

A Case Study of Student Evaluation of Teaching in University

Benjamin Chan Yin Fah

Research Associate, Centre of Excellent for Sustainable Consumption Studies, UPM

Research Associate, Institute of Gerontology, UPM

Syuhaily Osman

Research Associate, Centre of Excellent for Sustainable Consumption Studies

Faculty of Human Ecology, Universiti Putra Malaysia (UPM)

Abstract

This study aims to determine the factors (course characteristics, lecturer characteristics, and tutorial ratings) that affect student evaluation of teaching in university. A total of 88 undergraduates were selected and self-administered questionnaire was used as a tool for data collection. The study found that most of the respondents have high agreement level towards the evaluation of course characteristics, lecturer's characteristics, and tutorial ratings. Lecturer overall teaching performance ratings were correlated with course characteristics, lecturer characteristics and tutorial rating. Multiple hierarchical analyses found that course's overall performance ratings was mostly explained by course, followed by lecturer's characteristics but not from tutorial ratings. From the point of a student, the improved of the teaching effectiveness based on the evaluation process may ultimately enhance knowledge acquisition and, for the educators, the evaluation did provide information for an individual improvement. Faculties may benefit where the evaluation might lead to fairer promoting, tenure and pay increase decisions for academic staffs.

Keywords: Course characteristics, Lecturer characteristics, Tutorial, Lecturer overall performance rating

1. Introduction

Nowadays, faculties are being held responsible for how well they serve the student population, and it has become common practice in universities and colleges for students to evaluate their lecturers that grade them (Germain & Scandura, 2005). In a general view, students who have spent the whole semester of the course, observing the lecturer in the class and how he or she interacts with students, and thus the students would be able to evaluate both for the course contents and the lecturer's overall teaching performance as well. (Theall & Franklin, 2001; Mukherji & Rustagi, 2008). Various studies have proposed different criteria on how to assess teaching performance. They differ according to different schools specific objectives but in general, evaluation goes into three main areas, namely: lecturer characteristics, the course contents and tutorial, if any.

A lecturer's characteristics in the class may have significant influence towards the students' learning process. For example, students tend to like their lecturers who know what they are talking about, having a good sense of humor, are warm and friendly, and who are concerned about their students (Magno & Sembrano, 2008; Hattie, 2009) where students are stimulated to greater achievement and learning if the lecturers have related characteristics that make them effective teachers. Nevertheless, a lecturer's unfavorable characteristics or quality may elicit poor performance feedback from his or her students and directly link to student achievement.

Course characteristics are taken into consideration when the students evaluate their lecturer. Course difficulty is negatively related to student ratings, indicating that students who feel more capable in learning the subject rate their teachers of the course more favorably (Chang, 2000). Course effectiveness, as perceived by students, could not be attributed entirely to the competence of the instructor, in that he or she was dependent for effective instruction on both active students and supportive conditions (Rindermann & Schofield, 2001). Challenging and average-difficulty courses are evaluated higher than easy courses, and average or above-average workloads are valued more than below-average workloads (Mukherji & Rustagi, 2008). Most of the subjects in university level involves tutorial classes and the overall teaching performance or teaching quality of a lecturer were partially depend on the teaching behavior of their instructor/s, but, at the same time, it is a social process involving participation such as likes discussion and interaction between students (Rindermann & Schofield, 2001). The role of a teaching assistant is frequently different than the meaning the Graduate Teacher Assistant (GTA) title suggested. In many university settings, these graduate students are not merely assistants at all. Instead, they are frequently take responsibility into the role of college instructor with little or no training in instructional strategies (Roach, 1997).

2. Objective and Significance of Study

Generally, if the course characteristic, lecturer characteristic, tutorial ratings and the overall performance of the

lecturer are evaluated with good grades from the students, the students are assumed to learn more from the course being taught. Nevertheless, with more knowledge the students learned from the courses, the students may have better academic achievement in the university and brighter future career development. The purpose of the study was to investigate the factors that affect student evaluation of teaching in university. The respondents are collected from those who attended one of the personal finance classes in Faculty of Human Ecology, Semester 2009/2010. The specific objectives of the study are detailed below.

To determine the lecturer evaluation score based on course characteristics, lecturer characteristics, tutorial assessed by the respondents in this study.

To determine the lecturer overall performance level assessed by the respondents in this study.

To identify the predictors of good performance teaching rating among the respondents in this study.

3. Methodology

3.1 Sampling and Location

Since the primary concern in this study is to explore the student evaluation of teaching in university, the research should be framed within universities and those who had finished the entire course within the semester. Data from this study was drawn from a small dataset collected by two final year undergraduates as a partial requirement for their Bachelor degree study. The sample size was $N = 88$. For sample selection, the two undergraduate students, upon obtaining the approval from their department's lecturer went into the classroom and requested all there in attendance to fill-in the questionnaire in self-administered way. In brief, the class was mainly offered to second-year-student from various programs likes Bachelor of Science (Human Development), Bachelor of Science (Human Development and Management), Bachelor of Science (Human Development and Information Technology) and Bachelor of Science (Consumer Studies). It is a three credit hours class with two hours lecturing and three hours tutorial class in a week. The selection of the course subject and the lecturer was just convenient.

3.2 Measurement of Variables

Course characteristics were measured using the six items in the evaluation sheet. This self-administered questionnaire was designed to assess the agreement or disagreement of the students toward the course characteristics. The instrument was rated on a five-point ordinal scale, ranging from "strongly disagree" to "strongly agree". The overall score of the scale was obtained by summing the scale of the six items. A higher total score indicated a stronger level of agreement towards the course characteristics. A sample item of the course characteristics is "The course given information is just right". Using the mean score of the instrument as cut off point, those who had scored between 6 and 18 were considered to have low level of agreement; where as those with the scores above 18 were considered to have high level of agreement.

For lecturer characteristics, it consisted of 20 items related to lecturer characteristics, where the criteria included planning, delivery, assessment, and professionalism. The items were rated on a five-point ordinal scales ranging from 'strongly disagree' to 'strongly agree'. A higher scale score indicated a stronger level of agreement toward the lecturer characteristic. Some sample items are "Teaching plan given is clearly explained", "Learning activities are managed effectively", "Assessment are implemented as planned", and "Lecturer shows concerns and demonstrates responsibility". The minimum for the summation score was 20 and 100 the maximum while score ranging from 20 to 60 means low level of agreement towards the evaluation of lecturer characteristics and "61 to 100" was categorized as high level of agreement.

Tutorials ratings were measured using four items with a five-point ordinal scale, ranging from "strongly disagree" to "strongly agree". Total scores was the sum of four items. A higher total score indicated a stronger level of agreement towards the tutorial ratings. A sample item of the tutorials is "The scope of task is just right". Using the mean score of instrument as cut off point, those who scored between 4 and 12 were considered to have low level of agreement; where as those with the scores above 12 were considered to have high level of agreement.

The lecturer's overall performance rating was measured using the last two items in the faculty's evaluation form. The items were rated on a five-point scales ranging from 'strongly disagree' to 'strongly agree'. A higher scale score indicated a stronger level of agreement toward the lecturer overall performance rating. The samples of the two items are "Overall, the teaching of this course is effective" and "Overall, this lecturer is one of the best educators".

3.3 Data analysis

The data collected was analyzed using Statistical Package for Social Science for Windows (SPSS for Windows Version 13.0). Descriptive analysis used frequency and percentage to examine the demography characteristics of the

respondents. Four types of data analysis were used: Pearson Moment Correlation, Independent sample t-test, ANOVA and multiple hierarchical regressions. The significance level was determined at a probability level of 5%.

4. Research Findings

4.1 Demographic Characteristics of the Respondents

This study consisted of 18 male respondents and 70 female respondents.. There were a total of 38 students in the age group of 18 to 21; while the other 50 students were aged 22 and above. Among the respondents, 43.2 % were aged between 18 years old and 21 years old. 56.8% of the respondents were aged 22 years old and above. Regarding the respondents' ethnicity, 44.3% were Chinese, 42% were Malays, 8% were Indians, and the remaining 5.7% were other nationalities. A majority of the respondents were in their third year of study, 48.9%, followed by 44.3% of those who were in their second year of study, while fourth year students were 6.8%. More than half of the total respondents 62.5% were taking the program of Bachelor of Science (Human Development), while 25% of the total respondents were taking Bachelor of Science (Human Development and Management), 11.4% of the total respondents were taking Bachelor of Science (Human Development and Information Technology), and 1.1% of the respondents from other faculties.

4.2 Student Evaluation towards Course Characteristics

For each item analysis, most of the responses revealed that the course given information is just right (Q1), the level of difficulty is just right (Q2), the arrangement of information and topics is clear (Q3), and the given examples are helpful (Q6). However, there are some responses revealed that the illustration of the topics is merely descriptive (Q4) and the course's presentation is not that appealing (Q5). A cross tabulation analysis found that both gender rated the course with a high level of agreement towards course characteristics, amounting to 98.6% of the female students and 88.9% of the male students. For those who rated a high level of agreement, 81.2% were female students and 18.8% were male students. For the age group, 85 students rated a high level of agreement, where 42.4% of them were aged between 18 and 21; while 57.6% of them were 22 years old or above. In terms of ethnicity, among the 85 students who rated a high level of agreement, 44.7% of the 85 students were Chinese, 41.2% of them were Malays, 8.2% of them were Indians, and 5.9% of them were other ethnics. For the year of study, 49.4% of the 85 students who rated a high level of agreement were in their third year of study in university, 43.5% of them were second year students, while 7.1% of them were fourth year students. Results from the analyses showed that there was no significant difference in course characteristics by different age group ($t= 1.06, p > .05$), ethnicity ($\chi^2= 50.44, p > .05$), year of study ($\chi^2= 11.15, p > .05$, or even studied program ($\chi^2= 65.2, p > .05$).

[Insert Table 1 here]

4.3 Student Evaluation towards Lecturer Characteristics

Both gender rated the lecturer with high level of agreement, amounting to 100% of the female students and 88.9% of the male students. For those who rated a high level of agreement, 81.4% were female students and 18.6% were male students. For the age group, 86 students rated a high level of agreement, where 58.1% of them were 22 years old or above; while 41.9% of them were aged between 18 and 21. In terms of ethnicity, among the 86 students who rated a high level of agreement, 45.3% were Chinese, 40.7% were Malays, 8.1% were Indians, and 5.8% were other nationalities. For the year of study, 50% of the 86 students who rated a high level of agreement were in their third year of study in university, 43% of them were second year students, while 7% of them were fourth year students. Results from the analyses showed that there was no significant difference in course characteristics by sex, ($t= .565, p > .05$), age group ($t= .036, p > .05$), year of study ($\chi^2= 47.38, p > .05$) and studied program ($\chi^2= 501, p > .05$). However, there is a significant difference in lecturer characteristics by ethnicity ($\chi^2= 865, p < .05$) where the Indians in this study rated the highest evaluation to the lecturer ($\bar{x}=88.71$), followed by Chinese ($\bar{x}=85.77$), other nationality/ethnicity ($\bar{x}=83.40$) and Malay ($\bar{x}=79.92$).

4.4 Student Evaluation towards Tutorial Ratings

First of all, the teaching assistant of the class is different with the instructor. In fact, the teaching instructor might have more time (three hours tutorial time) with the students compared to the instructor (two hours lecturing time). In general, a majority of the respondents graded their teaching assistant with high compliments in terms of her level of task (Q1), teaching coordination with her instructor (Q2), tutorial handout (Q3) and support to students (Q4). A closer inspection of the cross tabulation between the tutorial rating level with selected variables might find that both gender rated the course with high level of agreement towards tutorial ratings, amounting to 97.1% of the female students and 88.9% of the male students. For those who rated a high level of agreement, 81% were female students and 19% were male students. For the age group, 84 students rated a high level of agreement, where 58.3% of them were 22 years old or above; while 41.7% of them were aged between 18 and 21. For the year of study, half of the 84

students who rated a high level of agreement were in their third year of study in university, 42.9% of them were second year students, while 7.1% of them were fourth year students. Results showed that there is no significant difference in tutorial rating between male and female respondents ($t = .149$, $p > .05$), respondents' age group ($t = .233$, $p > .05$), respondents' ethnicity ($\chi^2 = 27.71$, $p > .05$), respondents' year of study ($\chi^2 = 5.21$, $p > .05$), or their study program ($\chi^2 = 13.63$, $p > .05$).

4.5 Predictors of Lecturer Overall Performance

Pearson Product Moment Correlations were computed on selected pairs of variables to test for direction and strength of the correlation between variables. Examination of the correlation of the variables, the respondents' evaluations of a lecturer's overall performance rating was significantly related to the course characteristics ($r = .689$), lecturer characteristics ($r = .755$) and tutorial rating ($r = .661$).

[Insert Table 2 here]

Table 3 is the summaries table of a multiple regression analyses of the lecturers' overall performance rating. Two in three variables significantly contribute to lecturer overall performance rating, namely: course characteristics and the lecturer characteristics. This result demonstrates that students rated high for the course characteristics are more likely to give high ratings to the lecturer overall performance. This finding is consistent with findings of Rindermann and Schofield (2001), which stated that in the instructor scales, very high levels of stability were found, suggesting that in the same courses being taught by the same teachers, the perception of teaching behavior by different students was stable. The F statistic for the overall goodness of fit of model is 45.041, which is significant at $\alpha = 0.01$. After excluding the non-significant variables, the final regression model produced by enter method for lecturer overall performance rating: $-0.355 + 0.047(\text{Course Characteristics}) + 0.031(\text{Lecturer Characteristics}) + \text{error}$. This final model explained 60 per cent variance in Lecturer overall performance rating and this indicated that 60% of the dependent variable was explained by the two factors. Of the two variables, the quality of the course itself was the most dominant predictor for the overall rating compared to the lecturer's characteristics. Tutorial rating did not contribute to the final model even though the correlation with overall performance is significant in the bivariate level.

[Insert Table 3 here]

5. Conclusion, Implication and Recommendation

The major purpose of this study was to determine the relationship between course characteristics, lecturer characteristics, and tutorial ratings with lecturer overall performance of one of the personal finance classes. The results showed that respondents gave a high level of rating towards course characteristics, lecturer characteristics, and tutorial of the course being taught by one of the lecturers of the Faculty of Human Ecology. Based on the respondent's gender group, the findings showed that there was no significant difference in the ratings towards course characteristics, lecturer characteristics, and tutorial between males and females. On the other hand, for ethnicity differences, Indians tend to give a higher scoring compared to Chinese and Malay in evaluating their lecturer's characteristics. This could be because of the course was taught in English, where the English is more dominant among the Indian students. The findings also showed that there is a positive relationship between course characteristics, lecturer characteristics, and tutorial with lecturer overall performance.

This study has several implications towards certain groups. First, the lecturer may use the result of this study to have a better understand regarding what his or her students' evaluations on him or her, as well as the course characteristics, the lecturer characteristics and the tutorial. Furthermore, the teaching evaluation not only provided a lot of information which might be used for managing the studies but could also be used for individual improvement by the lecturer. The key function of the system of higher education quality assessment is a purposeful collection of information, its analysis, interpretation and generalizing and thus, faculties would benefit where the faculty administration might employ the same method in determining the teaching effectiveness and the overall performance of the lecturers. This approach will bring more relevance and understanding to how to interpret the teaching evaluations, which could lead to fairer promoting, tenure and pay increase decisions for academic staffs, improved job satisfaction, and ultimately a sense of justice. For the students, the improvement of the teaching effectiveness based on the evaluation process may ultimately enhance knowledge acquisition.

The sample size can be increased by involve more classes in the research in order to generate more reliable results. However, a word of caution should be noted. The data in this study were gathered from just one course in one university. So, we are not generalizing the results to other courses, lecturers, or universities across the nation. Plus, the accuracy of the data provided by the respondents depends on their honesty as well as their understanding of the questions asked. The respondents' skills of evaluating based on the criteria are different which is due to their

personal judgment of the scale used in the questions. Lastly, limited time and cost are the constraint faced when conducting this study where only two classes are selected and thus, the result may not as accurate and reliable.

References

- Beran, T., Violato, C., Kline, D. & Frideres, J. (2005). The Utility of Student Ratings of Instruction for Students, Faculty, and Administrators: A "Consequential Validity" Study. *The Canadian Journal of Higher Education*, 35 (2), 49-70.
- Chang, T.S. (2000). Student Ratings: What Are Teacher College Students Telling Us about Them? Paper presented at the Annual Meeting of the American Educational Research Association (New Orleans, LA, April 24-28, 2000).
- Emery, C. R., Kramer, T. R. & Tian, R.G. (2003). Return to Academic Standards: A Critique of Student Evaluations of Teaching Effectiveness. *Quality Assurance in Education*, 11(1), 37-46.
- Germain, M.L. & Scandura, T. A. (2005). Grade Inflation and Student Individual Differences as Systematic Bias in Faculty Evaluations. *Journal of Instructional Psychology*, 32(1), 58-67.
- John C. Hattie. (2009). *Visible Learning: A Synthesis of Over 800 Meta-Analyses Relating to Achievement*. Routledge, London & New York.
- Kupermintz, H. (2003). Teacher Effects and Teacher Effectiveness: A Validity Investigation of the Tennessee Value Added Assessment System. *Educational Evaluation and Policy Analysis*, 25 (3), 287-298.
- Lang, J. W. B. & Kersting, M. (2006). Regular Feedback from Student Ratings of Instruction: Do College Teachers Improve their Ratings in the LongRun? *Instructional Science*, 35, 187-205.
- Magno, C., & Sembrano, J. (2008). The Role of Teacher Efficacy and Characteristics on Teaching Effectiveness, Performance, and Use of Learner-Centered Practices. *The Asia Pacific-Education Researcher*, 16(1), 73-90.
- Morgan, D.A., Sneed, J. & Swinney, L. (2003). Are Student Evaluations A Valid Measure of Teaching Effectiveness: Perceptions of Accounting Faculty Members and Administrators. *Management Research News*, 26(7), 17-32.
- Mukherji, S. & Rustagi, N. (2008). Teaching Evaluations: Perception of Students and Faculty. *Journal of College Teaching & Learning*, 5(9), 45-54.
- Rindermann, H. & Schofield, N. (2001). Generalizability of Multidimensional Student Ratings of University Instruction Across Courses and Teachers. *Research in Higher Education*, 42 (4), 377-399.
- Roach, K. D. (1997). Effects of Graduate Teaching Assistant Attire on Student Learning, Misbehaviors, and Ratings of Instruction. *Communication Quarterly*, 45 (3), 125-141.
- Rocca, K. A. & McCroskey, J.C. (1999). The Interrelationship of Student Ratings of Instructors' Immediacy, Verbal Aggressiveness, Homophily, and Interpersonal Attraction. *Communication Education*, 48, 308-316.
- Stankeviciene, J. (2007). Assessment of Teaching Quality: Survey of University Graduates. Paper presented at the European Conference on Educational Research (University of Ghent, September 19-21, 2007).
- Theall, M. & Franklin, J., (2001) *Looking for Bias in All the Wrong Places: A Search for Truth or a Witch Hunt in Student Ratings of Instruction?* New Directions for Institutional Research, No.109. San Francisco: Jossey Bass.
- Worthington, A.C. (2002). The Impact of Student Perceptions and Characteristics on Teaching Evaluations: A Case Study in Finance Education. *Assessment and Evaluation in Higher Education*, 27(1), 49-64.

Table 1. Descriptive Statistics of Level of Agreement towards Evaluation of Course Characteristics, Lecturer Characteristics, Tutorial Ratings by Respondents

Variables		Course		Lecturer		Tutorial	
		Low	High	Low	High	Low	High
		%	%	%	%	%	%
Gender	Male	66.7	18.8	100	18.6	50.0	19.0
	Female	33.3	81.2	0	81.4	50.0	81.0
Age Group	18-21	66.7	42.4	100	41.9	75.0	41.7
	22 and above	33.3	57.6	0	58.1	25.0	58.3
Ethnicity	Malay	66.7	41.2	100	40.7	75.0	40.5
	Chinese	33.3	44.7	0	45.3	25.0	45.2
	Indian	0	8.2	0	8.1	0.0	8.3
	Others	0	5.9	0	5.8	0.0	6.0
Year of Study	Second Year	66.7	43.5	100	43.0	75.0	42.9
	Third Year	33.3	49.4	0	50.0	25.0	50.0
	Forth Year	0	7.1	0	7.0	0.0	7.1
Studied Programme	BS (PM)	66.7	62.4	100	61.6	75.0	61.9
	BS (PM & P)	33.3	24.7	0	25.6	25.0	25.0
	BS (PM & TM)	0	11.8	0	11.6	0.0	11.9
	Others	0	1.2	0	1.2	0.0	1.2

Table 2. Correlation between Lecturer Overall Performance Rating, Course Characteristics, Lecturer Characteristics and Tutorial Ratings

Variable	Lecturer Overall Performance		
	R	R ²	Sig (2-tailed)
Course Characteristics	.689**	.475	.000
Lecturer Characteristics	.755**	.570	.000
Tutorial	.661**	.437	.000

Table 3. Summary of Multiple Regression Analyses for Financial satisfaction

Model	Variable (N = 274)	B	SE B	β	p
1	(Constant)	.839	.393		.035
	Course Characteristics	.141	.016	.689	.000
2	Course Characteristics	-.282	.408		.490
	Lecturer Characteristics	.054	.022	.265	.015
3	Course Characteristics	.039	.008	.552	.000
	Lecturer Characteristics	.047	.022	.227	.037
	Tutorial Characteristics	-.355	.403		.380

Note. R² = 0.474 for model 1; Δ R² = 0.599 for model 2, Δ R² = 0.617 for model 3

* p < 0.05, ** p < 0.01

Appendix 1: Evaluation Form

Please tick according to your level of agreement regarding to the teaching and learning item using the scale below.

1= Strong disagree 2= Disagree 3= Partially agree 4= Agree 5= Strongly agree

No	Item	1	2	3	4	5
1	The course given information is just right	<input type="checkbox"/>				
2	The level of difficulty is just right	<input type="checkbox"/>				
3	The arrangement of information and topics is clear	<input type="checkbox"/>				
4	The illustration of the topics is descriptive	<input type="checkbox"/>				
5	The course's presentation is appealing	<input type="checkbox"/>				
6	The examples are helpful	<input type="checkbox"/>				
7	Teaching plan given is clearly explained	<input type="checkbox"/>				
8	Lecturer is always prepared for each teaching session	<input type="checkbox"/>				
9	Learning activities are managed effectively	<input type="checkbox"/>				
10	Teaching delivery is easily understood	<input type="checkbox"/>				
11	Important aspects are emphasized in teaching	<input type="checkbox"/>				
12	Lesson contents are summarized	<input type="checkbox"/>				
13	Delivery method stimulates my thoughts	<input type="checkbox"/>				
14	Delivery method enhances my interest in the subject	<input type="checkbox"/>				
15	Illustrations/ example used are helpful in the learning process	<input type="checkbox"/>				
16	Teaching materials/ aids used help in the learning process	<input type="checkbox"/>				
17	Students are encouraged to ask questions/ to express opinion	<input type="checkbox"/>				
18	Students are exposed to current issues/ industrial practices	<input type="checkbox"/>				
19	Assessment are implemented as planned	<input type="checkbox"/>				
20	The assignment given conformed with the course content	<input type="checkbox"/>				
21	Feedbacks on test/ laboratory work are given	<input type="checkbox"/>				
22	Assessment marks are announced within stipulated period	<input type="checkbox"/>				
23	Lecturer shows concerns and demonstrates responsibility	<input type="checkbox"/>				
24	I am being respected as a student	<input type="checkbox"/>				
25	Lecturer portrays a good image	<input type="checkbox"/>				
26	The lecturer's delivery is convincing	<input type="checkbox"/>				
27	The level of the task is just right	<input type="checkbox"/>				
28	The teaching assistant's coordination with the lecturer is good	<input type="checkbox"/>				
29	The tutorial handouts are helpful	<input type="checkbox"/>				
30	The given support is very good	<input type="checkbox"/>				
32	Overall, the teaching of this course is effective	<input type="checkbox"/>				
33	Overall, this lecturer is one of the best educators	<input type="checkbox"/>				

Note: Course's Characteristics (C1-C6); Lecturer's Characteristics (C7-C26); Tutorial's Characteristics (C27-C30)