2006

Observations on the Folly of Using Student Evaluations of College Teaching for Faculty Evaluation, Pay, and Retention Decisions and Its Implications for Academic Freedom

Terence Lau
University of Dayton, tlau01@udayton.edu

William Wines
Missouri Western State University

Follow this and additional works at: https://ecommons.udayton.edu/mgt_fac_pub

Part of the Business Administration, Management, and Operations Commons, Entrepreneurial and Small Business Operations Commons, and the Marketing Commons

eCommons Citation

This Article is brought to you for free and open access by the Department of Management and Marketing at eCommons. It has been accepted for inclusion in Management and Marketing Faculty Publications by an authorized administrator of eCommons. For more information, please contact frice1@udayton.edu, mschlangen1@udayton.edu.
OBSERVATIONS ON THE FOLLY OF USING STUDENT EVALUATIONS OF COLLEGE TEACHING FOR FACULTY EVALUATION, PAY, AND RETENTION DECISIONS AND ITS IMPLICATIONS FOR ACADEMIC FREEDOM

WILLIAM ARTHUR WINES*
TERENCE J. LAU**

ABSTRACT

Research on student teaching evaluations is vast. An examination of this research demonstrates wide disagreements but also substantial consensus of authority for the proposition that student evaluations should be used only with extreme care, if at all, in making personnel decisions. A number of reasons cause administrators to use teaching evaluations for personnel decisions. The literature, however, is virtually unanimous in its condemnation of norming student evaluations in order to rank classroom performances. Current cases on academic freedom indicate some retrenchment by the Circuits from broader pronouncements in earlier Supreme Court cases. This paper concludes that the use of non-validated student evaluations alone without any other criteria for teaching effectiveness raises substantial problems in faculty retention and promotion decisions. It also suggests that such an approach in the right case might violate academic freedom and the First Amendment.

Fifty years ago, students at Harvard University and the University of California, Berkeley, were publishing guides rating teachers and courses.

William Arthur Wines is Associate Professor of Business Ethics and Legal Environment in the Department of Finance, The Richard T. Farmer College of Business Administration, Miami University, Oxford, Ohio 45056. Professor Wines holds a B.S.B.A. with distinction from Northwestern University and a J.D. from the University of Michigan and is admitted to practice in Minnesota and Washington State. In 1999, Wines was the John J. Aram Professor of Business Ethics at Gonzaga University, Spokane, Washington.

** Terence J. Lau is Assistant Professor in the Management and Marketing Department at the University of Dayton. Prior to joining the University of Dayton, he served as Director for ASEAN Governmental Affairs for Ford ASEAN Operations in Bangkok, Thailand. Professor Lau's academic awards include recognition of his article Triggering Parent Company Liability Under United States Sanctions Regimes: The Troubling Implications of Prohibiting Approval and Facilitation as the best article presented at the Academy of Legal Studies annual meeting in August 2004 at Ottawa, Ontario, Canada.
Irreverent and funny, they featured pungent comments: . . . ‘Dr. Y communicated very closely with the blackboard, but I couldn’t tell you what he looks like, as he never faced the class.’ Unfortunately, what originated as a light-hearted dope sheet for the use of students has, at the hands of university and college administrators, turned into an instrument of unwarranted and unjust termination for large numbers of junior faculty . . . .

INTRODUCTION

Employee evaluations are widely used by various organizations for a variety of purposes. They can be used to make hiring and termination decisions, and are often the basis for making compensation decisions. The evaluation systems used by corporations vary widely in scope and application, and the media have widely reported on some for the controversy they generate. Most of these evaluation systems, however, have at their core a system wherein an identified supervisor (and, in some systems, a co-worker and/or subordinates) provides job performance information feedback to the employee, and significant employment actions (such as hiring, firing, promoting, pay increases, incentive compensation) may then produce results based on the content of the evaluations.

3. See id.
4. Perhaps the most criticized of these systems, and the one most litigated, is the system known commonly as “forced ranking,” wherein employees are graded and subsequently ranked amongst each other according to a numerical formula. A famous example is the system adopted by General Electric, which forced managers to rate the bottom ten percent of employees, the middle seventy percent, and then the top twenty percent. Katherine Reynolds Lewis, Companies Rank Their Employees; Proponents Say it’s a Good Way to Identify the Best and Worst Workers, THE POST-STANDARD, Apr. 24, 2005, at E1. As many as a quarter of U.S. companies use some form of forced ranking to evaluate employees, but many, including General Electric, have backed away from a numerical quota. Id. These systems are controversial mainly because they allegedly affect older workers disproportionately. See Marcia Coyle, Bell-Curve Rankings Reviled by Older Workers, Plaintiffs, RECORDER, Aug. 19, 2004, at 3. See also Meredith L. Myers, Grades are No Longer Just for Students: Forced Ranking, Discrimination, and the Quest to Attain a More Competent Workforce, 33 SETON HALL L. REV. 681 (2003) (tracing litigation surrounding forced ranking employee performance appraisal systems); Deborah L. Rhode, Occupational Inequality, 1988 DUKE L.J. 1207, 1228 (1988) (tracing history of job performance appraisals).
5. See Gray & Bergman, supra note 1.
Even the use of job performance evaluations in the commercial sector is not without limitations. Arbitrary and capricious tests that discriminate either directly or indirectly (statistically) against members of protected groups are illegal and generate avenues for relief under civil rights laws. Once a member of a protected class (race, color, religion, national origin, sex, or age) raises a prima facie case against such a performance standard, educational requirement, or job-qualification test, the employer has the burden of showing the screening or evaluating device/standard/test is job-related and, hence, valid. Validity tests are usually either based upon criterion validity, content validity, or construct validity.

At virtually all American universities, however, faculty members are evaluated very differently. Most faculty members are evaluated on their scholarship and service contributions by their peers. When it comes to teaching, however, faculty members are usually evaluated anonymously by their students. Students are asked to fill in machine-readable forms (commonly called “Scantron,” an apparent reference to the manufacturer) with their responses to survey questions regarding the course and the instructor. Research indicates that approximately eighty percent of the questions contained on student evaluation forms either are unclear, subjective, or ambiguous or are
unrelated to classroom teaching performance. Yet, the administrators’ push to use these unvalidated and mainly irrelevant results to determine career paths in higher education continues unabated.

One observer theorized about the attraction of student evaluations for administrators in the following terms:

For administrators, the attractiveness of student evaluations of faculty is that they provide an easy, seemingly objective assessment of teaching that does not require justification. The ease of student evaluations comes in reducing the complexities of teaching performance to a series of numbers, particularly when commercial forms are used. The most common type of commercial student evaluation form utilizes a Likert-type scale for students to rate faculty related to a series of statements about the course and instruction. Each point on the scale is assigned a numerical value which allows the computation of composite scores for individual items, groups of items, or all of the items. Finally, the student ratings are often normed nationally and locally in spite of the near universal recommendations in the literature against norming of student ratings.

For example, Miami University, in Miami, Ohio, has a standardized “School of Business Administration Instructor/Course Evaluation” form, reportedly designed by a faculty committee several years ago, containing twenty-seven questions. There is no evidence that the instrument has been either piloted or validated. It has questions

14. T.L. Simmons, *Student Evaluation of Teachers: Professional Practice or Punitive Policy?* 1 SHIKEN: JALT TESTING & EVALUATION SIG NEWSLETTER (Japanese Ass’n for Language Teaching), Oct. 1996, at 12-16, available at http://www.jalt.org/test/sim_1.html (citing findings contained in H.T. Tagomori, A Content Analysis of Instruments Used for Student Evaluation of Faculty in Schools of Education at Universities and Colleges Accredited by the National Council for Accreditation of Teacher Education, (unpublished Ed. Doctorate Dissertation, University of San Francisco, 1993) (on file with Gleeson Library, University of San Francisco)). Tagomori analyzed the content of 4,028 evaluation items contained in the 200 evaluation instruments he studied. *Id.* at 1. His analysis revealed 54.6% of the items were ambiguous, unclear and/or subjective. *Id.* at 11-14. Another 24.5% of the items did not correlate with classroom teaching performance. *Id.* at 16. His study concluded that student evaluation instruments, used in present form, are unreliable. *Id.* at 30-32.


16. Miami University is used merely because the authors are familiar with its practices and not because its practices are any better or worse than other schools. In fact, every college and university at which the authors have taught has used standardized student evaluations of instruction.

17. See MIAMI UNIVERSITY SCHOOL OF BUSINESS ADMINISTRATION, INSTRUCTOR/COURSE EVALUATION FORM (on file with authors).

such as: “4. Relative to other courses, I learned more from this class.”; “16. This course contains interesting subject matter.”; “20. The instructor was excellent (independent of how you feel about the course).”

At the end of each semester, these forms are distributed to the business students at Miami University, collected by the departments and then processed to yield a statistical summary of the results by course and by instructor. The results, calculated to the second decimal place, are then compared to: (a) the mean for all sections of the course taught by that instructor; (b) the mean for all courses taught by that instructor; (c) the mean for all sections of that course taught within the department; (d) the mean for all courses taught within the department; and (e) the mean for all courses taught in the School of Business Administration. In spite of the wealth of data provided, however, other missing key pieces of data mean that a complete picture of instructor effectiveness can be difficult to discern.

In this attempt to reduce to one-hundredths of a point the essentials of a classroom's dynamics, Miami University has embraced our culture's love affair with quantification. What started as a quest for rigor and scientific knowledge has devolved into a convenient style of numerically ranking unobserved teaching. As is typical at many institutions, the results of these evaluations are used for key personnel decisions. For example, at Miami University, faculty are required to undergo student evaluations and submit the results to their departments, and it is suggested that these results be included in the “dossiers” junior professors prepare in application for tenure and promotion.

(Defining "pilot survey" as "a survey, usually on a small scale, carried out prior to the main survey, primarily to gain information to improve the efficiency of the main survey").

19. MIAMI UNIVERSITY SCHOOL OF BUSINESS ADMINISTRATION, supra note 17.
20. Id.
21. For example sample sizes for the department or the division are not provided; also the standard deviations are not given for the department or the division. Without those measures, it is not possible to determine whether the differences between professors or the differences between a professor's scores and the means for the department or division are statistically meaningful. See, e.g., FREDERICK A. EKELAND, THE STATISTICAL METHOD IN BUSINESS, 8-9, 451-55 (1962); EDWIN MANSFIELD, STATISTICS FOR BUSINESS AND ECONOMICS: METHODS AND APPLICATIONS 388-90 (1983).
In Section I of this article, we review the practice of using student teaching evaluations as a basis for faculty employment decisions in higher education. We review existing research in this area across multiple academic disciplines, and present a summary of research findings that suggest that student evaluations can be fraught with inaccuracy and open to abuse. In Section II, we turn our focus to academic freedom and its role in the modern university. We analyze the Constitution’s mandate for freedom of speech as well as the implications for violating academic freedom on the First Amendment. In Section III, we argue that given the biases that are often evident in student teaching evaluations, over-reliance on these evaluations by employers may run afoul of civil rights legislation and important constitutional safeguards.

A review of the literature on the subject of student teaching evaluations shows that there are serious problems with an uncritical use of them for faculty pay, promotion, and retention decisions. More ominously, the continued use of teaching evaluations as a basis of promotion and tenure will inevitably cause faculty members to tailor their speech to obtain promotion and tenure (a conclusion supported by ample research), and that in effect, faculty are coerced into making speech with which they do not agree. We conclude by offering our thoughts on the future of academic freedom in the courts and legislatures, and our suggestions for future research into this topic.

I. USING STUDENT TEACHING EVALUATIONS IN HIGHER EDUCATION

In this section, the article reviews how student teaching evaluations are used in higher education to make personnel decisions. We start by reviewing the existing research into student evaluations of teaching, the biases inherent in anonymous student evaluations, and some of the misuses to which results of student evaluations of teaching are put.

A. Review of the Literature on Student Evaluations

1. Interdisciplinary Nature of Scholarly Research

A search on Google of the terms “Student Evaluation of Teaching at Colleges and Universities,” results in millions results generated in approximately one second of computer time. One article recently declared that “over 2000 studies have appraised...
student evaluations of college teachers.\textsuperscript{27} There is controversy surrounding student evaluations because of the manner in which evaluations are used.\textsuperscript{28} No one seems to doubt that student evaluations provide some measure of student satisfaction. Controversy arises when administrators seek to use student evaluations as a measure of instructional effectiveness.\textsuperscript{29} One author stated that a simple review of one database in 2002 generated “thousands of articles concerning various aspects of SETs [student evaluations of teaching]”\textsuperscript{30} and cited another article for the proposition that more than 170 studies were directed at the question of whether SETs represented valid measures of student learning.\textsuperscript{31} Consequently, this article makes no pretense of having made an exhaustive review of the literature. We have, however, made a strenuous effort to review the better and more recent studies available to us.

2. Literature Can Be Grouped into Following Areas:

a. Evaluations Used for Improper Purposes

In one case of which we are aware, a finance professor, denoted as Professor F, was denied tenure from the university. On an appeal, Professor F wrote:

In my conversation with [the Provost]. . . I learned that much of the information in my tenure dossier was discounted by the University P&T Committee because it was not documented by my department. Specifically, I was told that the multiple measures of my teaching effectiveness in my tenure dossier were not given weight in my tenure decision because they had not been specifically discussed by my department. These measures were therefore assumed to be of little significance. [The Provost] stated that this left only my teaching evaluation scores to be given weight by University P&T as an indication of my teaching effectiveness.\textsuperscript{32}

\textsuperscript{27} Alan Safer, Lesley S. J. Farmer, Angelo Segalla & Ashraf F. Elhoubi, Does the Distance from the Teacher Influence Student Evaluations? 28 EDUC. RES. Q. 28 (Mar. 2005).

\textsuperscript{28} Adams, supra note 9, at 10.

\textsuperscript{29} Id. (citing Phillip C. Abrami, Sylvia d'Apollonia & Peter A. Cohen, Validity of Student Ratings of Instruction: What We Know and What We Do Not, 82 J. EDUC. PSYCHOL. 219 (1990)).

\textsuperscript{30} Valen E. Johnson, Teacher Course Evaluations and Student Grades: An Academic Tango, 15 CHANCE 9-16 (2002).

\textsuperscript{31} Id. (citing Anthony Greenwald & Gerald Gillmore, Grading Leniency is a Removable Contaminant of Student Ratings, 52 AM. PSYCHOL. 1209 (1997)).

\textsuperscript{32} Letter from Professor F to Chair of Faculty Committee on Rights and Responsibilities, [State University] (Mar. 21, 2005) (on file with authors). The name of
Professor F stated that evidence showed teaching evaluation scores are gender biased against female faculty members, particularly in traditional male-dominated areas such as finance.\footnote{33} This gender bias was specifically recognized and noted by her department in its letter recommending that she be granted tenure.\footnote{34} Further, the gender-bias in student teaching evaluation scores has been well-documented in the literature.\footnote{35} Finally, the grievant noted the literature has numerous calls for “multiple measures” of teaching effectiveness because so many studies have found negative biases toward women professors — especially in fields traditionally dominated by men.\footnote{36} Thus, it can be considered a misuse of student evaluations to use them as an exclusive measure of teaching effectiveness for women teaching in traditional male areas — above and beyond the general argument that student evaluations are not a reliable measure of teaching effectiveness.

\subsection*{b. Student Evaluations Reveal Bias Against Certain Groups}

\subsubsection*{i. Double Standard}

Females in traditionally male-dominated areas, such as engineering and business, are held to higher standards than males.\footnote{37} Some researchers suggest this double standard, under which females are expected to be competent in their work but are also either rewarded for conforming to gender stereotypes or punished for not being warmer and more nurturing than their male counterparts, may be a result of the university involved having been redacted.

\begin{itemize}
  \item \textit{Professor F stated that evidence showed teaching evaluation scores are gender biased against female faculty members, particularly in traditional male-dominated areas such as finance.}\footnote{33}
  \item This gender bias was specifically recognized and noted by her department in its letter recommending that she be granted tenure.\footnote{34}
  \item The gender-bias in student teaching evaluation scores has been well-documented in the literature.\footnote{35}
  \item Finally, the grievant noted the literature has numerous calls for “multiple measures” of teaching effectiveness because so many studies have found negative biases toward women professors — especially in fields traditionally dominated by men.\footnote{36}
  \item Thus, it can be considered a misuse of student evaluations to use them as an exclusive measure of teaching effectiveness for women teaching in traditional male areas — above and beyond the general argument that student evaluations are not a reliable measure of teaching effectiveness.
  \item\textit{Females in traditionally male-dominated areas, such as engineering and business, are held to higher standards than males.}\footnote{37}
  \item Some researchers suggest this double standard, under which females are expected to be competent in their work but are also either rewarded for conforming to gender stereotypes or punished for not being warmer and more nurturing than their male counterparts, may be a result of the university involved having been redacted.
\end{itemize}
of male students in areas such as business and engineering being more “traditional” in their views of appropriate female roles.\(^{38}\)

\textit{ii. Beauty Bias}

Studies have found that instructors who are viewed as better looking receive higher instructional ratings on required student evaluations.\(^{39}\) It is fairly well established that, after adjustment for many other determinants, beauty affects earnings in the marketplace.\(^{40}\) Using a sample of 463 courses taught at the University of Texas at Austin between the academic years 2000 and 2002, one study looked at 16,957 completed student teaching evaluations from a population of 25,547 registered students.\(^{41}\) The ninety-four instructors who taught these courses were then rated by a representative panel of undergraduate students on beauty, and the researchers unit normalized each rating and summed these ratings “to create a composite standardized beauty rating for each instructor.”\(^{42}\)

The researchers, after analyzing the data, concluded professors who are perceived as “better looking” received higher instructional ratings, with the impact of a move on the composite beauty scale from the tenth to the ninetieth percentile being substantial.\(^{43}\) The impact was found within university departments and even within particular courses;\(^{44}\) and it was larger for male than female instructors and slightly asymmetrical — meaning the penalty was greater for bad looks than the reward for good looks.\(^{45}\) In addition to studying pulchritude among University of Texas at Austin teaching faculty, the researchers also found the following:

Minority faculty members receive lower teaching evaluations than do majority professors, and non-native English speakers receive substantially lower ratings than do natives. Lower-division
courses are rated slightly lower than upper-division courses. Non-tenure track instructors receive course ratings that are surprisingly almost significantly higher than those of tenure-track faculty. This may arise because they are chiefly people who specialize in teaching rather than combining teaching and research, or perhaps from the incentives (in terms of reappointment and salary) that they face to please their students.

Perhaps the most interesting result among the other variables in the vectors X and Z is the significantly lower ratings received by female instructors, an effect that implies reductions in average class ratings of nearly one-half standard deviation.\(^{47}\)

iii. Asian Bias

In late January of 2005, Representative Bette Grande of North Dakota, a substitute middle and high-school teacher when the legislature is not in session, proposed a bill in the North Dakota legislature giving tuition refunds to college students who complained their instructors did not speak English fluently.\(^{48}\) Representative Grande’s concerns were aroused when she spoke with her former students and with friends of her college-aged son, a student at North Dakota State University.\(^{49}\) These students complained that they had trouble in class understanding a professor’s accent.\(^{50}\) When Representative Grande took her concerns to the North Dakota State University provost, he defended the policy of using foreign-born teaching assistants and urged that North Dakotans, who can be “fairly provincial,” need to learn to live in an increasing global world.\(^{51}\) Representative Grande dismissed the provost’s responses as “more excuses.”\(^{52}\)

Research performed at the University of Georgia showed undergraduate students listening to the exact same lecture but watching a photo of an Asian woman instead of an equally well-dressed Caucasian woman missed significantly more questions about the lecture than the group watching the photo of the Caucasian woman.\(^{53}\)

\(^{47}\) Id. at 6.
\(^{49}\) Id.
\(^{50}\) Id.
\(^{51}\) Id.
\(^{52}\) Id.
Consider, for example, the case of Min Liu. In the summer of 2002, Min Liu traveled from Shanghai, China to Fargo, North Dakota to begin Ph.D. studies in communication at North Dakota State University. Ten days later, she was teaching her first course at NDSU. Although there are very faint traces in her speech of native Chinese language interference with a few English pronunciations, Ms. Liu completed a “widely used” oral examination for English proficiency and received the maximum possible score.

Notwithstanding Ms. Liu’s proficiency in English, her sophisticated vocabulary, and three years experience as a teaching assistant, she still receives undergraduate complaints about her teaching every time she teaches a course. These complaints have been “always on anonymous end-of-semester course evaluations and never from a student in person.”

iv. The “Miss Congeniality” Bias

Although recent studies have indicated that the relationship between “entertaining” or “expressive” (i.e., “witty, enthusiastic, theatrical, or engaging”) behavior on the part of instructors and high student ratings may have been overstated in initial research findings due to various flaws in the studies, more current research sustains the conclusion that “expressiveness” generates higher student ratings on teaching evaluations. Moreover, one article put forward by the Center for Research on Learning and Teaching declared instructor “expressiveness” in the case of students who were not highly motivated (e.g., those in introductory courses or required courses) had a more powerful effect on student achievement than the amount of content covered in the course.

A study conducted at Miami University showed that students’ concepts of good teaching were heavily based on the relationships

---

54. Gravois, supra note 48.
55. Id.
56. Id.
57. Id.
58. Id.
59. Id.
61. Id. The article gave no direct or supporting citation for this finding and merely listed bibliographic references on expressiveness, the most recent of which was published in an anthology in 1992. The perspective of the authors was that high student evaluations translated into higher student learning and hence almost directly into teaching effectiveness. This bias is not supported by a number of other studies and authors.
between faculty and students, whereas professors understood teaching as related to information transfer and testing fairness.\textsuperscript{62} One result of this study was that on more than half of the twenty-four items tested, student and faculty responses differed to a statistically significant degree.\textsuperscript{63} The disagreements were along coherent lines; faculty emphasized fairness in grading and pedagogical concerns whereas the students emphasized such items as “smiles and displays a friendly demeanor, treats students as equals, greets students encountered outside of class,” and “adjusts course material to fit student interests.”\textsuperscript{64}

\textit{v. The “thirty-Second Snapshot”}

Research conducted at Harvard University as part of a doctoral dissertation asked undergraduate student judges to view three ten-second clips of videotapes of thirteen graduate teaching fellows.\textsuperscript{65} These student judges were then asked to evaluate 15 dimensions of non-verbal behavior for each teacher on a linear scale from one (not at all) to nine (very).\textsuperscript{66} The results showed “ten of the fifteen non-verbal dimensions were significantly predictive of the criterion variable of teacher effectiveness” as contained in student evaluations of the same Teaching Fellows by students in their Harvard classes at the end of the semester.\textsuperscript{67} The researchers noted, “[i]t is interesting that these results are consistent with other findings based on far longer observations” and that both are related to student evaluations of teaching.\textsuperscript{68} A less technical observation would be that the first thirty seconds in a classroom — even a non-speaking thirty seconds — are reliable predictors of student evaluations at the end of the term. If this observation is correct, and the results certainly seem to support such a conclusion, the obvious next question is: what do student evaluations really measure?

\textsuperscript{62} David J. Walsh & Mary Jo Maffei, \textit{Never in a Class by Themselves: An Examination of Behaviors Affecting the Student-Professor Relationship}, 5 J. ON EXCELLENCE IN COLLEGE TEACHING 23, 38 (1994).
\textsuperscript{63} \textit{Id.}
\textsuperscript{64} \textit{Id.} at 38-39.
\textsuperscript{66} \textit{Id.}
\textsuperscript{67} \textit{Id.} at 434.
\textsuperscript{68} \textit{Id.}
vi. Measuring the Number of Rows in the Classroom

One study found, using multiple-regression analysis, that the number of rows of seats in the classroom was significantly negatively correlated to teaching evaluations. That study involved the ratings of instructors in seventy-five sections of college algebra freshman classes during the three semesters from fall 2001 through fall 2002 at California State University, Long Beach. The researchers reported

The main finding of this study is that student assessments of the effectiveness of their college algebra instructor are influenced by the dimensions of the classroom while taking into the account the number of enrolled students. The greater the number of rows in the classroom, the lower the average student evaluations.

vii. Correlation to Anticipated Grades in Course

The study at California State University also showed that mean student grades were positively correlated to student ratings of teachers. However, and perhaps more importantly, the authors noted that this finding has "been consistently reported in the literature." The defenders of student teaching evaluations occasionally make the argument that higher grades in classes taught by better teachers should be expected, apparently oblivious to both the unsubstantiated assumption of causation when data establish only correlation, as well as the blatant assumption of an answer to the most controversial issue in the literature, namely: whether student evaluations really measure classroom effectiveness. The literature shows higher mean grades in classes travel together with higher teacher evaluations. The contention is centered around the nature of the relationship. Fairly sophisticated research supports the position college
faculty engage in an implicit trade-off of grades for evaluations.\textsuperscript{77} One researcher concluded, “[T]he results from this analysis provide conclusive evidence of a biasing effect of student grades on student evaluations of teaching.”\textsuperscript{78} He added the following warning about policy implications for the quality of higher education in the United States:

From a policy viewpoint, the findings of this study are important. As an increasing number of universities use student evaluations of teaching in administrative decisions that affect the careers of their faculty, the incentives for faculty to manipulate their grading policies in order to enhance their evaluations increase. Because grading policies affect student enrollment decisions and the amount students learn in their courses, the ultimate consequence of such manipulations is the degradation of the quality of education in the United States.\textsuperscript{79}

\textit{viii. Smaller Classes Generate Better Evaluations}

A few researchers found no significant relation between class size and others argue that the relationship is curvilinear, “[m]ost studies report the influence of class enrollment on student evaluations, revealing that lower enrollment is associated with higher evaluations.”\textsuperscript{80} It should be noted that class size is something over which the individual faculty member has absolutely no control; yet, when administrators use student evaluation results in a normative fashion (e.g., those faculty above the mean are “good” teachers), the individual faculty gets the credit or bears the burden of the administrative decisions on class size. Reducing class sizes is one of the few actions that education research almost unanimously shows improves instruction;\textsuperscript{81} however, college and university administrators face strong pressures not to embrace this choice.\textsuperscript{82}

\textit{B. The Loneliness of a Heterodox Instructor Inside an Orthodox Circle}

Our review of the literature did not turn up any studies of how heterodox faculty members fare on their student evaluations in schools

\begin{itemize}
\item 77. \textit{Id.} at 16.
\item 78. \textit{Id.}
\item 79. \textit{Id.}
\item 80. Safer, Farmer, Segalla & Elhoubi, \textit{supra} note 27, at 28.
\item 82. See, \textit{e.g.}, Murray Sperber, \textit{Beer and Circus: How Big-Time College Sports is Crippling Undergraduate Education} (2000).
\end{itemize}
and colleges where the vast majority of faculty, staff, and students are orthodox. We suspect that, based on the studies of non-native English speakers and on the biases toward women instructors, such faculty members do not receive evaluations in line with their teaching. We are not referring to orthodoxy to mean religious preferences here but rather situations such as a Marxist economics professor in a business school’s economics department, or a business ethicist in a finance Department dominated by University of Chicago style economics, or a conservative legal positivist in a law faculty dominated by critical legal studies scholars. Our hypothesis is that at many schools, when the student body reflects and has absorbed the biases of the faculty — especially the case in first degree programs — the tendency will be for students to decline to engage the message at the level of ideas but rather to symbolically “shoot the messenger” on anonymous, end-of-semester teaching evaluations.

C. Examples of how Student Evaluation Results are Misused by College Administrators

In one undergraduate institution, the Dean’s Office ordered the department chairs to “rank order” every teaching member of their department using summaries of teaching evaluations. One department chair arbitrarily selected five questions, averaged them, and then used the statistical results to rank order the eighteen teaching faculty in his department. These ratings were then routinely used to distribute fifty percent of the new monies for annual pay raises. This takes “norming” to new levels because the information necessary to even determine if the numerical differences were statistically significant (i.e., 2.96 versus 2.99), was not provided.

“Norming” student evaluations on the Lake Wobegon 83 principle that all of our professors have to be above average makes an indefensible assumption that all faculty with student evaluation numbers over the mean are “good” teachers and those below the mean are “bad” teachers. 84 The truth may not even be close to that assumption. Suppose eighty percent of the teachers are doing a very good job in the classroom; such a normative assumption would discriminate against good teachers. If eighty percent were mediocre or worse in the classroom, such an assumption would reward bad teaching.

84. Gray & Bergman, supra note 1.
II. IMPLICATIONS FOR DUE PROCESS AND FIRST AMENDMENT

For society's good — if understanding be an essential need of society — inquiries into . . . problems, speculations about them, stimulation in others of reflections upon them, must be left as unfettered as possible. Political power must abstain from intrusion into this activity of freedom, pursued in the interest of wise government and the people's well-being, except for reasons that are exigent and obviously compelling.  

— Justice Frankfurter

We begin this section with a discussion of academic freedom and its role in shaping the modern university. We then analyze the constitutional mandate for due process and protecting a faculty member's First Amendment rights.

A. Academic Freedom and its Role in the Modern University

Universities hold special regard in modern society as places where freedom of thought and inquiry are encouraged. The system of granting tenure to faculty members, for example, is meant to encourage faculty to engage in research, teaching, and service that is new and daring. The typical argument for tenure is that without it, faculty “face excessive pressure to avoid risky, innovative, or controversial topics in their research and teaching.” Tenure captures the idea that within academe, “professional autonomy and collegial self-governance” are the reigning values. In spite of this ideal, recent events illustrate that the notion of academic freedom is constantly under attack.

Sometimes the controversies surrounding academic freedom have little to do with faculty, but rather with the campus environment. Recently, incidents at multiple institutions across the country surrounding the Israeli-Palestinian conflict have stretched and sorely tested the ideals surrounding academic freedom. Such was the case

86. For a commentary on tenure as it evolves to meet new challenges, see James J. Fishman, Tenure: Endangered or Evolutionary Species, 38 AKRON L. REV. 771 (2005).
at Carnegie Mellon University, for example, when Malik Zulu Shabazz was invited to campus in February, 2005. Shabazz leads the New Black Panther Party for Defense, an organization the Southern Poverty Law Center designated as a hate group for its anti-Semitic doctrine. University officials asked the sponsoring student organization to reconsider, but its “commitment to free expression” meant that the students made the final decision. The resulting speech by Shabazz, which involved his aides walking the aisles with a nightstick, taking photos of white people, and denunciations of Jewish persons, stirred an angry response from the community, which argued that someone of Shabazz’s ilk posed no value to the educational mission of a university. In this case, the use of academic freedom as a defense for University-sanctioned actions resulted in a fierce backlash from Jewish students.

Princeton University recently faced a similar backlash from the Jewish community when a prominent Middle East scholar from Columbia, Rashid Khalidi, was considered for an endowed Chair in Contemporary Middle East Studies. Princeton’s Center for Jewish Life reacted strongly, accusing the administration of “hiring an individual who has a political agenda rather than a scholarly approach to history.” The debate took a ferocious turn at Columbia itself, when pro-Israeli students complained they were intimidated by various pro-Palestinian faculty members.

Vanderbilt University’s use of “creating a more welcome environment” as a rationale for a controversial university action also failed when the university attempted to remove the word “Confederate” from the “Confederate Memorial Hall” stone monument in front of a student residence hall in early 2005. An appeals court ruled that if
Vanderbilt took such action, it would have to repay the monetary gift from the United Daughters of the Confederacy that donated fifty thousand dollars in 1935. 98

In other instances, the attacks on academic freedom are more sinister and are aimed squarely at a faculty member's teaching or scholarship. Examples of universities terminating employment of instructors and adjunct faculty are plentiful. At Southern Methodist University, for example, writing instructor Elaine Liner was fired after writing an anonymous blog critical of “snooty” Ph.D.-holding faculty and mocking wealthy students looking for a “Mrs.” degree. 99 At Georgetown College, an adjunct sociologist professor, who cross-dressed occasionally at area restaurants, sued the college when he was fired even though he never cross-dressed at the college. 100 Suppose he had cross-dressed at the college, does that mean Georgetown has a dress code for faculty?

The examples are more difficult to comprehend when they affect tenured or tenure-track faculty, who, after all, are supposed to be the biggest beneficiaries of a tenure system that promotes academic freedom. Consider for instance the case of Professor Jonathan Bean, a prize-winning author and past winner of the College of Liberal Arts Teacher of the Year award. 101 In the spring of 2005, he distributed an optional handout in his survey course on American history. 102 The handout was a reprint from a conservative newspaper on the “Zebra Killings,” a series of black-on-white murders in San Francisco in the 1970s. 103 Other readings in the course included works by Dr. Martin Luther King, Jr. and Malcolm X. 104 In response, six other faculty members published a letter in the student newsletter “denouncing him for handing out ‘racist propaganda’ that was full of ‘falsehood and innuendo.’” 105 The Dean then cancelled the discussion sections of his class for a week and informed his teaching assistants that they no longer had to work for Professor Bean. 106

98. Id.
102. Id.
103. Id.
105. Young, supra note 101.
106. Jaschik, supra note 104.
David Graeber, a tenure-track anthropology professor at Yale, was not so lucky.\footnote{David Epstein, \textit{Early Exit}, \textsc{Inside Higher Ed}, May 18, 2005, http://www.insidehighered.com/news/2005/05/18/yale.} After teaching six years at Yale, the university failed to renew his contract for the last two years prior to a tenure vote.\footnote{Id.} Graeber is a popular teacher with good teaching evaluations and a prolific scholar with two books and dozens of articles, papers, and essays.\footnote{Id.} He believes that Yale did not renew his contract because he is a self-professed anarchist who supported the unionization of graduate student employees — in other words, for his political beliefs.\footnote{Id.} An online petition in support of Graeber has collected over four thousand signatures,\footnote{Id.; see also In Support of David Graeber, at http://www.PetitionOnline.com/dgraebertpetition.html (last visited Sept. 11, 2006).} and rumors swirl at Yale that six other faculty members in the Anthropology department may leave.\footnote{Id.}

Consider the case of Professor Miguel De La Torre, a minister and tenured professor at Hope College in Michigan.\footnote{Julie Post, \textit{Anthropology Profs Fear Exodus}, \textsc{Yale Daily News}, Sept. 23, 2005, http://www.yaledailynews.com/article.asp?AID=29962 (last visited Sept. 11, 2006).} In a column he wrote for a local newspaper, Professor De La Torre wrote satirically about the religious right’s condemnation of homosexuality.\footnote{Miguel De La Torre, \textit{When the Bible is Used for Hate}, \textsc{The Holland Sentinel}, Feb. 1, 2005, http://www.hollandsentinel.com/stories/020105/020105018.shtml.} In response, the President of Hope College sent him a letter criticizing his writings and claiming that it made it more difficult for Hope to raise money.\footnote{Id.} Despite his achieving tenure, publishing several books, and strong teaching reviews, Professor De La Torre was passed over for a merit pay increase, and eventually quit his position at Hope.\footnote{Id.}

Sometimes, controversy follows faculty members even when they engage in speech outside the classroom, outside their scholarship, and indeed, outside of every sphere of being a faculty member. Brooklyn College professor Timothy Shortell was voted by his colleagues to become the new Chair of the Sociology Department.\footnote{Scott Jaschik, \textit{Did SpongeBob Article Cost Professor a Job?}, \textsc{Inside Higher Ed}, Apr. 28, 2005, http://www.insidehighered.com/news/2005/04/28/hope.} Within days, news erupted that Professor Shortell held views that were critical of organized religion and the its followers.\footnote{Id.} In one
online essay, he referred to religious people as “moral retards” and compared them to children. The New York Sun, Fox New’s “The O’Reilly Factor,” and the Washington Times were quick to denounce Shortell. Brooklyn College President Cristoph Kimmich appointed an investigative committee. Amidst the outcry, Shortell voluntarily withdrew from the Chair’s position.

Academic freedom has been described as “that aspect of intellectual liberty concerned with the peculiar institutional needs of the academic community.” The academic freedom of university professors and researchers is generally understood to be freedom from political, ecclesiastical, or administrative interference with investigation, discussion, and publication in their field of study. Apparently, no one gave much thought to corporations chilling academic freedom before the 1990s. However, the American Association of University Professors (“AAUP”) did consider politics and religion as potential enemies of free thought.

In 1940, the AAUP produced the classic statement on academic freedom, the 1940 Statement of Principles on Academic Freedom and Tenure. In relevant part, it reads as follows:

a. Teachers are entitled to full freedom in research and in the publication of the results, subject to the adequate performance of their other academic duties; but research for pecuniary return should be based upon an understanding with the authorities of the institution.

119. Id.
120. See Chairman Shortell, N.Y. SUN, May 18, 2005, at 8.
121. The O’Reilly Factor: Professor Rewarded for Slamming Religion (Fox television broadcast May 26, 2005).
123. Jaschik, supra note 117.
124. Jacob Gershman, Professor Who Belittled Believers Drops Bid to Head Up a Department, NY. SUN, June 8, 2005, at 1.
b. Teachers are entitled to freedom in the classroom in discussing their subject, but they should be careful not to introduce into their teaching controversial matter which has no relation to their subject. Limitations of academic freedom because of religious or other aims of the institution should be clearly stated in writing at the time of the appointment.

c. College and university teachers are citizens, members of a learned profession, and officers of an educational institution. When they speak or write as citizens, they should be free from institutional censorship or discipline, but their special position in the community imposes special obligations. As scholars and educational officers, they should remember that the public may judge their profession and their institution by their utterances. Hence they should at all times be accurate, should exercise appropriate restraint, should show respect for the opinions of others, and should make every effort to indicate that they are not speaking for the institution.120

In 1967, the U.S. Supreme Court had the opportunity to address academic freedom.129 The case involved a declaratory judgment action seeking injunctive relief brought by faculty members of Buffalo State University who were notified that they would be fired for refusing to sign the “Feinberg Certificate.”130 This certificate declared that the signee was not a Communist and that if he had ever been a Communist, he had communicated that fact to the President of the University of the State of New York.131 In a five to four decision, the Court, in an opinion by Justice William Brennan, held that the New York statutes requiring the Feinberg Certificate were unconstitutionally overbroad because the state could achieve its objectives, namely “protecting its education system from subversion,” through less sweeping prohibitions.132 Brennan wrote:

Our Nation is deeply committed to safeguarding academic freedom, which is of transcendent value to all of us and not merely to the teachers concerned. That freedom is therefore a special concern of the First Amendment, which does not tolerate laws that cast a pall of orthodoxy over the classroom.133

The pall of “orthodoxy” in the 1990s and early twenty-first century seems to be self-imposed by many colleges and universities, which

128. Id.
130. Id. at 589, 595.
131. Id. at 589.
132. Id. at 603, 609.
133. Id. at 603.
are now dependent on financial contributors to keep operations going. In the 1990s, some state universities’ presidents began referring to their universities as “state-assisted” universities, indicating more accurately their financial relationship with their states. Problems with loyalty oaths and seditious speech seem to have disappeared. Now the issue is whether a professor’s research, publications, or speeches will offend a major corporation, a powerful lobby such as the NRA or AAUP, or a political interest group with political clout. One source suggests strongly that the modern college administrators are, at least partially, at fault for this climate.

One sad conclusion is that the First Amendment means little when university administrators and university professors engage in self-censorship to appease various interests. As another author has noted, “[b]ecause the loss of employment is so damaging, the expectation that they will be fired for expressing their opinions could have a serious chilling effect on individuals’ political speech.”

On February 3, 2005 at a raucous meeting during which two protestors were arrested, the University of Colorado Regents ordered a virtually unprecedented investigation into whether the controversial ethnic studies Professor Ward Churchill should be fired for cause. Interim Chancellor Phil DiStephano, assisted by Arts & Sciences Dean Todd Gleason and School of Law Dean David Getches, will direct the investigation. Professor Churchill’s act which precipitated this crisis was that he wrote an essay about the September 11, 2001 terrorist attacks on the Trade Towers in NYC and on the Pentagon that failed all measures of political correctness and most measures of sensitivity. He compared “technocrats” working in the World Trade Center to notorious Nazi bureaucrat Adolf

134. See David W. Breneman, For Colleges, This is Not Just Another Recession, CHRON. HIGHER EDUC., June 14, 2002, at B7.
137. Dale E. Miller, Terminating Employees for Their Political Speech, 109 BUS. & SOCY. REV. 225, 229 (2004); see also Lee Bollinger, University President, Columbia University, Address at National Press Club Luncheon (Apr. 2, 2003) (explaining that an untenured faculty member who had called for the United States to lose the Iraq war and proclaimed hope for the deaths of U.S. soldiers could not be fired because the speech did not occur in the classroom but, rather, at an open “teach-in”).
139. Id.
Eichmann and declared that the United States had invited such attacks by a long history of “violent domination of other countries and cultures.”

In March 2005 as one outcome of the controversy involving Ward Churchill, the President of the University of Colorado, Betsy Hoffman, resigned after she had defended Mr. Churchill’s academic right of free speech against critics who wanted him fired. Mr. Churchill’s three-year-old essay exploded into a political firestorm in January 2005 before a scheduled appearance of Churchill at Hamilton College in New York. In December 2004, when the Kirkland Project for the Study of Gender, Society, and Culture at Hamilton College sent an e-mail message to Hamilton Professors mentioning Mr. Churchill’s speech, one professor found and read Churchill’s essay and urged Hamilton officials to cancel the speech. On January 21, 2005, the campus newspaper highlighted Mr. Churchill’s remarks. A Syracuse, New York newspaper reported on the controversy on January 26, 2005, including comments by Hamilton professors who called the event “morally outrageous” and “an act of utter irresponsibility.”

On January 28, Bill O’Reilly discussed Mr. Churchill on his Fox television show and suggested that Professor Churchill be arrested for sedition. On February 1, 2005, Hamilton College cancelled the event, which had been transformed from the initial speech by Churchill into a panel discussion on political prisoners and prisoner issues. On February 3, 2005, the Colorado Board of Regents held a special meeting to discuss Professor Churchill and announced its apology and the probe into Churchill.

A self-appointed group of hyper-active right-wingers managed to intimidate both Hamilton College and the University of Colorado. It

140. Id.
141. See Paul Fain, Under Fire on 2 Fronts, U. of Colorado Chief Resigns, CHRON. HIGHER EDUC., Mar. 18, 2005, at A10. “The continuing controversy over Mr. Churchill has received as much if not more attention as the football scandal.” Id. The football scandal escalated in March, 2005 when someone leaked details of a sealed grand-jury report indicating that the grand jury had found an assistant football coach “had sexually assaulted two female athletic trainers.” Id.
143. Id.
144. Id.
145. Id.
146. Id.
147. Id.
149. Hamilton College had already shown itself to be vulnerable to allowing outside pressure to dictate public discourse on campus with the imbroglio over a one-month course on memoir writing scheduled to be taught by Susan Rosenberg, a former associate of the Weather Underground who served seventeen years of a fifty-eight year sentence.
also enabled politicians disposed to being hostile to Ward Churchill’s style (if not his ideas) into leading a head-hunting expedition. The collateral damage, so far, seems to be limited to the career of Betsy Hoffman, President of the University of Colorado, already wounded by a sports scandal, and the fabric of the “profound national commitment to the principle that debate on public issues should be uninhibited, robust, and wide-open.”

Professor Wines remembers being an undergraduate on the Northwestern University campus in Evanston when the politics of civil rights were being hotly debated. Northwestern, not by any standard a radical outpost of free thinking, handled the situation in a model way. The head of the American Nazi Party, George Lincoln Rockwell, was invited to campus for a structured and even-handedly moderated debate on race and civil rights in Scott Auditorium. Rockwell was, among other things, a prominent denier of the Holocaust at that time. Security was present, but civility held sway. What that one undergraduate came away with was an indelible impression of the way universities have served society for a thousand years. In an open and fair debate with informed scholars, Mr. Rockwell showed himself to be a small, bigoted, and not very intelligent man whose ideas did not hold up when confronted with reason in the bright light of public discourse. Despite Mr. Rockwell’s skills in political advocacy and his slick demeanor, the most outstanding lesson that night was that his ideas had no support in science, history, justice, or reason. The structure of free universities suffered no harm that night, but the propaganda of the American Nazi Party took a serious body blow.

Tenure systems at most universities are ostensibly based on merit, but, as many untenured faculty know, the secrecy surrounding tenure votes means that controversial faculty, no matter how accomplished or meritorious, risk a negative decision. The ideals of

142. Ms. Rosenberg withdrew in the face of the opposition and the public relations mess at Hamilton College in Clinton, NY. Id.


151. Some faculty of that era at Northwestern recall President J. Roscoe Miller canceling a planned speech by George Lincoln Rockwell in 1963. The author has no memory of that event or whether it happened. However, a later appearance by Rockwell in Cahn Auditorium on December 7, 1965 is recalled by Wines and could, possibly, have resulted from protests over an earlier cancellation. For more details, see WILLIAM H. SCHMALTZ, HATE: GEORGE LINCOLN ROCKWELL AND THE AMERICAN NAZI PARTY 268 (2000).

152. Id. at 49-50.

153. For an excellent summary of the origins and functioning of free western universities, see CHARLES HOMER HASKINS, THE RISE OF UNIVERSITIES (1957).
academic freedom are therefore attenuated by the process, even before one considers the impact of student teaching evaluations. Student evaluations, with all the biases they embrace, put pressure on faculty to go slow and not rock the boat. In other words, do not push undergraduates to maximize their intellectual potential because that might fuel resentment, and do not confront the dominant political and religious beliefs of your particular subset of late adolescents even when such beliefs are patently false\textsuperscript{154} and when confronting them is supposedly part of the education process and is course-appropriate. Undergraduates might retaliate on evaluations. Even post-tenure, as the Churchill situation demonstrates, faculty face pressures from students, alumni, administrators, and sometimes even the public and the media,\textsuperscript{155} to conform their views to those views popularly held.

B. Judicial Treatment of the “Special Concern”\textsuperscript{156} of the First Amendment — Academic Freedom

American jurisprudence has had little occasion to address the issue of academic freedom. A dearth of willing plaintiffs, perhaps, makes appellate court opinions on the issue rare. When the courts have weighed in, however, they have been quick to uphold the popular notion that academe is, and should be, a bastion of freedom of thought and speech.

The United States Supreme Court began weighing in on academic freedom with a series of cases in the 1950s.\textsuperscript{157} In \textit{Wieman v. Updegraff}, the Court was faced with an Oklahoma statute that re-
quired all state employees take a loyalty oath, affirming that they were not Communists.\textsuperscript{158} The case involved faculty and staff of the Oklahoma Agricultural and Mechanical College who refused to take the oath, and resulted when a private citizen filed suit to enjoin the state from paying salaries of the faculty and staff, an action authorized by the state statute.\textsuperscript{159} The Supreme Court held that “indiscriminate classification of innocent with knowing activity” was an assertion of arbitrary power and overturned the statute.\textsuperscript{160} In a concurring opinion joined by Justice Douglas, Justice Frankfurter wrote:

To regard teachers . . . as the priests of our democracy is therefore not to indulge in hyperbole. It is the special task of teachers to foster those habits of open-mindedness and critical inquiry which alone make for responsible citizens, who, in turn, make possible an enlightened and effective public opinion . . . [Teachers] cannot carry out their noble task if the conditions for the practice of a responsible and critical mind are denied to them . . . They must be free to sift evanescent doctrine, qualified by time and circumstance, from that restless, enduring process of extending the bounds of understanding and wisdom, to assure which the freedoms of thought, of speech, of inquiry, of worship are guaranteed by the Constitution of the United States against infraction by National or State government.\textsuperscript{161}

Three years later, college professor Paul Sweezy found himself in the crosshairs of New Hampshire Attorney General Louis Wyman, who wanted to know if Sweezy had ever advocated the theory of dialectical materialism during lectures delivered at the University of New Hampshire.\textsuperscript{162} Sweezy was subpoenaed to reveal his “personal political views, the investigation of his political activities, and, finally . . . the names of others with whom he was active . . . [as] part of the larger ‘anticommunist’ purge against militant trade unionists . . .”\textsuperscript{163} Sweezy refused to answer the questions, observing that witnesses such as himself essentially only had two choices when faced with a subpoena such as the one he was facing.\textsuperscript{164} His first option was to plead the Fifth Amendment, an option he rejected because “one of the

\begin{itemize}
  \item \textsuperscript{158} Wieman v. Updegraff, 344 U.S. 183, 184-85 (1952).
  \item \textsuperscript{159} \textit{Id.}
  \item \textsuperscript{160} \textit{Id. at 191.}
  \item \textsuperscript{161} \textit{Id. at 196-97} (Frankfurter, J., concurring).
  \item \textsuperscript{163} \textit{Id.}
  \item \textsuperscript{164} See Paul M. Sweezy, \textit{Statement to the New Hampshire Attorney General}, MONTHLY REV., Apr. 1, 2000, at 37.
\end{itemize}
noblest and most precious guarantees of freedom, won in the course of bitter struggles and terrible suffering, has been distorted in our own day to mean a confession of guilt, the more sinister because undefined and indeed undefinable.\textsuperscript{165} The second option, and the one he chose, was to contest the “legitimacy of offensive questions” on constitutional grounds.\textsuperscript{166} Sweezy was convicted of contempt for his refusal to answer questions, and he appealed his case to the Supreme Court.\textsuperscript{167}

The Supreme Court, in a 6-2 decision in favor of Sweezy, observed that the power of legislative investigations should be “carefully circumscribed” when the investigation concerns “highly sensitive areas” such as “freedom of communication of ideas, particularly in the academic community.”\textsuperscript{168} The Court declared that the areas of academic freedom and political expression should be areas in which the government “should be extremely reticent to tread.”\textsuperscript{169} Most importantly, the Court added to the \textit{Wieman}\textsuperscript{170} understanding of academic freedom by writing:

\begin{quote}
No one should underestimate the vital role in a democracy that is played by those who guide and train our youth. To impose any strait jacket upon the intellectual leaders in our colleges and universities would imperil the future of our Nation . . . Scholarship cannot flourish in an atmosphere of suspicion and distrust. Teachers and students must always remain free to inquire, to study and to evaluate, to gain new maturity and understanding; otherwise our civilization will stagnate and die.\textsuperscript{171}
\end{quote}

In the ensuing years, the Supreme Court invalidated an Arkansas statute requiring teachers to annually file an affidavit listing every organization to which they “belonged or regularly contributed within the preceding five years,”\textsuperscript{172} a loyalty oath for all University of Washington employees,\textsuperscript{173} a requirement that State University of New York faculty members swear that they had never been Communists, known as the Feinberg law,\textsuperscript{174} a Maryland statute that required faculty to swear that they never advocated the “alteration” of the

\begin{footnotes}
\textsuperscript{165} Id.
\textsuperscript{166} Id.
\textsuperscript{167} Simon, \textit{supra} note 162, at 37.
\textsuperscript{169} Id. at 250.
\textsuperscript{170} \textit{Weiman}, 344 U.S. at 183.
\textsuperscript{171} Id.
\textsuperscript{172} Sheldon v. Tucker, 364 U.S. 479 (1960).
\textsuperscript{174} \textit{Keyishian}, 385 U.S. at 589.
\end{footnotes}
United States government,\textsuperscript{175} and an Arkansas statute\textsuperscript{176} that made it illegal for teachers to teach the theory of evolution. The Court’s jurisprudence also began to acknowledge that academic freedom was an important component of the University’s freedom, and not just the faculty members’ freedom.\textsuperscript{177} Most recently, in a case upholding the University of Michigan Law School’s use of race as a criterion in achieving a “critical mass”\textsuperscript{178} of minority students, the Court has reaffirmed the view that academic freedom is a “special concern” of the First Amendment.\textsuperscript{179}

\textbf{C. Modern Academic Freedom Under Attack}

In spite of the Supreme Court’s recognition of the special protection academic freedom deserves under the First Amendment, three recent cases, two of which were heard by the Supreme Court in the 2005 term, have the potential to reshape the boundaries of academic freedom. The three cases appear dissimilar on their face. In \textit{Urofsky v. Gilmore},\textsuperscript{180} the issue surrounds the use of state-owned computers for accessing sexually-related materials. In \textit{Garcetti v. Ceballos},\textsuperscript{181} retaliation actions against state employees for speech uttered in their capacities as state employees are at stake. Finally, in \textit{Forum for Academic and Institutional Rights v. Rumsfeld},\textsuperscript{182} several law schools challenged a federal law that punishes law schools which bar military recruiters. While the factual situations in each case are different, the

\textsuperscript{175} Whitehill v. Elkins, 389 U.S. 54 (1967).

\textsuperscript{176} Epperson v. Ark., 393 U.S. 97, 109 (1968). \textit{But see} Minn. State Bd. for Community Colleges v. Knight, 465 U.S. 271, 273 (1984) (holding that faculty members’ First and Fourteenth Amendment rights were not violated when the university conferred with only one teaching union, to which the plaintiffs did not belong, pursuant to state law).

\textsuperscript{177} \textit{See}, e.g., Regents of the Univ. of Mich. v. Ewing, 474 U.S. 214, 226 (1985) (holding academic freedom includes autonomous decision-making in ruling that University did not violate student’s constitutional rights by refusing to allow him to retake a test). \textit{But see} Univ. of Pa. v. EEOC, 493 U.S. 182 (1990) (holding that academic freedom and the First Amendment did not confer upon a university a right to withhold tenure review documents from EEOC subpoena during Title VII investigation).


\textsuperscript{180} Grutter, 539 U.S. at 324 (quoting Powell, J., Regents of the Univ. of Cal. v. Bakke, 438 U.S. 289, 312 (1978)). \textit{But see id.} at 362 (Thomas, J., dissenting) (arguing the First Amendment and academic freedom did not create the right for universities to violate the Equal Protection Clause). For an excellent analysis of the First Amendment and its foundation for the \textit{Grutter} decision, see Paul Horwitz, \textit{Grutter’s First Amendment}, 46 B.C. L. Rev. 461 (2005).

\textsuperscript{181} Garcetti v. Ceballos, No. 04-473, slip op. (U.S. 2005).

resulting impact of each have startling implications for academic freedom.

In *Urofsky*, six professors from various public universities and colleges in Virginia filed suit challenging a Virginia state law that restricted state employees from using state-owned computers or networks to access or download sexually explicit materials without prior written consent from the state agency for which the employees worked (in this case, public colleges and universities). The professors argued that they possessed a constitutional right to determine for themselves, “without the input of the university (and perhaps even contrary to the university’s desires), the subjects of [their] research, writing, and teaching.” In dismissing this argument, the Fourth Circuit held that “to the extent the Constitution recognizes any right of ‘academic freedom’ above and beyond the First Amendment rights to which every citizen is entitled, the right inheres in the University, not in individual professors . . . .” In his dissent, Justice Murnaghan pointed out that the statute would restrict professors from accessing “works of Toni Morrison and many themes found in Victorian poetry.” The fact that the statute also applies to email means that discussions by state professionals with their clients on abnormal sexual behaviors would also be implicated. In September 2005, the Sixth Circuit joined the Fourth Circuit in holding that any right of academic freedom belongs to the university and not to the individual faculty member. This result seems to conflict with the Second Circuit, which has held that government administrators should not “discipline a college teacher for expressing controversial, even offensive, views lest a ‘pall of orthodoxy’ inhibit the free exchange of ideas in the classroom.” Until the issue of whether academic freedom rests with the faculty or the institution is clarified by the Supreme Court, faculty must labor under differing First Amendment standards depending on where they teach. As the Court declined to hear *Urofsky*, professors’ struggle with this issue will continue.

The issue of how much protection the First Amendment affords public employees is revisited in *Garcetti*. The plaintiff was a deputy

---

184. *Urofsky*, 216 F.3d at 409-10.
185. *Id.* at 410.
186. *Id.* at 438.
187. *Id.* at 439.
188. *Id.*
190. Vega v. Miller, 273 F.3d 460, 467 (2d Cir. 2001) (citing Keyishian, 385 U.S. at 603).
district attorney who believed that the investigating police in a criminal case may have lied to obtain a search warrant.\textsuperscript{192} The defense counsel subpoenaed the plaintiff to testify at a suppression hearing.\textsuperscript{193} The plaintiff alleges that after he testified, he was demoted and subject to other retaliatory actions.\textsuperscript{194} A federal district court granted summary judgment to the defendants on the basis of governmental immunity.\textsuperscript{195} A unanimous Ninth Circuit reversed, holding that absolute immunity only applied to the district attorney’s office during litigation, and the actions taken against the plaintiff were administrative in nature, and therefore the prosecutor only had qualified immunity.\textsuperscript{196} The Ninth Circuit went on to address the plaintiff’s First Amendment claims. The Ninth Circuit applied Supreme Court precedents that first require a determination of whether the speech addresses a matter of public concern,\textsuperscript{197} and if it does, whether the plaintiff’s “interest in expressing himself outweighs the government’s interest in promoting workplace efficiency and avoiding workplace disruption.”\textsuperscript{198}

The defendant’s arguments in this case pose a clear threat to academic freedom. In essence, the defendant argues that speech made in routine reports or made in the performance of routine job functions is not protected by the First Amendment.\textsuperscript{199} In an amicus brief filed on behalf of the California State Association of Counties, the Association wrote:

the looming, yet uncertain threat of federal judicial scrutiny may well chill the exercise of what otherwise be rational, effective management practices. In order to avoid liability in litigation, the public employer is likely to err on the side of foregoing corrective criticism of government while carrying out job duties. As a result, the government will retain workers who have been identified as ineffective or unsatisfactory.\textsuperscript{200}
As the AAUP stated in its amicus brief in support of the plaintiff, adoption of the rule proposed by the defendants would introduce a perverse irony:

First Amendment academic freedom would extend only to those public statements on which faculty members were least well informed — matters that fell totally outside the fields in which they study and teach. Only those statements clearly beyond academic expertise would be considered ‘matters of public concern.’

The Ninth Circuit squarely rejected adoption of a per se rule stripping all First Amendment protection from an employee’s speech if it is made in the routine course of the employee’s business: “Whether a job duty is routine or non-routine is a far less important factor for purposes of First Amendment analysis than the content of the public employee’s speech.” This case was originally scheduled to be heard during the Court’s 2005-2006 term and oral arguments were heard on October 11, 2005. In early 2006, however, the Court took the extraordinary step of rescheduling the case so that newly sworn Justice Alito could participate in the decision. This case was the only case rescheduled for a new round of oral arguments to allow Justice Alito to vote. In ruling against the plaintiff, the Court held that “when public employees make statements pursuant to their official duties, the employees are not speaking as citizens for First Amendment purposes, and the Constitution does not insulate their communications from employer discipline.” The Court further held that Ceballos “did not speak as a citizen.” In *Rumsfeld v. FAIR*, the plaintiff was an association of thirty-six law schools and law faculties which do not permit military recruiters to recruit law students at their schools because of the military’s policy on homosexuals. The participating law schools challenged the
constitutionality of the Solomon Amendment, a Congressional measure that threatened to eliminate federal funding to any school that continued to restrict military recruiter access to their students.\textsuperscript{210} The Third Circuit Court of Appeals reversed the District Court’s denial of the Association’s motion for preliminary injunction, holding that FAIR had a “reasonable likelihood of success on the merits of its expressive association claim against the Solomon Amendment.”\textsuperscript{211} The court relied on \textit{Boy Scouts of America v. Dale},\textsuperscript{212} in which the Supreme Court held that New Jersey’s public accommodations law could not be used to force the Boy Scouts to accept a gay Scout leader.\textsuperscript{213} The Third Circuit also held that the Solomon Amendment amounted to “compelled speech,” a separate violation of the First Amendment.\textsuperscript{214} The court held that if the law schools were forced “to comply with the Solomon Amendment, the law schools must affirmatively assist military recruiters in the same manner they assist other recruiters, which means they must propagate, accommodate, and subsidize the military’s message.”\textsuperscript{215}

A unanimous Supreme Court overturned the Third Circuit.\textsuperscript{216} The Court addressed the expressive association issue by drawing the distinction that whereas the BSA was compelled to accept a gay member and leader, the law schools are not being required to accept the military as a member within their own ranks.\textsuperscript{217} In addition, the Court pointed out that “[s]tudents and faculty are free to associate to voice their disapproval of the military’s message.”\textsuperscript{218} In addressing the compelled speech issue, the Court concluded that the Solomon Amendment “does not dictate the content of the speech at all, which is only ‘compelled’ if, and to the extent, the school provides such speech for other recruiters.”\textsuperscript{219}

Finally, attacks on academic freedom have originated within the legislative branch of government as well. Legislatures in California,\textsuperscript{220}

\begin{itemize}
  \item 212. \textit{Id.} at 230 (citing \textit{Boy Scouts of Am. v. Dale}, 530 U.S. 640, 653 (2000)).
  \item 213. \textit{Id.}
  \item 214. \textit{FAIR}, 390 F.3d at 236.
  \item 215. \textit{Id.}
  \item 217. \textit{Id.} at 19.
  \item 218. \textit{Id.} at 20.
  \item 219. \textit{Id.} at 12.
\end{itemize}
Colorado, Florida, Georgia, Indiana, Kansas, Maine, Massachusetts, Minnesota, New York, North Carolina, Ohio, Pennsylvania, South Dakota, Tennessee, and Washington have all deliberated bills calling for an “Academic Bill of Rights.” For example, the Florida bill would set a “statewide standard that students can not be punished for professing beliefs with which their professors disagree.” On the federal level, some Congressional Republicans are pushing for a similar measure “to ensure that students hear ‘dissenting viewpoints’ in class and are protected from retaliation because of their politics or religion.” While ostensibly about academic freedom, these measures inject legislative oversight into what faculty members say in the classroom, at great expense to academic freedom.

These recent Supreme Court cases, along with the movement for legislative oversight of higher education curricula, lead us to believe

236. Vanlandingham, supra note 222.
238. University action to tailor faculty members’ speech in order to keep donors happy is already a serious problem in higher education. See, e.g., Daniel Golden, In Religion Studies, Universities Bend To Views of Faithful, WALL ST. J., Apr. 6, 2006, at A1.
that First Amendment protection of academic freedom, formulated as recently as *Grutter*,239 is very much at risk. More problematically, without this protection, university administrators gain more discretion to use student teaching evaluations improperly.

III. CIVIL RIGHTS STATUTES AS POSSIBLE ALTERNATIVE THEORIES

Title VII of the 1964 Civil Rights Act240 applies to private employers with 15 or more employees and to state governments. The use of student evaluation numbers, particularly if normed, would seem to be an illustration of disparate impact (neutral on their face but statistically discriminatory in impact on protected groups). Research could support a finding that such normed use of student evaluations discriminate against women, Asians, and older people. In addition to Title VII, there is also Title IX of the 1964 Civil Rights Act, as amended primarily in 1972, that covers all educational institutions receiving federal financial aid.241

The Civil Rights Restoration Act of 1987 prohibits discrimination based upon race, color, religion, sex, national origin, age, and handicap and again covers all educational institutions receiving federal funds.242 Finally, — maybe too obvious — The Equal Pay Act of 1963 outlaws sex differences in pay for substantially equal work.243 It could be a violation of all of these laws if the department or college used normed student evaluations to distribute either merit or discretionary pay on the basis of anonymous, unvalidated student numbers, which discriminate based on race or sex.

Initially, our interest was in the issue of whether the use of normed student evaluations presented First Amendment and Academic Freedom issues. However, the current decisions at the U.S. Circuit Court level, especially *Urofsky*,244 and at the Supreme Court level, especially *Garcetti*,245 raise the specter of a vastly reduced interpretation limiting Academic Freedom to the institutions. With such case law in the circuits and in light of the Supreme Court’s reluctance to grant certiorari and its recent jurisprudence, use of civil rights statutes may be an important alternative theory for plaintiffs to consider. However, an extensive review of possible applications of this theory is beyond the scope of this article.

---

244. *Urofsky*, 216 F. 3d, cert. denied, 531 U.S. at 1070.
245. *Garcetti*, No. 04-473, slip op.
IV. IMPLICATIONS FOR FUTURE RESEARCH & CONCLUSIONS

Academic freedom is a value that most Americans profess to cherish and celebrate. The Supreme Court has recognized the special value of academic freedom under the First Amendment, but in recent jurisprudence has signaled an unwillingness to extend speech boundaries for faculty members and their speech. This reluctance leaves faculty members very much at the mercy of administrators, who use student teaching evaluations for personnel decisions. The current normative use of student evaluations for personnel decisions in virtually all of higher education is, in large part, unfair. One might also reasonably question whether the practice is effective, given recent research that indicates that twenty percent of U.S. college students completing “4-year degrees . . . have only basic quantitative literacy skills”246 and that “more than 50% of students at 4-year colleges . . . lack the skills to perform complex literacy tasks, such as comparing credit card offers with different interest rates.”247 We believe, and have argued here, that the near-obssession with student teaching evaluations has led faculty to carefully tailor their pedagogy towards maintaining evaluations that continue employment rather than educate students.

Whether such practice is an unconstitutional violation of First Amendment rights or of academic freedom is much less clear. A federal district court recently called this area of law “admittedly amorphous.”248 The older Supreme Court cases, such as Keyishian,249 would probably support a finding that such use of non-validated, unpiloted and anonymous student violations constituted a First Amendment violation. The current case law supports a very restricted interpretation of academic freedom and raises the possibility that such uses of student evaluations might not violate Academic Freedom or the First Amendment. Such uses might, however, trigger relief for some professors under the Civil Rights laws. One thing is clear: as our society develops and matures, we are beginning to see the public marketplace of ideas dwindle to just a few speakers.250 If administrators continue to use student teaching evaluations as the sole criterion for judging teaching effectiveness, we can rely on academe to join the voices of those silenced.

247. Id.
250. The Project for Excellence in Journalism reports that even though the number of media outlets has increased in the United States, they are carrying fewer stories. The State of the News Media 2006, Project for Excellence in Journalism, http://www.stateofthenewsmedia.com/2006/.