# The Effects Of Different Sections And Students' Pre-Course Interest On An Instructor's Teaching Evaluations

Obeua S. Persons, (E-mail: persons@rider.edu), Rider University

## ABSTRACT

*This study has identified two important factors, unrelated to an instructor's teaching ability, which can affect an instructor's teaching evaluations. The first factor, which has never been examined in any prior studies, is the section effect. This study finds that teaching evaluations differ significantly across sections of the same course taught by the same instructor. This section effect cannot be explained by six student-related variables. The second factor, which is students' pre-course interest measured at the beginning of a course, is found to be positively related to teaching evaluations. These findings suggest that higher-education administrators may want to consider the section effect and the students' pre-course interest when they evaluate an instructor's teaching effectiveness for promotion, tenure and merit decisions.*

## INTRODUCTION

*M*any educators believe that student ratings of teacher effectiveness are affected by a number of factors unrelated to teaching performance (Wilson, 1998). This study investigates two such factors, the section effect and the pre-course interest. It addresses two specific questions. First, could an instructor who teaches more than one section of the same course receive different ratings across sections? Second, does students' pre-course interest affect their evaluations of an instructor's teaching effectiveness? Higher-education administrators have come to rely heavily on students' evaluations to assess the teaching effectiveness of an instructor. Such assessment is especially crucial for an instructor's career during promotion and tenure decisions (Gramlich & Greenlee, 1993). Therefore, it is highly important that instructors and higher-education administrators are aware of the effects of different sections and student pre-course interest on teaching evaluations.

No study has examined the section effect. On the other hand, several studies have investigated the relationship between students' pre-course interest and teaching evaluations. Unlike earlier studies, this study measures the pre-course interest at the beginning rather than at the end of a course. Using a regression analysis, this study finds a positive relationship between teaching evaluations and students' pre-course interest, i.e., an instructor is likely to receive high ratings if students have high pre-course interest. This study also finds that an instructor could receive significantly different evaluations across different sections. This section effect cannot be explained by a student's expected grade, the beginning-of-semester grade point average (GPA), the percentage of females, and the number of accounting majors, students using tutoring services, and honor students.

These findings have direct implications for higher-education administrators regarding their assessment of an instructor's teaching effectiveness for promotion, tenure and merit decisions. For example, the administrators may want to consider students' pre-course interest when they assess an instructor's teaching effectiveness. Additionally, the administrators should not immediately downgrade the assessment if an instructor receives lower evaluations in just one section. These findings also have a direct implication for designing a better teaching evaluation which should include a student pre-course interest assessed at the beginning of a course.

## HYPOTHESIS DEVELOPMENT

No study has investigated whether teaching evaluations of an instructor could differ across sections of the same course. It is possible that an instructor may receive better evaluations in one section than the others because of the different composition of students. An instructor could receive better teaching evaluations if a section has many students with strong academic motivation. Greenwald (1996) suggests that motivated students are more serious about their learning, more appreciative of the instructor's effort, and tend to rate an instructor higher. Therefore, teaching evaluations could differ across sections depending on how motivated the majority of students in the sections are. This leads to the following hypothesis.

**H1:** *An instructor could receive different teaching evaluations across different sections of the same course.*

Several studies (e.g., Greenwald & Gillmore, 1997a, 1997b; Marsh & Roche, 1997) have documented a significantly positive relationship between students' pre-course interest and teaching evaluations. These studies, however, measured students' pre-course interest at the *end* of a course by asking students to provide their levels of interest in the subject prior to the course. Therefore, students are expected to make retrospective judgments of pre-course interest. It is not clear whether such assessment of pre-course interest measured at the end of the course represents the actual pre-course interest independent of how effective an instructor is or the end-of-course interest which is likely influenced by an instructor's teaching effectiveness. To avoid this ambiguity, this study measures pre-course interest at the beginning of the first class meeting of a course. It is expected that an instructor would receive high teaching evaluations if students have high pre-course interest.

**H2:** *An instructor's teaching evaluation is positively related to students' pre-course interest.*

## RESEARCH DESIGN

This study uses the ordinary-least-square regression analysis to test the potential effects of different sections and students' pre-course interest on teaching evaluations. Teaching evaluations were measured by the student ratings of an instructor's teaching effectiveness. These ratings were provided by students from two sections of an introductory managerial accounting at a private university. This course is required for all undergraduate business majors during their sophomore year. These two sections were taught by the same tenured professor who has been teaching this course for several years. Students provided their pre-course interest at the beginning of the course after they had been exposed to a course syllabus which described course objectives and 12 topics to be covered in the course. At the end of the course, students provided their evaluations of the instructor's teaching effectiveness. There were a total of 61 student evaluations from both sections. Below is the regression model that has seven explanatory variables, two of which are the section and the pre-course interest, and the other five are control variables. All variables in the model are obtained from teaching evaluations which are anonymous.

- RATING = $b_0$ + $b_1$SECTION + $b_2$INTEREST + $b_3$HREXAM + $b_4$EXPGRADE + $b_5$TUTOR + $b_6$HONOR + $b_7$GENDER
- RATING is a student's rating of teaching effectiveness based on a scale of 1 to 10 where 10 is the highest rating.
- SECTION is a dummy variable which takes the value of 1 for the first section and 0 for the second section. The first hypothesis about section effect will be confirmed if SECTION has a significant coefficient.
- INTEREST is a student's pre-course interest measured at the beginning of the course and based on a scale of 1 to 10 where 10 represents the highest interest. The second hypothesis will be confirmed if INTEREST has a significant and positive coefficient.
- HREXAM is the number of hours a student usually studied for an exam. HREXAM could measure how serious a student is regarding his/her learning and grade. On the other hand, a good student who regularly reviews course materials might not need as much time to study for an exam as a student who does not at all review the materials until a few days before an exam. Therefore, no expectation is placed upon the sign of HREXAM coefficient.

- EXPGRADE is the grade a student expected to get in the course. A number of studies (e.g., Griffin, 2004; Ellis et al, 2003; Krautmann & Sander, 1999; Greenwald, 1996) have documented a positive relationship between student ratings of teaching effectiveness and students' expected grades. In other words, more lenient-grading instructors tend to receive higher student ratings. This positive relationship leads to a major concern of grade inflation. On the other hand, Marsh and Roche (2000) and Decanio (1986) find no relationship between student ratings and the grades students expect to receive. Because of these conflicting findings of the prior studies, no expectation is placed upon the sign of EXPGRADE coefficient.
- TUTOR is a dummy variable which takes the value of 1 if a student used any tutoring services for the course and 0 otherwise. Students, who use tutoring services, normally have hard time understanding course materials. These students may blame an instructor for their hard time in the course, and may give the instructor poor ratings. A negative relation is, therefore, expected between TUTOR and RATING.
- HONOR is a dummy variable which takes the value of 1 if a student is in an honor program and 0 otherwise. Honor students tend to be more serious about their learning and grades. They may also have a better ability to understand course materials than non-honor students. Therefore, they are more likely to enjoy a course, and therefore, give an instructor better ratings on teaching effectiveness. This implies a positive relationship between HONOR and RATING.
- GENDER is a dummy variable which takes the value of 1 if a student is female and 0 if male. Centra & Gaubatz (2000) and Feldman (1998) note that a same sex favorability in rating exists, i.e., students of the same sex as their instructor may provide higher ratings of the instructor's teaching effectiveness. Because the instructor in this study is female, a positive relationship between GENDER and RATING is expected.

**RESULTS**

Table 1 presents descriptive statistics of all variables in the regression model. The instructor received an average rating of 7.279 from 10. The mean SECTION of .459 represents the fact that there were 28 student evaluations in one section and 33 in the other. The majority of students had relatively high pre-course interest with the mean INTEREST of 6.607 and the median of 8. An average number of hours students studied for an exam were 3.829 hours. The mean TUTOR of 0.131 suggests that only eight out of 61 students used tutoring services. EXPGRADE has the mean of 2.739 and the median of 3, suggesting that B and B- were the grades the majority of students expected in the course. The mean HONOR of 0.115 suggests that only seven out of 61 students were in the honor program. The mean GENDER of 0.344 represents the fact that there were 21 female and 40 male students.

Table 2 reports regression results which indicate two variables with a significant and positive coefficient. They are SECTION with a significance level of less than 0.05 and INTEREST with a significance level of less than 0.001. These results support both hypotheses. That is an instructor could receive different teaching evaluations across different sections of the same course, and the teaching evaluations are positively related to students' pre-course interest. This positive relationship between the teaching evaluations and the pre-course interest is consistent with earlier studies even though this study measures the pre-course interest at the beginning instead of at the end of a course. The insignificant coefficient of EXPGRADE suggests that the expected grade is not related to the rating of teaching effectiveness, and that the grade inflation is not present in this study. The regression model has a very high explanatory power with an $R^2$ value of 0.634.

This study also attempts to explain the section effect by examining whether the two sections differ significantly in terms of six student-related variables. These variables are a student's expected grade, the beginning-of-semester GPA which could proxy for how academically motivated a student was, the percentage of females, and the number of accounting majors, students using tutoring services, and honor students. The GPA, which is not part of the anonymous course evaluation, is obtained from the university record. The analysis indicates no significant difference (at 0.05 level) between the two sections with respect to the mean expected grade (2.63 vs. 2.83), an average GPA (2.745 vs. 2.817), the percentage of females (39.3% vs. 30.3%), the number of accounting majors (1 out of 28 vs. 1 out of 33), students using tutoring services (2 out of 28 vs. 6 out of 33), and honor students (2 out of 28 vs. 5 out of 33). Therefore, these six student-related variables cannot explain the section effect. The class-size effect (Ellis et al, 2003) also cannot explain this section effect because both sections have similar size (28 and 33 student evaluations).

**CONCLUSIONS**

This study investigates two relevant hypotheses concerning teaching evaluations. The first one is that an instructor could receive different ratings of teaching effectiveness across different sections of the same course. The second one is that the rating is positively related to students' pre-course interest. The regression results provide support for both hypotheses. This regression analysis, which also controls for students' expected grade, does not indicate any grade inflation. The results suggest that higher-education administrators may want to: (1) assess students' pre-course interest at the beginning of a course, and (2) consider the section effect and the students' pre-course interest when evaluating an instructor's teaching effectiveness for the merit, promotion and tenure decisions.

A limitation of this study is its relatively small sample of teaching evaluations for only one undergraduate course taught by an instructor at a private university. These study design characteristics, however, are not so different from other research conducted in the area of student learning and student rating; that is, the use of undergraduates as participants and a single instructor (e.g., Connor-Green, 2000; Williams & Ceci, 1997). Additionally, to test the section effect, it is necessary to limit the analysis to only one course taught by the same instructor. Future studies may want to replicate this study using a larger sample, perhaps at a public university where the class size of each section of an introductory course tends to be much larger. Future studies may also want to examine potential explanations for the section effect which cannot be explained in this study.

**Table 1**
**Descriptive Statistics Of Regression Variables**

| Variables | Minimum | 25% | Mean | Median | 75% | Maximum |
|-----------|---------|-----|------|--------|-----|---------|
| RATING | 1 | 6 | 7.279 | 8 | 9 | 10 |
| SECTION | 0 | 0 | 0.459 | 0 | 1 | 1 |
| INTEREST | 1 | 5 | 6.607 | 8 | 9 | 10 |
| HOUREX | 0.25 | 2 | 3.829 | 3.75 | 5 | 10 |
| TUTOR | 0 | 0 | 0.131 | 0 | 0 | 1 |
| EXPGRADE | 1 | 2 | 2.739 | 3 | 3 | 4 |
| HONOR | 0 | 0 | 0.115 | 0 | 0 | 1 |
| GENDER | 0 | 0 | 0.344 | 0 | 1 | 1 |

RATING = Students' ratings of teaching effectiveness. SECTION = 1 for a section and 0 for the other. INTEREST = Students' pre-course interest. HREXAM = A number of hours a student studied for an exam. EXPGRADE = A grade a student expected to get in the course. TUTOR = 1 if a student used any tutoring services and 0 otherwise. HONOR = 1 if a student was in an honor program and 0 otherwise. GENDER = 1 for female and 0 for male. There are 61 observations of student evaluations.

**Table 2**
**Regression Results Of Sections And Pre-Course Interest On Teaching Evaluations**
$$RATING = b_0 + b_1SECTION + b_2INTEREST + b_3HREXAM + b_4EXPGRADE + b_5TUTOR + b_6HONOR + b_7GENDER$$

| Variables | Expected Sign | Est. Coeff. | Std. Error | t-Statistic | Prob. > t |
|-----------|---------------|-------------|------------|-------------|-----------|
| SECTION | n/a | 0.882 | 0.250 | 2.15 | 0.042[**] |
| INTEREST | + | 0.540 | 0.086 | 6.30 | 0.000[****] |
| HREXAM | n/a | -0.019 | 0.096 | -0.20 | 0.844 |
| EXPGRADE | n/a | -0.112 | 0.310 | -0.36 | 0.720 |
| TUTOR | - | -0.643 | 0.507 | -1.27 | 0.106 |
| HONOR | + | 0.965 | 0.931 | 1.04 | 0.153 |
| GENDER | + | 0.271 | 0.420 | 0.65 | 0.261 |
| Intercept | n/a | 3.703 | 1.073 | 3.34 | 0.007[***] |
| $R^2$ = 0.634 | | F-Statisitic = 20.55 | | Probability > F = 0.000[****] | |

RATING = Students' ratings of teaching effectiveness. SECTION = 1 for a section and 0 for the other. INTEREST = Students' pre-course interest. HREXAM = A number of hours a student studied for an exam. EXPGRADE = A grade a student expected to get in the course. TUTOR = 1 if a student used any tutoring services and 0 otherwise. HONOR = 1 if a student was in an honor program and 0 otherwise. GENDER = 1 for female and 0 for male. There are 61 observations of student evaluations.
[**], [***], [****] Statistically significant at $p < 0.05$, $p < 0.01$ and $p < 0.001$, respectively.

**REFERENCES**

1.      Centra, J. A. & Gaubatz, N. B. (2000). Is there gender bias in student evaluationsof teaching? *The Journal of Higher Education*, 70, 17-33.
2.      Connor-Green, P. A. (2000). Assessing and promoting student learning: Blurring the line between teaching and testing. *Teaching of Psychology*, 27 (2), 84-88.
3.      Decanio, S. J. (1986). Student evalutions of teaching—A multinominal logit approach. *Journal of Economic Education*, 17, 165-176.
4.      Ellis, L, Burke, D. M., Lomire, P., & McCormick D. R. (2003). Student grades and average ratings of instructional quality: The need for adjustment. *Journal of Educational Research*, September/October, 97, 35-40.
5.      Feldman, K. A. (1998). Reflections on the study of effective college teaching and student rating: One continuing question and two unresolved issues. In: Smart, J. C., Editor, 1998. *Higher education: Handbook of theory and research*, Agathon, New York, 35-74.
6.      Gramlich, E. M. & Greenlee, G. A. (1993). Measuring teaching performance. *Journal of Economic Education*, 24, 3-12.
7.      Greenwald, A. G. (1996). Applying social psychology to reveal a major (but correctable) flaw in student evaluations of teaching. Working Paper, University of Washington.
8.      Greenwald, A. G. & Gillmore, G. M. (1997a). Grading leniency is a removable contaminant of student ratings. *American Psychologist*, 52 (11), 1209-1217.
9.      Greenwald, A. G. & Gillmore, G. M. (1997b). No pain, no gain? The importance of measuring course workload in student ratings of instruction. *Journal of Educational Psychology*, 89 (4), 743-751.
10.     Griffin, B. W. (2004). Grading leniency, grade discrepancy, and student ratings of instruction. *Contemporary Educational Psychology*, October, 410-425.
11.     Krautmann, A. C. & Sander, W. (1999). Grades and student evaluations of teachers. *Economics of Education Review*, 18, 59-63.
12.     Marsh, H. W. & Roche, L. A. (1997b). Making student evaluations of teacher effectiveness effective. *American Psychologist*, 52 (11), 1187-1197.
13.     Marsh, H. W. & Roche, L. A. (2000). Effects of grading leniency and low workload on students' evaluations of teaching: Popular myth, bias, validity, or innocent bystanders? *Journal of Educational Psychology*, March, 202-228.
14.     William, W. M. & Ceci, S. (1997). How'm I doing? Problems with student ratings of instructors and courses. *Change*, September-October, 13-23.
15.     Wilson, R. (1998). New research cast doubt on value of student evaluations of professors. *The Chronicle of Higher Education*, A12-A14.

**NOTES**