Comparison of Teaching Quality Scores of Teachers as Evaluated by Students of Different Grade Point Averages

Zahra Kiani[1], PhD; Afagh Zarei[2*], MS

[1]Assistant Professor, School of Medicine, Birjand University of Medical Sciences, Birjand, Iran
[2]PhD Student in Medical Education, Education Development Center, Birjand University of Medical Sciences, Birjand, Iran

Abstract

Background: Student Evaluation of Teaching (SET) is a long-established tradition for teacher evaluation. Teaching evaluation has many proponents and opponents. Meanwhile, the probable effects of confounding factors such as student and teacher's gender, student grade, and teacher's personality traits have always been studied. Given the contradictory results reported in the literature and opinions of teachers regarding the potentially higher value to be diffused with the views of the top students, this study was conducted to evaluate the effect of total grade point average (GPA) of students on their evaluation scores of teachers. The study also investigated the associations between teacher's gender, teacher's academic degree and rank, and the scores of teaching quality.

Methods: In this descriptive study, analyses were made in four consecutive academic semesters from 2015 to 2016. A total of 7518 questionnaires, previously confirmed in terms of validity and reliability, were completed by students of medicine, dentistry, nursing and midwifery, allied medicine, and health schools. The questionnaires contained 15 items scored on a 5-point Likert scale. Data were analyzed in SPSS software version 16 using ANOVA and Scheffé's post-hoc test. The level of significance was considered at 0.05 for all statistical tests.

Results: The findings showed that teacher evaluation scores varied significantly across different GPA categories. The difference was evident between the categories A and C as well as B and C. The relationship between gender, academic rank and degree of the teacher, and teaching quality score was not significant in any discipline.

Conclusion: Significant differences were found between the categories A and C as well as B and C. However, given the concerns about inflation or trade of scores and the possibility of ignoring the students by teachers if only assessment by top students is considered, it is better to use 360-degree evaluation which is a multi-dimensional evaluation. This prevents from probable biases in student surveys.

Keywords: TEACHER EVALUATION, FACULTY EVALUATION, STUDENT EVALUATION, TEACHING QUALITY EVALUATION

Introduction

Education is a process that involves both teaching and learning. The main clients of this process in a university are teachers and students. To ensure learning in students, which is the main goal of education, the students are customarily evaluated in an academic process, and their grades are used to provide feedback. This helps solve educational problems and promotes their academic achievement (1). Every year, a wide range of studies is conducted on new methods of student assessment so that the optimal method of evaluation can be employed, and the fair grades of students' real learning can be achieved. On the other hand, how should teachers, as the other pole of the

***Corresponding author:** Afagh Zarei**, PhD Student in Medical Education, Education Development Center, Birjand University of Medical Sciences, Birjand, Iran.
Phone: +98 (56) 32395643; **Fax:** +98 (56) 32440447
Email: zareiafagh@yahoo.com

learning process, be evaluated?

Assessing teachers who are the pillars of education, similar to the case of students, requires scholarly research and the elimination of problems and confusions to provide a fair assessment of their performance. Importantly, their performance plays a key role in the efficiency of the educational system (2). Obviously, a member of the faculty as a classroom instructor has duties that should be assessed accordingly. What matters about the evaluation of a teacher is his/her ability to achieve the educational goals of interest. Assessing the teaching quality by students is a long-established tradition in this regard that has been used in the academia since the 1920s (3). The first report on the assessment of teachers dates back to 1915. However, the golden age of studies in this area were the 1970s when the validity and reliability of this method and its application for developmental goals were confirmed (4). It is expected that students, as the main target of the teaching process, be considered as the fundamental parties for evaluating the teaching quality of teachers. Accordingly, in many parts of the world, the only source of teacher evaluation are students (4). Nonetheless, this is posed by some teachers as a major challenge in the evaluation process where they posit that survey from students about the quality of teaching is associated with bias and that students lack sufficient knowledge for evaluation, and therefore the results can be unreliable (3-6). These discrepancies have made the multi-dimensional or 360-degree evaluation proposed by the American Accreditation Council of Graduate Medical Education (ACGME) for assessing residents to be used similarly for teachers. This type of evaluation evaluates the teacher's performance from a variety of aspects. Although multiple sources are surveyed in this approach, the role of students as one of the most important sources of evaluation remains intact (7, 8).

Research shows that factors such as academic status and popularity of the teacher, teacher and student's gender, teacher's personality traits and social behavior, level of difficulty of the course, the time at which the class is held, evaluation time, the student's expected score, the degree to which the teacher is serious or funny, the academic competency of the teacher, the necessity of the course, the employment status and tenure of the teacher, educational level of the student, type of course, and academic performance of the student contribute to students' evaluations (2, 5, 9). Meanwhile, an important concern for a large number of teachers is the impact of students' educational performance on the teacher's assessment score, which attracts greater attention than other factors. Some faculties believe that their evaluation should be made by top students who have good academic performance in order to yield reliable results. Several studies have been performed in this regard, some of which support this idea (10, 11) while a few are contrary to it (12-14). Some have also associated the use of student surveys with biases (3, 15, 16), whereas, importantly, all students are the audience to the teacher's teaching activities. The question is to what extent can one build on students' opinions in this regard?

On the other hand, in light of the importance and sensitivity of teacher assessment and its impact on crucial issues such as academic promotion of teachers, a student's educational performance can be an important contributor to evaluating the teachers (3) and requires further investigation. Taking this into consideration, this study aimed to explore the teaching quality scores of teachers in Birjand University of Medical Sciences and compare them across groups of students with varying categories of overall GPA. The association between the teacher's gender, teacher's academic degree and rank, and the score of teaching quality were also investigated.

Methods

This descriptive cross-sectional study was

done in Birjand University of Medical Sciences during all four semesters of 2015 and 2016 academic years. In this survey, all courses were assessed by census. The evaluation in this study was performed using the data available on Sama Samaneh educational management software, where students are allowed to electively fill out the questionnaires at the end of each semester before the final exams. The questionnaires are the ones used in the Education Development Center of the university for teacher evaluation. The questionnaire consisted of fifteen items in four main domains of teacher's dignity (3 items), educational management (3 items), academic competency (2 items), and teaching method (6 items), as well as an overall assessment of the teacher (1 items), which in total examined the quality of a teacher's teaching. The questionnaires are shown in appendix 1. The questionnaire was scored on a 5-point Likert scale as *always, very often, sometimes, rarely, and never.* Its validity and reliability were verified and documented in the Education Development Center. Briefly, the content validity of this questionnaire was confirmed in five expert panels. Then, the validated form was twice filled out by 42 students for which the Cronbach's alpha coefficient was calculated as 0.78 that was acceptable.

All the data were analyzed in SPSS (version 16). GPA categories in this study comprised of Group A (GPA ≥17 out of 20), Group B (GPA=14-16.99), and Group C (GPA <14). Descriptive statistical tests involved frequency and percentage of returned questionnaires in terms of degree levels of students, school of students, ranks of teachers, academic degrees of teachers, gender of teachers and GPA category of students. Analytical results were reported using independent t-test to compare the evaluation scores with respect to the teacher's gender, and one-way ANOVA and its post-hoc tests to both determine the relationship between teaching quality evaluation of students and students' GPA and to compare between evaluation scores in terms of the academic rank and degree of teachers, after the normal distribution of data was studied.

Results

This study was conducted online in four semesters of 2015-16 and 2016-2017 academic years in all schools of Birjand University of Medical Sciences including medicine, dentistry, nursing and midwifery, allied medicine, and health schools. A total number of 83204 students evaluated 1150 teachers in 4183 courses and finally 1516, 1436, 2702, and 1864 questionnaires were completed respectively from the first to the fourth semesters in this study, with each questionnaire being related to one lesson. The descriptive analysis results of evaluation in the four semesters are described in Tables 1 to 6.

ANOVA test was employed in three GPA categories to determine the relationship between teaching quality evaluation of students and students' GPA, the results of which is presented in Table 7 for each semester. In this test, after testing the homogeneity of variances, the result of Levene's test was significant in all four semesters (P=0.000). Consequently, as the ANOVA test results were

Table 1: Frequency of the returned evaluation questionnaires in terms of degree levels of students

		First semester of 2015	Second semester of 2015	First semester of 2016	Second semester of 2016
		N (%)	N (%)	N (%)	N (%)
Degree levels of the students	Associate	244 (13.1%)	248 (13.4%)	248 (13.4%)	245 (14.5%)
	Bachelor	1098 (59.1%)	1090 (58.7%)	1094 (58.9%)	978 (57.8%)
	Master	116 (6.2%)	116 (6.2%)	116 (6.2%)	59 (3.5%)
	General Practice	399 (21.5%)	403 (21.7%)	339 (21.5%)	411 (24.3%)

Table 2: Frequency of the returned evaluation questionnaires in terms of school of students

		First semester of 2015	Second semester of 2015	First semester of 2016	Second semester of 2016
		N (%)	N (%)	N (%)	N (%)
School	Medicine	513 (27.6%)	525(28.3%)	498 (26.8%)	245(25%)
	Nursing and Midwifery	537 (28.9%)	551 (29.3%)	567 (30.5%)	468 (27.6%)
	Allied Medicine	379 (20.4%)	376 (20.2%)	376 (20.2%)	390 (23%)
	Health	306 (16.5%)	283 (15.2%)	292 (15.7%)	291 (17.2%)
	Dentistry	122 (6.6%)	122 (6.6%)	124 (6.7%)	121 (7.1%)

Table 3: Frequency of the returned evaluation questionnaires in terms of ranks of teachers

		First semester of 2015	Second semester of 2015	First semester of 2016	Second semester of 2016
		N (%)	N (%)	N (%)	N (%)
Ranks	Professor	9 (0.5%)	9 (0.5%)	12 (0.6%)	9 (0.5%)
	Associate professor	140 (7.5%)	140 (6%)	142(7.6%)	138 (8.2%)
	Assistant Professor	559 (30.1%)	593 (31.9%)	583 (31.4%)	502 (29.7%)
	Instructor	1149 (61.9%)	1115 (7.5%)	1120 (60.3%)	1044 (61.7%)

Table 4: Frequency of the returned evaluation questionnaires in terms of academic degrees of teachers

		First semester of 2015	Second semester of 2015	First semester of 2016	Second semester of 2016
		N (%)	N (%)	N (%)	N (%)
Academic degrees of teachers	Bachelor	237 (12.8%)	0	420 (22.6%)	384 (20.6%)
	Master	852 (45.9%)	767 (41.3%)	664 (35.8%)	643 (38%)
	Specialty	336 (18.1%)	656 (35.3%)	320 (17.2%)	293 (17.3%)
	Sub-specialty	29 (1.6%)	29 (1.6%)	29 (1.6%)	24 (1.4%)
	General Practice	55 (3%)	61 (3.3%)	66 (3.6%)	321 (19%)
	PhD	384 (18.7%)	344 (18.5%)	358 (19.3%)	64 (3.8%)

Table 5: Frequency of the returned evaluation questionnaires in terms of gender of teachers

		First semester of 2015	Second semester of 2015	First semester of 2016	Second semester of 2016
		N (%)	N (%)	N (%)	N (%)
Gender of teacher	Woman	782 (42.1%)	779 (41.9%)	779 (41.9%)	704 (41.6%)
	Man	1075 (57.9%)	1078 (58.1%)	1078 (58.1%)	989 (58.4%)

Table 6: Frequency of the returned evaluation questionnaires in terms of GPA category of students

		First semester of 2015	Second semester of 2015	First semester of 2016	Second semester of 2016
		N (%)	N (%)	N (%)	N (%)
GPA category of students	A (≥17)	645 (34.7%)	645 (34.7%)	645 (34.7%)	564 (33.3%)
	B (14-16.99)	647 (34.8%)	647 (34.8%)	647 (34.8%)	584 (34.3%)
	C (<14)	565 (30.4%)	656 (30.4%)	656 (30.4%)	645 (32.2%)

Table 7: Comparison of mean evaluation scores in terms of semesters

Group	Mean	SD	F	P
First semester of 2015				
GPA category "A"	4.57	0.39	7.84	0.000
GPA category "B"	4.55	0.28		
GPA category "C"	4.47	0.57		
Second semester of 2015				
GPA category "A"	4.57	0.39	7.91	0.000
GPA category "B"	4.55	0.28		
GPA category "C"	4.48	0.57		
First semester of 2016				
GPA category "A"	4.57	0.39	7.84	0.000
GPA category "B"	4.55	0.28		
GPA category "C"	4.47	0.57		
Second semester of 2016				
GPA category "A"	4.47	0.36	4.86	0.008
GPA category "B"	4.45	0.32		
GPA category "C"	4.39	0.63		

meaningful, assuming inequality of variances, Scheffé's post-hoc test was used to examine the difference between groups.

The result of Scheffé's post-hoc test following ANOVA showed a significant difference in ANOVA between GPA groups A and C in the first and second semesters of 2015 (P=0.03 and P=0.04, respectively). Also, in the first semester of 2016, there was a significant difference between A and C groups and between the GPA pairs of A and C and B and C (P=0.04 and P=0.03 respectively). These findings indicate that evaluation scores vary from one GPA category to another.

The results of the ANOVA test in regard of the comparison between evaluation scores in terms of the academic rank and degree of teachers is presented in Table 8.

For more study, the results of Scheffé's post hoc analysis table of comparison of mean evaluation scores in terms of academic degree of teachers is attached in Appendix 2.

Table 8: Results of Scheffé's post hoc analysis table of comparison of mean evaluation scores in terms of semesters

Group	Group	Differences of means	P
First semester of 2015			
GPA category "A"	GPA category "B"	0.016	0/79
GPA category "A"	GPA category "C"	0.09	0/001
GPA category "B"	GPA category "C"	0.76	0.008
Second semester of 2015			
GPA category "A"	GPA category "B"	0.01	0/74
GPA category "A"	GPA category "C"	0.09	0/001
GPA category "B"	GPA category "C"	0.74	0.1
First semester of 2016			
GPA category "A"	GPA category "B"	0.01	0/79
GPA category "A"	GPA category "C"	0.09	0/001
GPA category "B"	GPA category "C"	0.07	0.008
Second semester of 2016			
GPA category "A"	GPA category "B"	0.01	0/89
GPA category "A"	GPA category "C"	0.07	0/01
GPA category "B"	GPA category "C"	0.06	0.04

For more information on the results of Scheffé's post-hoc analysis of comparison of mean evaluation scores in term of academic rank of teacher, the full table is shown in Appendix 3. According to Tables 9 and 10, the comparison of teaching quality scores in terms of academic rank and degree in the four academic semesters showed that the association was significant only in the 4th semester (P=0.000 and P=0.002, respectively), whereas for other semesters, it was not significant.

Independent t-test was used to compare evaluation scores with respect to the teacher's gender. The results showed that in none of the semesters, there was any association between the students' assessment scores and the gender of the teacher (P=0.09, P=0.07, P=0.1, and P=0.09 respectively). This means that students' evaluation was not associated with gender.

Discussion

The study was cross-sectional exploring 4183 courses in 4 semesters among almost all students in Birjand University of Medical sciences.

Our finding of this study can be summarized as follows:

1- There was a statistically significant difference between teaching quality evaluation of students and students' GPA such that evaluation scores varied from one GPA category to another. Students with greater GPA scores had higher evaluations for their teachers.

2- Among the four semesters, the comparison of teaching quality scores in terms of academic rank and degree of teachers in the four academic semesters showed that the comparison was significant only in one semester.

Table 9: Comparison of mean evaluation scores in terms of academic degree of teachers

Group	Mean	SD	F	P
First semester of 2015				
Bachelor	4.5	0.48		
Master	4.53	0.46		
Specialty	4.55	0.39	0.58	0.71
Sub-specialty	4.61	0.22		
General Practice	4.52	0.31		
PhD	4.54	0.39		
Second semester of 2015				
Bachelor	-	-		
Master	4.53	0.44		
Specialty	4.54	0.44	0.71	0.57
Sub-specialty	4.61	0.22		
General Practice	4.47	0.34		
PhD	4.55	0.38		
First semester of 2016				
Bachelor	4.52	0.48		
Master	4.53	0.45		
Specialty	4.55	0.39	0.77	0.56
Sub-specialty	4.61	0.22		
General Practice	4.47	0.35		
PhD	4.55	0.38		
Second semester of 2016				
Bachelor	4.39	0.41		
Master	4.40	0.54		
Specialty	4.49	0.35	3.8	0.000
Sub-specialty	4.59	0.25		
General Practice	4.43	0.38		
PhD	4.49	0.39		

Table 10: Comparison of mean evaluation scores in terms of academic rank of teachers

Group	Mean	SD	F	P
First semester of 2015				
Instructor	4.52	0.45	2.17	0/08
Assistant Professor	4.56	0.38		
Associate professor	4.5	0.39		
Professor	4.81	0.12		
Second semester of 2015				
Instructor	4.53	0.45	1.95	0.11
Assistant Professor	4.56	0.38		
Associate professor	4.5	0.39		
Professor	4.81	0.12		
First semester of 2016				
Instructor	4.53	0.46	2.28	0.77
Assistant Professor	4.55	0.38		
Associate professor	4.5	0.39		
Professor	4.81	0.11		
Second semester of 2016				
Instructor	4.40	0.5	6.80	0.000
Assistant Professor	4.5	0.37		
Associate professor	4.48	0.32		
Professor	4.65	0.12		

3- There was no association between the students' assessment scores and the gender of the teacher.

In spite of our finding on the relationship between teaching quality evaluation of students and students' GPA, a few studies indicated no association, whereas some studies agree with our finding. For example, Shakurnia and colleagues mentioned that there was a negative correlation between teacher's evaluation score and student's GPA (9). In contrast, Boring concluded that students' GPA influenced their evaluation of teachers (10). Griffin reported a moderate correlation between them (17) and Vakili in Semnan University of Medical Sciences found similar results (12) that agree with our findings. These contradictory results suggest that, regardless of high or low GPAs, some other factors can be considered as a contributor to the teacher evaluation score by students that could be considered in the teacher evaluation process.

With regard to the relationship between teaching quality scores and academic rank of teachers, it seems that higher ranks of the teacher correlated with greater evaluation scores. However, we found contrary results, except in one semester that could be accidental. According to Khalkheili and co-workers, the rank of the teacher is one of the most effective factors in predicting teacher evaluation (18). Also, Aghamirzayi in Mazandaran University of Sciences and Technology verified this finding (2). Gharatapeh's study in Kermanshah University of Medical Sciences showed similar results (19). She reported a significant difference between instructors and associate professors, but this difference in our study was found only in one semester. Therefore, this factor cannot possibly predict teacher evaluation.

Concerning the association between the students' assessment score and the gender of the teacher, our study showed no relationship, while other researchers considered the teacher's gender as an effective factor in teacher evaluation (2, 10, 11, 20, 21). The results of a number of studies do not correspond with our results including Asassfeh's study in which the teacher's gender was a less influential factor in their evaluation (22). Gender is context bounded as Boring mentioned this in her

study. That is why, we found diverse results in different studies.

Totally, although we did a longitudinal follow-up and found no meaningful difference between different GPA categories in our university, more investigation is needed with more details in other universities and different contexts. This is essential because there is no sufficient agreement concerning teacher evaluation by high or low GPA students. On the other hand, we cannot ignore the low GPA students. If it is assumed that a professor's effective performance has led to greater learning in high-grade students. This means that the professor has succeeded in establishing learning in only a limited number of students and that his/her teaching methods may have been appropriate for a limited number of students. In other words, other students were at the same time deprived of appropriate training in the class and that they would have the right to evaluate the professor. Accordingly, one should not quickly decide to eliminate students with a low grade from the evaluation process, and this requires more quantitative and qualitative studies.

Conclusion

Although some studies have rejected the relationship between learning and the teaching quality scores, the results for the university under study showed differences across different GPA groups. Given the potential cultural or social differences, the results of this study may differ from those of other universities. Therefore, it is suggested that all universities explore this issue by evaluating the teachers of the respective universities.

Conflict of Interest

The author declares no conflict of interest.

References

1. Saif AA. Modern educational psychology: Psycho-logy of learning and teaching. 7th edition. Tehran: Doran; 2013

2. Aghamirzayi T, Salehi Omran E, Rahimpour Kami B. Effective Factors on Student Evaluation of Faculty Members Performance. Educ Strategy Med Sci. 2014; 7(1): 57-62.

3. Shakurnia A, Fakoor M, Elhampour H, Taherzadeh M. Evaluation of validity and reliability of the questionnaire of student evaluation of teaching. JSMJ. 2011; 10(6): 583-93.

4. Wachtel HK. Student evaluation of college teaching effectiveness: A brief review. Assess Eval High Educ. 1998; 23(2): 191-212. doi:10.1080/0260293980230207.

5. Mason PM, Steagall JW, Fabritius MM. Student evaluations of faculty: A new procedure for using aggregate measures of performance. Economics of Education Review. 1995; 14(4): 403-16. doi:10.1016/0272-7757(95)00016-D.

6. Golding C, Adam L. Evaluate to improve: useful approaches to student evaluation. Assess Eval High Educ. 2016; 41(1): 1-14. doi:10.1080/02602938.2014.976810.

7. Berger JS, Pan E, Thomas J. A randomized, controlled crossover study to discern the value of 360-degree versus traditional, faculty-only evaluation for performance improvement of anesthesiology residents. J Educ Perioper Med. 2009; 11(2).

8. Metcalfe DH, Matharu M. Students' perception of good and bad teaching: report of a critical incident study. Med Educ. 1995; 29(3):193-7.

9. Shakurnia A, Malayeri AR, Torabpour M, Elhampour H. Correlation between student evaluation of teaching and students' grades. Iranian journal of medical education. 2006; 6(1): 51-58.

10. Boring A, Ottoboni K, Stark PB. Student evaluations of teaching (mostly) do not measure teaching effectiveness. ScienceOpen Res. 2016; 1. doi:10.14293/S2199-1006.1.

11. Nouhi E, Farajzadeh S, Adhami A. Factors Affecting facutty evaluation by students in

Kerman University of Medical Sciences. JME. 2002; 2(1): 19-23. doi:10.22037/jme. v2i1.890.

12. Vakili A, Hajaghajani S, Rashidy-Pour A, Ghorbani R. An investigation of factors influencing student evaluation of teacher performance: A comprehensive study in Semnan University of Medical Sciences. Faslnamahi Kumish. 2011; 12(2): 93-103.

13. Waller L. An Analysis of Grades, Class Level and Faculty Evaluation Scores in the United Arab Emirates. International Education Studies. 2016; 9(2): 66. doi:10.5539/ies.v9n2p66.

14. Kockelman KM. Student grades and course evaluations in engineering: What makes a difference. Annu Conf Expo (Am Soc Eng Educ). 2001.

15. Koper PT, Felton J, Sanney KJ, Mitchell JB. Real GPA and Real SET: two antidotes to greed, sloth and cowardice in the college classroom. Assess Eval High Educ. 2015; 40(2): 248-64. doi:10.1080/02602938.201 4.904272.

16. Shaw GP. Measuring teacher effectiveness- or not? J Am Podiatr Med Assoc. 2013; 103(1): 94-6.

17. Griffin TJ, Hilton III J, Plummerb K, Barret D. Correlation between grade point averages and student evaluation of teaching scores: taking a closer look. Assess Eval High Educ. 2014; 39(3): 339-48. doi:10.10 80/02602938.2013.831809.

18. Azizi Khalkheili T,Hosseinpor A. Factors Affecting Teaching Quality of Faculty Members in Sari Agricultural Sciences and Natural Resources University from Students' and Graduates' Perspective. Iranian Agricultural Extention and Education Journal. 2016; 12(1): 165-79.

19. Gharatapeh A, Rezaei M, Pasdar Y, Asadi F, Safari Y, Nazari M. Comparison of The Teaching Quality aspects by student Evaluation of Education Quality (SEEQ) and students Survey Quwstionnaires Health School, Kermanshah University of Medical Sciences. Educ Strategy Med Sci . 2015; 8(4): 247-54.

20. MacNell L, Driscoll A, Hunt AN. What's in a name: exposing gender bias in student ratings of teaching. Innov High Educ. 2015; 40(4): 291-303. doi:10.1007/s10755-014-9313-4.

21. Thawabieh AM. Students Evaluation of Faculty. International Education Studies. 2017; 10(2):35-43. doi:10.5539/ies.v10n2p35.

22. Asassfeh S, Al-Ebous H, Khwaileh F, Al-Zoubi Z. Student faculty evaluation (SFE) at Jordanian universities: a student perspective. Educational Studies. 2014; 40(2): 121-43. doi:10.1080/03055698.201 3.833084.

23. Uttl B, White CA, Gonzalez DW. Meta-analysis of faculty's teaching effectiveness: Student evaluation of teaching ratings and student learning are not related. Studies in Educational Evaluation. 2017; 54: 22-42. doi:10.1016/j.stueduc.2016.08.007.

Appendix 1: The detail contents of questionnaire have been used in this study.

No.	Subject	Evaluation Criteria	Scale				
1	Teacher's Dignity	Maintaining decent physical appearance (appropriate clothing, etc.) and observing teacher's dignity	Always	Often	Sometimes	Rarely	Never
2		Teacher's social behavior toward the student and establishing mutual respect	Very Good	Good	Average	Poor	Very Poor
3		Paying attention to moral, cultural and Islamic issues in dealing with students	Always	Often	Sometimes	Rarely	Never
4	Educational Management	Timely attendance and adherence to class time	Always	Often	Sometimes	Rarely	Never
5		Paying attention to students' attendance and absence	Always	Often	Sometimes	Rarely	Never
6		Ability to manage and control the class	Very Good	Good	Average	Poor	Very Poor
7	Academic	Mastery of the course content	Very Good	Good	Average	Poor	Very Poor
8	Ability	Using new and up-to-date scientific resources in teaching and encouraging the students to study such resources	Always	Often	Sometimes	Rarely	Never
9	Teaching Method	Describing the objectives of the course by presenting the course plan	Always	Often	Sometimes	Rarely	Never
10		Presenting the course content in a practical way and using appropriate and diverse teaching methods (along with appropriate examples) for conveying the course concepts and material.	Always	Often	Sometimes	Rarely	Never
11		Use of teaching aid tools based on the available facilities and in accordance with the course type	Always	Often	Sometimes	Rarely	Never
12		Motivating the students to study more and actively participate in class discussions	Always	Often	Sometimes	Rarely	Never
13		Summarizing the presented contents at the end of the class	Always	Often	Sometimes	Rarely	Never
14		Evaluating the students' learning during the semester through designing appropriate questions	Always	Often	Sometimes	Rarely	Never
15		In general, the teacher is regarded as a model for you in different dimensions	Always	Often	Sometimes	Rarely	Never

Appendix 2: The result of Scheffé's post hoc analysis table of comparison of mean evaluation scores in term of degree of teacher.

A. The result of Scheffé's post hoc analysis table of comparison of mean evaluation scores in terms of degree of teacher in the first semester of 2015

Group	Group	Differences of means	P
Bachelor	Master	-0.02	0.97
Bachelor	Specialty	-0.04	0.88
Bachelor	Sub-specialty	-0.10	0.9
Bachelor	General Practice	-0.01	1.00
Bachelor	PhD	0.04	0.94
Master	Specialty	-0.02	0.99
Master	Sub-specialty	-0.08	0.96
Master	General Practice	0.01	1.00
Master	PhD	-0.01	0.99

Specialty	Sub-specialty	-0.05	0.99
Specialty	General Practice	0.03	0.99
Specialty	PhD	0.00	1.00
Sub-specialty	General Practice	0.09	0.97
Sub-specialty	PhD	0.67	0.98
General Practice	PhD	0.02	1.00

B. The result of Scheffé's post hoc analysis table of comparison of mean evaluation scores in terms of degree of teacher in the second semester of 2015

Group	Group	Differences of means	P
Bachelor	Master	-	-
Bachelor	Specialty	-	-
Bachelor	Sub-specialty	-	-
Bachelor	General Practice	-	-
Bachelor	PhD	-	-
Master	Specialty	-0.00	0.99
Master	Sub-specialty	-0.07	0.91
Master	General Practice	0.05	0.89
Master	PhD	-0.02	0.96
Specialty	Sub-specialty	-0.07	0.93
Specialty	General Practice	0.06	0.86
Specialty	PhD	-0.01	0.98
Sub-specialty	General Practice	-0.13	0.72
Sub-specialty	PhD	-0.08	0.76
General Practice	PhD	0.08	0.76

C. The result of Scheffé's post hoc analysis table of comparison of mean evaluation scores in terms of degree of teacher in the first semester of 2016

Group	Group	Differences of means	P
Bachelor	Master	-0.004	1.00
Bachelor	Specialty	-0.03	0.96
Bachelor	Sub-specialty	-0.08	0.94
Bachelor	General Practice	0.46	0.98
Bachelor	PhD	-0.02	0.97
Master	Specialty	-0.02	0.97
Master	Sub-specialty	-0.08	0.95
Master	General Practice	0.05	0.97
Master	PhD	-0.02	0.98
Specialty	Sub-specialty	-0.05	0.99
Specialty	General Practice	0.07	0.87
Specialty	PhD	0.00	1.00

D. The result of Scheffé's post hoc analysis table of comparison of mean evaluation scores in terms of degree of teacher in the second semester of 2016

Group	Group	Differences of means	P
Bachelor	Master	-0.00	1.00
Bachelor	Specialty	-0.9	0.18
Bachelor	Sub-specialty	-0.19	0.51
Bachelor	General Practice	-0.03	0.99
Bachelor	PhD	-0.09	0.18

Master	Specialty	-0.09	0.12
Master	Sub-specialty	-0.19	0.52
Master	General Practice	-0.03	0.99
Master	PhD	-0.09	0.12
Specialty	Sub-specialty	-0.09	0.95
Specialty	General Practice	0.05	0.97

Appendix 3: The result of Scheffé's post hoc analysis table of comparison of mean evaluation scores in terms of academic rank of teachers

A: The result of Scheffé's post hoc analysis table of comparison of mean evaluation scores in terms of academic rank of teacher in the first semester of 2015

Group	Group	Differences of means	P
Instructor	Assistant Professor	-0.31	0.56
Instructor	Associate professor	0.24	0.9
Instructor	Professor	-0.28	0.27
Assistant Professor	Associate professor	-0.05	0.59
Assistant Professor	Professor	-0.25	0.39
Associate professor	Professor	0.3	0.23

B: The result of Scheffé's post hoc analysis table of comparison of mean evaluation scores in terms of academic rank of teacher in the second semester of 2015

Group	Group	Differences of means	P
Instructor	Assistant Professor	-0.02	0.72
Instructor	Associate professor	0.2	0.91
Instructor	Professor	-0.27	0.14
Assistant Professor	Associate professor	0.05	0.63
Assistant Professor	Professor	-0.25	0.37
Associate professor	Professor	-0.3	0.23

C: The result of Scheffé's post hoc analysis table of comparison of mean evaluation scores in terms of academic rank of teacher in the first semester of 2016

Group	Group	Differences of means	P
Instructor	Assistant Professor	-0.02	0.70
Instructor	Associate professor	0.02	0.95
Instructor	Professor	-0.28	0.17
Assistant Professor	Associate professor	0.04	0.70
Assistant Professor	Professor	-0.25	0.25
Associate professor	Professor	-0.30	0.14

D: The result of Scheffé's post hoc analysis table of comparison of mean evaluation scores in terms of academic rank of teacher in the second semester of 2016

Group	Group	Differences of means	P
Instructor	Assistant Professor	-1.00	0.00
Instructor	Associate professor	-0.08	0.22
Instructor	Professor	-0.25	0.42
Assistant Professor	Associate professor	0.014	0.99
Assistant Professor	Professor	-0.15	0.80
Associate professor	Professor	-0.16	-0.76