

A Study of the Correlation of the Improvement of Teaching Evaluation Scores Based on Student Performance Grades

Chi Yuan Chen¹, Shu-Yin Wang² & Yi-Fang Yang³

¹ Center of Teacher Education, Chinese Culture University, Taipei, Taiwan

² Admission affairs office, Chinese Culture University, Taipei, Taiwan

³ Center of Teaching and Learning Resource, Chinese Culture University, Taipei, Taiwan

Correspondence: Chi Yuan Chen, Center of Teacher Education, Chinese Culture University, Taipei, 55, Hwa-Kang Rd., Yang-Ming-Shan, Taiwan 11114, R. O. C. Tel: 886-2-2861-0511

Received: October 23, 2016

Accepted: March 25, 2017

Online Published: April 3, 2017

doi:10.5430/ijhe.v6n2p162

URL: <https://doi.org/10.5430/ijhe.v6n2p162>

Abstract

The purpose of the study is to explore the influence of teaching evaluations on teachers in that they might try to please their students by giving higher grades in order to get higher teaching evaluation scores. To achieve this purpose, the study analyzed the correlations between teaching evaluation scores, student's final grades and course fail rates, and it also examined whether students' final scores and course fail rates are important predictors of teaching evaluation scores. The study used teaching evaluation scores and students' final grades of the courses offered in the fall term of academic year 2014 and the spring term of academic year 2015 in one university in Taiwan as research samples. The results showed that both student's final grades and course fail rates are predictors of teaching evaluation scores. There is a positive correlation between teaching evaluation scores and students' final grades, and a negative correlation between teaching evaluation scores and course fail rates. Based on the findings, the study inferred that the implementation of teaching evaluations may influence teachers to give better grades and lower course requirements to please their students in order to get higher teaching evaluation scores.

Keywords: Student Evaluate Teaching (SET), Teaching evaluations, Student's final grades, Course fail rate, Higher education

1. Introduction

Since quality of teaching has become an important issue in Taiwan's academia, many universities in Taiwan have established teaching evaluation systems to maintain and improve teaching quality. The most prevalent teaching evaluation system is Student Evaluate Teaching (SET), in which students give feedback by questionnaire to evaluate teaching quality at the end of every semester. Many universities use the results of SET as the only reference for judging teachers' quality of teaching and as important evidence for teachers' promotions.

However, since the results of teaching evaluations are so important and influential on teachers' careers, the implementation of SET may only encourage teachers to try some methods, such as lowering course requirements or giving students good grades even though students are not so engaged in learning, in order to please students to get higher teaching evaluation scores instead of spurring students to work hard to improve students' learning outcomes. The ways teachers try to please their students can make students feel that it is easy to learn and that they have learned well, and this good feeling may lead to increased teaching evaluation scores, but this may not benefit teaching effectiveness because students might not have actual good learning outcomes. If the effects on increasing teaching evaluation scores are sustained, the way we evaluate teaching and the way we use the results of teaching evaluations might only encourage teachers to only please their students, but not to improve their teaching effectiveness.

Therefore, the study tried to discover the influences of teaching evaluations on teachers' attitudes on grading students' performance. The purpose of the current study is to explore whether teachers try to please their students to get higher teaching evaluation scores. To achieve this purpose, the study examined the prospect that giving student better scores would benefit teaching evaluation scores, and if setting stricter course requirements might decrease teaching evaluation scores. To verify this prospect, the study analyzed correlations between teaching evaluation scores and students' final grades, correlations between teaching evaluation scores and course fail rates, and the importance of student's final scores and course fail rates as the predictors of teaching evaluation scores. If student's final grades and

course fail rates correlate with teaching evaluation scores, and are important predictors of teaching evaluation scores, then the prospect that teachers try to get higher teaching evaluation scores by giving better grades and lower course requirements would be sustained. Then the influences of teaching evaluations on teachers are not making teachers become better teachers who try hard to improve student's learning outcomes, but instead are making teachers become a more likable teachers who try to please their students but not make students learn more.

2. Literature Review

There are several researchers who have explored the influences of teaching evaluations on teachers' grading attitudes. A lenient grading attitude was the topic that most researchers are interested in. Researchers hypothesized that teachers may give students better grades in order to get higher teaching evaluation scores. To verify this grading leniency hypothesis, researchers used several criteria such as students' grades or students' expected grades to explore the relationship between teaching evaluation scores and the performance grades teachers give to their students.

Students' grades are the prevalent student learning outcome variable that researchers have used to explore whether teachers have lenient grading attitude. Many researchers explored the correlation between teaching evaluation scores and students' grades. For example, Sauer (2012) explored the relationship between teaching evaluation scores and course grades, and found that teaching evaluation scores were a statistically significant predictor of students' course grades. Hofer, Yurkiewicz, and Byrne (2012) analyzed the relationship between teaching evaluation results and students' grades. They found that there was a significant but relatively small correlation between the teaching evaluation scores and average students' grades. Brockx, Spooren, and Mortelmans (2011) examined the influence of course grades on teaching evaluation scores. The results showed positive significant relationships between course grades and teaching evaluation scores. Johnson, Narayanan, and Sawaya (2013) examined the relationship between course characteristics and teaching evaluation scores. The results reported that average course grades were positively correlated with teaching evaluation scores and had statistically significant effects on teaching evaluation scores, but the effect sizes were small in some cases. Stroebe (2016) proved that teaching evaluations scores in basic courses positively correlated with students' grades. However, they are negatively correlated with students' performance grades in advanced courses. The author inferred that the good teaching evaluation scores of basic courses is the result of lenient grading rather than effective teaching.

Some researchers changed the variable of students' grades. For example, Stehle, Spinath, & Kadmon (2012) investigated the correlation of teaching evaluation scores and two different criteria of student learning: a multiple-choice test and a practical examination. The results showed that there was a strong positive correlation between teaching evaluation scores and the practical examination scores but no significant correlation between teaching evaluation scores and multiple-choice test scores. They inferred that strength of correlation of teaching evaluation scores and student performance varies with the criteria used to indicate student learning. Griffin, Hilton, Plummer, and Barret (2014) used grade point averages (GPAs) as a variable and found that the GPAs and teaching evaluation scores correlated moderately, but the correlation has a large variance.

Students' expected grades are another prevalent criterion that researchers have used to explore the influences of teaching evaluations on teachers' grading attitudes. They hypothesize that when students expected they would get better grades, they would give teachers higher teaching evaluation scores. Miles and House (2015) examined the impact of factors that teachers can't control on teaching evaluation scores. The results found that higher student' expected grades may lead to higher teaching evaluation scores. Eiszler (2002) examined the association between the percentage of students who expected to get grades of A or A- and the teaching evaluation scores they gave to their teachers in the courses offered between 1980 and 1999. The results showed that teaching evaluation scores were significantly correlated to student' expected grades and the authors inferred that the existence of the correlation may encourage grade inflation. Matos-Diaz and Ragan (2009) examined the correlation between the average grade expected in a class and teaching evaluation scores. The results show that teaching evaluation scores are significantly and negatively related to the variance of student' expected grades. They inferred that teachers may narrow the grade distribution, thus giving worse performing students higher grades than what they should get in order to get higher teaching evaluation scores. Some researchers didn't explore the correlation between students' grades and teaching evaluation scores but focused on grading leniency. Ewing (2009) found that teaching evaluation scores may be influenced by lenient grading practices and may not accurately reflect student learning. Gump (2007) thought that this lenient grading only existed in the limited samples the researchers used. Therefore, we should be careful of inferring the results of grading leniency.

Based on past research, student's grades, students' expected grades, and the distribution of grades have been used as the criteria to explore teachers' grading attitudes when they are concerned about teaching evaluation scores. However,

little research has used data from universities in Taiwan to examine teachers' grading attitudes. Taiwan began to use Student Evaluate Teaching systems just in the late twenty-first century. This study explores if teachers in Taiwan, where social relationships are considered to be important, would tend to have stronger lenient grading attitudes. Learning from other research, we tried to examine this effect by using individual student's final grades, whole class students' average grades, and course fail rates from one university in Taiwan.

3. Research Methodology

The study used institutional database analysis as the research method. One university's institutional database in Taiwan was used as the data analysis source. The data included teaching evaluation scores, students' final grades, and fail rates of the courses offered in the fall term of academic year 2014 and the spring term of academic year 2015. There were 4,646 courses offered in the fall term of academic year 2014, and used sample size of student grade profile was 204,537 observations. There were 4,501 courses offered in the spring term of academic year 2015, and used sample size of student grade profile was 208,368 observations. Course enrolled students were all asked to complete a student evaluate teaching questionnaire at the end of the semester. This questionnaire contained twenty-seven questions and these questions were divided into eight parts: teaching attitudes, teaching methods, teaching materials, assessment methods, student learning outcomes, students' interest in the course, students' satisfaction with the course, and other opinions. The questionnaire used a seven-points Likert scale to investigate students' opinions regarding the courses they enrolled in, and the results of the questionnaire were viewed as teaching evaluation scores.

Besides teaching evaluation scores, students' final grades, which including the individual student's grades and the whole class students' average grades at the end of the semester, and course fail rates, which was the percentage of the students who failed to pass the requirements of the course and can't get the credit, were also used as the data analysis source.

The study analyzed whether students' final scores and course fail rates are important predictors of teaching evaluation scores, whether there were correlations between teaching evaluation scores and students' final grades, and whether there were correlations between teaching evaluation scores and course fail rates. Furthermore, the study examined the correlations under different curriculum conditions in order to explore whether the correlations among teaching evaluation scores, students' final scores, and course fail rates were significantly different among different curricula.

4. Findings

4.1 The Correlation between Teaching Evaluation Scores and Students' Final Grades

Table 1 shows the analysis results of correlations between teaching evaluation scores and students' final grades. There is significant correlation between an instructor's teaching evaluation score and the final grade of each student or the whole class average grade. The correlation is positive, which means if the final grade of a student or the whole class average grade increases, the teaching evaluation scores also increases. Moreover, the strength of the correlation of the whole class average grade was higher than individual student's grade.

Table 1. Correlation between students' final grades and instructors' teaching evaluation scores

	Teaching evaluation scores linked to each student's grades	Teaching evaluation scores linked to whole class average grades
Year 2014		
Sample size	204537	4646
Sample average	73.139	75.020
Pearson correlation	0.080***	0.237***
Year 2015		
Sample size	208368	4501
Sample average	73.733	76.630
Pearson correlation	0.091***	0.258***

*p < .05, ** p < .01, *** p < .001

To achieve a more in-depth analysis, the study classified students' final grades into high or low average level groups and explored each correlation. Table 2 shows the correlation difference between high average and low average courses. Students' final average grades are found to be significantly correlated with their instructor's teaching evaluation score

in both high and low average groups.

Table 2. Correlation difference between high average and low average courses

	Grade profile sample size	Students' final average grades	Correlation	Statistics
Year 2014				
Low	61442	63.570	0.028	
Medium	112451	74.954	0.017	z=4.35
High	30644	85.667	0.045	P<0.001
Year 2015				
Low	71411	64.808	0.006	
Medium	113975	76.613	0.028	z=-4.61
High	22982	87.184	0.025	p<0.001

We also wanted to explore whether the correlation between teaching evaluation scores and students' final grades were different in different curriculum conditions; therefore, we adopted Fisher's z transformation to examine the differences (the highest and lowest correlation coefficients) in the correlations between teaching evaluation scores and students' final grades among course factors such as college, the selection type of the course (major or elective), the year and level of the course, and the course session (the time that the course was offered such as morning, afternoon, or evening). The results show that the correlation between teaching evaluation scores and students' final grades were significantly different in different colleges (fall term in year 2014, $z=3.35^{***}$; spring term in year 2015, $z=2.57^{**}$), major or elective course (only fall term in year 2014 significant, $z=2.37^{**}$), different course year (only fall term in year 2014 significant, $z=2.80^{**}$), and different course session (fall term in year 2014, $z=2.08^*$; spring term in year 2015, $z=2.21^*$).

Table 3. The statistical analysis results of the correlation in different curriculum conditions via Fisher's z (z)

	College	Major/ Elective	Year level	Course session
Year 2014	3.35***	2.37**	2.80**	2.08*
Year 2015	2.57**	1.32	1.17	2.21*

4.2 The Correlation between Teaching Evaluation Scores and Course Fail Rate

Table 4 shows the results of correlations between teaching evaluation scores and course fail rates. Failure of the course means that the student's final grade failed to pass the requirements of the course (in this case, did not get 60 points out of 100) and the student could not get the credit and needed to retake the course if necessary. The course fail rate means the percentage of the students who failed to pass the course requirements. The study explored the correlation between teaching evaluation and course fail rates. The results show that the correlation between teaching evaluation scores and course fail rate is significant ($r = -0.19$, $p < .001$) in both 2014 and 2015 years' samples. In addition, the correlation coefficient is negative, which indicates that if the final fail rate increases in a course, the instructor's teaching evaluation scores in that course will decrease.

Table 4. Correlation between instructor's teaching evaluation scores and course fail rates

Statistics	Failed student amount	Fail rate	Teaching evaluation scores
Academic year 2014			
Average	4.62 (per class)	9.01%	93.83
Pearson correlation coefficient	-0.20***	-0.19***	4646 courses
Academic year 2015			
Average	4.35 (per class)	7.99%	93.56
Pearson correlation coefficient	-0.25***	-0.19***	4501 courses

* $p < .05$, ** $p < .01$, *** $p < .001$

We also examined whether the correlation between teaching evaluation scores and course fail rates would be different in different curriculum conditions; therefore, we adopted Fisher's z transformation (z) to analyze the differences (the highest and lowest correlation coefficients) in the correlations between teaching evaluation scores and course fail rate among course factors such as college, the selection type of the course, the year level of the course, and the course session. The results show that the correlation between teaching evaluation scores and course fail rate were only significantly different in different colleges (fall term in year 2014, $z=-5.06^{***}$; spring term in year 2015, $z=-3.79^{***}$) and different course years (fall term in year 2014, $z=-4.70^{***}$; spring term in year 2015, $z=-2.75^{**}$).

Table 5. The statistical analysis results of the correlation in different curriculum conditions via Fisher's z (z)

	College	Major/Elective	Year level	Course session
Year 2014	-5.06***	0.01	-4.70***	-1.42
Year 2015	-3.79***	-0.32	-2.75**	-1.35

4.3 Whether Student's Final Grades and Course Fail Rates were Important Predictors of Teaching Evaluation Scores

In order to explore whether students' final grades and course fail rates were important predictors of teaching evaluation scores, the study adopted the regression model to examine if students' final grades and course fail rates were strong predictors to explain variations in teaching evaluation scores. Table 6 and 7 shows the results of regression analysis with students' final grades and course fail rates separately. At first, we put both students' final grades and course fail rates into the model and found that the correlation between students' final grades and course fail rates is too high ($r = -0.813$) to influence the results. Therefore, we reran the regression model to input these two factors separately, and the results show that students' final grades and course fail rates are both predictors of teaching evaluation scores.

The strength of partial correlation indicates the amount of explained variance of a certain predictor of teaching evaluation scores. The larger the partial correlation of a certain factor, the stronger that factor was in predicting teaching evaluation scores. Results show that students' final grades explain 6.02% of teaching evaluation scores variance, and course fail rate explains 3.56% of teaching evaluation scores variance.

Table 6. Teaching evaluation scores regression model with class mean grades

Variable	Regression coefficient	Partial correlation	Explained variance
Year 2014 & 2015			
Class mean score	0.14***	0.25	
Adjust R ²	6.02 %***		Sample size: 9100

* $p < .05$, ** $p < .01$, *** $p < .001$,

Table 7. Teaching evaluation scores regression model with course fail rates

Variable	Regression coefficient	Partial correlation	Explained variance
Year 2014 & 2015			
Fail rate (%)	-0.09***	-0.19	
Adjust R ²	3.56 %***		Sample size: 9100

* $p < .05$, ** $p < .01$, *** $p < .001$,

5. Discussion

The study's purpose is to explore the influence of teaching evaluations on teachers' grading attitude. We examined whether teachers try to give students better performance scores and lower course requirements to please their students in order to get higher teaching evaluation scores. To achieve this purpose, we examined whether students' final grades and course fail rates are important predictors of teaching evaluation scores, analyzed the correlation between teaching evaluation scores and students' final grades, and explored the correlation between teaching evaluation scores and course fail rates. The results show that students' final grades and course fail rate are both statistical significant predictors of teaching evaluation scores, which is similar to the research results of Sauer (2012), but this can only explain 6.02% and 3.56% of the teaching evaluation scores variance.

We also found that teaching evaluation scores are significantly positively correlated with individual student's final grades and whole class average grades, but both of their correlation coefficients are small. The correlation between

teaching evaluation scores and whole class average grades is stronger than the correlation between individual student's grades. The results are similar to the research results of Hofer, Yurkiewicz, and Byrne (2012), Brockx, Spooren, and Mortelmans (2011), and Johnson, Narayanan, and Sawaya (2013). The study also found that variation of examined correlations existed among different curricula conditions, which is similar to the research results of Griffin, Hilton, Plummer, and Barret (2014). We inferred that the correlation between teaching evaluation scores and students' grades occurs when teachers give students higher performance grades, and that their teaching evaluation scores would be higher, but the phenomenon did not necessary happen. The variance for this phenomenon explained by students' grades was not very high, and the correlations were different in different courses conditions.

The results also show that teaching evaluation scores are significantly negatively correlated with course fail rates, but the correlation coefficient was also slight and was different among different curricula. We inferred the results would show that when teachers failed fewer students, their teaching evaluation scores would be higher. However, just like the correlation between students' grades, the variance for this phenomenon explained by fail rates was not very high, and the relationship was different in different courses conditions.

Based on these results, we assumed that students' final grades and course fail rate both can predict teaching evaluation scores. When teachers give student better scores and lower course fail rates, teachers would get higher teaching evaluation scores. However, these correlations are not very obvious, and were different in different curriculum conditions, which means that there are still other factors that can predict teaching evaluation scores more. Teachers may be able to give students' better scores to slightly increase their teaching evaluation scores, but teachers can't necessary get higher teaching evaluation scores just by giving students higher grades or just by lowering their course requirements. There are other more important factors such as teachers' personalities, course conditions, and teaching methods that may influence teaching evaluation scores.

There are limitations to the study. Researchers who have found that there are correlations between teaching evaluation scores and students' grades have given inconsistent explanations to the correlation. Some of them inferred that the existence of the correlation means the evidence of teaching effectiveness, which means students' good performance and grades are the results of effective teaching; therefore, teachers get higher teaching evaluation scores. Others explained the existence of the correlation occurred because of the lenient grading attitude. This current study can't solve this conflict of explanations which exist in the correlations.

6. Conclusion

Based on the research results, we verified that teachers' grading attitudes are influenced by the implementation of teaching evaluations, which is the student evaluate teaching (SET). Teachers tend to give students better performance grades and lower course requirements to get higher teaching evaluation scores. However, the research results also show that although teachers may tend to please their students by giving better grades and lower course requirements, their teaching evaluation scores may not necessary increased just by doing these things. Besides the grades students receive from their teachers, students consider other factors when giving teaching evaluation scores. These factors may include the efforts students give in the process of learning, teachers personalities, teachers' enthusiasm, the time and type of the curricula, or maybe most importantly, whether the teaching methods and teaching materials match the needs of students. Therefore, teachers may higher their teaching evaluation scores by giving students better grades, but they can't get high teaching evaluation scores by only giving better grades. Besides giving high grades, teachers should try to understand what students really care about when learning and what are the needs of students. This is the best way to increase teaching evaluation scores.

Acknowledgements

We thank Center of Information and Communication at our research cooperation university for providing data of students' grade profiles and instructors' course evaluation, and thank Yu-Chuen Huang for his assistance with data preparation, statistical analysis and methodology suggestion.

References

- Brockx B., Spooren P., & Mortelmans D. (2011). Taking the "grading leniency" story to the edge. The influence of student, teacher, and course characteristics on student evaluations of teaching in higher education. *Educational Assessment, Evaluation and Accountability*, 23, 289–306. <https://doi.org/10.1007/s11092-011-9126-2>
- Davidovitch, N. & Soen, D. (2011). Myths and facts about student surveys of teaching the links between students evaluations of faculty and course grades. *Journal of College Teaching & Learning (TLC)*, 6. <https://doi.org/10.19030/tlc.v6i7.1124>

- Eiszler, C. (2002). College students' evaluations of teaching and grade inflation. *Research in Higher Education*, 43, 483-501. <https://doi.org/10.1023/A:1015579817194>
- Ewing, A. (2009). *Essays on measuring instructional quality in higher education using students' evaluations of teachers and students' grades*. University of Washington, ProQuest Dissertations Publishing.
- Griffin, T., Hilton III, J., Plummer, K., & Barret, D. (2013). Correlation between grade point averages and student evaluation of teaching scores: Taking a closer look. *Assessment & Evaluation in Higher Education*, 39, 339-348. <https://doi.org/10.1080/02602938.2013.831809>
- Heckert, T., Latier, A. & Ringwald-Burton, Amy. (2006). Relations among student effort, perceived class difficulty appropriateness, and student evaluations of teaching: Is it possible to "Buy" better evaluations through lenient grading? *College Student Journal*, 40, 588-596.
- Hoefler, P., Yurkiewicz, J., & Byrne, J. (2012). The association between students' evaluation of teaching and grades. *Decision Sciences Journal of Innovative Education*, 10, 447-459. <https://doi.org/10.1111/j.1540-4609.2012.00345.x>
- Johnson, M., Narayanan, A., & Sawaya, W. (2013). Effects of course and instructor characteristics on student evaluation of teaching across a college of engineering. *Journal of Engineering Education*, 102, 289-318. <https://doi.org/10.1002/jee.20013>
- Matos-D áz, H. & Ragan, J. (2009). Do student evaluations of teaching depend on the distribution of expected grade? *Education Economics*, 18, 317-330. <https://doi.org/10.1080/09645290903109444>
- Miles, P. & House, D. (2015). The tail wagging the dog: An overdue examination of student teaching evaluations. *International Journal of Higher Education*, 4. <https://doi.org/10.5430/ijhe.v4n2p116>
- Remedios, R. & Lieberman, D. (2008). I liked your course because you taught me well: The influence of grades, workload, expectations and goals on students' evaluations of teaching. *British Educational Research Journal*, 34, 91-115. <https://doi.org/10.1080/01411920701492043>
- Sauer, T. (2012). Predictors of student course evaluations. *University of Louisville, ProQuest Dissertations Publishing*.
- Sojka, J., Gupta, A., & Deeter-schmelz, D. (2002). Student and faculty perceptions of student evaluations of teaching: a study of similarities and differences. *College Teaching*, 50, 44-49. <https://doi.org/10.1080/87567550209595873>
- Stehle, S., Spinath, B., & Kadmon, M. (2012). Measuring teaching effectiveness: Correspondence between students' evaluations of teaching and different measures of student learning. *Research in Higher Education*, 53, 888-904. <https://doi.org/10.1007/s11162-012-9260-9>
- Stroebe, W. (2016). Why good teaching evaluations may reward bad teaching. *Perspectives on Psychological Science*, 11, 800-816. <https://doi.org/10.1177/1745691616650284>