Cheater's Proof: Excessive Judicial Deference toward Educational Testing Agencies May Leave Accused Examinees No Remedy to Clear Their Names

Aron E. Goldschneider

Follow this and additional works at: https://digitalcommons.law.byu.edu/elj

Part of the Courts Commons, and the Education Law Commons

Recommended Citation
Available at: https://digitalcommons.law.byu.edu/elj/vol2006/iss1/3

This Article is brought to you for free and open access by BYU Law Digital Commons. It has been accepted for inclusion in Brigham Young University Education and Law Journal by an authorized editor of BYU Law Digital Commons. For more information, please contact hunterlawlibrary@byu.edu.
CHEATER’S PROOF: EXCESSIVE JUDICIAL DEFERENCE TOWARD EDUCATIONAL TESTING AGENCIES MAY LEAVE ACCUSED EXAMINEES NO REMEDY TO CLEAR THEIR NAMES

Aron E. Goldschneider*

I. INTRODUCTION

We live in an age in which standardized testing has become the principal means to judge the capabilities, educational level, and potential of young Americans. As a result of Congress’s No Child Left Behind Act, many grade school students now learn in their tender years that numerically measurable achievement on standardized tests is the “be all and end all” of their schooling and the key to their educational fate. Increasingly, slower students are under great pressure to perform adequately on standardized tests to avoid being held back, while

---

* Associate, Law Firm of Daniel M. Jaffe, Philadelphia. B.F.A., New York University; J.D., Temple University Beasley School of Law. I would like to thank Diane Maleson, Temple University Professor of Law, for her encouragement in producing this paper. Thanks also to my loving family for their patience and support.


3. See Stager, supra n. 1 (stating that the No Child Left Behind Act “will compel states to test their students every year from grades 2–12 in order to rank schools and shut many of them down. Our Proctor-in-Chief, George W. Bush, is extending the joys of standardized testing into Head Start”).

4. Supporters of the No Child Left Behind law argue that it prevents schools from claiming good results based upon the performance of the brightest students, while ignoring the performance of slow students. Jay Mathews, Federal Law’s Effect: Raised Expectations, http://www.washingtonpost.com/ac2/wp-dyn/?pgname=article&node=&contentld=A54709-2003Jan14&notFound=true (Jan. 14, 2003). Another view, however, is that the law encourages schools threatened with closure to leave students back in order to raise scores in the grade level to which the slow student is denied admittance, and perhaps in the repeated grade as well if he drops out. One academic, holding this view, suggests that, “‘[i]t’s not ‘No Child Left Behind’... it’s ‘Leave Them Behind and Blame Them Too’ at its worst, and perhaps a little better than status quo at its best.’” Id. (quoting Tim Hacsi, a Harvard Graduate School of Education researcher).
brighter students must score in the higher percentiles in order to be admitted to mentally gifted programs.\textsuperscript{5}

Similarly, standardized testing dominates the consciousness of many college-bound high school students, who spend significant amounts of their energy, thought and time on preparing for college entrance exams. In response, a burgeoning movement in opposition to the Scholastic Aptitude Test (SAT) has challenged that test’s purposes and efficacy,\textsuperscript{6} with a significant number of colleges, including the entire state system of California, rejecting its use in their admissions criteria.\textsuperscript{7} Additionally, a growing number of students who resent the role the test plays in their lives have chosen to opt out of the SAT testing process altogether.

Aside from criticism regarding standardized testing’s worthiness as an educational tool, much of the controversy surrounding educational testing has been over the methodologies used in formulating standardized test questions. Charges that standardized test questions have inherent

\textsuperscript{5} For just one example of a school district’s use of test scores to determine entry to its gifted programs, see e.g. Providence Pub. Sch. Dist., Advanced Academic Programs, Testing, http://www.providenceschools.org/dept/gifted/index.html (accessed Jan. 24, 2006) (“students receiving a combined score at or above the 61st percentile on both the Naglieri Non-Verbal Ability Test and the district’s SAT10 will move to the next phase of the selection process”). See also FairTest Examr., Tests Misused for Enrichment Program Admissions, http://www.fairtest.org/examarts/spring96/enrich.htm (Spring 1996) (describing how academic enrichment programs such as Northwestern University’s Center for Talent Development, Stanford University’s Education Program for Gifted Youth, and The Institute for the Academic Advancement of Youth (IAAY) at Johns Hopkins University all use the PSAT, SAT, ACT, or in the case of younger students, I.Q. tests, to identify gifted students and set minimum scores for admission to their programs).

\textsuperscript{6} For example, Richard C. Atkinson, president of the University of California system has criticized those aspects of the exam that are vestiges of IQ tests, such as analogy sections that create a “perverse incentive” for students to spend time and money preparing for “idiosyncratic” SAT questions. Ben Gose & Jeffrey Selingo, The SAT’s Greatest Test: Social, Legal, and Demographic Forces Threaten to Dethrone the Most Widely Used College-Entrance Exam, Chron. of Higher Educ. A10 (Oct. 26, 2001).

\textsuperscript{7} In February of 2001, the University of California system made the surprising announcement that its 170,000-student, nine-school group would no longer require the SAT for admission to its colleges, a decision that has led other state universities with competitive admissions such as North Carolina’s public colleges and the University of Texas at Austin to reconsider their SAT requirements. \textit{ld.} Additionally, Harvard, M.I.T., and eleven other “top colleges” are now part of a study to determine whether state tests already administered in high schools could be relied upon to assess students for college admissions purposes. \textit{ld.}

\textsuperscript{8} For example, in one recent Colorado “testfest,” a number of high school students performed music and read books outside the building in which their fellow students were subject to the rigors of the SAT. \textit{ld.} Bill Wetzel, an NYU student and founder of an organization called Students Against Testing, has announced plans to craft similar protests. \textit{ld.} Though Wetzel personally received an excellent 1420 SAT score, which no doubt helped him enter NYU, he says he now wishes he had chosen an SAT-optional school. \textit{ld.} Wetzel criticizes the way in which many of his high school classes focused on achieving high SAT scores. “I noticed the difference between some classes, where the teachers and the students were trying to get the highest scores possible, and classes that emphasized curiosity and real critical thinking.” \textit{ld.} (quoting Bill Wetzel).
rational or gender biases have long been center stage in educational news, periodically refreshed by new studies and research. As a result, whether fair or not, the once-burnished image of the standardized test as the "great equalizer" against privilege is largely a thing of the past.

Despite the controversy around standardized testing, exams like the SAT, ACT, GRE, GMAT, and LSAT remain the gateway through which the vast majority of students must pass if they wish to be admitted to institutions of higher learning. Indeed, the College Board, while ostensibly admonishing colleges not to set minimum SAT requirements for admissions, continues to publish what appear to be minimum SAT "cut-offs" for more than twenty institutions in its College Board College Handbook.

As if there were not enough controversy surrounding standardized

9. See Gose & Selingo, supra n. 6, at A10 (noting that the College Board, owner of the SAT, "has done a good job holding onto the ball, [for close to twenty years] fending off critics who maintain that the test discriminates against female and minority students").

10. Id. The College Board has consistently pointed to evidence that the SAT actually predicts that minority students will do better than they actually do in college. Id. Recent studies, however, have shown that questions employing "difficult" vocabulary actually favor minorities while "easier" questions can be confusing to minority test takers because of cultural differences in their approach to basic language. Jeffrey R. Young, Researchers Charge Racial Bias on the SAT, Chron. of Higher Educ. A34 (Oct. 10, 2003).

11. See Nicholas Lemann, The Big Test: The Secret of the American Meritocracy (Farrar, Straus & Giroux 1999). Lemann explains that many educators involved in the evolution of standardized testing believed it to be a tool for democratizing change—from Clark Kerr, who as president of the University of California instituted a Master Plan based on standardized testing to "create a fair opportunity for anyone to join the elite," to Stanley Kaplan, founder of the Kaplan test-prep schools who began his career coaching Jewish middle and working-class students, much like himself, who prior to standardized testing were largely denied admittance to the higher echelons of educational opportunity. Id. at 110–12, 136. Though the current higher education landscape is undoubtedly far more egalitarian than it was in the pre-testing age, standardized tests seem to have earned a reputation for impeding the progress of minorities while favoring those from the established classes. Id.

12. "ACT," when referring to the exam American College Testing administers, stands for "American College Test."

13. "GRE" stands for "Graduate Record Exam."

14. "GMAT" stands for "General Management Aptitude Test."

15. "LSAT" stands for "Law School Admission Test."

16. See e.g. Eric Hoover, The Changing Environment for College Admissions, Chron. of Higher Educ. A30 (Nov. 29, 2002) (describing how the majority of colleges continue to raise their admissions standards, "including standardized-test scores and high-school grade-point averages" leading to a corresponding drop in the acceptance rate from sixty-eight percent in 1992 to sixty percent in 1999, despite decreasing numbers of students graduating high school).

17. The College Entrance Examination Board (CEEB) is a non-profit educational organization, chartered more than a century ago, which has primarily functioned as an administrator of tests intended to determine which students will succeed in college. See infra nn. 41–50 and accompanying text for a fuller description of this organization.

18. Gose & Selingo, supra n. 6, at A10.
admissions testing over its questionable pedagogical legitimacy and potential biases, it has become increasingly apparent that the integrity of the testing process itself may be less than reliable. Opportunities abound for students to gain advantages on standardized tests either through sophisticated test-prep coaching or various forms of cheating. In response, testing agencies reserve the right to, and do, invalidate scores they consider suspect, whether they be those of individual testees suspected of cheating or a class of test-takers, some of whom may have been privy to leaked test questions or information gathered from prior tests.

The following Article discusses the legal conflict that arises when a test-taker challenges an invalidation determination made by a testing agency which has cancelled, or intends to cancel the test-taker’s standardized test score. More specifically, the Article focuses on the legal hurdles the accused examinee must surmount in order to preserve the challenged score through legal action. The author argues that it is unduly burdensome for a test-taker to pursue a worthy claim under existing “testing law,” due to the excessive deference paid to testing services by the courts, the difficulties in bringing equitable actions, and the limited legal avenues available to plaintiffs. Simply put, an innocent test-taker who has been unjustly accused of cheating may not be able to clear his/her name.

Accordingly, the Article offers a critical look at the legal standards courts have set, primarily concerning the due process and contractual rights of test-takers, and offers suggestions for a more equitable approach. The author questions the established judicial policy of non-interference in testing determinations and doubts the prevailing assumption that interference in testing agency determinations would threaten testing integrity. The author reasons that judicial deference may ultimately undermine testing validity, because it reinforces testing agencies’ largely unfettered power to invalidate large score increases that might otherwise raise questions about testing reliability and fails to compel testing agencies to make meaningful improvements in testing administration and security. Part II of this Article discusses the scope of the standardized cheating problem, emphasizing emerging problems in maintaining testing validity. Part III provides background on the major institutions involved in standardized testing. Part IV presents a general overview of legal barriers to standardized testing invalidation challenges, with particular focus on the difficulties of bringing equitable actions. Part V focuses on the two main legal avenues available to standardized testing plaintiffs, with subsection A devoted to unconscionability and breach of contract claims, and subsection B discussing violation of due
process claims. Part VI provides a critical analysis of the chief assumptions underlying judicial deference to testing agency determinations. Finally, Part VII urges courts to reject formalistic enforcement of adhesion contracts and instead exercise powers of equity to validate examinee scores in appropriate cases. The author concludes that unless courts reconsider the unwarranted judicial deference accorded testing agencies, innocent test-takers wrongly accused by testing agencies will continue to be left without a meaningful legal remedy to clear their names and validate their scores.

II. SCOPE OF THE CHEATING PROBLEM

Standardized testing agencies are having an increasingly difficult time controlling exam security and maintaining testing validity. The scope and seriousness of recent security breaches indicate that standardized testing procedures are vulnerable to cheating on multiple fronts. Examinee access to re-used test questions prior to testing, high-tech spying, and test-taking by imposter examinees all threaten the integrity of test-taking results, in part justifying the vigor with which test-taking agencies respond to those potential instances of cheating that they manage to discover.

For example, the National Board of Podiatric Medical Examiners (NBPME) and its agent Chauncey, a subsidiary of the Educational Testing Service (ETS), not long ago invalidated an entire July 2002 computer-based test taken by podiatry students at the New York College of Podiatric Medicine (NYCPM).\textsuperscript{19} Claiming proof by ""documented evidence"" that NYCPM students had ""access to 'documents containing secure test questions,'"" the NBPME withheld score results and mailed notifications to students' intended recipients, mostly hospital residency programs, detailing its allegations.\textsuperscript{20}

At roughly the same time that ETS's Chauncey subsidiary was invalidating the NYCPM students' scores, ETS was also issuing bulletins to American universities that GRE scores from China, Taiwan, and South Korea might be suspect.\textsuperscript{21} Attached to these warnings were bemusing qualifications that most Asian test-takers were hard-working and their


\textsuperscript{20} Id. at **1-2 (quoting the NBPME web site). In an ensuing action in the United District Court for the Southern District of New York, sixty NYCPM student plaintiffs brought an action to compel Chauncey to release their scores but failed to win a preliminary injunction. Id. at *1.

scores genuinely earned. Concerns over Asian GRE scores were prompted by information ETS obtained that students were using the Internet to post questions (and sometimes answers) that they remembered from previous tests. With each student remembering just one or two questions, the aggregate question pool was apparently substantial. As a consequence of this question pooling, ETS claimed that national average test scores had increased one hundred points in China, and fifty points in Taiwan and South Korea.

Because testing agencies reuse exam questions, many standardized testing exams are vulnerable to question pooling. The same test questions may appear on multiple exams over a certain period of time, especially on computer-based tests. Even on paper and pencil exams, a significant percentage of test questions may be re-used, primarily for the purposes of maintaining consistency across successive tests. While Internet question pooling is of recent concern to testing agencies, attempts to gain an edge by studying prior test questions are nothing new. Testing agencies have for many years been locked in intellectual property battles with their traditional test-prep nemeses, Kaplan and Princeton Review, accusing them of improperly using copyrighted materials from previous exams and even sending in dummy test-takers with the express purpose of copying questions from exams. In an effort to curb the leaking of

---

22. Id. The GRE Board, which sets policies and oversees the use of the GRE, “felt it had no choice but to tell graduate schools about the problems with the test, but [was nonetheless] trying to publicly acknowledge that many Asian students [were] scoring high because of hard work, not cheating.” Id. Meanwhile, the ETS was “urging universities to look at the scores in the context of the students’ other achievements, their essays, and their letters of recommendation.” Id.

23. Id.

24. Id.

25. Id.

26. Id.


28. E.g. Lemann, supra n. 11, at 113 (recounting how Stanley Kaplan, founder of Kaplan schools, was treated as a “pariah” by the ETS, particularly for holding post-SAT parties where students remembered and discussed questions from the exam they just took), 229 (recounting how “[i]n the early days, [John] Katzman [founder of Princeton Review] obtained his prep materials by doing exactly what Kaplan had done in his early days. As [Katzman] put it: ‘We’d send people in to take tests, and send ourselves in. I’d take ten, fifteen kids and say: I’ll buy you Chinese food if you tell me as many questions as you can remember’”). See also Critics Blast College Board Requirement That SAT Proctors Take “Oath” on Coaching, 48 Chron. of Higher Educ. A35 (Oct. 19, 2001) [hereinafter Critics Blast] (reporting that the College Board discovered in Spring 2001 “that a secure copy of the SAT II biology exam had been copied without authorization and was used in private coaching”); Educ. Testing Serv. v. Stanley Kaplan Educ. Ctr., 965 F. Supp. 731, 734–35 (D. Md. 1997) (recounting how in 1994 a Kaplan GRE product director sent several students in to reconnoiter a new computer-based GRE-CAT to look for testing differences from the paper and
test questions, ETS recently made the controversial demand that its SAT proctors sign an oath that they would not work for test-prep schools because of concerns that proctors would filch or reproduce exam questions. 29

Naturally, reliable information on re-used questions is a valuable commodity, and a major chip to be parlayed in the game of high stakes testing, both for students and the cram schools that cater to them in an increasingly competitive market. This is particularly true when a small bump in a student’s standardized test score can mean the difference in that student gaining admission to an elite school or profession with all the attendant economic implications of that admission.30

High-tech spying techniques further expand the scope of the cheating problem. With sophisticated technologies increasingly available to the general public, the level to which determined test “spies” can go in their efforts to pilfer test questions is daunting. For example, two Columbia undergraduate students were recently arrested for using walkie-talkies and high-tech transmitters to steal questions from a computer-based GRE exam.31 Having gained admittance to a special disabilities testing room by feigning a respiratory illness, the ostensible examinee distracted the proctor, while his partner snuck into the testing room and attached a device that not only intercepted questions from the computer but also transmitted them to a van parked outside the building.32

This does not mean that a wily student need be particularly high-tech to succeed in dramatically raising a test score by cheating. While standardized exams, particularly graduate and professional tests, are increasingly being administered through computer-based testing (CBT),33 pencil and paper exams are still the norm for K-12 testing and

---

29. Critics Blast, supra n. 28, at A35.
30. For a harsh example of how the smallest test score fractions can determine an examinee’s future, see In re Singh, 576 S.E.2d 899, 900 (Ga. 2003), in which a Georgia Bar examinee contested the way examiner ACT rounded and scaled his raw scores on the Multistate Bar Exam, leaving him with a 269.9 total score, short of the 270 points he needed for admission. As the Georgia Supreme Court put it, bluntly and without a trace of sympathy: “applicant Singh achieved a score of 269.9, not a passing score.” Id.
32. Id.
33. See Allison Yang, A New and Improved Way to Fail: Why the Standardized Tests of the Future Won’t be Quite So Standard Anymore, 30 Yale Herald Online 1 ¶¶ 2, 7–10 (Oct. 27, 2000) http://www.yaleherald.com/archive/xxx/2000.10.27/features/exclusive.html (stating that both the GRE and the GMAT are now computer based, with the GRE exclusively so in America, and discussing how computer-based tests may favor younger, more tech-savvy examinees and those from the developed world); see also JustColleges, About the TOEFL, CBT,
college-entrance testing, and both are vulnerable to students copying from their neighbors and to imposter examinees taking tests as “ringers” for other students. The vulnerability lies not in the test-taking process itself but in the failure of identity checking procedures to stop impersonators before they sit for tests.

Proof of the extent of this imposter examinee problem can be found in the approximately sixty criminal cases pending in the District Court of New Jersey in which numerous foreign nationals are charged with attempting to defraud ETS in connection with its Test of English as a Foreign Language (TOEFL) exam by using imposters to test for them. In one of the first pre-trial determinations rendered in this slew of cases, United States v. Alsugair, the district court held that ETS had a property interest in both its trademark and goodwill, and that the defendant could be charged with mail fraud for depriving ETS of that interest through an elaborate scheme involving the substitution of photographs, forging of documents, and the mailing of a phony ETS envelope.

http://www.justcolleges.com/tests/index.phtml?no=tests_toeflcbt.htm (accessed Sept. 28, 2005) (indicating that the Test of English as a Foreign Language is now offered both as a CBT and a paper and pencil test and explaining some of the advantages and disadvantages of each); American College of Physicians, Board Exams About to Make the Move to Computers, ACP Observer (Nov. 2003) (available at http://www.acponline.org/journals/news/nov03/abim.htm) (announcing that the American Board of Internal Medicine plans to change all its certification and recertification exams to computer based testing by 2006); American Association of Critical Care Nurses, CCRN and CCNS Paper-and-Pencil Exams Offered as Option in Rural Areas, 17 AACN News (newsletter of the Am. Assoc. of Critical Care Nurses) 2, ¶¶ 2–3 (Nov. 2000) (available at http://www.aacn.org/AACN/aacnnews.nsf/0/d1a740f42a2bce6682569ce00812ae0?OpenDocument#paper) (announcing that pencil and paper exams are available to those living in rural areas but noting that the less expensive method of test-taking is at urban computer centers).

34. See Yang, supra n. 33, at ¶ 12 (indicating that the SAT, ACT, and MCAT are still administered by paper and pencil and that “[f]or the SAT and ACT, the obstacle in the high-tech conversion is the sheer number of test-takers each year and the still limited computer accesses in many secondary schools across the nation”).

35. See e.g. U.S. v. Alsugair, 256 F. Supp. 2d 306, 309 n. 2 (D.N.J. 2003). The court also cited to the United States Attorney’s Office for the District of New Jersey news release entitled “Dozens of Foreign Students Arrested Nationwide in English Language Testing Scam,” indicating that the problem was not limited to just New Jersey. Id.


37. Id. at 319. The court, however, granted the defendant's motion to dismiss allegations based on deprivation of ETS's property interest in its copyright, as well as administration and scoring services. Id.

38. The fraudulent scheme for which the defendant was charged was described by the court as follows:

The imposter allegedly appeared at a test site and falsely identified himself as the student who had to take the exam. This imposter, posing as the student, had his photo taken at the test site, sat for the TOEFL exam, and directed that the exam results be mailed to a predetermined location in California. Once the test results arrived, the real student's photograph was substituted for the imposter's photograph, and the fraudulent TOEFL exam results were then
The scale of these and other attacks on testing security and reliability suggest that the reputation of standardized testing agencies and the validity of the exams they administer are ever-more fragile. Surveying the cram-school landscape, John Katzman, CEO and founder of test-prep leader Princeton Review, boldly proclaims, “Let’s face it, there’s cheating at every level . . . Teachers cheat. Principals cheat. Sometimes, whole communities cheat . . .”.39

In the face of such charges, services like ETS will no doubt continue to manifest a very strong interest in protecting their image as providers of accurate information. As one court has observed, it is their “sole stock in trade.” 40

III. BACKGROUND: THE COLLEGE BOARD, ACE, ETS, AND ACT

In order to better understand the forces at work in standardized educational testing cases, some background on the principal testing players may be helpful, namely the College Entrance Examination Board, the American Council on Education, and the two major American testing agencies—Educational Testing Service and American College Testing.

The College Examination Entrance Board (CEEB), founded in 1900, is a non-profit membership organization “composed of more than 4,700 schools, colleges, universities, and other educational organizations,” which describes its mission as “connect[ing] students to college success and opportunity.”41 The College Board, as it is generally known, boasts of “serv[ing] over three and a half million students and their parents, 23,000 high schools, and 3,500 colleges through major programs,” and oversees the SAT, PSAT, and the Advanced Placement Program (AP), among others. 42

In its early years, the College Board primarily functioned as “a tweedy, clubby association of a few dozen private schools and colleges [ ] founded in 1900 to perfect the close fit between New England boarding schools and Ivy League colleges.”43 To effectuate this purpose, the Board administered “a weeklong battery of essay examinations in

mailed to schools requiring the real student’s exam results in a phony ETS envelope.

Id. at 309.

39. Critics Blast, supra n. 28, at A35.
42. Id.
43. Lemann, supra n. 11, at 28–29.
various subjects, called the ‘college boards,’ which served as “a uniform admissions test that all the [prestigious] colleges would accept.”

Later, however, the Board became the administrator of the Scholastic Aptitude Test, or SAT as it came to be known, to a much wider body of students. The SAT, which had its origins in Army intelligence tests, was first tried out on high school students in 1926, not for admissions purposes, but in order to establish the test’s validity in predicting test-takers’ freshman grades. By the mid-to-late 1930’s, however, the newly-automated SAT, along with a battery of additional subject-specific multiple choice tests, became the testing pathway for high school students seeking scholarship admission to Ivy League schools. Then, just weeks after the bombing of Pearl Harbor, the essay examinations of the college boards were suspended, never to be resurrected, and the SAT became the “admissions device” for all students seeking entry to Ivy League schools.

Several years later, at the close of World War II, the president of the Carnegie Foundation of philanthropic charities approached Henry Chauncey, head of testing at the College Board, about the College Board taking over the Carnegie-run Graduate Record Exam (GRE) and ultimately instituting a national testing agency to administer all standardized tests in America. But while Chauncey and the Carnegie Foundation were excited about new plans to expand and consolidate national testing, a powerful national educational interest had to be negotiated with: the American Council on Education.

Founded in 1918, the American Council on Education (ACE) is a non-profit membership organization that now claims to have enlisted more than “1,800 accredited, degree-granting colleges and universities and higher education-related associations, organizations, and corporations” in its ranks. The Council, which oversees the high school equivalency General Educational Development (GED) exam, describes its mission as “provid[ing] leadership and a unifying voice on key higher education issues” and “influenc[ing] public policy through advocacy.

44. Id.
45. Id.
46. Id. at 32.
47. Id. at 39.
48. Id. at 54.
49. Id. at 60–61.
50. Id. at 62.
research, and program initiatives.\textsuperscript{52}

In 1945, ACE was considered, according to its leader George Zook, "the leading educational organization in the country."\textsuperscript{53} When Zook was approached for his endorsement of a national testing agency plan, he sought to make the proposed agency a subsidiary of his own organization.\textsuperscript{54} Indeed, he was highly opposed to the College Board, which he considered to be a "tiny, regional, elitist" organization, taking charge of any national standardized testing plans.\textsuperscript{55} Soon after Zook made his position known, a committee commissioned by the president of Harvard, in order to obtain ACE's support, suggested that the College Board essentially "hand over all its tests to the new agency and then cease to exist."\textsuperscript{56}

Nonetheless, after much infighting, politicking, and some heavy arm-twisting from the Carnegie foundation, it was the College Board that eventually "won out."\textsuperscript{57} Consequently, when the Educational Testing Service was chartered and opened its doors on January 1, 1948 in Princeton, New Jersey, Henry Chauncey of the College Board presided over its operations.\textsuperscript{58}

For more than ten years following its inauguration in 1948, ETS "enjoyed what amount[ed] practically to a monopoly in college admissions testing."\textsuperscript{59} Not only did ETS own the copyrights to all the major higher education tests, but it had also "inherit[ed] from the College Board all the most prestigious private universities in the East as clients."\textsuperscript{60} And yet the activities of the College Board and ETS had hitherto been restricted almost entirely to the elite schools of the Northeast. If they were to become truly national players, they would have to gain entry to the state school systems of the Midwest and West, which were generally open-admissions, and resistant to standardized tests as admissions devices.\textsuperscript{61}

Unfortunately for ETS, in 1959, ETS's first serious competitor, American College Testing (ACT), was born in America's heartland of

\textsuperscript{53} Lemann, supra n. 11, at 62.
\textsuperscript{54} Id.
\textsuperscript{55} Id.
\textsuperscript{56} Id.
\textsuperscript{57} Id. at 63–65.
\textsuperscript{58} Id. at 65. ETS was chartered with Chauncey as president and James Conant, president of Harvard, as Chairman of the Board. Id.
\textsuperscript{59} Id. at 102–103 (quoting a confidential internal ETS report circulated in 1958).
\textsuperscript{60} Lemann, supra n. 11, at 96.
\textsuperscript{61} Id. at 97.
Particularly alarming to ETS’s plans of national dominance was
the fact that ACT was attempting to capture the same state universities
that ETS was shooting for. The approaches of the two organizations,
however, as they battled “state by state,” were quite different. While
ACT attacked from the bottom up, broadcasting a populist message that
emphasized “guidance and placement of the many,” ETS remained elitist
in approach, with its mission being to find the most gifted students and
place them at the top universities, while “guiding the rest to lower their
aspirations.”

In the coming years, the two organizations carved up the country,
with ETS seemingly capturing Texas, Michigan, Georgia, and Colorado,
and ACT claiming victories in Illinois and Ohio. The two companies’
battle for California’s huge state school system typified their contrasting
approaches—ACT succeeded in winning over “the low-prestige, high-
body-count . . . state[] colleges and junior colleges,” while ETS
eventually managed to capture “the crown jewel, the University of
California.”

Today, ETS proclaims itself “the world’s largest private educational
testing and measurement organization,” and ACT remains ETS’s chief
competitor, testing nearly as many high school seniors per year as ETS.
Though the two organizations retain their non-profit status, they are both
openly aggressive, competitive, and expansive in their orientation.

Indeed, ACT announces its vision “[t]o be the world’s leading
provider of information for educational and career decisions in support of
lifelong learning.” Meanwhile, ETS, which “develops and administers
more than twelve million tests [in almost two hundred countries]
worldwide,” has signaled its intention to “broaden its scope beyond the
U.S. measurement space into the worldwide education and training
space.” This testing goliath, which already formulates and administers
the SAT, PSAT, AP, GMAT, GRE, and TOEFL, specifically has its eyes on
“increasing its presence in certain education markets[.] K-12,
IV. OVERVIEW OF LEGAL BARRIERS TO SCORE INVALIDATION CHALLENGES

Daunting hurdles face standardized testing examinees who challenge the invalidation of their standardized test scores by testing agencies. Examinees lucky enough to have their day in court are often frustrated by the refusal of courts to rule on the merits of whether the particular examinee cheated instead of on factors that generally resolve favorably for the testing agency. Other difficulties for examinees derive from the nature of the equitable relief they seek. The majority of examinees seek an injunction barring the agency from canceling their scores or directing the testing agency to release or reinstate their scores. This equitable posture makes an examinee's case difficult to pursue for several reasons: 1) traditional deference on the part of courts toward educational testing agencies and their determinations, based on established equitable policies; 2) the expense in bringing an equitable claim; and 3) the high burden of proof required for the granting of injunctive relief.

As this Article's review of testing case law will show, plaintiffs have obtained limited success in only a handful of recorded invalidation cases, and even these "victories" did not ultimately result in meaningful relief for plaintiffs. An even bleaker picture for suspected examinees emerges if one considers that recorded cases likely showcase the claims of only a tiny fraction of examinees who maintain their scores were invalidated unfairly. The "tip of the iceberg" principle operates in many areas of litigation, since reported cases are only a fraction of the total number of cases brought to trial, which are only a fraction of the cases settled prior to trial, which are only a fraction of the claims never filed, and so on. 73

72. Id.

73. See Andre N. Moenssens, Novel Scientific Evidence in Criminal Cases: Some Words of Caution, 84 J. Crim. L. & Criminology 1, 9–10 (Spring 1993) (noting the operation of this "tip of the iceberg" principle in the context of fraudulent expert witnesses); Linda S. Mullenix, Taking Adequacy Seriously: The Inadequate Assessment of Adequacy in Litigation and Settlement Classes, 57 Vand. L. Rev. 1687, 1743 (2004) (invoking the principle in noting that it is difficult to estimate the true incidence of class action filings and their attendant litigation "since many class action suits involve ongoing litigation in which events and outcomes never result in reported orders, decisions, or appellate opinions."); See also Dow v. Donovan, 150 F. Supp. 2d 249, 271 (D. Mass. 2001) ("[T]he record of reported cases (probably only the tip of the iceberg) shows a history of slow recognition of [frequent conflicts of interest between liability insurers and policyholders sued in tort."]); Caroline R. Adams, Student Author, The Constitutional Validity of the Religious Land Use and Institutionalized Persons Act of 2000: Will Ruppa’s Strict Scrutiny Survive the Supreme Court’s Strict Scrutiny?, 70 Fordham L. Rev. 2361, 2381 (May 2002) (recounting scholar’s testimony before
For the following reasons, however, there is cause to believe that this “tip of the iceberg” principle applies with particular force to standardized testing claims, and that effective remedies are largely out of reach for examinees who have been wronged by invalidation determinations.

A. Refusal of Courts to Make Determinations on Whether Individual Test-Takers Cheated

As will be discussed below, some litigants have brought court challenges attacking contractual restrictions limiting dispute resolution to internal investigation or binding arbitration. Because invalidation matters are contractually subject to either the testing agency’s internal investigation or binding arbitration, most disputes will not reach the courts. Additionally, the litigants have challenged provisions limiting the scope and procedural depth of these inquiries. Nonetheless, courts have invariably refused to rule on the merits of whether particular test-takers cheated and have thereby denied plaintiffs any chance to attain their ultimate remedy. Rather than ruling on whether examinees cheated, courts limit their determination to whether the testing agency 1) had a substantial basis for challenging the examinee’s score, 2) provided adequate procedural avenues for the examinee to attempt to validate the score, and/or 3) made a good faith effort to investigate the alleged testing “irregularities” fully.

The first area of inquiry, a showing by the testing agency that it had a substantial basis to challenge a score, is not difficult for the testing agency to establish. A review of cases involving cancellation of individual examinees’ test results shows that the trigger that almost always set off the initial investigation was a very large increase in scores across successive tests. Because testing agencies can offer numbers

Congress in which he invoked the “tip of the iceberg” principle to assert that minority religions were “vastly over-represented in zoning litigation and frequently discriminated against.”

74. See infra nn. 168–176 and accompanying text.

75. See infra nn. 148–156 and accompanying text (discussing testing booklet contract provisions).

76. Id.


78. See Langston v. ACT, 890 F.2d 380, 381 (11th Cir. 1989) (noting that any score that increased by more than a certain amount of points would be “automatically flag[ged]” by ACT’s computer marking system); see also Koza v. ACT, Inc., 2001 WL 1191050, at *1 (Mich. App. Oct. 10, 2001) (reporting that ACT stated in a letter to plaintiff that it “reviewed ‘all unusual score increases’ “); Dalton, 663 N.E.2d at 290 (reporting that plaintiff’s score increase of more than 350 points across successive SAT exams fell into the category of what ETS termed “Large Score
demonstrating the rareness of such large increases between tests and studies maintaining they are virtually impossible to achieve fairly, a large increase by itself goes a long way to providing a substantial basis for the testing agency’s challenge. Add impressive-looking statistical odds against a test-taker’s answer sheet matching so many answers on another test-taker’s sheet, or expert opinion that the handwriting on the test-taker’s first test did not match her second, and the establishment of a substantial basis for the challenge is largely a foregone conclusion.

As to the second area of inquiry, the procedural avenues afforded the suspected testee to validate his or her score, there has hardly been a testing invalidation opinion that failed to mention that ETS and ACT gave examinees suspected of cheating a chance to retest at no expense in order to confirm their scores. Other procedures that courts have pointed to as adequate means of validating a suspected examinee’s scores include allowing test-takers to submit evidence that might help to confirm the validity of their test results, giving them the chance to be heard at a hearing, and offering them the opportunity to submit their dispute to binding arbitration.

Relying on this list of ostensible protections afforded the test-taker, courts have uniformly found that testing agencies offered suspected examinees adequate procedural avenues to confirm

---

79. See e.g. Scott, 600 A.2d at 502 (noting that in its internal review, ETS “took into consideration the fact that only three test takers out of a sample of over 7,000 recorded gains of more than 42 points in a General Knowledge test).

80. See Sarah Stockwell et al., The SAT Coaching Coverup: How Test Preparation Programs Can Raise Scores by 100 Points or More and Why the College Board and ETS Deny the Evidence 9 (Natl. Ctr. for Fair & Open Testing 1991) (noting that up until 1988, the College Board maintained that various studies showed that test-prep coaching only raised SAT scores from zero to thirty points).

81. In some cases, these statistics would appear virtually conclusive that the examinee cheated. See e.g. Tolleson v. Educ. Testing Serv., 832 F. Supp. 158, 159 (D.S.C. 1992) (citing as an established fact that the odds that plaintiff could have answered “all of the 98 correct responses the same and 31 out of 38 incorrect responses the same” as a neighboring examinee are “less than 1 in 100,000,000;” based on statistical evidence provided by the chairperson of the ETS Board of Review).


84. Murray, 170 F.3d at 516; Johnson, 615 F. Supp. at 637.


86. Langston, 890 F.2d at 387.
suspected scores.  

Only the third area of inquiry, whether the testing agency made a good faith effort to investigate the alleged testing "irregularities" fully, is likely to undergo close scrutiny by the court. This area of inquiry will be discussed at length later in this Article.

B. Judicial Deference to Testing Agencies

Recorded case law shows an established pattern of judicial deference to educational testing agency invalidation determinations based upon public policy concerns. In *Dalton v. Educational Testing Service*, an influential SAT score cancellation case, New York’s highest court opined that Educational Testing Service was due much the same judicial deference traditionally accorded academic institutions. The court summed up its position as follows:

The comparison between ETS and academic institutions is surely not exact, inasmuch as judicial restraint in matters of academic achievement is based, in part, on the inherently subjective nature of the evaluation to be made by professional educators. Still, similar policy concerns militate against directing ETS to release a questioned score. When a standardized testing service reports a score, it certifies to the world that the test-taker possesses the requisite knowledge and skills to achieve the particular score. Like academic credentials, if courts were to require testing services to release questioned scores, "the value of these credentials from the point of view of society would be seriously undermined."

In *In re K.D. v. Educational Testing Service*, an earlier, much-cited law school admissions testing case, the Supreme Court of New York County stated specific policy concerns, here echoed by another court:

[A] testing service "performs a highly valuable service not only to the law schools but to the public as well. Moreover, the accuracy of its predictions is defendant’s sole stock in trade. The less accurate as a forecaster its tests are, the less value they have . . ." Thus, when unable

---

87. This is not to suggest that courts addressing the adequacy of procedural avenues have necessarily conducted a constitutional due process inquiry. In a number of cases, courts have ruled only on whether the examinee was, in good faith, afforded the procedural opportunities guaranteed by the testing contract. See *Murray*, 170 F.3d at 516; *Johnson*, 615 F. Supp. at 638; *Dalton*, 663 N.E.2d at 291–95. For a review of those cases in which courts have ruled on violation of due process claims, see *infra* nn. 239–303 and accompanying text.

88. 663 N.E.2d 289.

89. *Id.* at 294.


91. 386 N.Y.S.2d 747.
to vouch for the integrity of test results, a testing service is "within its obligations and duties to the (college) and to the public in requesting . . . a reexamination." 92

Accordingly, under this policy approach, even if significant factual doubt is raised as to whether an examinee actually cheated, the court, in balancing the equities, will still stop short of directing a testing agency to report a suspect score. 93 Not only do the courts emphasize "the reliance that students, educational institutions, prospective employers and others place on the legitimacy of scores released by [the testing agency]," 94 but they also stress that "'[t]he other test-takers are entitled to assurance that no examinee enjoys an unfair advantage in scoring."' 95 Finally, going further, courts have declared that a testing agency's reputation itself is a weighty interest, and that "'it act[s] within its right to protect its own image,'" when it invalidates a suspect score. 96

C. Expense in Bringing Educational Testing Claims

Educational testing claims challenging invalidation determinations are likely to be very expensive to pursue. Though some standardized testing litigants do seek money damages in addition to equitable demands for release of their test scores, common sense suggests that such claims do not make attractive candidates for contingency representation. This is because the chances of success based on established case law are slim, and damages likely to be small, if capable of being proved at all. 97

92. Doe at *7 (quoting In re K.D., 386 N.Y.S.2d at 752) (citations omitted).
93. See Dalton, 663 N.E.2d at 291–95 (affirming the lower court's determination that ETS breached its contract with a high school student by failing to consider in good faith considerable factual evidence that the student had not cheated on an SAT exam, yet nonetheless modifying, on public policy grounds, the lower court's affirmation of the trial court's order directing ETS to release the student's SAT score, requiring only that ETS reconduct its investigation and duly consider plaintiff's evidence).
94. Id. at 294.
95. Murray, 170 F.3d at 517 (quoting Scott, 600 A.2d at 504).
96. Id. (quoting In re K.D., 386 N.Y.S.2d at 752).
97. A fruitless search by the author found no recorded cases even mentioning any money damages paid to a standardized testing examinee challenging an invalidation determination. This, of course, does not mean that damages have never been paid. Rather it suggests that testing agencies have done a good job of avoiding the establishment of legal precedents by wearing plaintiffs down, settling when advantageous, or both. See Allan Nairn et al., The Reign of ETS: The Corporation That Makes Up Minds 284 (The Ralph Nader Report on the Educ.Testing Serv. 1980) (describing how ETS "inundated" a 23-year-old pro se plaintiff seeking compensation for a lost semester resulting from "an ETS records foul-up" by filing "a docket of legal papers eight inches thick" before eventually settling with the plaintiff for $250.00 after "photo-copy costs alone had drained [the plaintiff's modest budget]"). Vincent F. Nicolis, Esq., plaintiffs' counsel in the Dalton and Cortale cases discussed infra nn. 204–217, 218–236 and accompanying text, recounted how on many nights he was "up until three in the morning" trying to answer ETS's legal onslaught during the litigation of
Moreover, testing invalidation cases can be factually complex, most likely requiring the production of multiple witnesses and costly experts. A good example of this may be found in *Dalton v. Educational Testing Service*, in which a high school student was accused by ETS of having an imposter sit in his place for an SAT test. The *Dalton* plaintiff sought an injunction prohibiting ETS from canceling his score and an order compelling release of the score based on specific performance of his testing contract. In successfully arguing that ETS breached its duty of good faith under the contract, the plaintiff presented a total of sixteen witnesses and ETS countered with nine. The trial took twelve days and "occupied more than 2,000 pages of transcript." The trial was then followed by two presumably costly appeals. In the end, the New York Court of Appeals effectively denied the plaintiff the relief he was seeking when it ruled that ETS would not be directed to release his score.

*Dalton* is hardly the only instance of a standardized testing plaintiff obtaining unfavorable results after presenting an extremely costly invalidation claim. Similarly complex claims put forward by plaintiffs, requiring the hiring of experts and the presentation of extensive documentation, have fared poorly.

---


99. Id.

100. Id. This equitable claim was the only cause of action originally brought by the plaintiff, though he later amended his complaint to seek money damages under other causes of action. Id. at 742–43. The equitable claim was separated from the damages action and, following discovery, an expedited bench trial was held on the original claim only. Id. at 743. A fruitless search by the author found no record of the damages action.

101. Id. at 744.

102. Id. at 743.

103. Id.

104. The case was first appealed to the New York Supreme Court Appellate Division and then to the New York Court of Appeals. Id.

105. Id. at 294. See infra nn. 212–214 and accompanying text for a discussion on the relief granted in *Dalton*.

106. See also *Johnson*, 615 F. Supp. at 637–39 (granting summary judgment for defendant ETS on all five of plaintiff’s claims and noting, even prior to trial, "extensive documentation" provided by plaintiff to the ETS Board of Review, four meetings with the Board in which plaintiff was represented by counsel, and the hiring by plaintiff of a handwriting expert, whose opinion, favorable to the plaintiff, nonetheless could not outweigh three unfavorable opinions provided by ETS’s handwriting experts); see e.g. *Langston*, 890 F.2d at 383–88 (describing complex litigation at
D. Heavy Burdens for Obtaining Injunctive Relief

When an individual standardized testing examinee receives word that, following a testing agency investigation, his score has been determined to be invalid, the only viable options he has to stop the agency from canceling or withdrawing his score under his contract are either to retest, which allows the examinee to validate his score if he is able to come within a certain amount of points of the questioned score, or to agree to binding arbitration, which will not address the district court level involving unsuccessful contractual, constitutional, and tort claims, and noting plaintiff's hiring of at least two statistical experts in vainly challenging ACT's mathematical methods for determining that plaintiff cheated.

107. See infra nn. 146–156 and accompanying text (providing a discussion on the test-taker's booklet contract with the testing agency).

108. Cases differ, in different contexts, as to the range of points an examinee's retest must come within in order to validate a suspected score. Langston, 890 F.2d at 382 n. 4 (test-taker at first informed by ACT that retest score had to "confirm" suspected score but later was informed that if retest score were "within a couple of points," then suspected score would be validated); Tolleson, 832 F. Supp. at 159 (ETS required examinee to test within a fifty to a hundred point range of suspected National Teacher's Exam (NTE) score of 650); Koza, 2001 WL 1191050 at *1 (ACT required examinee to retest within three points of suspected ACT composite score of 24 in order to validate questioned score); In re K.D., 386 N.Y.S.2d at 750 (examinee required to retest within fifty points of suspected LSAT score); DePina v. Educ. Testing Serv., 297 N.Y.S.2d 472, 474 (App. Div. 2d Dept. 1969) (testee told retest scores would have to "approximate[ ]" questioned CEEB test scores).

109. The current ETS bulletin for GRE testing states:

Invalid Scores. ETS may cancel scores if, in its judgment, there is substantial evidence that they are invalid for any reason. Evidence of invalid scores may include, but is not limited to, discrepant handwriting, unusual answer patterns, and inconsistent performance on different parts of the test. Before canceling scores pursuant to this paragraph, ETS notifies the test taker in writing about its concerns, gives the test taker an opportunity to submit information that addresses ETS's concerns, considers any such information submitted, and offers the test taker a choice of options. The options include voluntary score cancellation, a free retest, or arbitration in accordance with ETS's standard Arbitration Agreement.


Although upon the initial questioning of the testee's score, the testee is invited to submit evidence supportive of the validity of the score, it is, of course, difficult to prove a negative, and unlikely that a testing agency will change its determination (founded upon hard statistical evidence) purely upon affidavits as to the testee's character, school performance, or circumstances surrounding the testing event itself. In any case, once the testing agency has completed its internal investigation and decided the score is invalid, the testee must either retest or submit to binding arbitration. Another choice offered the examinee, allowing the recipient school to be the arbiter of the dispute, hardly seems a credible option, since the last thing the test-taker wants is for his chosen school to
issue of whether or not plaintiff cheated but rather focus on the propriety of the testing agency’s invalidation action, for example, “whether the testing agency acted reasonably and in good faith in deciding to cancel [the examinee’s] scores,”110 or “whether there [was] substantial evidence supporting cancellation of the scores in question based on the information available . . . .”111 Should the examinee refuse to retest or submit to binding arbitration, the testing agency will move to cancel the score. It is at this point that an examinee may have to seek a preliminary injunction if he wishes to block the testing agency from withdrawing the score or notifying an educational institution to which he has already been accepted. Due to the typically high standards required by courts before granting an injunction, it is nearly impossible for an examinee to achieve such relief.

An early testing case involving a request for a preliminary injunction, DePina v. Educational Testing Service,112 illustrates just how heavy the burdens for examinees are. A testee sought a court order to stop ETS from withdrawing his 1968 College Entrance Examination Board scores and notifying the United States Merchant Marine Academy that his scores had been invalidated.113 Though the testee was granted a preliminary injunction by the Supreme Court of Nassau County, the New York Supreme Court Appellate Division reversed the order, calling it “an improvident exercise of discretion.”114 While the court cited to “a thorough comparison” made by ETS of the examinee’s scores with the scores of another test-taker, “reveal[ing] circumstances which indicated, prima facie, that plaintiff had cheated,”115 it detailed no particular facts from this comparison.116 Nor did the court explicitly reference the sine qua non for the granting of a preliminary injunction—a showing by the plaintiff that without the injunction, the plaintiff would be irreparably harmed.117 Rather the court, in issuing its memorandum opinion,

become aware of doubts as to his score.

110. Koza, 2001 WL 1191050 at *3 (quoting ACT arbitration agreement form provided to examinee that outlined the terms of arbitration to be conducted by the American Arbitration Association).

111. Scott, 600 A.2d at 502 (quoting ETS arbitration agreement form submitted to examinee that provided for arbitration under American Arbitration Association Commercial Arbitration Rules).

112. 297 N.Y.S.2d 472.

113. Id. at 473.

114. Id. at 474.

115. Id. at 473.

116. Id.

117. As the Second Circuit Court of Appeals has stated, “[p]erhaps the single most important prerequisite for the issuance of a preliminary injunction is a demonstration that if it is not granted the applicant is likely to suffer irreparable harm before a decision on the merits can be rendered.”
generally "[took] into consideration the legal and equitable principles of law applicable to the granting of preliminary injunctions," 118 namely that, "[when ruling on such motions] courts must weight [sic] the interests of the general public as well as the interests of the parties to the litigation." 119 This "weighing" of the interests led the court, without comment on the interests of the examinee, to conclude that requiring the examinee to retest was "within [ETS's] rights and indeed within its obligations and duties to the Academy and to the public," 120 and therefore ETS would not be restrained from invalidating the examinee’s scores and notifying the Academy. 121

Seven years later, the Supreme Court of New York County, in In re K.D. v. Educational Testing Service, 122 faced with "[a]lmost the identical issue" decided in DePina, enlarged upon the policy grounds for denying a restraining injunction. It likewise found ETS’s offer to allow the LSAT examinee to retest an adequate remedy, calling it "eminently fair and reasonable under the circumstances." 123 Though In re K.D. was a more well-reasoned, factual, and expansive opinion than DePina, its policy basis for denying an injunction was based on the same assumption made by the DePina court—that the testing agency was fulfilling its duty to maintain reliability in providing "a highly valuable service." 124

Similarly, in 1990, the New York Supreme Court Appellate Division, in Yaeger v. Educational Testing Service, 125 affirmed the trial court’s summary judgment dismissal of an examinee’s suit seeking an injunction prohibiting ETS from canceling her National Teacher Examination scores. 126 As with In re K.D. and DePina, the court made no mention of the testee’s equitable interests in retaining her scores, citing only to ETS’s interests in maintaining testing validity. 127 Once again, a one-sided "weighing" of the interests at stake suggested a strong judicial presumption not only that the testing agency had good reason to challenge the invalidated scores but also that the examinee was actually

---

118. *DePina*, 297 N.Y.S.2d at 474.
120. *Id.*
121. *Id.*
122. 386 N.Y.S.2d 747.
123. *Id.* at 752.
124. *Id.*
125. 551 N.Y.S.2d 574.
126. *Id.* at 575.
127. *Id.* at 576–77.
guilty of cheating.\textsuperscript{128}

In the somewhat different context of a group invalidation dispute, \textit{Doe v. The National Board of Podiatric Medical Examiners (NBPME)},\textsuperscript{129} the United States District Court for the Southern District of New York recently provided an object lesson on the steep odds plaintiffs face when seeking preliminary injunctions against testing agencies.\textsuperscript{130} In \textit{Doe}, podiatry students, whose board scores were invalidated based on general allegations of leaked test questions, sought a court order compelling the National Board of Podiatric Medical Examiners and ETS subsidiary Chauncey to release their scores and certify them as valid.\textsuperscript{131} The court stated that in order to obtain a preliminary injunction, the plaintiffs would have to show: “(1) irreparable harm in the absence of the injunction and (2) either (a) a likelihood of success on the merits or (b) sufficiently serious questions going to the merits to make them a fair ground for litigation and a balance of hardships tipping decidedly in the movant’s favor.”\textsuperscript{132}

The court further declared that when the plaintiffs sought a mandatory (rather than prohibitory) injunction “to alter the status quo,”\textsuperscript{133} and a remedy that essentially provided the ultimate relief they were seeking (validation and release of their scores),\textsuperscript{134} then they had to meet the still higher burden of “substantial, or clear showing of, likelihood of success to obtain preliminary relief.”\textsuperscript{135} The court concluded that the plaintiffs were unable to meet this burden, citing prior case law in which courts, based on judicial deference to testing agencies, denied relief to suspected examinees.\textsuperscript{136} The court also held that the plaintiffs could not “show that ‘extreme or very serious damage [would] result from a denial of preliminary relief.’”\textsuperscript{137} Downplaying the havoc retesting had wreaked on the lives of the students,\textsuperscript{138} and the fact that

\begin{itemize}
\item[128.] See \textit{id.} (opining that “the record contains adequate evidence to support ETS’ determination to cancel the petitioner’s scores on the ground of questionable validity,” but offering no examples of that evidence).
\item[129.] 2003 WL 21403698.
\item[130.] \textit{id.} at *1.
\item[131.] \textit{id.}
\item[132.] \textit{id.} at *2.
\item[133.] \textit{id.}
\item[134.] \textit{id.}
\item[135.] \textit{id.} at *3 (quoting \textit{Tom Doherty Assocs. v. Saban Ent.}, 60 F.3d 27, 35 (2d Cir. 1995)).
\item[137.] \textit{id.} (quoting \textit{Doherty}, 60 F.3d at 35).
\item[138.] The examinees claimed a litany of hardships including adverse effects on “completing their graduation requirements; participating in externships, internships, and residency programs; and obtaining a license for the practice of podiatric medicine.” \textit{id.} at *2. The court admitted that the students had to sit for the retest while “taking a full load of classes and participating in internship
five of the sixty students who passed the invalidated first test failed on the retest, the court emphasized the harm that would befall the testing agency's ability "to make predictions of competency, acumen, and performance based upon its expertise," and stressed the "valuable service" that the agency provided to medical institutions and the public health in general.

Doe shows how high standards for preliminary injunctions, coupled with established judicial deference to testing agencies, converge to form a difficult barrier for plaintiffs to overcome. Consequently, in order for an examinee plaintiff to obtain a preliminary injunction, she may have to demonstrate blameworthy or morally questionable conduct on the part of the testing agency that will tip the balance of equities in her favor.

programs. Id. Nonetheless, the court opined that the "[d]efendants mitigated the harms suffered by plaintiffs by offering a specially-scheduled free of charge retest." Id. at *6. The court emphasized the positive stating that "[a]ll but five students passed [the retest]." and characterized seemingly reasonable concerns by some students that their passing but lowered scores would hurt them in their employment search as "speculative." Id.

139. Id.
140. Id. at *7.

141. In two cases in which testing plaintiffs succeeded in obtaining preliminary injunctions, Cortale v. Educ. Testing Serv., 656 N.Y.S.2d 154 (Sup. Ct. Suffolk County 1997) and Mindel v. Educ. Testing Serv., 559 N.Y.S.2d 95 (Sup. Ct. N.Y. County 1990), plaintiffs were able to show, at a minimum, irresponsible conduct on the part of ETS. In Cortale, an examinee sought a preliminary injunction preventing ETS from canceling her GRE score based upon allegations that she copied from another test-taker. 656 N.Y.S.2d at 156. The court granted her preliminary injunction and denied ETS's motion for summary judgment in large part on facts that ETS destroyed plaintiff's test booklet, in which plaintiff had made notes supporting her answers, after it began its investigation into the validity of her scores. Id. at 157. After finding that the plaintiff had offered persuasive proof that she would be irreparably harmed by the cancellation of her score, the court held that ETS's "destruction of evidence [...] created both a reasonable probability of success and balancing of equities in favor of the plaintiff." Id. In Mindel, the court granted a mandatory injunction ordering ETS to cancel the score of an SAT testee and provide her an expedited retest. 559 N.Y.S.2d at 99. As a factual predicate for its decision, the court cited several compelling, even shameful facts showing that the administering of the examinee's SAT test was seriously flawed and unfair, including undisputed evidence that the examinee's test booklet was borrowed from her by a proctor while she was taking the test, and that a small child was allowed to enter the testing room several times during the exam. Id. at 97. The court then found that the plaintiff's claim "met the three basic requirements for a preliminary injunction." Id. at 98. First, the court found the plaintiff would be injured irreparably by the "loss of the opportunity [to gain an early decision from the elite colleges of her choice] due to the passage of time" if she were not given an immediate special opportunity to retest and improve her already excellent scores. Id. Second, the court found that the plaintiff's case evinced "a likelihood of ultimate success." Id. Third, the court found that the balance of equities favored the plaintiff. Id. In support of this last conclusion, as well as its decision to decide the matter in favor of the plaintiff without need for a trial, the court pointed to the "the potential harm to the plaintiff with no corresponding detriment to defendant." Id. at 98-99.
V. AVAILABLE LEGAL AVENUES IN STANDARDIZED TESTING INVALIDATION CASES

The following analysis addresses the two principal legal approaches that standardized testing plaintiffs have taken in challenging invalidations—breach of contract (or unconscionability of contract) claims and violation of due process claims. Though some standardized testing plaintiffs have also brought tort actions, alleging damages for defamation, interference with contract, or even outrageous conduct, recorded cases largely show these claims to have been peripheral and unavailing.

A. Contract Claims

Though the circumstances surrounding most standardized testing cases raise obvious issues of adhesion and good faith dealing, contract law has not been an especially fruitful source of relief for examinees suspected of cheating. Still, contract law has produced some minor victories for examinees. This subsection will first discuss the nature of the testing "contract," and then discuss the efficacy of arguments based on principles of adhesion and good faith.

142. Unless a testing agency publicizes its invalidation of a testee's score to an individual or organization wholly unrelated to the testee's educational use of the test scores, publication will likely be deemed privileged. For example, in Langston, the Eleventh Circuit Court of Appeals stated that any communication made "pursuant to a duty owed either to the public or to a third party, or where the statement is one in which the speaker and the third party have corresponding interests," would be privileged. 890 F.2d at 387. Under this standard, the court held that ACT's communications to the examinee's guidance counselor, who had the duty of "posting ACT test scores on student's transcripts and forwarding them to colleges" were privileged. Id. Thus the plaintiff would have to show that the communications were made with malice. Id. The Langston court did not even bother to address the unlikely possibility that ACT could have acted with malice, but rather granted summary judgment to ACT on the plaintiff's defamation claim after its discussion of privilege and a bare statement of the malice standard. Id. See also Johnson, 615 F. Supp. at 615 (holding that defendant ETS's communications with law schools regarding a suspected examinee's questioned scores were privileged).

143. The court in Johnson summarily dismissed this claim, stating that a showing of interference with contract required the plaintiff to demonstrate that the interference was "intentional and without justification," and concluding without further discussion that ETS's communication with the recipient school in that case was justified, 615 F. Supp. at 639.

144. Considering the standard required for a showing of outrageous conduct as stated by the Langston court, that is, conduct "outside the bounds of decency and utterly intolerable in a civilized society," 890 F.2d at 387, it would appear that this claim, absent the most egregious misconduct imaginable by the testing agency, would not have the proverbial snowball's chance.

145. Supra nn. 142-144.
1. Testing “Contracts”

Before a prospective test-taker sits for a standardized college entrance, graduate entrance, or professional licensing exam, that test-taker must enter into a contract with the testing service formulating and administering the exam. This contract between the parties is formed when the test-taker signs the test registration form and ostensibly agrees to be bound by the terms of the registration booklet or bulletin.

The registration bulletin or booklet provided by the testing agency constitutes a standard form contract, which outlines its terms in language exactly the same for every examinee taking that version of the test. The booklet contract generally contains provisions warning test-takers against reproducing any part of the exam, or engaging in any number of prohibited actions during the exam, such as looking back at sections of the test already completed, giving or receiving help from other test-takers, or viewing other test-takers’ answer sheets or booklets.

Typically, the contract-booklet further informs prospective examinees: 1) that should it be suspected that information has been reproduced from an exam, scores from that exam may be invalidated; 2) that the testing agency has the exclusive right to determine the validity of an exam; 3) that the testing agency has the right to cancel a score if it has reason to suspect the score is invalid (for example, “if there is an apparent discrepancy in photo identification, if the student engages in misconduct, [or] if there is a testing irregularity ...”); and 4) that the test-taker must adhere to the testing agency’s procedural guidelines.

146. See Murray, 170 F.3d at 515 (addressing contract formation in the college entrance examination context); Doe, 2003 WL 21403698 at *4 (addressing contract formation in the licensing/residency admissions testing context); Koza, 2001 WL 1191050 at *7 (addressing contract formation in the college entrance examination context); Cortale, 251 A.D.2d at 529–530 (addressing contract formation in the graduate school entrance examination context).

147. E.g. Murray, 170 F.3d at 515.

148. It would appear that there may be some variations in testing contracts between jurisdictions. See Dalton, 663 N.E.2d at 290 (noting that an SAT test-taker agreed to conditions contained in a “New York State edition of the Registration Bulletin”).

149. E.g. Doe, 2003 WL 21403698 at *4 (noting the NBPME bulletin’s admonition that, “[a]ny attempt to reproduce all or part of an examination is strictly prohibited”).

150. E.g. Koza, 2001 WL 1191050 at *7 (reproducing a passage from ACT’s “‘Test Security Procedures’” which contains these and other prohibitions).

151. E.g. Doe, 2003 WL 21403698 at *4 (quoting the NBPME test bulletin’s notice that, “examination scores may be invalidated in the event of this type of suspected [sharing of information] breach”).

152. E.g. id. (quoting contract language stating that the NBPME reserved “‘the sole right to determine whether or not an examination is valid or invalid’”).

153. E.g. Koza, 2001 WL 1191050 at *7 (stating, “ACT reserves the right to cancel test scores when there is reason to believe the scores are invalid”).

154. Murray, 170 F.3d at 515 n. 1 (quoting the ETS’s SAT bulletin provisions).
Testing booklet contracts generally present a list of options available to a test-taker who wishes to challenge a testing agency determination. For example, when an SAT test-taker's result "[was] questioned because it may have been obtained unfairly," the test-taker, under the terms of the booklet contract, was given "five options: (1) the opportunity to provide additional information, (2) confirmation of the score by taking a free retest, (3) authorization for [the testing agency] to cancel the score and refund all fees, (4) third-party review by any institution receiving the test score or (5) arbitration."156

2. The Issue of Adhesion in Testing Contracts

A number of plaintiffs challenging testing agency determinations and seeking to void unfavorable dispute resolution provisions in their testing contracts have argued that standard-form agreements between testing agencies and prospective test-takers amount to adhesion contracts.157 At least according to general principles of contract law, this conclusion, recognized by several courts,158 seems inescapable because educational testing contracts amply meet both the generally-accepted definition of an adhesion contract and the individual elements which constitute that definition.

Black's Law Dictionary defines an "adhesion contract" as: "A standard-form contract prepared by one party, to be signed by the party in a weaker position, usu[ally] a consumer, who has little choice about the terms."159 No doubt, standardized testing contracts match this definition, since the contract booklets providing terms are prepared unilaterally by the testing services, and the accompanying registration forms are signed by student consumers who not only have little choice but no choice at all regarding the terms of the contract.160

155. E.g. Koza, 2001 WL 1191050 at *8 (quoting an explicit provision in an ACT registration booklet stating that the test-taker "agreed to 'abide by all procedures and requirements stated [in the booklet], including those concerning test score cancellation and binding arbitration'").

156. Dalton, 663 N.E.2d at 290. See also Murray, 170 F.3d at 515 n. 2 (quoting the ETS's SAT bulletin provisions).


160. In re K.D., 386 N.Y.S.2d at 751 (noting that an examinee could not conceivably "indicate to [the testing agency] that the terms contained in the Bulletin were not acceptable to him," since if he did so, he would not be allowed to take the exam).
A more detailed hornbook articulation of the elements implicit in the Black’s Law Dictionary definition also suggests that educational testing contracts are contracts of adhesion:

There are at least three distinct possibilities, which often appear in combination [in an adhesion contract]. First, bargaining over terms may not be between equals. The standardized contract may be used by an enterprise with such disproportionately strong economic power that it can dictate its terms to the weaker party. Second, there may be no opportunity to bargain over terms at all. The standardized contract may be a take-it-or-leave-it proposition in which the only alternatives are adherence or outright rejection. Third, one party may be completely, or at least relatively, unfamiliar with the terms. The standardized contract may be used by a party who has had the advantage of time and expert advice in preparing it while the other party may have no real opportunity to scrutinize it.

Upon examination, all three elements in this expanded definition closely mirror educational testing contracts.

First, regarding the relative bargaining power of the contracting parties, the contracting student is in a dramatically weaker position than the testing agency. To achieve the goal of entering a particular university, the individual student by absolute necessity must participate in the testing process and receive a favorable score. There is generally only one game in town—the testing agency’s test—by which the student can achieve this goal. By contrast, common sense suggests that the testing agency would suffer an infinitesimally negligible loss if the potential examinee were to choose to reject the contract booklet terms and forego the test.

Second, the testing agency contract is offered to the potential examinee on a “take it or leave it basis.” The examinee has no latitude to cross out provisions in the booklet contract or pencil in changes. The only choice available to the examinee is between accepting the contract in its totality or rejecting it outright.

Third, the average test-taker no doubt reads the booklet agreement as a set of rules to be followed rather than as a series of contract provisions to be thoughtfully considered before approval. By contrast, the other

162. As one court has stated, “[s]ince these exams are required by almost all accredited institutions, candidates have no choice but to take them on the terms offered.” Doe, 2003 WL 21403698 at *5 n. 3.
163. In re K.D., 386 N.Y.S.2d at 751 (noting that an examinee “could [not] contract with a party other than the [testing agency] to take a law school aptitude test, since no such entity exists”).
164. Id.
party to the contract, the educational testing agency, enjoys the combined benefits of time, legal assistance, and past experience to craft the terms of educational testing contracts to favor their interests and withstand challenges under the law.\textsuperscript{165}

Despite these clear indications that educational testing contracts are contracts of adhesion, many courts of appeal have not characterized them as such, much less made a finding that they should be “closely scrutinized” or invalidated as unconscionable.\textsuperscript{166} Rather, a number of courts have not addressed the issue at all,\textsuperscript{167} while others have summarily dismissed the allegation,\textsuperscript{168} noted the issue only in passing,\textsuperscript{169} or rejected the claim that the adhesion contract was unenforceable under a common law test requiring proof that the disputed adhesion provisions themselves were substantively unreasonable.\textsuperscript{170}

In a recent case representing this last approach, \textit{Koza v. ACT},\textsuperscript{171} the Court of Appeals of Michigan ruled that under Michigan law, the plaintiff could not meet his burden to show that an arbitration provision in an educational testing agreement should be invalidated as a term of an adhesion contract.\textsuperscript{172} The court applied “a two-prong test of procedural and substantive unconscionability” as follows: “1) What is the relative bargaining power of the parties, their relative economic strength, the

\begin{quote}
165. See Nairn, supra n. 97, at 266–67 (discussing how ETS carefully considers the language of its contracts and that “public definition of the contractual rights of ETS consumers has been judiciously avoided”).

166. Cf. Martin, 431 A.2d at 874–75 (stating that an ETS real estate licensing exam contract, as “a contract of adhesion [...] must at [the] very least be closely scrutinized by the court to determine its reasonableness”).

167. Murray, 170 F.3d at 516–17 (upholding a testing contract with no discussion of the adhesion issue); Langston, 890 F.2d at 385–86 (rejecting a plaintiff’s contract claims under Alabama law with no mention of the adhesion issue); Cortale, 251 A.D.2d at 529–530 (discussing the provisions of an ETS testing contract with no mention of the adhesion issue).

168. E.g. Scott, 600 A.2d at 503 (rejecting a motion judge’s determination that both “the ETS procedures for questioning scores and the arbitration agreement itself were unenforceable contracts of adhesion.” (quoting the motion judge of the New Jersey Superior Court Chancery Division, Bergen County)). The appellate court simply concluded that the “[p]laintiff was not compelled to arbitrate [...][w]hen she chose arbitration she bound herself by the arbitration agreement, including its terms as to the scope of the arbitrator’s authority and the procedures to be employed.” id.

169. See e.g. Doe, 2003 WL 21403698 at *5 n. 3 (quoting In re K.D., 386 N.Y.S.2d at 751 for the proposition that “contracts with an academic testing service would appear to fit the description” of an adhesion contract,” but not indicating whether or how this would affect the court’s analysis).

170. Koza, 2001 WL 1191050 at **9–10; see also In re K.D., 386 N.Y.S.2d at 752 (finding that an educational testing contract was a contract of adhesion, recognizing that the court could use various pretexts to invalidate its provisions if deemed unconscionable, but finding that provisions empowering the testing agency with the right to cancel the examinee’s test score if “there is a question about its validity,” and requiring the examinee to retake the test to confirm previous scores were not “so unfair and so unreasonable” that they should be nullified or disregarded).

171. 2001 WL 1191050.

172. Id. at *10.
alternative sources of supply, in a word, what are their options?; (2) Is the challenged term substantively reasonable? 173

The Michigan court found that the plaintiff had “presented evidence of procedural unconscionability,” 174 but held that the plaintiff had “provided no authority [showing] that the [disputed arbitration] provision [was] substantively unconscionable.” 175 The court therefore concluded that the plaintiff was bound by those provisions in the registration materials stating that any arbitration would be conducted solely through written submissions and limited to an inquiry into “whether [the] defendant [testing agency] acted reasonably and in good faith in deciding to cancel the [examinee’s] scores.” 176

Because courts addressing the adhesion issue have almost uniformly refused to invalidate educational testing provisions on grounds of unconscionability, they have either explicitly or implicitly accepted the validity of educational testing contracts in their formation and terms. 177

3. Good Faith Adherence to the Testing Contract

Once courts have affirmed the validity of standardized testing contracts, their attention is focused on whether the testing agency performed the terms under minimum standards of good faith and fair dealing. 178 Though the results of arguments based on bad faith have not been especially fruitful, minor victories in recent cases suggest that such arguments may still hold promise.

The legal contours of the implied covenant of good faith and fair dealing, a duty owed by an any party to a contract, 179 have been described as “shrouded in mystery.” 180 Standards for good faith and fair dealing

173. Id. at **9–10.
174. Id. at *10.
175. Id.
176. Id. at **8, 10. Though the Koza court agreed with the examinee that the defendant agency could not insist on arbitration in Dallas, Texas, this was based upon the fact that the location for arbitration was simply not a provision of the contract rather than on any unconscionability ground. The court of appeals reversed the circuit court on this issue only, holding that arbitration should be held in the Detroit, Michigan metropolitan area, the hometown of the examinee. Id. at **10–11.
177. Cf. Martin, 431 A.2d at 874–75 (putting the words “agreement” and “contract” in quotes when referring to an ETS real estate licensing exam application and booklet contract, respectively, id. at 870, 874, and suggesting that a booklet contract provision that denied test-takers the right to inspect their graded exams could be voided as unconscionable. Id. at 874–75).
178. See infra notes 179–236 and accompanying text (discussing various aspects of good faith inquiries in testing invalidation cases).
179. Dalton, 663 N.E.2d at 291. “Implicit in all contracts is a covenant of good faith and fair dealing in the course of contract performance.” Id.
180. Thomas A. Diamond & Howard Foss, Proposed Standards for Evaluating When the Covenant of Good Faith and Fair Dealing Has Been Violated: A Framework for Resolving the
are notoriously nebulous and many court decisions based on this principle amount to "ad hoc" determinations "yielding inconsistent results and depriving parties of the ability to predict what conduct will violate the covenant." Accordingly, some courts and legal scholars have tried to formulate a workable framework for applying the principle of good faith and fair dealing. A full examination of these standards would prove far too lengthy for the purposes of this article. However, brief mention of some areas of conduct considered violative of the covenant may be useful in the testing context.

In *Dalton v. Educational Testing Service*, certain general principles of good faith and fair dealing were announced by the Court of Appeals of New York to support its decision that ETS in that case violated the covenant. The court, discussing the scope of the covenant, declared that "[e]ncompassed within the implied obligation of each promisor to exercise good faith are 'any promises which a reasonable person in the position of the promisee would be justified in understanding were included.'" The court continued, "[t]his embraces a pledge that 'neither party shall do anything which will have the effect of destroying or injuring the right of the other party to receive the fruits of the contract.'" The court, defining the standard to be applied when the contract "contemplates the exercise of discretion," stated that the controlling party had a duty "not to act arbitrarily or irrationally in exercising that discretion." The court cautioned, however, that there were limits to even this low level of scrutiny, and that no good faith obligation could be implied that was at odds with the express terms of the contract.

The approach taken by the *Dalton* court is characteristic of the good faith inquiry applied by most courts in testing invalidation cases—to determine whether the testing agency "arbitrarily or irrationally" evaluated the facts and circumstances surrounding a cheating allegation. Unfortunately, the issue of relative materiality of injury to the examinee

---


181. *Id.* at 586.
182. *Id.* at 600–32.
183. 663 N.E.2d 289.
184. *Id.* at 294–92. According to this author's research, *Dalton* appears to be the only invalidation testing case where a court made an attempt to provide the standards under which it was making its good faith and fair dealing determination.
186. *Id.* (quoting *Kirke La Shelle Co. v. Armstrong Co.*, 188 N.E. 163, 167 (N.Y. 1933)).
187. *Id.*
188. *Id.*
and testing agency, as well as whether options other than invalidation were available to the testing agency, have not been part of courts’ good faith inquiry, though standards proposed by certain legal scholars perhaps suggest that they should be. Rather, courts, through the policy arguments outlined earlier in this article, have simply provided conclusory statements that the reputation and reliability of testing agencies would be harmed if they were forced to release suspected test scores. Courts have thus reasoned that agencies have no choice but to invalidate suspected scores once they have a substantial basis for believing them to be unfairly earned.

Meanwhile, courts rejecting plaintiffs’ bad faith arguments have invariably addressed the facts of the cases before them on an ad hoc, standardless basis. In doing so, they have focused on the procedural options afforded an examinee, particularly the opportunity to retest, as well as the quality and nature of the investigation conducted by the testing agency.

For example, in Johnson v. Educational Testing Service, the United States District Court for the District of Massachusetts held that ETS’s decision to cancel an examinee’s dramatically improved LSAT score based upon its Test Security Office (TSO) investigation, consideration of the examinee’s proffered evidence, and discussions at four Board meetings, “was a reasonable one, reached with deliberation in good faith.” The court characterized ETS’s decision-making process as “lengthy” and noted that ETS had relied on three handwriting analysts’ opinions in coming to the conclusion that it was not the examinee who sat for her test, but an imposter. The court pointed out that prior to ETS’s full investigation, ETS had offered the testee the opportunity to “retest under special supervision.” The court further noted that ETS had reviewed the examinee’s “extensive documentation”

189. Under standards proposed by law professors Thomas Diamond and Howard Foss, a party to a contract (defendant) violates the principle of good faith and fair dealing if he, inter alia: 1) "ha[s] reason to know that his conduct would cause plaintiff material contractual injury unless the conduct was necessary to avoid material contractual injury to himself," Diamond & Foss, supra n. 180, at 602; 2) "engages in conduct that injures plaintiff's contractual interests if he ha[s] reason to know that there was an alternative which would have provided him essentially the same benefits while substantially reducing plaintiff's contractual injuries," id. at 609; or 3) dishonestly evaluates facts or circumstances under a contract which "confers discretion upon defendant to determine whether particular facts or circumstances exist." id. at 615. Diamond and Foss suggest that "material contractual injury" under this proposed standard occurs "when conduct defeats a party’s essential purpose for entering into the contract." id. at 602.

191. Id. at 638.
192. Id.
193. Id. at 637.
before coming to its conclusion. The district court granted summary judgment for the defendant and dismissed the breach of contract claim. The First Circuit Court of Appeals affirmed the district court's holding, relying on the same "indicia" of reasonableness and concluding "that ETS went beyond the letter of its contractual promise."  

Similarly, in Langston v. ACT, the Eleventh Circuit Court of Appeals pointed to American College Testing's "extensive" investigation of a high school football player's sizeable increase in his ACT score over a previous test and the "alarming similarity" between his current test answers and another examinee's. The court emphasized that the student had been given the opportunity to retest or submit to arbitration but had refused to do so. Thus, in granting a defense motion for summary judgment, the court held there was "no genuine issue as to whether ACT breached its obligation to act in good faith under the contract."  

However, two New York decisions mentioned above, Dalton v. Educational Testing Service (1995), and Cortale v. Educational Testing Service (1998) held that ETS's failure to consider evidence offered by examinee plaintiffs violated the covenant of good faith and fair dealing. These two relatively recent cases suggest that, on appropriate facts, New York courts may now be willing to more closely analyze the merits of whether the examinee actually cheated, at least in coming to a determination as to whether the testing agency acted "arbitrarily or irrationally," in making its invalidation decision.  

In Dalton, the New York Court of Appeals quoted with approval the trial judge's decision that the ETS Board of Review breached its contractual good faith obligations when it "failed to make even rudimentary efforts to evaluate or investigate the information" furnished

194. Id. at 638.  
195. Johnson v. Educ. Testing Serv., 754 F.2d 20, 26 (1st Cir. 1985) [hereinafter Johnson II]. The court used colorful language in affirming the summary judgment dismissal of the plaintiff's claim, stating, "a plaintiff cannot force a trial by pointing to smoke but not fire, and '[h]ere we do not even see any smoke.'" Id. (quoting Packish v. McMurtrie, 697 F.2d 23, 27 (1st Cir. 1983)). The court went further, stating that "[w]hile Johnson 'is entitled to all favorable inferences, [s]he is not entitled to build a case on the gossamer threads of whimsy, speculation and conjecture.'" Id. (quoting Manganaro v. Delaval Separator Co., 309 F.2d 389, 393 (1st Cir. 1962)).  
196. 890 F.2d 380.  
197. Id. at 386.  
198. Id.  
199. Id.  
200. 663 N.E.2d 289.  
201. 674 N.Y.S.2d 753.  
203. Dalton, 663 N.E.2d 289 at 293.
by [the suspected examinee]...204 Though the examinee presented verification that he was ill with mononucleosis during his first poor SAT showing, as well as eyewitness statements that he was actually in the testing room for the questioned second test,205 assertions amply supported by detailed and highly compelling evidence, the trial court found that the ETS Board completely ignored these proofs.207 According to the trial court, ETS deemed evidence confirming the examinee’s presence at his second SAT exam to be “irrelevant” because it viewed the examinee’s “discrepant scores” between the two tests and alleged handwriting discrepancies to be sufficient proof that the examinee had achieved the scores unfairly.208 Consequently, ETS concluded that the

204. 663 N.E.2d at 291 (quoting Dalton, 588 N.Y.S.2d at 745).
205. Id.
206. The trial court detailed the impressive proofs the suspected examinee submitted to ETS’s internal investigation as follows:

[The testee] informed ETS that he had been ill with mononucleosis during the [first] examination and submitted evidence of his academic abilities; that he had maintained an average of 85 and received second honors while at [his high school]. He also informed ETS that he had completed a test preparation or coaching course and he submitted to ETS the diagnostic tests administered by the Princeton Review, a “coaching course” in verbal, math and SAT test-taking skills which he attended in the period between the two examinations. Those diagnostic tests indicated test results consistent with his subsequent performance on the [second, challenged] SAT.

To support his denial that an imposter had taken the [second] SAT, [the examinee] furnished ETS with the report of a document examiner retained by [his] family who, disagreeing with the ETS examiner, found that [the examinee’s] handwriting matched both the May and November SAT examination sheets. Submitted to the test security office was the statement of the ETS paid proctor who supervised the administration of the SAT in the classroom assigned to [the examinee]... the proctor informed ETS that she specifically recognized [the examinee] at a subsequent meeting, arranged by the ETS SAT test administrator... as being present in the classroom on [the second test date]. She recalled his photo identification card... and remembered having reprimanded him for talking during a break in the examination. She also recalled [his family name]... [and] further informed ETS in her statement that [the examinee] had recognized her at this subsequent meeting and was able to detail her unique classroom instructions and procedures, including her requirement that students sign the roster sheet in her presence beside their printed names.

The proctor’s statement was accompanied by the statements of two students who also identified [the examinee] as being the person in the classroom on the day of the exam. One student, previously unacquainted with [the examinee], specifically stated that [the examinee] had stood out in the classroom that day because he was fair-complexioned and blue-eyed and exhibited “an attitude”, while the majority of the other test takers were Asian, African-American or Hispanic. The statement by the ETS paid SAT test administrator... confirmed that the examinee had correctly identified the proctor from other individuals present at the subsequent meeting. Lastly, [the examinee] offered to submit to a lie-detector test, at his own expense, and to have his fingerprints taken and compared to those latent fingerprints which ought to be found on the second answer sheet and test booklet.

Dalton, 588 N.Y.S.2d at 745–46.
207. Dalton, 663 N.E.2d at 291.
208. Id.
testee plaintiff "could [only] controvert the Board's preliminary finding that the [second improved] score was invalid . . . by taking a retest."209

Although the Court of Appeals affirmed the lower courts' finding that ETS's refusal to properly consider relevant evidence was a breach of its contractual duty of good faith, and therefore mandated specific performance, it nonetheless "differ[ed] as to the scope of the relief" provided by the lower courts.210 While the trial court directed ETS to release the examinee's challenged score, and the Appellate Division affirmed,211 the Court of Appeals refused to do so. The court concluded that the plaintiff was only due the specific performance accorded him under the contract—"good faith consideration of the material he submitted to ETS."212 The court rejected the plaintiff's argument that this was "an empty exercise" rather than a meaningful remedy and that ETS would simply "rubber stamp its prior determination without good-faith attention to his documentation."213 The court pointed to provisions in the examinee's contract which provided him with the options of "third-party review by any institution receiving the test score as well as arbitration," if ETS once more rejected his explanations.214

If one looks beyond the surface of the Dalton decision, a fair reading is that the examinee plaintiff's "victory" was a loss. Though the court suggested that the examinee plaintiff would still have viable options if ETS once again refused to change its opinion, this is doubtful. First, it is highly unlikely that the examinee would wish to submit his case to a university inquiry for the obvious reason that this would defeat the purpose of his challenge—to prevent ETS from damaging his college prospects. And, as noted earlier in this article, his other option, arbitration, would be limited to a determination only on whether ETS had a substantial basis for challenging the scores, a low burden for ETS to meet. Thus, as a result of the court's decision, not only would the plaintiff be accorded no forum in which to clear his name, but also would quite likely end up having his score invalidated. On the whole then, the outcome in Dalton was not particularly encouraging for testing invalidation plaintiffs.

Dalton may also be a negative case for plaintiffs in that the court provided key limiting language accompanying the partial relief it granted. In particular, the court emphasized that, according to the testing

209. Id.
210. Id. at 291.
211. Id.
212. Id. at 294.
213. Id. at 291.
214. Id.
contract, "ETS was under no duty, express or implied, to initiate an external investigation into a questioned score." Additionally, the court stressed that "[n]othing in the contract compelled ETS to prove that the test-taker cheated," and that any such requirement "would be inconsistent with the contractual language placing the burden squarely on the test-taker to overcome the ETS finding of score invalidity." The court pointedly noted that the trial judge had declined to decide the issue of whether or not the examinee himself actually took the test, in essence, whether or not he cheated.

By contrast, the New York Supreme Court Appellate Division's *Cortale* opinion contains no such limiting language. In *Cortale*, a GRE testee improved the verbal portion of her score across successive tests by more than two hundred points, raising ETS's red flags and prompting a preliminary investigation. Based upon a statistical analysis comparing her answers to another test-taker's, who "may have been seated near her," ETS came to the conclusion that the *Cortale* plaintiff had copied her co-examinee's answers. After being accused of cheating, the plaintiff submitted considerable exculpatory evidence, but ETS declined to change its determination. In the meantime, the plaintiff had been accepted to a graduate university program. When ETS notified the examinee plaintiff that it intended to inform her university that it was canceling her GRE score, the examinee brought suit seeking a permanent injunction restraining cancellation and a declaratory judgment validating her score. ETS filed a cross motion for summary judgment dismissing the complaint, which was denied by the trial court. The plaintiff won a preliminary injunction pending a final determination of the matter, and ETS brought an appeal on the denial of summary judgment only.

The Appellate Division upheld the trial court's ruling denying ETS

215. *Id.* at 292.
216. *Id.* at 292.
217. *Id.* at 294.
218. *Cortale*, 251 A.D.2d at 528.
219. *Id.*
220. *Id.* For example, the examinee submitted evidence to ETS showing that she had suffered an injury to her hand prior to the first examination for which she was taking a prescription narcotic for pain. Additionally, she offered evidence showing that her academic abilities were more consistent with the higher scores she achieved on her second examination. *Id.*
221. *Id.*
222. *Cortale*, 656 N.Y.S.2d at 156.
223. *Id.; Cortale*, 251 A.D.2d at 529.
224. *Id.*
225. *Id.*
summary judgment.\textsuperscript{226} In markedly different language from its \textit{Yaeger} decision eight years earlier,\textsuperscript{227} the court began its analysis by bluntly proclaiming, “ETS, in reliance upon purely circumstantial evidence, determined that the plaintiff was guilty of cheating.”\textsuperscript{228} The court, denying summary judgment, held that a question of fact existed because ETS based its determination “largely upon a statistical analysis of [a] pattern of erasures and incorrect answers appearing on the plaintiff’s answer sheet,” and the plaintiff presented “expert evidence which attacked those statistical analytical methods as unreliable.”\textsuperscript{229} Moreover, the court noted that one of the studies submitted by ETS, “ostensibly in support of its motion” actually called into question the reliability of ETS’s own investigative policies, suggesting that ETS tended to ignore submissions offered by examinees in support of their questioned scores.\textsuperscript{230}

The \textit{Cortale} opinion, though providing helpful language for testing plaintiffs, should not be overestimated. The court ruled solely on the appropriateness of summary judgment dismissing the good faith claim,\textsuperscript{231} as opposed to the \textit{Dalton} court which ruled on the good faith claim itself.\textsuperscript{232} Thus, the \textit{Cortale} court was not compelled to reconfirm the limited remedies available to a test-taker showing breach of good faith by the testing agency—further investigation, third-party review by the plaintiff’s chosen university, or arbitration. Nonetheless, the court nowhere suggested, even in dicta, that the underlying issue was only whether ETS had a substantial basis for challenging the score.\textsuperscript{233} On the contrary, the court referenced “the plaintiff’s efforts to clear her name and prove her entitlement to her score.”\textsuperscript{234}

Moreover, the court, breaking with the tradition of earlier testing invalidation cases, stressed the interests of the test-taker, departing from well-worn policy mantras advocating deference to testing agencies. The court used strong language in this regard, citing “the serious educational and vocational ramifications that may flow from a finding that the plaintiff cheated.”\textsuperscript{235} Though this last statement was made in response to egregious facts that ETS destroyed the examinee’s test booklet

\begin{itemize}
  \item \textsuperscript{226} \textit{Id.}
  \item \textsuperscript{227} \textit{Supra} nn. 126–128 and accompanying text.
  \item \textsuperscript{228} \textit{Cortale}, 251 A.D.2d at 530.
  \item \textsuperscript{229} \textit{Id.}
  \item \textsuperscript{230} \textit{Id.}
  \item \textsuperscript{231} \textit{Id.}
  \item \textsuperscript{232} \textit{Dalton}, 663 N.E.2d at 291–95.
  \item \textsuperscript{233} \textit{Cortale}, 251 A.D.2d 528.
  \item \textsuperscript{234} \textit{Id.} at 530.
  \item \textsuperscript{235} \textit{Id.}
\end{itemize}
containing scratch work that might have substantiated her innocence,\textsuperscript{236} it nonetheless stood as an indicator that under appropriate facts, deference to the testing agency might give way to sympathetic consideration of the examinee’s interests.

\textit{Cortale} notwithstanding, even if the testing agency is compelled by its duty of good faith to honestly consider all of the plaintiff’s proffered evidence, the nature and scope of an invalidation inquiry will be limited to the options outlined in the booklet registration materials. This means that a suspected examinee may never be able to obtain a judgment on the underlying grounds for an invalidation—the accusation of cheating. As one court succinctly put it, “The issue before this court is not whether or not [the examinee] cheated on the test; the issue is whether or not [the testing agency] could refuse to release the score.”\textsuperscript{237} Or as another court stated, “Under the governing law, the outcome of plaintiff’s case does not turn on whether or not plaintiff cheated on his exam, but only on whether or not [the testing agency] carried out its contractual obligations in good faith.”\textsuperscript{238}

\textbf{B. Deprivation of Due Process Claims}

Perhaps not surprisingly, the narrow framing of the legal inquiry under contract law has led a number of plaintiffs to allege violations of their due process rights. These arguments likewise are generally unsuccessful, as many courts fail to classify testing agency conduct as state action subject to due process. In some instances, courts have acknowledged that testing agencies may be state actors; however, because of deference to testing agencies, courts are still likely to rule that testing agency conduct satisfies the demands of due process.

The benchmark due process case in educational testing invalidation

\begin{itemize}
\item \textit{Id.}\ ETS not only destroyed the suspected examinee’s booklet but also that of “Candidate B” who the examinee had allegedly copied from. \textit{Cortale}, 656 N.Y.S.2d at 156. The destruction of this evidence alone was a sufficient ground for the trial court to find that a question of fact existed as to whether ETS investigated in good faith:

The question of destruction of evidence was not addressed by the [New York Court of Appeals] in Dalton or other similar cases. Defendant has asserted that all test booklets are destroyed in the normal course of its business. Although litigation had not been commenced in connection with the examination (at the time the materials were destroyed) the propriety of destroying material which is arguably relevant to an ongoing investigation and a future arbitration proceeding (provided in the contract) is, at best, questionable. At the very least, this Court finds that it creates a triable issue as to defendant’s good faith efforts in discharging its obligation to fully investigate the question of plaintiff’s academic dishonesty. Accordingly, the cross-motion for summary judgment must be denied.

\textit{Id.} at 157.
\end{itemize}

\begin{itemize}
\item \textit{Langston}, 890 F.2d at 385 n. 9.
\end{itemize}
law is *Johnson v. Educational Testing Service*. In *Johnson*, the First Circuit Court of Appeals held that a test-taker accused by ETS of having an imposter sit for her LSAT exam failed to state a viable due process claim against ETS under the Fourteenth Amendment. Consequently, the First Circuit upheld the district court's grant of summary judgment to ETS, and affirmed the dismissal of the examinee's due process claim.

The First Circuit explained that in order for the plaintiff to present a cognizable violation of due process, she would first have to establish that "ETS [was] a state actor and that its conduct was state action." The court noted that under Supreme Court law, the Fourteenth Amendment provided no protection against wrongful or discriminatory private conduct. Thus, a plaintiff alleging a violation of due process would have to show that the defendant’s actions "caus[ed] the deprivation of a federal right [that was] fairly attributable to the State." The court further explained that its "state action inquiry [was] two-fold: (1) whether 'the deprivation [was] caused by the exercise of some right or privilege created by the State or by a rule of conduct imposed by the State or by a person for whom the State [was] responsible,’ and (2) whether ‘the party charged with the deprivation [was] a person who may fairly be said to be a state actor.’" Therefore, according to this dual inquiry, even if a defendant were held to be a state actor, his actions would not necessarily be state actions.

The plaintiff argued that ETS was a state actor because by administering the LSAT, "a prerequisite to admission to all law schools," the agency "exercise[d] virtual veto power" over which applicants would be admitted to those law schools, many of them state institutions. The plaintiff further argued that under ETS's contract with the Law School Admissions Council (LSAC), the LSAC retained "ultimate responsibility" for the content and administration of the LSAT, and consulted with ETS regarding the overall "conduct of the program." Finally, the plaintiff pointed out that 45% of LSAC member schools were state institutions, and that a number of trustees on the ETS board were either public officials or "representatives of bodies that included..."
public institutions."\textsuperscript{249}

The First Circuit rejected the plaintiff's arguments that ETS was a
state actor. Although admitting that the plaintiff's "conception of state
action was arguably tenable when her complaint was filed" (prior to a
lengthy hiatus in district court)\textsuperscript{250} it concluded that recent Supreme Court
holdings set standards for finding state action that were "fatal to her
theory."\textsuperscript{251} The court relied on \textit{Blum v. Yaretsky}\textsuperscript{252} and \textit{Rendell v. Baker},\textsuperscript{253} in which the Supreme Court held that a nursing home and
private school, respectively, were not engaged in state action despite the
fact that an overwhelming amount of their funds were derived from
government coffers.\textsuperscript{254} The First Circuit quoted \textit{Rendell-Baker} as
follows:

The school, like the nursing homes [in Blum], is not fundamentally
different from many private corporations whose business depends
primarily on contracts to build roads, bridges, dams, ships, or
submarines for the government. Acts of such private contractors do not
become acts of the government by reason of their significant or even
total engagement in performing public contracts.\textsuperscript{255}

The court then noted that the Fourth Circuit, in \textit{Arlosoroff v. NCAA},\textsuperscript{256}
had recently held that the NCAA was not a state actor,\textsuperscript{257} because in
light of the Supreme Court's \textit{Rendell-Baker} and \textit{Blum} holdings, indirect
state involvement was no longer enough to convert private activity into
state action.\textsuperscript{258} The First Circuit reasoned that if the NCAA, with half of
its members state or federal institutions, was not a state actor, then \textit{a fortiori}, ETS was not a state actor.\textsuperscript{259} The court opined, "Whereas the
NCAA is capable of disqualifying an athlete from intercollegiate
competition, ETS merely reports test scores and lacks authority to decide
who shall be admitted and who shall be rejected."\textsuperscript{260}

---

\textsuperscript{249} \textit{Id.} at 24.

\textsuperscript{250} \textit{Id.} at 27. The court made reference to the fact that ETS's motion for summary judgment
had been "left pending by the district court for a long period of time," indicating a lapse of more than
twelve years. \textit{Id.}

\textsuperscript{251} \textit{Id.} at 24.

\textsuperscript{252} 457 U.S. 830 (1982).

\textsuperscript{253} 457 U.S. 830 (1982).

\textsuperscript{254} \textit{Johnson II}, 754 F.2d at 24.


\textsuperscript{256} 746 F.2d 1019, 1020 (4th Cir. 1984).

\textsuperscript{257} The Supreme Court, in \textit{NCAA v. Tarkanian}, later held that the NCAA did not engage in
state action when it promulgated rules that a state university, unquestionably a state actor, followed

\textsuperscript{258} \textit{Johnson II}, 754 F.2d at 24.

\textsuperscript{259} \textit{Id.}

\textsuperscript{260} \textit{Id.} (internal citation omitted).
The court concluded its analysis by finding that "the formulation, grading, and reporting of standardized tests [was] not an exclusive public function," and that the examinee plaintiff had presented no evidence "that public institutions belonging to ETS took the lead in instigating the conduct she challenge[d]." Therefore, the court held that the examinee’s due process argument could not overcome the higher standard for state action set by the Supreme Court’s recent decisions.

The Johnson court, however, cautioned in a footnote that its decision did not mean that ETS could never be deemed a state actor engaged in state activity. The court noted that in Martin v. Educational Testing Service, in which ETS administered real estate licensing exams and admitted that it acted as an agent of the Commonwealth of Pennsylvania, ETS was found to be a state actor. The court also cited Golden Rule Life Insurance Co. v. Mathias, in which an insurance licensure examinee was able to state a claim meriting a trial on violation of due process. The Golden Rule plaintiff alleged "that ETS designed examinations for licensure of insurance agents and brokers, graded examinations, determined who passed, and printed state licenses." The court in Golden Rule deemed this activity sufficient to show that the defendant ETS was engaged in state action and that the plaintiff’s due process claim could go forward.

The First Circuit’s footnote qualification was apparently ignored by at least one federal court piggybacking on the Johnson decision. Eight years after Johnson, the United States District Court for the District of South Carolina boldly and erroneously stated in a testing invalidation case, Tolleson v. Educational Testing Service, that "[t]he case law is uniform in holding there is no due process violation when the ETS reports test scores to state agencies. Neither the ETS’s acting alone nor operating in concert with the state or state agencies constitutes the

261. Id. at 25.
262. Id.
263. Id.
264. Id. at 25 n. 2.
265. 431 A.2d at 871 n. 7.
266. Johnson II, 754 F.2d at 25 n. 2.
268. Johnson II, 754 F.2d at 25 n. 2.
269. Golden Rule, 408 N.E.2d at 317.
270. See Tolleson, 832 F. Supp. at 158, 161 (D.S.C. 1992). The United States District Court for the District of South Carolina characterized the Johnson opinion as "the seminal case, which is very similar to the case at bar." Id.
271. 832 F. Supp. 158.
requisite ‘state action.’” 272 Though it relied heavily on Johnson, the Tolleson court apparently chose to disregard the Johnson footnote, sweeping Martin and Golden Rule under the rug.

This was a significant omission by the district court, because for the purposes of its state action inquiry, the facts of Tolleson resembled those of Martin and Golden Rule, at least to the extent that the testee sought admission to a profession for which a standardized ETS test was the sole avenue for admission. 273 In Tolleson, an examinee wishing to teach social studies in South Carolina failed the National Teacher’s Exams in that subject four times, but on his fifth attempt improved his score by two hundred points and was flagged by ETS for cheating. 274 In bringing a due process claim against ETS and the South Carolina Department of Education under 42 U.S.C. § 1983, the examinee averred “that the State of South Carolina, through the Department of Education and [the relevant state statute] empowered ETS to act with the full authority of the state.” 275 The district court, rejecting the plaintiff’s due process claim, admitted that the South Carolina “statute provide[d] that applicants seeking to teach in South Carolina [had to] take the NTE and that the ETS [was] the entity charged with reporting the test scores.” 276 Nonetheless, the court found that “ETS [was] a mere vehicle for reporting information; the ETS act[ed] only as a medium for conveying examination results. The ETS [had] no authority to determine certification or make any judgments as to the qualifications of applicants.” 277

Not only did the Tolleson court ignore Martin, Golden Rule, and the First Circuit’s footnote pertaining to them, but it also failed to consider key language in the district court’s Johnson opinion, undisturbed by the First Circuit on appeal, that distinguished Martin and Golden Rule. Specifically, the district court in Johnson stated that Martin and Golden Rule involved examinations administered by ETS for state licensing authorities which were the sole requirement for and the sole method of obtaining a real estate and insurance broker’s license, respectively, for practice in the state. . . . Here ETS’s role [in administering the LSAT] more closely resembles [another standardized testing case], where satisfactory performance . . . was only one requirement of several for

272. Id. at 164.
273. Id. at 158–59.
274. Id. at 159.
275. Id. at 160.
276. Id. at 160 n. 8.
277. Id. at 161.
those seeking to be high school principals, an oral examination and experience being necessary as well.\textsuperscript{278} Clearly this language at least bore consideration by the \textit{Tolleson} court, since the NTE administered by ETS in \textit{Tolleson} was "a prerequisite to certification" that all applicants had to pass in order to teach a given subject.\textsuperscript{279} Moreover, the court pointed to no other qualifications or testing requirements essential to obtain teaching certification.

The District Court for the District of South Carolina also failed to consider \textit{Scott v. Educational Testing Service},\textsuperscript{280} a case factually on point and decided just eight months before \textit{Tolleson}, in which the assumption that ETS conduct rises to the level of state action is implicit. In \textit{Scott}, a New York temporary public school teacher who received passing scores on her NTE came under suspicion by ETS, which concluded that the examinee had copied her answers on two sections of the exam from another test-taker.\textsuperscript{281} After being offered various options by ETS, the examinee chose to submit to arbitration, but requested an oral hearing in order to fully defend herself against the accusations.\textsuperscript{282} The arbitrator refused the request for an oral hearing and then decided against the examinee, ruling that ETS was free to cancel the suspected NTE scores.\textsuperscript{283} The examinee then filed an action to vacate the arbitration award, alleging violations of due process and including a claim for damages under 42 U.S.C. § 1983.\textsuperscript{284} On the return by ETS of an order to show cause why the arbitration award should not be vacated and the examinee's NTE scores reinstated, the motion judge ruled that "ETS was obligated to afford plaintiff due process because it was 'acting as an agent of New York City and the New York licensing authority.'"\textsuperscript{285} The judge found that the terms of the testing contract and arbitration agreement were "'unenforceable contracts of adhesion'\textsuperscript{286} and that they violated the test-taker's due process rights because they did not afford the test-taker the opportunity to contest the underlying charge of cheating.\textsuperscript{287} The judge then "vacated the award and ordered that a 'new in-person oral arbitration' be conducted in which ETS would have.

\begin{itemize}
\item \textsuperscript{278} \textit{Johnson}, 615 F. Supp. at 635 n. 5.
\item \textsuperscript{279} \textit{Tolleson}, 832 F. Supp. at 158–59.
\item \textsuperscript{280} \textit{Scott}, 600 A.2d 500.
\item \textsuperscript{281} \textit{Id.}, 600 A.2d at 501.
\item \textsuperscript{282} \textit{Id.} at 502.
\item \textsuperscript{283} \textit{Id.}
\item \textsuperscript{284} \textit{Id.}
\item \textsuperscript{285} \textit{Id.} at 502–03 (quoting the trial court without citation).
\item \textsuperscript{286} \textit{Id.} at 503 (quoting the trial court without citation).
\item \textsuperscript{287} \textit{Id.} at 501.
\end{itemize}
the burden of establishing that plaintiff had in fact cheated."\(^{288}\) ETS was granted both leave to appeal and a stay on the trial court’s order.\(^{289}\)

The Superior Court of New Jersey Appellate Division reversed the trial court, holding that the examinee was not denied due process because her “interests [were] fairly accommodated by a procedure which permit[ted] ETS to cancel scores upon an adequate showing of substantial question as to their validity, without any necessity for a showing of actual cheating or other misconduct.”\(^{290}\) However, in order to reach the merits of the due process issue, the court “assume[d], without deciding, that ETS’s conduct was ‘state action’ subject to the due process clause of the Fourteenth Amendment.”\(^{291}\) In marked contrast to Tolleson’s statement that the case law was “uniform” on the issue, the court observed that other courts had come to “variant results” as to whether national testing agencies were involved in state action depending on the facts of individual cases.\(^{292}\) The court noted, however, that those precedents were not determinative since the court would not be deciding the issue of state action based upon the trial record.\(^{293}\)

Ultimately, the court concluded that ETS’s procedures satisfied due process on policy grounds. The court added a new wrinkle to standard policy refrains on maintaining testing reliability by asserting that the examinee’s interests actually overlapped with the testing agency’s, since the examinee’s “effort to preserve her score [was] bottomed on the proposition that they [were] presumptively reliable.”\(^{294}\) Citing impressive statistical evidence provided by ETS showing an infinitesimal chance that the examinee’s scores were properly earned, the court declared, “Proof of wrongdoing is one way of establishing unreliability; but if unreliability is otherwise shown, an absence of proof as to how it came about is of no matter.”\(^{295}\) Concluding its due process/policy analysis, the court quoted from Langston v. ACT on the impracticality of requiring a testing agency to show that an examinee actually cheated:

> To demand that ACT prove by eyewitness testimony that an individual cheated before invalidating a score would undermine ACT’s primary function of providing colleges with scores that are highly reliable. ACT could not possibly catch every student who cheats in its exams if it had

\(^{288}\) *Id.* at 501 (quoting the trial court without citation).

\(^{289}\) *Id.* at 503.

\(^{290}\) *Id.* at 504.

\(^{291}\) *Id.* at 503 (citing Tarkanian, 488 U.S. 179; Blum, 457 U.S. 991).

\(^{292}\) *Id.*

\(^{293}\) *Id.*

\(^{294}\) *Id.* at 504.

\(^{295}\) *Id.*
to produce an eye witness [sic] to confirm every instance of misconduct.

In spite of its ultimate conclusion, Scott showed that even after the Supreme Court’s Rendell-Baker and Blum decisions, a court could entertain the possibility that ETS, at least in a licensing exam, was sufficiently involved in state action that its investigation and arbitration conduct would have to accord due process to an examinee. The Supreme Court’s more recent 5-4 decision in Brentwood Academy v. Tennessee Secondary School Athletic Assn.,297 in which the majority held that a state interscholastic athletic association’s regulatory enforcement activity was state action,298 apparently has not foreclosed this possibility.

Justice Souter’s Brentwood opinion could be seen as widening the permissible state action inquiry beyond a more narrow analysis predicated on criteria developed from a limited number of the Court’s cases.299 The Brentwood majority “identified a host of facts that [could] bear on the fairness of [ ] an attribution [of state action] . . . for example . . . exercise ‘of coercive power,’ . . . [or] when the State provides ‘significant encouragement, either overt or covert,’ . . . when a private actor operates as a ‘willful participant in joint activity with the State or its agents . . . when it is controlled by an ‘agency of the state,’ when it has been delegated a public function by the State, when it is ‘entwined with governmental policies,’ or when government is ‘entwined in [its] management and control.’”300

The Supreme Court’s cases suggest that a proper state action analysis of whether a testing agency’s invalidation of an examinee’s test results (in the most likely context, a state licensing exam) amounted to state action, would likely review closely a number of key factors. Did the testing agency, in addition to designing an examination for state licensure to a particular profession and grading that examination, determine who passed, and even print the state licenses, as the court

296. Id. at 505 (quoting Langston, 890 F.2d at 386).
298. Id. at 298.
299. The Supreme Court, in reversing the Sixth Circuit Court of Appeals, noted that that court, although recognizing “that there [was] no single test to identify state actions and state actors [ ] applied three criteria derived from Blum . . . Lugar . . . and Rendell-Baker . . . and found no state action under any of them.” Id. at 294 (citations omitted). The three factors applied by the circuit court were whether there was a “symbiotic relationship between the State and the Association,” whether the Association was engaged “in a traditional and exclusive public function,” and whether the Association was “responding to state compulsion.” Id.
noted in *Golden Rule*, or could it be deemed a "mere vehicle for reporting information... [with] no authority to determine certification or make any judgments as to the qualifications of applicants," as the court found in *Tolleson*. Potential factors, perhaps relevant in determining the degree of "entwinement" between the activities of the state and the testing agency in a given case might be: whether a passing score on the standardized test was the only requirement and only method for determining whether the examinee could be licensed for a profession within the state; whether the state interacted with the testing agency in designing the exam or had a hand in shaping its content; whether the test was formulated for a particular state’s licensing process or was a "generic" national exam, such as the Multi-State Bar Examination (MBE), or whether the testing agency or the state itself administered the test.

In the final analysis, however, it seems unlikely, based on testing law to date, that a plaintiff in an invalidation case can succeed on a due process claim, even if that claim survives dismissal and is reviewed on the merits, as in *Scott*. The *Scott* decision indicates that, absent some major shift away from the traditional deference paid testing agencies by the courts, and a concomitant standard that those agencies only need show a substantial basis for invalidating an examinee’s score, testing invalidation procedures established and implemented by testing agencies

---

301. 408 N.E.2d at 317.
303. In *Hoofer v. Ronwin*, an interesting testing case, an unsuccessful bar examinee sued the Arizona Supreme Court’s Committee on Examinations and Admissions, alleging, *inter alia*, that members of the commission had violated the Sherman Act by conspiring to restrain trade by reducing the number of attorneys who could practice law in the state, 466 U.S. 558, 565 (1989). The U.S. Supreme Court, in a 4-3 decision, held that the denial of the examinee’s application is the decision of the Arizona Bar was ultimately the decision of the Arizona Supreme Court and thus the doctrine of sovereign immunity applied to the state action of the Committee, rendering proper the dismissal of the examinee’s claim. *Id.* at 581–82. The dissent recognized that the Arizona Supreme Court had "delegated to [the Committee] the task of administering the bar exam, and retained the authority to review or revise any action taken by" the Committee, but argued that those factors were insufficient to confer immunity on the Committee under the Sherman Act. *Id.* at 593–94 (Stevens, J., dissenting).

Although the Court did not address whether ETS, which reported the multi-state portion of the exam, was a state actor, it did provide an explanation of the scaling process conducted by ETS, which was at the heart of the examinee’s complaint—i.e., that there was not a pre-set number of questions that an examinee had to get right to succeed on the exam, but a number determined after the test was "scaled," a process "viewed as the fairest" by ETS. *Id.* at 570. The Court quoted from material published by the National Conference of Bar Examiners, which stated that, in order to ensure "the same level of competence from test to test," tests needed to be scaled, "since the level of difficulty varie[d] from test to test." *Id.* (quoting the Bar Examiner’s Handbook 61–62 (2d ed. 1980)). The Handbook further explained that it was a "statistical analysis on [ ] reused questions [that] determine[d] how many points [were] to be added to or subtracted from the raw score to provide an applicant’s scaled score." *Id.*
are likely to continue to pass constitutional muster.

VI. CRITICAL ANALYSIS OF THE PREVAILING JUDICIAL APPROACH TO INVALIDATION CASES

Good name in man and woman, dear my lord,
Is the immediate jewel of their souls.
Who steals my purse steals trash; 'tis something, nothing;
'Twas mine, 'tis his, and has been slave to thousands;
But he that filches from me my good name
Robs me of that which not enriches him
And makes me poor indeed.  

Testing invalidation cases are complex disputes involving many factors, not least of which is the weighing of equities between the reputation of the test-taker and the reputation of the testing agency. Unfortunately for test-takers, courts have long emphasized the interests of testing agencies without a corresponding nod to suspected test-takers’ interests. The likely reason for this is that ETS and ACT are hugely influential corporations with carefully-crafted beneficent and infallible images that influence educational policymaking at all levels of government. Another reason is that, as a function of stare decisis, courts reviewing invalidation determinations have tended to parrot the policy statements of prior court opinions, without considering whether fresh evidence and social research may have eroded the assumptions underlying those policy positions.

A closer look at the standardized testing environment reveals that a number of key assumptions made by courts in forming their opinions may be based on fallacies. Perhaps the most important of these


305. See Naim, supra n. 97, at 276 (noting that despite ETS's gross inefficiencies in servicing examinee consumers, an ETS public relations director managed to turn a mechanical error blamed on a “crinkle in the corner” of an SAT testee's paper into an opportunity to tout publicly that “to [this director’s] knowledge it was the first time a mechanical error interfered with the accuracy of computer corrections at the 31-year old testing service, which processes 12 million pieces of paper a year.”) Id. (quoting In Short: Wrinkled by a Crinkle, Newsday 2 (July 31, 1979)). This infallible image has very recently suffered great damage as the result of widespread, substantial errors ETS made in scoring the Praxis teacher-certification exam, and Pearson Educational Measurement made in scoring the October, 2005 SAT exam. See Infra text accompanying nn. 344–348.

306. See infra n. 308 (discussing the current ETS President and CEO's efforts to lobby Congress on behalf of the No Child Left Behind Act); see also Lemann, supra n. 11 (describing numerous situations in which ETS sought to influence government policy in regard to education and testing); Nairn, supra n. 97, at 278–302 (describing ETS’s success in maintaining a non-profit status that allows it to elude almost any form of government oversight in conducting its operations).
assumptions are that: 1) testing agencies are altruistic non-profits whose primary purpose is to serve education and the public good; 2) testing agencies’ sole motive in invalidating scores is to maintain the integrity of scores relied upon by colleges, employers, and licensing institutions; 3) testing agency internal investigations are thorough and can be relied upon; 4) suspected examinees’ interests are not seriously damaged by invalidations, and in any event test-takers are offered the option to simply retest and confirm their scores; and 5) invalidating large score increases is an effective safeguard against cheating. The following section examines these suspect assumptions.

A. Testing Agencies’ Status as Altruistic Non-Profits Serving the Public Good

Contrary to popular characterization of testing agencies as altruistic non-profits serving the public good, standardized testing is very big business. This past year, ETS, the world’s leading testing agency, earned a record one billion dollars in revenues.\(^{307}\) No doubt, ETS did not earn this astounding figure merely by fulfilling obligations to test fairly and accurately, but by aggressively pursuing emerging markets and by seeking to promote the influence of testing in society.

As an example, in 2001, ETS lobbied Congress in favor of President George W. Bush’s No Child Left Behind proposal.\(^{308}\) Soon after, ETS boosted its overall yearly dollar revenues by 200 to 250 million dollars upon capturing a significant portion of the emerging K-12 testing market created by passage of the Act.\(^{309}\) The testing service’s subsidiary “ETS K-12 Works,” which has benefited so handsomely from this huge boom in K-12 standardized testing, is a for-profit corporation, as are two other ETS subsidiary corporations.\(^{310}\)

---


308. See H.R. Subcomm. on Educ. Reform of the Comm. on Educ. and the Workforce, Hearings on Measuring Success: Using Assessments and Accountability to Raise Student Achievement, 107th Cong. (Mar. 8, 2001) (available at http://cdworkforce.house.gov/hearings/107th/edr/account3801/landgraf.htm) [hereinafter Landgraf Educational Testimony]. Testifying before the House Education Reform Subcommittee, ETS President and CEO Kurt M. Landgraf opined, “I believe in the President’s plan. It is the right thing for our country, and it is doable.” \(\text{Id}\).

309. Interview by Marketplace, supra n. 307. According to Landgraf, this newfound market now constitutes “about twenty percent” of ETS’s overall revenue. \(\text{Id}\).

310. See Landgraf Educational Testimony, supra n. 308 (recording Landgraf’s testimony that ETS has “three for-profit subsidiary corporations. The Chauncey Group International develops and
As further evidence of this big business bent, one need merely consider the types of people attracted to serve on the boards of directors of these testing agencies, along with their impressive compensation packages. Though testing agency heads have historically been well-compensated for their efforts,311 recent pay scales are still more impressive. Kurt Landgraf, who took over as president and CEO of ETS in 2000, received almost $800,000 in compensation for his first ten months on the job, and the corporation has doled out giant bonuses to other officers and employees as well.312 Mr. Landgraf, a former CEO of DuPont Pharmaceuticals, typifies the new leadership at ETS—once “an entity staffed mostly by academics” but one that is now “run by executives recruited from the corporate world.”313

The incentive pay systems instituted by these executives are raising some eyebrows and calling into question the continued non-profit status of the ETS corporation.314 According to tax experts, it can be improper for a non-profit to “[cash] out most of its excess revenues in the form of bonuses.”315 Critics have charged that large incentive bonuses are particularly improper when they are derived from the fees paid by examinees who are captives of the standardized testing system.316

Further reinforcing the big business bent of testing agencies, competitors recognize that ETS may be benefiting unfairly by avoiding taxes and oversight from regulatory agencies.317 As John Katzman, CEO of Princeton Review put it, in characterizing the College Board and

administers occupational certification and professional assessment programs. ETS Technologies is devoted to developing and advancing technologies to support on-line learning and assessment applications. Our third subsidiary, called ETS K-12 Works, was created to provide testing and measurement services to the nation’s elementary and secondary schools”).


This is a billion-dollar commercial entity,” Mr. Landgraf said. “We’re an organization with a very strong social mission, but we are also a very large commercial enterprise. Our compensation is based on the simple principle that we have to attract people who can help us grow, and while we can never pay what DuPont or General Electric does, because we don’t have tools like stock options, we can use incentive pay and other cash payments.

Id.

313. Id.

314. Id.

315. Id.

316. Id.

317. See Nairn, supra n. 97, at 278–93 (describing how ETS’s non-profit status, clever public relations, and adept legal maneuvers provide it a shield against review by agencies like the FTC, and insure that “ETS remains almost immune to public oversight,” Id. at 284).
ETS’s entry into the test-prep market, “[T]hey’re the same sniveling, money-grubbing for-profit businessmen that the rest of us are.”318

Clearly, a characterization of testing agencies as altruistic non-profits serving the public good, if not misguided, does not tell the whole story. Were courts to perceive testing agencies as big business instead of as entities only serving the public good, they might be forced to reconsider many of the policy arguments behind their deference to testing agencies.

B. Motives Behind Invalidations

Courts have assumed that when testing agencies invalidate scores, they do so in order to protect the validity of representations made to universities, employers or institutions relying upon the agency’s ability to “accurately predict the aptitude of a candidate.”319 Courts, however, have failed to consider or address the fact that ETS also has a very strong interest in invalidating scores that might undermine the public’s belief in the reliability of the standardized testing process itself.

For many years ETS maintained that, as accurate indicators of knowledge and skills built up over years of study, standardized tests were very stable measures. Important to ETS’s contention that standardized tests accurately indicated knowledge and skills rather than test-taking ability was the notion that tests were impervious to test-prep coaching.320 Indeed, ETS made claims that even after many hours of test-prep classes, a student could improve his standardized score by only a couple of points at most, and therefore test-prep courses were largely a waste of time.321

Nonetheless, in living proof against these claims, companies like Kaplan and Princeton Review continued to grow and prosper, and privileged teenagers, ostensibly well-informed by their wealthy parents and peers, continued to take good advantage of test-prep coaching.322 At the same time that influential opinions like In re K.D. v. Educational Testing Service323 were being written by judges in the 1970’s, ETS still maintained that standardized test scores on their tests could not be increased by coaching.324 Since that time, however, ETS’s position has changed. In the face of numerous studies that coaching can improve scores significantly, even an average of 100 points on SAT’s, ETS softened its stand. Indeed, ETS not long ago went into business against

318. Critics Blast, supra n. 28, at A35.
320. See Stockwell, supra n. 80, at 6–10.
321. Id.
322. Id. at 3–5.
323. 386 N.Y.S.2d 747.
324. Id. at 7.
its test-prep adversaries and began to market its own test-prep materials. Yet even though ETS now admits that an examinee can improve SAT scores by perhaps 30 to 40 points through coaching, they continue to maintain, against the best evidence, that score increases of 100 points as a result of coaching are exceedingly rare.

Courts should recognize that score increases large enough to raise testing agency red flags are not only disturbing to testing agencies because they might indicate cheating by students, but also because, if legitimate, they strike at the testing agencies’ claims to reliability, stability, and the effective predictability of their tests. Additionally, they call into question the fairness of a system that allows those who can afford coaching to gain a significant advantage over those who cannot.

Perhaps it makes sense for courts to grant testing agencies the discretion necessary to protect themselves and their clients against cheaters. No test-taker should be allowed to gain unfair advantage by cheating at the expense of another test-taker, nor should universities or other institutions be misled into thinking that someone’s testing abilities were better than they actually were.

On the other hand, courts should not be in the business of protecting testing methodologies or empowering ETS to unfairly invalidate unusual scores simply because they upset ETS’s long-held contention that scores across successive tests cannot be significantly improved. This is a business interest, or at best an ideological interest, but certainly not an equitable interest. To the extent that courts are protecting such interests they are simply favoring one party’s well-being over the other party’s. Moreover, to assist testing agencies in covering up data that casts doubt on the reliability of testing methodologies makes a mockery of the very policy interest courts seek to protect—"the reliance that students, educational institutions, prospective employers and others place on the legitimacy of scores released by [the testing agency]."}

C. Reliability of Agency Evidence Indicating Cheating

Although courts have generally assumed the reliability of agency evidence indicating cheating, there are many reasons to suspect the validity of the most popular forms of evidence offered by testing agencies. Evidence concerning the statistical rarity of large score

325. Id. at 10–11.
326. Id.
327. Id. at 1–2.
328. Id. 1–2, 20–21.
increases and evidence concerning handwriting comparisons are particularly suspect.

Courts reviewing testing invalidation cases have largely accepted the testing agency position that large score increases across successive standardized tests are in and of themselves suggestive of wrongful activity. However, in light of the inherent conflict of interest testing agencies manifest in investigating large score increases, discussed in the previous subsection, courts should be more skeptical of this proposition. Moreover, as a matter of logic, one does not prove the unnaturalness of an event by its rarity alone. For example, it may be extremely rare that a falling brick strikes a pedestrian, but that does not lead to a conclusion that a malicious human being dropped the brick from a rooftop. In the same way, it is quite possible that an extremely large and correspondingly rare score increase achieved by an examinee across tests may be a result of a combination of unusual but innocent factors. Some factors that could account for an increase include prep-school studying following the first exam, the examinee’s lack of motivation on the first exam, the taking of medication before the first exam, an injury sustained before the first exam, test anxiety on the first exam, improper testing conditions on the first exam, personal problems

330. See e.g. Scott, 600 A.2d at 502 (noting that in its internal review, ETS “took into consideration the fact that only three test takers out of a sample of over 7,000 recorded gains of more than 42 points in [a] General Knowledge test”).

331. Koza, 2001 WL 1191050 at *2 n. 6 (quoting an examinee’s statement, “My first score was a mindless effort. This resulted from someone who didn’t care about his score and didn’t think it mattered.”).

332. Cortale, 251 A.D.2d at 529 (noting plaintiff’s claim that she was taking prescription narcotics prior to her first exam to lessen the pain from a hand injury).

333. Id.

334. Id.

335. Nairn, supra n. 97, at 273 (describing the case of an examinee who suffered from this condition).

336. See Law School Discussion, Anybody Have Testing Irregularities? I Did and I’m Complaining, http://70.84.78.174/prelaw/index.php?PHPSESSID=2d633edfe52b7a1de64c4e955fa6888b&topic=16273.msg246216 (accessed Mar. 1, 2006) (posting recent LSAT test-takers reports of irregularities and improper testing conditions at 2004 LSAT examinations, such as one student’s account as follows: “At my site the lead administrator came to the door and spoke to the proctor, then the proctor interrupted us in Section 2 without stopping the clock to tell us she needed for us to tear the bottoms of our tickets off and give her the top. Several people started to do this, others looked around in distress until someone finally said can we do this after the section? At least 1–2 minutes gone. I was working on RC, got distracted and had to start the passage over and didn’t finish 2 questions. She also interrupted us at least once per section to tell us to have our IDs out for checks and to clear the aisles so she could walk around “without tripping”); see also infra nn. 350, 351 (discussing website postings of students detailing numerous instances of testing irregularities and unfair testing conditions during standardized exams); supra n. 141 (discussing the improper test conditions described in Mindel, 559 N.Y.S.2d 95); infra n. 351 (discussing the author’s experience at an LSAT exam).
interfering with the examinee's concentration on the first exam, or allowance for the use of a calculator on the second exam.\textsuperscript{336}

The focus of the cheating inquiry, therefore, should be over the quality of the evidence the agency presents \textit{in conjunction} with a large score increase. Where this evidence consists of statistical evaluations regarding comparisons between the suspected examinee's sheet and another test-taker's, cases range from those in which the numerical odds against the integrity of the score are truly daunting\textsuperscript{337} to those in which courts have recognized that agency data may be questionable or even discredited\textsuperscript{338}.

Testing agency accusations of cheating by impersonation are generally far less convincing than accusations of copying, given the fact that they are based almost exclusively on supposed handwriting discrepancies between successive tests. Besides any doubts about the credibility of opinions given by agency handwriting experts,\textsuperscript{339} there are several reasons to view such accusations with skepticism. First, handwriting analysis, conducted as it too often has been by "experts" with dubious credentials, can be a notoriously inexact science.\textsuperscript{340} Second, handwriting samples produced months apart could vary due to changes in emotional immaturity, prescribed drug use,\textsuperscript{341} or nervousness. Third, far more reliable evidence, ignored by testing agencies, may be virtually conclusive of whether an examinee took his own test, for example, fingerprints on the testing materials themselves, or the credible eyewitness testimony of the proctors who administered the test.\textsuperscript{342}

\textsuperscript{336} Koza, 2001 WL 1191050 at *1 (The plaintiff claimed he "was able to make short cuts with [his] TI-82 calculator [that] cut the amount of time [he] spent on problems in half.").

\textsuperscript{337} See supra n. 81 (discussing convincing statistical evidence suggesting cheating by an examinee).

\textsuperscript{338} See Cortale, 251 A.D.2d at 530 (noting that "plaintiff adduced expert evidence which attacked [ETS's] statistical analytical methods as unreliable").

\textsuperscript{339} See Nairn, supra n. 97, at 273 (recounting how one examinee accused by ETS of cheating by impersonation showed at trial that ETS’s handwriting expert "was the same individual who had attested to the authenticity of the forged Howard Hughes signature on the celebrated Clifford Irving check").

\textsuperscript{340} Moenssens, supra n. 73, at n. 21 ("The appalling fact that graphologists without training in scientific questioned document examination succeed in hoodwinking the legal profession is further exemplified by the fact that the respected publisher of the multi-volume \textit{PROOF OF FACTS} recently commissioned the writing of the new chapter on questioned document examination—handwriting identification, to a graphoanalyst, Dorothy Lehman, who lacks standing in the field of forensic document examiners."); see \textit{U.S. v. Santillan}, 1999 WL 1201765, **2–5 (N.D. Cal. Dec. 3, 1999) (providing a discussion on the admissibility of handwriting analysis after \textit{Daubert}).

\textsuperscript{341} Nairn, supra n. 97, at 273 (recounting how a testee who suffered from test anxiety and took tranquilizers to improve her score believed that the tranquilizers she took affected the quality of her handwriting and led to ETS’s suspicions).

\textsuperscript{342} See \textit{Dalton}, 588 N.Y.S.2d at 745–46 (noting that consideration of these and other proofs were ignored by ETS in its internal investigation).
The heavy weight afforded these suspect forms of evidence is quite disturbing, especially in the context of a determination limited to the issue of whether a testing agency had a substantial basis to challenge a score or conducted its investigation in good faith. Were courts to afford examinees' proofs their proper weight within determinations on the merits of whether instances of cheating actually occurred, a more searching inquiry into the testing agencies' methods of investigation would likely result.

D. Weightiness of Test-takers' Interests and False Assumptions Concerning Re-testing

Contrary to general assumptions by courts that a test-taker's interest in relief is minimal, especially in light of the possibility of re-testing, the interests of suspected test-takers in the honesty of their scores are weighty. Aside from "the serious educational and vocational ramifications that may flow from a finding that [an examinee] cheated," an examinee also has an interest in his or her good name and self-esteem. These intangible interests stand apart from the worldly consequences flowing from score invalidation, but are no less worthy of consideration by a court ruling in equity. Thus far, however, few courts have even addressed what a suspected examinee stands to lose by an accusation of cheating (or in the case of group invalidation, by a mere association with an accusation of cheating).

First and foremost, courts undervalue a test-taker's interest in his or her good name. It may be said that an accusation of cheating is akin to a criminal charge, and in some cases it actually may be far worse, since certain passionate crimes may be morally justified, whereas cheating for self-advantage almost never is.

Though courts have repeatedly noted or assumed that the requirement that examinees retest in order to validate questioned scores (and their good name) is "eminently reasonable," and suggested that it is the suspected examinee who by refusing to retest is acting unreasonably, this presumption of a test-taker's guilt endorses a kind of "cheater's proof" logic and authoritarianism that might be objectionable to many Americans. In light of a presumption that the test-taker is innocent rather than guilty, a viewpoint yet to be explicitly entertained by any court, the requirement of re-testing takes on an altogether different cast. Certainly, to a test-taker wrongly accused, agreeing to re-test may feel like a validation of the testing agency's wrongful accusations. Though the test-taker who stands on principal and refuses to re-test might be displaying

343. Cortale, 251 A.D.2d at 530.
an overabundance of pride or even stubbornness, her reaction is nonetheless understandable.

Moreover, even if an innocent test-taker has no principled objection to re-testing, she or he may have legitimate practical concerns about submitting to a reexamination. For example, the test-taker may not have the same confidence in the reliability and consistency of the testing process as the agency has in itself. After all, the innocent test-taker who has been wrongly accused by an authority that evinces absolute certainty in its investigatory methods and proofs may justifiably doubt that company's promises of a fair and consistent retest as well.

Indeed, recent events suggest that there may be cause to mistrust the accuracy of current scoring processes in general. As this article goes to press, the College Board has had to acknowledge that widespread and extremely large computer-scanning errors were made in scoring the October 2005 SAT exam, which were only discovered after two students requested that their tests be re-scored by hand.344 To date, the Board, which has several times had to go back on earlier reports that understated the problem, admits 4,400 scores from that exam were marked too low, with the largest error coming in at 450 points.345 The Board reported that 1,600 exams from the October test that had been "separated for special processing because of security and other questions" had not yet been rescored but would be.346

This disturbing SAT debacle comes on the heels of an $11.1 million settlement announced in March, 2006 between ETS and thousands of teachers who were incorrectly graded on their Praxis teacher-licensing exams, many of whom claimed serious damage to their teaching careers as a result.347 Altogether, roughly 27,000 teachers were incorrectly graded between January 2003 and April 2004, including 4,100 who received erroneous failing scores.348

But regardless of whether the prospective re-tester has doubts about the integrity of the exam, she may simply doubt her own ability to come within the required range necessary to prove the questioned score. For the test-taker who suffers from test anxiety, low self-esteem, or serious depression (no doubt likely to be worsened by an accusation of cheating),

345. Id.
346. Id.
348. Id.
this is a legitimate concern. Requiring this sort of examinee to re-test under great pressure at a comparable level might be like asking a shortstop who just hit a homerun to duplicate the feat in his next at-bat.

Although, as discussed earlier in this article, the Scott court suggested that the test-taker’s interests and the testing agencies’ interests overlapped in “a common concern that [the testing agency’s] scores be reliable,”\(^\text{349}\) this is a specious argument. The truth is that the examinee who suffers from depression or test anxiety has no such stake in the objective validity of her test score, let alone the standardized testing process itself. This test-taker already knows that standardized tests results are an unreliable measure of her abilities. Her only interest lies in proving that despite difficulties in testing well, at least once, she did succeed.

\textit{E. Effectiveness of Invalidation in Combating Cheating and Maintaining Reliability}

Because testing agencies generally invalidate only those scores which have been red-flagged for showing large increases over preceding scores, there are reasons to doubt the final key assumption of courts—that invalidating scores is an effective way to protect against cheating and maintain testing reliability.

To begin with, it stands to reason that if a test-taker can raise his score drastically by glancing over at another test-taker’s paper, then lower levels of cheating must be easier to achieve and far more prevalent. Surely, if a cheater can copy enough answers to raise his score by \(x\) number of points, he can cheat with a little less industriousness, copy fewer answers, and raise his score by a significant but less suspicious amount. Similarly, imposters who are capable of scoring drastically higher than the individuals for whom they test can play it smart by testing to a lower level, thereby managing to fly under the agency radar.

Moreover, catching violators by red-flagging large increases in scores does nothing to catch those inveterate cheaters who cheat the first time around. Again, if an examinee can raise his score by a large number of points on his second or third test through copying, he may also be able to do so on his first test and never be detected.

Apparently, testing agencies are not terribly concerned about those who cheat the first time or who escape detection by cheating to a lesser degree. Perhaps this is because such cheating does not produce embarrassingly large score increases across tests—score increases that

\[349\] Scott, 600 A.2d at 504.
testing agencies might have trouble explaining to educators and employers who have reservations about the validity of standardized testing. How else can one justify agency complacency in failing to institute safeguards that could make such drastically successful score increases by cheating more difficult to achieve?\footnote{350}

Further evidence that authentic reliability is not a priority for testing agencies can be found in the failure of those agencies to institute even rudimentary safeguards to insure against unfair and inconsistent testing conditions, a far more prevalent and serious problem than cheating. For example, the fact that ETS does not bother to provide proctors with reliable digital timepieces that would at least improve the chances that all examinees receive the same amount of time in which to take their tests,\footnote{351} alone speaks volumes.\footnote{352} Anecdotal evidence suggests that

\footnote{350. For a sense of what goes on during high school PSAT, SAT, and AP exams, see College Confidential, \textit{ETS Needs to Step up Its Test Security} . . . ,} http://talk.collegeconfidential.com/showthread.php?t=9852 (accessed Mar. 1, 2006), a chat room thread in which an owner of a test preparation tutoring company attempts to engage students in a conversation regarding irregularities and cheating on such standardized high school exams, and expresses dismay at ETS’ failure to send in independent proctors or monitors to ensure proper testing conditions at high schools. Conveying the generally indifferent sentiment a number of students manifested in response to their own reports of, \textit{inter alia}, lazy, incompetent proctors, various forms of blatant cheating, and large timing irregularities, one student remarked: “I’ve never seen blatant (sic) cheating at the SATS besides little things like flipping back, flipping ahead, or starting a minute early . . . .” \textit{Id}.

\footnote{351. See Law School Discussion, \textit{supra} n. 335 (posting recent LSAT examinees’ numerous reports of timing irregularities occurring at 2004 LSAT exams where proctors made incorrect time announcements, ended exam sections too early, or allowed students to look ahead at sections or start sections early). \textit{See also supra} n. 350 (discussing a chat room thread that details students’ numerous anecdotal reports of blatant cheating and testing irregularities, including timing irregularities during PSAT, SAT, and AP exams); Silent Technology, \textit{The Silent Timer\textsuperscript{TM}: Testimonial \\& Stories}, http://silenttimer.com/stories/ (accessed Mar. 1, 2006) (promoting a device that allows one to time one’s exam silently and flashes a light when a test section’s time is running out; and providing testimonials from examinees who were able to improve their scores significantly through use of the device, which has the advantage over regular timers of not beeping and disturbing other test-takers). While Silent Technologies warns purchasers of its test-oriented timer: “Some tests do not allow timers at all, and others do. The best thing to do is contact your testing administration to be sure,” Silent Technologies, \textit{Legality FAQ}, http://silenttimer.com/timer/faq2.php?cat=7 “Can I use THE SILENT TIMER on test day?” (accessed Mar. 1, 2006), its own consumer testimonials make clear that examinees are using such devices regardless of any rules against them. Silent Technology, \textit{The Silent Timer\textsuperscript{TM}: Testimonial \\& Stories, supra} 351. Although rules against timers, if in place to prevent examinees from disturbing others with beeping sounds and causing false stoppages, have a legitimate purpose, it seems clear that if they stop most examinees from using timing devices, while others enjoy the advantages of such devices, inequities obviously will result.

The author’s own experience provides first-hand confirmation of the kind of timing irregularities that apparently occur at LSAT exams. In December of 2000, I took the LSAT exam administered at the University of Pennsylvania. I used a digital watch that accurately reflected the time remaining in any given section of the exam. Upon discovering that the proctor ended the first two sections of the exam close to a minute early, I spoke with her. She assured me she was carefully monitoring the time and that her watch, which she showed me, was accurate. To my amazement, it was a dial watch, with no second hand! Fortunately for me, I took the advice of test-prep books I had
egregious testing conditions may not be so rare, and that less egregious but nonetheless significant lapses are common. 353

Testing agency complacency regarding unfair test-site conditions should cause courts to question whether empowering testing agencies with greater discretion to invalidate large score increases serves the purpose of increasing overall testing validity. In fact, the opposite may be true. If testing agencies are given unbridled freedom to invalidate large score increases they label suspicious, they can protect their image of reliability while evading closer scrutiny of their testing procedures.

By contrast, if testing agencies in appropriate cases were required to prove cheating by a suspected examinee, the deficiencies of their testing apparatus would be exposed for all to see. For example, in those cases where an accused examinee argues that he could not possibly have cheated because he was seated too far away to see the dots on another test-taker’s sheet, 354 or that he thought each person’s test sheet was numbered differently, it would not only be ironic to hear the testing agency argue to the contrary, but also helpful to hear the agency’s explanations for not providing the test-taker a more secure environment in which to test. Such an airing of crucial facts might prompt a discussion of why testing agencies which earn tens of millions of dollars in “profits” over expenses, choose not to spend greater resources on hiring more proctors, providing better training for proctors, seating examinees further apart, instituting more sophisticated identity checks, or developing other means to greatly lessen or even eliminate the potential for copying or impersonation.

---

read by Kaplan and Princeton Review advising that proctors commonly shorted examinees on time and that one should therefore try to finish a minute early. Others present at that exam, however, may have failed to fill in any number of answers in time.

352. To allow proctors to time by their own timepieces, which may not even have second hands, see supra note 351, when LSAT examinees, for example, have paid more than 100 dollars each for the privilege of testing is nothing short of astounding. One can only conclude that every year a good number of law students arc accepted or rejected by their chosen law school based simply upon whether a proctor happened to give their testing group a minute more or a minute less on an LSAT test section. The fact that ETS does not take simple measures to eliminate this problem, which must surely have been called to its attention numerous times, cannot be explained by cost concerns. Reliable digital watches could be purchased in bulk at a cost of only a couple of dollars each. Moreover, they could even be specially manufactured to provide just one countdown function for those exams, like the LSAT, where each section is of equal length, virtually eliminating the chance for timing errors. All the proctor would have to do is push the button. The stopwatch could of course be re-used on the next exam.

353. Supra nn. 335, 351, 352.

354. See Koza, 2001 WL 1191050 at *2 n. 6 (noting that examinee claimed he could not possibly have copied so many problems while “sitting in a five feet radius” from other test-takers).
VII. SUGGESTIONS FOR A MORE EQUITABLE APPROACH TO TESTING INVALIDATION CASES

Given the fact that so many of the assumptions underlying judicial deference to testing agencies are grounded on fallacies, courts should rethink their approach to testing invalidation cases. A new starting point for courts should be to recognize that standardized testing is just one means to predict the abilities of students and employees—indeed a means that has been deemed untenable and unreliable by many respected educational authorities. The standardized testing industry should no longer be accorded the mantle of a quasi-educational apparatus engaged in unquestionably meritorious activities. A fairer assessment would be that testing agencies are self-interested corporate giants providing products and services, at a cost underwritten by testee consumers, to those institutions that subscribe to their methods of determining aptitude.

Courts should also recognize that as candidates for certain institutions, examinees must submit to the monopoly control exercised by testing companies over the principal means of entry to those institutions. Consequently, consumers have no choice but to enter into adhesion contracts written by the testing agencies themselves—contracts that are lacking in the most basic aspects of consideration that would be accorded in arm’s length negotiations.355

Though ruling on equitable matters, courts have taken an overly formalistic approach to invalidation cases. Courts have granted testing corporations the status of public institutions for equitable purposes, yet have emphasized their private nature for the purposes of due process inquiries. They have ruled that even when a testing corporation has utterly refused to consider exonerating evidence in good faith, the examinee must find a remedy under the terms of an adhesion contract written by that very same testing corporation.

When a court finds that a testing agency investigation of a flagged score was not conducted in good faith, it should not hesitate to exercise its equitable powers in ordering that the questioned score be validated and released. The approach taken by the court in Martin v. Educational Testing Service,356 where the court put quotes around the word “contract” when referring to the adhesion agreement entered into by a standardized test-taker, and doubted the enforceability of some of its

355. See Nairn, supra n. 97, at 264 (noting the dearth of conditions in the testing contract “that the consumer can legally compel ETS to honor,” for example, “scoring the tests accurately, getting the information out on time, or protecting confidentiality”).
356. 431 A.2d 868.
provisions,\textsuperscript{357} is entirely appropriate when the testing agency has breached its duty of good faith. Contract scholars, proposing a good faith standard in which willful refusal to consider exculpatory evidence by a party exercising discretion under the contract "indicates that he does not want to know the truth and amounts to dishonesty,"\textsuperscript{358} have concluded:

Defendant should not escape [the good faith/willful refusal] standard by asserting that plaintiff's evidence would not have induced him to make a contrary determination. Defendant's post hoc assertion is of doubtful veracity and unverifiable. Therefore, the determination of whether the relevant criteria are satisfied or the relevant factual conditions exist should be made not by the defendant but by the court. In the SAT case, \textit{the determination of whether plaintiff cheated would be made by the court based on all the evidence, including that which plaintiff was denied the opportunity to present.}\textsuperscript{359}

\section*{VIII. Conclusion}

Justice goes wanting when a court reviews the established facts of a given testing invalidation case, acknowledges that those facts amply support the innocence of an accused examinee, but nonetheless stops short of granting the examinee a meaningful remedy.\textsuperscript{360} Ruling as they are in equity, courts should rethink their policy of judicial deference to testing agencies, choose substance over form, and in appropriate cases nullify those terms of testing adhesion contracts which produce inequitable results. Only then will courts give meaning to the principle that, "Where a contract is breached . . . and the injured party is entitled to specific performance, the remedy must be a real one, not an exercise in futility."\textsuperscript{361}

\begin{itemize}
\item \textsuperscript{357} \textit{Id.} at 874.
\item \textsuperscript{358} Diamond \& Foss, \textit{supra} n. 180, at 616 (citation omitted).
\item \textsuperscript{359} \textit{Id.} at 617 (emphasis added). The authors made this statement after reviewing the facts of \textit{Dalton}.
\item \textsuperscript{360} Unfortunately, this is exactly what the New York Court of Appeals did in \textit{Dalton}. In that watershed case, the Court of Appeals did not disturb the trial court's finding that ETS demonstrated an egregious and utter disregard for the relevant and exculpatory proofs provided by the suspected examinee. \textit{Dalton}, 663 N.E.2d at 291. Yet the court, ruling in equity, nullified the relief granted by the trial court and consigned the plaintiff to an empty remedy dictated by the defendant under the terms of an adhesion contract. \textit{See id.} at 294-95 (seemingly whistling in the wind, the court stated, "We cannot agree with Dalton's assumption that ETS will merely rubber-stamp its prior determination without good-faith attention to his documentation and that reconsideration by ETS will be an empty exercise"). Vincent F. Nicolosi, Esq., plaintiff's counsel in \textit{Dalton}, confirmed that plaintiff's SAT score was never validated by ETS. Telephone Interview with Vincent F. Nicolosi. Counsel for Plaintiff in \textit{Dalton} (Mar. 26, 2004).
\item \textsuperscript{361} \textit{Dalton}, 663 N.E.2d at 294.
\end{itemize}