Student Consensus on RateMyProfessors.com

April Bleske-Rechek, Amber Fritsch, and Brittany Henn
University of Wisconsin Eau Claire

Results

Instructors Vary in Perceived Quality and Easiness

The histogram to the left shows the percent of instructors with each mean quality rating. The instructors under investigation here show a similar distribution of ratings as found in other research on RateMyProfessors.com: instructors differ widely in how high or low in quality they are rated, but the distribution of means is slightly negatively skewed, such that, on average, instructors are rated more positively than negatively in terms of their quality.

The histogram to the right displays the percent of instructors with each mean easiness rating. Again, instructors differ widely in how low or high in easiness they are rated. Ratings follow a normal distribution.

Student Consensus About Instructors is of Similar Degree Across Number of Student Raters

The scatter plot at left displays the association between number of quality ratings and degree of consensus in those ratings; the scatter at right displays the association between number of easiness ratings and degree of consensus in those ratings. As we expected and counter to negative assumptions about the reliability of student ratings on the site, degree of variance in a given instructor’s ratings is not associated with how many students have rated them. In other words, instructors with 10 ratings show the same degree of consensus in their ratings as do instructors with 50 ratings. The scatter at left shows an R² = .03, p < .09. If anything, the trend at right suggests that more variance in easiness ratings is tied to more raters, not fewer.

Within each sex and within each discipline, number of ratings was not associated with degree of variance in instructor quality ratings, all p > .07, values for r ranged from -.18 to .15.

Students Show the Most Consensus about Instructors with Really High or Low Mean Ratings

Although degree of variance in student ratings is not tied to how many students have provided ratings, degree of variance in student ratings about instructors is tied to the overall perception of their quality (shown at left, quadratic F(2, 364) = 424.85, p < .001, R² = .70) and easiness (shown at right, quadratic F(2, 364) = 66.25, p < .001, R² = .27). The effect is very robust for quality, and replicated by sex and discipline of instructor. Instructors with very high mean quality ratings showed very little variance in students’ ratings (or strong consensus) – in some cases essentially no variance at all.

Variance in Student Ratings follows Quality, not Easiness

The bar graph at right displays a finding we have documented previously. The 1031 instructors in Math & Natural Sciences departments are rated as less easy than are instructors in each of the other disciplines (335 Arts & Humanities, 66 Social Sciences, and 63 Pre-professional). F(3, 363) = 53.12, p < .001, partial η² = .27. However, instructors in Math & Natural Sciences are rated similarly in quality, F(3, 363) = 0.21, p = .89, partial η² = .002. This finding implies that students are distinguishing between easiness and quality.

In further support of students as objective judges of instruction, the bar graph at right shows that variance in students’ ratings of both easiness and quality do not differ by discipline. That is, students show similar consensus about their instructors, regardless of the type of discipline these instructors are in (easiness F(3, 363) = 0.06, p = .96, partial η² = .0006; quality F(3, 363) = 0.10, p = .96, partial η² = .0001).

Discussion

Research suggests that a majority of students use RateMyProfessors.com to either view or post ratings about instructors at their institution. However, scholars have suggested that student ratings are not valid – that students are biased, grade-oriented consumers.

In this study we have documented several effects that, taken together, suggest that students provide valid judgments of instruction: (1) the typical instructor is rated as higher in quality than in easiness; (2) student consensus about instructors can be achieved with as few as 10 ratings; (3) students show very strong consensus about those instructors who are perceived to be high in quality; and (4) although one might presume that disciplines differ in how susceptible they are to student subjectivity or bias, the degree of variance (and therefore, consensus) in students’ ratings is similar across disciplines.

Conclusion

In conclusion, we have shown that student consensus can be achieved with relatively few ratings. Our data place common assumptions about RateMyProfessors.com in question, and reinforce previous findings to suggest that, in the aggregate, students are providing objective ratings of quality of instruction.1-3.

References


Acknowledgements

We thank the Office of Research and Sponsored Programs at the University of Wisconsin-Eau Claire for supporting this research, and the McNair Foundation for providing financial support to Brittany Henn.