

What is Ethical About Grade Inflation and Coursework Deflation?

Donald Larry Crumbley · Ronald E. Flinn ·
Kenneth J. Reichelt

© Springer Science+Business Media B.V. 2010

Abstract Recent research questions the validity of student evaluation of teaching (SET) data to measure teaching and learning. Yet, there is extensive use of this instrument around the world, which arguably contributes to a decline in the rigor of college classes. This performance measurement has led to both unethical grade inflation and coursework deflation as faculty try to entertain students rather than educating them. These unethical teaching techniques used by many faculties are on the same plane as the unethical practices of executives “cooking their books.” Ethical and unethical SET management techniques of professors are discussed herein, along with incentive and structural pander pollution of administrators and universities.

Keywords Student evaluation of teachers · Pander pollution · Grade inflation · Dysfunctional behavior · Impression management · Earnings management

As more and more research questions the validity of summative student evaluation of teaching (SET), higher education is experiencing the simultaneous widespread use of this control device (Flinn and Crumbley 2009; Adams 1997; Haskell 1997; Newton 1988, Wright et al. 1984; Powell 1977; Ditts 1983; DuCette and Kenney 1982; Howard and Maxwell 1982; Worthington and Wong 1979; Brown 1976; Porcano 1984; Dowell and Neal 1983; Stumpf and Freedman 1979). A persuasive argument can be made that this increased use of SETs for administrative control has caused grade inflation, coursework deflation, and a reduction in student learning as a result of unethical behavior of professors and administrators.

D. L. Crumbley (✉) · K. J. Reichelt
E. J. Ourso College of Business, Louisiana State University, Baton Rouge, LA 70803, USA
e-mail: dcrumbl@lsu.edu

K. J. Reichelt
e-mail: reichelt@lsu.edu

R. E. Flinn
College of Business Administration, Creighton University, Omaha, NE 68178, USA
e-mail: RonaldFlinn@creighton.edu

The typical SET questionnaire treats the student as a customer and measures the satisfaction of the student with his or her professor, and not learning. SET scores are as much an expression of customer satisfaction rather than anything approaching meaningful, valid, and reliable measures of what a student actually learns in a college class (Saje 2005; Wallace and Wallace 1998; Zimmerman 2002). A committee at the Ohio State University reported that

the SET process measures student satisfaction instead of the quality of instruction, has questionable reliability, and does not take into account certain factors that influence the course (in particular the effect of grades). Whether consciously or not, many faculty do pander to students in terms of rigor and grades in order to influence SET results. Untenured faculty in particular often feel pressured to adopt this course of action when placed in situations where the use of a designated SET instrument is mandated (Zimmerman 2002).

A student earning an A in a course is much more satisfied with an instructor than one earning a F. A number of researchers have found a positive relationship between expected grades and SET ratings (Aigner and Thum 1986; Ditts 1983; Eiszler 2002; Isely and Singh 2005; Kau and Rubin 1976; Krautmann and Sander 1999; McPherson 2006; Mehdizadeh 1990; Nichols and Soper 1972; and Zangenehzadeh 1988). A survey by Crumbley and Reichelt (2009) suggests that instructors ease grading, inflate grades and deflate course work, when SET data is used for faculty evaluation purposes. By inflating grades, easing grades, and deflating coursework, an instructor games the system and, thus, is more likely to receive positive evaluations.

Consequently, students often use the threat of giving their professors low SET scores and will complain about the professor's teaching effectiveness in an attempt to intimidate the professor to accept late assignments, sloppy work, and all forms of excuses and laziness (Shapiro 2002).

If SET scores are gamed by instructors, then as Krautmann and Sander (1999) assert, they are a flawed instrument for teacher evaluation purposes.

Gamesmanship in Higher Education

The extensive use of SET scores for administrative control purposes to evaluate teaching effectiveness (Crumbley and Fliedner 2002; Davies et al. 2007) has adverse consequences for student learning, mainly coursework deflation (Bok 2003; Cochran 1992; Wallace and Wallace 1998) and grade inflation (Eiszler 2002; Germain and Scandura 2005; Pressman 2007; Wallace and Wallace 1998; Zangenehzadeh 1988). Particularly, untenured faculty members will pander to students (Zimmerman 2002). Lynn Sherr (2010), a Columbia University professor, asserts that "over 70 percent of college teachers—even at top schools like Yale, Harvard, and Stanford—are graduate students or adjuncts or gypsy visiting professors." Wallace and Wallace (1998) argue that the use of SETs for summative purposes accounts for much of the coursework deflation and grade inflation that has occurred in U.S. college and university education. Martin (1998, 1087) believes that "using student opinion surveys as the basis for evaluating teaching effectiveness causes the very problems that a good system's design would prevent (i.e., a decline in the quality of teaching.)"

The decline in the rigor of many college classes is not a recent event (Zimmerman 2002) and it is not confined to the United States (Dumbing; Jacobs and Van der Ploeg 2006). Simon (1996) observes several instances where college rigor has declined. First, while an

increasing number of young Americans find it important to earn a college degree, “more and more colleges are offering courses that the great majority of Americans can only view as devoid of genuine intellectual content.” Second, many colleges no longer have a core curriculum built around English, literature, the history of Western civilization, foreign languages, science, and mathematics. Third, the number of required courses in liberal arts has been reduced and the foreign language requirement has been eliminated even at elite U.S. colleges and universities. Fourth, in 1964 the average school year was 191 classroom days and in 1993 the average school year had dropped to 156 classroom days. Lastly, in 1964 79% of colleges and universities had classes on Saturdays while in 1993 only 6% did so.

In a recent study involving law school professors who teach legal writing, Fischer (2007) found that 25% of the respondents stated that they had refrained from doing something that they thought was pedagogically sound because it might negatively affect their student ratings, and 31% stated that student ratings had contributed to a lessening of rigor in law school classes. The critical Spellings Report (2006) states that “over the past decade, literacy among college graduates has actually declined. Unacceptable numbers of college graduates enter the workforce without the skills employers say they need in an economy in which, as the truism holds correctly, knowledge matters more than ever.”

The atmosphere in higher education today has been illustrated by Peter Sacks in his 1996 book titled *Generation X Goes to College: An Eye-Opening Account of Teaching in Postmodern American*. Sacks, an unemployed reporter who accepted a job as a college instructor, was so concerned with what he saw in the teaching profession, that he wrote an investigative account about it (Sacks 1996, ix-x, 3). He finds that neither administrators, faculty, nor students seem to care about learning or academic standards. Administrators want students to be satisfied, and students want good grades with little effort. They blame the faculty if they do not get good grades.

Sacks’s chairperson suggests he take an acting course to improve his class-satisfying ability. His peers advised him to teach to the evaluation instrument. He engaged in a so-called Sandbox Experiment by lowering standards and expectations and learning to entertain. He raised his evaluations and survived. Sacks outlined these beliefs and attitudes of today’s college students:

- They are consumers purchasing education and grades as a commodity.
- Hard work outside the classroom should not be a requirement.
- Teachers should be entertainers.
- They are disrespectful and often contemptuous of their teachers.

The situation may be worse today, based upon Bale and Dudney (2000), who say that Generation-X students should be taught as one would a child and not as an adult, consequently why should they be allowed to evaluate professors? Thus, a persuasive case can be made that the increased use of SET has caused higher education to become dysfunctional, resulting in a steep, slippery slide in the output quality of student learning. (Winsor 1977; Renner 1981). Fellingham (2007) worries that nonacademic measures, such as *student evaluations* and school rankings, are increasingly being used to decide who teaches and what they teach. Hill (2010) states that SETs have destroyed higher education and should be outlawed for promotion, tenure, and merit pay purposes.

In discussing the negative consequences of using SETs for summative purposes, Wallace and Wallace (1998, 446) conclude “Is it not ironic that institutions of higher learning that profess to have a core product of education have focused on a calibration tool [SETs] that is at virtual odds with that product?”

Other researchers have found that an instructor's grading policy and course difficulty can be significant factors in determining an instructor's evaluations. Nichols and Soper (1972, 1071–72) found that an increase of 1 point (on a 4.0 scale) in the mean expected grade will result in an increase of 0.53 in the mean rating of a professor. As an instructor inflates grades, he or she will be much more likely to receive positive evaluations (Worthington and Wong 1979). Trying to facilitate high achievement in students may result in low rating and vice versa (Sullivan and Skanes 1974). Snyder and Clair (1976) found that students who received higher grades than they expected tended to give very positive evaluations (a happy birthday effect). McPherson (2006) found that in some situations professors can “buy” better SET scores by inflating grades.

Current SET teaching evaluation procedures have not been changed to reflect important differences in attitude and ability between various generations such as Generation X and Generation Y (Johnson and Johnson 2010) and between student learning styles (Vermunt 2005). The process of having *student opinions* dictate how professors and administrators act is contributing to faculty burnout at all levels (Stripling 2010).

Even if this negative SET research is wrong, many professors believe that easy grading improves SET scores, so they grade inflate and deflate their coursework. Defenders of SETs argue about the statistically significant relationship between leniency and grades, but there is never an inverse relationship. Further, Weinberg et al. (2009) found a strong relationship to grades, but a weak relationship between learning and SETs, because students are unaware of how much they have learned in a course. They state that SETs “are affected by grade leniency and do not reflect learning produced in a course (p. 254).”

Student Perceptions

Most SET studies address the quantitative aspects, such as validity and reliability, and ignore such behavioral aspects as impression management, self-presentation, lying, and dysfunctional behavior. Also, the placement of such authority in the hands of students may be problematic for other reasons, but there is a paucity of research from the students' perception. Dwinell and Higbee (1993) surveyed over 150 students to gather opinions regarding the value of teaching evaluations. They observed that 67 percent of the respondents believed that rating forms were an effective means of evaluating instructors, and 83 percent believed that instructors changed their behavior as a result of weaknesses identified by the evaluations. However, the respondents did not believe that their evaluations affected faculty salaries, or promotion, and tenure decisions. Dwinell and Higbee (1993) concluded that students may not understand the importance of evaluations, because the instructions do not specify how evaluation results will be used. Thus, an important question is what variables do students use to evaluate professors?

Crumbley et al. (2001) administered questionnaires to 530 accounting students over two years at a large SET-driven public, southwestern university. The students were asked to rate the importance of 18 factors on instructor rating on a five-point Likert-type scale, and the degree of agreement/ disagreement with 17 statements related to instructor traits and behaviors that would lower overall teaching evaluation scores. Their sample was composed primarily of seniors and graduate students. The study found that the grading policy of an instructor was extremely important.

From a different perspective, many of the valid teaching techniques used to help educate students may cause students to punish professors. These teaching techniques will cause students to punish professors, as shown below in Table 1, from the Crumbley et al. (2001) study.

Table 1 Techniques that cause students to punish professors, but help students learn

Technique	Percentage impacted (%)	Ranked by mean
Asking embarrassing questions	41.9	2
Grades hard	28.2	4
Pop quizzes	27.1	6
Significant homework	20.0	5
Uses humor	14.5	13
Un-typed overheads	14.3	9
Calling on students	11.2	10
Grading homework	9.0	11

Crumbly et al. (2001); note that all techniques are not listed above.

Impression Management and Earnings Management

Impression management is a conscious or unconscious process in which individuals (professors) attempt to improve the perceptions of other people (students) about themselves. Thus, professors use and conceal information in order to increase their SET scores just like executives use earnings management (EM) to improve their financial statements.

Any evaluation system based upon student evaluations will cause instructors to present themselves in a “defensive” manner to control the impression students have of them (Rosenfeld et al. 1994). Most instructors desire to be viewed in a favorable manner by others in order to maximize their reward (Schlenker 1980), maintain their self-esteem (Schneider 1969), and create a desired self-identity (Swann 1987). In our litigious age, there is always the possibility of lawsuits by students related to grading (Schweitzer 1992).

Crumbly (1995) coined the term pander pollution (PP) behavior, defined as the purposeful intervention by a professor inside and outside the classroom with the intention of increasing SET scores and which is counterproductive to the learning process. Two examples include allowing students to copy from other students in an in-class exam and providing the exam ahead of time in multiple-section classes (Jones and Sprakman 2010). Another example of PP behavior is Duke University’s Cathy Davidson plan to get out of the grading business by turning over the grading to the students (Jaschik 2010).

Semester after semester, accounting and finance professors churn out many articles on earnings management. Executives caught engaging in abusive earnings management are fined and even sent to prison (Fuller and Jensen 2010). A new profession of forensic accountants is growing to catch these unethical executives and fraudsters. Yet, semester after semester, professors are engaging in abusive SET management, resulting in grade inflation and coursework deflation. These same professors are rewarded, given tenure, and promoted. Why is earnings management illegal and unethical, and yet SET management by professors is not condemned by academics? Forensic accountants should turn their talents and research toward SET management.

Moore (2009, 19) lists 35 collegiate corruptions in U.S. colleges and universities. He indicates that academic freedom gives professors the “right to loot the public’s resources: to dumb down classes, to inflate grades....” He describes course infantilizers, grade inflation, gratuity grades, law of reciprocal grading, and others. Moore (2009, 25) states that “instructors, and especially part-timers, adjuncts, and even tenured and non-tenured full-time faculty will think twice about giving low grades to any student.” According to Sherr (2010), undergraduate

teaching is often an afterthought, and most of the educational money “is going not into better teaching but rather into institutional growth, empire building and full-time pay and perks.”

Ethical Impression Management vs. Pander Pollution

There are, of course, ethical behaviors or techniques by professors to improve SET scores. Few people would argue with these sound practices:

- Fast turnaround time on homework and exams.
- On time or early for classes.
- Appropriate dress, make-up, hairstyle, and other visual appearance traits.
- Well-prepared and organized for each class.
- Be economic with the truth and do not tell lies.
- General behavior and manner (i.e., pleasant).
- Answer questions from students correctly and politely.
- Manage body language in order to conceal anxieties.
- Show openness.
- Downplay negative factors that may make you look unprofessional.
- Be personable and charismatic.
- Treat students with respect.

Just as there is abusive (illegal) earnings management (sometimes called non-discretionary and discretionary when dealing with total accruals), there are unethical techniques or behaviors (pander pollution) in and outside the classroom. Sometimes the dividing line depends upon the over-all facts or materiality.

Fuller and Jensen (2010) provide an excellent description of the earnings management process:

When earnings appear to be coming in short of projections, top managers often react by suggesting or demanding that middle and lower level managers redo their forecasts, plans, and budgets. In some cases, top executives simply acquiesce to increasingly unrealistic analyst forecasts and adopt them as the basis for setting organizational goals and developing internal budgets. But in cases where external expectations are impossible to meet, either approach sets up the firm and its managers for failure and in the process value is destroyed.

There is a real parallel between earnings management and SET management.

With so much riding on the SET scores for so many professors, caused by extensive use of summative SETs, it is not surprising that there is so much unethical behavior of professors. For example, consider a west coast professor who hands out brownies on the day he/she gives out the SET questionnaires. Is this ethical behavior or pander pollution? Yet there is a paucity of research (even in ethics' literature) on unethical SET management; although there are thousands of articles on earnings management?

Crumbley and Smith (2009) provides a list of 28 probable unethical behaviors of professors. The top ten are: 1) first and foremost, inflate your student's grades; 2) reduce the course material covered and drop the most difficult material first; 3) give easy examinations (e.g., true-false; broad, open-ended discussion questions; take home exams; and open book exams); 4) join the college party environment by giving classroom parties on SET day; sponsor student's officially-approved class skipping days to ball games, etc., as a means to increase student satisfaction; 5) give financial rewards such as establishing connections to

potential employers; 6) spoon-feed watered-down material to the students; 7) give answers to exam questions beforehand—either pass them out in class or if you want the students to work harder, put them on reserve in the library or on the Internet; 8) do not risk embarrassing students by calling on them in the classroom; 9) hand out sample exams, or take your examination questions from the student’s online exercises provided by the textbook publisher; and 10) grade on a wide curve.

Just like earnings management, there is a continuum from ethical SET management that encourages student learning to unethical pander pollution that only is intended to raise SET scores. Major organizations such as the American Accounting Association, American Management Association, American Economic Association, as well as academic researchers, should take a position on this issue.

Some techniques may be borderline between ethical versus unethical. For example, Olds and Crumbley (2003) found that “giving six examinations rather than three allows a student to memorize less of the course material before each examination, resulting in higher overall grades.” Of course, here higher grades did result in higher evaluations and possibly more student learning.

Ian Neath (1996) and Paul Trout (1997) provide other techniques for professors to increase SET scores without increasing learning. We list them below in Table 2, and rate them as follows: EB (ethical behavior), BL (borderline), and PP (pander pollution). Make your own judgment.

Administrators, colleges, and universities engage in pander pollution on a macro level by using incentive PP (IPP) and structural PP (SPP). Moore (2009, 22) calls these types of

Table 2 Techniques for increasing SET scores without increasing learning

Technique	Rating
1. Use confident words with a confident delivery.	EB
2. Dress casually like students.	EB
3. Demonstrate with stories that you are warm and nurturing.	EB
4. Remember that first day impressions are important.	EB
5. Smile, use gestures, be relaxed, move among students, and look them in the eyes.	EB
6. Closely review your evaluations because there can be mistakes that you need to identify.	EB
7. Administer tests after evaluation, not before.	EB
8. Suck up to your students (this is more intense than impression management).	BL
9. Inform students of their potential “high” grades before administering the questionnaires.	BL
10. Play games, sing songs, tell jokes, etc.	BL
11. Prevent poor students from taking your course.	BL
12. Show lots of films, perform demonstrations, and use the latest technology. In other words, cover less material and make exams easier.	PP
13. Entertain the students with as little content as possible. Play music. Play games. Encourage a great deal of class discussion.	PP
14. Teach the course less rigorously than the normal expectations of the students.	PP
15. Be like your students. Be more liberal at Berkeley, but more conservative at Baylor. Teach what they want, and how they want it.	PP
16. Do not leave the classroom while administering your evaluations.	PP
17. Give plenty of high grades.	PP
18. Teach less demanding classes.	PP

Ratings: *EB* Ethical, *BL* Borderline, *PP* Pander Pollution

corruption techniques “Economics of Grade Inflation.” He explains that inflating grades earns scarce resources:

- Making money by the educational unit because of a high retention rate. (A university can move into the elite status with a 98 to 99 percent retention rate (learning be damned).
- Saving time for faculty since they avoid wasting time on coursework, complaints from students, office hours, etc.
- High grades allow faculty to stay in favor with many administrators.

An excellent example of IPP is Michael McKinney’s new dysfunctional performance system in the Texas A&M System. The Chancellor instituted a bonus system giving the top 15% of professors up to \$10,000 based solely on student evaluation ratings; certainly, a fine system to increase retention rates and destroy student learning. Incidentally, in 2009, the Texas A&M faculty senate passed a resolution of “no confidence” in McKinney.

Other IPP would be at universities who give specific SET guidelines, for example one private university (which puts its primary emphasis on teaching and evaluates 65% of the faculty member’s overall teaching performance based on SETs) provides specific guideline scores on the SET data collected from anonymous student questionnaires, presented below in Table 3.

Since Question 19 is the typical overall teaching effectiveness, a professor has an incentive to engage in pander pollution.

Universities and colleges are beginning to introduce Structural PP (SPP). For example, students at Loyola Law School Los Angeles will have 0.333 tacked on every grade recorded in the last few years to make them more competitive in the job market. New York University, Georgetown, Golden Gate University, Tulane University, and six other law schools have deliberately made their grading systems more lenient. Even worse, Harvard, Stanford, Yale, and University of California—Berkeley law schools have eliminated traditional grading altogether, moving to a pass/fail system (Rampell 2010).

Rampell (2010) correctly asserts the following about these SPP moves:

These moves can create a vicious cycle like that seen in chief executive pay: if every school in the bottom half of the distribution raises its marks to enter the top half of the distribution, or even just to become average, the average creeps up. This puts pressure on schools to keep raising their grades further.

Conclusion

Gaming or dysfunctional behavior, spawned by the use of SETs for personnel decisions, must be addressed in order to maintain credibility in the grading system, to better align

Table 3 One private university’s SET guideline scores

	For Q1-15	For Q-19
1. Outstanding	Below 1.75	Below 1.5
2. Very Good	Below 2.25	Below 2.0
3. Successful	Below 2.75	Below 2.5
4. Needs Improvement	Between 2.75–3.25	Between 2.5–3.0
5. Unacceptable	Above 3.25	Above 3.0

performance with rewards, and to safeguard the integrity of the learning process. Common sense suggests that accounting faculty, accounting students, and all other stakeholders in higher education deserve a control/evaluation system that promotes, rewards, and values appropriate and good behavior (such as forcing students to work hard, prepare for class, and take some responsibility for their own learning). Clearly, research suggests that as the SET has “evolved” from a formative to a widely used summative teaching evaluation tool, precisely the opposite often occurs (as predicted by Martin 1998 and Wallace and Wallace 1998). Accounting faculty are often punished (certainly not rewarded) for doing the right things in class and encouraging the appropriate and right student attitudes and behaviors, and, therefore, student learning suffers. Corporate executives are put in prison for abusive earnings management; yet professors are given pay increases, promoted, and rewarded for abusive SET management.

Faculty should not be immune from punishment for unethical teaching techniques they use to “cook” their SET scores. More research is needed to determine the difference between ethical and unethical impression management techniques used today by many instructors to increase their SET scores at the expense of learning. Research should return some focus to the use of student feedback on a professor’s teaching effectiveness for formative and not summative purposes (Yorke 2003). Administrators must take appropriate action against professors who inflate grades and dumb-down their course coverage to manage their SET scores.

SET scores should not be used solely by administrators for tenure, promotion, and merit pay. Someone criticized John Maynard Keynes for reversing himself on a particular issue. He responded, “when the facts change, I change my mind. What do you do, sir?” (Malabre, Jr. 1994). Our dysfunctional academic control system is not working, so we need to change it.

References

- Aigner, D. J., & Thum, F. D. (1986). On student evaluations of teaching ability. *Journal of Economic Education*, 17(4), 243–266.
- Adams, J. V. (1997). Student evaluations: the rating game. *Inquiry*, 1(2), 10–16.
- Bale, J., & Dudley, D. (2000). Teaching generation X: do andragogical learning principles apply to undergraduate finance education? *Financial Practice & Education*, 10(1), 216–227.
- Bok, D. (2003). *Universities in the marketplace: The commercialization of higher education*. Princeton: Princeton University Press.
- Brown, D. L. (1976). Faculty ratings and student grades: a university-wide multiple regression analysis. *Journal of Educational Psychology*, 68, 573–578.
- Cochran, L. (1992). *Publish or perish: The wrong issue. Agendas, alternatives, applications, actions*. Cape Girardeau: StepUp Inc.
- Crumbley, D. L. (1995). The Dysfunctional Atmosphere of Higher Education: Games Professors Play. *Accounting Perspectives*, 1(1), 67–77.
- Crumbley, D. L., & Fliedner, E. (2002). Accounting administrators’ perceptions of student evaluation of teaching (SET) information. *Quality Assurance in Education*, 10(4), 213–222.
- Crumbley, D. L., & Reichelt, K. (2009). Teaching effectiveness, impression management, and dysfunctional behavior. *Quality Assurance in Education*, 17(4), 377–392.
- Crumbley, D. L., & Smith, S. (2009). The games professors play in the dysfunctional performance evaluation system used in higher education: Brainstorming some recommendations. In: *Measuring learning rather than satisfaction in higher education* (pp. 41–58). Sarasota, FL: American Accounting Association.
- Crumbley, D. L., Henry, B. K., & Kratchman, S. (2001). Students’ perceptions of the evaluation of college teaching. *Quality Assurance in Education*, 9(4), 197–207.
- Davies, M., Hirshberg, J., Lye, J., Johnson, C., & McDonald, I. (2007). Systematic influences on teaching evaluations: the case for caution. *Australian Economic Papers*, 46(1), 18–38.

- Ditts, D. A. (1983). A statistical interpretation of student evaluation feedback. *Journal of Economic Education* (Spring), 10–15.
- Dowell, D. A., & Neal, J. A. (1983). The validity and accuracy of student ratings of instructors: a reply to Peter A. Cohen. *Journal of Higher Education* (July/August), 459–463.
- DuCette, J., & Kenney, J. (1982). Do grading standards affect student evaluations of teaching? Some new evidence on an old question. *Journal of Education Psychology*, 308–314.
- Dumbing. Available at: http://nomuzak.co.uk/dumbinmg_down.html.
- Dwinell, P. L., & Higbee, J. L. (1993). Students' perceptions of the value of teaching evaluations. *Perceptual and Motor Skills*, 76, 995–1000.
- Eiszler, C. F. (2002). College students' evaluation of teaching and grade inflation. *Research in Higher Education*, 43(4), 483–501.
- Fellingham, J. (2007). Is accounting an academic discipline? *Accounting Horizons*, 21(2), 159–163.
- Fischer, J. (2007). Implications of recent research on student evaluations of teaching. University of Louisville. Law School. Legal Studies Research Paper No. 2007-20. Available at: <http://ssrn.com/abstract=1005681>.
- Flinn, D. E., & Crumbley, D. L. (2009). *Measuring learning rather than satisfaction in higher education* (p. 222). Sarasota: American Accounting Association.
- Fuller, J., & Jensen, M. C. (2010). Just say no to Wall Street: putting a stop to the earnings game. *Journal of Applied Corporate Finance*, 22(1), 59–63.
- Germain, M. L., & Scandura, T. A. (2005). Grade inflation and student individual differences as systemic bias in faculty evaluations. *Journal of Instructional Psychology*, 32(1), 58–67.
- Haskell, R. E. (1997). Academic freedom, tenure, and student evaluations of faculty. *Education Policy Analysis Archives*, 5(16), 1–36.
- Hill, J. (2010). Rating your professors: scholars test improved course evaluations. *Chronicle of Higher Education*, April 25, 2010.
- Howard, G. S., & Maxwell, S. E. (1982). Do grades contaminate student evaluations of instruction? *Research in Higher Education*, 175–188.
- Isely, P., & Singh, H. (2005). Do higher grades lead to favorable student evaluations? *Journal of Economic Education*, 36(1), 29–42.
- Jacobs, B., & Van der Ploeg, F. (2006). Guide to reform of higher education: a european perspective. *Economic Policy*, 21(47), 535–592.
- Jaschik, S. (2010). No grading, more learning. *Insider Higher Ed* (May 3). Available from: <http://www.insidehighered.com/news/2010/05/03/grading>.
- Johnson, M., & Johnson, L. (2010). *Generations, Inc.: From boomers to linksters—managing the friction between generations at work*. New York: AMACOM.
- Jones, J. C., & Sprakman, G. (2010). *A case of academic misconduct: Does self-interest rule?* Working paper, January 2010.
- Kau, J. B., & Rubin, P. H. (1976). Measurement techniques, grades and ratings of instructors. *Journal of Economic Education*, 8(1), 59–62.
- Krautmann, A. C., & Sander, W. (1999). Grades and student evaluations of teachers. *Economics of Education Review*, 18(1), 59–63.
- Malabre, A. L., Jr. (1994). *Lost prophets: An insider's history of the modern economists* (p. 220). Boston: Harvard Business Press.
- Martin, G. (1998). Evaluating faculty based on student opinions: problems, implications and recommendations from deming's theory of management perspective. *Issues in Accounting Education*, 13(4), 1079–1094.
- McPherson, M. (2006). Determinants of how students evaluate teachers. *Journal of Economic Education*, 37(1), 3–20.
- Mehdizadeh, M. (1990). Loglinear models and student course evaluations. *Journal of Economic Education*, 21(1), 7–21.
- Moore, P. (2009). Why we should measure learning: A glossary of collegiate corruption. In: *Measuring learning rather than satisfaction in higher education* (pp. 17–28). Sarasota: American Accounting Association.
- Neath, I. (1996). How to improve your teaching evaluations without improving your teaching. *Psychological Reports*, 78, 1363–1372.
- Newton, J. D. (1988). Using student evaluation of teaching in administrative control: the validity problem. *Journal of Accounting Education*, 6(1), 1–14.
- Nichols, A., & Soper, J. C. (1972). Economic man in the classroom. *Journal of Political Economy*, 80, 1069–1073.

- Olds, P. R., & Crumbley, D. L. (2003). Higher grades = higher evaluations: impression management of students. *Quality Assurance in Education*, 11(3), 172–177.
- Porcano, T. M. (1984). An empirical analysis of some factors affecting student performance. *Journal of Accounting Education* (Fall), 111–126.
- Powell, R. W. (1977). Grades, learning, and student evaluation of instruction. *Research in Higher Education*, 193–205.
- Pressman, S. (2007). The economics of grade inflation. *Challenge*, 50(5), 93–102.
- Rampell, C. (2010). In law schools, grades go up, just like that. *New York Times*, June 21.
- Renner, R. R. (1981). Comparing professors: how student ratings contribute to the decline in quality of higher education. *Phi Delta Kappa* (October), 128–131.
- Rosenfeld, P. R., Giacalone, A., & Riordan, C. A. (1994). Impression management theory and diversity. *American Behavioral Scientist*, 12(4), 601–620.
- Sacks, P. (1996). *Generation X goes to college: An eye-opening account of teaching in postmodern America* (p. 208). Chicago: Open Court.
- Saje, N. (2005). Teaching for tips. *Liberal Education*, 91(1), 48–50.
- Schlenker, B. R. (1980). *Impression management: The self-concept, social identity, and interpersonal relations*. Monterey: Brooks/Cole.
- Schneider, D. J. (1969). Tactical self-preservation after success and failure. *Journal of Personality and Social Psychology*, 13, 262–268.
- Schweitzer, T. A. (1992). Academic challenge' cases: should judicial review extend to academic evaluations of students? *The American University Law Review*, 41, 267–367.
- Shapiro, M. (2002). *Irreverent commentary on the state of education in America Today*. (July 15). Available at: <http://irascibleprofessor.com/comments-07-15-02.htm>.
- Sherr, L. (2010). Is college worth the cash? *More*, 13(6), 63–64.
- Simon, W. E. (1996). The dumbing down of higher education. *The Wall Street Journal* (March 19, 1996), A18.
- Snyder, C. R., & Clair, M. (1976). Effects of expected and obtained grades on teacher evaluation and attribution of performance. *Journal of Educational Psychology*, 68(1), 75–82.
- Spellings, M. (2006). A test of leadership: Charting the future of U.S. higher education, *U.S. Dept. of Education*.
- Stripling, J. (2010). Burning out, and fading away. *Inside Higher Ed* (June 20).
- Stumpf, S. A., & Freedman, R. D. (1979). Expected grade co-variation with student ratings of instruction: individual vs. class effects. *Journal of Education Psychology*, 273–302.
- Sullivan, A. M., & Skanes, G. R. (1974). Validity of student evaluation of teaching and the characteristics of successful instructors. *Journal of Educational Psychology*, 66(4), 584–590.
- Swann, W. B. (1987). Identity negotiation: where two roads meet. *Journal of Personality and Social Psychology*, 52, 1038–1051.
- Trout, P. (1997). How to improve your teaching evaluation scores without improving teaching. Available at: <http://mtprof.msun.edu/Fall1997/HOWTORAI.html>.
- Vermunt, J. B. (2005). Relations between student learning patterns and personal and contextual factors and academic performance. *Higher Education*, 49, 205–244.
- Wallace, J., & Wallace, W. (1998). Why the costs of student evaluations have long exceeded their value. *Issues in Accounting Education*, 13(2), 443–447.
- Weinberg, B. A., Hashimoto, M., & Fleisher, B. M. (2009). Evaluating teaching in higher education. *Research in Economic Education*, 40(3), 227–261.
- Winsor, J. L. (1977). A's, B's, but not C's?: A comment. *Contemporary Education* (Winter), 82–84.
- Worthington, A. G., & Wong, P. T. P. (1979). Effects of earned and assigned grades on student evaluations of an instructor. *Journal of Educational Psychology*, pp. 764–775.
- Wright, P., Whittington, R., & Whittenburg, G. E. (1984). Student ratings of teaching effectiveness: what the research reveals. *Journal of Accounting Education* (Fall), 5–30.
- Yorke, M. (2003). Formative assessment in higher education: moves toward theory and the enhancement of pedagogic practice. *Higher Education*, 45, 477–501.
- Zangenehzadeh, H. (1988). Grade inflation: a way out. *Journal of Economic Education*, 19(3), 217–226.
- Zimmerman, A. (2002). Reduced rigor and grade inflation diminish the quality of higher education. *NACTA Journal* (December). Available at: http://findarticles.com/p/articles/mi_qu4062/is_200212/ai_n9151533/print.