The Art of Dubious Student Evaluations

Rubina Vohra, PhD
New Jersey City University
Department of Economics
2039 Kennedy Blvd., Jersey City, NJ 07305
United States of America

Abstract

This study analyzes if the outcomes of the student evaluation survey conducted by the Faculty of Economics were influenced by factors which are not directly related to teaching effectiveness. This paper shows that expected grades, course difficulty and course workload have inappropriately influenced students' judgements on evaluations of course instructor. Our research model differ from others by using faculty evaluation forms twice during the semester – first before the midterm and then before the final exam – to analyze the bias in ratings. Our findings indicate that students who evaluate their professors' teaching quality are significantly affected by how easy the course is and how much workload is required in the course. It should be totally inappropriate to consider using data derived from student evaluations to supplement decisions with regard to faculty tenure, promotion, and merit procedures.

Key Words: Student evaluations; Teaching effectiveness; Overall quality of teaching.
JEL A10

Introduction

Over a long period of time, several studies have focused on student evaluations of teaching instruments utilized by Higher Education institutions (Centra, 2003; Felton et al, 2004; Harrison et al, 2004; Hobson et al, 2001; Liaw et al, 2003; McKeachie et al, 1971; Otto et al, 2005; Read et al, 2001; Tatro, 1995; and Wachtel, 1998).

The reliability, consistency, and rationality of student evaluation survey (SES) as a gauge of teaching performance has been the focus of much research. Increasingly, universities and colleges are relying more on scores from SESs to demonstrate verification of an instructor's success in teaching effectiveness. It has become normal practice to give out SES at the end of every semester. These surveys are expected to measure professors' teaching effectiveness and achievement in their classes.

SES scores can influence the jobs of thousands of college/university faculty in the USA. According to Yining and Hoshower (2003), SES results are used by the university as the most significant aspect in assessing teaching performance and students seldom recognize that these surveys operate in this manner or bear so much importance. McGhee et al (2003) and Hardy (2003) noted that the nature of student ratings were similar for both paper and online surveys. Others argued a somewhat lower response rate for online as opposed to paper surveys. Ballantyne (2003) and Johnson (2003) suggested that to reduce this response disparity it is necessary that both students and faculty accept online ratings and approve the advantages of online ratings with helpful criticism.

Colleges and universities have adopted standardized evaluation forms, such as the IDEA (Instructional Development and Effectiveness Assessment), the CIEQ (Course Instructor Evaluation Questionnaire), or the SIR (Student Instructional Report) as means of teaching evaluation. Although standardized instruments are used to award tenure, promotion and merit-based raises, their chief function is to offer the instructor with informational data that could be employed to enhance teaching style. This added enhanced use of SES's results suggests that a better awareness of factors influencing the outcomes is more than necessary.

In the past, research-oriented colleges and universities have placed less importance on teaching than on research to assess faculty for merit raises, promotion, and tenure.

However, business schools have started to reexamine the role of teaching in faculty evaluations in response to new accreditation rules and guidelines from the AACSB (American Assembly of Collegiate Schools of Business) and pressure from other factions such as state administrators responsible for their funding. Moreover, many business school instructors and administrators, as well as, corporate experts also have advocated that more attention be paid to teaching in faculty evaluation (Wallace, 1990; Weis, 1990).

Critics argue that some college instructors use an easy approach to obtain superior student endorsement by giving better grades and limiting course work in the belief that this will result in greater student approval. For some, a challenging instruction, lack of grading leniency, and a heavy course workload would likely negatively impact students' evaluation (Marsh 1987). Given the relevancy that colleges place on student course evaluations for merit-based raises, tenure, and promotion decisions, a professor's impulse to sway grades or course workload is an undesirable outcome.

Some argued that the attributes of both the course and instructor such as class size, level of instruction, the type of subject, and the time a course is taught have an undue effect on the outcome of SES. Gender and instructional proficiency of the instructor are other factors to consider.

Our university, like most others, evaluates teaching performance to award faculty promotion, tenure, or merit-based raises. This growing importance and significance of student evaluation in tenure, promotion and merit decisions warrants thoughtful consideration from all faculty members and college administrators.

We agree with most leading scholars in this field that student learning is the most compelling rationale for the evaluation of faculty performance in classes. Further, it is important for faculty to encourage and foster student learning than to influence student approval and satisfaction. Student learning should be the ultimate outcome for faculty to accomplish when teaching a class.

This study analyzes if the outcomes of the SES conducted by the Faculty of Economics at our university were influenced by factors not directly related to teaching effectiveness. Although student evaluations were introduced to acquire feedback from students for course and teaching improvement, they are undoubtedly used as a measure of teaching effectiveness at our institution. This study aims to present that such dubious practice can lead to exploitation of the SES instrument given the presence of biasing factors such as easy grading, course organization, course material covered, and difficulty level.

The main question of our research is when students assess the overall teaching quality of their professors, are they impartial, or are they biased by such extraneous factors as expected grades, course workload, course difficulty, and course organization? Our research model differ from others by using faculty evaluation forms twice during the semester – first before the mid-term and then before the final exam - to analyze the bias in ratings

The remainder of the paper is organized as follows: Section two reviews some past studies to address factors influencing class evaluation scores. The third section explains the data and methodology employed. The findings are reported in section four. The final section summarizes the results and discusses the implications.

Literature Review

Student evaluations of teaching continue to be an important and frequently controversial research area. Numerous studies have underlined the wide dependency on SESs and multiple problems associated to their application. For example, biases have been identified for such factors as academic ability of students, the effect of leniency in grading policy, course workload and actual grades obtained, as well as, potential biases in student ratings and class size (Greenwald, 1997; McKeachie, 1997; Wallace & Wallace, 1998; Liaw & Goh, 2003).

A growing body of research on student evaluation of teaching suggests that several aspects can influence the scores that a faculty member receives (Calderon et al, 1994; Calderon et al, 1996; Calderon et al, 1997; Wallace et al 1998). Analysis shows that SES scores are affected by peripheral factors not related to teaching effectiveness. Calderon et al. (1996) and Martin (1998) research provides a wide-ranging list of unimportant factors showing that SESs can be influenced by student attributes such as: a) student rationale for taking a class (disposition toward instructor and courses, expected course grade), b) instructor characteristics (such as faculty gender, experience, personality), and c) course characteristics (such as class size, grading leniency, course difficulty).

Based on the analysis of the student evaluation survey models of eight Australian universities, Bedggood et al, (1999) argue that student evaluation of teaching can be critically biased when used as the exclusive measure of teaching performance. They content that SES is valuable as a teaching evaluation method only if the ratings impartially reveal the essence it is intended to determine. They claim that several empirical reports showed that SES reveals not only traits of instructional effectiveness, but also features that are subjectively foreign and external. Thus, showing biases in the SES outcomes and diminishing its reliability as a measure of teaching quality. Other studies show similar conclusions in their findings (Marsh et al 1981; Goldberg et al 1991; and Langbein 1994).

Critics argue that course final grades are usually not available to students at the time they submit their evaluation form and thus should not be assumed to bias course evaluations. Forms are commonly submitted during the last two weeks of a semester when students are aware of only the grade they contemplate they will receive based on their observed accomplishment thus far.

Others found that expected grades are moderately associated with student evaluations (Feldman 1997). According to them, this moderate but significant relationship between expected grades and evaluation ratings has several likely justifications: The primary explanation is the rationality argument whereby students who received high grades in a course indicate how well they have learned and, therefore, evaluate the course or instructor favorably. Another likely justification for the positive correlation is based on students' scholastic motivation or their fascination in the subject matter. Classes that appeal to strongly inspired and academically driven students should display better grades because students work harder and learn more. Those classes should get superior evaluation ratings because motivated students value the course and the instruction they have received (Howard et al 1980; Marsh, 1987). Lastly, students achieving high grades would attribute those to their hard work and intelligence, and low grades and low ratings to poor instruction (Greenwald, 1980).

Several other researchers also found similar associations between professional grading criterions and student evaluations. Among them, Stumpf et al (1979) and Drake (1984) reported that grading leniency is correlated to superior evaluations of teaching performance. For Tang (1999), the differences between expected and actual grades affect evaluations. Likewise, Cerrito (2000) reports a positive association between a student's satisfaction with the instructor and his/her expected grade. Engdahl et al (1993) note that student evaluations of their instructor plunged after the instructor clearly assessed the student's course performances before the final exam.
Wallace and Wallace (1998) indicate that SESs are a measure of students "happiness index" at the end of a semester before receiving final grades, rather than a measure of teaching quality. Both argue that SES exhibits biased results because of the propensity to reduce course workload and raise grades by the instructor.

Some suggest that properly devised SESs can be a valuable choice in assessing faculty teaching performance, but warn that SESs do not provide complete collection of information essential to evaluate all facets of effective teaching (Calderon et al. 1996).

Feldman (1984) offers a broad analysis of past research on the correlation between class size and student evaluations. Most studies report a positive relationship between a smaller class size and evaluation ratings. Cranton et al (1986) note the existence of a complex connection between class size and teaching evaluation ratings interrelating with other factors such as the course difficulty and type of course (such as major course), By contrast, some studies find only a weak or no significant relationship between the two variables (Marsh 1987; and Langbein (1994).

Langbein (1994) and Goldberg et al (1991) discover that courses taught at advanced levels often received higher SES ratings, apparently because by being more driven in their studies upper level students are also more refined in their assessments.

Students take courses that can be generally divided into two groups, namely, quantitative and non-quantitative. The broad perception is that quantitative courses are more challenging to teach and learn, and henceforth typically receive lower evaluation ratings than non-quantitative subjects. Some argue that instructional techniques differ widely between them. While Deberg et al (1990) notes that the type of courses can influence teaching evaluation, Marsh et al (1981) pointed that the influence is somewhat less significant than the effect of the instructor who teaches them. Further, Langbein (1994) also discover no significant relationship between the types of course and overall instruction ratings.

Additionally, other researchers argue that course properties and faculty traits also influence student evaluations. Basow et al (1987) testing if male and female instructors were rated differently found that female faculty are being rated more adversely than their male colleagues. Similar conclusions are reported by Langbein (1994).
The research outcomes of the latent biasing factors are somewhat mixed. Analysis appears to be contingent on the framework, substance, as well as, method of teaching evaluation. Despite this what matters is the existence of one or more of these biases, their extent, and how to address them.

Data and Methodology

To date, a large body of research has been conducted on data derived from faculty evaluation forms that are filled out anonymously by students two weeks before the end of the semester. Our research focused on a similar mode of faculty evaluation, but we used evaluation forms twice during the semester – once before the mid-term and another before the final exam - to analyze the bias in ratings. The findings derived from before midterm evaluations were utilized as a channel for a dialog with the class as a whole, and to address any difficult areas that were uncovered before the semester was over. By evaluating teaching at intervals during the semester, faculty gained feedback needed to pinpoint would-be problematic areas. They also gathered proof of effective teaching. Our argument is simple: just as we do not give our student one final exam as the only measure of their performance in a class, why then should we offer only one evaluation concerning an instructor's teaching?

The sample for this study included a total of 20 classes in which the SES was administered. We collected data for the survey both before the midterm and before the final exam. The average size of each class was 25. The data used in this study includes information collected from the summer, fall and spring semester of the 2015/2016 academic session for the Bachelor of Economics program. In the middle of the semester, before midterm, and in the last two weeks of these semesters, responses were obtained from the students using a questionnaire. The questionnaire required responses to five variables commonly associated directly to teaching that were used for analysis including: course workload, course difficulty, course organization, expected grades and overall quality of professor. Overall quality of the instructor was not a derived variable. It was one of the variables actually rated by the students. Each variable was rated on a five-point scale. The ratings ranged from 1 (low) to 5 (high). We assumed that the data were at an interval level or that the difference between a 1 and a 2 was approximately the same level of difference as between any other two ratings. This assumption rationalizes our use of correlation and regression analysis. Before the midterm, for each class and each variable, we added the ratings and divided them by the total number of student evaluations to compute an average rating. We repeated this method for the survey given before the final exam. Then we added the two averages from both surveys and took the mean value for each variable for all 20 courses.

Analysis

Information on the variables used in our analysis is displayed in Table 1, including variable mean and standard deviation. For example, average workload is 3.20 on a 1 -5 scale, which means average rating of the faculty for outside workload is higher than the middle. The standard deviation shows the variation in the variable and depicts that about 68% of ratings are between 2.19 and 4.21.

Table: 1
Descriptive Statistics

	Mean	Std. deviation	N
Overall Quality	3..50	0.95	20
Workload	3.20	1.01	20
Course Organization	4.30	0.92	20
Course Difficulty	3.80	0.77	20
Expected Grade	4.90	0.31	20

"Data Source: Own calculations." N = 20.

A Pearson correlation coefficient was calculated for the relationship between variables as presented by correlation matrix in Table 2.

Table: 2
Correlation Matrix

	Overall Quality	Workload	Course Organization	Course Difficulty	Expected Grade
Pearson Correlation					
Overall Quality	1.00	.554	.060	.290	.542
Workload	0.554	1.000	-.295	.873	.408
Course Organization	0.060	-.295	1.000	-.208	.111
Course Difficulty	0.290	.873	-.208	1.000	.356
Expected Grade	0.542	.408	.111	.356	1.000
Sig. (1-tailed)					
Overall Quality		.006	.400	.108	.007
Workload	.006		.103	.000	.037
Course Organization	.400	.103		.190	.320
Course Difficulty	.108	.000	.190		.062
Expected Grade	.007	.037	.320	.062	
N=20					

"Data Source: Own calculations." N = 20.

As expected, positive correlation was found between overall quality of professor and expected grades ($r = .54$, $p<.05$). Relationship between overall quality of professor and workload and overall quality and expected grades indicate a significant linear relationship. A weak positive correlation that is not significant was discovered between overall quality and course organization and overall quality and course difficulty.

If the professor was organized and gave clear lectures, their ratings were expected positive- meaning they delivered well prepared lectures, were organized in the class presentation, and had delivered concepts employing appropriate methodology of presentation. Our analysis show if the course was better organized, it was less difficult ($r = - . 208$) and students felt workload was less demanding ($r = -.295$).
Hence, our regression model is as follows:

Overall Quality of Professor = β_0 Constant + β_1 Workload + β_2 Course Organization + β_3 Course Difficulty + β_4 Expected Grade

The model summary and ANOVA results are reported in Table 3.

Table 3:
Model Summary and ANOVA Results

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
	.788	.620	.519	.656102

	Sum of Squares	df	Mean Square	F	Sig.
Regression	10.543	4	2.636	6.123	.004
Residual	6.457	15	.430		
Total	17.000	19			

"Data Source: Own calculations." N = 20.

We obtained adjusted R square of .519. The ANOVA resulted in F = 6.123 with 4 and 15 degrees of freedom (F (4, 15) = 6.123, $p < .05$) suggesting we have a significant linear regression equation.

Table 4 reports the regression results for explaining overall teaching ratings. The main finding of this analysis is that workload, course difficulty and expected grade have a significant effect on the evaluations of teaching. The Tolerance value is more than .10, suggesting there is no multicollinearity concern. The β weights for workload (1.158), course difficulty (-1.054) and expected grade (0.988) are noteworthy. The coefficients of workload and course difficulty are significant at $p < .05$, and the coefficient of expected grade is significant at $p < .10$. In general, indicating they have a greater effect on the dependent variable (overall quality of professor).

Table 4:
The Regression Model for Explaining the Overall Teaching Ratings

	Unstandardized Coefficients				
	B	Std. Error	t	Sig.	Tolerance
Constant	-1.965	2.558	-.768	.454	
Workload	1.158	0.329	3.524	.003*	.208
Course Organization	.215	0.178	1.206	.246	.838
Course Difficulty	-1.054	0.404	-2.607	.020*	.235
Expected Grade	.988	0.556	1.777	.096**	.774

Dependent Variable: Overall Quality
Predictors: (Constant), Workload, Course Organization, Course Difficulty, Expected Grade
"Data Source: Own calculations." N = 20. *p < .05. ** p < .10.

Interestingly, course organization is not significantly associated with overall quality of professor. What these findings indicate is that professors will receive better evaluations when their courses are manageable for students. In other words, students will assess teaching as most effective when a) it is at their level of preparation and aptitude rather than too difficult, b) when the course workload is close to what other courses demand and, again, at their level of competence, and c) when the course is covered at the speed just about right for them and not necessarily at the level that the course requires. They would give better reviews only when they expect to be given, not necessarily earn, a better grade in the course.

Conclusion

Our analysis demonstrate that students' ratings of an instructor based on workload, course difficulty and expected grade strongly correlate to overall quality of the professor. Course organization was not significantly associated with overall quality of the professor. These findings are consistent with our expectation that the ratings do not reflect student learning. They do not reflect honest and true assessments of instructors but reflects students' desire to get a good grade without serious hard work. Further, our findings indicate that students who evaluate their professors' teaching quality are significantly affected by how easy the course is and how much workload is required in the course. They are willing to accept more workload outside the class if it is not difficult. However, if the course is difficult, they tend to rate instructor as less effective.

Additionally, our findings accurately portray the views of American college/university students. It is no surprise, given the importance for faculty to gain endorsement from students seeking an easy A grade that our academic standards have tumbled and grades are inflated. As long as student opinions are used for faculty teaching evaluation, colleges/universities run the risk of encouraging their faculty to succumb to their students' wishes by awarding easy grades and lowering teaching standards. It should be totally inappropriate to consider using data derived from student evaluations to supplement decisions related to faculty tenure, promotion, and merit procedures.

The learning experience comes not only from attending classes taught by instructors, but also by following course requirements such as homework, assignments, and outside readings. If asking students what they think about their courses parallels what some agencies such as Consumer Report do when they provide information on various goods and services, then SES should include questions asking students whether they fulfilled their share of course work responsibility such as readings, homework, and assignments. Better yet, survey results should include this information on each students survey form. Expecting a higher grade without making extra effort is like blaming a doctor for not healing because we did not take the medicine.

We should remember that the course instructor is the expert in his/her discipline, and the students are the learners. Students cannot absorb the entire difficulty of the subject matter and the broader context in one semester. They will only conceptualize the subject as their knowledge expands beyond the confines of the course. After the conclusion of their degree, they should be able to do a much better job of analyzing information all together in real- life context.

Since evaluations are conducted two weeks before the final exam, that gives the opportunity to the students who performed worse than their expectations to give poor reviews and skew the results significantly, particularly in small size classes. It should be evident that the overall teaching ratings acquired from student evaluation surveys are only a measure of the instructor's presentation skills and organization of the course and do not fully reveal the value of teaching or how much the student is learning, and should, therefore, be applied cautiously. Instructors who can inspire students to work hard warrant a positive evaluation. Unfortunately, student ratings do not seem to appreciate it.

References

Ballantyne, C. (2003). Online evaluations of teaching: an examination of current practice and considerations for the future, *New Directions for Teaching and Learning Winter*, (96): 103-112.

Basow, S.A. and Silberg, N.T. (1987). Student evaluations of college professors: Are female and male professors rated differently? *Journal of Educational Psychology*, 79: 308-314.

Bedggod, R.E.and Pollard, R.J. (1999). Uses and misuses of student opinion surveys in eight Australian universities, *Australian Journal of Education*, Vol. 43 No. 2, 129-41.

Caldron, T.G., Gabbin, A. and Green, B.P. (1996). Summary of promoting and evaluating effective teaching, *Journal of Accounting Education*, Vol. 14. No. 3, 367-383.

Caldron, T.G. and Green, B.P. (1997). Use of multiple information types in assessing accounting faculty teaching performance, *Journal of Accounting Education*,15(2), 221-239.

Caldron, T. G., Green, B.P. and Reider, B.P. (1994). Using real-world cases to illustrate the power of analytical procedures, *Journal of Accounting Education*, Vol. 12, No. 3: 245 – 268.

Centra, J.A. (2003). Will teachers receive higher student evaluations by giving higher grades and less course work? Research in Higher Education 44, No. 5: 495-518.

Cerrito, P. B. (2000). An examination of college-wide student teaching evaluations, *College Student Journal*, 24(2): 165-171.

Crenton, P.A. and Smith, R.A. (1986). A new look at the effect of course characteristics on student ratings of instruction. *American Educational Research Journal*, Vol. 23 No. 1, 117-128.

Deberg, C.L. and Wilson, J.R. (1990). An empirical investigation of the potential confounding variables in student evaluation of teaching. *Journal of Accounting Education*, Vol. 8 No. 1, 37-62.

Drake, J. (1984). Improving teacher performance through evaluation and supervision, paper presented at *Annual meeting of National Association of Secondary School Principals*, ED 250-782.

Engdahl, R., Keating, R. and Perrachione, J. (1993). Effects of grade feedback on student evaluation of instruction, *Journal of Management Education*, 17 (2), 174-184.

Feldman, K. A. (1997). Identifying exemplary teachers and teaching: Evidence from student ratings. In: Perry, R. P. and Smart, J. C. (eds), Effective Teaching in Higher Education: Research and Practice, *Agathon*, New York, 368-395.

Feldman, K.A. (1984). Class size and college students' evaluations of teachers and courses: a closer look. *Research in Higher Education*, Vol. 21, No. 1, 45-116.

Felton, J., Mitchell, J. and Stinson, M. (2004). Web-based student evaluations of professors: the relations between perceived quality, easiness and sexiness. *Assessment & Evaluation in Higher Education* 29, No. 1: 91-108.

Greenwald, A. G. (1980). The totalitarian ego: Fabrication and revision of personal history. *American Psychologist*, 35: 603-618.

Gillmore, G.M. and Greenwald, A.G. (1999). Using statistical adjustment to reduce bias in student ratings. *American Psychologist* 54(7): 518-519.

Goldberg, G. and Callahan, J. (1991). Objectivity of student evaluations of instructors. *Journal of Education for Business*, Vol. 66, No.6, 377-379.

Hardy, N. (2003). Online ratings: fact and fiction. *New Directions for Teaching and Learning Winter* (96): 31-38.

Harrison, P.D., Douglas, D.K. and Burdsal, C.A. (2004). The relative merits of different types of overall evaluations of teaching effectiveness. *Research in Higher Education* 45, No. 3: 311-323.

Hobson, S. M. and Talbot, D. M. (2001). Understanding student evaluations: What all faculty should know. *College Teaching* 49(1): 26-31.

Howard, G. S., and Maxwell, S. E. (1980). Correlation between student satisfaction and grades: A case of mistaken causation? *Journal of Education Psychology*, 72: 810-820.

Johnson, T. (2003). Online student ratings: will students respond? *New Directions for Teaching and Learning Winter* (96): 49-59.

Langbein. I. (1994). The validity of student evaluations of teaching, *Political Science and Politics*, Vol. 27 No.3, 545-553.

Liaw, S. and Goh. K. (2003). Evidence and control of biases in student evaluations of teaching, *International Journal of Educational Management* 17, No. 1: 37-43.

Marsh, H.W. (1987). Students' evaluation of university teaching: research findings, methodological issues, and directions for future research. *International Journal of Educational Research*, 11: 253-283.

Marsh, H. W. and Overall, J. U. (1981). The relative influence of course level, course type, and instructor on students' evaluations of college teaching, *American Educational Research Journal*, Vol. 18 No. 1: 103-112.

Marsh, H. W. and Roche, L. A. (1999). Rely upon SET research. *American Psychologist*, 54 (7): 517-518.

Martin, J.R. (1998). Evaluating faculty based on student opinions: Problems, implications, and recommendations from Deming's theory of management perspective. *Issues in Accounting Education*, November, 1079-1094.

McGhee, D., and Lowell, N. (2003). Psychometric properties of student ratings of instruction in online and on-campus courses. *New Directions for Teaching and Learning Winter* (96): 39-48.

Mckeachie, W. J., and Lin, Y. (1971). Sex differences in student response to college teachers: teacher warmth and teacher sex. *American Educational Research Journal*, 8: 221-226.

Otto, J., Sanford, D., and Wagner, W. (2005). Analysis of online student ratings of university faculty. *Journal of College Teaching and Learning* No. 6, 25-30.

Read, W., Rama, D., and Raghunandan, K. (2001). The relationship between student evaluations of teaching and faculty evaluations. *Journal of Education for Business* 76, No. 4: 189-193.

Stumpf, S.A. and Freedman, R.D. (1979). Expected grade covariation with student ratings of instruction: individual versus class effects, *Journal of Educational Psychology*, 71, 293-302.

Tang, S. (1999). Student evaluation of teachers: effects of grading at college level. *Journal of Research and Development in Education*, 32(2), 83-88.

Tatro, C. N. (1995). Gender effects on student evaluations of faculty. *Journal of Research and Development in Education* 28: 169-173.

Wachtel, H. (1998). Student evaluation of college teaching effectiveness: A brief overview. *Assessment and Evaluation in Higher Education*, 23(2): 191-211.

Wallace, W. (1990). One educator's view of how to respond to the challenges faced by higher education in business. *Issues in Accounting Education*, Fall, 302-306.

Wallace, J. J. and Wallace, W. A. (1998). Why the costs of student evaluations have long since exceeded their value. *Issues in Accounting Education*, May, 443-448.

Weis, W. (1990). What's going on in business schools? *Management Accounting*, 71(11), 49-52.

Yining, C. and Hoshower, L. (2003). Student evaluation of teaching effectiveness: an assessment of student perception and motivation, *Assessment of Evaluation in Higher Education*, 28(1), 71-88.