



Evaluating Students' Evaluation of Teachers' Effectiveness- 'A Study of M.D. University Management Department': An Overview

*Somveer **Anupama

* Research Scholar, Institute of Management Studies and Research (IMSAR), Maharshi Dyanand University, Rohtak

** Junior Research Fellow, Institute of Management Studies and Research (IMSAR), Maharshi Dyanand University, Rohtak

ABSTRACT

The competition among institutions imparting professional education is getting fierce and this has made universities and other educational institutions more cautious about their status and ranks and more conscious about the use of Students' Evaluation for Teachers' Effectiveness (S.E.T.E.). But the relevance of using this tool is still debatable among researchers, a quantum of researches claim it as very useful tool to measure and improve the performance of teachers and at the same time it is criticized heavily for its negative effects on teachers. This paper is a result of an extensive literature review of earlier work in this field and an empirical research on students and teachers in a business school of one of the university in Haryana, India. Research discussion and suggestions offer a comprehensive guide to improve S.E.T.E. and suggest refocusing teaching effectiveness.

Keywords : SETE, Performance appraisal, effectiveness

1. INTRODUCTION

A current practice among colleges and universities in India is for the administration to use a student evaluation instrument of teaching effectiveness as part of the faculty member's performance evaluation. If these instruments are used in isolation, as they frequently are, and without alternative or collaborative measures, then students become the primary determinant of a lecturer's success or failure in his or her academic career.

The collection of student ratings should be combined with data collected from different sources using various methods such as peer review, teaching portfolios, classroom-observations, or self-evaluation (Ory, 2001). At institutions that emphasize teaching (as opposed to research), higher-than-average levels of teaching effectiveness are often expected. Marsh (1987) commented that SETs may be the most studied form of personnel evaluation. In other words, the lecturers are the immediate customers and industry/society is the ultimate customer. From this position, it is clear that the use of SETE, which implicitly captures lecturer popularity, is inappropriate for measuring instructional effectiveness (i.e. learning). Cohen, (1981), Theall and Franklin, (2001) has found that the most accepted criterion for measuring good teaching is the amount of student learning that occurs. Those who learned more gave their teachers higher ratings. Ironically, while business departments purport to use student appraisals to increase total quality, Deming (1986) has suggested that the practice is inaccurate and demoralizing. In addition to criticisms of the evaluation philosophy and the validity of the instrument, there is reason to criticize the use of SETE as the only method of evaluating teaching effectiveness. Conversely, the majority of business schools use it as either the only method of teaching effectiveness or the most heavily weighted method (Abrami et al., 1990).

2. LITERATURE REVIEW

Over 2000 articles and books have been written on this topic over the past 70 years. In a study Ory (2001) found that the use of students' ratings for evaluating teacher effectiveness

is the single most researched issue in all of higher education.

Research, however, indicates that SETE is not the only possible source of information about teaching effectiveness, and it is certainly not the best source of that information. Globally, researchers have conducted hundreds of academic exercises on the reliability and validity of SETE. We will summarize some of the findings from that research and add several anecdotal cases to make this academic review more personal.

It is widely believed that SETE is only a popularity contest that has little to do with learning. Effective teaching is multidimensional in nature. Therefore, it is not surprising that countless factors influence students' insights on effective teaching. Emery (1995) found in a study of 2,673 students at a major state university that instructors who bring foodstuff to class receive the highest ratings of teaching effectiveness.

Abrami et al. (1982) suggest that instructional ratings should not be used in decision making about faculty promotion and tenure, because charismatic and enthusiastic faculty can receive favorable student ratings regardless of how well they know their subject matter. Further, these instructor attributes were not related to how much their students learned.

Few would argue with the notion that measuring student achievement is the purest form of assessing teacher effectiveness. Abrami, Leventhal and Perry (1982) performed a meta-analysis of existing literature on relationships between course content, instructor expressiveness, and student learning and SET scores.

Dowell and Neal (1982) found that student achievement accounted for only 3.9 percent of between-teacher student rating variance. In a well-controlled meta-analysis, Cohen (1983) found that student achievement accounted for 14.4 percent of overall instructor rating variance. Other analyses have turned up even lower estimates of student rating validity. In a meta-analysis of 14 multi-section validity studies,

Finally, in a more comprehensive study, Damron (1996) found that it is likely that most of the factors contributing to student instructional ratings are unrelated to an instructor's ability to promote student learning.

In short, there are three fundamental reasons to account the validity problems of SETE (Damron, 1996). First, validation studies that do not properly control for biasing factors (e.g. student characteristics, instructor characteristics, class characteristics) yield internally invalid and un-interpretable estimates of rating validity.

Second, when appropriate controls are implemented, resulting validity estimates account for only a small fragment of between instructors rating variance. The proportion of variance accounted for appears to be inversely related to the scope of the controls.

Third, even among well-designed validity studies, validity coefficients tend to be highly variable and mediated by situational factors to such a degree that coherent context-independent estimates of validity are not possible. The latter two problems have weighty implications for the accuracy and developmental utility of student ratings.

Damron (1996) suggests that even if a sufficiently valid rating questionnaire existed, there are no guarantees that interpretations of ratings data will be valid or consistent (or reasonable, coherent or fair).

Additionally, there are differences in what specific instruments are intended to measure, how appropriate they are to different institutional settings, and how they should be used (e.g. teaching improvement or personnel decisions).

Additionally, scholars have repeatedly indicated that students are not qualified to evaluate their lecturers. For instance, Adams (1997, p. 31) stated: "[Are] students, who are almost universally considered as lacking in critical thinking skills, often by the administrators who rely on student evaluations of faculty, able to critically evaluate their instructors? There is substantial evidence that they are not". They may, however, be subject to legal challenge, because student ratings lack a certain degree of behavioural specificity (i.e. a five-point Likert scale) (Cascio and Bernardin, 1981).

3. CASE ANALYSIS

Each faculty member has at some point in their career questioned the reliability of SETE and the propriety of using the ratings for promotion and tenure decisions. The following cases illustrate the inconsistencies of both student ratings and administrative interpretations of them. These examples are factual and drawn from personal experience of the authors.

Case 1. Teacher who meets lectures regularly

One of the teachers investigated in the study was very punctual and always join the class before time and left the class after completion lecture duration. The students have rated him at 4.12 (on a five-point scale), in class A and 4.04 in class B, while the college average is 4.60. So it was difficult to accept why a teacher devoting maximum possible time of a lecture to students is rated below the school's average.

Case 2. Teacher who ensures availability beyond lecturing hours

The above discussed teacher also ensures his availability beyond his lectures. While in his office room he used to welcome interaction with students compromising his research work. But again the students' evaluations of his performance on this variable were 4.32; these were lower than the college's average 4.55. Again, this poses a serious question on dedication of the teacher.

Case 3. Teacher who assess students fairly on the basis of their all-round performance

One of the teachers used to assess the performance of the

students on predefined parameters very fairly by devoting time and energy and always avoid all kind of biases. Though it become so tedious to teacher to evaluate the assignment, term papers and presentations of all students but the teacher always ensured fair assessment of all students and also ensured transparency of his assessment of students. Even though, students still rated him much lower than the college's average. He was rated comparatively low by the students giving him a score of 4.23 while the college average was 4.39. So kind of score is certainly justifiable in the perception of the teacher observed.

Case 4. Teacher who devotes time in preparing for the lectures

Another teacher who devotes enough time in sorting appropriate case studies and assign very useful topics to enhance their practical learning of the students. Also the teacher uses to share notes and power point slides with students. Even though she achieved a score of 4.55 but on the other hand the school average score was 4.68.

Case 5. Teacher who is known for his knowledge

Also the above discussed teacher who is commonly accepted as an expert possessing a great knowledge about business world and management concepts. Again he is rated lower by students by giving individual score of 4.57 comparing school averages of 4.68. Again it poses a very serious question regarding the teacher.

4. DISCUSSION

It is extremely difficult to design and implement a performance appraisal program that is accepted by all subordinates. The following comments raised by industry can be legitimately made by higher education:

- They tend to foster mediocrity and discourage risk taking. The lecturer mentioned in Cases 1-3 has retreated from his rigorous expectations in order to receive higher student ratings. Unfortunately, student achievement has also retreated.
- They focus on the individual and therefore tend to discourage or destroy teamwork within and between departments.
- The process is detection-oriented rather than prevention-oriented.
- They are often unfair, since administrators frequently do not possess observational accuracy.
- They fail to distinguish between factors that are within the faculty members' control and system-determined factors that are beyond their control.

If higher education is going to fully embrace total quality, it requires a closely monitored performance appraisal process that is oriented toward "best practices" and continuous improvement of quality.

5. CONCLUSION

Our observations and the literature suggest that the design of SETs matter. As such, we suggest few best practices for the administration and use of SETs.

- Student ratings should not be used as the only way to measure teaching effectiveness, use numerous sources of data.
- Make the phrasing on SETE instruments more "achievement" oriented rather than "satisfaction" oriented.
- All universities should ensure that their surveys questions are valid and reliable.
- Ensure that the system is legal. This is a complex topic that may require the consideration of several attorneys and precedence within education.
- Ensure that the system celebrates diversity. Every person have different cultural background so faculty from different cultures may honestly and justifiably have different concepts of what is acceptable and effective teaching behavior.

In conclusion, we endorse the notion that "no one has taught anything, unless someone has learned something". We recognize that the activity of teaching is essentially one of human interaction, and as such is inextricably tied to the student's perception of a lecturer's personality.

REFERENCES

- Abrami, P.C., d'Apollonia, S. and Cohen, P.A. (1990), "Validity of student ratings of instruction: what we know and what we do not", *Journal of Education Psychology*, Vol. 82 No. 2, pp. 219-31. | Abrami, P.C., Leventhal, L. and Perry, R.P. (1982), "Educational seduction", *Review of Educational Research*, Vol. 32, pp. 446-64. | Adams, J.V. (1997), "Student evaluations: the ratings game", *Inquiry*, Vol. 1 No. 2, pp. 10-16. | Arreola, R.A. (1995), *Developing a Comprehensive Faculty Evaluation System*, Anker Publishing, Boston, MA. | Braskamp, L.A. and Ory, J.C. (1994), *Assessing Faculty Work: Enhancing Individual and Institutional Performance*, Jossey-Bass, San Francisco, CA. | Cascio, W.F. and Bernardin, H.J. (1981), "Implications of performance appraisal litigation for personnel decisions", *Personnel Psychology*, Vol. 34, pp. 211-26. | Cohen, P.A. (1983), "Comment on a selective review of the validity of student ratings of teaching", *Journal of Higher Education*, Vol. 54, pp. 448-58. | Crumbley, D.L. (1995), "The dysfunctional atmosphere of higher education: games professors play", *Accounting Perspectives*, Spring, Vol. 1 No. 1, pp. 27-33. | Damron, J.C. (1996), "Instructor personality and the politics of the classroom", available at: www.mankato.msus.edu/dept/psych/Damron_politics.html (accessed May 2001). | Delucchi, M. and Pelowski, S. (2000), "Liking or learning? The effect of instructor likeability on overall ratings of teaching ability", *Radical Pedagogy*, Vol. 2, pp. 1-15. | Deming, W.E. (1986), *Out of the Crisis*, MIT Center for Advanced Engineering Study, Cambridge, MA. | Dowell, D.A. and Neal, J.A. (1982), "A selective review of the validity of student ratings of teaching", *Journal of Higher Education*, Vol. 53, pp. 51-62. | Dowell, D.A. and Neal, J.A. (1983), "The validity and accuracy of student ratings of instruction: a reply to Peter A. Cohen", *Journal of Higher Education*, Vol. 54, pp. 459-63. | Emery, C. (2001), "Professors as customers", in Lamb et al. (Eds) *Great Ideas for Teaching Marketing*, 6th ed., South-Western College Publishing, New York, NY. | Franklin, J. and Theall, M. (1990), "Communicating student ratings to decision makers: design for good practice", in Theall, M. and Franklin J. (Eds), *Student Ratings of Instruction: Issues For Improving Practice*, Jossey-Bass, San Francisco, CA. | Kaplin, W.A. and Lee, B.A. (1995), *The Law of Higher Education: A Comprehensive Guide to Legal Implications of Administrative Decision Making*, 3rd ed., Jossey-Bass, San Francisco, CA.