WCLTA 2011

What do higher education instructors consider useful regarding student ratings of instruction? Limitations and recommendations

Seyedeh Azadeh Safavi [a] *, Kamariah Abu Bakar [b], Rohani Ahmad Tarmizi [c], Nor Hayati Alwi [d]

[a b c d] *Faculty of Educational Studies, Universiti Putra Malaysia (UPM), 43400 Selangor, Serdang, Malaysia*

Abstract

Instructional improvement is a main purpose of performing Student Ratings of Instruction (SRI) in higher education. Although issues on the subject of student ratings have been extensively researched, there have been limited studies dealing with the question that what types of ratings information are useful for instructors in order to make improvements in their instructional practices? This study explored the utility of ratings information as perceived by instructors, examined significant differences in utility of ratings information based on instructors' characteristics, and identified instructors' reactions and recommendations. The population comprised 1139 instructors from 15 faculties of a major Malaysian research university. 330 of them were surveyed, with a response rate of 75.1%. The responses were subjected to descriptive statistics, principal component analysis, one-way ANOVA, and independent t-test. The results showed that the ratings information which focused on instructor's efforts in student learning enhancement was considered useful for instructional improvement. The instructors acknowledged that there were limitations in the current student ratings being used. They shared the views that faculties/departments should not limit themselves to the common set of rating items, and that instructors with teaching problems should be recommended to use of instructional consultant services.

1. Introduction

Student Ratings of Instruction (SRI) is a popular procedure in higher education (Pounder, 2007; Smith, 2007). Indeed, instructors are the major users of ratings information. They use of ratings information to enhance quality of their instructional practices. Although issues on the subject of student ratings have been extensively researched, missing in the literature is research on the instructional utility of ratings information (what information of student ratings is most likely to induce instructors to make improvements in their instructional practices?). As a step toward instructional improvement, in 1992, the University Putra Malaysia (UPM) designed a student ratings instrument to evaluate teaching performance of instructors. The main purpose of this instrument was enhancing quality of teaching performance. Its content validity was verified by an appointed committee, headed by the Dean of the Faculty of Educational Studies. Currently, UPM's rating instrument contains 22 instructional practices (items). Every semester, it is filled out by all students (undergraduates and postgraduates alike) from all 15 faculties at UPM. The quantitative data are usually analyzed by the deputy dean's office in each faculty; then the information is submitted

---

* Seyedeh Azadeh Safavi. Tel.: +6-017-348-4543
 E-mail address: safaviazadeh@yahoo.com

to the respective departments. In addition, a copy of the result is directly sent to instructors including professors, associate professor, senior lecturers, and lecturers with exception of tutors. This paper explores the utility of ratings information as perceived by instructors within the context of a major research university, examines significant differences in utility of ratings information based on instructors' characteristics, and identifies instructors' reactions and recommendations.

2. Method

2.1. *Population and sampling*

The present study was conducted at a public Malaysian research university, namely University Putra Malaysia (UPM). The population comprised 1139 professors, associate professors, senior lecturers, and Lecturers from 15 faculties. In order to determine the adequate sample size, the G Power 3.1 statistical software (Faul, Erdfelder, Buchner, & Lang, 2007) as well as the rule of thumb proposed by Hair, Anderson, Tatham, and Black (1998) were utilized and the largest sample size between the two was then selected as the final sample size of the study (n= 330). The samples drawn from each faculty were calculated based on the stratified proportionate random sampling.

2.2. *Design and instrumentation*

This study employed a self developed questionnaire. It comprised 34 items which were grouped into three sets of variables. The first set (4 items) addressed the instructors' demographic characteristics, the second set addressed the 22 items which found in the current UPM's rating instrument, and the third set (8 items) addressed instructors' reactions toward SRI as well as their recommendations for more effective use of student ratings. The eight items of third set were inspired and generated by research literature (i.e. Shemecklin et al., 1997; Cashin, 1990). The levels of content and construct validity were established by a panel of experts who were experienced and specialized in the area. Moreover, a trial run was conducted with fifty instructors in the pilot study to ensure its validity. To further verify the strengths of the result, the reliability of the instrument was estimated with the measures of internal consistency.

3. Results

3.1 Instructors' demographic characteristics

Instructors' demographic characteristics solicited the information about their gender, academic rank, levels of course they were teaching, and their teaching experiences. The results showed that the majority of the instructors were female (59.9%, n= 148). Of the respondents, 29.6 % (n= 73) were lecturers, 47.3% (n=117) were senior lectures, 17.4% (n= 43) were associate professors, and 5.7% (n= 14) were professors. The majority of the respondents (58.7%, n=145) reported that they taught in both postgraduate and undergraduate levels. Also, the respondents' teaching experiences ranged from 6 months to more than 30 years, with a mean about 12 years.

3.2 Perceived utility of ratings information

This section explored the types of ratings information that were useful for instructional improvement as perceived by instructors. For this purpose, a principal component analysis with Varimax rotation was utilized on the second set of variables (22 items). Finally, based on the methods of latent root criterion and Scree plot, three components with eigenvalues greater than 1.0 were extracted. Each component reflected one type of ratings information (Table 1).The first component (items 1 to 12) reflected a type of ratings information which focused on instructors' efforts in student learning Enhancement namely LE. The second component (items 17 to 22) reflected a type of ratings information pertaining to the Interpersonal Skills and Overall personality/ performance of instructors (ISO). The third component (13 to 16) addressed another type of ratings information regarding instructor's Assessment and Grading practices (AG).

According to Table 1, the Cronbach's alpha for the components of LE, ISO, and AG was .92, .82, and .85 respectively. A component mean was produced by averaging the means of the individual items in a component and was the base for interpreting the utility of ratings information. The results indicated that the ratings information which focused on instructors' efforts in learning enhancement (LETP) was perceived to be useful for instructional improvement while two other types of the ratings information (ISO & AG) were perceived to be not useful.

Table 1: Perceived utility of ratings information

Item	Usefulness	Component*		
	Mean	LE	ISO	AG
1-Teaching plan is clearly explained.	2.57	.67		
2- Instructor is always prepared for each teaching session.	2.69	.52		
3- Learning activities are managed effectively.	2.62	.52		
4- Student easily understood teaching delivery.	2.72	.74		
5- Important aspects are emphasized in the teaching.	2.74	.65		
6- Lesson contents are summarized.	2.51	.64		
7- Delivery method stimulates student's thought.	2.62	.73		
8- Delivery method enhances student's interest in the subject.	2.65	.68		
9- Examples are helpful for students in their learning process.	2.91	.63		
10- Teaching materials are helpful in learning process.	2.74	.53		
11- Students are encouraged to ask question.	2.65	.57		
12- Students are exposed to current issues.	2.58	.47		
13- Assessments are implemented as planned.	2.43			.85
14- Tests/assignments given conformed to the course content.	2.53			.68
15- Feedback is given on assignment.	2.48			.69
16- Assessment marks are announced within the stipulated period.	2.42			.65
17- The instructor shows concerns and demonstrates responsibility.	1.84		.61	
18- Student is being respected.	1.70		.73	
19- The instructor portrays a good image.	1.82		.72	
20- The instructor's delivery is convincing.	2.72		.63	
21-Overall, the teaching of this course is effective.	1.71		.71	
22- Overall, this instructor is the best instructor.	1.31		.65	
Cronbach's alpha		.92	.82	.85
Inter-item correlations mean		.49	.54	.50

*Component: LE= Learning Enhancement, ISO=Interpersonal Skills- Overall, AG= Assessment and Grading

The independent t-test analysis and one-way ANOVA were conducted to assess significant differences between instructional utility of ratings information and demographic characteristics constituting the dependent and the independent variables respectively. The results of these analyses showed that there were no significant differences in instructors' views of the utility of ratings information based on their demographic characteristics. In other words, instructors agreed on the utility of the ratings information immaterial of gender difference between them, academic position held, length of teaching experience, and level of courses they taught.

3.3 Instructors' reactions and recommendations

The results showed that the majority of instructors who participated in this study had positive reactions to SRI (Table 2).For instance, the majority of them (a total of 67.6%, n=167) disagreed with the statement that instructional performance SHOULD NOT be evaluated by student (Item 1) with about 30% of them strongly disagreeing to the statement and a majority of the instructors (total of 64.7%, n= 160) also were disagreeing with the statement that they DON'T take student ratings seriously (Item 2) with about 21% of them strongly disagreeing to the statement.

Instructors' recommendations were presented by their levels of agreement with the six items in Table 2 (items 3 to 8). As can be seen from Table 2, more than half of the respondents (total of 59.5%, n=147) concurred that

instructors with instructional problems should be recommended to use of instructional consultant services (Item 3), more than three quarters of them (total of 80.1%, n=198) agreed that the items of the current form need revision (Item 4), and a total of 80.9 % (n= 200) agreed that administration of the current form need revision (Item 5).Also, the majority of them (total of 88.3 %, n=218) agreed with the suggestion that a written explanation be prepared to explain how results of the student ratings should be interpreted by them (Item 6). Moreover, there was strong consensus (29.1% strongly agreed, 66% agreed; total of 95.1%; n=235) that faculties/departments should not limit themselves to a common set of items as currently practiced (Item 7), and a total of 96% (n= 237) agreed that they should be allowed to choose additional items especially for their courses with regard to their course objectives and teaching methods (Item 8).

Table 2: Instructors' reactions and recommendations

Item	Mean	Percentage (%)			
		SD	D	A	SA
1-Instructional performance should not be evaluated by students.	2.78	22.7	44.9	20.2	12.1
2-I don't take student ratings seriously.	2.78	20.6	44.1	27.9	7.3
3-Instructors with instructional problems should be recommended to use of instructional consultant services.	2.59	10.9	29.6	48.2	11.3
4-The items of the current form need revision.	3.06	2.0	17.8	51.8	28.3
5- Administration of the current form needs revision.	3.08	1.2	17.8	52.6	28.3
6-A written explanation should be prepared to explain how results of the student ratings should be interpreted.	3.10	2.4	9.3	64.0	24.3
7-Faculties/departments should not be limited to the common set of items as currently practiced.	3.23	.80	4.0	66.0	29.1
8-Instructors should be allowed to choose additional items relevant to their course objectives and teaching methods.	3.25	0.0	4.0	66.0	30.0

Scale: SD= Strongly Disagree, D= Disagree, A= Agree, SA= Strongly Agree

4. Discussion

The findings indicated that the ratings information which focused on instructors' efforts in learning Enhancement was considered useful for instructional improvement. Hence, it can be concluded that effective student learning is one of the more important concerns of the instructors at UPM. Instructional practices which were included in this type of information focused on instructor's efforts in enhancing student's interests, in stimulating students' thought, and in facilitating the learning process. This finding is consistent with those of Schemlkin et al. (1997) who found that the most useful feedback items for instructors are those that deal with the instructor's presentation and the encouragement of questions. Also, the ratings information related to Assessment and Grading practices (AG) was perceived to be not useful for instructors. This type of ratings information contained instructors' assessment and grading practices such as providing feedback on graded work, reflecting course content on tests, and scoring practices. This result confirmed the previous studies which found that students are unable to evaluate instructors' grading practices (Hoyt & Pallett, 1999; Keig & Waggoner, 1994). In addition, the ISO was not considered by instructors as a useful type of ratings information. This finding confirmed the idea that if student ratings are being used for instructional improvement, items that concern overall ratings will not provide useful information and instructors should use information from specific items rather than overall items (Algozzine et al., 2004, McKeachie, 1997, and Marsh & Roche, 1997).

The present study revealed that more than three quarters of instructors concurred that the items of the current form need revision. Moreover, there was strong consensus that faculties/departments should not limit themselves to the common set of items as currently practiced, and that instructors should be allowed to choose additional items relevant to their course objectives and teaching methods. Their views concurred with those of Franklin (2001) and Cashin (1990), who suggested that for instructional improvement, student ratings form should not contain a single set of items for all the classes. Hence, these findings offer a clear message for the university to design and use

different set of items for addressing differences in course objectives in all faculties. In addition, the majority of the instructors agreed that the administration of the current form need revision**.** Obviously, administration of student ratings contributes to effective use of student ratings. Hence, this finding offers to the university to develop targeted interventions aimed at assessing the administration of the current student ratings procedure in order to address the revisions that needs to be made.

Furthermore, majority of the respondents concurred that instructors with instructional problems should be recommended to use of instructional consultant services. The present finding is further supported by the previous studies which recommended that the most effective use of student ratings occurs when instructors are assisted by a professional teaching consultant (e.g. Marsh, 2007; Brinko, 1993). Currently, those instructors who receive low score in student ratings are recommended by administrators to attend in some classes/workshops regarding teaching methods at UPM. Although attending in some educational classes might be more useful for beginning instructors, but very experienced instructors may benefit from individualized consultations. In other words, the effect of student ratings on instructional improvement will be increased if they are accompanied by appropriate simultaneous consultations (Marsh, 2007).

5. Conclusion

In view of the findings in this study and the implications arising from them, some recommendations were made. It is expected that the findings of this research will bring valuable insight and understanding to help the university authorities at UPM specifically, and other universities generally, as well as all instructors to achieve more success in using the student ratings, and consequently, in enhancing the quality of teaching performance. Although our results indicate that the information gathered from the rating instrument would be useful for improving the quality of instructional practices, however, the nature of instructional practices which would be improved by student ratings need to be explored in future research.

References

Algozinne, B., Beattie, J., Bray, M., Flowers, C., Gretes, J., Howley, L., Mohanty, G., & Spooner, F. (2004). Student evaluation of college teaching: A practice in search of principles. *College Teachin.* 52 (4): 134-141.
Brinko, K. T. (1993). The practice of giving feedback to improve teaching. *Journal of Higher Education* 64: 574-593.
Cashin, W.E. (1990). Student Ratings of Teaching: Recommendations for Use. *Center for Faculty Evaluation & Development.* Kansas State University.
Faul,F., Erdfelder, E., Buchner, A., & Lang, A.G. (2007).Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41* (4): 1149-1160.
Franklin, J. 2001. Interpreting the numbers: Using narrative to help others read student evaluations of your teaching accurately. *New Directions for Teaching and Learning,* no. 87: 85-100.
Hair, J.F., Anderson, R.E., Tatham, R.L., & Black, C. (1998). *Multivariate data analysis.* Prentice Hall.
Hoyt, D & Pallett, W.H. (1999). Appraising teaching effectiveness: beyond student ratings. *IDEA Center.*
Keig, L., & Waggoner, M.D. (1994). Collaborative peer review: The role of faculty in improving college teaching. *Higher Education Report, No. 2.*Washington, DC: George Washington University.
Marsh, H.W. (2007). Do university teachers become more effective with experience? A multilevel growth model of students' evaluations of teaching over 13 years. *Journal of Educational Psychology, 99*(4): 775-790.
Marsh, H.W., & Roche, L. (1997). Making students' evaluations of teaching effectiveness effective: The critical issues of validity, bias and utility. *American Psychologist, 52*(11): 1187-1197.
Mckeachie, W.J. (1997). Student ratings: The validity of use. *American psychologist.* 52(11): 1218-1225.
Pounder, J. (2007). Is student of teaching is worthwhile? An analytical framework for answering the question. *Quality Assurance in Education.* 115(2).
Schmelkin, L.P., Spencer, K.J., & Gillman, E.S. (1997). Faculty perspectives on course and teacher evaluations. *Research in Higher Education.* 38: 575–592.
Smith, B.P. (2007). Student ratings of teaching effectiveness: An analysis of end-of-course faculty evaluation. *College Student Journal.* 41(4): 788-801.