

# Student Mood And Teaching Evaluation Ratings

Mary C. LaForge, Clemson University

## Abstract

*When student ratings are used for summative evaluation purposes, there is a need to ensure that the information adequately reflects the accomplishments of the instructor, and is not tainted by factors outside the control of the instructor. The present study was designed to explore how student ratings of university faculty might be affected by the mood of the student at the time of the evaluation.*

## INTRODUCTION

Formal student evaluation of teaching performance is prevalent among colleges and universities all over the world. Faculty evaluations completed by students in the classroom very often provide instructors with important feedback regarding teaching effectiveness. In many instances, student ratings are a significant factor in administrative decisions regarding faculty compensation, tenure, and promotion (Neath 1996). In some cases, teaching evaluations are made available to the student body for decisions about course selection.

Given the importance of student evaluations to the careers of academic faculty, questions about the accuracy of student evaluations in measuring teaching effectiveness are not uncommon (Hobson and Talbot 2000). The validity of formal undergraduate teaching evaluations is a concern for faculty, administrators, and students. Validity refers to the extent to which student evaluations actually measure what they are intended to measure -- instructor effectiveness. Validity is difficult to establish because researchers concede that there is no universally accepted criteria for what constitutes effective teaching. Research has thus tended to compare the results of student evaluations to either a measure of student learning such as student grades (actual or expected) or to other evaluations of teacher effectiveness that are presumed to be sound such as instructor self-evaluation, peer evaluations, and alumni evaluations (Hobson and Talbot 2000).

In response to ongoing concerns about the validity of student evaluations, research on potential biases has investigated whether factors unrelated to teaching skills may explain variability in ratings. There is much evidence suggesting that student ratings are influenced by extraneous factors such as student characteristics (e.g., race, sex, age, academic major, motivation for taking a course, expected course grade, learning style, knowledge of prerequisite material, biases toward the course and the instructor, GPA etc.), instructor characteristics (e.g., race, gender, rank, experience, weight, dress, etc.), course characteristics (e.g., course difficulty, whether the class is required or an elective, etc.), and other environmental characteristics (e.g., physical attributes and the ambiance of the classroom) (Martin 1998, Read, Rama, and Raghunandan.2001). Clearly, these types of variables "appear to mediate student ratings and must be accounted for in models of teacher effectiveness and in situations where personnel decisions are based upon student ratings" (Small, Hollenbeck, and Haley 1982, p. 205).

Moods are "feeling states that are subjectively perceived by individuals" (Gardner 1985, p. 282). These feeling states associated with moods are affective states that are general and pervasive and are not directed toward anyone or anything in particular. Instead, these states

"provide the affective coloring for day-to-day events" (George and Brief 1992, p. 31). Previous studies (e.g., Isbell and Wyer 1999, Watson and Clark 1984) have indicated that individuals who experience negative mood states tend to engage in automatic (unintentional) cognitive processes in which they encode and retrieve more negative than positive information on the target person. In contrast, individuals who experience positive mood states tend to encode and retrieve more positive information on the target person. These findings have been observed in the context of performance evaluation by Sinclair (1988) who found in a laboratory setting that raters experiencing positive moods gave more lenient ratings of ratees, while raters experiencing negative moods gave more severe ratings of ratees.

The mood of the student at the time of the evaluation rating may play a role in teaching evaluations (Munz and Munz 1997). Small, Hollenback and Haley (1982) found that student emotional states were related to instructor ratings. The correlational findings of this study indicated that the more negative the emotional state (anxious, hostile, and depressed) of the student (at the end of the semester, but not at the beginning of the semester) the lower evaluation rating. The purpose of the present study was to explore mood state as a potential influence on student ratings of instructors' teaching performance. An earlier investigative study by Munz and Munz (1997) revealed that positive mood state, but not negative mood state, measured at the time of the evaluation and two weeks before the evaluation assessments correlated positively with course and instructor ratings.

Based on this literature, we hypothesized that positive mood state would correlate positively with evaluations, while negative mood state would correlate negatively. To test these hypotheses, we selected a somewhat unconventional methodology. Our approach was to ask students to rate instructors from whom they had taken a college course in the recent past. We utilized an instrument garnered from the research literature on student evaluation and teaching effectiveness that contains high-inference items that have been shown to be strong correlates of teaching effectiveness (Young and Shaw 1999). Self-reported positive affect and negative affect were assessed through the Positive and Negative Affect Schedule (PANAS) developed by Watson, Clark, and Tellegen (1988).

## **METHODOLOGY**

The present study was conducted in a state-supported, land-grant university located in the southeast with an enrollment of about 17,000. Participants in the study were either majors in the Marketing degree program or the Management degree program in the College of Business which is a major division of the University. All business students at this university must complete a common curriculum for the freshman year before being admitted into a business baccalaureate degree program. Spring semester 2002, students enrolled in Consumer Behavior (Marketing) and Advanced Statistics (Management) courses were asked to participate in the study. Although an occasional student declined to cooperate when apprised of the nature and purpose of the study, virtually all students elected to participate. A training/briefing session was held for each group of students participating in the study.

Two hundred and forty-one students rated one qualitative and one quantitative instructor of their choice. Students were asked to select an instructor of a college course that they had taken in the recent past; one who in their judgment, could be highly effective, highly ineffective, or anywhere in between; and someone whom they thought they could recall in sufficient detail so that they could respond accurately to specific rating items about that instructor and that

instructor's instructional effectiveness. The instrument contains six specific items (e.g., "The instructor communicated effectively") and two global measures. The two overall measures were used in the present study -- course value ("The course was valuable to me") and teaching effectiveness ("Compared to other college instructors I have had, I would rate this instructor as extremely effective"). All items are rated on a scale from one to nine, where one is "not at all descriptive" and nine is "very descriptive" (Young and Shaw 1999).

Students completed the Positive and Negative Affect Schedule (PANAS) (Watson, Clark, and Tellegen 1988). The PANAS is a mood adjective checklist that contains 20 adjectives, 10 assessing positive affect and 10 assessing negative affect. Respondents are asked to indicate the extent to which each word describes their feeling and emotions "at this very moment". A five-point scale with labels of very slightly or not at all (1), a little (2), moderately (3), quite a bit (4), and extremely (5) is used. The two primary dimensions of mood: Positive Affect (PA) and Negative Affect (NA) are not opposites of each other, but are two highly distinct dimensions of mood that have consistently been represented as orthogonal dimensions in factor analytic studies (Watson, Clark, and Tellegen 1988). PA reflects one's enthusiasm, activity, and alertness. High PA is a state of high energy and pleasurable engagement, whereas low PA is characterized by sadness and lethargy. Negative Affect (NA) is associated with feelings such as anger, contempt, disgust, guilt, fear, and nervousness, with low NA being a state of calmness. Examples of PA adjectives are: interested, excited, and strong. Examples of NA adjectives are: distressed, upset, and scared.

## RESULTS

Item means (standard deviations) for qualitative and quantitative instructors are shown in Table 1 along with results of t-tests. Inspection of the table reveals that qualitative versus quantitative course evaluation mean scores did not significantly differ and qualitative versus quantitative instructor evaluation mean scores did not significantly differ. The sample mean score with regard to the value of qualitative content courses was 6.34 (2.47); for quantitative content courses it was 6.24 (2.44). The sample mean score for the instructor effectiveness was 6.50 for qualitative instructors, for quantitative instructors it was 6.26 (2.29).

For analysis, we asked participants to report their gender, the gender of the course instructor, and the grade they received in the course. One hundred and thirty-six males and 105 females participated in the study. In qualitative courses, 204 reported earning an "A" or "B", 33 a "C" or less; 135 were taught by male instructors and 106 by females. In quantitative courses, 155 students reported they earned an "A" or "B", 81 a "C" or less; 130 indicated their instructor was male, 111 indicated their instructor was female. Table 2 presents a correlation matrix showing relationships between the teaching evaluation items and measured variables (positive affect, negative affect, student gender, instructor gender, and course grade) by course type.

Table 2 shows that for both qualitative and quantitative courses, neither course value nor instructor effectiveness ratings were correlated with mood measures or with student gender. Qualitative instructor teaching effectiveness ratings were positively and significantly related to instructor gender ( $r=.17, p < .05$ ). This indicates that students tended to rate male quantitative instructors higher than they rated female quantitative instructors. What is perhaps most interesting about Table 2 is the pattern revealed in the last column. In every case, strong statistically significant positive relationships can be observed between ratings and course grades. The magnitude of the correlation coefficients in quantitative courses is especially large.

## DISCUSSION

Results of our study cannot be generalized beyond the focal business students and institution and may be limited by the selective nature of students' memory. Findings of our study suggest our concern that student moods influence formal student evaluations is unwarranted. On the other hand, Martin (1998) states that student characteristics such as attitudes and feelings are "perhaps critical to the student ratings" but these are "system-related variables" and the effects are "virtually impossible to measure" (p.1084). Because of the many technical difficulties involved, our approach was to conduct an observational study in which qualitative professors as a group were assumed to be similar and quantitative professors were assumed to be similar as a group. Obviously, this is not the case.

Our study does hint that the type of course an instructor teaches is relatively less important than student grades in determining the outcome of ratings. The relationship between student evaluations and grades has received more attention in the literature than any other variable (Martin 1998). Higher grades correlate very highly with individual ratings of teaching effectiveness and grading harshly lowers evaluations (Heath 1996).

"The central issue in evaluating teachers is not whether there is a relationship between grades and ratings of effectiveness but rather interpreting that relationship. While it is clear that effective instruction, which should increase grades, can give rise to positive evaluations, lenient grading standards can also give rise to more positive evaluations. Although correlation does not imply causation, in this case the result -- higher ratings -- is the same in both cases." (p. 1365).

## CONCLUSION AND IMPLICATIONS

Student moods are uncontrollable, or only partly controllable by the instructor and potentially bias faculty and course evaluations. Even though hypotheses were not supported here, we hope our study has generated interest and encourages others to investigate further this potential influence on ratings of teaching effectiveness. Students are in a unique position to evaluate some aspects of teaching effectiveness. However, student ratings appear to measure a host of factors that may or may not relate to students' learning. To be fair to instructors, the effects of background variables on student ratings need to be thoroughly investigated and properly dealt with.

## REFERENCES

- Chandler, T. A. (1978). "The Questionable Status of Student Ratings of Faculty." *Teaching of Psychology*, 150-152.
- Cohen, P.A. (1990). "Bringing Research into Practice." in M. Theal and J. Franklin (Eds) *Student Ratings of Instruction: Issues for Improving Practice* (San Francisco, Jossey Bass), pp. 123-132.
- DeBerg, Curtis and J. R. Wilson (1990). "An Empirical Investigation of the Potential Confounding Variables in Student Evaluation of Teaching." *Journal of Accounting Education*, 8, 36-62.
- Feldman, K. A. (1978). "Course Characteristics and College Students Ratings of Their Teachers: What We Know and What We Don't." *Research in Higher Education*, 9, 199-242.

- Hobson, SM. And D. M. Talbot (2000). "Understanding Student Evaluations: What all Faculty Show Know." *College Teaching*, 48(1), 26-31.
- Kwan, Kam-Por (1999). "How Fair are Student Ratings in assessing the Teaching Performance of University Teachers?" *Assessment and Evaluation in Higher Education*, 24(2), 181-195.
- Martin, J. R. (1998). "Evaluating Faculty Based on Student Opinions." *Issues in Accounting Education*, November, 1079-1094.
- Munz, D. C. and H. E. Munz (1997). "Student Mood and Teaching Evaluations." *Journal of Social Behavior and Personality*, 12, 1, 233-242.
- Neath, Ian (1996). "How to Improve Your Teaching Evaluations Without Improving Your Teaching." *Psychological Reports*, 78, 1363-1372.
- Neumann, L. and Y. Neumann (1985). "Determinants of Students' Instructional Evaluation: a Comparison of Four Levels of Academic Areas." *Journal of Educational Research*, 78, 152-158.
- Read, William J. D. Rama, and K. Raghunandan. (2001). "The Relationship Between Student Evaluations of Teaching and Faculty Evaluations." *Journal of Education for Business*, March/April, 189-192.
- Small, A. C. A. R. Hollenbeck, and R. Haley (1982). "The Effect of Emotional State on Student Ratings of Instructors." *Teaching of Psychology*, 9, 205-209.
- Wright, Penny, R. Whittington, and G. E. Whittingburg (1984). "Student Ratings of Teaching Effectiveness: What the Research Reveals." *Journal of Accounting Education*, Fall, 2(2), 5-29.
- Wallace, J.J. and W. A. Wallace (1998). "Why the Costs of Student Evaluations Have Long Since Exceeded Their Value." *Issues in Accounting Education*, May, 443-448.
- Young, Suzanne and D. G. Shaw (1999). "Profiles of Effective College and University Teachers." *Journal of Higher Education*, 70(6), 670-680.

**Table 1**  
**Results of Paired t-tests**

	Mean (Standard Deviation)	t	df	Significance
<b>The course was valuable to me.</b>				
Qualitative	6.34 (2.47)	0.46	239	.648
Quantitative	6.24 (2.44)			
<b>Compared with other college instructors I have had, I would rate this instructor as extremely effective.</b>				
Qualitative	6.50 (2.23)	1.03	187	.303
Quantitative	6.26 (2.29)			

**Table 2**  
**Correlation Coefficients**

	NA	PA	Student Gender	Professor Gender	Course Grade
<b>The course was valuable to me.</b>					
Qualitative	-.05	.06	-.09	-.01	.27**
Quantitative	-.05	.08	.02	.10	.43**
<b>Compared with other college instructors I have had, I would rate this instructor as extremely effective.</b>					
Qualitative	-.05	.06	-.08	.07	.30**
Quantitative	-.02	-.05	-.08	.17*	.46**

**NOTES:**

\* Indicates significantly different from 0,  $p < .05$  (two-tailed).

\*\* Indicates significantly different from 0,  $p < .0001$  (two-tailed).

Gender is coded 1 for male and 0 for female.

Course Grade is coded 1 for "A" or "B" and 0 for "C", "D", or "F".