unfairness, one could argue that they should be prohibited.

28. See, e.g., EDUCATIONAL TESTING SERVICE, 1980-81 LAW SCHOOL ADMISSION BULLETIN 14-15 (1980); EDUCATIONAL TESTING SERVICE, GRE INFORMATION BULLETIN 1980-81, at 17-18 (1980); EDUCATIONAL TESTING SERVICE, GRE INFORMATION BULLETIN 1979-80, at 13-14 (1979); EDUCATIONAL TESTING SERVICE, GMAT BULLETIN OF INFORMATION 1980-81, at 14-15 (1980).

29. However, it would undoubtedly be possible to compile the disclosed information into a highly inflammatory exposé, which would provide incentive by means of public embarrassment. But the process of making inflammatory the already known fact that test scores are partially invalid would necessitate so much distortion as to make the project of dubious social value.

However, a prohibition of admission testing would cause more injustice than it would remedy. Test scores, when combined with an applicant's previous grade point average, are the most valid indicator of the applicant's future academic performance that has yet been developed.³⁰ Although test scores are admittedly partially invalid, they are still a more valid criterion than any currently available replacement. If a legislature were to prohibit the use of test scores, it would only force colleges to use less valid criteria, resulting in even more injustice than presently exists. Thus, prohibition of admission testing is not a desirable solution despite the partial invalidity of admission tests.

B. Legal Standards for Testing

Short of an outright prohibition, it would be possible to put more teeth in truth-in-testing statutes than the present disclosure provisions afford. This intermediate approach would probably require setting a legal standard for test validity. For example, the legislature could determine an acceptable standard of validity for college admission testing or delegate the task of making such a determination to a regulatory agency.³¹

However, for such a standard to be more than rhetorical an enforcement mechanism capable of imposing sanctions would have to be provided. Sanctions would create serious drawbacks, however, for they interfere with the most valid, presently available criterion for making college admission decisions. Fines or damages would deprive testers of funds that could otherwise be used to improve the tests. An injunction or cease-and-desist order would amount to a prohibition and would also cut off any funds that would be gained from the fees of test takers. Thus, the imposition of sanctions necessary to enforce a legal standard for validity would not only interfere with use of the fairest admission criterion available but would also hobble the efforts of the status quo toward developing a more valid criterion.

Moreover, sanctions are unnecessary. Colleges already have an incentive to improve test validity as much and as quickly as possible. Legislative efforts to accomplish the same goal would be superfluous, if not completely counterproductive.

30. EDUCATIONAL TESTING SERVICE, *TEST USE AND VALIDITY* 16-19 (1980).

31. Alternatively, by setting no standard but creating a cause of action for aggrieved test takers, the legislature could leave the business of standard setting to the courts.

III. CONCLUSION

In summary, the partial invalidity of current college admission tests is a serious problem, but New York's attempt to remedy it through disclosure legislation is ineffective. Efforts to put more teeth in the law, such as a categorical prohibition or a legal standard with sanctions for noncompliance would only aggravate the situation. Moreover, such efforts are unnecessary since the colleges as owners and users of the tests have the power and incentive to bring about the desired improvements as rapidly as they can be achieved.

However, it is important for colleges, legislatures, and their critics to remember that not all problems are readily soluble, if soluble at all. Our efforts to describe the mind by a simple set of numbers may suffer from some very fundamental limitations, as Kant believed. The quixotic search for improvements in this area may not yield great practical results, and in the end we may have to content ourselves with the best we have been able to produce, even though it may be far from ideal.

Alan B. Asay