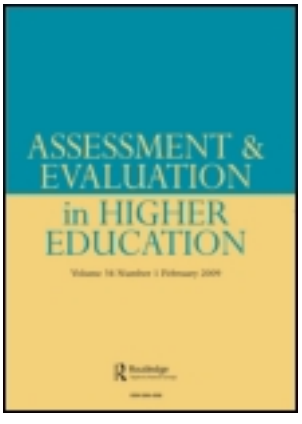

Assessment & Evaluation in Higher Education

Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/caeh20

Web-based student evaluations of professors: the relations between perceived quality, easiness and sexiness

James Felton [a], John Mitchell [a] & Michael Stinson [a]

[a] Central Michigan University, USA

Available online: 14 Sep 2010

To cite this article: James Felton , John Mitchell & Michael Stinson (2004): Web-based student evaluations of professors: the relations between perceived quality, easiness and sexiness, Assessment & Evaluation in Higher Education, 29:1, 91-108

To link to this article: http://dx.doi.org/10.1080/0260293032000158180

Assessment & Evaluation in Higher Education
Vol. 29, No. 1, February 2004

Carfax Publishing
Taylor & Francis Group

Web-based student evaluations of professors: the relations between perceived quality, easiness and sexiness

James Felton*, John Mitchell & Michael Stinson
Central Michigan University, USA

College students critique their professors' teaching at RateMyProfessors.com, a web page where students anonymously rate their professors on Quality, Easiness, and Sexiness. Using the self-selected data from this public forum, we examine the relations between quality, easiness, and sexiness for 3190 professors at 25 universities. For faculty with at least ten student posts, the correlation between quality and easiness is 0.61, and the correlation between quality and sexiness is 0.30. Using simple linear regression, we find that about half of the variation in quality is a function of easiness and sexiness. When grouped into sexy and non-sexy professors, the data reveal that students give sexy-rated professors higher quality and easiness scores. If these findings reflect the thinking of American college students when they complete in-class student opinion surveys, then universities need to rethink the validity of student opinion surveys as a measure of teaching effectiveness. High student opinion survey scores might well be viewed with suspicion rather than reverence, since they might indicate a lack of rigor, little student learning, and grade inflation.

Introduction

In 1999, RateMyProfessors.com opened a free web site where college students can rate their professors' teaching quality using an ordinal scale from 1 (worst) to 5 (best) for Clarity and Helpfulness. The web site then calculates an overall Quality (of teaching) score by averaging the Clarity and Helpfulness data for each professor. Students also evaluate their professors in other ways: rating them on Easiness from 1 (hardest) to 5 (easiest), whether or not they are sexy, and through written comments up to 255 characters long. All student posts are anonymous, voluntary and publicly available. The web site had its one-millionth student post on 26 August, 2003. At that time, the site had ratings for 223,746 faculty at 3310 universities in the USA and Canada.

In many universities, it has become standard practice to give out Student Opinion Surveys (SOS) at the end of every semester. These surveys are assumed to measure a professor's performance and success in his or her various classes. While Rate-

*Corresponding author: Department of Finance & Law, Central Michigan University, Mount Pleasant, Michigan 48859, USA. Email: james.felton@cmich.edu

ISSN 1234-5678 (print)/ISSN 8765-4321 (online)/04/010091-18

MyProfessors.com is not a university SOS, it is certainly a survey conveying the opinions of an enormous student body, and we believe that results obtained from RateMyProfessors.com expose larger implications about how students judge professors on university-sponsored SOS forms. For our analysis, we downloaded the scores posted on RateMyProfessors.com for over 3000 professors at 25 universities in the USA, and we examined the relations between teaching quality, course rigor and the sexiness of faculty members.

Student opinion survey scores affect the careers of thousands of college professors in the USA. According to Yining and Hoshower (2003), SOS scores are used by the academy as the most important factor in determining teaching effectiveness, and students rarely know that these forms function in this way or carry such weight. The ratings at RateMyProfessors.com are entirely self-selected, and we realize that some students go to the web site to 'slam' or praise professors in order to affect the professors' scores. However, our sample data are from over 65,000 student posts, and the data give us an opportunity to assess the factors influencing students' rating practices using a nation-wide sample. Our research question is this: when college students evaluate the teaching effectiveness of their professors, are they unbiased, or are they biased by such peripheral factors as course workload, course grades and the appearance of the instructor?

Literature review

Much study has been devoted to factors (outside of successful teaching strategies) that can affect students' evaluations of instructional quality. Greenwald (1997), McKeachie (1997), and Stapleton and Murkison (2001) provide reviews for much of this literature. Several authors find relationships between professorial grading standards and student evaluations. Among these authors, Tang (1999) finds that the differences between expected and received grades affect evaluations. As observed by Engdahl, Keating and Perrachione (1993), student evaluations of their professors dropped after their professors explicitly and summarily judged the students' course performances. Stumpf and Freedman (1979) and Hudson (1984) likewise indicate that grading leniency is related to higher evaluations of teacher quality. And Cerrito (2000) notes a positive correlation between a student's satisfaction with the instructor and his or her expected grade.

Perhaps the strongest evidence of a significant relationship between high grades and higher student satisfaction is presented by Greenwald and Gillmore (1997a, b). They report an average $+0.45$ correlation between expected grade and course evaluations. Greenwald and Gillmore (1997a) note the following five theories regarding the possible relationship between evaluations and grades:

(1) Teaching Effectiveness influences grades and ratings. More effective teachers have reason to give higher grades and students recognize the teacher's effectiveness.
(2) Student general motivation affects grades and ratings. Better students give higher evaluations.

(3) Student course-specific motivation affects grades and ratings. Students with an interest and therefore higher performance in a specific course give higher evaluations.
(4) Students infer course quality and their own ability from received grades.
(5) Students give high ratings in appreciation for lenient grading.

Greenwald and Gillmore note that although the significance of this high-ratings-for-lenient-grading theory is largely disputed in the literature, it best fits their data.

Other research indicates additional factors affecting student evaluations. McKeachie (1997) counters Greenwald and Gillmore, stating that although leniency affects ratings, it may backfire if undeserved, and it depends on organization culture. The challenge, or rigor, of the course has also been linked to evaluations. Although Heck, Todd and Finn (2002) find that the degree of challenge in a course has little impact on evaluations, Greenwald and Gillmore (1997a, b) find a 0.33 correlation between evaluations and challenge; they also note a -0.11 correlation between evaluations and hours of homework per credit hour. Further, Greenwald and Gillmore express surprise at finding a negative relationship (an average -0.49 correlation) between workload and expected grade.

Trout (1997, 2000) offers anecdotal evidence of instructors lessening the degree of course challenge in an attempt to bolster students' evaluations of their course; this evidence comes largely from universities that rely heavily on student evaluations for faculty promotion and tenure decisions. Trout's argument is echoed by Ryan (1980), who shows that, as a result of university's using teaching evaluations for promotion and tenure, 22% of instructors reduced material covered (7% increased) and 40% made exams easier (9% harder).

Immediacy has also been cited as a significant factor in teaching effectiveness. The Principle of Immediacy, as developed by Mehrabian (1971), states that people are drawn towards persons and things they like and give them high evaluations. Rocca and McCroskey (1999) found positive correlations between Immediacy and three attractiveness categories: Task Attractiveness (desire to work with the subject matter) (0.36), Physical Attractiveness (appearance of the instructor) (0.45), and Social Attractiveness (desire to socialize with the instructor) (0.58). Attitude Similarity (0.58) was found to be more important than Background Similarity (0.22) in increasing attractiveness.

Feeley (2002) finds a positive correlation between physical attractiveness and effectiveness ratings (0.40). He attributes this, at least in part, to a Halo Effect, which can be described as a 'can do no wrong' perception on the part of the rater. Halo Effects have also been attributed to affecting evaluations of career success, satisfaction, and tasks by Dion, Berscheid and Walster (1972), and Landy and Sigall (1974). Other research by Moritsch and Suter (1988) finds a Halo Effect in teaching evaluations, and Pike (1999) notes a Halo Effect in other student evaluations. When attractiveness impacts measures of effectiveness, Nussbaum (1992) and Rice, Stewart and Hujber (2000) conclude that a Halo Effect is the cause.

Cooper (1981) discusses four reasons for a Halo Effect: one factor may overwhelm (Engulf) other factors; the rater may not have enough information (Under-

sampling) to make an informed decision; rating factors are not well-specified and subject to generalization (Insufficient Concreteness); and Carelessness, both in terms of time and effort, in the rating process. Feeley (2002) and d'Apollonia and Abrami (1997) claim that Halo is not necessarily bad, because they find a positive correlation to overall accuracy.

The data

The rating categories for RateMyProfessors.com (2003) are described at the web site as follows:

> **Easiness**—This is definitely the most controversial of the three rating categories, which is why it is NOT included in the 'Overall Quality' rating. Although we do not necessarily condone it, it is certainly true that many students decide what class to take based on the difficulty of the teacher. When rating a teacher's easiness, ask yourself 'How easy are the classes that this professor teaches? Is it possible to get an A without too much work?'
>
> **Helpfulness**—This category rates the teacher's helpfulness and approachability. Is the teacher approachable and nice? Is he rude, arrogant, or just plain mean? Is he willing to help you after class?
>
> **Clarity**—This is the most important of the three categories, at least to most people. How well does the teacher convey the class topics? Is he clear in his presentation? Is he organized and does he use class time effectively?
>
> **Overall Quality**—The Overall Quality rating is the average of a teacher's Helpfulness and Clarity ratings, and is what determines the type of 'smiley face' that the Professor receives. Due to popular demand, a teacher's Easiness rating is NOT used when computing the Overall Quality rating, since an Easiness of 5 may actually mean the teacher is TOO easy.

The 25 institutions (universities, colleges, and community colleges) with the most student posts at RateMyProfessors.com on 25 November, 2002 are listed in Table 1. These institutions vary considerably from one another in a variety of ways. Based on Carnegie Classification, the institutions cover a wide range, from doctoral extensive universities (University of Delaware, University of Maine, Boston University, and Case Western Reserve University) to an Associate's College (Grand Rapids Community College). The total student posts vary from 19,223 at Grand Valley State University to 1014 at Kenyon College. The number of faculty listed at RateMyProfessors.com varies from 1553 at the University of Delaware to 130 at Lebanon Valley College, and enrollments vary from 38,600 at the University of Central Florida to 1500 at Kenyon College. Based on posts per faculty member and posts per 100 students, RateMyProfessors.com is the most popular with students at Grand Valley State University and Pace University.

On 25 November, 2002, we downloaded data from RateMyProfessors.com for all of the professors at these 25 institutions who had at least ten student posts, giving us a self-selected sample of 3190 faculty with a total of 65,678 posts. The data are not equally weighted by university. Instead, they include 631 faculty from Grand

Table 1. The 25 universities, colleges, and community colleges with at least 1000 posts at RateMyProfessors.com on November 25, 2002 (ranked by posts)

University, College, or Community College	Carnegie Classification	Posts	Faculty	Enrollment	Posts Per Faculty	Posts Per 100 Students
Grand Valley State Univ.	Masters I	19,223	1303	18,500	14.75	103.91
Univ. of Delaware	Doctoral Ext.	14,947	1553	18,000	9.62	83.04
Pace Univ.	Doctoral Int.	9451	723	8900	13.07	106.19
James Madison Univ.	Masters I	6345	849	15,000	7.47	42.30
Towson Univ.	Masters I	5221	975	16,000	5.35	32.63
San Diego State Univ.	Doctoral Int.	5143	1358	32,600	3.79	15.78
Univ. of N. C. Charlotte	Masters I	5142	807	18,900	6.37	27.21
N.W. Missouri State Univ.	Masters I	3885	320	6200	12.14	62.66
Univ. of Maine	Doctoral Ext.	3632	553	11,100	6.57	32.72
Grand Rapids Comm. Col.	Associate's Col.	2565	371	12,400	6.91	20.69
Marist College	Masters I	2420	413	4800	5.86	50.42
Central Michigan Univ.	Doctoral Int.	2307	663	19,400	3.48	11.89
Kettering Univ.	Engineering	2116	174	3200	12.16	66.13
Boston Univ.	Doctoral Ext.	1981	567	29,100	3.49	6.81
Christopher Newport Univ.	Liberal Arts	2090	245	5600	8.53	37.32
Millersville Univ. of Penn.	Masters I	1791	341	5700	5.25	31.42
Siena College	Liberal Arts	1445	275	2400	5.25	60.21
Franklin & Marshall Col.	Liberal Arts	1331	206	1900	6.46	70.05
Seton Hall Univ.	Doctoral Int.	1307	347	4800	3.77	27.23
Loyola College(Maryland)	Masters I	1234	279	2100	4.42	58.76
Case Western Reserve Univ.	Doctoral Ext.	1211	259	3600	4.68	33.64
Saginaw Valley State Univ.	Masters I	1163	235	7200	4.95	16.15
Univ. of Central Florida	Doctoral Int.	1066	414	38,600	2.57	2.76
Lebanon Valley College	Masters II	1045	130	1600	8.04	65.31
Kenyon College	Liberal Arts	1014	158	1500	6.42	67.60
Totals, Averages		98,985	13,518	289,100	7.32	34.24

Valley State University, 471 faculty from the University of Delaware, 329 faculty from Pace University, down to 15 faculty from the University of Central Florida.

We calculated an Average Sexiness score for each faculty member by dividing the number of posts noting that the professor is sexy by the total number of posts for a professor. The range for the sexiness score is therefore 0% to 100% and the score represents the percentage of sexy indicators.

Results aggregated by institution, discipline, and Carnegie Classification

The mean Quality, Easiness, and Sexiness scores for the faculty from the 25 institutions are listed in Table 2. Christopher Newport University has the highest Average Quality at 3.718 for 100 professors. It ranks 19th for Average Easiness and 14th for Average Sexiness. Seton Hall University is ranked lowest in Average Quality at 3.082 for 17 faculty. The University of Central Florida is considered easiest (ranked 1st for easiness), while Franklin & Marshall College is ranked the most

difficult (ranked 25th for easiness). The sexiest professors are at Kenyon College and Lebanon Valley College, while the least sexy professors are at the University of Delaware and Grand Valley State University.

Using averages by institution, we find no significant relations between Quality, Easiness, and Sexiness. The correlations (r) and p-values (p) are as follows (N = 25): Quality-Easiness, $r = 0.25$ ($p = 0.22$), Quality-Sexiness $r = 0.28$ ($p = 0.17$), and Easiness-Sexiness, $r = -0.23$ ($p = 0.27$).

The mean Quality, Easiness, and Sexiness scores for the 3,190 faculty from the 25 institutions are ranked in Table 3 by discipline. The highest Average Quality scores appear in Law (4.152), Languages (4.012), and Education (3.869), while the lowest Average Quality scores appear in Engineering (3.267), Computer Science (3.203), and Math (3.193). The easiest disciplines are Humanities (3.925), Education (3.725), and Music (3.478), while the most difficult disciplines are Ethnic Studies (2.777), Accounting (2.771), and Marketing (2.626). The disciplines with the sexiest professors are Medicine and Nursing (25.51), Marketing (10.59), and Languages (8.15), and the least sexy professors are in Math (1.80), Computer Science (1.78), Accounting (1.49), and Humanities (0.00 for only four professors).

Using averages by discipline (N = 31), we find two significant positive correlations. These correlations are as follows: Quality-Easiness, $r = 0.52$ ($p = 0.003$), Quality-Sexiness $r = 0.47$ ($p = 0.008$), and Easiness-Sexiness, $r = 0.06$ ($p = 0.74$). Therefore, students rate professors in the Accounting, Engineering, Computer Science, and Math departments lowest for quality, and they consider courses in these departments to be among the most difficult. Students also consider many of the professors teaching these courses to be the least sexy on campus. At the other end of the spectrum, students rate professors of Law, Languages, and Education as being the highest quality on campus. They are also among the easiest and sexiest.

The mean Quality, Easiness, and Sexiness scores for the 3190 faculty from these institutions are ranked in Table 4 according to the seven Carnegie Classification levels. The faculty at the four liberal arts schools (Christopher Newport University, Siena College, Franklin and Marshall College, and Kenyon College) are rated highest for Average Quality. They are also considered to be the most difficult (ranked 7th in Average Easiness), and they are ranked second for Average Sexiness. Using averages by Carnegie Classification (N = 7), we find that Average Quality and Average Easiness are almost perfectly negatively correlated. The correlations are as follows: Quality-Easiness, $r = -0.96$ ($p = 0.000$), Quality-Sexiness, $r = 0.52$ ($p = 0.19$), and Easiness-Sexiness, $r = -0.51$ ($p = 0.20$).

Results at the individual level

The correlations between Average Quality, Average Easiness, and Average Sexiness (N = 3190) for faculty with at least 10 student posts are listed in Table 5, and they are as follows: Quality-Easiness, $r = 0.61$ ($p = 0.000$), Quality-Sexiness, $r = 0.30$ ($p = 0.000$), and Easiness-Sexiness, $r = 0.17$ ($p = 0.000$). The relation between Quality and Easiness is graphed in Figure 1. For faculty with at least 20 posts (N = 1148), the correlations, which are listed in Table 6, are higher: Quality-Easi-

Table 2. Quality, Easiness, and Sexiness averages by institution (ranked by Average Quality at RateMyProfessors.com. Faculty from 25 institutions with at least ten posts on 25 November, 2002*)

University, College, or Community College	Sample Faculty	Sample Posts	Ave. Quality (Std. Dev.) Rank	Ave. Easiness (Std. Dev.) Rank	Ave. Sexiness (Std. Dev.) Rank
Christopher Newport Univ.	100	1574	3.718 (0.93) 1	2.975 (0.79) 19	4.89 (10.20) 14
Franklin & Marshall Col.	46	620	3.644 (0.89) 2	2.628 (0.72) 25	7.58 (14.72) 7
Loyola College (Maryland)	37	504	3.638 (0.93) 3	3.065 (0.90) 9	8.77 (21.32) 4
Boston Univ.	30	453	3.637 (0.99) 4	3.183 (0.59) 4	7.85 (20.88) 5
Lebanon Valley College	47	748	3.636 (0.84) 5	2.983 (0.80) 18	10.57 (22.67) 2
Siena College	44	618	3.607 (1.00) 6	2.807 (0.73) 22	9.64 (18.54) 3
Saginaw Valley St. Univ.	30	386	3.587 (0.90) 7	3.057 (0.68) 10	3.62 (10.48) 21
Grand Rapids Comm. Col.	88	1561	3.584 (0.90) 8	3.049 (0.76) 11	3.60 (10.74) 22
Kenyon College	29	376	3.555 (0.90) 9	3.035 (0.69) 12	10.61 (22.09) 1
Grand Valley State Univ.	631	16,825	3.549 (0.87) 10	3.030 (0.77) 14	2.88 (10.20) 25
N.W. Missouri State Univ.	152	3278	3.546 (0.89) 11	3.143 (0.77) 5	7.05 (17.62) 8
Univ. of Delaware	471	11,342	3.545 (0.87) 12	3.133 (0.78) 6	3.18 (9.25) 24
Univ. of N.C. Charlotte	169	3082	3.530 (0.94) 13	3.024 (0.81) 15	3.32 (10.34) 23
James Madison Univ.	231	4199	3.529 (0.95) 14	3.070 (0.83) 8	4.06 (13.27) 16
Univ. of Central Florida	15	266	3.507 (1.30) 15	3.633 (0.82) 1	4.05 (11.28) 17
Univ. of Maine	122	2183	3.499 (0.82) 16	3.033 (0.71) 13	4.44 (14.33) 15
Towson Univ.	171	2610	3.496 (1.03) 17	3.099 (0.88) 7	4.01 (12.80) 19
San Diego State Univ.	140	2255	3.437 (1.05) 18	3.008 (0.88) 16	7.71 (18.46) 6
Pace Univ.	329	7666	3.434 (1.01) 19	3.274 (0.85) 2	6.04 (15.71) 9
Ff Central Michigan Univ.	45	634	3.364 (1.00) 20	2.987 (1.00) 17	4.02 (12.69) 18
Marist College	75	1178	3.339 (0.98) 21	2.827 (0.81) 21	5.14 (14.01) 11
Kettering Univ.	88	1667	3.319 (1.02) 22	3.271 (0.91) 3	3.63 (13.42) 20
Millersville Univ. of Penn.	59	792	3.215 (1.04) 23	2.839 (0.79) 20	4.90 (16.64) 13
Case West. Reserve Univ.	24	614	3.154 (1.08) 24	2.792 (0.83) 23	5.10 (15.04) 12
Seton Hall Univ.	17	247	3.082 (1.06) 25	2.700 (0.90) 24	5.47 (9.34) 10
Totals	3190	65,678	3.510 (0.94)	3.070 (0.81)	4.60 (13.57)

*The correlations by university between Quality, Easiness, and Sexiness are as follows (N = 25): Quality-Easiness, r = 0.25 (p = 0.22), Quality-Sexiness r = 0.28 (p = 0.17), Easiness-Sexiness r = −0.23 (p = 0.27).

Table 3. Quality, Easiness, and Sexiness averages by discipline (ranked by Average Quality at RateMyProfessors.com. Faculty from 25 institutions with at least ten posts on 25 November, 2002*)

Discipline	Sample Faculty	Sample Posts	Ave. Quality (Std. Dev.) Rank	Ave. Easiness (Std. Dev.) Rank	Ave. Sexiness (Std. Dev.) Rank
Law	25	536	4.152 (0.80) 1	3.376 (0.87) 6	7.06 (17.38) 5
Languages	109	881	4.012 (0.87) 2	3.465 (0.74) 4	8.15 (16.73) 3
Education	52	994	3.869 (1.05) 3	3.725 (0.83) 2	6.53 (18.90) 9
Medicine & Nursing	7	133	3.857 (0.98) 4	3.271 (1.14) 10	25.51 (39.09) 1
Criminal Justice	30	728	3.800 (0.83) 5	3.237 (0.58) 11	6.63 (17.19) 7
Psychology	194	4274	3.720 (0.89) 6	3.154 (0.78) 13	5.49 (14.64) 14
Political Science	110	2266	3.665 (0.87) 7	3.045 (0.66) 20	6.64 (14.77) 6
English	374	7016	3.646 (0.90) 8	3.123 (0.81) 15	6.35 (15.44) 11
Communications	148	2840	3.585 (0.89) 9	3.276 (0.81) 8	4.66 (14.47) 20
Humanities	4	49	3.575 (0.85) 10	3.925 (0.43) 1	0.00 (00.00) 31
Music	76	1311	3.570 (0.85) 11	3.478 (0.72) 3	5.14 (15.27) 15
History	234	4898	3.551 (0.86) 12	2.921 (0.73) 27	3.45 (11.81) 24
Fine Arts	69	1426	3.544 (0.91) 13	3.117 (0.87) 16	4.85 (13.65) 18
Athletics	12	176	3.508 (1.10) 14	3.375 (1.04) 7	7.70 (17.58) 4
Not Specified	54	861	3.507 (0.90) 15	3.274 (0.84) 9	6.48 (14.89) 10
Business	193	4121	3.499 (0.96) 16	2.997 (0.83) 23	4.83 (12.63) 19
Geography	45	874	3.489 (0.86) 17	3.229 (0.72) 12	4.92 (14.84) 17
Science	329	7343	3.483 (0.91) 18	2.834 (0.77) 28	4.38 (14.03) 21
Sociology	89	1864	3.470 (0.85) 19	3.117 (0.74) 17	2.47 (9.37) 26
Anthropology	34	693	3.462 (1.00) 20	3.024 (0.98) 22	4.18 (12.69) 22
Philosophy	119	2595	3.458 (0.85) 21	3.071 (0.71) 18	5.99 (14.58) 13
Literature	10	215	3.440 (1.00) 22	2.940 (0.98) 25	2.37 (5.27) 27
Social Work	5	98	3.420 (1.32) 23	3.400 (0.83) 5	6.21 (13.88) 12
Ethnic Studies	13	193	3.415 (1.04) 24	2.777 (0.74) 29	6.57 (13.86) 8
Religion	27	388	3.407 (0.95) 25	3.048 (0.68) 19	4.05 (11.37) 23
Economics	142	2707	3.406 (0.96) 26	2.978 (0.80) 24	4.95 (13.76) 16
Marketing	19	333	3.390 (0.85) 27	2.626 (0.82) 31	10.59 (26.51) 2
Accounting	88	2074	3.361 (0.99) 28	2.771 (0.76) 30	1.49 (7.08) 30
Engineering	69	1262	3.267 (0.97) 29	3.030 (0.82) 21	2.77 (10.55) 25
Computer Science	182	3824	3.203 (1.07) 30	3.125 (0.93) 14	1.78 (7.10) 29
Math	328	7580	3.193 (0.93) 31	2.926 (0.79) 26	1.80 (9.17) 28
Totals	3190	65,678	3.510 (0.94)	3.070 (0.81)	4.60 (13.57)

*The correlations by discipline between Quality, Easiness, and Sexiness are as follows (N = 31):
Quality-Easiness, r = 0.52 (p = 0.003), Quality-Sexiness r = 0.47 (p = 0.008), Easiness-Sexiness r = 0.06 (p = 0.74).

Table 4. Quality, Easiness, and Sexiness averages by Carnegie Classification (ranked by Average Quality at RateMyProfessors.com. Faculty from 25 institutions with at least ten posts on November 25, 2002*)

Carnegie Classification	Sample Faculty	Sample Posts	Ave. Quality (Std. Dev) Rank	Ave. Easiness (Std. Dev.) Rank	Ave. Sexiness (Std. Dev.) Rank
Baccalaureate–Liberal Arts	219	3188	3.658 (0.93) 1	2.876 (0.76) 7	7.17 (16.24) 2
Masters Colleges & Univ. II	47	748	3.636 (0.84) 2	2.983 (0.80) 6	10.57 (22.67) 1
Baccalaureate-Associate's	88	1561	3.584 (0.90) 3	3.049 (0.76) 4	3.60 (10.74) 7
Doctoral/Research-Extensive	647	14,592	3.526 (0.88) 4	3.104 (0.77) 3	3.71 (11.38) 5
Master's Colleges & Univ. I	1,555	32,854	3.518 (0.92) 5	3.038 (0.80) 5	3.97 (12.74) 4
Doctoral/Research-Intensive	546	11,068	3.420 (1.03) 6	3.174 (0.88) 2	6.23 (15.98) 3
Spec.-Engineering & Tech.	88	1667	3.319 (1.02) 7	3.271 (0.90) 1	3.63 (13.42) 6
Totals	**3190**	**65,678**	**3.510 (0.94)**	**3.070 (0.81)**	**4.60 (13.57)**

*The correlations by Carnegie Classification between Quality, Easiness, and Sexiness are as follows (N = 7): Quality-Easiness r = −0.96 (p = 0.000), Quality-Sexiness r = 0.52 (p = 0.19), Easiness-Sexiness r = −0.51 (p = 0.20).

Table 5. The correlations between Quality, Easiness, and Sexiness for all faculty with at least 10 posts, N = 3190

		EASINESS	SEXINESS	QUALITY
EASINESS	Pearson Correlation	1	.17**	.61**
	Sig. (2-tailed)	.	.000	.000
SEXINESS	Pearson Correlation	.17**	1	.30**
	Sig. (2-tailed)	.000	.	.000
QUALITY	Pearson Correlation	.61**	.30**	1
	Sig. (2-tailed)	.000	.000	.

**Correlation is significant at the 0.01 level (2-tailed).

ness, r = 0.67 (p = 0.000), Quality-Sexiness, r = 0.31 (p = 0.000), and Easiness-Sexiness, r = 0.23 (p = 0.000). The relation between Quality and Easiness for faculty with at least 20 posts is graphed in Figure 2.

Using simple linear regression with Average Quality as the dependent variable and Sexiness and Easiness as predictors, the model is significant (F = 1,134.77, p = 0.000) for the 3,190 faculty with at least 10 posts, and the Adjusted R^2 for the model is 0.42. For the 1,148 faculty with at least 20 posts, the regression model is

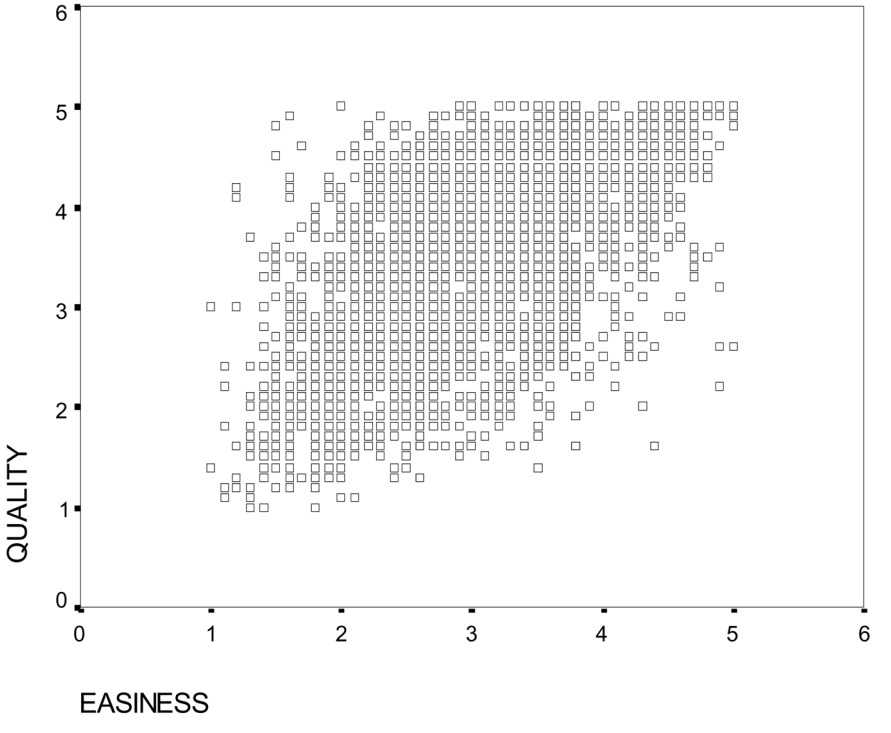

Figure 1. The relation between Quality and Easiness for all faculty with at least 10 posts, N = 3190

Table 6. Correlations between Quality, Easiness, and Sexiness for all professors with at least 20 posts, N = 1148

		EASINESS	SEXINESS	QUALITY
EASINESS	Pearson Correlation	1	.23**	.67**
	Sig. (2-tailed)	.	.000	.000
SEXINESS	Pearson Correlation	.23**	1	.31**
	Sig. (2-tailed)	.000	.	.000
QUALITY	Pearson Correlation	.67**	.31**	1
	Sig. (2-tailed)	.000	.000	.

**Correlation is significant at the 0.01 level (2-tailed).

significant (F = 522.10, p = 0.000), and the Adjusted R^2 for the model is 0.48. For the 520 faculty with at least 30 posts, the regression model is significant (F = 273.54, p = 0.000), and the Adjusted R^2 is 0.51. As such, for the professors in our sample with at least 30 student posts, 51% of the variation in Average Quality is due to Easiness and Sexiness.

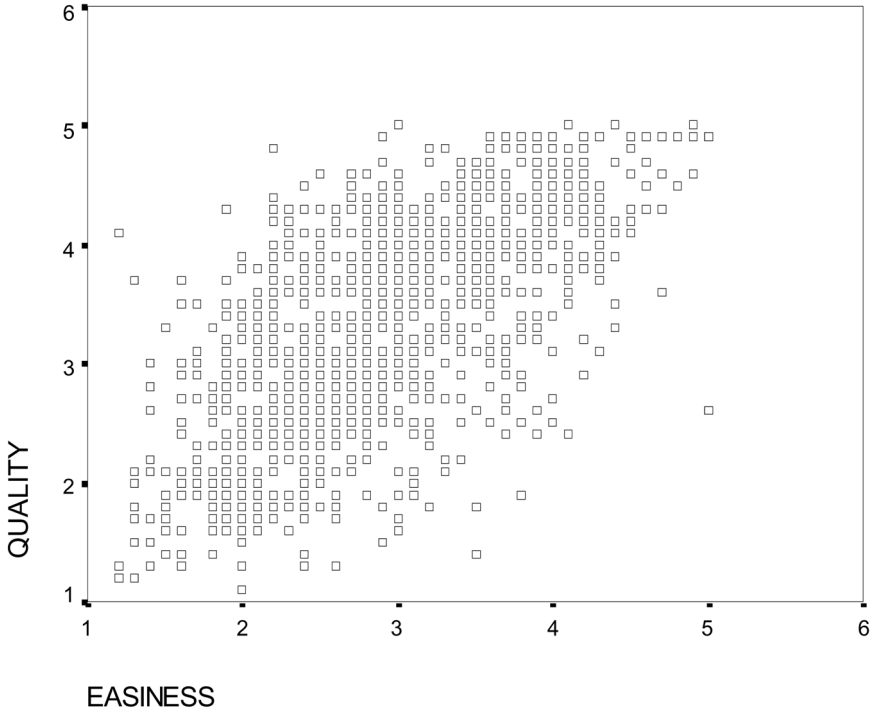

Figure 2. The relation between Quality and Easiness for all professors with at least 20 posts, N = 1148

Student comments

To find additional information regarding how students view their professors' teaching, we read thousands of comments at RateMyProfessors.com. First, we reviewed professors at the top right of Figure 2. Professors charted in this area are considered both high-quality and highly easy by their students. The students' comments are consistent with the scores given these professors, and the following comments from RateMyProfessors.com (2003) are indicative of many student posts regarding these professors: 'Can redo papers as many times as you need. Class is always done early'; 'Awesome teacher. Highly recommended. Cancels many classes and lets you out early'; 'The easiest class I have ever taken'; 'Honest and sincere man, awesome teacher, easy class'; 'Very nice person, down to earth'; 'She's the best, her classes are easy and she's a very sweet person'; 'Show up for class and you'll at least make a B'.

We suspect that many of the professors at the top right of Figure 2 are teaching-award candidates, since their students hold them in such high regard. Occasionally, these professors have student comments such as 'learned a lot,' but the majority of the posts describe the professors as 'nice', 'easy', 'cool', 'caring', 'understanding', 'laid back', or even 'a push over'. The fact that students consistently indicate that they like these professors because of their personalities and easy expectations should not go unnoticed.

Next, we considered student comments for professors at the bottom left of Figure 2. Professors falling into this area constitute those who are probably not teaching-award candidates. They are considered low-quality and difficult by their students. Again, these professors receive comments from students that are consistent with their scores. Characteristic posts on RateMyProfessors.com (2003) for this group of professors differ significantly from the previous: 'This class is absolutely ridiculous. She tells you so much information and doesn't tell you what will be on the test. Why do I feel like I want to kill myself every time class is over?' 'This is the worst teacher on the entire campus. Every class she covers at least 2 sections a day plus all the homework from each section is graded. Even if you go to class it's such an information overload that you can't understand anything. STAY AWAY!' 'This lady has no compassion or willingness to teach anything to students. She is the WORST. I will pay people not to take her. Save your GPA and sanity and take someone else;' 'He is a horrible professor. He doesn't know what he's doing and he is VERY hard to understand. He ignores the students' questions and does not explain anything. Take this class with this him if you want to really struggle and learn everything on your own!'

From student comments about these low-quality and difficult professors, it is difficult to determine whether real learning is taking place. What students make exceptionally clear is that they are unhappy about the workload, grade distributions, and teaching practices in these courses.

We then reviewed professors with an Average Quality score that is much higher than his or her Average Easiness score; these were some of the comments collected at RateMyProfessors.com (2003) for an English professor with Quality and Easiness scores of 3.3 and 1.8 respectively: 'This class is a lot of work, but it's certainly not

impossible. He really knows his stuff and if you ask for help, you'll get it. One of the best professors I've had yet'; 'He is the best. His class requires some actual effort, but it is worth it'; 'Very interesting, very intelligent. Not difficult if you study and go to class'; A sociology professor with Quality and Easiness scores of 3.7 and 2.2 respectively at RateMyProfessors.com (2003) received the following posts: 'This man changed my life through his teaching. Highly recommended for anyone seeking to think outside of the box'; 'Challenging. Good. Spoon-fed students looking for easy grade will be disappointed'.

These faculty are liked by most of their students, even though students note a rigorous course load and a high-level of learning. It holds that faculty with Quality scores far exceeding their Easiness scores ought to be appreciated for achieving a significant task: getting students to enjoy learning in the midst of high expectations and difficult curricula.

Finally, we analysed posts for professors at the bottom right of Figure 2. These are faculty considered to be low-quality and easy. A history professor received these posts at RateMyProfessors.com (2003): 'Absolute joke! Never went to class 2nd half of semester, ended up with an A-. Wrote a paper about my ethnicity that was completely made up, got A + . FREE A'S RULE!' 'Easy grader, lectures are boring and you don't really learn anything'; 'The dude is crazy. I was bored to tears. The easy A is not worth your sanity!' Clearly, these comments consistently reinforce the instructors' scores, and faculty at the bottom right of Figure 2 are not candidates for teaching awards.

Non-sexy professors

We then divided our sample into sexy professors and non-sexy professors, where sexy professors have Average Sexiness scores greater than zero. The relations between Quality and Easiness for non-sexy professors are detailed in Table 7. For non-sexy professors with at least ten posts, the correlation between Quality and Easiness is 0.61 ($N = 2{,}709$). Figure 3 shows the relation between Quality and Easiness for the 2,709 non-sexy professors. For faculty with at least 20 posts ($N = 990$), the correlation between Quality and Easiness is 0.66, and for faculty with at least 30 posts ($N = 454$), it is 0.68. Using simple linear regression with Average Quality as the dependent variable and Easiness as the only predictor (since their Average Sexiness scores are zero), the Adjusted R^2 for the model is 0.37 for 10-or-more posts, 0.44 for 20-or-more posts, and 0.46 for 30-or-more posts. The results indicate that, for the 454 non-sexy professors in our sample with at least 30 posts, 46% of the variation in Quality scores results from the professors' easiness. The regression model has an intercept of 1.04 ($t = 8.88$, $p = 0.000$) and a beta of 0.68 ($t = 19.57$, $p = 0.000$).

Sexy professors

The correlations between Average Quality, Average Easiness, and Average Sexiness for the 481 sexy professors are listed in Table 8. The correlations are as follows:

Table 7. Relations between Quality and Easiness for non-sexy professors as number of posts varies, N = 2709 to N = 11

Number of Posts	Number of Professors	Quality-Easiness Correlation	Regression Adj. R^2	Regression F-Value	Regression Intercept	Regression Easiness t-value
10 or More	2,709	0.61**	0.37	1,606**	1.30**	40.08**
20 or More	990	0.66**	0.44	774**	1.10**	27.83**
30 or More	454	0.68**	0.46	383**	1.04**	19.57**
40 or More	194	0.69**	0.47	170**	0.96**	13.06**
50 or More	106	0.76**	0.57	140**	0.71**	11.85**
60 or More	66	0.85**	0.71	161**	0.08	12.71**
70 or More	37	0.85**	0.71	88**	0.16	9.36**
80 or More	18	0.93**	0.85	97**	−0.27	9.85**
90 or More	11	0.93**	0.86	62**	−0.85	7.86**

**Significant at the 0.01 level (2-tailed).

Quality-Easiness, r = 0.46 (p = 0.000), Quality-Sexiness, r = 0.22 (p = 0.000), and Easiness-Sexiness, r = 0.10 (p = 0.027). The relation between Quality and Easiness for the 481 sexy professors is graphed in Figure 4.

For sexy professors, the correlation between Quality and Easiness is lower than for

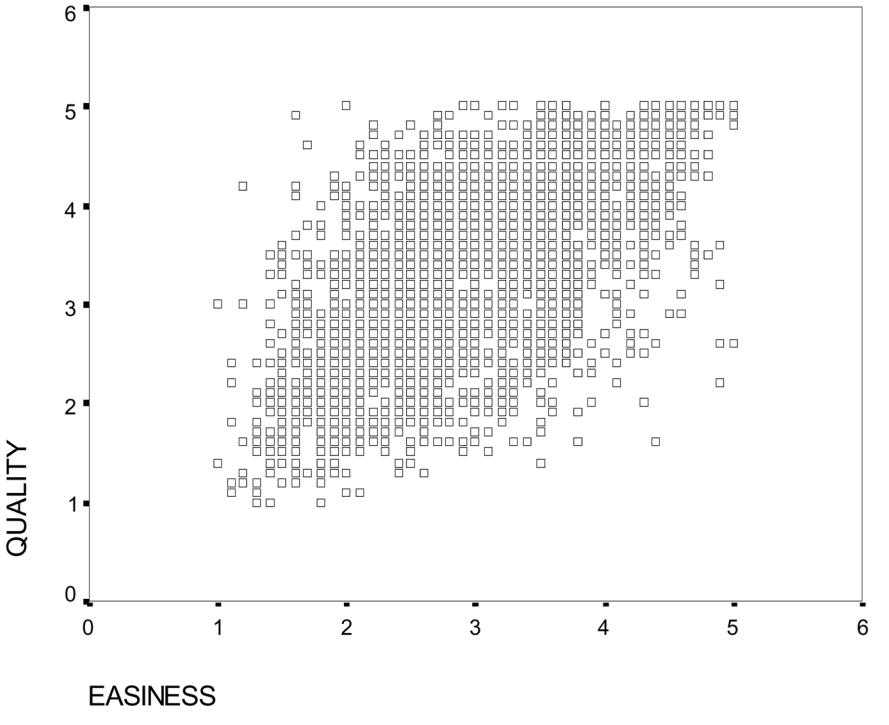

Figure 3. The relation between Quality and Easiness for non-sexy professors with at least 10 posts, N = 2709

Table 8. The correlations between Quality, Easiness, and Sexiness for sexy professors with at least 10 posts, N = 481

		EASINESS	QUALITY	SEXINESS
EASINESS	Pearson Correlation	1	.46**	.10*
	Sig. (2-tailed)	.	.000	.027
QUALITY	Pearson Correlation	.46**	1	.22**
	Sig. (2-tailed)	.000	.	.000
SEXINESS	Pearson Correlation	.10*	.22**	1
	Sig. (2-tailed)	.027	.000	.

**Correlation is significant at the 0.01 level (2-tailed).
*Correlation is significant at the 0.05 level (2-tailed).

their non-sexy colleagues, and the mean scores for Average Quality and Average Easiness are higher as well. The Average Quality score is much higher for sexy professors at 4.255 (standard deviation of 0.60), as is the Average Easiness at 3.430 (standard deviation of 0.70). These inflated scores indicate a Halo Effect for sexy professors as found in previous research. To further measure the importance of

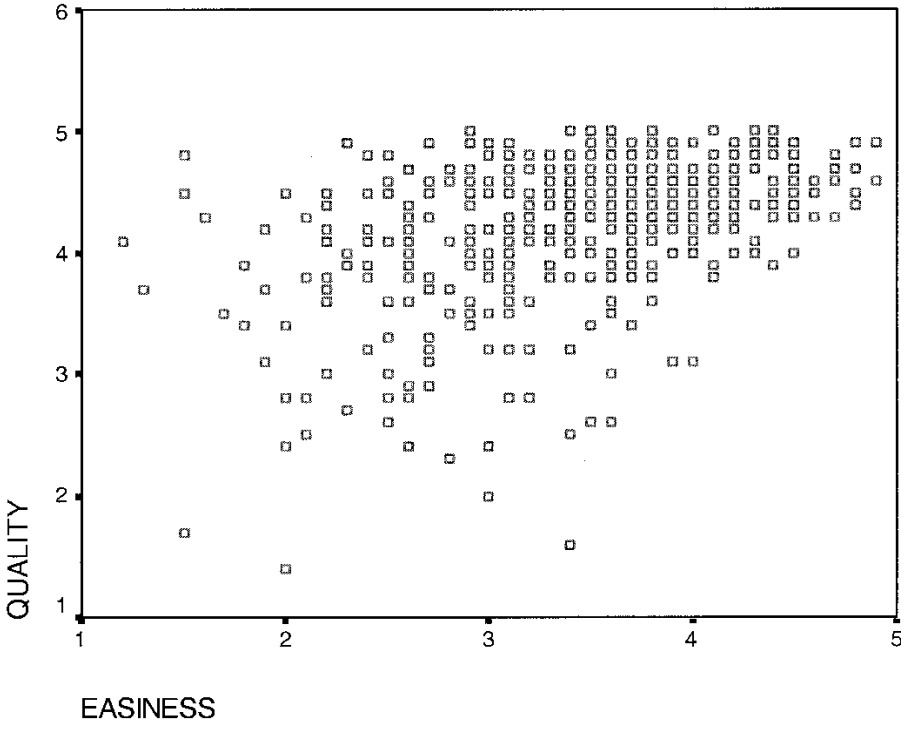

Figure 4. The relation between Quality and Easiness for sexy professors with at least 10 posts, N = 481

sexiness to student evaluators, we examined the relation between Quality and Easiness for professors with Average Sexiness scores over 20% (and at least 10 posts). The correlation between Quality and Easiness drops to 0.38 ($N = 288$, $p = 0.000$). Using linear regression with Quality a function of Easiness, the regression model is significant ($F = 49.45$, $p = 0.000$). The regression model has an intercept of 3.23 ($t = 20.29$, $p = 0.000$) and a slope of 0.38 ($t = 7.03$, $p = 0.000$), which further illustrates the Halo Effect.

It seems that the sexier the instructor, the more difficult his or her class can be while obtaining high-marks on student evaluations. Student comments support this finding. A science professor received the following comments at RateMyProfessors.com (2003) from a student who gave him a 1 ranking for Easiness and rankings of 5 for Helpfulness and Clarity: 'I loved just staring at this man. Who cares what he had to say. The hottest man in town'. Another student rated him a 2 for Easiness and a 5 for both Helpfulness and Clarity, and the student commented: 'His long, black hair is sooooooo HOT! I failed the first two exams but stayed in the class just so I could see him regularly. He's a pretty good teacher too'.

Conclusions

Our findings indicate that students who voluntarily evaluate their professors' teaching quality in a public forum are significantly affected by how easy the course and how sexy the instructor. Instructors offering easy courses tend to be rated more highly. Similarly, instructors perceived as sexy tend to receive higher quality scores. The relation between quality and easiness for sexy professors represents the Halo Effect reported by Feeley (2002) and other authors. While the Principle of Immediacy suggests that physically attractive instructors may be more effective, we believe, based on written student comments, that it is more likely that students merely perceive sexy instructors to be more effective. Further research might be useful in investigating this relationship.

Future research should also investigate whether quality instruction can make courses seem easier. This research might determine what percentage of the easiness score is attributable to great instruction. Clarifying this relationship would provide clearer indication about how universities might use student evaluations fairly and discernibly in decision-making. RateMyProfessors.com could adjust their ranking system slightly in order to provide students (and researchers) with better information. The Easiness ranking could be replaced with three choices such as Work Load from 1 (high) to 5 (low), Grade Distribution from 1 (low grades) to 5 (high grades), and Makes Learning Easy from 1 (below average) to 5 (above average).

Our research corroborates the work of Greenwald and Gillmore (1997a, b), because the strong relations in these data suggest that outside factors significantly bias the results of student evaluations. Specifically, these data raise the possibility that high-quality ratings may have more to do with an instructor's appearance and how easy he or she makes a course than with the quality of teaching.

If our findings using data at RateMyProfessors.com hold true for university-sponsored student opinion surveys as well, then the problems this situation poses for the

academy are unmistakable and alarming. If our results accurately depict the thoughts of American college students, then it is no wonder that, given the pressure for faculty to win approval from students seeking an easy A, academic standards have fallen. It is also no wonder that many in the academy are now in search of improved academic standards and reduced grade inflation. Further, our findings point to larger questions regarding how universities should treat student opinion surveys and how they might delineate better which faculty members are truly deserving of tenure, teaching awards, and promotions. As long as student opinions carry enormous weight, universities run the risk of encouraging their faculty to pander to their students' wishes by being easy graders and making sub-par curricula choices.

Acknowledgement

We are grateful to RateMyProfessors.com for the data, graduate students Berk Koprulu and Jessica Gordon for their contributions, and anonymous referees for their helpful comments.

Notes on contributors

James Felton and John Mitchell are at the Department of Finance and Law at Central Michigan University. Michael Stinson is at the Department of Computer Science at Central Michigan University.

References

Cerrito, P. B. (2000) An examination of college-wide student teaching evaluations, *College Student Journal,* 24(2), 165–171.

Cooper, W. H. (1981) Ubiquitous halo, *Psychological Bulletin,* 90, 218–244.

D'Apollonia, S. & Abrami, P. C. (1997) Navigating student ratings of instruction, *American Psychologist,* 52, 1198–1208.

Dion, K., Berscheid, E. & Walster, E. (1972) What is beautiful is good, *Journal of Personality and Social Psychology,* 24, 285–290.

Engdahl, R., Keating, R. & Perrachione, J. (1993) Effects of grade feedback on student evaluation of instruction, *Journal of Management Education*, 17(2), 174–184.

Feeley, T. H. (2002) Evidence of halo effects in student evaluations of communication instruction, *Communication Education,* 51(3), 225–236.

Greenwald, A. G. (1997) Validity concerns and usefulness of student ratings of instruction, *American Psychologist,* 52(11), 1182–1186.

Greenwald, A. G. & Gillmore, G. M. (1997a) Grading leniency is a removable contaminant of student ratings, *American Psychologist,* 52(11), 1209–1217.

Greenwald, A. G. & Gillmore, G. M. (1997b) No pain, no gain? The importance of measuring course workload in student ratings of instruction, *Journal of Educational Psychology,* 89(4), 743–751.

Heck, J. L., Todd, J. & Finn, D. (2002) Is student performance enhanced by perceived teaching quality? *Journal of Financial Education,* 28, 54–62.

Hudson, J. (1984) Improving teacher performance through evaluation and supervision, paper presented at *Annual Meeting of National Association of Secondary School Principals*, ED 250:782.

Landy, D. & Sigall, H. (1974) Beauty is talent: task evaluation as a function of the performer's physical attractiveness, *Journal of Personality and Social Psychology,* 29, 299–304.

McKeachie, W. J. (1997) Student ratings: the validity and use, *American Psychologist,* 52(11), 1218–1225.

Mehrabian, A. (1971) *Silent messages* (Belmont, CA, Wadsworth Publishing Company).

Moritsch, B. G. & Suter, W. N. (1988) Correlates of halo error in teacher evaluation, *Educational Research Quarterly,* 12, 29–34.

Nussbaum, J. F. (1992) Effective teacher behaviors, *Communication Education,* 41, 167–180.

Pike, G. R. (1999) The constant error of halo in educational outcomes research, *Research in Higher Education,* 40, 61–86.

RateMyProfessors.com (2003).

Rating Categories, RateMyProfessors.com (2003).

Rice, R. E., Stewart, L. P. & Hujber, M. (2000) Extending the domain of instructional effectiveness assessment in student evaluations of communication courses, *Communication Education,* 49, 253–266.

Rocca, K. A. & McCroskey, J. C. (1999) The interrelationship of student ratings of instructors' immediacy, verbal aggressiveness, homophily, and interpersonal attraction, *Communication Education,* 48, 308–316.

Ryan, J. (1980) Student evaluation: the faculty responds, *Research in Higher Education,* 12(4), 317–333.

Stapleton, R. & Murkison, G. (2001) Optimizing the fairness of student evaluations: a study of correlations between instructor excellence, study production, learning production, and expected grades, *Journal of Management Education,* 25(3), 269–291.

Stumpf, S. A. & Freedman, R. D. (1979) Expected grade covariation with student ratings of instruction: individual versus class effects, *Journal of Educational Psychology,* 71, 293–302.

Tang, S. (1999) Student evaluation of teachers: effects of grading at college level, *Journal of Research and Development in Education,* 32(2), 83–88.

Trout, P. (1997) How to improve your teaching evaluation scores without improving your teaching! *The Montana Professor,* 7(3), 17–22.

Trout, P. (2000) Teacher evaluations, *Commonweal,* 127(8), 10–11.

Yining, C. & Hoshower, L. B. (2003) Student evaluation of teaching effectiveness: an assessment of student perception and motivation, *Assessment of Evaluation in Higher Education,* 28(1), 71–88.