

EVALUATING RATEMYPROFESSORS.COM AS A VALUABLE TOOL FOR BOTH MARKETING STUDENTS AND EVALUATORS

Gregory S. Black, Metropolitan State University of Denver
Ann B. Murphy, Metropolitan State University of Denver

Abstract

This study uses data from official student evaluations in a comparison with comparable data from RateMyProfessors.com to assess the similarity of the data. Care is taken to assure that data used is directly comparable in terms of the same instructor/professor, the same class, and the same semester. Analyzing a total of 108 cases suggests that the data from RateMyProfessors.com is comparable to the data from official student evaluations. These findings imply that research that has previously been impossible to perform because of the difficulty of getting access to official student evaluations can be performed using unofficial student evaluations with the confidence that similar findings will result.

Introduction

If your complaint about a service or a service provider would never have any visible impact, would you go to the trouble of complaining? Conversely, thanking someone anonymously without ever knowing if that person whom you were thanking would even be aware of your gratitude may also be a deterrent. However, that is exactly what we are asking our students to do when we have them complete confidential student evaluations about us and our classes. Despite these problems, many universities and colleges consider these student evaluations to be so important that faculty members may be rewarded (retained, tenured, promoted) or punished (denied retention, tenure or promotion) because of them. Further, students have no way of knowing whether other students are being benefited from their evaluation of a professor indirectly (professors being rewarded or punished in the long term). And because of the confidentiality of these evaluations, they have no direct impact (other students having access to these ratings to help them select classes, professor, schools).

While many marketing educators, especially those in evaluative supervisory positions, realize there is value in student evaluations and that students are being benefited indirectly, the evaluations are not directly available to students, providing no short-term benefits to other students. They cannot use them as a tool to help them plan their educational experience. Thus, services like RateMyProfessor.com (RMP) are the tools to which an increasing proportion of students are turning for these planning purposes.

RMP is the largest of these service websites and allows students to rate both professors and colleges/universities. The website is open to the public and has user-generated ratings. A student who takes a class from a particular professor is able to rate that professor and post comments about why they do or do not recommend the professor. The same goes for the college/university ratings on RMP. In addition, every year, RMP compiles a list of top-rated professors and colleges nationwide.

The internet enhances our ability to do product research and we often consider other consumers' ratings of a product to be among the most useful data when making purchase decisions. RMP enables customers of higher education to accomplish these objectives. RMP is becoming increasingly important to students in setting expectations for classes (DeJong 2008) as well as helping them make more informed decisions. Not only are the students paying attention to the quantitative ratings on RMP, but the findings of one study suggest that the qualitative comments in RMP serve as word-of-mouth social communication (Hartman and Hunt 2013).

Despite the increased use of RMP information by students, is this information truly valuable and accurate? If the information the students are receiving is so important to them, should it also be valuable to faculty members and to those who are in positions of evaluating them? Evaluators of faculty performance often place maximum importance in official student evaluations administered by the colleges and universities. Since RMP also provides student ratings and comments, should these have value to faculty members and their evaluators? This study is designed to begin to answer these questions.

Literature Review and Hypothesis

Official Student Evaluations

Official student evaluations are routinely required at most universities (Chen et al. 2004), and most of these universities use these student evaluations as an important part of evaluating teaching performance of faculty members (e.g., Ahmadi et al. 2001; Hobson and Talbot 2001). In addition, these official evaluations are often used as a way to assess quality assurance (Kwan 1999). Concerns have arisen concerning the development of the instrument (Marsh 1987), its validity and reliability (Cohen 1981; Feldman 1977; Marsh 1987), and the potential of bias (Chen et al. 2004; Tollefson et al. 1989). However, these official student evaluations continue to be viewed as an important part of evaluating professors.

Perhaps the most popular belief about official student evaluations by faculty is that grades have a strong impact on these evaluations (Clayson 2004). However, after many years of attempting to make this link absolutely, there is still no clear answer as to whether student grades have any impact on student evaluations (e.g., Clayson 2004; Gotlieb 2009; Grant 2007; Grimes et al. 2004; McPherson and Todd 2007; Paswan and Young 2002). Professors may even be reluctant to give low grades in fear of low evaluations, particularly before they achieve tenure (Benton 2006). Other research has suggested that grades do not have a direct impact, but may have an indirect impact because of attribution (Gotlieb 2009) or the principle of reciprocity (Clayson 2004).

Other research has focused on other factors impacting official student evaluations. For example, Candill (2002) examined class size and found that it did not have an impact on overall student evaluations. Lee (2011) found that students who feel they have some sort of control over at least a portion of the class and how it is run are more likely to rate the class more positively. Wheeler (2008) found evidence that student evaluations tend to be more positive in classes that include experiential learning. Other research has shown that quality professor-student interaction (Wheeler 2008), class length (Reardon et al. 2008), and class rigor (Clayson et al. 1990) all impact official student evaluations.

One definitive literature review indicates several importance findings about these official student evaluations. First, these evaluations are reliable and stable. Second, they are primarily a function of the professor who teaches a course rather than the course being taught. Third, they are relatively unaffected by such factors as grading leniency, class size, workload in the class, and prior subject interest. Fourth, they are useful in improving instructional effectiveness when coupled with appropriate consultation. Finally, official student evaluations should adopt a broad approach rather than a narrow approach including only a small number of effectiveness variables (Marsh and Roche 1997; Theall and Franklin 2002).

Another set of studies has identified four dimensions of importance for evaluating teaching effectiveness. The first of these four dimensions is content expertise which includes formally recognized knowledge, skills and abilities a faculty member possesses in a chosen field by virtue of advanced training, education and/or experience. The second dimension is instructional design which determines how students interact with the content, and includes designing, sequencing, and presenting experiences intended to induce learning. The third dimension is

instructional delivery and includes those human interactions that promote or facilitate learning, as well as various forms of instructional delivery mechanisms. The final dimension is instructional assessment that includes developing and using tools and procedures for assessing student learning, both to provide feedback and to assign grades (e.g., Arreola 2007; Berk 2006). This same body of research also suggests that only three of the four dimensions are appropriately assessed with official student ratings. Students are not in an appropriate position to evaluate content expertise (e.g., Arreola 2007; Berk 2006).

RateMyProfessors.com Evaluations

RMP is the largest online destination for professor ratings. Whether the parent company, MTV, lends credence to the site or not, it at least shows that it has a heavy hitter behind it (Marcus 2011). Students have contributed 14 million ratings of 1.3 million professors and 7,000 schools. It includes ratings for professors from the United States, Canada and the United Kingdom. It is estimated that more than 4 million college students use the site each month. According to the website, “The site does what students have been doing forever – checking in with each other – their friends, their brothers, their sisters, their classmates – to figure out who’s a great professor and who’s one you might want to avoid” (www.RateMy Professors. com).

RMP uses a five-point scale similar to a semantic-differential format where the higher number indicates a more positive score for the professor being rated. Five primary factors about a professor are measure and each evaluation is for a specific class in a specific semester. The five factors used are helpfulness, clarity, easiness, interest, and hotness.

Helpfulness and clarity are considered to be the serious factors and it is these two factors that are considered in the overall quality calculations for the professor. Also, these two factors are most relevant and similar to official evaluation items and are basically assessing the dimension of instructional delivery. Helpfulness is measured by having students respond to the following questions. “Is this professor approachable, nice and easy to communicate with? How accessible is the professor and is he/she available during office hours or after class for additional help?” The end points for the scale are 1 = useless and 5 = very helpful. Clarity is measured by having students respond to the following questions. “How well does the professor teach the course material? Were you able to understand the class topics based on the professor’s teaching methods and communication style?” The end points for the scale are 1 = confusing and 5 = crystal clear.

Some students may want to know how easy or difficult a class is before they register, so RMP has an easiness item to provide this information to students. Easiness is measured by having students respond to the following questions. “Is this class an easy A? How much work needs to be done in order to get a good grade? The end points for the scale are 1 = hard and 5 = easy.

Another interesting item is student interest before enrolling in the class. It is measured by having students respond to the following question. “Is this subject one of you passions, or are you only taking the class to fulfill mandatory credits?” The end points for the scale are 1 = no prior interest and 5 = Obsessed with the subject before attending class. One last fun item is to allow the students to evaluate a professor’s hotness. What makes a professor hot is left to individual interpretation and is measure by responses to the simple question “Is your professor hot?”

Recent research indicates that RMP evaluations are useful and comparable to the official evaluations given by colleges and universities (Clayson 2014). Even the verbal comments students make on RMP are similar to those made on official instruments (Silva et al. 2008). This is good news for students because they are relying on RMP (and similar sites) more and more (Field et al. 2008). One criticism of RMP is self-selection; in other words, only those

students with extreme opinions, either negative or positive, go to the trouble of evaluating. Other critics claim there is too much of RMP that is for pure entertainment and students do not take it as seriously as they should (Davison and Price 2009). Other research has claimed serious flaws with the validity of the information provided on RMP because of sampling errors (Otto et al. 2005), ratings biases (Otto et al. 2005), and results dissimilar from official student evaluations (Albrecht and Hoopes 2009).

On the positive side, however, one study found that “easiness” was actually considered by students to be of less importance when using RMP as a tool (Landry et al. 2010). Other research found evidence that students do take the task of evaluation seriously and students with extreme opinions are not necessarily the only ones to use RMP to evaluate professors (Peterson et al. 2011). Though some research has been conducted to compare different types of official evaluations (student, peer, and self) (Webster 1990), only a few studies have shown that RMP ratings are positively correlated with official student evaluations (e.g., Brown et al. 2009; Sonntag et al. 2009; Timmerman 2008). None of this small body of research quantitatively comparing official evaluations with RMP evaluations has been conducted in marketing. Thus, the hypothesis for this study follows.

H₁: Official evaluations and RMP evaluations are statistically comparable.

Methodology

RMP evaluations are specific not to a professor, but also to a specific class and semester. In the sample of 20 professors used in this study, no class and semester for a professor was evaluated by more than one student. Thus, each evaluation was considered a separate case for this study. As with RMP evaluations, official evaluations are professor, course and semester specific, allowing the matching up of both types of evaluations. Therefore, using all RMP evaluations for these 20 professors who had a corresponding official evaluation resulted in 108 cases. RMP ratings utilize a five-point scale and the official university ratings used in this study are completed using a six-point scale. Thus, the scale items for both types of evaluations were converted to percentages of the total possible (five or six) to make these ratings directly comparable. The resulting official ratings were then compared to RMP ratings on several dimensions. First, all four RMP factors (helpfulness, clarity, easiness and rater interest) were combined for comparison with the official ratings. Second, since RMP itself considers helpfulness and clarity as the determinants of overall quality (a variable somewhat comparable to what universities are measuring in official evaluations), these two were combined for comparison. Finally, each RMP factor was compared individually to the official evaluations for that faculty member for the same class and semester. All comparisons were performed by simple correlation analysis.

Results

As can be seen in Table 1, correlation analysis results in close relationships between official evaluations and all RMP evaluations. The largest, most significant correlations are between official evaluations and the RMP quality factors, including helpfulness ($r = .540, p \leq .01$), clarity ($r = .576, p \leq .01$), and the variable that combines those two factors ($r = .586, p \leq .01$). However, correlations between official evaluations and RMP evaluations go beyond these factors.

Correlations between the other factors used by RMP and official evaluations were also significant (see Table 1). The correlations between official correlations and the RMP easiness factor ($r = .240, p \leq .05$), the RMP interest factor ($r = .217, p \leq .05$), and the variable that combines all four RMP factors ($r = .534, p \leq .001$) are also significant.

Table 1: Results of Correlation Analysis

	Official Evals	RMP Quality	RMP Total	RMP Helpful	RMP Clarity	RMP Easy	RMP Interest
Correlation Coefficient	-	.586	.534	.540	.576	.240	.217
Significance	-	.000	.000	.000	.000	.012	.024

Discussion and Conclusion

Gaining access to official student evaluations is sometimes difficult, if not impossible, rendering research with this data only an unfulfilled wish. However, research using student evaluation data can clearly be important and useful. The results of this study add to the literature that verifies that utilizing publicly-available student evaluations, such as those provided on RMP, for research could be comparably enlightening.

In addition to using the quantitative data provided on RMP, it may also be useful to examine the qualitative data that is also found there. So far, there has been little research done on examining the written comments on RMP or comparing them with written comments found on official student evaluations. Research is just beginning to be performed and one paper makes suggestions about how to interpret and use the official written comments in evaluating faculty and helping faculty improve their teaching performance (Candill 2002); however, it fails to actually perform any analysis using the data.

With verification of the near equality of RMP data to official student evaluation data, future research using this unofficial data could be performed in any area where student evaluation data analysis could be useful. One such area could be comparing departments across an individual university or across different universities. Another may be to examine the notion that lower performing instructors get a higher quantity of comments and these comments are more negative than higher performing instructors.

If future research continues to verify the equality of unofficial student evaluations, universities may even consider eliminating official evaluations altogether and encouraging students to utilize RMP or similar sites. This measure would conserve valuable resources of the universities. In addition, encouraging these unofficial evaluations would provide more data to our students as they attempt to design their educational program that will maximize their academic experience.

References

- Ahmadi, M., Helms, M. M., & Raiszadeh, F. (2001). Business students' perceptions of faculty evaluations. *The International Journal of Educational Management*, 15(1), 12-22.
- Albrecht, W. S. & Hoopes, J. (2009). An empirical assessment of commercial web-based professor evaluation services. *Journal of Accounting Education*, 27(September), 125-132.
- Arreola, R. A. (2007). *Developing a Comprehensive Faculty Evaluation System: A Guide to Designing, Building, and Operating Large-Scale Faculty Evaluation Systems*, 3rd edition. San Francisco, CA: Anker Publishing.
- Benton, T. H. (2006). A tough love manifesto for professors. *Chronicle of Higher Education*, 52(June 9), C1-C4.
- Berk, R. A. (2006). Should student outcomes be used to evaluate teaching? *Journal of Faculty Development*, 28(2), 87-96.
- Brown, M. J., Baillie, M. & Fraser, S. (2009). Rating RateMyProfessors.com: A comparison of online and official student evaluations of teaching. *College Teaching*, 57(2), 89-92.

- Candill, D. W. (2002). A model for interpreting written comments on student evaluations of marketing courses. *Marketing Education Review*, 12(1), 63-77.
- Chen, Y., Gupta, A., & Hoshower, L. (2004). Marketing students' perceptions of teaching evaluations: An application of expectancy theory. *Marketing Education Review*, 14(2), 23-36.
- Clayson, D. E. (2004). A test of the reciprocity effect in the student evaluation of instructors in marketing classes. *Marketing Education Review*, 14(2), 11-21.
- Clayson, D. E. (2014). The consequences of sampling level error in educational research: The uses and misuses of RateMyProfessor.com. *Proceedings of the Marketing Educators' Association Conference*, 191-192.
- Clayson, D. E. & Haley, D. A. (1990). Student evaluations in marketing: What is actually being measured? *Journal of Marketing Education*, 12(3), 9-18.
- Davidson, E. & Price, J. (2009). How do we rate? An evaluation of online student evaluations. *Assessment & Evaluation in Higher Education*, 34(1), 51-65.
- DeJong, P. F. (2008). Setting learner expectations. *Proceedings of the Marketing Management Association Conference*, 1-2.
- Donmeyer, C. J. (1997). Class size in an introductory marketing course: Student attitudes, evaluations, and performance. *Marketing Education Review*, 7(1), 13-25.
- Feldman, K. A. (1977). Consistency and variability among college students in their ratings among courses: A review and analysis. *Research in Higher Education*, 6(3), 223-274.
- Field, R. J., Bergiel, B. J., & Viosca, R. C., Jr. (2008). An exploratory study of student use and perceived accuracy of faculty evaluations on RateMyProfessors.com. *Competition Forum*, 6(2), 378-383.
- Gotlieb, J. (2009). Justice in the classroom and students' evaluations of marketing professors' teaching effectiveness: An extension of prior research using attribution theory. *Marketing Education Review*, 19(2), 1-14.
- Grant, M. M. (2007). Good teachers good grades. *Chronicle of Higher Education*, 53(June 29), A47.
- Grimes, P. W., Millea, M. J., & Woodruff, T. W. (2004). Grades – who's to blame? Student evaluation of teaching and locus of control. *Journal of Economic Education*, 35(Spring), 129-147.
- Hartman, K. B. & Hunt, J. B. (2013) What RateMyProfessors.com reveals about how and why students evaluate their professors: A glimpse into the student mind-set. *Marketing Education Review*, 23(2), 151-162.
- Hobson, S. M. & Talbot, D. M. (2001). Understanding student evaluations: What all faculty should know. *College Teaching*, 49(1), 26-31.
- Kwan, K. P. (1999). How fair are student ratings in assessing the teaching performance of university teachers. *Assessment & Evaluation in Higher Education*, 24(June), 181-195.
- Landry, N. M., Kurkul, K. E., & Poirer, C. R. (2010). Online faculty rating sites: Examining how students perceive and use RateMyProfessors.com. *Psi Chi Journal of Undergraduate Research*, 15(3), 139-142.
- Lee, J. K. (2011). Effects of perceived control on college students' evaluation of higher education institutions. *Business Education & Accreditation*, 3(1), 51-60.
- Marcus, J. (2011). They're here to stay: Ratings demand a change of tactics. *Times Higher Education*, 1987(1), 18-19.
- Marsh, H. W. (1987). Students' evaluations of university teaching: Research findings methodological issues and directions for future research. *International Journal of Educational Research*, 11(3), 353-388.
- Marsh, H. W. & Roche, L. A. (1997). Making students' evaluations of teaching effectiveness effective. *American Psychologist*, 52(11), 1187-1197.
- McPherson, M. A. & Todd, J. R. (2007). Leveling the playing field: Should student evaluation scores be adjusted? *Social Science Quarterly*, 88(July), 868-881.

- Otto, J., Sanford, D., & Wagner, W. (2009). Analysis of online student ratings of university faculty. *Journal of College Teaching & Learning*, 2(June), 25-30.
- Paswan, A. K. & Young, J. A. (2003). Student evaluation of instructor: A nomological investigation using structural equation modeling. *Journal of Marketing Education*, 24(December), 192-202.
- Peterson, R. L., Berenson, M. L., & Aijaz, R. (2011). An exploratory study: Is there anything good about RateMyProfessors? *Proceedings for the Northeast Region Decision Sciences Institute Conference*, 1178-1201.
- Reardon, J., Payan, J., Miller, C., & Alexander, J. (2008). Optimal class length in marketing undergraduate classes: An examination of preference, instructor evaluations, and student performance. *Journal of Marketing Education*, 20(1), 12-20.
- Silva, K., Silva, F., Quinn, M., Draper, J., Cover, K., & Munoff, A. (2008). RateMyProfessor: Online evaluations of psychology instructors. *Teaching of Psychology*, 35(2), 71-80.
- Sonntag, M. E., Bassett, J. F., & Snyder, T. (2009). An empirical test of the validity of student evaluations of teaching made on RateMy Professors.com. *Assessment & Evaluation in Higher Education*, 34(October), 499-504.
- Theall, M. & Franklin, J. (2002). Looking for bias in all the wrong places: A search for truth or a witch hunt in student ratings of instruction. *New Directions for Institutional Research*, 2001(109), 45-56.
- Timmerman, T. (2008). On the validity of RateMyProfessors.com. *Journal of Education for Business*, 84(September-October), 55-61.
- Tollefson, N., Chen, J. S., & Kleinsasser, A. (1989). The relationship of students' attitudes about effective teaching to students' ratings of effective teaching. *Educational and Psychological Measurement*, 49(Autumn), 529-536.
- Webster, C. (1990). Evaluation of marketing professors: A comparison of student, peer, and self-evaluations. *Journal of Marketing Education*, 12(1), 11-17.
- Wheeler, R. J. (2008). Experiential learning: Impact of two instructional methods on student-instructor interaction, student critical thinking, and student course evaluations. *Journal for Advancement of Marketing Education*, 12(Summer), 63-78.