Accounting Faculty Teaching Ratings: Are They Lower Because Accounting Classes Are More Difficult?

Richard L. Constand
University of West Florida

R. Daniel Pace
University of West Florida

Nicholas Clarke
Florida State University

This study examines two important issues; first, whether accounting professors earn lower ratings than professors in other disciplines and second, whether those lower ratings are related to students' perceptions of how easy those professors and classes are. The results of the statistical analyses indicate that accounting students perceive their professors to be significantly more difficult than students in non-accounting disciplines and this perception is related to lower teaching evaluations. The results should be of interest to any accounting professor who is preparing performance review or promotion and tenure applications and is required to include student generated faculty evaluations in their applications.

INTRODUCTION

Accounting professors (all teachers are referred to as professors regardless of tenure status or rank) would typically agree that their discipline is perceived by students as being more rigorous and demanding when compared to many other academic disciplines. This perception of greater difficulty could result in relatively lower teaching ratings for accounting professors when compared to professors in other, less demanding, disciplines. This study examines teaching ratings data from the RateMyProfessors.com website (RMP) and examines whether accounting professor teaching ratings are systematically lower than ratings for professors in other disciplines and if those lower ratings are associated with student perceptions of the relative easiness or difficulty of those classes being rated.

RMP is an online faculty rating system where students rate their college professors and classes. As of the summer of 2015, RMP reports having over 15 million ratings of over 1.4 million professors at over 7,000 different schools on their site. The ratings data reflect student perceptions of professor helpfulness, clarity, and easiness that are reported on a five point Likert-type scale. RMP also publishes an "overall quality" rating for individual professors that is calculated by averaging the helpfulness and clarity ratings from all student ratings submitted for a particular professor. Appendix A shows the RMP ratings names and definitions. This study focuses on this RMP quality rating as the single comprehensive measure of faculty teaching performance as reported by students. A similar study by Constand and Pace (2014, 2015) provides a detailed analysis of the differences in ratings between finance professors and professors in

other disciplines both in business and in non-business disciplines. The authors test for differences between finance professor ratings and ratings from five broad non-business academic discipline areas and repeat the tests comparing finance against 32 more narrowly defined departmental disciplines. They report that for most academic disciplines outside of the formal and natural sciences, finance professor ratings are significantly lower than ratings for professors in these other disciplines and that finance students perceive their classes as being significantly more difficult than students in most other disciplines do. They then use regressions to show that the differences between finance professor ratings and ratings for professors in other disciplines can be largely explained by the differences in student perceptions of professor and class easiness. They also report that the one business discipline with professor ratings and student perceptions of difficulty similar to finance is accounting. This current study uses a larger sample to focus on the differences between ratings for accounting professors and professors of other disciplines and explores the relationship between differing student perceptions of professor/class difficulty and accounting professor ratings.

This study is designed as follows. Section 2 reviews the relevant literature and the research questions that are addressed. Section 3 describes the data sample and Section 4 discusses the methodology employed in the analysis. Section 5 presents the results of the analysis. Finally, Section 6 summarizes the major findings of the study and discusses the implications for accounting faculty who must include university administered teaching evaluations for consideration of promotion and/or tenure.

LITERATURE REVIEW

The Importance of Teaching Ratings for Accounting Professors

Teaching ratings are widely used by administrators for both performance review and tenure and promotion (T&P) decisions. Calderon and Green (1997) report that 95% of accounting department administrators from 172 schools use evaluations to assess teaching ability and Calderon, Gabbin and Green (1996) discuss the widespread use of SET data for teaching evaluation purposes in an overview of work undertaken by the Teaching and Curriculum section of the American Accounting Association. In another set of survey results, Yunker and Sterner (1988) report that in 241 accounting departments surveyed, over 90% of department heads reported relying on SET for assessing teaching effectiveness. Across the broader university, Seldin (1993) reports that 86% of 600 liberal arts colleges surveyed use ratings to evaluate teaching ability indicating the potential for comparisons of ratings across disciplines is always present.

This widespread use of ratings data is also reflected in the views of accounting professors themselves. Green, Calderon, Gabbin and Habeggar (1999) report that 45% of accounting instructors believe department heads use only ratings data to evaluate teaching effectiveness and in a more recent article, Morgan, Sneed and Swinney (2003) use survey data to compare the attitudes of both accounting faculty members and accounting department administrators and report that administrators place greater faith in ratings as a reflection of teaching effectiveness than do faculty. Accounting faculty, on the other hand, believe that more difficult quantitative classes result in lower ratings while more lenient grading schemes in non-quantitative classes result in higher ratings. Given the widespread use of student evaluations and the importance placed on them by administrators, it is surprising there is not more research focused on ratings of accounting professors. In Watson, Apostolou, Hassell and Webber (2007), the authors review all articles published from 2003 to 2005 in five accounting educational journals and the subject of student generated evaluations arise in only 6 of 223 articles.

Comparison of RMP Data to Traditional SET Data

Since this paper uses RMP data the question arises as to whether RMP faculty ratings data are comparable to traditional school administered Student Evaluations of Teaching (SET) data. A number of articles compare RMP faculty ratings to school based SET survey ratings and most conclude there are no significant differences across the two ratings platforms. Coladarci and Kornfield (2007), Timmerman (2008), and Silva, Silva, Quinn, Draper, Cover and Munoff (2008) all compare RMP ratings and SET

ratings for individual professors and report no significant differences in ratings levels or the distributions of ratings. Albrecht and Hoopes (2009) focus on business professors and do report quality ratings from the RMP data are significantly lower than SET ratings for their sample but, more importantly, they also report no difference in the ranking of professors from the two systems. Bleske-Rechek and Michels (2010) report that RMP data follow a near-normal distribution, that average ratings are not affected by the number of ratings, and that students who use RMP are similar to non-RMP users in terms of GPA's, class level, and attitudes towards grades and learning. This research on the comparability of RMP ratings and SET ratings supports the position that the RMP data is a good proxy for ratings data from traditional SET surveys and justifies its use in this current study.

Teaching Ratings and Differences across Academic Disciplines

A number of studies report differences in ratings across disciplines and many relate these differences to the relative difficulty of different disciplines. Feldman (1978) reviews 11 different studies that examine SET ratings across disciplines and concludes that professors in the humanities, arts, literature and language disciplines are consistently given higher ratings than professors in the math, engineering, and hard science disciplines. Cashin (1990, 1995) reports that professors in math, science and business disciplines consistently earn lower ratings and suggests this is because students doing the ratings often lack the quantitative skills necessary to perform well in those types of classes. Franklin and Theall (1992) also note differences across disciplines and report that quantitative disciplines rely more on lectures and exams while humanities classes (which have higher ratings) focus more on papers, group discussions and practice quizzes. Felton, Mitchell and Stinson (2005) examine RMP data for 6,852 professors from across 36 different departments and report significant positive correlations between quality ratings and easiness ratings. They also rank all disciplines by quality ratings and easiness ratings and report that accounting ranks 31 out of 36 in terms of quality (ranked high to low) and 35 out of 36 in terms of easiness (ranked easy to difficult). While they report significant positive correlations between easiness and quality ratings, they do not present any statistical tests of differences across disciplines. Other studies reporting similar findings using RMP data include the works of Coladarci and Kornfield (2007), Barth (2008), Otto, Sanford and Ross (2008), Felton, Koper, Mitchell and Stinson (2008), and the previously discussed Constand and Pace (2014, 2015) articles. A number of articles report similar findings using SET data instead of RMP data and Germain and Scandura (2005) provide an extensive review of this literature.

It should also be noted that a number of the articles discussed above that examine RMP data and report differences across disciplines also examine the relationship between student perceptions of the physical attractiveness of their professors using the RMP "Hot" variable. Studies such as Felton, Mitchell and Stinson (2004, 2005), Hamermesh and Parker (2005), Riniolo, Johnson, Sherman, and Misso (2006), Felton, Koper, Mitchell and Stinson (2008), and Freng and Webber (2009) all examine the relationship between attractiveness and faculty ratings and report strong positive correlations between these two variables.

Conclusion of Literature Review

The existing literature indicates that RMP ratings and school administered SET ratings are consistent with each other and justifies the use of RMP data in this current study. The literature also indicates that faculty teaching ratings, when measured by the RMP quality variable, are positively related to students' perceptions of professor and class easiness and that there are significant differences between ratings levels and perceptions of easiness across disciplines. Both Greenwald and Gillmore (1997) and McKeachie (1997) discuss this issue of the positive relationship between easiness and higher faculty ratings and they express concern for how administrators use ratings data for making personnel decisions without regard to differences in the relative difficulty of different disciplines. This current study is designed to integrate the literature on the easiness effect and discipline differences by documenting the differences in perceived difficulty of accounting professors and classes when compared to other disciplines and to examine the importance of student perceptions of easiness in explaining those differences.

DATA

The professor ratings data is collected from the RMP website for a sample of universities that offer the top undergraduate accounting programs in the nation. The final sample reflects ratings data from 29 different universities (see Appendix B). Within each university, professors are ranked by the number of student evaluations they have received and the RMP data and discipline data for all professors with at least 10 student ratings are collected for all university disciplines. The final sample represents RMP average data for 1,604 professors teaching in business disciplines and 12,106 professors teaching in non-business disciplines for a total of 13,710 sets of average professor data for ratings posted between March of 2003 and December of 2013. The RMP data includes the "overall faculty rating" (Quality), the average perceived degree of professor easiness (Easiness) and whether the professor had been designated as "Hot" (HotDum) by student raters. An additional variable, the number of ratings for each professor used in the calculation of the variable averages (Nrate) is also examined in order to control for the possible impact of this factor.

The final sample represents average individual faculty ratings data from eight business disciplines; General Business, Economics, Finance, Hospitality, Information Science, Management, Marketing, and Accounting. Since Economics is often taught in the Business School and it is more similar to quantitative business courses then other fields in the Social Sciences discipline, Economics is included as a business discipline. For non-business disciplines, the department names are used to categorize departments into one of six broad academic disciplines; Applied Sciences, Formal Sciences, Humanities, Natural Sciences, Social Sciences, and non-Business Professional Studies. The data also allow the assigning of individual professor ratings data to either accounting or one of 64 other non-business disciplines across the university.

METHODOLOGY

The focus of this study is to compare the ratings of accounting professors with those of professors from other academic disciplines, to document differences across disciplines, and to determine if student perceptions of professor and class easiness are responsible for differences in overall professor quality ratings. Given the existing research, the results are expected to show that accounting professors earn lower "Quality" teaching ratings than professors in many other (often less quantitative) disciplines and that accounting students perceive their classes as being less easy than students do in many other disciplines.

After sample statistics and correlations for the RMP variables are presented, the statistical analysis is presented in three stages. First, the ratings variable relationships are examined using regression analysis in three different samples (the full sample, a non-accounting business discipline sample, and an accounting sample) in order to compare the results both between samples and to past studies. Second, a series of mean difference tests are performed in order to compare the average accounting professor ratings to average professor ratings in a number of other broad non-business college level discipline areas and other departmental level disciplines. Finally, the mean differences between the accounting ratings and ratings for non-business department level disciplines are used in a regression analysis in order to examine the relationship between students' perceptions of professor and class easiness and the teaching ratings given their professors.

Variable Relationships in Accounting, Other Business Disciplines, and in the Full Sample

After statistics and correlations for the RMP variables in the full sample and two business samples are presented, the results of a series of hierarchal regressions are reported for these samples. These regressions allow comparison of the RMP variable relationships in these samples to each other and to the results reported in past research. The regression models, which are similar to those of Freng and Webber (2009), are of the following form.

$$Quality = \alpha + \beta1*(Easiness), \text{ and} \quad (1)$$
$$Quality = \alpha + \beta1*(Easiness) + \beta2*(HotDum), \quad (2)$$
$$Quality = \alpha + \beta1*(Easiness) + \beta2*(HotDum) + \beta2*(LnNrate), \quad (3)$$

Where:
Quality = the RMP overall quality rating,
Easiness = the RMP easiness rating,
HotDum = a variable equal to 1 if the RMP rating is "Hot" or 0 otherwise, and
LnNrate = natural log of the number of ratings for each professor.

The results are expected to show the positive relationships between faculty Quality ratings and both perceived Easiness and HotDum variables as have been reported in past studies.

Accounting Professor Ratings Compared to Professor Ratings in Other Disciplines

The second analysis approach tests the differences in the means of the RMP variables for accounting professors to the means in other disciplines. The first analysis focuses on the differences between the ratings variables (Quality, Easiness and HotDum) of accounting professors and professors in other business school disciplines. The analysis is then repeated for a comparison of ratings from accounting and six broad non-business discipline areas and is repeated again comparing ratings from accounting and 64 different department level disciplines. All comparisons are of the mean accounting professor ratings variables to mean professor ratings variables in the other discipline groups. Since a series of individual t-tests might result in Type I errors (false positive errors) in some of the tests, non-accounting business disciplines are compared against accounting as a control group using a multiple comparison of the means t-test approach developed by Dunnett (1955). The testable hypotheses are variations of the following:

H_0: professors teaching in non-accounting disciplines have the same level of Quality and Easiness ratings as do accounting professors, and
H_a: professors teaching in non-accounting disciplines have significantly higher Quality and Easiness ratings than do accounting professors.

Given the evidence and discussions presented in the literature review above, it is expected that professors in disciplines that are relatively less demanding and less quantitative than accounting will have higher quality teaching ratings and will be considered easier by students while professors teaching in disciplines known for rigorous and quantitative content may have similar or lower quality ratings.

Explaining the Differences in Teaching Quality Ratings across Disciplines

Finally, the key focus of this study is whether differences between accounting professor teaching ratings and non-accounting professor ratings are related to differences in student perceptions of the relative easiness of the professors and their classes. To explore this issue, a regression analysis of the following form is performed with the variables used in the regression being the calculated differences between the mean accounting discipline ratings and the mean discipline ratings for the 64 department levels.

$$Q\text{-difference} = \alpha + \beta1*(E\text{-difference}), \text{ and} \quad (4)$$
$$Q\text{-difference} = \alpha + \beta1*(E\text{-difference}) + \beta2*(H\text{-difference}), \quad (5)$$

Where:
Q-difference = mean difference between the overall quality ratings of accounting professors and the quality ratings for professors in the non-accounting disciplines,
E-difference = mean difference between the overall easiness ratings of accounting professors and the easiness ratings for professors in the non-accounting disciplines, and

H-difference = mean difference between the overall Hotdum ratings of accounting professors and the hotness ratings for professors in the non-accounting disciplines.

The first regression shows the importance of differences in student perceptions of class and professor easiness on the overall Quality teaching ratings while the second regression will control for differences in student perceptions of the physical attractiveness of their professors. These tests are designed to show whether differences in perceived easiness between accounting and other disciplines is responsible for the majority of the observed differences between accounting professor ratings and ratings for professors in other disciplines.

RESULTS

Descriptive statistics for the data are presented in Table 1. Data are shown for the full sample of 13,710 professors from all disciplines, the 1,604 professors from non-accounting business disciplines, and the 286 accounting professors. The data show that accounting professors have lower average Quality ratings and lower average Easiness ratings then either the full sample or the sample of other business professors.

TABLE 1
DESCRIPTIVE STATISTICS

Panel A: Full Sample (n=13,710)

	Mean	Std. Dev.	Minimum	Maximum
Quality	3.60	0.85	1.00	5.00
Easiness	3.04	0.77	1.00	5.00
HotDum	0.24	0.42	0.00	1.00
Nrate	23.78	21.54	10.00	359.00

Panel B: Business Disciplines (n=1,604)

	Mean	Std. Dev.	Minimum	Maximum
Quality	3.44	0.85	1.00	5.00
Easiness	2.90	0.74	1.20	5.00
HotDum	0.19	0.40	0.00	1.00
Nrate	30.68	31.94	10.00	359.00

Panel C: Accounting Discipline (n=286)

	Mean	Std. Dev.	Minimum	Maximum
Quality	3.39	0.86	1.20	4.90
Easiness	2.56	0.63	1.20	4.50
HotDum	0.18	0.38	0.00	1.00
Nrate	35.58	30.95	10.00	204.00

Pearson correlation coefficients are shown for the ratings variables (Quality, Easiness, and HotDum) and for the number of ratings variable (Nrate) in Table 2. For all three samples there is a high positive correlation (ranging from 45% to over 53%) between Quality and Easiness (significant at .0001).

TABLE 2
CORRELATION COEFFICIENTS FOR RATINGS VARIABLES

Panel A: Full Sample (n=13,710)

	Quality	Easiness	HotDum
Easiness	0.5363		
	(<.0001)		
HotDum	0.4246	0.2019	
	(<.0001)	(<.0001)	
Nrate	-0.0267	-0.0258	-0.0536
	(-0.0017)	(-0.0025)	(<.0001)

Panel B: Business Disciplines (n=1,604)

	Quality	Easiness	HotDum
Easiness	0.4856		
	(<.0001)		
HotDum	0.4079	0.1652	
	(<.0001)	(<.0001)	
Nrate	-0.0468	-0.025	-0.0499
	(-0.0612)	(-0.3179)	(-0.0458)

Panel C:- Accounting Discipline (n=286)

	Quality	Easiness	HotDum
Easiness	0.4514		
	(<.0001)		
HotDum	0.3938	0.204	
	(<.0001)	(-0.0005)	
Nrate	-0.0469	-0.0086	0.0244
	(-0.4297)	(-0.8854)	(-0.6816)

The p-values are shown in parentheses and reflect tests of the null hypotheses that the correlation is = 0.

The HotDum variable is also positively correlated to the Quality ratings in all three samples with correlations of about 40% (all significant at the .0001 level). Easiness is also positively correlated with the HotDum variable in all three samples but the correlations are much lower (16% to 20%). Finally, results are similar to Freng and Webber (2009) in that there is a small significant negative correlation between

the Nrate and Quality variables in the entire sample and business sample. The correlation, however, is not significant in the accounting sample.

Table 3 presents the regressions on the dependent variable Quality for all three samples. All nine regression models exhibit significant f-statistics (.0001 level) and have adjusted r-squares ranging from 20% to 39%. The regressions also show that perceived Easiness and professor attractiveness (HotDum) are positively related to Quality (at the .0001 level) and that the number of ratings is not significantly related to average Quality ratings.

TABLE 3
REGRESSION RESULTS FOR VARIABLE "QUALITY"

	Full Sample (n=13,710)			Business Disciplines (n=1,604)			Accounting Discipline (n=286)		
Intercept	-1.7908	-1.8594	1.8618	1.8273	1.8711	1.9903	1.8212	1.9183	2.2456
t-statistic	71.38	80.08	49.12	24.41	27	18.92	9.63	10.74	8.26
p-value	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
Easiness	0.5958	0.5218	0.5217	0.5568	0.4930	0.4922	0.6112	0.5254	0.5169
t-statistic	74.39	69.07	69.04	22.24	20.99	20.95	8.53	7.62	7.51
p-value	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
HotDum		0.6620	0.6619		0.7244	0.7215		0.7050	0.7120
t-statistic		48.49	48.39		16.44	16.37		6.19	6.26
p-value		<.0001	<.0001		<.0001	<.0001		<.0001	<.0001
LnNrate			-0.0008			-0.0372			-0.0941
t-statistic			-0.08			-1.51			-1.59
p-value			0.9363			0.1325			0.1119
Adjusted R^2	28.75%	39.18%	39.18%	23.54%	34.54%	34.59%	20.10%	29.38%	29.76%
F-statistic	5,533.71	4,416.97	2,944.43	494.43	423.9	283.58	72.69	60.29	41.26
p-value	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001

Each dependent variable observation represents the average rating for an individual professor.

For the full sample, comparison of the three regressions shows the model with only Easiness included explains just over 28% of the variation in Quality while addition of the HotDum variable increases the model's explanatory power to just over 39%, as indicated by the adjusted r-square values. For the business sample the results are similar with the Easiness variable explaining about 23% of the variation in Quality and the expanded model with the HotDum variable explaining about 34% of the variation in Quality and for the accounting professors, the same pattern of significant relationships observed in the other two samples is repeated. The first model with Easiness alone explains about 20% of the variation in

Quality ratings and with the addition of the HotDum variable the model explains just over 29% of variation in Quality.

The results of these regressions are consistent with what has been reported in past studies and shows there is a strong positive relationship between students' perceptions of professor easiness and physical attractiveness and the overall quality ratings given those professors. When the three sets of regressions are considered it is interesting to note that while the Easiness and HotDum variable coefficients are always highly significant the size of the t-statistics for the Business and Accounting disciplines are noticeably lower than those for the full sample. It should also be noted that in all regressions, the intercept terms are all highly significant suggesting there might be additional important variables not included in these regressions.

Table 4 presents the results of the tests of the differences in the mean ratings of accounting professor ratings (the control group) against the ratings in non-accounting business disciplines. The tests are based on Dunnett's one tailed t-statistic approach. Critical values of the Dunnett's t-statistic are presented at the bottom of the table.

TABLE 4
T-TESTS: OTHER BUSINESS DISCIPLINES AND ACCOUNTING

	Differences Between the Means		
	Quality	Easiness	HotDum
Business (n=396)	0.1384	0.4962***	0.0035
Economics (n=466)	-0.0819	0.2662***	0.0298
Finance (n=131)	0.0213	0.0633	0.0049
Hospitality (n=13)	-0.3871	0.6997***	-0.1014
Information Science (n=52)	-0.0832	0.3516***	-0.0245
Management (n=133)	0.1520	0.5750***	-0.0204
Marketing (n=127)	0.4255***	0.8121***	0.1209**
F-statistic	7.02	26.66	1.90
p-value	<.0001	<.0001	0.0657

*** Significant at the .01 level, ** significant at the .05 level.
Differences calculated as the department mean minus the accounting department mean.
Test controls type I experiment-wise error for comparisons of all treatments against control.
Critical Values of Dunnett's t-statistic; 3.18 at p-value = .01 and 2.66 at p-value =.05.

For Easiness ratings, results indicate that in six of the seven business disciplines (General Business, Economics, Hospitality, Information Science, Management and Marketing) students perceive their professors to be easier than students perceive their accounting professors to be (significant at the .01 level). Finance is the only other business discipline that is perceived to be not significantly different than accounting in terms of easiness (or difficulty). Within these business disciplines, however, this perception of relative easiness does not appear to translate into significantly greater Quality ratings except for marketing professors (significant at the .01 level). Marketing professors are also perceived as being more physically attractive than accounting professors (significant at the .05 level).

Table 5 reports the Dunnett t-tests results for the comparison of accounting professor ratings to professor ratings from six broad university disciplines. On average, students in all other broad discipline areas rate their professors as being easier than students rate their accounting professors (significant at the .01 level) while students in four of the six areas (Applied Sciences, Humanities, Professional Studies, and Social Sciences) also rate the teaching Quality of their professors as significantly higher than accounting students do. Finally, students in the Humanities and Social Sciences rate their professors as "hotter", on average, than accounting students do.

TABLE 5
T-TESTS: NON-BUSINESS BROAD AREA DISCIPLINES AND ACCOUNTING

	Differences Between the Means		
	Quality	Easiness	HotDum
Applied Sciences (n=1,245)	0.2477***	0.5588***	0.0554
Formal Sciences (n=1,544)	-0.0237	0.3524***	-0.0261
Humanities (n=4,531)	0.4019***	0.5924***	0.1282***
Natural Sciences (1,847)	-0.0356	0.1718***	-0.0505
Social Sciences (n=2,468)	0.2882***	0.6038***	0.0915***
Professional Studies (n=471)	0.2002***	0.5165***	0.0276
F-statistic	93.38	109.85	55.55
p-value	<.0001	<.0001	<.0001

*** Significant at the .01 level.
Differences calculated as the discipline area mean minus the accounting department mean.
Test controls type I experiment-wise error for comparisons of all treatments against control.
Critical Values of Dunnett's t-statistic; 3.18 at p-value = .01.
The Professional discipline area excludes all business disciplines and economics.

Table 6 reports the Dunnett t-test results for the comparison of mean ratings for accounting professors to professor ratings from 64 non-business departmental disciplines. Students in 51 of the 64 disciplines (80%) give their professors higher Quality ratings than accounting students. In 15 of these comparisons (23%) the difference is significant at the .01 level while in another 6 they are significant at the .05 level and in the remaining 3 they are significant at the .10 level. When the Easiness rating is considered, students in 60 of the 64 disciplines (94%) rate their professors as easier than accounting students do and in 45 of these 64 comparisons (70%), the mean difference is significant at the .01 level. Only 4 disciplines show lower easiness ratings than accounting and none of these are significant at the .10 level. Finally, when the HotDum variable is considered, in 43 of the 64 discipline comparisons students rate their professors as more attractive, but the mean difference is only significant in 10 of these 43 comparisons (at the .10 level or better).

TABLE 6
T-TESTS: NON-BUSINESS DEPARTMENTAL DISCIPLINES AND ACCOUNTING

	Differences Between the Means		
	Quality	Easiness	HotDum
Agriculture (n=45)	0.5641***	0.7759***	-0.0450
Anthropology (n=306)	0.1499	0.6803***	0.0472
Architecture (n=49)	-0.0463	0.3483	-0.0355
Art (n=23)	0.5695*	0.9468***	0.1695
Art History (n=110)	0.2257	0.3790***	0.1035
Astronomy (n=41)	-0.0017	0.3186	-0.0808
Chemical Engineering (n=13)	-0.1332	0.2535	-0.0245
Biology (n=485)	0.0313	0.1653	-0.0649
Chemistry (n=572)	-0.1296	-0.0049	-0.0332
Civil Engineering (n=6)	-0.4371	-0.0119	-0.0117
Classics (n=114)	0.6094***	0.6864***	0.1726**
Communication (n=437)	0.3242***	0.6054***	0.1051*
Computer Science (n=286)	-0.0885	0.3976***	-0.0385
Criminal Justice (n=92)	0.5271***	0.6099***	0.2239***
Cultural Studies (n=24)	0.0421	0.7298***	0.0300
Dance (n=14)	0.4272	0.6095	0.1074
Design (n=20)	-0.0371	0.4031	0.0717
Education (n=148)	0.3048**	0.7584***	0.0649
Electrical Engineering (n=22)	-0.3007	0.1518	-0.1329
Engineering (n=383)	-0.0213	0.3486***	-0.0634
English (n=1,416)	0.4213***	0.5892***	0.1706***
Ethnic Studies (n=40)	0.5404***	0.8706***	0.1717
Family Studies (n=37)	0.4751*	0.6678***	0.0920
Film (n=46)	0.1282	0.5794***	0.2130*
Fine Arts (n=132)	0.2789*	0.4116***	0.1020
Geography (n=199)	0.2225	0.7301***	0.0277
Geology (n=116)	0.1905	0.4373***	-0.0404
Graphic Arts (n=14)	-0.1728	0.4453	-0.1069
Health Science (n=101)	0.5496***	0.8480***	0.2177***
History (n=757)	0.3135***	0.3442***	0.0396
Human Development (n=11)	0.8857**	1.2563***	-0.0874
Humanities (n=106)	0.2799	0.6523***	0.0953
Interdisciplinary (n=7)	0.4558	0.7238	0.2503
International Studies (n=48)	0.3359	0.4319***	0.0925
Journalism (n=101)	0.1149	0.3678***	-0.0199
Languages (n=740)	0.5590***	0.7234***	0.1987***
Law (n=113)	0.2528	0.4222***	0.0518

TABLE 6 (continued)
T-TESTS: NON-BUSINESS DEPARTMENTAL DISCIPLINES AND ACCOUNTING

	Differences Between the Means		
	Quality	Easiness	HotDum
Linguistics (n=16)	0.0754	0.5194	0.1342
Literature (n=68)	0.3968**	0.8616***	0.0864
Mathematics (n=1,185)	-0.0217	0.3185***	-0.0298
Mech. Engineering (n=15)	-0.4604	-0.1352	-0.1117
Medicine (n=18)	0.5018	0.0826	0.0439
Music (n=226)	0.3240***	0.9319***	0.0562
Natural Sciences (n=11)	0.6129	1.0108***	0.2762
Nursing (n=15)	0.3663	0.3981	0.0217
Nutrition (n=27)	0.2907	0.9974***	-0.0302
Philosophy (n=334)	0.2267**	0.4276***	0.0852
Physical Education (n=41)	0.3178	0.9503***	0.1875
Physics (n=286)	-0.1930	0.1622	-0.0629
Political Science (n=656)	0.2789***	0.4206***	0.0839
Psychology (n=664)	0.4230***	0.6652***	0.1259***
Public Administration (n=13)	0.3360	0.5381	-0.1014
Recreation (n=7)	0.7129	1.2810***	0.1074
Religion (n=295)	0.5092***	0.7385***	0.0793
Science (n=315)	0.0631	0.3673***	-0.0609
Social Science (n=103)	0.0955	0.6517***	0.0644
Social Work (n=37)	0.2265	0.6733***	0.0920
Sociology (n=326)	0.2332**	0.6470***	0.1008
Speech Sciences (n=16)	0.1567	0.8006***	0.1342
Statistics (n=73)	0.1965	0.7272***	0.0820
Theater (n=92)	0.4706***	1.0533***	0.1369
Women's Studies (n=64)	0.3770**	0.6600***	0.1967**
Writing (n=21)	0.7653***	0.3381	0.3455**
Zoology (n=8)	0.1879	-0.0369	-0.0533
F-statistic	13.60	19.64	8.65
p-value	<.0001	<.0001	<.0001

*** Significant at .01 level, ** significant at .05 level, * significant at .10 level.
Differences calculated as the department mean - the accounting department mean.
Test controls type I experiment-wise error for comparisons of all treatments against control.
Critical Values of Dunnett's t-statistic; 3.75 at p-value = .01, 3.31 at p=.05, and 3.09 at p=.10.

Table 7 reports the results of the two regressions that show how differences in Quality ratings between accounting and other disciplines are related to differences in student perceptions of professor easiness (and physical attractiveness) in the 64 discipline comparisons to accounting. In the first regression the coefficient for the E-difference variable, representing differences in easiness across disciplines, has a t-statistic of 7.25 and is significant at the .0001 level. The adjusted R-square for the model is 44.6%. In the second regression model, with the inclusion of the H-difference variable, the model's adjusted R-square rises to 58.4% and both the E-difference and H-difference variable coefficients are significant at the .0001 level with the expected positive sign. It should also be noted that the intercept terms are only marginally significant (at the .10 level) suggesting there are few important omitted variables in this analysis. These results provide strong evidence that the differences in the teaching Quality ratings between accounting professors and professors in other disciplines that are documented in Table 6 are related to student perceptions of the relative easiness or difficulty of the professors and classes in their respective disciplines. Stated simply, accounting students perceive their professors to be relatively more difficult and they give them lower teaching ratings.

TABLE 7
REGRESSION RESULTS FOR VARIABLE "Q-DIFFERENCE"

Intercept	-0.1024	-0.0792
t-statistic	-1.89	-1.67
p-value	0.064	0.0993
E-difference	0.6178	0.4678
t-statistic	7.25	5.81
p-value	<.0001	<.0001
H-difference		1.0364
t-statistic		4.67
p-value		<.0001
Adjusted R^2	44.62%	58.38%
F-statistic	52.57	45.88
Significance	<.0001	<.0001

Individual observations are the 64 differences in mean ratings between accounting and the other disciplines shown in Table 6.

SUMMARY AND CONCLUSIONS

This study examines professor teaching ratings from the online ratings site RateMyProfessors.com. The empirical results find both additional supporting evidence for relationships reported in past research and provide new evidence relating to the accounting discipline. When accounting professor ratings are compared to professor ratings in other business disciplines, students in all other disciplines (except finance) perceive their professors to be easier but this perceived relative easiness does not appear to be

related to a significant difference in professor teaching Quality ratings. But when the accounting professors are compared to broad discipline areas outside of business, in four of the six areas students rate their professors as significantly higher in overall teaching Quality and significantly easier in all six non-business discipline areas. These significant differences between perceived easiness and teaching ratings are also apparent when accounting professors are compared to professors in 64 departmental disciplines across the university. Finally, the regression results in Table 7 show differences in teaching ratings for the Quality variable across accounting and other disciplines are significantly related to both differences in the perceived easiness of professors and in the perceived physical attractiveness of those professors. These empirical findings lend strong support to the McKeachie (1997) argument that any attempt to compare teaching ratings across different classes and disciplines by administrators and personnel committees should be discouraged.

Given that the literature shows the RMP ratings are valid reflections of university SET survey ratings, these results should be of great interest to any accounting professor who is required to include ratings in their supporting documents when applying for promotion and/or tenure. Since administrators and committees typically compare ratings for accounting professors to those of professors in other disciplines, it should be made clear in the applicant's dossier that significant differences across disciplines do exist and these differences are related to student perceptions of the easiness or difficulty of the disciplines themselves and may not reflect teaching ability.

REFERENCES

Albrecht, W.S. and Hoopes J. (2009). An Empirical Assessment of commercial web-based professor evaluation services. *Journal of Accounting Education*, 27, 125-132.

Barth, M.M. (2008). Deciphering Student Evaluations of Teaching: A Factor Analysis Approach. *Journal of Education for Business*, 84, 40-46.

Bleske-Rechek, A. and Michels K. (2010). RateMyProfessors.com: Testing Assumptions about Student Use and Misuse. *Practical Assessment, Research and Evaluation*, 15, Available online at http://pareonline.net/pdf/v15n5.pdf.

Calderon, T.G., Gabbin A.L., and B.P. Green (1996). Summary of promoting and evaluating effective teaching: American Accounting Association teaching and curriculum section *Journal of Accounting Education*, 15, 367-383.

Calderon, T.G. and Green, B.P. (1997). Use of multiple information types in assessing accounting faculty teaching performance. *Journal of Accounting Education*, 15, 221-239.

Cashin, W.E. (1990). Students do rate different academic fields differently., in M. Theall and J. Franklin (eds.), *Student Ratings of Instruction: Issues for Improving Practice* (New Directions for Teaching and Learning No. 43). San Francisco: Jossey-Bass.

Cashin, W.E. (1992). Student Ratings: The Need for Comparative Data. *Instructional Evaluation and Faculty Development* 12, 1-6.

Cashin, W. E. (1995). Student Ratings of Teaching: The Research Revisited (IDEA Paper No. 32), Manhattan, KS: Kansas State University, Center for Faculty Evaluation and Development.

Coladarci, T., and Kornfield, I. (2007). RateMyProfessors.com versus formal in-class student evaluations of teaching. *Practical Assessment, Research and Evaluation*, 12, 1-15.

Constand, R. L. and Pace, R. D. (2014). Student Teaching Evaluations of Finance Faculty: Perceived Difficulty means Lower Faculty Evaluations, *Journal of Financial Education*, v. 40, Spring / Summer, 14-44.

Constand, R. L. and Pace, R. D. (2015). Addendum - Student Teaching Evaluations of Finance Faculty: Perceived Difficulty means Lower Faculty Evaluations, *Journal of Financial Education*, v. 40, Fall / Winter, 177-178.

Dunnett, C. W. (1955). A Multiple Comparison Procedure for Comparing Several Treatments with a Control. *Journal of the American Statistical Association*, 50, 1096-1121.

Feldman, K. A. (1978). Course Characteristics and College Student's Ratings of Their Teachers: What We Know and What We Don't. *Research in Higher Education*, 21, 199-242.

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

Felton, J., Mitchell, J., and Stinson, M. (2005). Cultural differences in student evaluations of professors. *Academy of Business Education Conference Proceedings,* retrieved December 2, 2007, from http://www.abe.villanova.edu/proc2004/felton2.pdf .

Felton, J. F., Koper, P. T., Mitchell, J. B. and Stinson, M. (2008). Attractiveness, easiness, and other issues: Student evaluations of professors on RateMyProfessors.com. *Assessment & Evaluation in Higher Education*, 33, 45-61.

Freng, S. and Webber, D. (2009). Turning Up the Heat on Online Teaching Evaluations: Does Hotness Matter? *Teaching of Psychology*, 36, 189-193.

Franklin, J. and Theall, M. (1992). Disciplinary Differences: Instructional goals and activities, measures of student performance, and student ratings of instruction. Working paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, 1992.

Green, B.P., Calderon, T.G., Gabbin, A.L., and Habeggar, J.W. (1999). Perspectives on implementing a framework for evaluating effective teaching. *Journal of Accounting Education* 17, 71-98.

Germain, M.L. and Scandura, T. A. (2005). Grade inflation and student individual differences as systematic bias in faculty evaluations. *Journal of Instructional Psychology*, 32, 58-67.

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

Hamermesh, D. S. and Parker, A. M. (2005). Beauty in the Classroom: Instructors' Pulchritude and Putative Pedagogical Productivity. *Economics of Education Review*, 24, 369–376.

McKeachie, W. J., (1997). Student ratings: The validity of use, *American Psychologist*, 52, 1218-1225.

Morgan, D., Sneed, J., & Swinney, L. (2003). Are student evaluations a valid measure of teaching effectiveness: Perceptions of accounting faculty members and administrators. *Management Research News*, 26 (7), 17-32.

Otto, J., Sanford, D. A. and Ross, D. N. (2008). Does ratemyprofessor.com really rate my Professor? *Assessment and Evaluation in Higher Education*, 33, 355-368.

Public Accounting Report (2008). 27th Annual Professor's Survey – 2008, Top 25 undergraduate Programs. *Public Accounting Report - The Independent Newsletter of the Accounting Profession Since 1978*, October, 7-8.

RateMyProfessors (2015). http://www.ratemyprofessors.com//categories.jsp (accessed November 2015).

Riniolo, T.C., Johnson, K. C., Sherman, T. R., and Misso, J. A. (2006). Hot or not: Do professors perceived as physically attractive receive higher student evaluations? *The Journal of General Psychology*, 133, 19-35.

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

Timmerman, T., (2008). On the Validity of RateMyProfessor.com, *Journal of Education for Business*, 84, 55-61.

Yunker, P.J. and Sterner, J. (1988). A Survey of Faculty Performance Evaluation in Accounting. *Accounting Educators Journal*, Fall, 63-71.

Seldin, P. (1993). How Colleges Evaluate Professors: 1983 versus 1993. *American Academy of Higher Education Bulletin* (Oct.), 12, 6-8.

Watson, S.F., Apostolou, B., Hassell, J.M. and Webber, S.A. (2007). Accounting education literature review (2003–2005). *Journal of Accounting Education*, 25, 1-58.

APPENDIX A: VARIABLE DEFINITIONS FROM RMP

Helpfulness - Helpfulness is defined as a professor's helpfulness and approachability. Is this professor approachable, nice and easy to communicate with? How accessible is the professor and is he/she available during office hours or after class for additional help?

Clarity - A professor's organization and time management skills can make a great difference on what you get out of the class. How well does the professor teach the course material? Were you able to understand the class topics based on the professor's teaching methods and style?

Overall Quality - The Overall Quality rating is determined by the average rating of the Helpfulness and Clarity given by all users. The Easiness rating is NOT included when calculating the Overall Quality rating.

Easiness - Some students may factor in the easiness or difficulty of the professor or course material when selecting a class to take. Is this class an easy A? How much work do you need to do in order to get a good grade? Please note this category is NOT included in the "Overall Quality" rating.

Source: RateMyProfessors.com

APPENDIX B: TOP UNDERGRADUATE ACCOUNTING PROGRAMS

1	University of Texas – Austin
2	University Of Illinois At Urbana-Champaign
3	Brigham Young University
4	University of Notre Dame (no ratings)
5	University of Southern California
6	Ohio State University
7	University of Georgia
8	Indiana University Bloomington
9	Pennsylvania State University (no ratings)
10	Texas A&M (Corpus Christi, College Station, Galveston, and Kingsville)
11	University of Wisconsin - Madison
12	Michigan State University
13	University of Virginia
13	Northern Illinois University
15	University of Alabama
15	University of Washington
17	College of William and Mary
18	University of Mississippi
19	University of Iowa
20	Baruch College
20	University of Missouri - Columbia
21	University of Florida
22	Arizona State University
22	Miami University – Ohio
23	Virginia Tech
24	Case Western Reserve University
24	Florida State University
24	North Carolina State University
24	University of Connecticut
24	University of Tennessee - Knoxville
24	University of Utah

Source: Public Accounting Report, 27th annual survey of accounting programs dated October 31, 2008. Because of ties in the rankings the number of universities is greater than 25.