

December 2022

The impact of grade inflation on teachers' evaluation: A quantitative study conducted in the context of five Lebanese universities

Bassem E. Maamari

Prince Mohammad Bin Fahd University, drbassem.maamari@gmail.com

Hiba S. Naccache

Qatar University, hnaccache@qu.edu.qa

Follow this and additional works at: <https://digitalcommons.usf.edu/jger>

 Part of the [Higher Education Administration Commons](#)

This Refereed Article is brought to you for free and open access by the M3 Center at the University of South Florida Sarasota-Manatee at Digital Commons @ University of South Florida. It has been accepted for inclusion in Journal of Global Education and Research by an authorized editor of Digital Commons @ University of South Florida. For more information, please contact scholarcommons@usf.edu.

Recommended Citation

Maamari, B. E., & Naccache, H. S. (2022). The impact of grade inflation on teachers' evaluation: A quantitative study conducted in the context of five Lebanese universities. *Journal of Global Education and Research*, 6(2), 192-205. <https://www.doi.org/10.5038/2577-509X.6.2.1169>

Corresponding Author

Bassem E. Maamari, Prince Mohammad Bin Fahd University, College of Business Administration, P.O.Box 1664 Al Khubar 31952, Kingdom of Saudi Arabia

Revisions

Submission date: May 1, 2020; 1st Revision: Aug. 13, 2020; 2nd Revision: Feb. 10, 2021; 3rd Revision: Aug. 26, 2021; Acceptance: Sep. 24, 2021

The Impact of Grade Inflation on Teachers' Evaluation: A Quantitative Study Conducted in the Context of Five Lebanese Universities

Bassem E. Maamari¹ and Hiba S. Naccache²

¹College of Business Administration
Prince Mohammad Bin Fahd University, Saudi Arabia
bmaamari@pmu.edu.sa

²Faculty of Education
Qatar University, Qatar
hnaccache@qu.edu.qa

Abstract

Asking students to evaluate teaching faculty by every ending semester in modern education is an established trend. In the higher education circles, it is validated based on a large body of research showing a relationship between these evaluations and students' achievement. The arising problem is whether this relation is positively associated or not, and the presence of a growing debate pertaining to the many factors influencing this correlation. Most of the cited research shows a link between the attitude of students and their achievement. This research studies the effect of students' grade point average (GPA), together with the type of university as public or private, and students' major, on their attitude towards faculty teaching evaluations. The results of the multiple regression show a strong relationship between GPA and students' attitude towards faculty evaluations, suggesting an ethical duality affecting grade inflation.

Keywords: teaching effectiveness, GPA, university, student achievement, Lebanon

Introduction

The tracing of the last forty years of GPA of university students reveals a large increase in average GPAs among them (Nikolakakos et al., 2012; O'Halloran & Gordon, 2014). The reason behind this sharp increase could be associated with improved learning and better standards among both students and teachers, but also to other factors, including grade inflation (Tyler, 1969; Hall, 2012; Wilkes, 2020), ringing alarms in the policymaking and research circles (Gershenson, 2020; Kizito et al., 2016). Revisiting the topic through a quantitative study based on a large sample and thorough statistical analysis, confirms the role of grade inflation as a tool for teacher marketability, at a time where business schools are turning into the business of education, with increasing competition for students' money (Wilkes, 2020).

The following sections provide the theoretical grounds of the study and hypotheses formulation; followed by the methodology used; the results and discussion; to conclude with suggested recommendations.

Literature Review

Several researchers studied the relationship between Student Evaluations of Teachers (SET) and grade inflation. Three categories concerning students' evaluations are identified: valid, biased, and neither (Greenwald, 1997). The valid category is resolved by a set of published articles across a five-year period, according to Greenwald (1997), "prominent reviews published since about 1980 give a clear impression that major questions of the 1970's about ratings validity were effectively answered and largely put to rest by subsequent research" (p. 1184). Moreover, the studies of Marsh (1982; 1984), Abrami (1985), and Abrami et al. (1990), Howson and Buckley (2020), and Shephard (2020) proved the evidence that students' evaluations can be constructed to explore, and mirror required goals and effectiveness in teachers' activities in class.

Teacher Effectiveness

Teaching institutions use the evaluation instrument SET or *Student Evaluations of Teaching* in order to conduct better appraisals of their faculty members, using these as basis to decide on promotions and tenure (Schneider, 2013). According to Vaillancourt (2013), using the SETs creates a problem called *Grading Leniency Hypothesis* in which students tend to give better evaluations for teachers who give them higher grades (Ellis et al., 2003), while giving lower ratings for teachers who give them lower grades, irrespective of the actual teaching effectiveness of the teacher (Schneider, 2013).

The literature is bi-polar, largely split between researchers that believe that SET is a valid and reliable method, and those who attack it is a biased instrument, in which students' self-interest emphasis is placed on their respective end-grade, and not on the real performance of the teacher (Schneider, 2013). Yet many experimental studies approve the Grading Leniency Hypothesis (Vaillancourt, 2013), and as a result, teaching institutions should consider the problem of grade inflation and accordingly adjust teachers' SET scores.

According to Ewing (2012) the SET problem can be classified under the *Principal-Agent-Client* problem in which the principal or the teaching institution, perceives that higher SET scores are indicative of effective and efficient teacher performance and better education quality of its faculty members (Olds & Crumbley, 2003; Tewari & Ilesanmi, 2020). Whereas for the agents or the teachers, the SET determines their rank and accordingly their tenure and promotion, thus it is less costly for them to increase their SET ratings by giving higher grades than by improving their teaching skills. Research found that lecturers and teaching assistants are more inclined to give higher grades than assistant professors and tenured professors (Ewing, 2012). Yet for the clients, the students, their GPA grade is a measure of the quality and the amount of learning they acquired from the course, and it is also a requirement for their degree and employment after graduation, thus they attempt to reward the instructors who give them higher grades and to take revenge on those giving them lower grades through lower SET scores.

According to Braga et al. (2014) effective teachers, those who deliver good teaching quality, require more effort, concentration, and assignments from their students. In return, and in general, students dislike exerting extra effort, and thus tend to dislike demanding teachers and consequently punish the good teachers by giving them bad SET evaluations due to their feeling of entitlement (Redding, 1998). Moreover, Jewell and McPherson's (2012) meta-study on public universities and

1,871 instructors, over a period of twenty years, found that the GPA of the students increased over the years without an associated increase in the learning quality of the students, which indicates grade inflation as a response to the application of the SET.

Grade Inflation

Grades are used by students, undergraduate and graduate schools, and subsequently by employers as an indicator of the students' performance and abilities, which is why students tend to choose courses that earn them higher grades and obviate courses that result in potentially lower grades, in order not to detriment their GPA (Butcher et al., 2014). In addition, students base their future decisions about subsequent courses and whether to switch majors or not, on the grades they receive in the introductory courses (Adelman, 2008; Main & Ost, 2014; Rojstaczer & Healy 2010). Research has found that certain courses have lower grading standards than others (Ellis et al., 2003), particularly engineering, economics, math, and sciences have lower average grades than humanities and social sciences (Arcidiacono et al., 2012).

In an interesting research study by Butcher et al. (2014) at Wellesley College, grade inflation was causing a problem to the reliability of the information produced by grading reports and to the credibility of the college. This is due to the fact that grade inflation might not reflect the real efforts and performance of students, which affects the degree to which employers and higher education institutions might certify the academic performance of the students' graduating from such institutions. As a result, students who possess inflated grades would not often be able to keep their averages in the new institutions they join, as their performance would not match their grade figures (Shephard, 2020). Accordingly, Wellesley College imposed a new policy to cap grade inflation in which average grades should not be more than B+, as a result the number of students in capped courses as well as the number of majors in such departments decreased (Butcher et al., 2014). Moreover, Arcidiacono et al. (2012) state that many academic institutions assess the differentiation performance among transferred students to identify if their new grades would highly be differentiated from the submitted grades to identify any possibility of grade inflation. As a result, Caruth and Caruth, (2013) and King (2015) concluded that academic institutions that witness concurrent cases of grades inflation would lose their credibility, image and respect among the other institutions and the industry. Furthermore, research shows a positive correlation between grade inflation and the number of new students registered in such institutions, as the students show a positive tendency towards receiving higher grades and achieve a higher academic performance. However, this trend is not proven across all majors, as Kinsler and Pavan (2015) state that grade inflation might not be considered as a motivator for some students in certain majors, especially in scientific majors such as medicine, engineering, pharmacy, and economics.

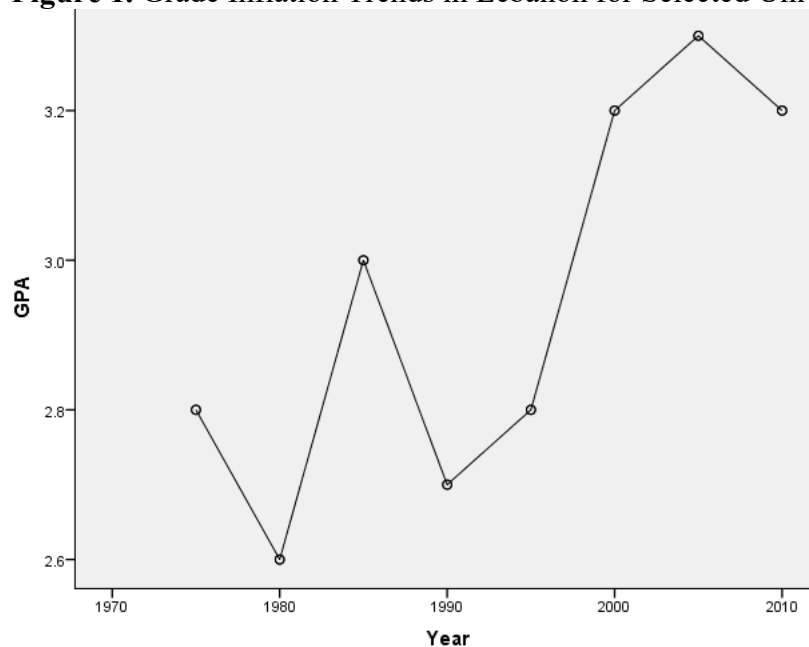
Theoretical Background

In the education literature, researchers report a strong positive correlation between students' grades and their corresponding faculty members. Two theories on grade inflation are discussed in the literature, the Grade-Leniency Theory, and the Attribution Bias Theory, which receives empirical support in Johnson's (2003) exhaustive study of grade inflation. Other theories involve a third variable such as the Teacher Effectiveness Theory (Centra, 1975; Cohen 1981; Costin et al., 1971; Feldman, 1989; Marsh, 1987; McKeachie, 1997; Murray, 1980) which suggests that effective teachers create more successful students (Gershenson, 2020) and reward those teachers with higher

rating grades. Thus, the empirical evidence clearly proves that grade inflation exists due to one of two characteristics, university-level, or instructor-specific factors (Jewell et al., 2013), and that grades influence the students' evaluation of their teachers as well as their choice of teaching institution, as grades correlate with their respective future employability and wages.

The researchers collected longitudinal data on student GPA from the national public university, and several private universities. The participating universities account for more than 50% of the total enrolled university students. The data included an interval of more than 30 years. A clear increase in the GPA of students is obvious in the data (see Figure 1).

Figure 1. Grade Inflation Trends in Lebanon for Selected Universities



Methods

Many universities in Lebanon follow the American higher education model. Given the widespread grade inflation along with the application of SET in several of these universities in Lebanon, the present study investigates the following research question: do students' evaluations of teaching effectiveness affect grade inflation at the university level in Lebanon?

The relationship between student's grade and student's evaluation of teachers was examined through a questionnaire. The questionnaire included three elements: (1) demographic and academic characteristics; (2) response to attitudinal statements; and (3) response to attitude toward teacher's evaluation.

The first part of the questionnaire was on attitude towards student evaluation of classroom teaching, containing 26 items with a five-point Likert scale response where 1 = Strongly agree to 5 = Strongly disagree. The second, on the effect of teaching practices, measured with a list of 16 instructional activities preceded by a statement *All activities below help me increase my understanding of the course content and improve my grade*. For each activity item, one of the five Likert scale were to be chosen with 1 = Strongly agree to 5 = Strongly disagree. Most of the items

included in the two questionnaires were borrowed from Ryan et al. (1980) and from Jacobs (1987). The last part contained a set of questions aimed at capturing the demographic and academic personal characteristics of the respondents. These scale items were reviewed by several faculty members of business and education to ensure clarity and effect on teacher attitudes toward student ratings. A review of the pertinent literature was done on attitude of students toward teacher evaluation and its relationship with grade inflation to make sure that these scale items were consistent with, and valid for answering, the research question of this study.

Sample

The distributed questionnaires targeted a dataset containing a total of 800 students in two faculties, business studies and education. The unit of analysis was collected from five different universities including one public and four private universities, all following the same American educational model. Being National Institute of Health (NIH) certified, the authors took all the necessary ethical steps to allow respondents the choice in participating in the study, and in preserving their right to anonymity and freedom of expression. The questionnaire was distributed during the Fall semester of 2019 to complete classes of randomly selected teachers in the five different universities, with a response rate of 56.25% (441 complete questionnaires).

Analyses were conducted using the SPSS 24.0 program revealing descriptive statistics, chi-square tests, correlations, and multiple regression with Sheffe's multiple tests for comparison, p -value less than .05 for significance. The questionnaire's reliability of the attitude scale was .86 and that of the effect of teaching practices .588, without retaining any item. The results of the demographic details included gender, age, academic characteristics type of institution, student majors and GPA are listed in Table 1. The sample contained 47% males and 53% percent female students. The gender of the teachers under investigation were 52% males and 48% females, including a combination of full-time and part-time teachers employed in public and private institutions. Of the students, 53% were ages 20-22 years, while 52% of the teachers were 30-45 years old, with a mean age of 49.1 ($SD = 9.6$). Students were, for the majority (78.5%), from 4 private universities, almost equally distributed (17-20% each), most of them majoring in business (65.5%) and the remainder were in education (34.2%) (Table 1).

Empirical Model

Two steps of data analysis were applied. First, the mathematical model of the relationships among the variables was laid out and then correlations, multivariate analysis and multiple regressions were used to test the hypotheses at hand.

Attitudes of students were modelled using an interaction between their GPA, major, and academic institution. The *Function* was maximized according to students' attitude that tended to improve course evaluations, which led to an increase in the ease of grading and getting lenient grades.

The framework demonstrated the interactions between students' attitude, GPA, major and academic institutions. Students' attitude was represented by $A^S = A^S(u(g,m,i))$ where the function was assumed to be strictly increasing, concave and additively separable. This was consistent with both theories of grade inflation, namely the Grade Leniency Theory and the Grade Attribution Theory. The Grade Leniency Theory posits that easier grading professors are rewarded with good

student evaluations, whereas the Grade Attribution Theory refers students' good grades to their respective performance and bad grades to poor instruction.

Table 1. Selected Demographic and Academic Characteristics

Characteristic	Sample	Percent
Student Gender		
Male	208	47.1
Female	233	52.7
Teacher Gender		
Male	228	51.6
Female	205	46.4
University		
Lebanese University (Public)	95	21.5
Lebanese American University (Private)	96	21.8
Rafic Hariri University (Private)	103	23.4
American University Beirut (Private)	85	19.3
Sagesse University (Private)	62	14.1
Student Age		
17_19	169	38.2
20_22	235	53.2
23+	35	7.9
Teacher Age		
25_45	229	51.8
46_55	172	38.2
55+	32	7.2
Major		
Accounting	57	12.9
International Business	26	5.9
Economics	16	3.6
Management	83	18.8
Finance	38	8.6
MIS	13	2.9
Marketing	49	11.1
Hospitality	5	1.1
Family Business	3	0.7
Education	151	34.2

Furthermore, the relationship between each variable: the academic institution, GPA, and the major, with the students' attitude, had conditions as follows: $A_i^S > 0$, $A_m^S > 0$, $A_g^S > 0$, where μ , η and φ are the multipliers associated with the Lagrangians L, and where Kuhn–Tucker first order conditions can be written as follows:

$$\begin{aligned} \frac{\partial L}{\partial g} = A_m^S u_g - \mu \leq 0 & \quad g \geq 0 & \quad g \frac{\partial L}{\partial g} = 0 & \quad 1 - g \geq 0 & \quad \mu(1 - g) = 0 \\ \frac{\partial L}{\partial m} = A_i^S u_m - \varphi \leq 0 & \quad m \geq 0 & \quad m \frac{\partial L}{\partial m} = 0 & \quad 1 - m \geq 0 & \quad \varphi(1 - m) = 0 \\ \frac{\partial L}{\partial i} = A_g^S u_i - \eta \leq 0 & \quad i \geq 0 & \quad i \frac{\partial L}{\partial i} = 0 & \quad 1 - i \geq 0 & \quad \eta(1 - i) = 0 \end{aligned}$$

To account for grade target, an extension of the model was used with the professor's choice of grading. The actual policy within the institution was to send a grade distribution report for each course taught by the professor from the Registrar's Office. Therefore, the grade target affected the results, believing the policy to be binding and the professor's choice of grade to be equal to the target. Applying these constraints, maximizing the mathematical model, and applying the Implicit Function Theorem, the researchers obtained results on the implication for students' behavior. The results showed that increasing grade expectations led to a decrease in students' efforts. Finally, after relaxing the grade-target constraint the teaching professors were able to inflate grades and

sustain high evaluations, maximizing research time and minimizing teaching and preparation time. These results were fundamentally different from what was reported in the reviewed literature.

Overall, our findings indicated an unfavorable relationship between students' evaluations and professors grading, which can jeopardize the credibility and benefits of the SET system used by universities currently.

Findings

Attitudes

Overall attitude toward SET for the faculty was computed using the Method of Summated Rating. This method involved reversing the scores of negatively worded scale measures to reflect a consistent positive score of attitudes. Then total scores were computed by summing individual scale measures scores and averaging them (see Table 2).

Table 2. Response to Attitudinal Statements

Statement	<i>M</i>	<i>SD</i>
1. I regard students' evaluations of classroom teaching as a benefit for teacher personally	3.97	0.840
2. I regard students' evaluations of classroom teaching as annoying for teacher personally	2.90	1.102
3. Student ratings help improve classroom teaching	3.85	0.919
4. The process by which student evaluations of classroom teaching are administered does not allow students to give adequate thought to the items	3.35	0.861
5. Items in student ratings of classroom teaching are too general	3.29	0.895
6. Even when faculty can interpret their ratings, they often do not know what to do in order to improve their classroom teaching	3.38	1.046
7. Ratings by non-students (e.g., peers, alumni) would be more effective than student ratings in improving classroom instruction	3.23	1.188
8. I believe that student evaluations lead to positive changes in the quality of teaching at this school	3.79	1.028
9. I believe faculty try to manipulate student judgments of their classroom teaching	3.21	1.040
10. I have confidence in students' ability to judge good classroom teaching.	3.78	2.139
11. I dislike having the quality of my classroom teaching assessed by numerical scores	3.33	1.065
12. Student evaluation of classroom teaching increases the feelings of distance between faculty and administration (e.g., department chairs, deans	3.32	2.245
13. Evaluation of classroom teaching gives students an inflated view of their own power	3.74	1.784
14. Student rating of classroom instruction is reliable for use in personnel decisions such as promotion and tenure.	3.44	1.047
15. Student evaluation of classroom teaching increases the feelings of distance between faculty and students	3.15	1.197
16. Student evaluation of classroom teaching has a negative effect on instructor morale that outweighs positive effects.	2.98	1.232
17. The use of student ratings of classroom teaching should be optional.	3.52	1.253
18. Students are objective when evaluating classroom teaching.	3.55	1.933
19. The way ratings from student evaluations of faculty teaching are analyzed/reported makes it easy to distinguish good teaching from poor teaching.	3.63	1.222
20. Student ratings represent the best procedure for evaluating instructors' classroom teaching.	3.48	1.225
21. We evaluate teachers on semester basis	3.68	1.259
22. Grade inflation is common in Lebanon	3.51	1.308
23. Teachers care about how we evaluate them	3.55	1.401
24. Ethics are important when it comes to teachers' evaluation	3.98	1.358
25. Teachers have enough power to manipulate grades to enhance how students evaluate them	3.36	1.523
26. Our administration takes the teachers' evaluation we do very seriously	3.55	3.167

The mean attitude was 3.34 ($SD = 1.37$). Thus, overall, the students were neither favorable nor unfavorable toward their classroom teaching evaluation in spite of a few outlier respondents. Mean responses ≥ 3.9 indicated that the students agreed with the 3 statements *I regard student evaluations of the classroom teaching as a benefit to teachers, not as an imposition; Student ratings help improve classroom teaching; and Ethics are important when it comes to teacher's evaluation.*

Mean responses < 3 indicated that the students disagreed with the following statement *Student's evaluation of classroom teaching has a negative effect on instructor morale that outweighs positive effects*.

Multivariate Analysis

The researchers used the multivariate statistical analysis to examine the relationships among multiple variables at the same time. In this study, several dependent variables were examined together with several independent variables. The relationships examined in this study check first whether one or more variables such as students' GPA, majors, and type of university, affected the attitude of students towards teacher evaluations and its effect regarding grade inflation. A significance effect appeared when the independent variables are type of university, GPA, and major of study, each variable alone and the three variables together (see Table 3).

Table 3. Response to Attitude Toward Teachers' Evaluation

Attitude toward teachers' evaluation	GPA	Major	Type of University	Model	R ²
I regard students' evaluations of classroom teaching as a benefit for teacher personally	.292	.860	.672	.268	.786
I regard students' evaluations of classroom teaching as annoying for teacher personally	.053	.313	.503	.009	.834
Student ratings help improve classroom teaching	.628	.713	.843	.554	.764
The process by which student evaluations of classroom teaching are administered does not allow students to give adequate thought to the items	.140	.146	.754	.062	.812
Items in student ratings of classroom teaching are too general	.064	.288	.770	.022	.825
Even when faculty can interpret their ratings, they often do not know what to do in order to improve their classroom teaching	.023	.325	.437	.033	.820
Ratings by non-students (e.g., peers, alumni) would be more effective than student ratings in improving classroom instruction	.788	.995	.981	.218	.791
I believe that student evaluations lead to positive changes in the quality of teaching at this school	.246	.426	.906	.211	.792
I believe faculty try to manipulate student judgments of their classroom teaching	.380	.538	.859	.151	.798
I have confidence in students' ability to judge good classroom teaching.	1.000	.964	.973	1.000	.588
I dislike having the quality of my classroom teaching assessed by numerical scores.	.122	.268	.024	.061	.812
Student evaluation of classroom teaching increases the feelings of distance between faculty and administration (e.g., department chairs, deans	1.000	.989	.999	1.000	.533
I regard students' evaluations of classroom teaching as a benefit for teacher personally	.000	.230	.408	.000	.957
I regard students' evaluations of classroom teaching as annoying for teacher personally	.028	.003	.185	.009	.834
Evaluation of classroom teaching gives students an inflated view of their own power	.003	.041	.660	.000	.865
Student rating of classroom instruction is reliable for use in personnel decisions such as promotion and tenure.	.531	.762	.830	.224	.790
Student evaluation of classroom teaching increases the feelings of distance between faculty and students.	.063	.158	.801	.050	.815
Student evaluation of classroom teaching has a negative effect on instructor morale that outweighs positive effects.	.000	.192	.958	.000	.936
The use of student ratings of classroom teaching should be optional.	.395	.353	.319	.483	.769
Students are objective when evaluating classroom teaching.	.022	.758	.859	.034	.820
The way ratings from student evaluations of faculty teaching are analyzed/reported makes it easy to distinguish good teaching from poor teaching.	.270	.345	.657	.195	.793
Student ratings represent the best procedure for evaluating instructors' classroom teaching.	.012	.585	.111	.013	.830
We evaluate teachers on semester basis	.010	.046	.015	.000	.857
Grade inflation is common in Lebanon	.018	.205	.638	.017	.828
Teachers care about how we evaluate them	.026	.665	.346	.013	.830
Ethics are important when it comes to teachers' evaluation	.841	.999	.503	.009	.628
Teachers have enough power to manipulate grades to enhance how students evaluate them	.026	.230	.843	.554	.786
Our administration takes the teachers' evaluation we do very seriously	.841	.003	.346	.993	.834

Moreover, the MANOVA reflected a significant effect of GPA on the attitude of students regarding grade inflation. These results are summarized in Table 4.

The design formula of the testing Students' Attitude is:

$$SET = Intercept + GPA + Type\ of\ University + Major + (GPA \times Type\ of\ University) + (GPA \times Major) + (Type\ of\ University \times Major) + (GPA \times Type\ of\ University \times Major).$$

In the multivariate case analysis, two matrices with no single number representing the ratio of the two matrices were divided. As a result, several multivariate tests were developed, usually based on different aspects of the between-group to within-group matrix ratio. There were four multivariate statistical tests commonly applied: Pillai's Criterion, Hotelling's Trace Criterion, Wilk's Lambda, and Roy's Largest Root. The first three give identical results in a two-group analysis but differed in more complex analyses. They all tested the null hypothesis of no group mean differences in the population. In this research, all the four tests: Pillai's Trace, Wilk's Lambda, Hotelling's Trace and Roy's Largest Root revealed a high significance concerning the model as detailed in Table 4.

Table 4. Pillai's Trace, Wilk's Lambda, Hotelling's Trace and Roy's Largest Root

Multivariate Statistical Test	GPA*Major	GPA * Type of University	Major * Type of University	Model
Pillai's Trace	.000	.064	.468	.000
Wilks' Lambda	.000	.017	.527	.000
Hotelling's Trace	.000	.002	.588	.000
Roy's Largest Root	.000	.000	.064	.000

A high significance appeared when GPA was used in combination with each variable alone, and the interaction of the three variables together highly affected the attitude of students. Although there are situations where one multivariate test has an advantage over another, the four standard tests gave similar results for this model.

In addition to MANOVA, a multiple regression is run to check the effect of GPA, type of university, and major, on the attitude of students toward SET. This effect will lead to prove whether students' attitude toward the SET does lead to a grade inflation where the teachers tend to increase the grades to make sure students will give positive feedback. The difference between MANOVA and multiple regression is the coding of the items of the attitude of students. The researchers coded items that support the high attitude toward the power of the student survey on administration which led to a change in the teacher methods, as positive items. Negative items are those that decrease the effect of SET on teacher's methods. The researchers summed the items after reverse coding and tested using multiple regression (see Table 5). The results are $R = .214$; $R^2 = .058$; $Adj. R^2 = .052$; $F\ Change = 8.939$, $df_1 = 3$; $df_2 = 435$, and $p < .000$

Table 5. Multiple Regression Coefficients

Model	Unstandardized Coefficients		Standardized Coefficients	t	p
	B	Std. Error	B		
Constant	84.246	3.865		21.796	.000
GPA	.366	1.181	.015	.310	.057
Type of University	2.528	.621	.205	4.069	.000
Major	-.212	.154	-.068	-1.375	.070

Note. Dependent Variable: Sum Attitude

Moreover, when checking for the effect of GPA alone, a high significance (see Table 6) was obtained, which led to suggest the existence of the effect of GPA on students' attitude toward SET, as it provides a proof of its effect on grade inflation.

Table 6. GPA Regression Coefficients

Model	Unstandardized Coefficients		Standardized Coefficients	t	p
	B	Std. Error	Beta		
Constant	85.399	3.584		23.826	.000
GPA	1.521	1.188	.061	1.281	.001

Note. Dependent Variable: Sum Attitude

Conclusions

The reviewed literature features much analysis of teachers' evaluation. In this study, and due to the importance of the potential relationship's effect on student grades, an attempt was made to check if the attitude of students towards teachers' effectiveness was affected by their GPA, type of university (public or private), and students' major.

Although students reported that teacher evaluation is annoying for teachers personally, they exhibited an overall neutral attitude toward how teachers use these evaluations for their own benefit, as much as for position and tenure which is in line with Schneider (2013) while contradicting other findings. This highlighted the role of students in attempting, or not, to manipulate their teachers' behavior. Evaluations are typically done by students after the end of the semester. The attitude of students toward teacher evaluation and its relationship with grade inflation is supported by the high correlation of the students' GPA with their attitude toward these evaluations, especially the high significance with the item regarding their point of view toward grade inflation, which is in line with Olds and Crumbley (2003), Vaillancourt's (2013), Schneider (2013), and Tewari and Ilesanmi (2020). The grade inflation hypothesis was also supported by the multivariate effect of the items showing that students have an inflated view of their own power and attempt to retaliate at teachers who push them to exert higher effort levels or who do not embellish graded (Redding, 1998; Braga et al., 2014).

Theoretical Implications

The model in this research puts forth a base for the grade inflation problem and shows how students' attitude affected by their respective institutions and their GPA, does lead to grade inflation, which raises an alarm for administrators who need to seek a mending solution for the problem. As these institutions compete for students' funds (Wilkes, 2020), they tend to lure students through other factors than quality of learning as per Tyler (1969), Hall (2012), Wilkes (2020), thus resulting in higher GPAs that are not associated with higher learning quality or levels (Jewell & McPherson, 2012). These higher grades negatively affect the brand and long-term image of the institution which will be labelled as an *easy grade* by the hiring community, thus foregoing the quality of education. Moreover, the model highlights some findings that are consistent with the extant literature and motivate researchers for further empirical studies. Finally, the attitude of students can affect their overall work and how they feel about teachers' evaluations, but the question to be answered remains, how this attitude can be improved, by the role of institutions or by the teachers themselves, and whether policy-makers can interfere (Kizito et al., 2016; Gershenson, 2020) at the height of this budgetary competition, where they are in charge of providing sound budgets for sustaining and developing their respective institutions. Thus, these policymakers seem to be balancing the losses between grade inflation with increased monetary income (Kinsler & Pavan, 2015) and its effect on the earning experience, learning quality and learning levels on one side, and institutional reputation on the other (Caruth & Caruth, 2013; Butcher et al., 2014; King, 2015). These market-to-budget strategies will have long-term effects

on the reputation of these institutions. These effects will be under the spotlights as their graduates reach the workplace and find themselves ill-equipped to perform/handle the job requirements, perceive recognition or promotion (Shephard, 2020). Moreover, these effects are exaggerated by the intensified integration and use of technology in the process (Pow & Lai, 2021).

Practical Implications

The effect of change is now unambiguously positive; that is, with grade targets in place, the institution can unambiguously increase the professor's teaching effort by increasing its emphasis on course evaluations. Therefore, we suggest the following to educational administrators. First, students are indifferent to, and do not mind the process of SET; therefore, university education administrators should not be afraid to formulate questionnaires or measurement tools to assess classroom teaching effectiveness, as suggested by Marsh (1984), Abrami (1985), Abrami et al. (1990), Howson and Buckley (2020) and Shephard (2020).

Second, classroom teaching effectiveness is positive from the students' perspective and helps improve classroom teaching (Costin et al., 1971; McKeachie, 1997; Gershenson, 2020) allowing faculty to highlight positive student behavior, reinforce selected attitudes, or to vent their concerns and dislikes. Therefore, the tools used should be inclusive of different measurement methods that assess the same variable, such as reverse questions.

Third, students believe that ethics are important when assessing teachers. This warrants a deeper look into the ethical commitment of the students filling the SET. Moreover, this should alarm educational administrators that some students may also be using the SET process as a pressure-tool on the respective teachers to indirectly blackmail them for grades. Thus, there might loom a need to protect the teachers too.

Fourth, in spite of the above, students do believe that SET does affect teacher morale. Their response that SET has a larger negative effect than the positive effect sought, shows that students feel/believe that they have a powerful tool that can be used to manipulate the classroom environment and eventually their grades, making the SET a biased tool (Greenwald, 1997) that needs adaptation.

Fifth, the strong relationship between SET, university type, and grade inflation on one side, and with GPA on the other side, shows that across different university types (public/private), the students' GPA is an engine or instigator of action on the students' side, to behave, act, or reflect certain attitudes in course/major selection (Adelman, 2008; Rojstaczer & Healy, 2010; Butcher et al., 2014; Main & Ost, 2014), as well as during the SET.

Sixth, effective teachers and teachers who are tenured and doing their best to teach efficiently and provide a positive and rewarding learning experience for the students should be involved in the process of SET formulation. These teachers should be invited to reflect on their personal professional practice in adult learning (Tenuto, 2021), and to take part in the preparation phases of the SET administration, such as for example suggesting new questions or reformulating previously used ones. With novel technologies, it is also easy to use qualifying questions that move respondents to answer different sets of questions based on their response to that qualifying question or use simulations or even avatars (Bosch & Ellis, 2021).

Seventh, as the SET is affected by many instructor characteristics (Jewell et al., 2013), the SET should include a section on personality, emotional intelligence, teaching effectiveness, etc., tools that more distinctly differentiate one teacher from another. Thus, it is recommended that administrators revisit their recruitment and retention strategies (Landerlinger et al., 2021).

Finally, with the increasing dependency on interactive, activity-based and case-study teaching methods, different versions of the SET need to be provided for different courses. Courses where interaction is minimal (such as in mathematics, economics, accounting, and finance) require different assessment tools than those requiring higher levels of interaction and innovation in teaching techniques, such as management, marketing, and entrepreneurship. In the latter the teaching process relies heavily on class scenario-discussions, case analyses, opinion building, leadership, and teamwork, in which the teacher is more of a dynamic coach rather than a feeder. Thus, requiring a different type of assessment, and put forth a new set of challenges for education administrators (Seery et al., 2021).

Limitations and Future Research

The study limitations were, first, the number of cooperating private universities was not large, although the four participating ones represented more than 40% of the student population, the public university represented 100% of the public-sector student body. Second, the cross-sectional nature of the study allowed for a number of biases. Finally, the scope of the study was limited to a pre-set number of variables, whereas many others may also have had a role in this interplay. The analysis of student attitude data showed how difficult it is to accurately evaluate teaching effectiveness where many factors play a role in this attitude that hence leads to grade inflation. From one side, teaching effectiveness involves multiple factors in itself, and from another it seems to be affected by students' behavioral and perceived learning outcomes. One should conclude from this study that teaching evaluation data should be used cautiously to separate ineffective teachers from effective teachers. One can only be sure that with a high GPA, the student is not unfairly categorizing good teachers as ineffective.

References

- Abrami, P. C. (1985). Dimensions of effective college instruction. *The Review of Higher Education*, 8(3), 211-228.
- Abrami, P. C., d'Apollonia, S., & Cohen, P. A. (1990). Validity of student ratings of instruction: What we know and what we do not. *Journal of Educational Psychology*, 82(2), 219-231.
- Adelman, C. (2008). Undergraduate grades: A more complex story. In L. H. Hunt (Ed.), *Grade inflation: Academic standards in higher education* (pp. 13–44). State University of New York.
- Arcidiacono, P., Aucejo, E. M., & Spenner, K. (2012). What happens after enrollment? An analysis of the time path of racial differences in GPA and major choice. *IZA Journal of Labor Economics*, 1(1), 5-29.
- Bosch, C., & Ellis, T. (2021). Using avatars to address teacher self-efficacy. *Journal of Global Education and Research*, 5(1), 15-35.
- Braga, M., Paccagnella, M., & Pellizzari, M. (2014). Evaluating students' evaluations of professors. *Economics of Education Review*, 41, 71-88.
- Butcher, K. F., McEwan, P. J., & Weerapana, A. (2014). The effects of an anti-grade-inflation policy at Wellesley College. *The Journal of Economic Perspectives*, 28(3), 189-204.
- Caruth, D. L., & Caruth, G. D. (2013). Grade inflation: An issue for higher education? *Turkish Online Journal of Distance Education*, 14(1), 102–110.
- Centra, J. A. (1975). Colleagues as raters of classroom instruction. *Journal of Higher Education*, 46(3), 327–337.
- Cohen, P. A. (1981). Student ratings of instruction and student achievement: A meta-analysis of multisection validity studies. *Review of Educational Research*, 51(1), 281–309.

- Costin, F., Greenough, W. T., & Menges, R. J. (1971). Student ratings of college teaching: Reliability, validity, and usefulness. *Review of Educational Research, 41*(5), 511-535.
- Ellis, L., Burke, D. M., Lomire, P., & McCormack, D. R. (2003). Student grades and average ratings of instructional quality: The need for adjustment. *Journal of Educational Research, 97*(1), 35-40.
- Ewing, A. (2012). Estimating the impact of relative expected grade on student evaluations of teachers. *Economics of Education Review, 31*(1), 141-154.
- Feldman, K. A. (1989). The association between student ratings of specific instructional dimensions and student achievement: Redefining and extending the synthesis of data from multisection validity studies. *Research in Higher Education, 30*(1), 583-645.
- Gershenson, S. (2020). End the “easy A”: Tougher grading standards set more students up for success. *Education Next, 20*(2), 18-25.
- Greenwald, A. G. (1997). Validity concerns and usefulness of student ratings of instruction. *American Psychologist, 52*(11), 1182-1186.
- Hall, R. A. (2012). A neglected reply to grade inflation in higher education. *Global Education Journal, 2012*(3), 144-165.
- Howson, C. K., & Buckley, A. (2020). Quantifying learning: Measuring student outcomes in higher education in England. *Politics and Governance, 8*(2), 6-14.
- Jacobs, L. C. (1987). *University faculty and students' opinions of student ratings* (ED291291). ERIC. <https://eric.ed.gov/?id=ED291291>
- Jewell, R., & McPherson, M. (2012). Instructor-specific grade inflation: Incentives, gender, and ethnicity. *Social Science Quarterly, 93*(1), 95-109.
- Jewell, R. T., McPherson, M. A., & Tieslau, M. A. (2013). Whose fault is it? Assigning blame for grade inflation in higher education. *Applied Economics, 45*, 1185-1200.
- Johnson, D. W. (2003). Social interdependence: Interrelationships among theory, research, and practice. *American Psychologist, 58*(11), 934-945.
- King, B. (2015). Changing college majors: Does it happen more in stem and do grades matter? *Journal of College Science Teaching, 44*(3), 44-51.
- Kinsler, J., & Pavan, R. (2015). The specificity of general human capital: Evidence from college major choice. *Journal of Labor Economics, 33*(4), 933-972.
- Kizito, R., Munyakazi, J., & Basuayi, C. (2016). Factors affecting student success in a first-year mathematics course: A South African experience. *International Journal of Mathematical Education in Science and Technology, 47*(1), 100-119.
- Landertinger, L., Tesaro, D., & Restoule, J. (2021). We have to get more teachers to help kids: Recruitment and retention strategies for teacher education programs to increase the number of indigenous teachers in Canada and abroad. *Journal of Global Education and Research, 5*(1), 36-53.
- Main, J. B., & Ost, B. (2014). The impact of letter grades on student effort, course selection, and major choice: A regression-discontinuity analysis. *The Journal of Economic Education, 45*(1), 1-10.
- Marsh, H. W. (1982). SEEQ: A reliable, valid, and useful instrument for collecting students' evaluations of university teaching. *British Journal of Educational Psychology, 52*(1), 77-95.
- Marsh, H. W. (1984). Students' evaluations of university teaching: Dimensionality, reliability, validity, potential biases, and utility. *Journal of Educational Psychology, 76*(5), 707-754.
- Marsh, H. W. (1987). Students' evaluations of university teaching: Research findings, methodological issues, and directions for future research. *International Journal of Educational Research, 11*(1), 253-388.
- McKeachie, W. J. (1997). Student ratings: The validity of use. *American Psychologist, 52*(11), 1218-1225.
- Murray, H. G. (1980). *Evaluating university teaching: A review of research*. Ontario Confederation of University Faculty Associations.
- Nikolakakos, E., Reeves, J. L., & Shuch, S. (2012). An examination of the causes of grade inflation in a teacher education program and implications for practice. *College and University, 87*(3), 2-14.
- O'Halloran, K. C., & Gordon, M. E. (2014). A synergistic approach to turning the tide of grade inflation. *Higher Education, 68*(6), 1005-1023.
- Olds, P. R., & Crumbley, D. L. (2003). Higher grades = Higher evaluations: Impression management of students. *Quality Assurance in Education: An International Perspective, 11*(3), 172-177.
- Pow, W., & Lai, K. (2021). Enhancing the quality of student teachers' reflective teaching practice through building a virtual learning community. *Journal of Global Education and Research, 5*(1), 54-71.
- Redding, R. E. (1998). Students' evaluations of teaching fuel grade inflation. *American Psychologist, 53*(11), 1227-1228.

- Rojstaczer, S., & Healy, C. (2010). Grading in American colleges and universities. *Teachers College Record*, 4, 1-6.
- Ryan, J., Anderson, J. A., & Birschler, A. B. (1980). Student evaluation: The faculty responds. *Research in Higher Education*, 12, 317-333.
- Seery, K., Barreda, A. A., Hein, S. G., & Hiller, J. L. (2021). Retention strategies for online students: A systematic literature review. *Journal of Global Education and Research*, 5(1), 72-84.
- Schneider, G. (2013). Student evaluations, grade inflation and pluralistic teaching: Moving from customer satisfaction to student learning and critical thinking. *Forum for Social Economics*, 42(1), 122-135.
- Shephard, K. (2020). On assessment and evaluation, and researching the practices of higher education. In J. C. Lee, R. Maclean, & P. B. Corcoran (Eds.), *Higher education for sustainability, education for sustainability* (pp. 143-155). Springer.
- Tenuto, P. L. (2021). Teaching in a global society: Considerations for university-based educational leadership. *Journal of Global Education and Research*, 5(1), 96-110.
- Tewari, D. D., & Ilesanmi, K. D. (2020). Teaching and learning interaction in South Africa's higher education: Some weak links. *Cogent Social Sciences*, 6(1), 1-16.
- Tyler, R. (1969). *Basic principles of curriculum and instruction*. University of Chicago.
- Vaillancourt, T. (2013). Students aggress against professors in reaction to receiving poor grades: An effect moderated by student narcissism and self-esteem. *Aggressive Behavior*, 39(1), 71-84.
- Wilkes, K. (2020). Questions of value for higher education: The case of luxury student accommodation. In S. Dawes, & M. Lenormand (Eds), *Neoliberalism in Context* (pp. 255-270). Palgrave Macmillan.

Acknowledgement

The authors would like to thank all their colleagues for their support in the questionnaire preparation, during the data collection process, and in reviewing the manuscript for them. The authors declare that they did not receive any funding from any private, government or not-for-profit entity for completing this work.