

SEVERAL BIASES IN EVALUATION PROCESS OF PROFESSORS BY UNDERGRADUATE STUDENTS

Patricia Nasser Carvalho¹, Marcus V. H. Carvalho²

¹Universidade Federal de Minas Gerais (Department of Economics), Universidade Federal de Minas Gerais, Brazil.

²Department of Surgery, Faculdade de Medicina de Jundiaí, Sao Paulo, Brazil

ORCID

Patricia Nasser Carvalho: 0000-0002-8152-9779

Marcus V. H. Carvalho: 0000-0003-3911-370X

Correspondence:

Patrícia Nasser Carvalho, Universidade Federal de Minas Gerais, Faculdade de Ciências Econômicas,
Av. Pres. Antônio Carlos, 6627 - Pampulha, Belo Horizonte - MG, 31270-901

"If you prick us do we not bleed? If you tickle us do we not laugh? If you poison us do we not die? And if you wrong us shall we not revenge?" Shakespeare in Merchant of Venice.

Abstract

The evaluation of the teaching process uses questionnaires applied to the students, giving them the power of be anonymous in their responses. However, if at the end of the course some students get low grades, at the moment of their responses to the questionnaires they can have an attitude of revenge against the teacher, manifested by bad evaluation to the teachers. There are other situations that interfere with students responses leading to biases in evaluation process. We reviewed factors that interfere in the teacher evaluation through questionnaires causing biases and, we also tried to find alternatives to diminish biases in evaluation. We search original articles and reviews in web sites databases about biases in professor evaluation process and comments about avoid these biases. We also conducted a manual search for relevant articles in the reference lists of articles. Several factors can be biases in teaching evaluation process as feeling of revenge, more difficult course, younger students and gender male students evaluators. Non-white teachers, female teachers, older teachers, non-charism teachers, non-good look teachers also get lower grades. Teacher evaluations through questionnaires must be viewed with caution due to several biases. Since questionnaires are difficult to be replaced in the evaluation of teachers, we discuss alternatives in order to avoiding injustice to the teachers.

Keywords: Teaching evaluations bias, student evaluations of teaching, teacher effectiveness, education quality.

Introduction

Evaluation of teacher performance is essential for educational institutions, students, and for professors themselves as a way to improve the teaching process. The most common tool to perform these evaluations are end-of-course questionnaires prepared specifically for this purpose. Questionnaires are used worldwide and are justified due to their easy application and mode of response, being feasible and economical in terms of time and effort to collect a range of views from the whole population to be studied (Wall, 2010). To achieve teacher evaluation, well-elaborated questionnaires, that avoid being merely generic, were and have been offered consistently to students allowing them to give their opinion about the teachers. These questionnaires allow students to present their personal opinion about professor's classes while remaining anonymous. These evaluations are generally carried out at conclusion of all classes of a specific course and after the final grades for each subject. At that moment students are invited to evaluate separately each teacher.

However, assuming that students' responses to these questionnaires are an accurate and unquestionable path to teacher evaluation could lead to a misjudgments and unfair treatment of professors. Under certain conditions, undergraduate students may consider offensive to them to receive a low grade in a test of a course. In that situation, one of the most instinctive human feelings, revenge, can appear and manifest itself in the end-of-course teacher evaluation. Thus, the debate about validation of teacher evaluations through students' responses to questionnaires continues to be debated.

Objectives

This study aims to identify the factors that intervene in students answers to teacher evaluations and point out the criteria used by students when responding to teacher evaluations questionnaires, that means, identified the factors that lead to biases in interpretation the teacher evaluation process. Also, we tried to find alternatives to diminish effect of biases in the evaluation process.

Methods

The terms "teaching evaluations biases AND questionnaires" was used in searching original articles and reviews in web sites databases. All articles selected in the search were in English. Abstracts for oral presentations and letters to the editors were ignored. We also conducted a manual search for further relevant articles in the reference lists of articles.

The search for papers about teacher evaluating was done concisely in the last 20 years. First, titles and abstracts were read to know whether they fit the purpose of reviewing the issue. If their eligibility remained unclear, the full-text reports were then read to decide.

Results

Fourteen factors that can cause biases in teacher evaluation were found: students revenge against their low grades in final exams, course difficulty, students age, students gender, teacher race, teachers gender, teacher character traits, teachers age, teacher physical appearance, class size, response rate to questionnaires. The results are discussed as follow. Also we discuss few alternatives to avoid unjust teacher evaluation by students, considering that bad evaluation mostly is linked to low grades received by students in their own evaluation in subject tests.

Discussion

Revenge

Age of 18 to 25 is the period characterized by self-focused and emotional instability. These characteristics, normally encountered in young adults, make undergraduate students stringent critics and uninhibited evaluators. Thus, if a student for some reason has not obtained the desired grade in a course, he or she can use the teacher evaluation to punish the professor under the cover of the anonymity of their responses to the questionnaires (Stroebe, 2016). These vengeful attitudes are the result of resistance to criticism, a feeling of inferiority compared to their peers, a feeling of time wasted and money wasted as he need to spend more money to repeat the discipline, etc. For these young adults, if the grades given to them were considered unfair, this is the teacher's failure, not their responsibility. However, it should be considered that students' performance equally depends on their own application, ability, and motivation, not only the teaching (Morrison, 2003). Therefore, these questionnaires reflect the level of satisfaction of the students who completed the course with the service/teaching rather than being an evaluation of the teacher as such (Beechman, 2009).

Also, there is low response rate of students to teachers evaluation questionnaires (Dommeyer et al., 2004; Goss & Salomons, 2017), and certainly those who are most interested in respond to questionnaires are those with the revenge feeling. This fact intervenes, causing poor teacher evaluation by the fact that the proportion of responders are higher between those with the feeling of revenge.

Students' Expectations of Grades

Furthermore, psychology research has suggested that the way in which people react to events is often strongly influenced by their expectations and not well conceived expectations of the students to a specific course contribute to their response to the questionnaires (Pinquart & Pietzsch, 2022). People generally, in particular young people, tend to overestimate their abilities. Thus, overconfidence may lead students to go to their courses evaluations without be well-prepared. Authors analyzed the difference between students expect and actual grade and how teacher pedagogies can influence overconfidence in students of an economics curriculum. Their findings were that students of lower divisions classes have a greater tendency to be overconfident than those in upper divisions classes and, it increases the importance of tests to reduce overconfidence (Novell & Alston, 2007), meaning that results of tests improve students to have a more rational behavior. Also, authors in a study with undergraduate students in a statistic course, comparing their grades with their forecast, the author's findings were students expectations are not rational and that most students are overconfident (Magnus & Peresetsky, 2018). However, overconfidence decreases during the

course and is smallest at the third exam, showing that students adjust their expectations accrues (Burst, 2007).

Questionnaires and Anonymity

The criteria that students use to answer to questionnaires are often different from those of university focus. Students may simply care about their grades, whereas in the most cases, faculty cares about student learning (Braga et al., 2014). These two different approaches can compromise the validity of questionnaires.

It is also difficult to analyze several of the most frequent answers to these questionnaires, for instance:

In a circumstance where 50% of students have answered that a teacher behaved respectfully towards students and the other 50% responded that the teacher did not demonstrate a respectful attitude towards the students, these answers would be impossible of being coherently interpreted.

Experience has shown that to avoid a bad teacher' evaluation in these questionnaires, the teacher should give good grades for all or at least for most of the students, which would be a lenient attitude on the part of the teacher (Stroebe, 2020; Griffin, 2004). Such approach, however, unfeasible given that following graduation, students will in turn become professionals with deficiencies and possibly cause harm to the society.

Certainly, in some circumstances, these questionnaires can cause sorrow to the teachers since students had the power to unjustly evaluate their professors. This is not a desirable outcome either for the teachers or for the evaluation process in general, so, caution is needed when interpreting questionnaires responses.

Thus, students' satisfaction is a complex phenomenon influenced by a number of variables (Hornstein, 2017). Furthermore, authors found some studies showing biases in questionnaires evaluation against female instructors, discipline types, student genders and grade expectations. For example, gender biases can be enough to cause more effective instructors to get lower grade than less effective instructors (Boring, 2021). Some alternatives have been proposed to avoid misinterpretation students answers of those questionnaires as a mid-term evaluation that consist of a conventional student questionnaire, but frequently augmented with interview, debriefing sessions and follow-up questionnaires (Senior, 2000). That suggestion comes from the fact that there is evidence that students are more satisfied with mid-term evaluations (Warner et al. 2015), once mid-term evaluations are substantial improvements over end-of-term questionnaires.

A key to good teacher evaluation is to collect data from multiple sources, such as teacher's peers, students, instructors, administrators, etc (Wall,2010). However, this would be much more laborious and time-consuming. Because questionnaires are cheap instruments, easy to be applied and a simple tool to get information, they are difficult to be replaced in students' evaluation of teaching. To make the advance of the questionnaires feasible and to avoid their defects, we consider three approaches:

1. apply the questionnaires before the final exams or
2. the opinions of the particular good students could be given more weight in the evaluation of teachers' performance (Braga, 2014) or
3. if the decision is to apply the questionnaires after the final exams, one drastic solution would be applying the teacher evaluation to those students who received below-average grades in interview setting between the students and other teachers who were not directly involved with that subject or course (current paper authors' suggestion).

4. apply one first test to students without to wait too long because the result of that test will count to student adjust better their expectations at an early stage and this will be of use to them in their effort for the course (Magnus & Peresetsky, 2018).

Student evaluation questionnaires must be used with caution also because if the rate of questionnaires answers by the students are low, the extrapolation of the results for the whole class is unreliable (Hornstein, 2017).

Table 1 provides multiple factors that influence the students` responses to questionnaires quoting the references of these studies/papers. Feeling of revenge, low grade in final exams, more difficult, course, younger students, gender male students give lower grades to teacher in their evaluations. Non-white teachers, female teachers, older teachers, non-charism teachers, non-good look teachers also get lower grades. When the class is large and when the lower number of students address teacher evaluations, the teachers receive lower grades too.

The teachers could avoid a poor evaluation if they leniently reward students with better grades. Additionally, the teacher could be influenced to offer an easier course. However, obviously, none of these attitudes are positive to the teaching.

Conclusion

In conclusion, an attitude of revenge when students undertake teacher evaluation brings injustice to instructors but, others factors can function as biases in process of teacher evaluation. Some alternatives suggested to ameliorate teacher evaluation and diminish biases in these evaluation processes are: 1. apply the questionnaires before final exams; 2. giving more weight to opinions of the best students, or 3. listen to the lower grades students teacher evaluation in a personal interview. The two last alternatives were not evaluated yet certainly due to the anonimity concerns. However, in some cases this anonymity allowable to the students does not seem to be fair to the teachers.

Furthermore, in a positive attitude, institutions could create stimulus to enhance students response rate to questionnaires, giving to the students an extra point to final grade to stimulate the responsiveness average to questionnaires. This could improve the response rate of those students, avoiding reasons to have an attitude of revenge.

Some ideas have been proposed to improve questionnaires, but none of them can avoid this dimension of revenge and others students` perceptions that bring biases to evaluation process, therefore, we propose applying the questionnaires before final exams.

Teacher evaluation by the students depends on multiple factors beyond the control of the teachers, and administrators must look at the results of these evaluations with caution to avoid injustice to the teacher.

Ethical Approval: No ethical approval is required for this work.

The authors declare no conflict of interest for this work.

References

1. Bedard K, Khun P. Where class size matters: class size and student ratings instructors effectiveness. *Economics of Education Review* 2008; 27:253-265
DOI: [10.1016/econodurev.2006.08.007](https://doi.org/10.1016/econodurev.2006.08.007)
2. Beechman R. Teaching quality and student satisfaction: Nexus or Simulacrum? *London Review of Education* 2009;7(2):135-146.
DOI: [10.1080/14748460902990336](https://doi.org/10.1080/14748460902990336)
3. Boring A, Ottoboni K, Stark PB. Student evaluations of teaching (mostly) do not measure teaching effectiveness. 2021.
<https://www.scienceopen.com/hosted-document?doi=10.14293/S2199-1006.1.SOR-EDU.AETBZC.v1>
DOI: [10.14293/S2199-1006.1.SOR-EDU.AETBZC.v1](https://doi.org/10.14293/S2199-1006.1.SOR-EDU.AETBZC.v1)
4. Braga M, Paccagnella M, Pellizari M. Evaluating student`s evalaution of professor. *Economic of Education Review* 2014. 41:71-88.
DOI: [10.1016/j.econedurev.2014.04.002](https://doi.org/10.1016/j.econedurev.2014.04.002)
5. Dommeyer CJ, Baum P, Hanna RW, Chapman KS. Gathering faculty teaching evaluations by in-class and online surveys: their effects on response rates and evaluations. 2004. 29(5):611-623
DOI: [10.1080/02602930410001689171](https://doi.org/10.1080/02602930410001689171)
6. Fan Y, Shepherrd LJ, Slavich E, Waters D, Stone M, Abel R, Johnston EL. Gender and cultural bias in student evaluations: why representation matters 2018; *PLoS ONE* 14(42): e0209749
DOI: [org/10.1371/journal.pone.0209749](https://doi.org/10.1371/journal.pone.0209749)
7. Goos M, Salomons A. Measuring teaching quality in higher education: assessing selection bias in course evaluations. *Res High Educ* 2017; 58:341-364
DOI: [10.101007/s1162-9429-8](https://doi.org/10.101007/s1162-9429-8)
8. Griffin B W. Grading leniency, grade discrepancy, and student ratings of instruction. *Contemp Educat Psychology* 2004;29(4):410-425
doi.org/10.1016/j.cedpsych.2003.11.001
9. Hornstein, HA. Student evaluations of teaching are an inadequate assessment tool for evaluating faculty performance. *Cogent Education* 2017;4:1304016 DOI:[10.1080/2331186X.2017.1304016](https://doi.org/10.1080/2331186X.2017.1304016)
10. Joye SW, Wilson JH. Professor age and gender affect student perceptions and grades. *J Scholarship of teaching and learning* 2015;15(4):126-138
DOI: [10.14434/josotl.v15i4.13466](https://doi.org/10.14434/josotl.v15i4.13466)

11. Khon J, Schoenfeld-Tacher R, Helleyer. The role of gender in teaching effectiveness ratings of faculty. *Academy of Educational Leadership J*; 2006. 10:121-137

doi.org/10.3102/0034654313496870

12. Langbeins L. Management by results: student evaluation of faculty teaching and the mis-measurement of performance 2008.; *Economics of Education Ver* 27:417-428

DOI: [10.1016/econedurev.2006.12.003](https://doi.org/10.1016/econedurev.2006.12.003)

13. Luo, Mingchu Neal. Student Response Rate and Its Impact on Quantitative Evaluation of Faculty Teaching". 2020. *The Advocate* 25 (2)

DOI: [10.4148/2637-4552.1137](https://doi.org/10.4148/2637-4552.1137)

14. Magnus JR, Peresetsky A A. Grade expectations: Rationality and overconfidence. *Front in Psychol* 2018; 8:1-10

DOI: [10.3389/fpsyg.2017.02346](https://doi.org/10.3389/fpsyg.2017.02346)

15. MacNell L, Driscoll A, Hunt Na. What´s in a name: exposing gender bias in student rating of teaching 2015; 40:291-303

DOI: [10.1007/s10755-014-9313-4](https://doi.org/10.1007/s10755-014-9313-4)

16. McPherson MA, Todd Jewell R, Kim, M. What determines student evaluation scores? A random effects analysis on undergraduated economics classes. *Eastern Economic J.* 2009, 35:37-51

DOI: [10.1057/palgrave.eej.9050042](https://doi.org/10.1057/palgrave.eej.9050042)

17. Morrison J. ABC of learning and teaching in medicine: Evaluation. *BMJ* 2003.326(7385):385-387.

DOI: [10.1136/bmj.326.7385.385](https://doi.org/10.1136/bmj.326.7385.385)

18. Novell C, Alston RM. I thought I got na A! Overconfidence across the economics curriculum. *The J of Economic Education* 2007;38(2):131-142

[DOI.org/10.3200/JECE.38.2.131-142](https://doi.org/10.3200/JECE.38.2.131-142)

19. Pinquart M, Pietzsch MC. Change in Students' Educational Expectations – A Meta-Analysis. *J Educational and Developmental Psychology* 2022.12(1):43-53

DOI: [10.5539/jedp.v12n1p43](https://doi.org/10.5539/jedp.v12n1p43)

20. Ponzio M & Scopp V. The good, the bad, The ugly: teaching evaluations, beauty and abilities. 2012; Michela Ponzio & Vincenzo Scoppa, 2012; Working Papers 201204, Università della Calabria, Dipartimento di Economia, Statistica e Finanza "Giovanni Anania" – DESF

<https://EconPapers.repec.org/RePEc:clb:wpaper:201204>

21. Remedios R, Lieberman DA. 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 J* 2008; 34(1):91-115

DOI: [10.1080/01411920701492043](https://doi.org/10.1080/01411920701492043)

22. Senior BA. Student teaching evaluations options and concerns. *J Construction Education*. 2000. 5(1):20-29

<http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.527.5468>

23. Shelvin M, Banyard P, Davies M, Griffiths MD. The validity of student evaluation of teaching in higher education: love me, love my lectures? *Assessment & Evaluation in Higher Education* 2000; 25(4):396-405

DOI: [10.1080/713611436](https://doi.org/10.1080/713611436)

24. Stroebe W. Why Good Teaching Evaluations May Reward Bad Teaching: On Grade Inflation and Other Unintended Consequences of Student Evaluations. *Perspect Psychol Sci*. 2016 Nov;11(6):800-816.

DOI: [10.1177/1745691616650284](https://doi.org/10.1177/1745691616650284)

25. Stroebe W. Student evaluations of teaching encourages poor teaching and contributes to grade inflation: a theoretical and empirical analysis. *Basic and Applied Social Psychology* 2020. 42(4):276-294.

DOI: [10.1080/01973533.2020.1756817](https://doi.org/10.1080/01973533.2020.1756817)

26. Spooren P. On the credibility of the judge. A cross-classified multilevel analysis on student evaluations of teaching. *Studies in Educational Evaluation* 2010. 36:121-131

DOI: [10.1016/j.stueduc.2011.02.001](https://doi.org/10.1016/j.stueduc.2011.02.001)

27. Spooren P. On the credibility of the judge. A cross-classified multilevel analysis on student evaluations of teaching. *Studies in Educational Evaluation* 2010. 36:121-131

DOI: [10.1016/j.stueduc.2011.02.001](https://doi.org/10.1016/j.stueduc.2011.02.001)

28. Wall D. Evaluation: Improving Practice, Influencing Policy *in Understanding Medical Education: Evidence, Theory and Practice* - Chapter 23, pages 336-351. Editor Tim Swanwick. Wiley Online Library 2010.

DOI: [10.1002/9781444320282.ch23](https://doi.org/10.1002/9781444320282.ch23)

29. Warner J, Simmons, A. Giving voice to students: a preliminar analysis of informal mid-term evaluations & procedural justice. *Academy of Educational Leadership Journal*; Arden 2015; 19(1): 71-79.

[p584373.pdf \(decisionsciences.org\)](https://www.decisionsciences.org/p584373.pdf)

TABLE 1. FACTORS WITH THE POWER TO INFLUENCE IN STUDENTS ANSWER TO QUESTIONNAIRES FOR TEACHER EVALUATION

FACTOR	TYPE OF INFLUENCE	STUDY/PAPER
Students revenge against their low grades in final exams	Students who got low grades in exams attribute poor performance to the teachers	Wall D, 2010 Stroebe W, 2016 Morrison J, 2003
Final exam grades	The higher final grades, the higher teacher scores	Langbein L 2008 Spooren 2010
Course difficulty	More difficult is the course, lower is teacher evaluation scores	Remedios R & Lieberman DA 2008
Students age	The older is the student, the better is the teacher scores	Spooren 2010
Female X Male students grades on teacher evaluation scores	Female students rated faculty effectiveness higher than male students	Khon J & Hatfield L, 2006
Teacher race	White instructors receive better scores	McPherson et al. 2009
Teachers gender	Students rated male teacher higher than female teacher	McNell et al. 2014 Fan Y et al. 2019 Joye SW & Wilson JH, 2015
Teacher character traits	Teachers with “charisma” get the higher score	Shevlin et al. 2000
Teachers age	Younger teachers receive better scores	McPherson et al. 2009 Spooren 2010
Teacher physical appearance	Better looking instructors receive higher rates	Ponzo M & Scoppa V, 2012
Class size	Large class size has a negative impact on student evaluations of teacher	Bedard K & Khun P, 2005
Response rate to questionnaires	The lower response rate, the lower teacher scores in questionnaires	Luo MN, 2020
Private X Public institutions	Grade inflation is greater in Private institutions	Langbein L 2008
Class attendance	Students who attend most classes provide higher rates to teacher	Spooren 2010