Case No. 03-1479

UNITED STATES COURT OF APPEALS
FOR THE SEVENTH CIRCUIT

CHICAGO SCHOOL REFORM BOARD OF TRUSTEES, an Illinois municipal corporation,	)	
Plaintiff-Appellee,	)	Honorable Magistrate Judge
v.	)	Edward A. Bobrick, Presiding
SUBSTANCE, INC., an Illinois corporation, and GEORGE N. SCHMIDT,	)	
Defendants-Appellants.	)	

Appeal From The United States District Court
For The Northern District of Illinois,
Eastern Division
Case No. 99 C 440
The Honorable Charles R. Norgle

BRIEF OF AMICI CURIAE
ON BEHALF OF DEFENDANTS-APPELLANTS
SUBSTANCE, INC. AND GEORGE SCHMIDT

On the Brief:
Gerald Bracey
1797 Duffield Lane
Alexandria, VA 22307

PURE
407 S. Dearborn, #515
Chicago, IL 60605

FAIRTEST
342 Broadway
Cambridge, MA 02139

Attorney for Amici Curiae:
Sarah Vanderwicken
Chicago Lawyers' Committee for Civil Rights Under Law, Inc.
100 North LaSalle Street, Suite 600
Chicago IL 60602

Dated: June 25, 2003

CIRCUIT RULE 26.1 DISCLOSURE STATEMENT

Appellate Court No: 03-1479

Short Caption: Chicago School Reform Bd. of Trustees v. Substance, Inc., *et al.*

(1) The full name of every party that the attorney represents in the case (if the party is a corporation, you must provide the corporate disclosure information required by Fed.R.App.P. 26.1 by completing item #3):

Gerald Bracey, Parents United for Responsible Education (PURE), National Center for Fair and Open Testing (FAIRTEST) and Chicago Lawyers' Committee for Civil Rights Under Law

(2) The names of all law firms whose partners or associates have appeared for the party in the case (including proceedings in the district court or before an administrative agency) or are expected to appear for the party in this court:

Chicago Lawyers' Committee for Civil Rights Under Law

(3) If the party or amicus is a corporation:

i) Identify all its parent corporations, if any; and

Not Applicable

_____ii) list any publicly held company that owns 10% or more of the party's or amicus' stock:

Not Applicable

DATED: June 25, 2003

CHICAGO LAWYERS' COMMITTEE FOR
CIVIL RIGHTS UNDER LAW, INC.

By: _____
Sarah Vanderwicken
Counsel of Record for Above

Sarah Vanderwicken
Chicago Lawyers' Committee
For Civil Rights Under Law, Inc.
100 North LaSalle Street, Suite 600
Chicago, IL 60602
(312) 630-9744

TABLE OF CONTENTS

Page:

CIRCUIT RULE 26.1 DISCLOSURE STATEMENT i

TABLE OF CONTENTS ii

TABLE OF AUTHORITIES iv

STATEMENT OF INTERESTS. 1

SUMMARY OF ARGUMENT. 3

STATEMENT OF FACTS 4

ARGUMENT 7

I. THE TRIAL COURT ERRED IN REFUSING TO CONSIDER FACTS SUPPORTING DEFENDANTS' FIRST AMENDMENT AND FAIR USE DEFENSES 7

A. THE FIRST AMENDMENT PROTECTED THE SUBSTANCE ARTICLE 8

B. DEFENDANTS MADE "FAIR USE" OF THE CASE 9

II. PUBLICATION OF CASE EXAM EXCERPTS IN SUBSTANCE SERVED THE PUBLIC INTEREST AND RAISED MATTERS OF GREAT PUBLIC CONCERN 10

A. TESTING FOR ACCOUNTABILITY AND HIGH STAKES PURPOSES HAS BECOME THE CENTRAL STRATEGY FOR IMPROVING SCHOOLS. 10

B. CHICAGO HAS BEEN AT THE FOREFRONT OF USING STANDARDIZED TESTS FOR HIGH-STAKES PURPOSES–BACKGROUND TO CASE DEVELOPMENT 12

C. HIGH STAKES TESTING POLICIES HAVE HAD DEVASTATING CONSEQUENCES FOR THOUSANDS OF CHILDREN IN CHICAGO AND ELSEWHERE 14

D. TO ASSURE QUALITY, TESTS SHOULD BE OPEN TO PUBLIC SCRUTINY 16

III. THE CASE EXAMINATIONS WERE OF LITTLE UTILITY TO THE BOARD AND COULD NOT HAVE BEEN USED AGAIN EVEN ABSENT THE *SUBSTANCE* PUBLICATION 19

A. THE CRESST STUDY SHOWS THAT MOST OF THE MULTIPLE-CHOICE ITEMS WERE FLAWED 20

B. THE CASE WAS IMPROPERLY DESIGNED 21

1. THE CASE WAS BIASED AND DISCRIMINATORY 22

2. THE CASE WAS NOT DESIGNED FOR USE IN HIGH STAKES DECISIONS 23

3. THE TESTS WERE POORLY RELATED TO COURSE CONTENT . . 25

IV. THE CONFIDENTIALITY OF THE CASE EXAMS DID NOT DESTROY DEFENDANTS' FIRST AMENDMENT RIGHTS TO PUBLICATION 27

CONCLUSION 30

CERTIFICATE OF COMPLIANCE WITH F.R.A.P. RULE 32(a)(7). 31

CIRCUIT RULE 31(e) CERTIFICATION 32

CERTIFICATE OF SERVICE 33

TABLE OF AUTHORITIES

Page:

CASES:

Campbell v. Acuff-Rose, 510 U.S. 569 (1994). 9

Conner v. Reinhard, 847 F.2d 384 (7th Cir. 1988). 28

Council 31, AFSCME v. Ward, 978 F.2d 373 (7th Cir. 1992). 22, 23

Donahue v. Staunton, 471 F.2d 475 (7th Cir. 1972). 28

Hanneman v. Breier, 528 F.2d 750 (7thCir. 1976) 27, 28

Pickering v. Board of Education, 391 U.S. 563 (1968). 7-8, 27

Ty, Inc. v. Publications International Ltd., 242 F.3d 512 (7th Cir. 2002). . . 9, 29

Williams v. Jader Fuel Co., Inc., 944 F.2d 1388 (7th Cir. 1991). 7

Wainscott v. Henry, 315 F.3d 844 (7th Cir. 2003). 8, 9, 19, 27, 29

STATUTES:

17 U.S.C.A. §107. . . . . .8, 9

F.R.C.P 12(c) 4

No Child Left Behind Act, Pub.L. No. 107-110, Title I, §§1111 & 1116, 20 U.S.C.A. §§6311 & 6316 11

U.S. CONST. amend. I 4

OTHER AUTHORITIES:

American Educational Research Association, American Psychological Association, National Council on Measurement in Education, Standards for Educational and Psychological Testing, 54, 146-147 (1985) 13

Christopher Jencks, Inequality, A Reassessment of the Effect of Family and

Schooling in America, New York, Basic Books (1972) 11

Consortium on Chicago School Research,
Update:Ending Social Promotion 10-12 16

FairTest Examiner 7 (Summer 2002) 18

FairTest Examiner 21 (Winter-Spring 2003). 18

FairTest Examiner 18 (Winter-Spring 2003) 18

FairTest Examiner 4 (Summer 2002) 17

G.F. Madaus, M.M. West, M.C. Harmon, R.G. Lomax, K.A. Viator, The Influence of Testing on Teaching Math and Science in Grades 4-12 (Center for the Study of Testing, Evaluation and Educational Policy. 1992) 15

George F. Madaus, A Brief History of Attempts to Monitor Testing. National Board on Testing and Educational Policy v. 2, n.2., (Feb. 2001) 16,17

H.D. Hoover, et al., Interpretive Guide for
School Administrators, 15-18 (1996) 12

J.P. Heubert and Robert M. Hauser, eds., Board on Testing and Assessment of the National Research Council of the National Academy of Sciences, High Stakes: Testing for Tracking, Promotion and Graduation (Nat. Academy Press, 1998) 13, 15

Kathleen Rhoades and George Madaus, Errors in Standardized Tests: A Systemic Problem (2003) 17, 19

Peter Sacks, Standardized Minds (1999) 10,11

Lorrie A. Shepard and Mary Lee Smith, Flunking Grades: Research and Policies in Retention (1989) 15

George F. Madaus, A Technological and Historical Consideration of Equity Issues Associated with Proposals to Change the Nation's Testing Policy, v.64 n.1 Harvard Educational Review, 75-76 (Spring 1994) 10

Office of Civil Rights of the US Department of Education, The Use of Tests as Part of High-Stakes Decision-Making for Students: A Resource guide for

Educators and Policy-Makers, 127 (July 2000) 14

The National Commission on Excellence in Education, A Nation At Risk: The Imperative for Education Reform 5 (1983) 11

Tonya R. Moon, Carolyn M. Callahan, Carol A. Tomlinson, "Effects of State Testing Programs on Elementary Schools with High Concentrations of Student Poverty–Good News or Bad News" *Current Issues in Education*, Vol. 6, No. 8, (April 28 , 20003) 15

STATEMENT OF INTEREST

Amici curiae are persons and organizations that have long cautioned against the over-use of standardized tests in Illinois schools and elsewhere. [1]

1. Parents United for Responsible Education (PURE) is a non-profit organization founded in 1987 and dedicated to improving Chicago's public schools. PURE provides information, support, training, and advocacy to both Chicago public school parents and to local school councils. In 1999, PURE filed a complaint with the U.S. Department of Education concerning the Chicago Board of Education's use of standardized tests to restrict promotion opportunities, resulting in a change to the Board's promotion policy.

2. The Chicago Lawyers' Committee for Civil Rights, Inc. (CLC) is an Illinois non-profit organization started in 1969 by a group of local attorneys to provide high quality legal counsel to those clients whose civil rights cases and projects would benefit the community at large. Since its inception, CLC has worked to increase resources and opportunities in public education for minority and poor children. For the past several years, CLC has focused on the number of students "pushed out" of school because of the Chicago Board of Education's reliance on standardized tests for promotional and other "high

[1] Counsel for Plaintiff-Appellee Chicago Board of Education and Defendants-Appellants have consented to the filing of this brief.

stakes" decisions.

3. The National Center for Fair & Open Testing (FairTest) is an national advocacy organization working to end the abuses, misuses and flaws of standardized testing and ensure that evaluation of students and workers is fair, open, and educationally sound. Fair Test places special emphasis on eliminating the racial, class, gender, and cultural barriers to equal opportunity posed by standardized tests, and preventing their damage to the quality of education.

4. Gerald Bracey is an Associate Professor in the Graduate School of Education at George Mason University who holds a Ph.D. in psychology from Stanford University. Professor Bracey has extensive experience with standardized educational tests, as he served as a Research Psychologist in the Early Childhood Education Research Group at the Educational Testing Service, was the Associate Director of the Institute for Child Study at Indiana University, was the Director of Research, Evaluation and Testing for the Virginia Department of Education, held a similar position with the Cherry Creek (Colorado) School District, and has written numerous books on the issues of educational reform, educational assessment, educational standards, and educational outcomes.

SUMMARY OF ARGUMENT

As alleged in defendants' amended counterclaims and third-party complaint, defendant George Schmidt and *Substance* newspaper have performed in the best tradition of rigorous investigative journalism for almost three decades, confronting bureaucrats both within the Chicago Board of Education ("Board" or "CPS") and within the teachers' union, exposing fraud and mismanagement, and informing teachers and others interested in Chicago schools about poor educational decisions in the administration of Chicago schools. The latter included the development, starting in the mid-1990's, of "accountability" policies that punitively targeted schools and students whose performance on the tests primarily reflected the effects of poverty, as well as the racial and ethnic segregation prevalent in Chicago. When the CPS administrators installed in 1995 began shortly thereafter to use standardized tests for promotional and other "high stakes" decisions, *Substance* provided ammunition to the many teachers, parents and other stakeholders distressed by such policies, in the form of facts, analysis and editorial commentary.

The Chicago Academic Standards Examination ("CASE") – a standardized set of examinations that were developed by in the late 1990's – is an excellent example of why such tests should not be used for high-stakes purposes. As explained below, the CASE tests were shoddily constructed and

their planned use to determine whether students had achieved minimum competency in various high school courses, was completely unjustified, amounting to educational malpractice.

Immediately after *Substance* published portions of the CASE exams in order to expose the absurdity of reliance on them, the Board sought a temporary retraining order based on alleged copyright infringement and trade secrets and claimed over $1 million in damages. Although Schmidt and *Substance* asserted First Amendment and "fair use" affirmative defenses, and later sought to bring counterclaims and a third-party complaint, they were never allowed to offer into evidence the facts supporting those claims and defenses; the trial court *sua sponte* and erroneously entered judgment on the pleadings under Rule 12(c), refused to reconsider these rulings, and rejected defendants' efforts to develop their legal arguments, or support them with evidence. R.ApxE.. Had the court taken evidence, *amici curiae* submit that defendants would have proved that their publication of portions of the CASE exam constituted "fair use" under the Copyright Act and that their activities were protected by the First Amendment.

STATEMENT OF FACTS

Substance has been continuously published in Chicago since 1975. R.97 ¶12. It is a bona fide newspaper, qualifying for the Periodicals Rate from the United States Postal Service. Id. ¶19. It was originally founded as a newsletter

for substitute teachers, but grew over time to cover all aspects of education in Chicago, including curricular programs, Board cover-ups of criminal activity by a Board official, union malfeasance, and facilities mismanagement. Id., ¶¶ 21, 23m, 23s, 26b-f. Its approach has been "no holds-barred" and in the best tradition of rigorous investigative journalism. Id. ¶20. It has often provided a voice for people, such as substitute teachers who were not represented by the teacher's union and whose jobs were vulnerable, whose views would have otherwise been left unexpressed. The newspaper has frequently received information, tips and documents regarding matters of public concern in relation to CPS, some of which have been sent anonymously to the newspaper. Id. ¶29. George Schmidt has been a reporter for *Substance* since its inception, and has been its Editor since June 1996. Id. ¶16-17. Until his dismissal in August , 2000, he had also been a CPS teacher for more than 26 years, with high evaluations and ratings. Id. ¶14.

Beginning in 1996, *Substance* began reporting on the mis-use of standardized testing in Chicago schools under the leadership of then-CEO Paul Vallas, and the devastating and discriminatory impact of that testing on Chicago's mostly minority children. Id. ¶26g. In January 1999, the newspaper received, anonymously, copies of some of the CASE examinations, the authenticity of which could be verified. Id. ¶30-31. *Substance* subsequently published, in the January/February 1999 issue, several articles harshly

critical of the examinations, and, to illustrate their poor quality, verbatim excerpts. Id. ¶32. Shortly after that, Schmidt was removed from his teaching position and the Board brought this lawsuit against him and *Substance*, claiming that the publication of the exam excerpts had cost the Board $1.4 million. Id. ¶¶33, 35. After further investigation revealed that that claim could not be substantiated, the Board entered a settlement agreement with Schmidt by which, in the event he loses entirely on appeal, he will pay the Board $500. R.102. In addition, based undoubtedly on the poor quality of the exam, the Board has now decided to discontinue the CASE altogether. R.101

ARGUMENT

I THE TRIAL COURT ERRED IN REFUSING TO CONSIDER FACTS SUPPORTING DEFENDANTS' FIRST AMENDMENT AND FAIR USE DEFENSES.

By entering judgment on the pleadings, the district court apparently decided that there were *no facts that could be proven* that would establish that defendants had any right to publish excerpts from the CASE exam. *Williams v. Jader Fuel Co., Inc.,* 944 F.2d 1388, 1400 (7th Cir. 1991). But where, as here, claims and defenses involve a balancing of interests, and where the facts that have been developed since entry of that judgment show that the plaintiff's claims were insubstantial at best, that decision was clearly in error.

Amici curiae submit that, had the trial court taken evidence in this case, that evidence would have first shown, as discussed below, that the *Substance* publication of exam excerpts constituted speech on a matter of great public concern, protected by the First Amendment, and that the Board lost little if anything of value by virtue of the *Substance* publication after the tests had already been given. The First Amendment interests of defendants, and the public, in exposure of the shoddy construction of the tests outbalanced the interests of the Board of Education in using the tests again in the future, especially since the Board knew that it could not use the same tests anyway without drastic changes and improvements. *Pickering v. Board of Education*,

391 U.S. 563 (1968). Second, such evidence could establish that the publication of portions of these examinations was a "fair use" under Section 107 of the Copyright Act, 17 U.S.C. §107.

A. The First Amendment Protected the *Substance* Article

In *Wainscott v. Henry*, 315 F.3d 844, 848 (7th Cir. 2003), this Court reiterated the framework for determining the protections extended by the First Amendment. In that case, a city worker who made critical remarks about the city administration was terminated, a penalty that was reduced to a 45-day suspension without pay, and a directive not to make "statements concerning management issues, decisions, or policies of the administration." *Id.* Once it is established that speech is on a matter of public concern, and here it indisputably was, the court then must "balance the employee's interest in commenting upon such matters and the employer's interest in efficient public services." *Id.*, *citing Pickering v. Board of Education*, 391 U.S. 563, 568 (1968). This multi-faceted balancing test involves the following criteria:

> (1) whether the speech would create problems in maintaining discipline or harmony among co-workers; (2) whether the employment relationship is one in which personal loyalty and confidence are necessary; (3) whether the speech impeded the employee's ability to perform his responsibilities; (4) the time, place and manner of the speech; (5) the context in which the underlying dispute arose; (6)

> whether the matter was one on which debate was vital to informed decision making; and (7) whether the speaker should be regarded as a member of the general public.

Id. at 851 (citations omitted).

Had the trial court considered evidence relevant to defendants' affirmative defense under the First Amendment, it would have been clear that the Board's interest in providing efficient educational services was counterbalanced by Schmidt's interest in commenting on matters of pre-eminent public concern. Indeed, the evidence would show that Schmidt's commentary was calculated to *improve* the quality of those services. The public information, moreover, was vital to informed decision-making.

B. DEFENDANTS MADE "FAIR USE" OF THE CASE

A reproduction of copyrighted material may constitute a "fair use," a judicial doctrine codified under Section 107 of the Copyright Act, 17 U.S.C.A. §107. As explained in *Campbell v. Acuff-Rose*, 510 U.S. 569, 577-78 (1994), there are four factors for consideration:

> (1) the purpose and character of the use, including whether such use is of a commercial nature or is for nonprofit educational purposes; (2) the nature of the copyrighted work; (3) the amount and substantiality of the portion used in relation to the copyrighted work as a whole; and (4) the effect of the use upon the potential market for or value of the copyrighted work.

See also Ty, Inc. v. Publications International Ltd., 242 F.3d 512 (7th Cir. 2002) (fair use involves a complementary usage, which does not function as a

substitute for the original creative work).

On a proper "fair-use" analysis, defendants' publication in *Substance* would be shown to be a fair use.

II. PUBLICATION OF CASE EXAM EXCERPTS IN SUBSTANCE SERVED THE PUBLIC INTEREST AND RAISED MATTERS OF GREAT PUBLIC CONCERN

A. TESTING FOR ACCOUNTABILITY AND HIGH STAKES PURPOSES HAS BECOME THE CENTRAL STRATEGY FOR IMPROVING SCHOOLS.

In the view of *amici*, informed public debate about educational assessment has never been more important. Over the past several decades, standardized testing has moved from being primarily an adjunct tool used by teachers to being the central mechanism for student and school accountability. Peter Sacks, Standardized Minds (1999), see esp. Ch. 4; George F. Madaus, A Technological and Historical Consideration of Equity Issues Associated with Proposals to Change the Nation's Testing Policy, v.64 n.1 Harvard Educational Review, 76-95 (Spring 1994). States have increasingly used tests to determine whether schools are doing an adequate job, to apply awards and sanctions to those schools based on their absolute test scores or changes in those scores, and to determine whether students will graduate or be promoted to the next grade. Sacks, *supra*.

With passage of the most recent re-authorization of the federal Elementary and Secondary Education Act, the No Child Left Behind Act, this trend has been strengthened. Now, all public schools in every state will be

evaluated primarily based on their score gains on state tests. No Child Left Behind Act of 2001, Pub.L. No. 107-110, Title I, §§ 1111 and 1116, 20 U.S.C.A. §§ 6311 & 6316.

The use of tests has moved to a central position in U.S. education, just as education itself has loomed ever-more-important in the public's mind. In 1983, the National Commission on Excellence in Education, appointed by Terrence H. Bell, President Reagan's Secretary of Education, released A Nation at Risk, which declared the US "is being overtaken by competitors throughout the world," that "the educational foundations of our society are presently being eroded by a rising tide of mediocrity that threatens our very future," and that there should be greater use of both state and local tests. The National Commission on Excellence in Education, A Nation At Risk: The Imperative for Education Reform 5 (1983), in Sacks at 75 (1999).

Even before the publication of A Nation At Risk, at least since the years of Lyndon Johnson's "great society," efforts to address social issues of poverty have concentrated ever more intensely on schools. Despite the lack of evidence that schools are a good mechanism for addressing poverty, most national discussions about how to solve poverty and other social ills have put education at the center. Christopher Jencks, Inequality, A Reassessment of the Effect of Family and Schooling in America, New York, Basic Books (1972).

Education reform is thus an issue of great public importance,

perceived by policy makers and the public as powerfully affecting individual and national well-being. As it increasingly has become seen as the primary means to change education, standardized testing has arisen to the status of one of the most important public policy concerns of our time.

B. Chicago Has Been at the Forefront of Using Standardized Tests for High-Stakes Purposes – Background to CASE Development

Beginning in 1996, the CPS administration began a policy of using the standardized tests that were given annually in elementary and middle schools, the Iowa Tests of Basic Skills (ITBS) and the high school level Tests of Achievement and Proficiency (TAP) for determining whether students would be promoted and whether schools would be subject to various levels of intervention. This "high stakes" use of the tests was incongruent with purposes for which those tests were constructed; indeed, the tests were not valid or validated for those purposes but rather were intended to give a "snap shot" overall view of how the tested population compared with a national group of students on whose performance the tests were normed. H.D. Hoover, et al., Interpretive Guide for School Administrators, 15-18 (1996). They did not purport to indicate how well students in any particular district learned material they had been taught and were not based on ("aligned with") the state standards that were intended to drive the curriculum. Hoover, *supra*.

Although proponents of high stakes testing claim it is necessary to prevent "social promotion" and to assure that students will graduate with basic skills, critics have shown that those goals are not justified or assured by the tests that have been used for those purposes. Indeed, the high stakes use of standardized tests to create a barrier to promotion or graduation has been decried by almost every professional testing and educational association in the country, as well as by innumerable academics and other professionals. *See* for example, as well as by innumerable academics and other professionals. *See* for example, American Educational Research Association, American Psychological Association, National Council on Measurement in Education. Standards for Educational and Psychological Testing. 54, 146-147 (1985); Jay P. Heubert & Robert M. Hauser, eds., Board on Testing and Assessment of the National Research Council of the National Academy of Sciences, High Stakes: Testing for Tracking, Promotion and Graduation (Nat Academy Press, 1998).

Although testing and educational professionals, as well as parents and other stakeholders, were alarmed at Chicago's misuse of the tests, the CPS administration continued to use them both for high stakes purposes and to tout purported progress because they served a useful political purpose – they provided readily reported "numbers" that the public could look to for evidence of school performance. *See*, Office of Civil Rights of the US Department of Education, The Use of Tests as Part of High-Stakes Decision-

Making for Students: A Resource Guide for Educators and Policy-Makers, 127 (July, 2000). Efforts to conduct a more complex but accurate examination of student and school performance and the factors affecting it were outgunned by the power of the inaccurate but readily accessible comparison of Chicago student performance to nationally normed "grade-level" scores. *Id.*

It is in this context that the CASE exams were developed. Despite the juggernaut of testing it had helped start, CPS knew that the TAP tests were highly inappropriate for assessing high school performance since they were so little connected to either state standards or classroom instruction in Chicago. The CASE was intended to assess student achievement of the curricular objectives of actual CPS courses. An initial small piloting of the test was conducted in June 1998 and in the fall of 1998 high school teachers were required to go to meetings about the CASE and use pre-CASE preparatory materials. Despite teachers' objections that the materials were of poor quality, and premonitions that the tests would be as well, CPS pressed on with its plans to administer the next, large scale piloting of the tests in January, 1999. It was after those pilot tests were administered in January that *Substance* published the excerpts at issue in this case.

C. High Stakes Testing Policies Have Had Devastating Consequences for Thousands of Children in Chicago and Elsewhere

A major criticism of the high stakes use of tests to create barriers

to promotion and graduation is that such use of the tests has such widespread and serious effects on individual students, with little or no justification for such use with any individual person. Tens of thousands of students have been denied diplomas in states that use tests as barriers to graduation, and many more have been retained in grade. Heubert and Hauser (1998). Grade retention has been extensively studied and found to increase the dropout rate as well as damage motivation and self-esteem without increasing academic achievement. *Id;* Lorrie A. Shepard & Mary Lee Smith, Flunking Grades: Research and Policies in Retention (1989). Research also indicates that minority and low-income students are more likely to be in classes that emphasize test preparation rather than richer teaching, thereby lowering, not enhancing, the quality of education for those students even though test scores may rise . G.F. Madaus, M.M. West, M.C. Harmon, R.G. Lomax, K.A. Viator, The Influence of Testing on Teaching Math and Science in Grades 4-12 (Center for the Study of Testing, Evaluation and Educational Policy. 1992); Tonya R. Moon, Carolyn M. Callahan, Carol A. Tomlinson, "Effects of State Testing Programs on Elementary Schools with High Concentrations of Student Poverty–Good News or Bad News" Vol.6, No.8, Current Issues in Education, (April 28, 2003) available at http://cie.asu.edu/volume6/number8/ accessed by amicus Neill 5/15/2003].

In Chicago, many thousands of students have been retained in

grade based on ITBS scores since the CPS policies were instituted in 1996. However, these children were unlikely to have been helped by retention in grade. According to research by the Consortium on Chicago School Research,

> [f]ew students who were retained in 1997 made adequate progress the next year.... Students who were retained in 1998 had an extra chance to meet the test cutoffs in January.... Even with this extra chance, however, the performance of retained students was only slightly better than in 1997.... Only about 38 percent of eighth graders who were retained or sent to Transition Centers in 1998 were able to raise their scores to the test cutoff by August.

Consortium on Chicago School Research, Update:Ending Social Promotion 10-12. In addition, the Consortium reported, the retention policy led to huge numbers of students dropping out. "By the end of 1999, fully 29 percent of student retained or sent to Transition Centers in 1997 had dropped out." *Id.* Such results are an obvious basis for considerable public concern.

D. TO ASSURE QUALITY, TESTS SHOULD BE OPEN TO PUBLIC SCRUTINY

With testing coming to occupy a central place in school "reform" efforts that are themselves a primary public concern, the question of what is in the tests has become important. Testing is a technology that relies on increasingly complex statistical analyses that are ill-understood by the public or even most educators. George F. Madaus. A Brief History of Attempts to Monitor Testing, Vol.2 No.2 National Board on Testing and Educational Policy (Feb. 2001), accessed by amicus Neill on June 14, 2003, at http://www.bc.edu/research/nbetpp/publications/v2n2.html. As discussed

below, tests may have errors in construction and scoring that, unless caught, can have devastating effects on both students and schools. While standardized tests are often secret, many jurisdictions are coming to understand that, because the tests have such significant effects on students, schools and districts, and because there is no federal or state regulation of the test development industry, tests need to at least be open to public scrutiny once they have been given. *Id.* Thus, several jurisdictions, including Maine, New Hampshire, New York, and Massachusetts have voluntarily chosen to release exams once they are given. In other districts, including Ohio, Texas, Arizona and Florida, lawsuits have led to parents (and a newspaper in the case of Arizona) being able to access the exams. Kathleeen Rhoades and George Madaus, Errors in Standardized Tests: A Systemic Problem (2003).

This access has had several important results. First, individual flawed items can be flagged. For example, students and researchers have done so in Massachusetts, leading to removal of some items from the test and rescoring of the exams, enabling students who had "failed" to now pass. *Id.* at 40-43.

Second, public scrutiny has led to improved quality of test items or ways in which the state has presented items. For example, a mother in New York City found that passages from various pieces of literature had been bowdlerized on the state exam, sometimes to the point of incomprehension.

The issue was reported nationally, FairTest Examiner 4(Summer 2002), which caused living authors to publicly complain, and led the state to promise to halt the practice.

Third, the overall quality of the exams can be reviewed and improved. For example, a series of panels of university professors, published authors of fiction and non-fiction, and university admissions officers, were convened by the Carnegie Foundation to examine the quality of the Regents exams in Language Arts and History. Their reports excoriated the quality of the tests FairTest Examiner, 18 (Winter-Spring 2003).

None of these means of improving the tests would have succeeded had the tests been secret; nor would the students whose diplomas were originally denied have graduated. While it may be more costly to construct new items after errors are noted, it may be more costly to keep errors away from public scrutiny. In Minnesota, a test company error caused 300 students to be wrongfully denied a diploma. FairTest Examiner, 7 (Summer 2002). The unrelenting, many months effort of the father of one such student led to his being allowed to examine the exam. The error he found led to score increases for thousands of students and graduation for many. In the subsequent lawsuit, the trial judge initially denied plaintiffs demand for punitive damages, until he saw test company documents indicating negligent and lax behavior. The company, NCS Pearson will pay up to $12 million in damages. FairTest

Examiner, 21 (Winter-Spring 2003).

Test errors are not new, but with increased testing and increased stakes attached to the tests, errors have been increasing. Rhoades and Madaus 29-30 (2003). While some errors have been brought to light, it is all-too-probable that many more have not been, with unknown but undoubtedly significant harm to many thousands of students. Thus modest increases in testing expenses to provide additional items once tests have been made public may be a relatively small investment in ensuring fairness and in ensuring the possibility of public debate as to the quality of tests and what should be on them.

III. THE CASE EXAMINATIONS WERE OF LITTLE UTILITY TO THE BOARD AND COULD NOT HAVE BEEN USED AGAIN EVEN ABSENT THE SUBSTANCE PUBLICATION

Even the best standardized tests available would constitute an inappropriate basis for evaluating what students know or how schools or districts or states are performing. And certainly, the CASE tests came nowhere close to constituting the best tests available. Indeed, one might consider CASE a worst case. The trial court's failure to take evidence about this was error. In determining whether or not the defendants made "fair use" of the CASE, and whether the First Amendment protected its publication, the court should have considered evidence regarding the validity and legitimacy of the CASE, and not allowed the Board to wall off secrecy about its poor construction. *Wainscott,*

supra, at 851.

Central to a proper analysis of defendants' defenses and claims would be a set of validation studies performed by the National Center for Research on Evaluation, Standards, and Student Testing ("CRESST"), based at the University of California at Los Angeles, with major funding from the MacArthur Foundation. S.Apx.H, C01214-15. (CRESST's *Technical Review of CASE Semester I and Semester 2 Field Test Data 1998-1999* appears at S.Apx.H.) These studies established a serious lack of integrity, both in the overall design, and in the formulation of the individual questions, on the forms of the CASE published in *Substance*.

A. The CRESST Study Shows That Most of the Multiple-Choice Items were Flawed

The CRESST validation studies found the following number of flaws in the individual multiple-choice CASE items that were published in the January, 1999 *Substance:*

26 of 28 Algebra items

24 of 30 English I items

20 of 30 U.S. History items

29 of 30 World Studies items

S.Apx.H, App. C, G. The largest number of flaws were those showing a bias against racial or ethnic subgroups, usually in favor of the majority, or

where the items fell out of the range of difficulty appropriate for this kind of examination. Some items also showed gender bias. The CRESST analysis shows that 5 out of 6, or 84%, of those items had to be reformulated or discarded.[2]

The racial and ethnic bias in items overwhelmingly favored the majority, meaning that the test would have disparate impact, depressing the scores of minority students. It must be noted that these bias calculations are not based on the overall performance of minority and majority students. On the contrary, they are based on the performance for each of the racial and ethnic subgroups, examining students with comparable ability. S.Apx. H, at C01222.

B. THE CASE WAS IMPROPERLY DESIGNED

The fact that CRESST found the overwhelming majority of items to be flawed is critical. To make people aware of the systemic flaws in the tests, isolated questions would be insufficient. If defendants had published only some selected egregious examples, they could be dismissed as anomalies in otherwise sound instruments. In fact the entire examination was flawed as

---

[2]Similarly, the rate of flawed items for the multiple-choice sub-tests that were not published are as follows: Geometry, 27 of 28 items flawed; Physics: 28 of 30 items flawed; Biology: 28 of 30 items flawed; Chemistry, 28 of 30 items flawed; English II, 25 of 30 items flawed; Earth Science: 26 of 30 items flawed; Environmental Science: 27 of 32 items flawed. S.Apx.H, App. C. The unpublished examinations had a 90% error rate. For all the multiple-choice CASE sub-tests administered in January, 1999, the error rate was 88%. *Id.*

well.

1. THE CASE WAS BIASED AND DISCRIMINATORY

The large number of test items with racial, ethnic or gender bias translated into examinations with disparate impact. "Adverse impact may result if several items comprising the test are found to be biased during the item analysis phase." S.Apx. H, at C01223. "Decisions made using this test at this cut point may have detrimental effects" for the disfavored racial/ethnic or gender subgroup. *Id.*

Adverse impact is determined using the "Four-Fifths" Rule, familiar from employment discrimination cases. *Council 31, AFSCME v. Ward*, 978 F.2d 373, 376 (7th Cir. 1992). Under this test, "the ratios between the passing rates of two groups of students are calculated..." and the two pass rates are compared, the lower pass rate divided by the higher. "Instances where the ratio is below .80 indicate a potentially adverse effect on a given subpopulation of students." S.Apx.H at C01222-01223.

To illustrate, assume that, as a hypothetical example, the pass-rates on a certain sub-test were 47% for minorities, but 71% for majorities. Dividing 47 by 71, one gets .66, or approximately 2/3, below the 4/5ths (.80) required by the "Four-Fifths" rule of thumb.

Applying this analysis, at the 50% cut-point, disparate impact was found on the following CASE multiple-choice sub-tests: Algebra and World

Studies: disparate impact by gender and racial/ethnic subgroup; English I and U.S. History: disparate impact by racial/ethnic subgroup. S.Apx.H at 01226-01227: C01253; C01386**;** C01288; C01295.

The CRESST analysis concludes, "it is advisable to explore alternative test compositions and decision point scenarios so as to minimize any disparate effects." *Id.* What these "disparate effects" refer to is the increased likelihood that minority students would have significantly higher failure rates, lessened academic success and reduced life opportunities, due to racially-discriminatory measurement errors built into the test.

2. THE CASE WAS NOT DESIGNED FOR USE IN HIGH STAKES DECISIONS

The large number of flawed items resulted in an extremely high failure rate on the CASE. Over and over, one reads in the CRESST analyses of individual tests that the test is "far too difficult for the intended student population and the stated purpose of the CASE examinations…Individual items do not provide adequate coverage across the desired range, with too many difficult items and much too few easier items comprising each semester's exam. For this subject area test, very few students would qualify." S.Apx.H, C01218 *et seq.*

Substance reported a statement by CEO Paul Vallas that in a prior administration of the examination, there had been a failure rate of 75% on the Algebra sub-test. S.Apx.A at 25. The CRESST analysis reflects that the

January, 1999 administration continued to fail students at an extremely high rate. For the multiple-choice sub-tests published in *Substance*, the pass-rates were as follows.

Algebra: on average, 35% of the items answered correctly (chance is 25%). 4.4% pass rate at a 70% cut-point; 19.4% pass rate at a 50% cut-point. S.Apx.H, at C01225-26.

English I: on average, 47% of items answered correctly. 11.4% pass-rate at a 70% cut-point, 46.5% pass-rate at 50%. *Id.* at C01252-53.

U.S. History: on average, 46% of items answered correctly. 8.4% pass-rate at 70% cut-point; 36% pass-rate at 50% cut-point. *Id.* at C01287-88.

World History: on average, 39% of items answered correctly. Fewer than 4.5% pass rate at 70% cut-point: 25.7% pass rate at 50% cut-point. *Id.* at C01294-95.

In short, the CASE was targeted at such a high level of mastery, at a cut-point of only 70%, none of the sub-tests published in *Substance* had a pass-rate higher than 11.4%; the failure rates ran as high as 95.5%. Even when the cut-point was lowered to 50%, not even half the students passed these sub-tests. At this cut-point, the failure rates were very high, *ranging*

from 53.5% to 80.6%.[3] Had the Board implemented the CASE as a minimum competency test, all of those students would have failed the course, which they would have to repeat.

3. THE TESTS WERE POORLY RELATED TO COURSE CONTENT

The justify the staggering failure rates on the CASE, the evidence would need to support a strong relationship between test and course content. The CRESST validation study shows that such evidence was lacking.

CRESST performed statistical analyses called "exploratory factor analysis" on the CASE sub-tests. This involves the identification of certain "constructs" to be measured by the examination. This might be a certain topic covered by the course in issue. (For example, in the geometry examination, one of the "constructs" might be "measurement".) These constructs, and not some random factor, should account heavily for the differences in scores. The individual test items should relate to these constructs. If the test is properly constructed, the factors should explain the variability of the test scores. The CRESST validation study explains, "it is desirable that the factor structure accounts for a large proportion of the total variance. The judgmental standard used to assess the quality and usefulness of the factor structure is 70% of the total variation in student scores is accounted for by the observed factor

[3] Failure rates for the multiple-choice sub-tests that were not published in *Substance* were comparable. S.Apx.H, App. C.

structure. Instances where the factor structure accounts for 70% or more of the total variation in student scores suggests that there is a strong relationship between the items, which in turn suggests that they measure the same construct." S.Apx.H, at C01220, 01223-24. Thus, where the test betrays a weak factor structure, it means that the test is not properly focused on the course concepts. The whole test must be restructured so that the individual items relate appropriately to the basic factors of the course.

The CASE sub-tests published in *Substance* had weak factor structures. The factor analysis identified 4 to 6 factors in these sub-tests, accounting for only 31-34% of the variance among the multiple-choice items. S.Apx.H, at C01227,C01254,C01261,C01289; C01296. Far from reaching the 70% standard, *i.e.,* establishing that certain definable standards account for the variability of at least 70%, the converse occurred: in each of these four sub-tests, approximately one-third of the variability was *not* explained by the identified constructs.

Expert testimony could establish that the technology of constructing multiple-choice questions has been known for many years; the multiple-choice item was invented in 1915. CRESST had provided a handbook of how to construct multiple-choice questions, and workshops. The CRESST analysis reveals a remarkably low level of integrity in the framing of the examinations.

IV. THE CONFIDENTIALITY OF THE CASE EXAMS DID NOT DESTROY DEFENDANTS' FIRST AMENDMENT RIGHTS TO PUBLICATION

The CRESST validation study, and the considerations of public policy outlined by *amici* above, establish facts and issues demonstrating the urgent public importance of defendants' speech, and its clear entitlement to First Amendment protection. The appalling CASE was not fit for use, and Schmidt acted in the best interests of the public in making this abusive practice known.

The trial court acknowledged that Schmidt's publication had but minimal impact on his performance as a teacher, R.Apx.E at 31. It concluded, however, that under *Pickering*, Schmidt acted improperly in publishing materials that, as a teacher, he knew to be confidential. *Id.*

This is not the law. The confidentiality of otherwise protected speech is not a dispositive factor under *Pickering*. This Court considered this precise issue in *Hanneman v. Breier*, 528 F.2d 750 (7th Cir. 1976), cited favorably in *Wainscott*, *supra.* In *Hanneman*, police officers violated a department confidentiality rule in making public criticisms of police department operations. Finding the speech protected, the court held:

> In order to justify application of the confidentiality rule in this case, however, the defendant chief of police must still meet the requirements of *Pickering v. Board of Education, supra*. He must show a state interest in confidentiality applicable on these facts which outweighs the public and individual interests in the

> particular statements made....

Id. at 754. The court determined that under the facts of that case, "the state interests which support the police department's confidentiality rule [do not] justify its application in this case."

Similarly, in *Donahue v. Staunton*, 471 F.2d 475 (7th Cir. 1972), a chaplain in a mental hospital waged a public campaign, including media coverage, in opposition to what he regarded as untenable conditions in the facility. The court held that his discharge violated his First Amendment rights to free speech:

> We conclude and hold in the case at bar that the interest of society in "uninhibited and robust debate" on matters of public concern, such as mental health care, and plaintiff's individual interest in being free to speak out on matters of concern to him, outweigh those of the State as an employer. This is a proper guarantee of "the public's right to know."

Id. at 481-482. That interest of the public is no less in this action. "[B]oth the public and government officials have a first amendment interest in being informed of the potential misuse of public funds[]" for the improper assessment practices exposed in the *Substance* publication. *Connor v. Reinhard*, 847 F.2d 384, 390 (7th Cir. 1988).

Moreover, unlike the plaintiffs in *Hanneman* and *Donahue*, Schmidt's publication activities were not related to his classroom performance.

His press was off the premises, and his journalism involved no conflict whatever with his work at his school. In keeping with *Wainscott*, *supra*, Schmidt should be treated as a member of the general public insofar as his editorial work is concerned, not as a teacher.

These issues are not, and cannot be, exorcized by the Board's invocation of the Copyright Act. This Court has recently cautioned against "the impairment of freedom of expression that would result from giving a copyright holder control over public criticism of his work..." *Ty, supra,* at 517. The "fair-use" doctrine should be interpreted consistent with First Amendment questions. *Id.* at 523. As the *Ty* court concluded, this is "the kind of fact-laden issue appropriate for summary judgment only in extreme cases, which this case is not", and defendants are "entitled to a trial on [their] defense of fair use." *Id.* at 519, 516.

CONCLUSION

For the foregoing reasons, this Court should reverse the decisions of the District Court striking and denying leave to amend or reinstate defendants' Affirmative Defenses, Counterclaim and Third-party Complaint and the permanent injunction against defendants should be dissolved.

Respectfully submitted,

_____
Sarah Vanderwicken

For amici curiae: Parents United For Responsible Education (PURE)
407 South Dearborn
Chicago IL 60605

National Center for Fair & Open Testing (FAIRTEST)
342 Broadway
Cambridge MA 02139

Gerald Bracey
1797 Duffield Lane
Alexandria, VA 22307

Chicago Lawyers Committee for Civil Rights Under Law
100 North LaSalle Street Room 600
Chicago IL 60602

Sarah Vanderwicken
Chicago Lawyers' Committee for Civil Rights Under Law
100 North LaSalle Street Room 600
Chicago IL 60602
(312) 630-9744

CERTIFICATE OF COMPLIANCE WITH F.R.A.P. RULE 32(a)(7)

The undersigned, counsel of record for Amici Curiae Bracy, PURE, FairTest and Chicago Lawyers' Committee for Civil Rights Under Law, Sarah Vanderwicken, furnishes the following in compliance with F.R.A.P. Rule 32(a)(7):

I hereby certify that this Brief of Amici Curiae conforms to the rules contained in F.R.A.P. Rule 32(a)(7) for a brief produced with a proportionally spaced font. The length of the principal portion of this Brief, from Statement of Interest through the Conclusion, is 6,404 words.

DATED: June 25, 2003

CHICAGO LAWYERS' COMMITTEE FOR
CIVIL RIGHTS UNDER LAW, INC.

By: _____
Sarah Vanderwicken
Attorney for Gerald Bracey,
PURE and FAIRTEST

Sarah Vanderwicken
Chicago Lawyers' Committee
For Civil Rights Under Law, Inc.
100 North LaSalle Street, Suite 600
Chicago, IL 60602
(312) 630-9744

CIRCUIT RULE 31(e) CERTIFICATION

The undersigned, counsel of record for Gerald Bracey, PURE, and FAIRTEST, Amici Curiae, Sarah Vanderwicken, furnishes the following in compliance with Circuit Rule 31(e):

I hereby certify that the computer disk filed concurrently herewith is an accurate copy of Brief of Amici Curiae (also filed concurrently herewith) converted to a PDF format. The disk is virus free, and includes the entire contents of the brief from cover to cover. The contents of the disk as provided comply in full with Circuit Rule 30(e).

DATED: June 25, 2003

CHICAGO LAWYERS' COMMITTEE FOR
CIVIL RIGHTS UNDER LAW, INC.

By: _____
Sarah Vanderwicken
Attorney for Gerald Bracey,
PURE and FAIRTEST

Sarah Vanderwicken
Chicago Lawyers' Committee
For Civil Rights Under Law, Inc.
100 North LaSalle Street, Suite 600
Chicago, IL 60602
(312) 630-9744

CERTIFICATE OF SERVICE

The undersigned, counsel of record for Gerald Bracey, PURE, and FAIRTEST, Sarah Vanderwicken, hereby certifies that on June 26, 2003, two copies of the Brief of Amici Curiae On Behalf of Defendants-Appellants Substance, Inc. and George Schmidt, as well as a digital version containing the brief, were delivered by Chicago Lawyers' Committee for Civil Rights, Inc. to counsel for the Defendant-Appellee, Nancy K. Laureto and Patricia A. Felch via U.S. Mail.

DATED: June 26, 2003

CHICAGO LAWYERS' COMMITTEE FOR
CIVIL RIGHTS UNDER LAW, INC.

By: _____
Sarah Vanderwicken
Attorney for Gerald Bracey,
PURE and FAIRTEST

Sarah Vanderwicken
Chicago Lawyers' Committee
For Civil Rights Under Law, Inc.
100 North LaSalle Street, Suite 600
Chicago, IL 60602
(312) 630-9744