ISSN: 2473-3857
v2 n4060

RateMyInformationSystemsProfessor: Exploring the factors that influence student ratings

Mark Sena
sena@xavier.edu

Elaine Crable
crable@xavier.edu

Department of Management Information System
Xavier University Williams College of Business
Cincinnati, OH 45207-1215, USA

Abstract

Based on 820 entries on Ratemyprofessors.com, we explore whether information systems course ratings differ from those in marketing or management courses, whether lower level course ratings differ from those in senior or graduate level courses, whether course ratings differ between genders, and whether perceived course difficulty impacts course ratings. Our findings did not reveal significant differences between information systems and other subjects. However we did find a substantial relationship between perceived course difficulty and overall course ratings. Rating differences between genders and across course levels was not found to be statistically significant for information systems courses given our sample size.

Keywords: student evaluations, university teaching, ratemyprofessors.com, student opinion

1. INTRODUCTION

The evaluation of faculty teaching by students has been occurring for decades. It remains a major consideration as a measure of teaching effectiveness and quite often a major decision in promotion and tenure for faculty. These have typically been evaluations based on written forms filled out anonymously by the students in a classroom with controlled processes (Cashin, 1995; Centra, 2003). This research on student evaluation of faculty was extended by a number of authors (Otto, Sandford, Jr. & Ross, 2008; Bleske-Rechek & Michels, 2010; Felton, Mitchell, & Stinson, 2004) when a different source of evaluation came on the scene with the World Wide Web. Online faculty rating sites included in the early 2000s were RateMyProfessors.com, PassCollege.com, ProfessorPerformance.com, RatingsOnline.com and Reviewum.com (Foster, 2003). RateMyProfessors.com (RMP) has been the most enduring and most used site while the others have lost their popularity over the past decade.

RMP is a student review site, founded in May 1999 by John Swapceinski, a software engineer from Menlo Park, California. RMP allows college and university students to assign ratings to professors in America, Canada, and United Kingdom institutions. The RateMyProfessor (RMP) site was originally launched as TeacherRatings.com and converted to RateMyProfessors.com in 2001. According to RMP it has been around for over a decade and as of July 2016 it contained 8,000+ schools and 1.4 million rated professors with over 15 million student ratings. RMP has altered the landscape of information available to students and claims to be the biggest online listing of faculty ratings. This site allows students to assign numeric ratings to instructors for Easiness, Clarity, and

Helpfulness and the latter two scores become averaged to provide a rating of Overall Professor Quality.

Past research on RMP has primarily focused on the reliability and validity of the information posted at the site and the results have been mixed. Some research has indicated that students just focus on the exceptionally good or exceptionally poor faculty (Kindred & Mohammed, 2005) while other research has indicated students focus on issues unrelated to learning like course difficulty or workload (Davison & Price, 2009) plus faculty sexiness (Silva K, Silva F, Quin, Draper, Cover & Munoff, 2008). Even with all this one study found that RMP had reasonable correlations with traditional in-class evaluations (Coladarci & Komfield, 2007).

Regardless of the validity or reliability of RMP's results, students still flock to the site to make course selection decisions. Kindred and Mohammed (2005) found that students used RMP frequently to discover what other students had to say about a professor in order to use it for course selection purposes and also found there was a jump in frequency of use around registration times. The students reported that it was a good way to evaluate a potential instructor without having to talk to numerous other students and advisors to find out similar useful information.

The Hayes and Prus (2014) study found that students look for reliable and useful information to help them make course selection decisions and their study suggested that students believed that RMP was as useful and reliable as more traditional sources. While their data indicated that students do critically evaluate sources and the information these provide, that information may be biased by factors that students are not aware, such as halo effects and difficulty bias, and therefore, could be less valid. A confounding issue when using RMP for course selection was discussed by Felton et al., (2004). They found that RMP ratings could be affected by perceived difficulty. The perceived easier instructors received higher scores on Helpfulness, Clarity, and Overall Quality ratings. Since students perceive these ratings to be useful and reliable when making course selection, difficulty may indirectly affect course selection decisions. In addition, students who read the negative reviews on RMP often will form less positive expectancies for a course, which could result in less effort on the part of the students in selected courses (Kowai-Bell et al., 2011).

2. FACTORS AFFECTING STUDENT OPINIONS

As briefly mentioned in this Introduction, research has found that students are affected by a number of factors when selecting courses. Students wants courses that will fit their schedule but gender has always been a significant factor (Wilson, Stocking, & Goldstein, 1994) and students also have preferred instructors considered to be extroverted (Radmacher & Martin, 2001) and sexy (Silva et al., 2008). Other researchers have found students consider factors like course difficulty and workload (Davison & Price, 2009) to be important. Babad & Tayeb (2003) found that students will choose more difficult courses if the evaluations indicated a high level of perceived learning value even if the course was considered difficult.

RMP gives students access to the type of information they seek within the qualitative student comment area as well as in the quantitative course evaluation area. Hayes and Prus (2014) found in their study that students believe that RMP is as useful and reliable as more traditional sources such as other students and their advisors. They found that the students consider all the available information, weighing numeric averages equally with any anecdotal comments. Students use the evidence to make course selections regardless of any bias being posted by others. Interestingly, one study found that RMP correlated quite well with traditional in-class evaluations (Coladarci & Kornfeld, 2007) so the students might be using relatively valid in-class evaluations for their course selections.

3. RESEARCH QUESTIONS AND METHODOLOGY

Student evaluations have come under fire for their potential unreliability in measuring teaching effectiveness (Boring et al, 2016). In addition to the potential for gender bias, some faculty perceive that student evaluations may vary according to subject matter or the degree of rigor imposed by the instructor. Information systems courses are a requirement for business degrees in nearly every AACSB accredited undergraduate degree program. Faculty may believe that students who are required to take a particular course may be less interested in the material. In addition, due to computer anxiety

ISSN: 2473-3857
v2 n4060

and the inherent challenges of teaching information systems to students with varying degrees of skill and aptitude, faculty may feel that strong student evaluations may be more difficult to achieve in introductory or core classes.

These factors are important to study, not only because they add to the body of research on perceived teaching effectiveness and online reputation systems, but they may also inform faculty and administrators about potential biases in annual merit review or tenure and promotion decisions.

Using data collected from RMP, this study examines the impact of course subject (Information Systems vs. other business subjects), course level (as designated by the course number), gender, and the perceived level of course difficulty on instructor ratings.

We believe this study will have practical contributions to faculty and administrators regarding patterns and potential bias of student ratings while adding to the growing body of research in the areas of student evaluations and, more broadly, online reputation systems.

Specifically, we will explore the following research questions:
1. Does the mean of overall ratings for Information Systems courses differ from the mean of Marketing or Management courses?
2. Does the mean of overall ratings for Information Systems courses differ by course level (100-300 level, vs 400 level vs grad level)?
3. Does the mean of overall ratings for Information Systems courses differ by the gender of the instructor?
4. Is the perceived difficulty of information systems courses negatively correlated with the overall ratings of courses?
5. Does the mean of perceived course difficulty differ for information systems courses vs. Marketing or Management courses?
6. Does the mean of perceived course difficulty by course level?
7. Does the mean of perceived course difficulty differ by the gender of the instructor?
8. Is the correlation between overall ratings and course difficulty impacted by gender, course level, or by discipline?

In order to examine these questions, 820 ratings were collected from RMP. Potential ratings were identified by searching RMP for ratings from a randomized list of AACSB accredited universities. Thirty-four universities were included in the sample. The most recent rating for up to ten information systems, marketing, and management instructors was collected. In total, the sample included 290 information systems ratings, 266 management ratings, and 264 marketing ratings. There were 532 males and 281 females in the sample (there were seven observations where the gender was not able to be determined). For each observation, the course discipline, course level (100, 200, 300, 400 or graduate), overall rating, difficulty rating, and gender were collected.

4. FINDINGS

As shown in Table 1, there was not a significant difference in the overall mean between the subject of Information Systems as compared with two other business subjects, Management and Marketing.

As shown in Table 2, there is a modest difference in mean ratings of senior and graduate level Information Systems courses as compared with those of 100 thru 399 level courses. However, the t-test for difference of means is not significant with a p-value of .13. Perhaps with additional observations (there were only 79 senior and grad entries), this difference would be statistically significant. Interestingly, there was a substantial difference in Management ratings but none in Marketing ratings.

Table 1: Overall Mean Rating by Subject

Subject	Mean Overall Rating	T-stat *	P-value
INFO	3.61		
MGMT	3.68	-.58	.28
MKTG	3.64	-.30	.38

* one tailed two sample t-test INFO vs. other subjects

Table 2: Overall Mean Rating by Course Level

Subject	100-399 Level	Senior /Grad	T-stat*	P-value
INFO	3.55	3.76	-1.12	.13
MGMT	3.55	3.93	-2.09	.02
MKTG	3.64	3.65	-.02	.49

** one tailed two sample t-test by course level*

Table 3: Overall Mean Rating by Gender

Subject	Female	Male	T-stat*	P-value
INFO	3.49	3.65	-.92	.18
MGMT	3.77	3.64	.69	.24
MKTG	3.56	3.69	-.71	.24

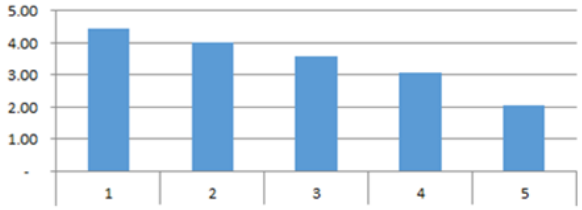
** one tailed two sample t-test by gender*

ISSN: 2473-3857
v2 n4060

The results shown in Table 3 indicate that males received a higher overall mean rating than females in information systems. However, again the t-test for difference of means is not significant with a p-value of .18. Note that in management, females actually had a higher (though insignificant) mean than did males.

In Figure 1, the chart shows that there is a notable pattern between the overall rating for information systems courses and the perceived difficulty of the course. This is supported by a significant correlation (R = -.49). The mean ratings vary substantially from a mean of 4.43 for courses with a difficulty rating of 1 to a mean of only 2.05 for courses with a difficulty rating of 5. A similar pattern was found for marketing and management courses with correlations of R = -.48 for each of those subjects.

Figure 1: Overall Mean Rating by Difficulty Level for Information Systems Courses



Difficulty Level (1 = easy, 5 = difficult)
Correlation (R) = -.49

Interestingly, Table 4 shows that Information Systems is actually rated overall as less difficult (average of 2.72) than courses in Management (2.93)and Marketing (3.02). The t-test difference in means are both statistically significant at p< .05. Perhaps because there are an abundance of introductory courses offered in information systems, students view them as less difficult overall as compared to management and marketing subjects.

Tables 5 shows that we found virtually no difference in perceived difficulty across course levels. Table 6 shows that males are considered more difficult than females in management courses. However, in information systems females had a higher mean, although the difference was not significant given the sample.

Table 4: Overall Mean Difficulty by Subject

Subject	Difficulty Rating	T-stat *	P-value
INFO	2.72		
MGMT	2.93	-1.90	.03
MKTG	3.02	-2.81	.002

** one tailed two sample t-test INFO vs. other subjects*

Table 5: Overall Mean Difficulty by Course Level

Subject	100-399 Level	Senior /Grad	T-stat*	P-value
INFO	2.72	2.73	-.08	.47
MGMT	2.93	2.92	.07	.47
MKTG	3.00	3.08	-.47	.32

** one tailed two sample t-test by course level*

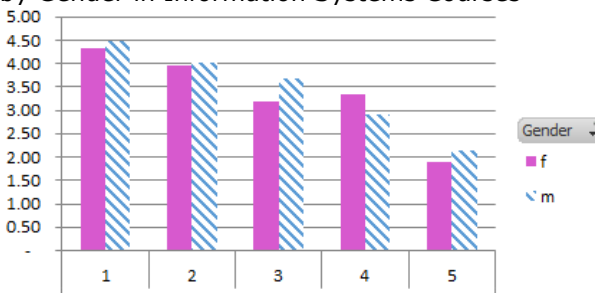
Table 6: Overall Mean Difficulty by Gender

Subject	Female	Male	T-stat*	P-value
INFO	2.78	2.70	.52	.30
MGMT	2.70	3.03	-1.98	.02
MKTG	2.97	3.07	-.68	.25

** one tailed two sample t-test by gender*

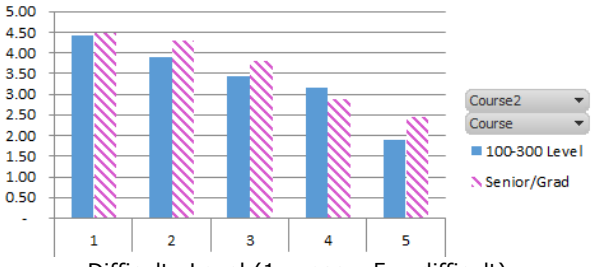
As shown in Figures 2 and 3, the relationship between perceived difficulty levels and overall ratings in information systems courses is similar for different genders and across course levels. The correlation between difficulty and ratings is significant for both females (r=-.48) and males (r=-.50) and for 100-399 level courses (r=-.48) and senior or grad level courses (r=-.49).

Figure 2: Overall Mean Rating vs Difficulty Level by Gender in Information Systems Courses



Difficulty Level (1 = easy, 5 = difficult)
Correlation (R) = -.48 females; -.50 males

Figure 3: Overall Mean Rating vs Difficulty by Course Level in Information Systems Courses



Difficulty Level (1 = easy, 5 = difficult)
Correlation (R) = -.48 100-399 level; -.49 senior/grad

5. CONCLUSIONS

As student course evaluations remain a common yet controversial method of assessing the quality

ISSN: 2473-3857
v2 n4060

of instruction, it is important to examine any factors that might influence these measures. This study explored potential differences in student ratings by course subject, course level, gender, and perceived course difficulty. Our findings indicate that information systems courses are not rated lower than those of marketing or management courses. We found moderate but statistically insignificant differences in ratings across different course levels and gender. We did find a substantial relationship between perceived course difficulty and student ratings. In terms of course difficulty, our findings indicated that information systems courses were viewed as less difficult than those of marketing and management courses. There were little differences in perceived difficulty between course levels and gender. The significant negative correlation between perceived course difficulty and course ratings was consistent across course levels and different genders.

This study provides evidence to support or refute some anecdotal claims by instructors regarding student ratings. The claim that information systems courses are harder or rated lower as compared to marketing or management courses was not supported. Conversely, our study would support any claim that a more difficult class results in lower student ratings. Any claim regarding course level and gender bias in student evaluations should require addition study as there were not statistically significant results in this study given the sample sizes.

This study has some inherent limitations given the use of RMP as a means of data collection. Clearly RMP data could suffer from non-response bias and lack of controls for the subject pool. While we collected a large overall sample size of 820 observations, when broken down by subject, class level, gender, and difficulty levels, some measurements could have used additional observations to better examine the effects. This study could certainly be extended to other course subjects or to measure additional effects such as course subjects within information systems, demographic differences (age, ethnicity, etc,) of instructors, research productivity of faculty, and many other potentially interesting factors that may influence student ratings.

6. REFERENCES

Babad, E. & Tayeb, A. (2003). Experimental analysis of students' course selection. British Journal of Educational Psychology, 73, 373-393. doi: 10.1348/000709903322275894 .

Bleske-Rechek, A. & Kelsey Michels (2010) Ratemyprofessors.com: testing assumptions about student use and misuse. Practical Assessment, Research & Evaluation. 15, no. 5: 1-11.

Boring, A., Ottoboni, K., & Stark, Ph. B. (2016). Student evaluations of teaching (mostly) do not measure teaching effectiveness. ScienceOpen Research. https://www.scienceopen.com/document/vid/818d8ec0-5908-47d8-86b4-5dc38f04b23e

Cashin, W. (1995) Student ratings of teaching: the research revisited. Idea Paper No. 32, Center for Faculty Evaluation and Development, Kansas State University.

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

Coladarci, T., & Komfield, I. (2007). RateMyProfessors.com versus formal in-class student evaluations of teaching. Practical Assessment, Research & Evaluations, 12(6) www.pareonline.net/getvn.asp?v=12&n=6

Davison, E., & Price, J. (2009). How do we rate? An evaluation of online student evaluations. Assessment & Evaluation in Higher Education, 34, 51-65.

Felton, J., Mitchell, J., & Stinson, M. (2004). Web-based student evaluations of professors: The relations between perceived quality, easiness and sexiness. Assessment & Evaluation in Higher Education, 29, 91-108. doi: 10.1080/0260293032000158180

Foster, A. (2003). Picking apart pick-a-prof: does the popular online service help students find good professors, or jut easy A's? Chronicle of Higher Education 49, no. 26: A33-34.

Hayes, M. W. & J. Prus (2014). Student use of quantitative and qualitative information on

ISSN: 2473-3857
v2 n4060

ratemyprofessors.com for course selection. College Student Journal; Winter 2014, 48, No 4, 675-688.

Kindred, J., & Mohammed, S. N. (2005). He will crush you like an academic ninja!: Exploring teacher ratings on Ratemyprofessors.com. Journal of Computer-Mediated Communication, 10 (3), article 9. Retrieved from: http://jcmc.indiana.edu/vol 10/issue3/ kindred.html

Kowai-Bell, N., Guadagno, R. E., Little, T., Preiss, N., & Hensley, R. (2011). Rate my expectations: How online evaluations of professors impact students' perceived control. Computers in Human Behavior, 27, 1862-1867.

Otto, J., A.S. Sandford, Jr., and N. Ross (2008) Does ratemyrofessor.com really rate my professor? Assessment & Evaluation in Higher Education, 33, no. 4: 355-368.

Silva, K. M., Silva, F. J., Quinn, M. A., Draper, J. N., Cover, K. R., & Munofif, A. A. (2008). Rate my professor: Online evaluations of psychology instructors. Teaching of Psychology, 35, 71-80

ISSN: 2473-3857
v2 n4060