



International Journal of Educational Methodology

Volume 5, Issue 3, 401 - 406.

ISSN: 2469-9632

<http://www.ijem.com/>

Potential for Blocking Advancement: Teaching Surveys for Student Evaluation of Lecturers

Eyal Eckhaus

Ariel University, ISRAEL

Nitza Davidovitch*

Ariel University, ISRAEL

Received: May 14, 2019 • Revised: July 21, 2019 • Accepted: August 5, 2019

Abstract: In the current study we examined the relationships between student evaluations of lecturers (teaching surveys) and faculty members' perceptions of these surveys as capable of blocking and limiting their professional advancement. Faculty members are judged and evaluated by academic authorities for their academic performance in research and teaching. 178 questionnaires were collected from the faculty of several academic institutions. We employ a mix method analysis, and form a model that reflects the factors perceived by faculty members as having the potential to block their professional advancement in academia. The research findings show that lecturers are of the opinion that teaching load has a detrimental effect on students' evaluations in the surveys. Lecturers at the beginning of their academic life, those in lower ranks: senior teacher and senior lecturer, address the negative aspects of the surveys more than others. The research findings indicate that although more hours are taught in colleges than at universities, it is harder to receive positive survey ratings at colleges. Moreover, since in Israeli academia research is still the main criterion for promotion – faculty members born in Israel were found to teaching less than those born elsewhere. Hence, faculty members think that student surveys are destructive and entail risks for their professional advancement. Assuming that students' voice and opinions on teaching are important – how can a balance be achieved between the research achievements of faculty members and student satisfaction?

Keywords: *Blocking advancement, teaching surveys, student evaluation.*

To cite this article: Eckhaus, E., & Davidovitch, N. (2019). Potential for blocking advancement: Teaching surveys for student evaluation of lecturers. *International Journal of Educational Methodology*, 5(3), 401-406. <https://doi.org/10.12973/ijem.5.3.401>

Introduction

Teaching evaluation surveys a century ago, however they have increased from the 1970s (Hativa, 2008). In the first years of their use, these questionnaires were an evaluation tool with no significant consequences, however over time the status of the surveys increased and from a mere evaluation tool providing recommendations for improvement, they were transformed into a tool with high value implications for academic careers – awarding tenure, promotion, and academic appointments (Ehie & Kararhanos, 1994; Harrison, Douglas & Burdsal, 2004; Smith & Pollak, 2008; Eckhaus & Davidovitch, 2019).

These aspects have an inevitable effect on salaries, perceived prestige, and career development options. Surveys results are accorded great significance and provide superiors with a measure. The considerable dominance of the evaluation surveys, which some call the "evaluation mandate" (Chandler, 1978), dictates and determines the career development course of lecturers to a not insignificant degree.

These implications have transformed the surveys into a source of much criticism by academic faculty (Smith & Pollak, 2008) and a focus of many studies, seeking to explore the reliability and validity of the surveys – to what degree students' perceptions should be granted the power to influence the career of academics. This topic is largely debated, with great controversy in all research on teaching (Kulik, 2001).

On one hand there are those who see the very existence of the survey as a judgmental means that is harmful to the lecturer's academic status, puts him at a disadvantage, and is forced on him (Eckhaus & Davidovitch, 2019). Advocates of this approach feel that they are supposed to "provide the goods" for the students, while the latter in return give them credit points. This perspective usually derives from a lack of dialogue between the lecturers and their superiors, among the lecturers themselves, and between the lecturers and the students (ibid.). These are countered by those who support

* Corresponding author:

Nitza Davidovitch, Ariel University, Education Department, ISRAEL. ✉ d.nitza@ariel.ac.il



the surveys and find them significant for advancing and improving the quality of teaching and the satisfaction of the students-consumers (Marsh, 1987; Marsh & Roche, 1994).

Despite the controversy, everyone agrees that improving the quality of teaching and promoting "efficient teaching" is important (National seeks to improve retention, graduation rates, 2002). The dispute is not about the need for control rather about how it should be achieved and how efficient it is (Wachtel, 1998; Feldman, 1978; Chen & Hoshower, 1998). Studies show that teaching load, as reflected by classroom size and number of hours in the curriculum (Feldman, 1978; Chen & Hoshower, 1998; Wachtel, 1998), is one of the significant factors in students' evaluations of faculty – the higher the teaching load, the more negative the evaluations.

Hence, this element of teaching load is perceived as a potential bias factor that might affect lecturer's teaching quality grades, as the quality of the teaching is not unaffected by context and does not isolate skills, proficiencies, and abilities of the lecturer from additional elements active in the student-lecturer-classroom interaction. There is a certain logic to the claim made by faculty, particularly those at the beginning of their academic life, as evident in their academic rank, who do not see student feedback as a valid means for evaluating their work's quality, not to mention as a means of dictating their future. On the relationship between personal and professional background features of academic faculty and their evaluation by students in teaching surveys

Thousands of studies have been conducted on teaching evaluations, in the attempt to comprehend the different elements influencing students when grading lecturers. A range of studies attempted to regulate and influence student ratings, while providing suggestions and advice on how to bridge potential biases (Baldwin & Blattner, 2003). However, while some try to offer means of defense, others propose examining those factors that individuals are unable to change – features such as ethnicity, age, and gender. These demographics may theoretically also have a part in students' evaluation of lecturers.

Seniority and rank

Many longitudinal studies conducted in the different departments have found that teaching quality diminishes as a function of seniority. The studies found that, lacking external intervention, there is a moderate and constant decrease in teaching quality (Ryans, 1960; Barnes, 1985), with a negative correlation between teaching experience and efficacy. Seniority and rank variables are often investigated together, due to the clear interface between all the variables.

Feldman (1983) explored how these variables affect evaluations of lecturers. The research results indicated negative correlations and an inverse ratio between teaching quality and seniority (Feldman, 1997; Marsh, 1987; Renaud & Murray, 1996). Feldman (1983) also found that grading of lecturers are positively correlated with rank, with higher ranks generating higher evaluations by students.

Ethnicity

Ethnicity's effect on student evaluations is an area that has been neglected relative to other background characteristics. Most of the studies explore ethnicity as secondary to other background variables perceived as more central (Feldman, 1997). Ethnicity was found to bias evaluations of lecturers and to reflect discrimination by students. Ethnicity and evaluations of lecturers were found to be positively correlated (Worthington, 2002; Centra, 1993).

In light of the research literature on teaching load, we sought to focus and to examine in the current study how much truth there is in lecturers' claims of bias stemming from this external factor of teaching load, which is not directly related to the quality of their teaching.

Research hypotheses

H1. Institution type (*INT*) negatively affects *TEACH*

H2. Institution type (*INT*) negatively affects *HOURS*

H3. *TEACH* positively affects *NEGATIVE*

H4. *LECTURER* positively affects *TEACH*

H5. *LECTURER* positively affects *NEGATIVE*

H6. *NEGATIVE* positively affects *SURVEY*

H7. *STR* and *SLT* positively affect *NEGATIVE*.

H8. *SLT* positively affect *NEGATIVE*

H9. *STR* positively affects *Hours*

H10. *ISR* negatively affects *TEACH*.

Methodology

Initial Sample

We employed Google Docs to distribute the surveys to faculty members of 7 academic institutions. 178 questionnaires were collected from Ariel University (88), Ben-Gurion University (21), Emek Yezreel Academic College (20), while 7 respondents did not mention their affiliation. 47.3% of the respondents were female and 52.7% male. Respondents' age ranged from 22-39 (16.8%), 40-49 (41.3%), and 50+ (41.9%). Most of the respondents were born in Israel (83.4%). Academic ranks were as follows. Teacher 5.9%, Senior teacher 4.1%, Instructor Dr. 1.8%, Lecturer 31.4%, Senior lecturer 34.3%, Associate professor 11.2%, Full professor 2.4%, and Other 8.9%. The survey included the following open-ended questions: "Do you think the teaching evaluation survey format should change, and how?", "In what way do you think the teaching evaluations cause harm?", "In your opinion, are there alternatives for the teaching evaluation survey, in order to assess teaching? What are they?", "What is the value of the teaching evaluations?" along with demographic questions, the type of the institution (university or college), and the average weekly number of hours taught.

Survey Analysis

We used Structural Equation Modeling (SEM) to test the model's goodness-of-fit (Eckhaus & Sheaffer, 2018a, 2018b; Mueller & Hancock, 2018). Model fit was estimated using CFI, TLI, RMSEA, and CMIN / DF. Values of CFI and TLI close to .95 or higher are an indication of good fit (Nascimento, Baelum, Dahlen, & Lopez, 2018), a CMIN / DF ratio lower than 3 is considered a good fit (Eckhaus & Davidovitch, 2019; Wu, Chen, & Yuan, 2018). RMSEA values should be $\leq .06$ (Hu & Bentler, 1999).

Automated text classification

We employed TEXTIMUS, a software that supports automated content analysis and text mining (Eckhaus & Ben-Hador, 2018, 2019). In the first step, n-gram frequencies were generated. N-gram refers to a contiguous sequence of n words from a given sequence of text (Davidovitch & Eckhaus, 2018b; Mehta, Parekh, & Karamchandani, 2018). Next, the 'Bag-of-Words' (BoW) technique was employed (Davidovitch & Eckhaus, 2018a; De Vries, Schoonvelde, & Schumacher, 2018; Eckhaus & Davidovitch, 2018a, 2018b), which is based on the concept of representation of documents as a collection of words, regardless of grammar and order. We analyzed the frequency of a set of keywords in the collection of documents, and compiled groups of the most frequent words employed for each of the research variables. Similar to studies that employed BoW in an organizational context (Eckhaus & Sheaffer, 2018c; Eckhaus, Taussig, & Ben-Hador, 2018), we then summed the words' occurrences to create the variables. *TEACH* – referring to teaching; *LECTURER* – referring to the teacher in the classroom; *THINK* - thinking before writing and serious approach towards the survey evaluation; *NEGATIVE* – negative feedback, difficulties, and challenges; *SURVEY* - referring to the teaching evaluation survey. Institution type (INT) was binary coded - university =1 or college= 0, and *HOURS* - the average weekly number of hours taught. The other demographic variables were similarly binary coded: Senior teacher rank (STR), Senior lecturer rank (SLR), and Israel as birth country (ISR), were coded 1=yes and 0=no.

Results

Spearman correlations, means, and standard deviations are presented in Table 1.

Table 1. Correlation matrix: Means, STD

	TEACH	LECTURER	NEGATIVE	SURVEY	INT	HOURS	STR	SLR	ISR
TEACH	-								
LECTURER	.21*	-							
NEGATIVE	.26***	.19*	-						
SURVEY	-.02	.12	.04	-					
INT	-.23**	-.03	-.17*	.03	-				
HOURS	.94***	.06	.04	0	-.24**	-			
STR	-.05	-.12	.16	-.003	.09	.15	-		
SLR	.14	.03	.14	-.02	.06	.17*	-.15	-	
ISR	-.11	.21*	.01	-.03	.08	-.11	.002	.02	-
Mean	.74	.68	.70	.50	.65	8.55	.04	.34	.83
SD	1.06	1.21	1.10	1.41	.48	4.01	.20	.48	.37

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

The demographic variables gender and age were added as controlled variables to *NEGATIVE*, *SURVEY*, *TEACH*, and *LECTURER*. *ISR* was added as a control variable to *TEACH* since most respondents were born in Israel.

Figure 1 illustrates the results. The hypothesized model showed a very good fit: CMIN / DF = 0.97, $p > .05$, CFI = 1, TLI = 1.02, RMSEA = 0. All hypotheses were supported. *INT* negatively affects *TEACH* (H1), *INT* negatively affects *HOURS* (H2), *TEACH* positively affects *NEGATIVE* (H3), *LECTURER* positively affects *TEACH* (H4) and *NEGATIVE* (H5), respectively. *NEGATIVE* positively affects *SURVEY* (H6), *STR* and *SLT* positively affect *NEGATIVE* (H7-H8 respectively), *STR* positively affects *HOURS* (H9). Since teaching methods may vary in different countries, we hypothesized a difference regarding *TEACH*, and as hypothesized *ISR* negatively affects *TEACH* (10)

Gender and age had no statistically significant effect on the variables.

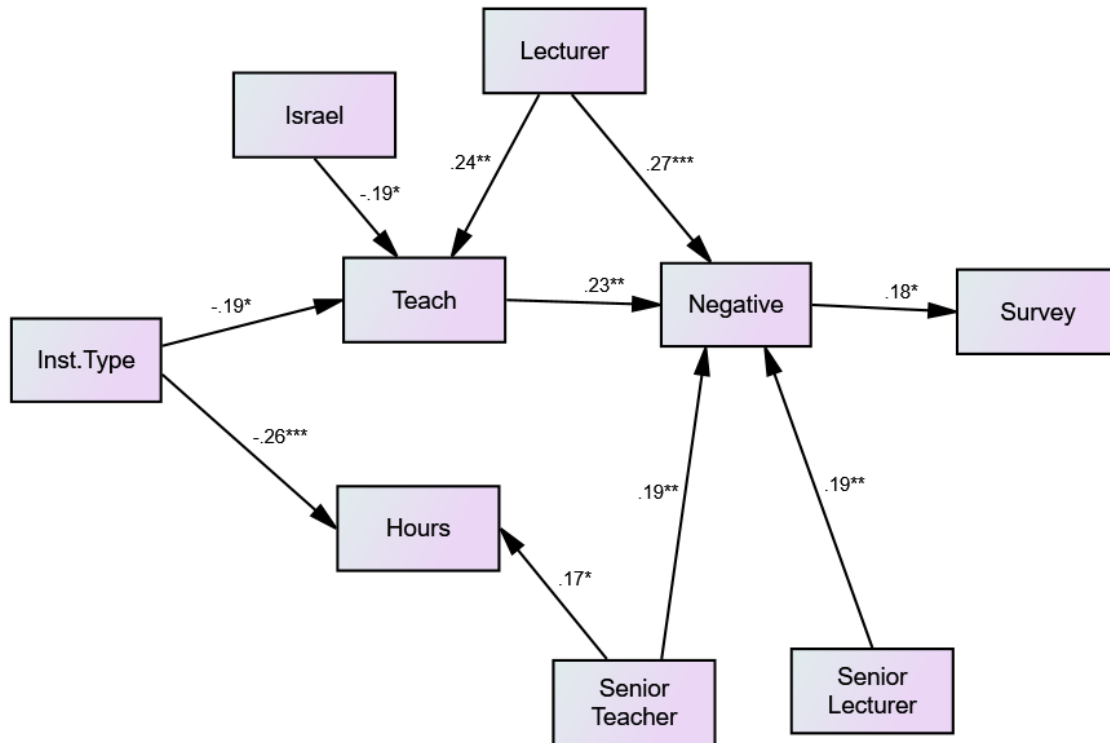


Figure 1. Model path and weights

The findings indicate that, as perceived by faculty members, **teaching load** (number of teaching hours per week) affects students' degree of satisfaction, as evident in the surveys, and generates less positive evaluations. According to the faculty, in addition to teaching it is also necessary to manage students' complaints, and more teaching means taking care of a larger quantity of complaints and problems.

The implications of teaching load are also evident in quotations from the statements of faculty members teaching at universities and colleges. Due to the teaching load, college lecturers reduce the level of their teaching. As one lecturer explains:

"At colleges evaluation surveys receive a disproportionate weight in evaluating lecturers. This is simply because they are the only measure. This adds unnecessary pressure and distorts the relationship between lecturers and students, harming the lecturer's ability to set limits and manage the class for fear of vengefulness".

Another lecturer further explains the implications:

"Many students who compare their knowledge level to that of university graduates complain about the low level and the minimal demands made of them."

These quotations express faculty members' perceptions of the surveys as having the potential to block and limit their professional advancement. Faculty members are judged and assessed by the academic authorities for their performance in research and teaching. Since the dominant criterion for promotion in academic systems is research, faculty members estimate that of all their activity, research receives considerable weight as a decisive factor in their promotion. In addition, academic teaching has an impact as well, where faculty members are judged for how they convey knowledge.

Summary and Discussion

The research literature indicates that evaluation of academic faculty by the students has far reaching implications for academic careers, implications that have been growing in recent years. Student evaluations of lecturers inevitably affect salaries, perceived prestige, and career development options. Therefore, these implications transform the surveys into

a source of much criticism by the academic faculty and an object of many studies that question the reliability and validity of the surveys, as a tool which gives students the power to influence the careers of academics.

This article attempts to deal with the criticism concerning the relationship between the teaching load, reflected in the faculty's teaching hours, and student evaluations of lecturers in teaching surveys. Many lecturers feel that they must "provide the goods" for the students, who in return award them credit points.

The research findings show that, according to the lecturers, the *teaching load* (number of teaching hours a week), which is an external factor that is not dependent on them and on the quality of their teaching, is detrimental to students' evaluations in teaching surveys. In addition to the teaching load, the lecturers take care of student complaints, and the more they teach they must take care of more complaints and problems. Lecturers who are at the beginning of their academic life, those who are in lower ranks: senior teacher and senior lecturer, address the negative aspects of the surveys more than the others. The research findings show that although more hours are taught at colleges than at universities it is harder to receive good evaluations at colleges. Moreover, since in Israeli academia research is still the main criterion for promotion – faculty members born in Israel were found to address the topic of teaching less than those born elsewhere.

Hence, faculty members think that student surveys are destructive, entailing a real risk to their professional advancement. Assuming that students' voice and opinions about teaching are important – the question is how to arrive at a balance between the research achievements of faculty members and student satisfaction.

References

- Baldwin, T., & Blattner, N. (2003). Guarding against potential bias in student evaluations. *College Teaching*, 51(1), 27-33.
- Barnes, B. (1985). *About science*. Oxford, UK: Basil Blackwell.
- Chen, Y., & Hoshower, L. B. (1998). Assessing student motivations to participate in teaching evaluations: an application of expectancy theory, *Issues in Accounting Education*, 13(3), 531-549.
- Centra, J. (2003). Will teachers receive higher student evaluations by giving higher grades and less course work? *Research in Higher Education*, 44(5), 495-518.
- Chandler, T. (1978). The questionable status of student evaluations of teaching. *Teaching of Psychology*, 5(3), 150-152.
- Davidovitch, N., & Eckhaus, E. (2018a). Effect of faculty on research cooperation and publication: Employing natural language processing. *Economics and Sociology*, 11(4), 173-180. <http://dx.doi.org/10.14254/2071-789X.2018/11-4/11>
- Davidovitch, N., & Eckhaus, E. (2018b). The influence of birth country on selection of conference destination-employing natural language processing. *Higher Education Studies*, 8(2), 92-96.
- De Vries, E., Schoonvelde, M., & Schumacher, G. (2018). No longer lost in translation: Evidence that Google Translate works for comparative bag-of-words text applications. *Political Analysis*, 26(4), 417-430.
- Eckhaus, E., & Ben-Hador, B. (2018). To gossip or not to gossip: Reactions to a perceived request to gossip – a qualitative study. *Trames: A Journal of the Humanities and Social Sciences*, 22(3), 273-288. <https://doi.org/10.3176/tr.2018.3.04>
- Eckhaus, E., & Ben-Hador, B. (2019). Gossip and gender differences: a content analysis approach. *Journal of Gender Studies*, 28(1), 97-108. <https://doi.org/10.1080/09589236.2017.1411789>
- Eckhaus, E., & Davidovitch, N. (2018a). Impact of gender and conference size on conference preferences – employing natural language processing. *International Journal of Educational Methodology*, 4(1), 45-52. <https://doi.org/10.12973/ijem.4.1.45>
- Eckhaus, E., & Davidovitch, N. (2018b). Improving academic conferences – criticism and suggestions utilizing natural language processing. *European Journal of Educational Research*, 7(3), 445-450.
- Eckhaus, E., & Davidovitch, N. (2019). How do academic faculty members perceive the effect of teaching surveys completed by students on appointment and promotion processes at academic institutions? A case study. *International Journal of Higher Education*, 8(1), 171-180.
- Eckhaus, E., & Sheaffer, Z. (2018a). Factors affecting willingness to contribute goods and services on social media. *The Social Science Journal*. <https://doi.org/10.1016/j.soscij.2018.08.001>
- Eckhaus, E., & Sheaffer, Z. (2018b). Happiness enrichment and sustainable happiness. *Applied Research in Quality of Life*. <https://doi.org/10.1007/s11482-018-9641-0>

- Eckhaus, E., & Sheaffer, Z. (2018c). Managerial hubris detection: the case of Enron. *Risk Management*, 20(4), 304-325. <https://doi.org/10.1057/s41283-018-0037-0>
- Eckhaus, E., Taussig, R., & Ben-Hador, B. (2018). The effect of top management team's tacit persuasion on the stock market. *e - Journal of Social & Behavioural Research in Business*, 9(2), 9-22.
- Ehie, I. & Karathanos, D. (1994). Business faculty performance evaluation based on the new AACSB accreditation standards. *Journal of Education for Business*, 69(5), 257- 262. doi: 10.1080/08832323.1994.10117695
- Feldman, K. A. (1983). The seniority and instructional experience of college teachers as related to the evaluations they receive from their students. *Research in Higher Education*, 18, 3-124.
- Feldman, K.A. (1978) Course characteristics and college students' ratings of their teachers: what we know and what we don't. *Research in Higher Education*, 9(2), 199-242.
- Feldman, K. A. (1997). Identifying exemplary teachers and teaching: Evidence from student ratings. In R. P. Perry & J. C. Smart (Eds.), *Effective teaching in higher education: Research and practice* (pp.368-395). New York: Agathon.
- Harrison, P., Douglas, D., & Burdsal, C. (2004). The relative merits of different types of overall evaluations of teaching effectiveness. *Research in Higher Education*, 45(3), 311-323.
- Hativa, N. (2008). Myths and facts about evaluation surveys by students. *Al Hagova*, 7, 13-14. [In Hebrew]
- Hu, L.-t., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1-55.
- Kulik, J. A. (2001). Student ratings: validity, utility, and controversy. *New Directions for Institutional Research*, 2001(109), 9-25.
- Marsh, H. W. (1987). Students' evaluations of university teaching: Research findings, methodological issues, and directions for future research. *International Journal of Educational Research*, 11(3), 253-288.
- Marsh, H. W. & Roche, L. A. (1994). The use of students' evaluations of university teaching to improve teaching effectiveness. *Final project report for the Evaluations and Investigations Program of the Department of Employment and Education*. Canberra, Australia: Australian Government Printing Service.
- Mehta, A., Parekh, Y., & Karamchandani, S. (2018). Performance evaluation of machine learning and deep learning techniques for sentiment analysis. In V. Bhateja, B. L. Nguyen, N. G. Nguyen, S. C. Satapathy, & D.-N. Le (Eds.), *Information systems design and intelligent applications* (pp. 463-471). Singapore: Springer.
- Mueller, R. O., & Hancock, G. R. (2018). Structural equation modeling *The reviewer's guide to quantitative methods in the social sciences* (pp. 457-468). Abingdon, UK: Routledge.
- National survey seeks to improve retention, graduation rates. (2002). *Black Issues in Higher Education*, 19(14), 18.
- Nascimento, G. G., Baelum, V., Dahlen, G., & Lopez, R. (2018). Methodological issues in assessing the association between periodontitis and caries among adolescents. *Community Dentistry and Oral Epidemiology*, 46(3), 303-309. <https://doi.org/10.1111/cdoe.12367>
- Renaud, R. D., & Murray, H. G. (1996). Aging, personality, and teaching effectiveness in academic psychologists. *Research in Higher Education*, 37(3), 323-340.
- Ryans, D. G. (1960). Prediction of teacher effectiveness. In C. W. Harris (Ed.), *Encyclopedia of educational research* (pp. 1486-1491). New York, NY: Macmillan.
- Smith, K., & Pollak, M. W. (2008). What can they say about my teaching? Teacher educators' attitudes to standardised student evaluation of teaching. *European Journal of Teacher Education*, 31(2), 203-214.
- Wachtel, H. K. (1998). Student evaluation of college teaching effectiveness: a brief review, *Assessment and Evaluation in Higher Education*, 23(2), 191-211.
- Worthington, A. (2002). The impact of student perceptions and characteristics on teaching evaluations: a case study in finance education. *Assessment & Evaluation in Higher Education*, 27(1), 49-64.
- Wu, K., Chen, S., & Yuan, Y. (2018). Research on the customer loyalty of bicycle-sharing company based on PLS-SEM model. In *Proceedings of the 2018 2nd International Conference on Management Engineering, Software Engineering and Service Sciences* (pp. 68-72). Newyork, NY: The Association for Computing Machinery.