

What Research has to Say about Student Ratings of University Classroom Instructional Effectiveness

Jameel Hasan, Eastern Washington University

The paper provides the research findings about the relationship between research productivity and teaching effectiveness: Are faculty with excellent publication records the only qualified to teach? The average college correlation between scholarly productivity and instructional effectiveness - as perceived by students - was .12 (Feldman, 1987). Feldman concluded that 'in general for all practical purposes, the relationship between the two is essentially unrelated.'

The last seventy years of extensive research and more than 2,000 articles on student ratings and university classroom instruction are hotly debated on many campuses. Now there is little doubt that the procedure can provide valid and useful information for both faculty members and faculty managers. The paper answers the frequently voiced concerns about student ratings of faculty teaching effectiveness: Are ratings based solely on professors' popularity? Are the forms valid and reliable? What course characteristics are relevant to student ratings of instruction? What characteristics of students, and what characteristics of professors are relevant to student ratings?

Professors are the "lone rangers" of education; they are sequestered in their classrooms, unable to share in what their colleagues are doing. Professor Hatch (2006) argues for making the classroom more public. Is it practical - given the faculty and faculty managers dynamics in a given public university?

American Association of University Professors (AAUP) Statement of Teaching Evaluations states: "The kind of teaching that distinguishes itself in colleges and universities is integral with scholarship, has a way of getting outside classroom confines, and may exemplify the highest meaning of service. A judicious evaluation system would recognize the broad dimensions of teaching, be sensitive to different kinds and styles of instruction, and be as useful in distinguishing superior teaching from the merely competent as in identifying poor teaching." (AAUP: Statement on Teaching Evaluation, April 1990).

Introduction and Intent

Student ratings of college instruction are hotly debated on many campuses; UOA and EWU are no exceptions. There is far more research on student evaluation of teaching than any other form of evaluation; the use of student evaluation has increased dramatically, from 10-15% in the mid-sixties to over 85% in 1990's; a fairly high level of agreement exists between student ratings and experts' ratings of the same instructor (.70); validity is particularly high on aspects to which students can most appropriately respond, e.g., clarity of objectives, organization, interest in students, and comparison with other teachers; no matter how it is measured - split-half, test-retest, parallel forms, etc. - student ratings are highly reliable; a high correlation exists between end of term evaluation and 3-10 year subsequent evaluation; also, high correlation between current ratings and those of graduates. In other words, students are either victims or beneficiaries of our expertise in our subject matter and mode of communicating and designing the teaching-learning environment.

AACSB-sponsored 3-year most extensive research project concluded that "across all respondent categories, teaching is perceived to be the most important function.... The standard on scholarly activities, while designed to improve the quality of graduate level teaching, may instead cause misallocation of resources by diverting teaching effort into research of dubious value. What is needed is scholarly activity which prepares the faculty member to do better teaching; this does not necessarily imply publication in academic journal" (pp. 158-159). The AACSB researchers further observed, "To use an analogy, is there enough superior baseball talent to double the number of major league teams...the school gains poor research and loses good teaching" (p. 175).

Unfortunately, these debates are often uninformed by the extensive research. The controversies over the last thirty five years [In 1920 students at the University of Washington filled out what were arguably the first student ratings forms] caused student ratings to become the most extremely studied aspect of any other form of evaluation. Now, more than seventy years of research and more than 2,000 articles, there is little reason to doubt that procedure can provide valid and useful information for both faculty members and faculty managers. The purpose of student ratings is to provide information that can be used by university managers to make program and personnel decisions.

The current system of evaluation at universities is not implemented in a uniform manner across campus. Some faculty, students, and managers find this lack of consistency quite problematic - the need for a uniform student rating form (attached), utilizing commonly accepted principles and best practice of selecting or designing the Form. The author is confident that the University Task Force on Teaching Effectiveness has surveyed: 1) The University of Colorado, Faculty Course Questionnaire, 2) The University of Wisconsin, Student Assessment of Learning Gains, and 3) The University of Washington, Elements of CEC, 4) The University of Wisconsin, Student Assessment of Learning Gains, and 5) University of Michigan's highly respected Professor Wilbert McKeachie's form given in his famous book, *Teaching Tips*; also, please see his seminal article in Volume 52, No.11, 1218-1225, November, 1997 *American Psychologist* on the subject of: "Student Ratings; The Validity of Use." The following are frequently voiced concerns about student ratings of faculty teaching effectiveness:

1. Are Ratings Based Solely on Popularity? Easy Teachers/Easy Graders Get the Highest Ratings? The term popular is never defined; the assumption that popularity somehow implies a lack of substance, knowledge, or challenge is entirely without merit and there is no research to substantiate it. In fact, several empirical studies indicate students learn more in courses in which teachers demonstrate concern/interest for students and their learning - and of course teachers also earn higher ratings. For the last 41 years at EWSC/EWU the author has continuously known the names and a brief background of all his students on the first/second day of the class, without any name tag or photographs, etc. A highly respected researcher Aleamoni says: ...the students are not easily fooled. In rating their instructors, students discriminate among various aspects of teaching ability: If a teacher tells great jokes and has the students in the palm of his or her hand in the classroom, he or she will receive high ratings in humor and classroom manner, but these ratings do not influence students' assessment of other teaching skills (p. 27). Faculty members who assign more work and more difficult work tend to be rated as most effective. Some empirical studies show no effect of grading practices on overall student ratings, however; according to good professor McKeachie, if students learn more from a teacher one would expect their grades and their ratings to be higher.
2. Are Student Rating Forms Reliable and Valid? Empirical research on student evaluation of teaching generally concludes that student ratings tend to be reliable, valid, relatively unbiased and useful. Professors Murray and Felder are frequently referenced in literature review in terms of:
 - a. Evaluations are generally consistent across raters, rating forms, courses and time periods for a given semester;
 - b. The correlate moderately to highly with evaluations made of the same instructor by independent observers;
 - c. They correlate significantly with various objective indicators of student performance, such as performance on standardized exams; and
 - d. There are low correlations with extraneous factors such as class size, severity of grading, etc. In short, the empirical studies show that student ratings agree with other measures of teaching effectiveness: learning measures, student comments, expert observations, and alumni ratings. One of the major problems is day-to-day practice; student ratings are often misused, misinterpreted, and not accompanied by other information (multiple-source and multiple-method of teaching evaluation) that allows users to make good decisions. As a result, there is a suspicion, anxiety, and even dignified hostility toward student ratings.

3. What Course Characteristics Are Relevant/Irrelevant to Student Ratings of Instruction?

- a. Discipline: Research has shown that highest ratings are given to courses in the arts and humanities, followed in descending order by biological and social sciences, business and computer science, mathematics, engineering, and physical sciences.
- b. Reason for Taking the Course: Students tend to give slightly higher ratings to courses in their major fields and/or to courses that are elective rather than required: it may be a good idea to include an item that assesses student interest in the course.
- c. Course Level and Difficulty Level of Class: Ratings in higher-level courses tend to be higher than in lower-level courses. Within a discipline, the courses that are more difficult or have greater workloads tend to receive higher ratings from students.
- d. Time of Day Class is Taught: The time of the day the class is taught has no effect on ratings.

4. What Student Characteristics Are Relevant/Irrelevant to Student Ratings of Instruction?

- a. Motivation: If students are motivated, instructors are more likely to receive higher ratings in those classes.
- b. Gender: The research on the effect of students' gender on the ratings they give has not been conclusive. But, overall, there was no significant difference in the ratings. The following four student characteristics are not related to the ratings they give instructors:
 - i. Academic ability,
 - ii. Age,
 - iii. Class level (freshman or senior), and
 - iv. Personality.

5. What Instructor Characteristics Are Relevant/Irrelevant to Student Ratings of Instruction?

- a. Personality: Research shows that students appreciate instructors who are knowledgeable, warm, outgoing, and enthusiastic. These same traits are likely to make the person more effective as a teacher, so that students are stimulated to greater achievement and learning. According to frequently quoted research studies conducted by Professors Braskamp and Ory: "Neither the 'stand-up comic' with no content expertise nor the 'cold-fish expert' with only content expertise receives the highest ratings consistently."
- b. Faculty Rank: Regular faculty tend to receive higher ratings than teaching assistants.
- c. Factors Not Related:
 - i. Age of instructor,
 - ii. Years of teaching experience, and
 - iii. Gender: Analysis of classroom studies indicates no practical difference in the overall ratings of male and female instructors. In 28 studies conducted by world-renowned higher education researcher Professor Kenneth Feldman, the correlation between gender and overall evaluation of the teacher was 0.2.

Relationship between Research Productivity and Teaching Effectiveness: Are Faculty with Excellent Publication Records the Only Persons Qualified to Teach?

First, let us take a look at the insights of seminal thinkers and philosophers about the relationship between research publication productivity as compared to teaching. Second, take a look at the most exhaustive meta-analysis - more than 40 studies of relation between faculty research publication productivity and teaching effectiveness-by world renowned higher education researcher our good old Professor Kenneth Feldman.

The president of the American Council on Education points out that “we have been inappropriately focused on research, ignoring the value of approaches that focus on teaching or applications.” Professor Peter Drucker - the winner of the U.S. Medal of Freedom, and probably the most famous and influential published author in business and management areas, observes: “Academia defines knowledge as what gets printed. But surely this is not knowledge; it is raw data. Knowledge is information that changes something or somebody - either by becoming grounds for action, or by making an individual or an institution capable of different and more effective action. And this, little of the new “knowledge” accomplishes. We no longer accept the old axiom that it is the duty of the people of knowledge to make themselves understood. But until this has been done, no knowledge will have been produced.” (Drucker, 1989, p. 251).

Soji Shiba, a Japanese management guru and a professor at Japan’s University of Tsukuba, and a former visiting professor at M.I.T., remarks that “Western business schools put theory first, reality second. Business is the reality. We have to learn from business. Research and theory have a part play, but it is smaller than western schools think.” (The Economist). Professor Shiba’s remarks were fully underscored by a Business Week cover story titled “Is Research in the Ivory Tower ‘Fuzzy, Irrelevant, Pretentious’? The Dean of New York University’s Graduate School of Business said, “They say nothing in these articles, and they say it in a pretentious way. If I wasn’t the dean of this school, I’d be writing a book on the bankruptcy of American management education.”

Over 80 years ago, Alfred Whitehead observed (1929): “Mankind is as individual in its mode of output as in the substance of its output as in the substance of its thoughts...In every faculty you will find that some of the more brilliant teachers are not among those who publish. Their originality requires for its expression direct intercourse with their pupils in the form of lectures, or of personal discussion...Thus it would be the greatest mistake to estimate the value of each member of a faculty by the printed work signed with his name.” (pp. 148-149)

Porter and McKibbin write: “Given the fact that our data show, that across all respondent categories, teaching is perceived to be the most important function.... Market forces and the rapidly rising costs of education, both public and private, along with a generally rising public insistence on accountability of all institutions, will drive business/management school in general to give more serious attention to their teaching practices and commitments - whether they want to or not - than has been true in the past. This will occur, in our view, despite the countervailing pressures within the university - across a wide spectrum of institutions - for a greater emphasis on research.”

Professor K.A. Feldman Findings About Faculty Research Publication Productivity and Instructional Effectiveness-as Perceived by Students: The average correlation between scholarly productivity or accomplishment and instructional effectiveness was +.12; Feldman concluded that “in general, the likelihood that research productivity actually benefits is extremely small or (alternatively) that the two, for all practical purposes, are essentially unrelated.” (Research in Higher Education, Vol., 27, 1987, pp. 227-298).

The author is pleased with EWU’s mission of primacy of teaching and learning excellence. Professor Trimble vs. WSU, Washington State Supreme Court Ruling, 2000. Our state supreme court, in a case dealing with business professor Trimble of Washington State University (140 Washington 2d 88, February 24, 2000) made some incisive observations, which underscores the inordinate emphasis on scholarship at the cost of class room teaching effectiveness: “Due partially to his campus assignment in Vancouver, Trimble taught more video classes and had more preparations than average for a WSU Pullman based faculty member. Trimble claims this adversely affected student teaching evaluations, and required greater time by Trimble to prepare, thus decreasing the amount of time available for research and publication. The College viewed this issue a purely speculative claim. However, viewing the facts in the light most favorable to Trimble, it is reasonable that the more one devoted to scholarly activity the less time one can devote to being effective in the classroom. In fact, there has been recent scholarship devoted to this topic.” Various law review articles have addressed the effect of an over emphasis of scholarship at the expense of classroom effectiveness. (Patrick J. Schiltz, Legal Ethics in Decline: The Elite Law Firm, the Elite Law School, and the Moral Formation of the Novice Attorney. (82 MINN L. REV. 705, 1998).

Dean Schiltz's 94-page article recommended by our supreme court; Dean Schiltz observes the following about the teaching and scholarship dilemma: "Scholarship, not teaching, is the be-all; and end-all in academia. Those doing this work ask questions like, 'What clients may an ethical lawyer represent? Which of a client's ends may an ethical lawyer pursue?' Typically, the central concern of this scholarship is the good attorney who is asked to represent the bad client, or the good attorney who is asked to help a bad client achieve an immoral or unjust end. A person with a strong enough moral compass may very well resist the pressures of the legal profession on her own, but a mentor can help. This 'conflict' or 'tension' between teaching and scholarship confronts every law professor with a different dilemma. In any event, my concern here is not with the impact of the academization of the law school on scholarship, but its impact on mentoring. At bottom, professors are discouraged from teaching; Both take time away from scholarship, and in today's academy, scholarship is fast becoming the be-all and end-all." (pp. 1-2).

Conclusions

The literature clearly demonstrates that student ratings forms that are psychometrically sound are reliable, valid and relatively free from bias, and very useful in improving teaching. It is quite clear that student ratings are not only valid, but often the "only way to get much of the information needed for most evaluations." Professors Marsh and Roche conclude that "despite ill-founded fears on the part of the faculty, and claims based on research fraught with methodological weakness...the bulk of the research, however, has supported student ratings of instruction continued use as well as advocating further scrutiny."

What is the position of any university's Task Force on Teaching Effectiveness concerning: 1) The publication of student rating information? 2) What are the fundamental principles that inspire the centrality of learning of student learning at a given university? and 3) What is the guiding perspective for determining and enhancing the quality of teaching at a given university?

REFERENCES

- Aleamoni, L. 1989. Typical faculty concerns about evaluation of teaching. In L. Aleamoni, (Ed.), *Techniques for evaluating and improving instruction*. San Francisco: Jossey-Bass.
- Behrman, N., & Levin, R. 1984. Are business schools doing their job? *Harvard Business review*, January-February: 140-147.
- Braskamp, L., Brandonberg, D., & Ory, J. 1984. *Evaluating teaching effectiveness: A practical guide*. Thousand Oaks, CA: Sage.
- Braskamp, L., & Ory, J. 1984. *Assessing faculty work: Enhancing individual and institutional performance*. San Francisco: Jossey-Bass.
- Byrne, J. 1990. Is research in the ivory tower 'fuzzy, irrelevant, and pretentious'? *Business Week*, October 29: 62-64.
- Drucker, P. 1989. *The new realities: In government and politics, in economics and business, in society and world view*. New York: Harper & Row.
- Feldman, K. 1987. Research productivity and scholarly accomplishment of college teachers as related to their instructional effectiveness: A review and exploration. *Research in Higher Education*, 26: 227-298.
- Feldman, K. 1992. College students' view of male and female college teachers: Part I - Evidence from the social laboratory and experiments. *Research in Higher Education*, 33: 317-375.

- Hatch, T. 2005. Into the classroom: Developing the scholarship of teaching and learning. San Francisco: Jossey-Bass.
- Haynes, P. 1991. Management education: Passport to prosperity. *The Economist*, March 2: 3-26.
- Marsh, H., & Roche, L. 1997. Making students' evaluations of teaching effectiveness effective: The critical issues of validity, bias, and utility. *American Psychologist*, 52: 1187-1197.
- McKeachie, W. 1997. Student ratings: The validity of use. *American Psychologist*, 52: 1218-1225.
- McKeachie, W. 1994. Teaching tips: A guidebook for the beginning college teachers. Lexington, MA: D. C. Heath.
- Porter, L., & McKibbin, L. 1988. Management education and development: Drift or thrust into the 21st Century? New York: McGraw-Hill.
- Schlitz, P. 1998. Legal ethics in decline: The elite law firm, the elite law school, and the moral formation of the novice attorney. *Minnesota Law Review*, 82: 725-726.
- Statement on Teaching Evaluation. 1990. American Association of University Professors.
- Trimble v. Washington State University, 140 WA 2d 88, 1-67409-4, (2000, February 24).
- Whitehead, A. 1929. The aims of education and other essays. New York: Macmillan.

Jameel Hasan is a professor of management at Eastern Washington University. He has taught international business, multinational people management, organizational behavior, and business and society for the past 41 years. Prior to his academic career, he was an industrial analyst for the U.S. Department of State.