

Archived version from NCDOCKS Institutional Repository <http://libres.uncg.edu/ir/asu/>

Davison, Elizabeth L. and Jammie Price. 2009 "How Do We Rate? An Evaluation of Online Student Evaluations" *Assessment & Evaluation in Higher Education*. 34(1) pp 51-65. (Sept. 2009) (ISSN: 0260-2938) This is an Accepted Manuscript of an article published by Taylor & Francis Group, available online: <http://www.tandfonline.com/10.1080/02602930801895695>.

## How Do We Rate? An Evaluation of Online Student Evaluations

Elizabeth Davison and Jammie Price

**ABSTRACT:** This paper analyses the popular RateMyProfessors (RMP) website where students evaluate instructors in higher education. A study was designed to measure (1) the awareness and utilisation of the RMP website, (2) the internal and external validity of the RMP ratings in measuring teaching effectiveness, and (3) variation in the above across disciplines. It is concluded that the category of ratings, created by the website, establishes an anti-intellectual tone that manifests itself in comments about instructors' personality, easiness of workload and entertainment value rather than knowledge attained.

The Rate My Professors (RMP) website ([ratemyprofessors.com](http://ratemyprofessors.com)) offers a community forum where 'students do the grading', anonymously sharing evaluations of instructors with other students. The website aims to make information available so students can choose better instructors and courses, thereby improving their education. The merits of the website are often contested within the walls of academe, especially by the faculty that are directly scrutinised, but largely this website lacks academic or public evaluation. Although the debate concerning the website usually centres on the quality of data provided by the website, there is little information to substantiate the credibility of the data. This paper contributes to the emerging literature (see Felton et al. 2004; Coladarci and Kornfield 2007) that provides information on the services and disservices of this popular website to the academic community.

There are many similar websites, such as [pickaprof.com](http://pickaprof.com), [campusdirt.com](http://campusdirt.com), [myprofessorsucks.com](http://myprofessorsucks.com), and [rateyourprof.com](http://rateyourprof.com), but RMP is by far the most popular student rating site. Since RMP started in 1999,<sup>1</sup> the for-profit website has observed great success and now sells ads to numerous international corporations such as VISA, Monster, MSN, Citibank and Allstate. In early 2004, RMP had almost 1.5 million ratings from nearly 4000 schools. Three years later, the number of postings increased by 300% to over six million. Currently over 6000 colleges and universities and nearly 800,000 instructors are listed across United States, Canada, England, Scotland and Wales. With over eight million student members, the daily traffic averages over 200,000 unique visitors per day. One thing is for sure, the website is not going away. Like it or not, the information on the website is very profitable.

The numbers do not lie: Students like this website. Out of curiosity, we initially set out to explore how we and our colleagues were being portrayed on the RMP website and perhaps use the information to improve our own teaching. In reading the various postings, we immediately observed that students focus on how easy, nice, helpful and entertaining we are. Subsequently, we questioned what the website actually measures, and whether those measures are valid, and how often students use these data in decision-making. Given the popularity of the website, we believe discussions of how this website impacts on the learning environment are important. We began our study by investigating the RMP website and rating process, which we summarise below.

## **How RMP works**

RMP lists several categories which visitors may use to rate an instructor: easiness, helpfulness, clarity, sexual hotness, and the student's overall interest in the course prior to taking it.<sup>2</sup> All are measured with five-point Likert scales. An overall score is averaged from the helpfulness and clarity ratings. High overall scores produce the symbol of a smiling face next to an instructor's name; low overall scores produce a frowning face. Similarly, instructors with high hotness scores receive a chilli pepper next to their names.

The RMP website invites students to offer open-ended comments, within certain parameters such as discouraging racist, sexist and homophobic remarks. RMP asks for a student volunteer from each school to monitor the postings for adherence to these guidelines, though we found many violations. Perhaps most blatantly, RMP itself encourages discrimination by asking students to rate instructors' 'hotness' which sexualises instructors, students and the learning environment. Felton et al. (2004) find that the website's sexiest professors benefit from the 'halo effect' in that students are blinded by their good looks and rate them with high marks regardless of other factors.

What is most striking about the rating process is what is missing. Few safeguards exist to ensure the external validity of the responses. The onus is on the student to log in, register and take the time to post a rating on a particular instructor. This process lends itself to bias, with students who either loved or hated an instructor more likely to post. It does not matter when a student posts a rating. It could be on the first day of the semester, or years after graduation (both of which occur frequently). Even people who never actually took courses with an instructor have access to post ratings. We also detected evidence of several false instructors and students within our own department.<sup>3</sup>

The internal validity of the ratings is highly suspect. Although there is little agreement about what measures are best for analysing teaching effectiveness (see discussion below), we argue that the limited questions on the RMP site are not robust measures of teaching effectiveness. For example, although the website is very clear that easiness is not used to calculate the instructor's particular overall score, the fact remains that easiness is offered as a category to evaluate instructors whereas there is no measure for how much students learn from instructors or how knowledgeable the instructors are about the course topic. Helpfulness and clarity are important dimensions of a classroom experience and certainly should be among the possible categories used to evaluate instructors. However, to have the overall score (posted prominently by each instructor's name), which is perceived as overall effectiveness, defined only by clarity and helpfulness, while not asking anything about the learning process or attained knowledge, calls into question the validity of the website.

From our perspective as instructors, and sociologists, the measures offered on the RMP website appear less academically rigorous and more consumer oriented. They are designed for consumers – students – who are literally shopping for instructors and, subsequently, a degree. Next in our study, we turned to the literature on student evaluations of instructors for explanations of increasing consumerism in college today (Baker and Copp 1997; Delucchi

and Smith 1997a, 1997b; Shepperd 1997; Smith 2000).

## Literature review

Over the last 20 years, many social scientists have studied and critiqued student evaluations of teaching (Abrami and Mizener 1985; Dowd 1988; Goldman 1990; Blunt 1991; Benson and Lewis 1994; Wachtel 1998). Recently, Smith and Pino (2005) observe a declining commitment to an academic work ethic, specifically learning and self-development, among students today. The number of hours that students study today exemplifies this trend. Delucchi and Korgen (2002) disclose that 70% of sociology majors study 10 or fewer hours a week. Similarly, Rau and Durand (2000) estimate an average of eight hours of study a week among a sample of all majors at Illinois State University.<sup>4</sup>

Within this context, what is teaching effectiveness? And if we cannot agree on what teaching effectiveness is (Lucal et al. 2003), what do student evaluations of instructors measure? Numerous studies identify that students define instructors as effective when an instructor is nice, caring, understanding, helpful, enthusiastic, entertaining and, above all, easy (Talley and Timmer 1992; Delucchi and Smith 1997a, 1997b; Basow, 2000; Delucchi 2000). Unfortunately, these characteristics do not predict actual learning (Williams and Ceci 1997), as documented in the infamous Dr Fox study by Naftulin et al. (1973). Students who want ease, niceness and entertainment cannot assess effectiveness in teaching in any reliable or valid way (Wachtel 1998; Sprague and Massoni 2005), as the true measure of effectiveness is learning. Hence, most student evaluations of instructors really measure customer satisfaction (McKeachie 1994; Shepperd 1997; Fries and McNinch 2003).

Nevertheless, at nearly all schools in the United States, student perceptions of instructors serve as the main measure of teaching effectiveness (Wagenaar 1995; Lucal et al. 2003). Administrators use these data in determining retention, promotion and compensation for instructors (Basow 1998). Utilising student evaluations of an instructor's teaching effectiveness ignores the social fact that students and instructors together construct the reality, and the subsequent meaning, of each course. As Baker and Copp (1997, 41) explain, 'course evaluations assume that the instructor "controls" the class and likewise controls what goes on in students' heads'.

Moreover, student evaluations of instruction are empirically biased against a number of groups: gay men and lesbian instructors (Ewing et al. 2003), older instructors (Arbuckle and Williams 2003), less experienced instructors (Fries and McNinch 2003), instructors anticipated to issue lower grades (Sinclair and Kunda 2000), instructors whom students perceive to hold different values and politics from their own (Kemmelmeyer et al. 2005), and women instructors (Centra and Gaubatz 2000).

The gender effect itself is gendered (Sprague and Massoni 2005). As Das and Das (2001, 673) sum: 'males and females may be using different rules to evaluate the communication patterns (and other behaviors) of professors'. The gender effect varies across men and women students and instructors, ages of students and instructors and, most importantly, the perceived synchrony in gender performance between student and instructor (Basow 2000). In general, students like androgynous instructors best (Basow 1998; Das and Das 2001), as well as instructors whom students perceive to match their own sex (Schroeder and Mynatt 1999) and gender performance (Feldman 1993). Specifically, students like their male teachers to be funny and their female teachers to be nurturing (Sprague and Massoni 2005). They expect women teachers to do 'emotion' work and rate women instructors poorly if they fail to meet this expectation (Baker and Copp 1997). Further, Sinclair and Kunda (2000) detect a positive correlation between students' evaluations of women instructors and student grades.

## Methods

Having gained a better understanding of the ratings at the RMP website and student evaluations of instruction more generally, we designed a study to assess: (1) the awareness among our students of the RMP website, (2) their utilisation of the website, (3) validity of the RMP ratings in measuring teaching effectiveness, and (4) variation in the above across disciplines. Our analysis primarily centres on the validity of the RMP data, but if students do not visit or use the RMP data regularly, any such bias is immaterial.

We evaluated the website postings for our place of employment, Appalachian State University, for two reasons. One, because our familiarity with most instructors allows us to better evaluate the validity of comments; and two, our campus consistently lists in the top 100 most rated schools on the RMP website. Our knowledge of our colleagues and access to campus directories allows us to account for the gender of each instructor and confirm his/her department association. Appalachian State University is a comprehensive university of around 15,000 undergraduate and graduate students located in a rural mountain community of western North Carolina in the United States. In 2005, at the time of our study, Appalachian State University employed 601 tenured and tenure-track instructors and 102 non-tenure track instructors.<sup>5</sup>

To assess awareness and utilisation of the RMP website, we administered a short questionnaire to 216 students enrolled in sociology, economics, statistics, and women studies courses (a convenience sample) in 2004. The questionnaire asked students if they had heard of, visited or posted a rating on the RMP website, and if and how they found the RMP data useful. Additionally, we asked students to rank the importance of the information on the website. Here, we added three categories that are not currently asked by RMP: (a) instructor's interest in topic; (b) instructor's knowledge of topic; and (c) how much students learned in the course.

To begin assessing the validity of RMP data in measuring teaching effectiveness, we collected and analysed the quantitative data and qualitative data listed on the website for our university. We analysed the data that are collected by the website to address whether the measures offer the best overall picture of teaching effectiveness. Regarding the quantitative data, we gathered the individual instructor scores on all scales listed on the RMP website and the overall instructor score that RMP calculates from the clarity and helpfulness ratings. Beside each instructor's name is the overall score, date of last posting, number of ratings and easiness score. If you click on an instructor's name, you will see all the individual postings for that instructor. Each individual posting lists the date of posting, how the individual rated the instructor on easiness, helpfulness, clarity and interest in subject, and any comments that the individual made about the instructor or course. RMP listed overall scores for 1007 instructors at our university at the time we extracted the data. We computed correlation coefficients between easiness, helpfulness, clarity, hotness, interest, overall score, instructor gender, course discipline, and common themes identified in our qualitative analysis.

Collecting and analysing the qualitative data – students' written comments – was a laborious task, but perhaps the most informative. Not restrained by the website's categories, students are free to post comments about teaching qualities that matter to them. We believe these comments better indicate what students value in an instructor and in a course.

Among the 1007 instructors at our university, some have well over 50 separate comments, with some comments several sentences long. Analysing all of these RMP postings proved too time consuming. Instead, we drew a two-stage systematic random sample. First, we copied all of the RMP postings for half of all instructors at our university. From that list, we then drew every seventh posting, ending in 1166 comments to analyse. After

reading all the comments several times, we identified 13 common themes into which we then coded the data for prevalence: entertainment, ease, politics, informative, student centredness, persona, preparedness, discrimination, and intellect (see Appendix 1 for descriptions). We then documented whether each comment coded was positive or negative in meaning.

The last step in our analysis was to compare the frequency of the 13 themes across sociology instructors and those in economics, political science and psychology. We selected economics, political science, and psychology because sociology is commonly compared with these disciplines. To ensure an adequate sample size for each discipline, we went back to the RMP website and copied every fifth posting for every instructor listed in these four disciplines, ending in a sample size of 467 comments. We then selected excerpts from the comments to highlight the patterns across disciplines.

## Results

### *Awareness and utilisation: student survey results*

As seen in Figure 1, 92% of the 216 surveyed students have heard of the RMP website. Students usually hear about the website from friends or classmates (88%). Most (80%) have visited the website more than once, but only around 30% of students claim to actually post ratings. Most of the respondents deem the website helpful (84%) and credible (95%). Three-quarters of the students use the website to decide whether to take a particular instructor or not, while only a third use it to select or avoid a course completely, regardless of instructor. Since the popularity of this website has grown in the last three years, we expect a similar survey would indicate even more usage today.

As observed in Figure 2, 55% to 75% of students discern instructor's interest, instructor knowledge and how much they learned in a course as important characteristics in selecting an instructor. Unfortunately, these indicators are not part of the RMP website. Whereas the website does ask students to rate easiness, hotness and their own interest in the course topic, these measures are reported as far less informative to students. Having learned that most of

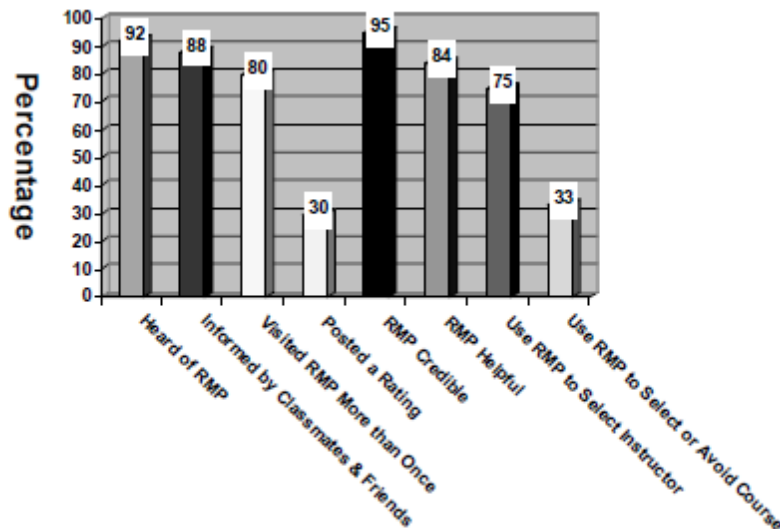


Figure 1. Awareness and utilisation ( $n = 216$ ).

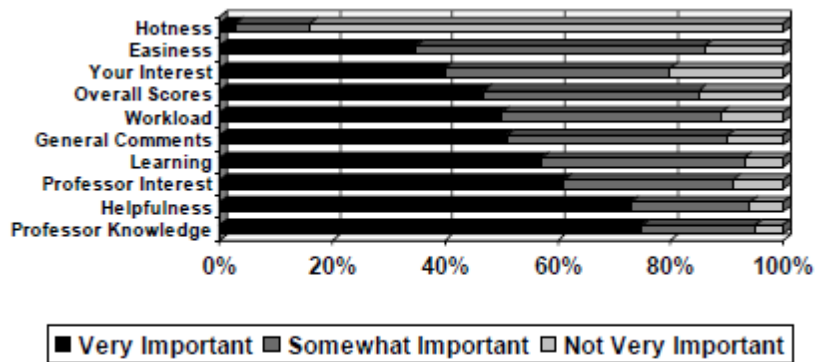


Figure 2. Importance of RMP categories ( $n = 216$ ).

our students did in fact use the RMP website, we moved on to the next phase in our study: analysing quantitative and qualitative data from the RMP website to assess the validity of using these data in decision-making.

### **Quantitative analysis of overall score**

As mentioned before, RMP posts a prominent overall score beside each instructor's name. The only ratings that calculate into this score are clarity and helpfulness. A significant correlation of 0.85 exists between clarity and helpfulness which denotes that these two characteristics are not two distinct measures in the eyes of students. The significant relationship between easiness and overall score ( $r = 0.51$ ) conveys that the easier the course, the higher the overall score. This finding affirms the same relationship found between easiness and overall score in the Felton et al. 2004 study ( $r = .61$ ) and the Coladarei and Kornfield 2007 study ( $r = .40$ ).

The significant inverse relationship between number of ratings and overall score ( $r = -0.13$ ) implies that students are more likely to post ratings on the website to whine or complain about instructors. For the postings where students rated their interest in the course content (a measure added recently to the RMP website), a significant and positive correlation ( $r = 0.30$ ) shows that the greater the interest in the subject matter, the greater the overall score of the instructor. In our study we identified the gender of each instructor to control for possible gender biases. The correlation coefficients show that gender does not have a bearing on overall scores (see Table 1). However, gender is significantly associated with certain personality traits as will be discussed in the qualitative analysis below.

### **Qualitative analysis of RMP postings**

We found some common themes among the open student comments that students included with each posting. For each theme we present the percentage of comments that are negative or positive about each characteristic. We also indicate if the trait significantly correlates to overall score.

As reflected in Figure 3, the most common theme (45% of all comments) in the students' qualitative remarks entails easiness. Students want easy courses and reward instructors accordingly. They prefer low workload demands (little reading and few papers),

Table 1. Pearson correlation coefficients with overall score.

Quantitative scores	Overall score (1 = lowest/bad to 5 = highest/good)
Easiness scale 1 = very hard 5 = very easy	0.508***
Gender 0 = male 1 = female	0.056
Number of ratings	-0.130***
Interest	0.301***
Qualitative comments	
Entertainment	0.102***
Easiness	-0.023
Informative	0.131***
Personality – general	0.027
Student centredness	0.107***
Professor intellect	0.014
Politics	-0.062*
Student chiding	0.033
Sexual	0.079**
Language	-0.082**
Physical traits	0.001
Discrimination/bias	-0.103***
Preparedness	-0.189***

\* – p<.05.  
 \*\* – p<.01.  
 \*\*\* – p<.001.

### Prevalence of Categories

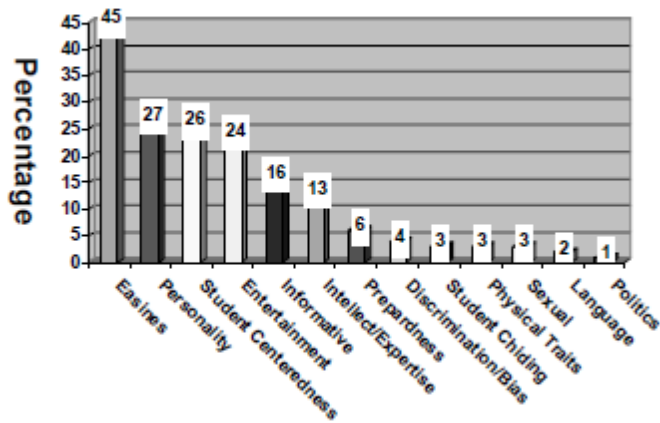


Figure 3. Prevalence of qualitative categories among student comments (n = 1166).

Table 2. Significant gender correlations.

	Gender (0 = male 1 = female)
Easiness	0.071*
Helpfulness	0.069*
Intellectual/expertise	-0.062*
Student centredness	0.086**

cancelled classes and little effort required to get an A. The qualitative comments about easiness do not significantly correlate to the overall score (Table 1). However, comments about easiness are slightly ( $r = 0.07$ ), but significantly, more likely to be made about female instructors (see Table 2).

About 40% of the easiness comments praise instructors for their easiness, as in the comment below:

Extremely easy. All u do is just sit there and talk about crap for an hour straight. U do NOTHING. TAKE THIS CLASS, I SWEAR.

Slightly more of the easiness comments (around 60%) are negative evaluations, alerting students that they will have to study or to avoid an instructor because it is impossible to get an A (a so called 'GPA Buster'). Examples of students' negative evaluations of an instructor's easiness include:

I graduated with a 3.7 overall GPA, and the biggest reason it was that low was THIS class. Stay away.

The quizzes are VERY hard. You must read every single story or you won't do good in the class.

The next two most common themes in students' written comments on the RMP website involve instructor personality in general (27%) and, specifically, student centredness (26%). Most of these comments (about 75%) are positive descriptions. Many students refer generally to an instructor being 'nice' or 'cool', such as in 'She's as cool as a fan on high. I'm thinking of failing on purpose just so I can take her again.' One student reflected, 'I wouldn't mind calling him Dad'. Most negative comments concern an instructor's arrogance.

Student centredness, a specific aspect of instructor personality, deals with how helpful or accessible an instructor is to students, such as taking 'extra time to explain papers and projects'. It refers to the extent to which an instructor cares about students, such as in remembering their names. Student centredness often boils down to how far an instructor is willing to accommodate a student's personal schedule, as in:

He will go out of his way to help. And will always be there to listen.  
I was out sick a lot and he worked with me every step of the way.

When students post negative comments on an instructor's personality it usually relates to the instructor being mean to students publicly, such as 'he makes you feel stupid by saying you need to go to learning assistance when you ask questions'. Or, the instructor 'picks out individuals and tells them in front of everyone that they need to drop the class'.

Whereas general personality statements do not significantly correlate with overall scores, the student centredness comments significantly impact on the overall scores in that the more student centred the comments, the higher the overall score. Correlating student centredness with gender, we discover (see Table 2) that students are slightly more likely to make student centredness comments about females than males (0.09), supporting Sprague and Massoni's (2005) finding that female instructors are expected to be more nurturing.

The fourth most frequent type of comment emphasises how entertaining an instructor is during class time (24%). Given a generation raised on TV and video games, we are surprised that entertainment remarks are not more common. Entertainment comments are significantly related to overall score ( $r = 0.10$ ). The presence of an entertainment statement is likely to lead to a higher overall score. Most of the entertainment comments (66%) are positive, often praising an instructor for being enthusiastic and spontaneous, such as:



He shows an amazing enthusiasm for what he does. As a result, the students in his class have no choice but to be equally enthusiastic.  
Really fun class. He gets a little off the subject sometimes, but his stories are always entertaining.  
I couldn't wait to go to class.

Negative comments on entertainment often stress how an instructor bores students such as 'the answer to insomnia' and 'make sure you bring a Coke, and a snack to stay awake'. Students reveal on the RMP website that they consider PowerPoint presentations dull, such as 'The man could put a crack addict to sleep. PPT presentations ... coma-inducing torture'. Unlike, Sprague and Massoni (2005), we did not identify a gender effect with entertainment comments, which signifies that students expect entertainment from both male and female instructors.

Less common themes in the students comments include how much, or if, students learn anything from an instructor (16%), an instructor's expertise or intellect (13%), and instructor preparedness (or lack thereof) for class (6%). Few students accuse instructors of discrimination against them (4%) or being politically biased (1%).

There were three unexpected trends: rarely do students sexualise instructors with their comments on RMP (3%) or complain about instructors who do not speak English as their primary language (2%); and, students occasionally chide one another (2%) with comments such as, 'This is college guys, not middle school'. Of these other factors, the more informative the course, the less political the content, the more sexually appealing the instructor, and the fewer issues with language, bias or instructor preparedness, the significantly higher the overall score.

### ***Ratings across disciplines***

We next compare sociology instructors with those in three other closely related disciplines (economics, political science and psychology) across the 13 themes found in the RMP postings (see Table 3).

There is not a significant difference in the number of easiness comments across the four disciplines. However, students are more likely to comment about the entertainment value, political orientation and bias of sociology instructors than of instructors in economics, political science, or psychology. The comments about the political orientation and bias of sociology instructors imply that sociologists, more so than the instructors in the other disciplines, are perceived to be indoctrinating students. The excerpts below reinforce this contention:

I think if you have conservative values or if you are a Christian you should avoid his [a sociologist] classes.

Table 3. Department comparisons: counts.<sup>1</sup>

Factor	Sociology <i>n</i> = 122	Economics <i>n</i> = 83	Political Science <i>n</i> = 77	Psychology <i>n</i> = 185	<i>x</i> <sup>2</sup>	df	<i>p</i>
Entertainment	63 (52)	22 (27)	28 (36)	67 (36)	8.91	3	0.031
Ease	59 (48)	47 (57)	39 (51)	74 (40)	3.84	3	0.279
Informative	24 (20)	4 (5)	22 (29)	41 (22)	13.1	3	0.004
Personality – general	33 (27)	19 (23)	19 (25)	51 (28)	.580	3	0.901
Student centredness	18 15	21 25	15 19	45 24	4.01	3	0.261
Professor intellect	10 (8)	9 (11)	17 (22)	30 (16)	7.69	3	0.053
Politics	10 (8)	0 (0)	3 (4)	0 (0)	20.6	3	0.000
Student chiding	5 (4)	3 (4)	3 (4)	2 (1)	3.23	3	0.357
Sexual	2 (2)	1 (1)	5 (6)	13 (7)	7.53	3	0.057
Language	1 (1)	5 (6)	0 (0)	3 (2)	9.57	3	0.023
Preparedness	0 (0)	0 (0)	2 (3)	10 (5)	11.1	3	0.011
Physical traits	6 (5)	2 (2)	2 (3)	5 (3)	1.51	3	0.681
Discrimination	13 (11)	0 (0)	5 (6)	6 (3)	13.1	3	0.005

Notes: <sup>1</sup>Percentages are given in parentheses. \**p* < 0.05; \*\**p* < 0.01; \*\*\**p* < 0.001.

This guy uses his class as a soap box for his liberal anti-American garbage. It appears as though he has a hidden agenda to brainwash students.... I hope John Ashcroft is watching him and arrests his ass for treason.

A strong leftist and persistent to convince class of his liberal views. Beware of the brainwashing.

At the core of most sociology courses are inequality issues that challenge students to think critically of structural systems and processes that benefit positions of privilege. These explanations are not always widely accepted among students, causing the sociological perspective to be more criticised and politicised than other disciplines. Students commonly reflect upon the nature of sociology negatively, as in:

Her classes are very femi-Nazi! If you aren't a hard core feminist, avoid like the plague!  
He can be very demeaning to anyone who's not white, male, and straight.

## Discussion

Our study finds that most students at our university are aware of the RMP website, most visit it, most think it is credible, and, alarmingly, most students use it to choose instructors. Although the goal of the RMP website is to generate profit rather than education, the fact is that the website is taken very seriously by students, primarily because there are few other online alternatives, in most institutions of higher education, for exchanging information about courses and instructors.

Although information concerning courses at most colleges and universities is gathered from students through the end of the semester evaluations, the data are normally not available to those students. Since students apparently have a strong desire to exchange information

about their instructors and courses, and have few alternatives to publicly do so, we are not surprised by the popularity of the RMP website. However, our findings suggest that the information provided by the RMP website is not valid.

A huge problem with the site is the lack of external validity. There is no way to control who posts a message (whether they actually took an entire course with the instructor) or when a message is posted (we found students posting on the first day of course or as a 10-year alumnus). Although RMP claims to have a person monitoring the ratings for each school, we found that many comments violate RMP's guidelines. RMP is not effective in discouraging comments about physical traits, especially since a selling point of the website is to rate an instructor's sexual appeal (see Felton et al. 2004 for the importance of sexual hotness ratings on overall scores). We detect student chiding as the only real quality control on the RMP website.

Moreover, the overall score that RMP computes for each instructor is a combination of two highly correlated variables (helpfulness and clarity), and is void of any measure concerning the learning process or knowledge attained. Nevertheless, RMP prominently displays a smiling or frowning face symbol, accordingly, next to each instructor name. The lack of comprehensive measures of teaching effectiveness calls into question the internal validity of the information provided by the website. Few would argue that the essence of teaching effectiveness is captured by the measures of helpfulness, easiness, clarity, interest in subject and sexual appeal (the current rating categories offered by the website). Yet, despite the obvious validity issues, our study finds that students inaccurately perceive this overall score as a measure of teaching effectiveness. Thus, it is difficult to compare official institutional evaluations collected at the end of every course with the RMP site since there are so few measures in common (see Coladarci and Kornfield 2007 for their comparison of RMP evaluations with institutional evaluations).

When commenting about instructors in open-ended statements, students do not offer many rigorous indicators of teaching effectiveness. In our evaluation of qualitative remarks, almost half (46%) of the students refer to the easiness aspect of a course. We show that easiness strongly correlates with overall score; the easier the course, the higher the overall score. The emphasis on easiness suggests that students today are not interested in the learning process or the end product of knowledge. Rather, they seek the path of least resistance to obtaining a college degree. Since easiness is rarely measured on institutional evaluations, we see a disconnect between student, professor and administrative expectations for courses and higher education.

This finding upholds a growing body of literature which argues that college students today adopt an increasingly consumerist goal for their education (Baker and Copp 1997; Delucchi and Smith 1997a, 1997b; Shepperd 1997; Smith 2000). Students perceive themselves primarily as customers (Higher Education Research Institute 2000), purchasing a product: a degree. The fact that most of the comments on student centeredness describe an instructor's willingness to accommodate a student serves as more evidence of students' emphasis on customer satisfaction. Even the positive comments are consumerist, such as 'he works for the students'.

As Delucchi and Smith (1997a) and Delucchi and Korgen (2002) note, we can hardly change the students' consumerist values and expectations upon entering college. Websites like Rate My Professor will continue to cater to these demands. In a consumerist environment, though, student evaluations are not 'good' data. They measure how easy the instructor is, how fun, and sometimes, as in the case of the Rate My Professor website, how sexy he or she is. Such data should not be used by students or organisations to evaluate an instructor's ability to teach.

Instead, we urge colleges and universities to develop their own RMP websites. Well designed

student evaluations of teaching can be informative, and motivate necessary pedagogical change (Gallagher 2000). Questions should focus on learning (Delucchi 2000; Fries and McNinch 2003), reinforce the purpose and values of the educational organisation, and be relevant to the discipline. In our student survey, students rank the more serious academic factors (such as how much they learn and instructor expertise) as the most important information, leading us to believe that if principal measures of teaching effectiveness are provided, students will utilise and benefit from these measures. Building upon Basow (2000), Koeber (2005) develops such a model which includes six dimensions of teaching effectiveness that students can measure: course design, rapport, grading quality, course value, difficulty, and workload.

We recommend institutionalising Koeber's (2005) model online, to collect data quickly and efficiently, and with data summaries immediately available to students. Such a strategy would probably displace websites like Rate My Professor, at least at the college level. Technical controls would improve data quality, such as only allowing registered students to evaluate a course during the last week of the term via their student email account. Several schools in the United States are testing similar models. We are currently encouraging our university to follow suit.

## Notes

1. John Swapceinski started RMP. Currently, Will DeSantis and Patrick Nagle, co-founders of the SwitchTextbooks.com business, serve as CEO and COO, respectively. RMP recently merged with Rate My Teacher (a website that offers ratings of middle and high school teachers) and intend to compete with FaceBook and MySpace, two other popular websites targeting young people.
2. RMP offers explanations of each category on its website (<http://www.ratemyprofessors.com/faq.jsp> and [http://www.ratemyprofessors.com/rater\\_guidelines.jsp](http://www.ratemyprofessors.com/rater_guidelines.jsp)).
3. According to ratemyprofessors.com, our department faculty at one time included classical Sociologists Georg Simmel and Emile Durkheim. One student considered Durkheim 'so hot, he makes me melt when I see him in the halls...', while another commented, 'good class, great guy, but I don't get the stuff about totems'. We also found evidence of colleagues either rating themselves or each other, sometimes in jest but also, perhaps, in spite.
4. In contrast, Rau and Durand (2000) found an average of 25 hours of study per week among students at the University of Michigan. Clearly there is an institutional effect, with more competitive schools attracting the more academically engaged student.
5. More than 600 instructors are listed on the RMP website for Appalachian State University, because an instructor's name will remain even if the instructor moves to another institution.

## References

- Abrami, P.C., and D.A. Mizener. 1985. Student/instructor attitude similarity, student ratings, and course performance. *Journal of Educational Psychology* 77: 693–702.
- Arbuckle, J., and B.D. Williams. 2003. Students' perceptions of expressiveness: age and gender effects on teacher evaluations. *Sex Roles: A Journal of Research* 49 (9–10): 507–516.
- Baker, P., and M. Copp. 1997. Gender matters most: the interaction of gendered expectations, feminist course content, and pregnancy in student course evaluations. *Teaching Sociology* 25, no. 1: 29–43.
- Basow, S.A. 2000. Best and worst professors: gender patterns in students' choices. *Sex Roles* 43, nos 5–6: 407–417.
- . 1998. Student evaluations: the role of gender bias and teaching styles. In *Arming Athena: Career strategies for women in academe*, eds L.H. Collins, J.C. Chrisler, and K. Quina, 135–

156. Thousand Oaks, CA: Sage.

Benson, D.E., and J.M. Lewis. 1994. Students evaluations of teaching and accountability: implications from the Boyer and the ASA reports. *Teaching Sociology* 22: 195–199.

Blunt, A. 1991. The effects of anonymity and manipulated grades on student ratings of instructors. *Community College Review* 18: 48–54.

Centra, J.A., and N.B. Gaubatz. 2000. Is there gender bias in student evaluations of teaching? *Journal of Higher Education* 71: 17–33.

Coladarci, T., and I. Kornfield. 2007. RateMyProfessors.com versus formal in-class student evaluations of teaching. *Practical Assessment Research and Evaluation* 12: 6.

Das, M., and H. Das. 2001. Business students' perceptions of best university professors: does gender role matter? *Sex Roles* 45, nos 9–10: 665–676.

Delucchi, M. 2000. Don't worry, be happy: instructor likability, student perceptions of learning, and teacher ratings in upper-level sociology courses. *Teaching Sociology* 28, no. 3: 220–231.

Delucchi, M., and K. Korgen. 2002. We're the customer – we pay the tuition: student consumerism among undergraduate sociology majors. *Teaching Sociology* 30, no. 1: 100–107.

Delucchi, M., and W.L. Smith. 1997a. A postmodern explanation of student consumerism in higher education. *Teaching Sociology* 25, no. 4: 322–327.

———. 1997b. Satisfied customers versus pedagogic responsibility: further thoughts on student consumerism. *Teaching Sociology* 25, no. 4: 336–337.

Dowd, J. 1988. Sociology is different: the misevaluation of teaching effectiveness. *Sociological Inquiry* 58: 393–413.

Ewing, V.L., A.A. Stukas Jr, and E.P. Sheehan. 2003. Student prejudice against gay male and lesbian lecturers. *Journal of Social Psychology* 143, no. 5: 569–579.

Feldman, K.A. 1993. College students' views of male and female college teachers, Part II: evidence from the students' evaluations of their classroom teachers. *Research in Higher Education* 34: 151–211.

Felton, J., J. Mitchell, and M. Stinson. 2004. Web-based student evaluations of professors: the relations between perceived quality, easiness and sexiness. *Assessment and Evaluation in Higher Education* 29: 91–108.

Fries, C.J., and J.R. McNinch. 2003. Signed versus unsigned student evaluations of teaching: a comparison. *Teaching Sociology* 31, no. 3: 333–344.

Gallagher, T.J. 2000. Embracing student evaluations of teaching: a case study. *Teaching Sociology* 28, no. 2: 140–147.

Goldman, L. 1990. Student evaluations of their professors rarely provide a fair measure of teaching ability. *Chronicle of Higher Education* August 8, B2.

Higher Education Research Institute. 2000. *The American freshman: national norms for fall 1999*. Los Angeles: University of California, Los Angeles.

Kemmelmeier, M., C. Danielson, and J. Basten. 2005. What's in a grade? Academic success and political orientation. *Personality and Social Psychology Bulletin* 31, no. 10: 1386–1399.

Koeber, C. 2005. Introducing multimedia presentations and a course website to an introductory sociology course: how technology affects student perceptions of teaching effectiveness. *Teaching Sociology* 33: 285–300.

Lucal, B., C. Albers, J. Ballantine, J. Burmeister-May, J. Chin, S. Dettmer, and S. Larson. 2003.

Faculty assessment and the scholarship of teaching and learning: knowledge available/knowledge needed. *Teaching Sociology* 31, no. 2: 146–161.

McKeachie, W.J. 1994. *Teaching tips: strategies, research and theory for college and university professors*. 9th ed. Washington, DC: Heath.

Naftulin, D.H., J.E. Ware, and F.A. Donnelly. 1973. The Doctor Fox lecture: a paradigm of educational seduction. *Journal of Medical Education* 48: 630–635.

Rate My Professor. July 10, 2006. <http://www.ratemyprofessors.com>

Rau, W., and A. Durand. 2000. The academic ethic and college grades: does hard work help students to make the grade? *Sociology of Education* 73, no. 1: 19–38.

Schroeder, D.S., and C.R. Mynatt. 1999. Graduate students' relationships with their male and female major professors. *Sex Roles* 40: 393–420.

Shepperd, J.W. 1997. Relevance and responsibility: a postmodern explanation of student consumerism in higher education. *Teaching Sociology* 25, no. 4: 333–335.

Sinclair, L., and Z. Kunda. 2000. Motivated stereotyping of women: she's fine if she praised me but incompetent if she criticized me. *Personality and Social Psychology Bulletin* 26, no. 11: 1329–1342.

Smith, W.L. 2000. Teaching in a consumeristically charged environment. *Michigan Sociological Review* 14, fall: 58–72.

Smith, W.L., and N.W. Pino. 2005. College students, the academic ethic, and academic achievement. *Virginia Social Science Journal* 40, no. winter: 52–67.

Sprague, J., and K. Massoni. 2005. Student evaluations and gendered expectations: what we can't count can hurt us. *Sex Roles: A Journal of Research* 53, no. 11–12: 779–793.

Talley, K., and D. Timmer. 1992. A quantitative methods exercise: student and faculty interpretations of classroom teaching. *Teaching Sociology* 20: 75–79.

Wachtel, H.K. 1998. Student evaluation of college teaching effectiveness: a brief review. *Assessment and Evaluation in Higher Education* 23, no. 2: 191–211.

Wagenaar, T. 1995. Student evaluations of teaching: some cautions and suggestions. *Teaching Sociology* 23: 64–68.

Williams, W.M., and S.J. Ceci. 1997. How'm I doing: problems with student ratings of instructors and courses. *Change* 29, no. 5: 12–23.

## **Appendix 1: Qualitative coding categories**

*Entertainment*: Indications of whether or not and how the course/instructor is entertaining during class.

*Ease*: Indications of whether or not and how difficult the instructor/course is, usually focusing on the likelihood of yielding an effortless A.

*Politics*: Indications of whether or not and how the professor conveys his or her political identity – this was always concerning the professor's liberal identity. Though it could have included other political perspectives.

*Informative*: Indications of whether or not and how the course/instructor provides valuable

or relevant information.

*Professor's student centredness:* Indications of whether or not and how the instructor is concerned about student welfare or needs.

*Professor persona:* Indications of the instructor's personality – such as mean, easygoing, nice, angry, etc....

*Preparedness:* Indications of how well the instructor prepared or organised the course and/or lectures.

*Discrimination/Bias:* Indications of whether or not and how the instructor/course treats students unequally concerning sex, race, gender, sexuality, religion.

*Professor intellect/expertise:* Indications of the instructor's knowledge of the topic.

*Student chiding:* Indications of students policing, reprimanding, or correcting previous student postings.

*Sexual:* Indications of the professor's good looks or sexual innuendos.

*Physical traits:* Indications of the professor's gender, disabilities, race, ethnicity or age.

*Language:* Indications of whether or not English is the instructor's native language.