

# When Profs Get Graded

August 27, 2012



As the popularity of teaching evaluation websites is growing, so is concern over whether ratings on such sites provide an accurate representation of instructors' performance. Because many students rely on such websites such as [Ratemyprofessors.com](http://Ratemyprofessors.com) (RMP) when making course decisions, it is important that we examine how these ratings compare to student evaluations of teaching (SET).

Founded in 1999 by John Swapceinski after a particularly bad experience with a faculty member, RMP is a free website that allows students to anonymously rate an instructor's "easiness," "helpfulness," and "clarity" on a 5-point scale and leave a comment of up to 350 characters. The website generates an "overall quality" rating by averaging an instructor's helpfulness and clarity scores. Students can also rate an instructor's physical appearance (hot or not) by assigning him/her a "chili pepper." Fairly recently, the website began asking students about the attendance policy and textbook used in the class. Students can now also rate and comment on their college as a whole.

Ratings for over 1.7 million instructors from 7,500 different colleges across the United States, Canada, and the United Kingdom, plus 13 million comments, appear on the website as of August 2012 (RMP, 2012). But many have questioned the validity of the website's ratings (e.g., Martin, 2003), despite these high numbers. One concern is that students are not qualified to evaluate instructors in general (e.g., Ahmadi, Helms, & Raiszadeh, 2001), and there are a host of concerns about online evaluations in particular.

Students who use RMP are a self-selected sample; thus ratings on the website may not be representative of the broader student population. A common concern among instructors is that RMP ratings are biased by a disproportionate number of negative ratings. RMP (2012) rejects that criticism, claiming that "well over half of the ratings on this site are positive." However, there is evidence to suggest that students are not motivated to rate an instructor unless they have had a particularly good or particularly bad experience. Kindred and Mohammed (2005) found that most of the comments students reported reading on RMP represented extreme positions, very positive or very negative, with little in between.

The anonymity of RMP also raises concerns. It is impossible to ensure that raters have actually taken a course with the instructor they are evaluating, that professors are not rating themselves or each other, or that multiple submissions from disgruntled students will not skew an instructor's ratings. Thus, there is a potential for ratings to be based on hearsay or experiences outside of the classroom.



The nature of online communication in general may also be a factor in the validity of RMP ratings. Research suggests that people tend to behave differently while online than in person. For example, Siegel, Dubrovsky, Kiesler, and McGuire (1986) found that computer-mediated groups exhibited more hostile behavior (including name-calling, swearing, and insults) than did groups that interacted face-to-face. Thus, students may be inherently more antagonistic in RMP's anonymous online ratings than in on-paper SET ratings. But anonymity can have positive effects too. Research suggests that people are generally more likely to give and receive support (Whitty, 2002), display less social anxiety (Scealy, Phillips, & Stevenson, 2002), and disclose more information (Whitty & Gavin, 2001) while communicating online rather than in person. Thus, students may actually be more open and honest in RMP ratings.

In a content analysis of 1,054 RMP ratings, Kindred and Mohammed (2005) found that the majority of written comments left by users were related to instructors' competence and personality, rather than inconsequential factors such as race, gender, or attractiveness. This finding is encouraging given that the researchers also found that students preferred written comments to the numerical or graphical components of RMP ratings. However, students tended to be suspicious of RMP ratings in general and placed greater trust in the information provided to them by other students directly (Kindred & Mohammed, 2005).

Students continue to use RMP ratings when making academic decisions because they typically have no alternative means to learn about an instructor, though the ratings are susceptible to bias and false submissions. Most colleges conduct their own evaluations to assess instructors' performance, but the results are generally not made available to students. In a study my colleagues and I conducted, we found that among 110 students surveyed, 83% said they have visited Ratemyprofessors.com, 36% said they

have rated an instructor on the website, and 71% said they have avoided taking an instructor's class based on his/her ratings (Brown, Baillie, & Fraser, 2009). With regard to the validity of RMP ratings, 47% of respondents believed that RMP ratings are more representative of instructors' performance than official student evaluations of teaching, 34% believed that both are equally representative, and just 17% believed that official ratings are more representative.

But are RMP ratings really valid? The answer depends on what we mean by "valid." Are students qualified to evaluate college faculty? Do student evaluations simply reflect expected grades? Are student evaluations influenced by non-relevant faculty characteristics (such as race, gender, attractiveness, and sexual orientation)? An abundance of research has examined these questions, often with mixed results (Marsh & Roche, 2000). However, critics and supporters are not easily persuaded by the data.

In our study, we examined the validity of RMP ratings by comparing the ratings for 312 instructors to their SET ratings. Statistical comparisons between RMP and SET ratings revealed moderate to strong correlations. Furthermore, regression analysis found that RMP ratings are significant predictors of instructors' performance as measured by SET ratings. These findings are consistent with those of similar studies (Timmerman, 2008). Overall our results suggest that, in lieu of SET ratings, RMP ratings may serve as a viable alternative for students.

A seemingly obvious way to reduce students' use of RMP is to give them access to official evaluations — a move likely to draw much opposition from instructors (Nasser & Fresko, 2002). However, the results of our survey suggest that students actually believe that RMP ratings are more honest and more representative than SET ratings.

Although it is important that colleges refrain from using RMP ratings to make administrative decisions, instructors may be able to use these ratings to their advantage in some way. Students can leave RMP ratings and comments at any point throughout the semester — not only when the semester ends. Thus, instructors can benefit from mid-semester feedback from students and make adjustments to the course as needed. Constructive interim evaluations can provide valuable information about ways to improve teaching (Algozzine et al., 2004). The net result can be improved SET ratings, which are routinely used by administrators when making decisions about the hiring, tenure, and promotion (Newport, 1996).