STUDENT RATINGS IN VIETNAM HIGHER EDUCATION: HOW ARE INSTRUCTORS' REACTIONS?

Hao T. Nguyen

University of Social Sciences and Humanities

Vietnam National University

Ho Chi Minh City 12 Dinh Tien Hoang St., Dist.1, Ho Chi Minh City, Vietnam

haonguyenpy2@yahoo.com

ABSTRACT. *This work focuses on investigate how instructors react to the use of students rating of instruction whose purpose is to help instructors to improve their teaching skill. We collect data with questionnaires and illustrate this issue with statistical methods. Our findings are that instructors of lower division courses, instructors with less teaching experience and female instructors tend to appreciate the usefulness of the student rating procedure.*

Keywords: Student Rating; Teaching Improvement; Instructors Perceptions

1. Introduction. First used in Canadian and American universities in the mid-1920s, student ratings of instruction have become integral to accountability in higher education (Zabaleta 2007). Besides peer evaluation, student rating of teaching in Vietnam's colleges and universities is the requirement from Ministry of Education and Training in effort of instructional improvement. This kind of evaluation is one of the most popular approaches to teacher evaluation in which students express their opinions and feelings concerning their teachers instructional processes and activities during one semester. However, there still have different points of view surround student ratings of teaching; especially instructors express their concerns on reliability and validity of student ratings results. Given these concerns, the target of present study is to examine instructors' perceptions of using student ratings for instructional improvement, and their attitudes toward student ratings of instruction as well as utilizing student rating results for improving teaching practice based on empirical evidences at University of Social Sciences and Humanities, Vietnam National University, Ho Chi Minh City (HCM-USSH). Preliminary results contribute evidences for school administrators in effort of improving and assuring training quality. This study also provides a source of reference in the use of student ratings for personnel decision-making.

1. Do students' technical and scientific areas differ in their problem-solving efficacy? And How?
2. How are technical and scientific areas students of problem-solving efficacy affected by their college learning experiences?

2. Review of Literature.

2. 1. Reliability and Validity of Student Ratings. Reliability and validity are two factors

of measurement, student ratings of teaching is not an exception. Reliability of student ratings of teaching is defined as consistency across time and across student rating instruments for an individual faculty member (Hooper and Page 1986) and related to the accuracy and precision of a measurement procedure (Emory 1985). In comparison with reliability, validity is more complicated and controversial as well as important. As a common sense of validity, a measurement instrument is valid if it reflects what it intends to do. For student ratings of teaching is actually measure teaching effectiveness. If a device of student ratings is valid that means the positive correlation of student ratings and teaching effectiveness is high. An effective teacher will receive high ratings and an ineffective one will get low rating results from students. Many researchers show that student ratings of instruction are related to effective teaching (Marsh 1987). However, there are still some researchers who distrust the validity of student ratings of teaching effectiveness (Lori, Regina & Peter 2010). In generally speaking, researchers have generally supported the reliability and validity of student ratings even though the empirical results are not always consistent across studies (Greenwald 2002)

2. 2. Using Student Ratings of Teaching for Improving Teaching. The most important purpose of student ratings is improving instructors teaching and their courses. Unfortunately, according to Centra (1993) teaching improvement occurs only if a faculty member knows how to make changes and is motivated to do so. Even if faculty members know how to interpret student evaluation results, they may not know what to do in order to improve their teaching (Jacobs 1987). Regarding using student ratings of teaching for improving teaching staffs instruction, many studies have found mixed findings and not being stables. A significant influence of student ratings of teaching on improving teaching has been found in some studies. In a meta-analysis of Cohen (1980:339), he pointed out that: "student ratings are a valuable source for improving instruction at the college level." On the other hand, some studies reveal have no such effects. As Yao and Gradys (2005:507) citation, Rotem and Glasman (1979) reported that "feedback from student ratings does not seem to be effective for the purpose of improving performance of university teachers."

2. 3. Instructors Attitudes on Using Student Ratings. Studying attitudes about using student ratings in higher education among faculty members received diversity picture in term of rating purposes. Beran and Rokosh (2009:1973) asserted: contrary to anecdotal reports, which tend to emphasize instructors negative views of student ratings, the empirical literature to date has revealed a more positive outlook. Studies of Schmelkin et al. (1997) and Beran et al. (2002, 2005) showed instructors positive attitudes on utilizing of student evaluations of teaching in general. However, Kulik & McKeachie (1975) find that there is no convincing evidence that the information helps teachers improve their effectiveness. Instructors tend to support the formative application rather than summative application of student ratings. Having small number of teaching staff, from 8 percent to 23 percent, agrees with using student ratings of teaching for administrative decision-making (Nasser and Fresko 2002). In generally speaking, instructors tend to approve that student ratings of teaching is an acceptable means of instructional improvement, but there still exist skeptical attitude of using student ratings of teaching for summative evaluation regarding to

personnel decisions.

Since 2008 undergraduate student rating of teaching has been implemented in HCM-USSH for improving instruction quality. Student ratings of teaching surveys are administered in the end of each term and completed before final exams are taken, and typically anonymous without the presence of lecturers. HCMUSSH uses paper and pencil survey instead of web-based survey. Questionnaire includes following primmary sections: student background information (6 items); their ratings on course information (5 items), course contents (7 items), lecturer teaching activities (16 items); overall ratings (2 items) and respondents comments on course con-tent, course reference material, lecture teaching activities, performance evaluation methods, facilities. After data of student ratings are collected, Office of Educational Testing and Quality Assurance analyze, summarize statistics and reports are made available across instructors, the president and faculty deans, department heads and viewed as evidences of teaching activities.

3. Research Method.

3. 1. Participants. This study was conducted at HCMUSSH where has 822 employees in which 503 serve as full time teaching staff (Centra, 1993) The questionnaire was sent to all full-time faculty (N=503). 283 out of 503 faculty members, accounting for 56.3% completed the survey. The sample quite parallels the university population in terms of demographic variables. The demographic information for participants is summarized in Table 1 below.

3. 2. Instrument. This study used questionnaire survey to collect data. The survey, consisted of three sections, was developed by the researcher with counseling from assessment experts and referring to other studies concerning student ratings of teaching or student evaluation of teaching. The first section was participant background information, such as: gender, academic title, teaching experience. The second one focused on instructor perceptions on using student ratings of courses for purposes of teaching with eight items. The last section examined instructor perceptions of using student ratings for instructional improvement, and their attitudes toward student ratings of instruction as well as utilizing student rating results for improving teaching practice. The third section included 16 items. Exception demographic variables, 24 items of two last sections use five - point Likert scale, which score 5 presenting the highest level of agreement.

3. 3. Procedure. The questionnaire survey was twice-mailed to all currently full-time teaching staff from May 15th to June 5th. The first time sent copy of survey and a cover letter explaining the study purpose and calling their help for completing the survey. Response rate for the first time is 39.2% (n=197). With unresponsive mails received the reminder email for the purpose of following-up. The following-up letter along with an additional copy of survey was mailed to them. Finally, there were 88 lecturers returning their feedback with valid responded surveys. There were only two responses excluded from data analysis because more than 35% items not completed, leaving a total sample of 283. The data from the survey was entered into SPSS.

Descriptive statistics were employed for this study. Frequencies and percentages were described the demographic information of sample. Meanwhile average and standard deviation of each item, except demographic information, were computed. In addition, inferential statistics, independent sample T-test and one-way ANOVA were applied to find out significantly differences of each item among various groups in terms of gender, academic title, years of teaching experience.

3. 4. Research Questions. Based on literature review, two research questions were formulated that served as the foundation for this study:

1. To what extend instructors agree with using student ratings of teaching for purposes of improving teaching?
2. Are there differences in instructor perceptions of using student ratings of teaching for teaching purposes based on gender, academic title, and years of teaching experience?

4. Findings. This section presents the findings of the research. Descriptive statistics of demographic information and instructor perceptions on using student ratings for teaching purposes and improving teaching quality is first displayed. Then independent sample T-test examines the influence of gender on the respondents to each of eight statements. Finally, ANOVA figures out significantly difference of independent variable, like academic title, years of teaching experience, for each statement. The questionnaire asked for demographic information about each participants, including gender, academic title, and years of experience. As shown in table 1, female sample (53%) was more slightly than male sample (45%), a common practice in disciplinary of social sciences and humanities. Majority of instructors in HCMUSSH has been holding master degree (53%), and only 27% of participants have been doctorate whereas 18% have been still bachelor. The years of teaching experience ranged from 1 to 34 with an average of 10.85 years and were grouped into 5 categories:≤5years, 6-20 years, and> 31years. So, most of instructors in HCMUSSH are young generation.

The first research question was to what extend instructors agree with using student ratings of teaching for purposes of improving teaching and which was addressed in following section. Table 2 displays participant perceptions on using student ratings for teaching purposes at HCMUSSH. The results indicate that in general, the majority of instructors agreed and strongly agreed using student ratings for refining teaching method (n = 241, 87%). Most respondents (n = 212, 75%) agreed and strongly agreed with the statement that using student ratings for improving instructors treatment of students. It is important to note that only 53% (n = 150) supported the use of student ratings for improving overall teaching quality. Having 47% (n = 135) respondents refined teaching contents and 43% (n = 121) of them refined instruction objectives based on student ratings. Approximately 35% lecturers felt student ratings help to modify mid C term and final exams as well as alter course textbooks. Lastly, there was only 32% (using student rating for selecting support material. In sum, agreement proportion of using student ratings for teaching purposes is not really high

TABLE 1. Distribution of faculty members background variables (N=283)

Background Variables		Frequency	Percentage	Mean SD
Gender	Male	126	45	
	Female	150	53	
	Not specified	7	2	
Academic title	Bachelor degree	50	18	
	Masters degree	149	53	
	Doctorate degree	77	27	
	Not specified	7	2	
Years of teaching experience	5 years	106	37	
	6-20 years	107	38	10:85 ± 9:3
	≥21 years	56	20	
	Not specified	14	5	

TABLE 2. Mean and frequency of using student ratings for teaching purposes
according to lecturer perceptions (N=283)

Student ratings used	Mean	Strongly disagree	Disagree	Neutral	Agree	Strongly agree	Rank (mean)
Improving overall teaching quality	3.58	14 (5%)	28 (10%)	91 (32%)	79 (28%)	71 (25%)	3
Improving instructor's treatment of students	3.94	0	14 (5%)	57 (20%)	142 (50%)	70 (25%)	2
Refining instructional objectives	3.25	22 (8%)	35 (12%)	105 (37%)	93 (33%)	28 (10%)	5
Refining teaching contents	3.35	8 (3%)	63 (22%)	77 (27%)	86 (30%)	49 (17%)	4
Refining teaching methods	4.20	0	0	35 (13%)	150 (54%)	91 (33%)	1
Modifying midterm and final exams	3.16	21 (7%)	56 (20%)	106 (36%)	58 (21%)	42 (15%)	6.5
Altering course textbooks	3.16	21 (7%)	64 (23%)	99 (35%)	64 (23%)	35 (12%)	6.5
Selecting support material	3.10	21 (7%)	57 (20%)	113 (40%)	57 (20%)	35 (12%)	8

Table 3 presents the results of instructors mean, frequency and percentage of agreement ratings on using student ratings of teaching for instructional improvement and practical issues related to implement student ratings of them. Of the 16 items in table 3 below, the highest scores of instructor ratings included I feel my career has been harmed to some degree by student ratings I have received (4.05), The same instrument of student ratings cannot be appropriate for all courses of different disciplines (4.05), Student ratings

should conduct every semester (3.68). The lowest scores were obtained on the following items: Student should not evaluate instructors (1.75),I do not really know how to use the results of student ratings to improve my teaching or my course (1.89), I know enough about statistics to interpret the results of student ratings without assistance (1.97), Student ratings cant really provide useful feedback needed for improving the quality of instructors teaching (2.03), Faculty members in general to tend to water down their requirements in order to get favorable ratings (2.42). Remaining items kept in medium scores, comprise: I am improving my teaching or my course from semester to semester based on ratings of students (3.47), Student ratings provide reliable feedback for planning changes in teaching (3.10), I have never received necessary assistance from a teaching improvement specialist or master (2.94), If the evaluation was given at an earlier point in the semester I would use the student feedback right away (2.89), An otherwise poor teacher can get higher rating by lenient grading (2.81), Most of students take the evaluation process seriously (2.75), The use of student ratings in my institutes provides no or little benefit to the quality of instruction students receive (2.59), Ratings of students consistent with peer observations or/and own assessment (2.55). In sum, instructors approve using of student ratings for instructional improvement with different instrument for various disciplinary but they feel a little bit harmed regarding student feedback results. It is important to pay attention that they are difficult with interpreting statistic- related report if having no assistant or support.

Regarding answer research question 2 "There are differences in instructor perceptions of using student ratings of teaching for teaching purposes based on gender, academic title, and years of teaching experience?" has multiple parts. Separated part of this question will be addressed independently. For gender, we used independent sample T-test to ascertain whether there were any statistically significant differences between male and female teachers for each item. Two out of eight items were found statistically significant difference at level 0.05: Selecting support material (p=0.020), (Meanmale=2.99, Meanfemale =3.29); and Altering course text-books (p=0.015), (Meanmale=2.98, Meanfemale =3.29). It could be said that female instructors found usefulness of student ratings for selecting teaching materials than did male counterparts. Analysis of variance (ANOVA) was employed to process first 8 items about lecturer perceptions of using student ratings for purposes of teaching with background variables, academic title and years of teaching experience, using as independent variables. The results are presented in following table with statistically significant at the 0.05 level for each factor. There were eight statements that differed significantly in regards to academic title. Specifically, bachelor degree holding instructors rated seven out of eight statements higher than did those with master degree holding, except statement Selecting support material. They also scored significantly higher than did those with doctorate degree holding on six items: Refining instructional objectives, Refining teaching contents, Refining teaching methods, Modifying mid-term and final exams, Altering course textbooks, Selecting support material. These statements: Modifying mid-term and final exams, Altering course textbooks, Selecting support material had significantly higher ratings for master degree holding instructors in comparing with doctorate degree holding ones. It is easily to recognize that the lower level of educational degree instructor holding, the higher rating of student feedbacks usefulness for teaching purposes. Regarding years of teaching experience, six out of eight statements were found to

vary significantly between the levels. Less than 5 - year of teaching experience and 6-20 year of teaching experience instructors rated statement Refining instructional objectives, Refining teaching contents, Refining teaching methods, Modifying mid-term and final exams, Altering course textbooks, Selecting support material higher than did those with more than 20 years of teaching experience. This finding confirms that experienced instructors found less meaningfulness of student feedback for their teaching improvement.

TABLE 3. Instructors mean frequency and percentage of agreement ratings

Student ratings used	Mean (rank)	Strongly disagree	Disagree	Neutral	Agree	Strongly agree	N
The same instrument of student ratings cant be appropriate for all courses of different disciplines	4.05 (1.5)	7 (3%)	15 (6%)	56 (21%)	65 (25%)	119 (45%)	262
I feel my career has been harmed to some degree by student ratings I have received	4.05 (1.5)	99 (35%)	120 (42%)	43 (15%)	14 (5%)	7 (3%)	276
Student ratings should conduct every semester	3.68 (3)	21 (7%)	35 (12%)	43 (15%)	99 (35%)	85 (30%)	283
Refining teaching contents	3.35	8 (3%)	63 (22%)	77 (27%)	86 (30%)	49 (17%)	4
I am improving my teaching or my course from semester to ratings of students	3.47 (4)	7 (3%)	21 (7%)	98 (36%)	136 (49%)	14 (5%)	276
Student ratings provide reliable feedback for planning changes in teaching	3.10 (5)	7 (3%)	56 (20%)	142 (50%)	64 (23%)	14 (5%)	283
I have never received necessary assistance from a teaching improvement specialist or master	2.94 (6)	49 (18%)	43 (15%)	107 (40%)	28 (10%)	42 (16%)	283
If the evaluation was given at an earlier point in the semester I would use the student feedback right away	2.89 (7)	22 (8%)	64 (30%)	113 (70%)	77 (27%)	7 (2%)	269

An otherwise poor teacher can get higher	2.81 (8)	50 (18%)	64 (23%)	85 (31%)	42 (15%)	35 (13%)	276
Most of students take the evaluation process seriously	2.75 (9)	14 (5%)	98 (35%)	115 (41%)	56 (20%)	0	283
The use of student ratings in my institutes provides no or little benefit to the quality of instruction students receive	2.59 (10)	49 (18%)	99 (36%)	72 (26%)	28 (10%)	28 (10%)	276
Ratings of students consistent with peer observations or/and own assessment	2.55 (11)	14 (5%)	112 (43%)	122 (47%)	7 (3%)	7 (3%)	262
Faculty members in general to tend to water down their requirements in order to get favorable ratings	2.42 (12)	64 (21%)	92 (33%)	78 (28%)	42 (15%)	7 (3%)	283
Student ratings cant really provide useful feedback needed for improving the quality of instructors' teaching	2.03 (13)	85 (30%)	120 (42%)	63 (22%)	15 (5%)	0	283
I know enough about statistics to interpret the results of student ratings without assistance	1.97 (14)	0	7 (3%)	64 (23%)	114 (41%)	91 (33%)	283
I do not really know how to use the results of student ratings to improve my teaching or my course	1.89 (14)	84 (31%)	136 (51%)	42 (16%)	7 (3%)	0	269
Student should not evaluate instructors	1.75 (16)	141 (50%)	85 (30%)	43 (15%)	14 (5%)	0	283

TABLE 4. ANOVA results of instructor perceptions among groups of academic
title and years of teaching experience

Factor	Background variables	Mean difference	P	F (Between groups)
Improving overall teaching quality	Bachelor (A) Master (B) PhD (C)	.6307(A-B)	.002*	6.866
Improving instructors treatment of students	Bachelor (A) Master(B) PhD (C)	.3540(A-B)	.027*	4.990
Refining instructional objectives	Bachelor (A) Master (B) PhD (C)	0.9256(A-B)	.000**	21.232
		.9891(A-B)	.000**	
	<5 years (A) 6-20 years (B) >21 years (C)	.6612(B-C)	.000*	8.755
Refining teaching contents	Bachelor (A) Master (B) PhD (C)	1.0187(A-B)	.000*	19.593
		.9473(A-C)	.000*	
	<5 years (A) 6-20 years (B) >21 years (C)	.6344(A-C)	.003**	6.905
Refining teaching methods	Bachelor (A) Master (B) PhD (C)	.5107(A-B)	.000**	12.773
		.3782(A-C)	.004**	
	<5 years (A) 6-20 years (B) >21 years (C)	.4286(B-C)	.000**	8.421
Modifying mid-term and final exams	Bachelor (A) Master (B) PhD (C)	.6387(A-B)	.001*	14.701
		1.0218(A-C)	.000*	
		.3832(B-C)	.033*	
	<5 years (A) 6-20 years (B) >21 years (C)	.4481(A-C)	.041*	3.292
Altering course textbooks	Bachelor (A) Master (B) PhD (C)	.8123(A-B)	.000*	27.414
		1.3236(A-C)	.000*	
		.5113(B-C)	.001*	
	<5 years (A) 6-20 years (B) >21 years (C)	.7547(A-C)	.000**	10.097
		.5093(B-C)	.005**	
Selecting support material	Bachelor (A) Master (B) PhD (C)	-.8123(A-B)	.000**	28.946
		1.3236(A-C)	.000**	
		.5113(B-C)	.001**	
	<5 years (A) 6-20 years (B) >21 years (C)	.8797(A-C)	.000*	15.451
		.6998(B-C)	.00*	

*   Schffe post-hoc multiple comparisons statistically significant at 0.05 level

**   Dunnett T3 post-hoc multiple comparisons statistically significant at 0.05 level

5. Conclusion. Student rating of teaching is not a new topic regarding higher education. Yet, in Vietnamese tertiary education this issue is still unexplored by researchers. This current study with sample of 283 instructors at HCMUSSH indicates that: (a) instructors approve and support usefulness of student rating results for instructional improvement at moderation degree; (b) they show disagree with anecdotal controversy surrounding student ratings of teaching (eg. Faculty members in general to tend to water down their requirements in order to get favorable ratings, Student should not evaluate instructors). By using ANOVA, the findings show that Lower level of educational degree and less teaching experience instructors have higher rating of student rating usefulness for teaching purposes. In addition, female lecturers rate helpfulness of student feedback for altering instructional material significantly higher than did male ones. Limitations to this study include the fact that the data was collected at a single university, which could not generalize to other universities or colleges. Therefore, the researcher suggested further researches will be needed with larger, multi-institutional samples to better understand instructors perceptions about the use of student ratings for instructional improvement. In future work, we could improve our questionnaire design by using linguistic labels to ease the rankings. Linguistic labels can be quantified by using fuzzy set theory and statistics with fuzzy numbers.

REFERENCES

[1] A. G. Greenwald (2002), Constructs in student ratings of instructors, *The role of constructs in psychological and educational measurement, New York*.

[2] A. Rotem and N. S. Glasman (1979), On the effectiveness of students evaluative feedback to university instructors, *Review of Educational Research*, no.49, pp.497-511.

[3] F. Nasser and B. Fresko (2002), Faculty views of student evaluation of college teaching, *Assessment and Evaluation in Higher Education*, vol.27, no.2, pp.187-198.

[4] F. Zabaleta (2007), The use and misuse of student evaluations of teaching, *Teaching in Higher Education*, vol.12, pp.55-76.

[5] H. W. Marsh (1987), Students evaluations of university teaching: Research findings, methodological issues, and directions for future research, *International Journal of Educational Research*, vol.11, pp.253-388.

[6] J. A. Centra (1993), Refflective Faculty Evaluation, *Jossey-Bass, San Francisco*.

[7] J. A. Kulik and W. J. McKeachie (1975), The evaluation of teachers in higher education, *Review of research in Education,* vol.3, pp. 210-240.

[8] L. C. Jacobs (1987), University Faculty and Students Opinions of Student Rating, *Bureau of Evaluative Studies and Testing, Indiana University, Bloomington.*

[9] L. P. Schmelkin, et. al. (1997), Faculty perspectives on course and teacher evaluations, *Research in Higher Education*, vol.38, no.5, pp.575-592.

[10] Lori R. Kogan, et. al. (2010), Student evaluations of teaching: perceptions of faculty based on gender, position, and rank, *Teaching in Higher Education*, vol.15, no.6, pp.623-636.

[11] P. Hooper and J. Page (1986), Measuring teaching effectiveness by student evaluation, *Issues in Accounting Education*, Spring, pp.56-64.

[12] Peter A. Cohen (1980), Effectiveness of student-rating feedback for improving college instruction: a meta-analysis of findings, *Research in Higher Education*, vol.13, pp. 321-341

[13] T. Beran, C. Violato and T. Collin (2002), The Universal Student Ratings of Instruction Instrument at the University of Calgary: A review of a three-year pilot project, *Submitted to the Office of the Provost and Vice-President (Academic) at the University of Calgary.*

[14] T. Beran, et. al. (2005), The utility of student ratings of instruction for students, faculty, and administrators: A "consequential validity" study, *The Canadian Journal of Higher Education*, vol.35, no.2, pp.49-70.

[15] T. N. Beran and J. L. Rokosh (2009), Instructors perspectives on the utility of student ratings of instruction, *Instructional Science*, vol.37, pp.171-184.

[16] Y. Yao and M. L. Grady (2005), How Do Faculty Make Formative Use of Student Evaluation Feedback?: A Multiple Case Study, *Journal of Personnel Evaluation in Education,* no.18, pp.107-126.