

# Student Evaluation of Lecturers – What do Faculty Members Think about the Damage Caused by Teaching Surveys?

Prof. Nitza Davidovitch<sup>1</sup> & Dr. Eyal Eckhaus<sup>2</sup>

<sup>1</sup>Department of Education, Ariel University, Ariel, Israel

<sup>2</sup>Department of Economics & Business Administration, Ariel University, Ariel, Israel

Correspondence: Eyal Eckhaus, Department of Economics & Business Administration, Ariel University, Ariel, Israel. Tel: 972-3-906-6399. E-mail: eyale@ariel.ac.il

Received: April 5, 2019

Accepted: May 10, 2019

Online Published: June 4, 2019

doi:10.5539/hes.v9n3p12

URL: <https://doi.org/10.5539/hes.v9n3p12>

## Abstract

Many studies have been conducted on teaching evaluations and student surveys. The current study is unique for examining, by means of direct questions, the meaning of teaching surveys as perceived by academic faculty in Israel. Senior faculty members at academic institutions completed questionnaires, with a total of 182 questionnaires collected. We employed mixed research methods, beginning with qualitative analysis followed by Structural Equation Modeling (SEM), with the goal of developing a model that reflects faculty members' beliefs on teaching surveys. The research findings show that the lecturers find that student evaluations are detrimental to their relationship with their students, and adversely affect their teaching practice and interpersonal interactions with their students. In view of the importance attributed to students' voices and their opinions of teaching, the question is how should these evaluations be addressed, Do teaching surveys constitute a reliable managerial tool and a foundation for improving teaching – or should other tools be developed to improve teaching practices, independent of students' opinions?

**Keywords:** higher education, teaching, teaching evaluation, performance

## 1. Introduction

Evaluation of faculty members' activities and performance-based compensation have become important topics in recent decades. Faculty at academic institutions perform a wide range of activities, which include teaching, research, academic administration, and community service. However, academic freedom makes it hard to supervise and report the performance of faculty in the various fields of activity (Eckhaus, 2019, Mizrachi, 1994). In the last decade, awareness of the significance of quality in academic institutions has grown, in Israel as well as in Western Europe and the US. This heightened awareness is a result of the considerable competition between institutions, as well as of the high number and diversity of the students (Davidovitch & Eckhaus, 2019, Eckhaus & Davidovitch, 2018a, 2018b).

One topic that is gradually receiving more research attention is the appraisal of faculty performance. In the academic system, the primary criteria for evaluating academic faculty are research, teaching, and contribution to the community (Mizrachi, 1994). In the last decade, attention has been drawn in academia to quality assessment measures in general and to evaluation of academic teaching in particular (Wu, Chen, & Yuan, 2018).

Several methods are used to evaluate the performance of faculty members at academic institutions. The main tool for evaluating teaching is teaching surveys, administered to students at the conclusion of each semester. These performance evaluations are also used as inputs in faculty promotion decisions.

The current study examines the opinions of faculty members in Israeli academia on the significance and effects of teaching surveys. The study explores the pernicious effects that faculty members attribute to teaching surveys, and whether there is an association between demographic variables and the perceived harm caused by teaching surveys.

## 2. Literature Review

Lecturer evaluation surveys were first introduced to academia in the early 20th century, however they gained status and significance mainly beginning from the 1970s (Hativa, 2008). Use of surveys is most prevalent in the

United States and Canada, where extensive efforts are devoted to ensure the professional status of teaching and setting strict standards of knowledge and skills (Hativa, 2003a). Teaching surveys are perceived as a means of realizing these goals. Leading European countries have also come to embrace this method of evaluation (Rinderman & Schofield, 2001).

The uses of teaching surveys have expanded in the last decade from an evaluation tool that highlights areas for improvement to assessments whose results have significant effects on academic careers and decisions concerning appointments, tenure, and promotions (Harrison, Douglas, & Burdsal, 2004). These decisions have an inevitable impact on salaries, bonuses, status, and options for career development. These "assessment mandates" (Chandler, 1978) largely determine the development of faculty members' careers.

As a result, teaching surveys have become heavily criticized by academic faculty (Smith & Pollak, 2008) and a subject of many studies that examine the reliability and validity of the surveys, with the aim of determining the extent to which students' opinions and assessments should be given the power to influence the careers of academics. Some claim that this is the greatest controversy in all research on teaching (Kulik, 2001).

On one side are those who are in favor of the surveys and who recognize their significance as a means for promoting and improving the quality of teaching as well as the satisfaction of the students-clients (Marsh & Roche, 1993). On the other side are those who see the mere existence of the surveys as a judgmental means that is detrimental to the lecturer's academic status, sets him or her on an inferior level, and is applied forcibly from outside (Avdor, 2006). Proponents of this approach feel that they are expected to "provide the goods" for students, and the latter in return award them credits. This point of view often derives from a lack of dialogue between the lecturers and their superiors, among the lectures, and between the lecturers and the students (Avdor, 2006). Many faculty members argue against student evaluations (Hativa, 2008), contending that these are biased for various reasons that include areas that are not under the lecturer's control, such as the course features: degree of difficulty, mass of material; (Chen & Hoshower, 1998); lecturer features: appearance, sex, age, ethnicity (Anderson & Siegfried, 1997; Wachtel, 1998); student features: personal expectations, sex, motivation, class attendance (Anderson & Siegfried, 1997; Chen & Hoshower, 1998; Davidovitch & Notzer, 2004); and administrative features: classroom size, number of curriculum hours (Chen & Hoshower, 1998; Wachtel, 1998). Each of these factors is perceived as a potential cause of bias that might affect ratings of the lecturer's quality of teaching, as the nature of teaching is not unrelated to context and does not separate a lecturer's skills, capabilities, and capacities from other factors that operate within the student-lecturer-class interaction. Hence, there is a certain amount of logic to the claim of faculty members who do not recognize student evaluations as an efficient means of assessing the quality of their work and certainly not as a means of dictating their future.

Although the views on this issue are divided between supporters and objectors, all agree that it is important to promote "effective teaching", improve the quality of teaching, and reduce the number of students who drop out of school. While legitimate teacher authority is a traditional foundation for effective teaching, governments and universities try to develop new practices and policies to encourage, measure and reward "good teaching" (Ballantyne, Borthwick, & Packer, 2000).

Thus, the question is whether academia relies excessively on teaching surveys. Are the surveys valid and reliable or do the results of teaching surveys simply reflect mutual manipulations of students (who seek to receive high grades) and lecturers (who seek high ratings).

### *2.1 Research Rationale*

One study on student attendance of courses and their evaluation of lecturers (Davidovitch & Notzer, 2004) found a significant positive association between students' course attendance and their evaluations in all five faculties included in its case study of an academic institution. Other background variables checked found associations only between students' year in their program, attendance, and teaching evaluations.

Another study (Hativa, 2003b) examined the association between student evaluations and grades in courses. The research data reinforce the findings on course-level data by Hativa (2008), indicating the lack of an association between grades and student evaluations of lecturers.

A study by Hativa, Many, and Dayagi (2010) on the whys and wherefores of teacher evaluation by their students, involved a case study of Tel Aviv University. The study identified the main factors related to classroom teaching that affect students' perceptions of the quality of teaching. The study also identified differences in teachers' ratings when comparing poor and outstanding students, low-rated and high-rated teachers, differences in participation rates in the online teaching survey by poor versus outstanding students.

The qualitative and empirical analyses investigate typical faculty members' complaints, as described in the

following research hypotheses.

1. Complaints about evaluations by non-attending students

The teaching evaluations of students with poor course attendance contain more complaints in the comment section.

2. The fact that students who do not attend the course complete surveys generates vengeful responses by students (for instance, they do not understand the material because they were not present in class).

Complaints that students use the teaching surveys to venting their frustration

3. The fact that evaluations are utilized as a sphere for venting frustration lead to taking revenge against lecturers through harmful comments.

4. The fact that evaluations are utilized as a sphere for venting frustration leads to irrelevant comments in the evaluations.

5. Disrespectful language

valuations that contain offense comments are irrelevant for an evaluation of teaching.

6. Only those who complain complete surveys

The fact that only those who complain complete surveys results creates a disproportionate impact by a small number of people.

The fact that only those who complain complete surveys, results in completion of surveys by a small group of people.

7. Disproportionate impact

Since negative evaluations by a small number of people have a disproportionate impact, evaluations are not a reliable tool for improving teaching.

8. Comments that are irrelevant for teaching

Since the evaluations include comments that are irrelevant for teaching, they are not a reliable tool for improving teaching.

9. Small number of respondents

The fact that only a small number of respondents complete the evaluations gives these individuals undue influence.

### 3. Method

#### 3.1 Initial Sample

A questionnaire was used for data collection, incorporating an open-ended question: "How do you think that the student teaching evaluations might be damaging to you?" and several demographic questions, such as gender and age, and information on the academic institution.

We employ a mixed-methods design (Ghasempour, Bakar, & Jahanshahloo, 2014), incorporating both qualitative and empirical techniques, with a main focus on an explanatory approach. The questionnaires were distributed online to the senior faculty members of seven academic institutions using Google Docs. One-hundred and eighty-two questionnaires were completed and returned. The majority of the respondents were from three institutions. Ninety-one respondents were from Ariel University, 21 from Ben Gurion University, and 20 from the Jezreel Valley Academic College. Seven respondents were from four other institutions, and the remainder did not complete their affiliation. Regarding gender, 47.9% of the respondents were female and 52.1% male. Respondents were distributed by age as follows: 22-39 (17.1%), 40-49 (41.4%), and 50 and older (41.4%).

#### 3.2 Data Analysis

First, major themes were identified in the texts of faculty members' responses to the open-ended question (Eckhaus & Davidovitch, 2018a, 2018b). This process of identifying categories within texts is known as text categorization. Next, a model was constructed based on the identified categories (Table 1).

Table 1. Main themes

Variable	Theme	No. of respondents
<i>Complain</i>	Only students who complain complete the teaching survey	13
<i>Attend</i>	Only students with poor attendance completed the survey, due to academic incentives, often causing them to fabricate ratings	18
<i>Insult</i>	Insults and abuse, disrespectful language	16
<i>Revenge</i>	The survey is used as a tool for taking revenge against the lecturer	21
<i>Unreliable</i>	The survey is an unreliable tool for improving teaching	32
<i>Irrelevant</i>	The comments are irrelevant for teaching	19
<i>Small</i>	Only a small number of students complete the survey and therefore the feedback does not truly reflect the quality of teaching	27
<i>Excessive</i>	Negative evaluations by a small number of individuals have a disproportionate influence on the overall evaluation	19
<i>Frustration</i>	The survey serves as sphere for venting frustration	17

Categories were binary coded: Each statement was marked 0 if the text did not belong to the category, and 1 if it did belong (Eckhaus & Sheaffer, 2018b). We employed a mixed-methods design (Eckhaus & Ben-Hador, 2017), a combination of strength, integrating qualitative and empirical analyses, with the aim of gaining better understanding of the research problem (Tomas, 2017).

Structural Equation Modeling (SEM) was applied to test the model's goodness-of-fit (Eckhaus, 2019; Eckhaus & Sheaffer, 2018a). Model fit was estimated using CFI, TLI, RMSEA, NFI, and minimum discrepancy divided by their degrees of freedom (CMIN/DF). Values of CFI and TLI close to .95 or higher are an indication of good fit, the ratio CMIN/DF should be as small as possible, lower than 3 is typically considered a good fit (Hinz et al., 2017). RMSEA should be .06 or smaller (Hu & Bentler, 1999).

#### 4. Qualitative Analysis

Below are several examples of respondents' responses, by hypothesis:

H1a. The teaching evaluations of students with poor course attendance contain more complaints in the comment section.

"Students who do not attend provide evaluations. Students who did not get what they wanted give devastating evaluations. Then [you have to try to] justify this to your superiors."

"In courses I teach attendance is not mandatory. Many students who do not attend class complete surveys and fabricate lecturer ratings."

"The surveys let students who did not even attend class express their opinion and thus distort the picture."

H1b. The fact that students who do not attend the course complete surveys generates vengeful responses by students (for instance, they do not understand the material because they were not present in class).

"Particularly harmful are evaluations by students who do not attend most of the lectures, [whose low ratings are] a type of revenge for their lack of success [in the course]."

"Students who did not attend provide evaluations. Students who did not get what they wanted give devastating evaluations. Then [you have to try to] go and justify this to your superiors."

H2a. The fact that evaluations are utilized as a sphere for venting frustration leads to taking revenge against lecturers.

"...And those who complete them can be the same students whose made requests [for exemptions regarding] course conditions and the syllabus that you did not approve, or who were reprimanded in class for disrupting the class (discipline problems)."

"The surveys are usually administered in the three final classes, after many students have received letters of warning before expulsion and subsequently attend class. These student complete the surveys although they were in fact hardly present during the course (sometimes for barely 2 classes), and [their evaluations reflect their] anger at the attendance requirement."

H2b. The fact that evaluations are utilized as a sphere for venting frustration leads to irrelevant comments in the evaluations.

"...They complain about the level of difficulty, but they don't even make an effort to come and ask questions during office hours. They complain about the workload, etc. In other words, they lack a serious attitude to their studies and they have no understanding of what it means to take responsibility for their actions. Regrettably, I have nothing to learn from these comments if I am not willing to lower the standards of my course."

"Students who find the course difficult will give negative evaluations."

H3. Evaluations that contain offense comments are irrelevant for an evaluation of teaching.

"Abusive comments, particularly regarding the lecturer's accent, which is not something that can be changed... When students complete the surveys after they were graded on an assignment, the ratings were unusually severe and unrelated to how the class was conducted."

H4a. The fact that only those who complain complete surveys creates a disproportionate impact by a small number of people.

"Particularly harmful are evaluations by students who do not attend most of the lectures, [their ratings serve] as a type of revenge for their failure."

"Many times most of the feedback is completed by students at the extreme ends of the spectrum – either those who did very well in the course or those who failed the course, and not by the silent majority. Therefore, relying on evaluations portrays an incomplete and not a very accurate picture of the class's feelings about me."

"Most students who are satisfied with the course do not complete the survey (that's what they tell me) and this generates an incomplete picture. In any case, the number of evaluations is very small and you cannot really compare teaching a class of 20-30 students to a class of 100 students or more."

H4b. The fact that only those who complain complete surveys results in completion of surveys by a small group of people.

"Usually those who make a point of completing the survey are those who had problems with the course – low grades, behavior problems, etc."

"As I see it, verbal comments (negative or positive) are completely unnecessary. They reflect the personal opinion of a single student who probably does not represent what others think (one way or the other)."

"Sometimes there is a sense of dissonance between the results of the survey and the actual situation, perhaps because it is more important for those with negative opinions to complete the survey than it is for those with positive opinions."

H5. Since negative evaluations by a small number of people have a disproportionate impact, evaluations are not a reliable tool for improving teaching.

"...I have often seen that only a very small proportion of students in the course complete the survey and therefore the results, which are based on such a small number of participants and the manner of calculation, which utilizes averages, do not truly reflect anything."

"Sometimes the survey presented includes a small proportion of students and then it does not sufficiently reflect reality."

H6. Since the evaluations include comments that are irrelevant for teaching, they are not a reliable tool for improving teaching.

"Abusive comments, particularly regarding the lecturer's accent, a topic on which is not something that can be changed there is nothing to change or improve... When students complete the surveys after they were graded on Evaluations provided after receiving a grade for an assignment, the ratings were unusually severe and unrelated to how the class was conducted. The fact that no specific questions are asked and that students can only give a general rating allows those with a chip on their shoulder to take revenge on the lecturer. This also provides support for charismatic lecturers and comes at the expense of good lecturers who are less charismatic, particularly in classes with a large number of students. As a result, the teaching survey is more like rating a 'reality show' and has become less practical."

"Sometimes there are comments from which it is hard to understand how to improve."

H7. The fact that only a small number of respondents complete the evaluations gives these individuals undue influence.

"They are completed by a small group and therefore do not reflect the class's true feelings."

"Sometimes the survey includes a small proportion of students and then it does not sufficiently reflect reality."

"Surveys completed by a small proportion of the students do not necessarily reflect the opinion of all students."

### 5. Empirical Results

Spearman's correlations, means, and standard deviations are presented in Table 2. Figure 1 illustrates the model and the results.

Table 2. Correlation matrix: means, SD, and reliability

	<i>Complain</i>	<i>Attend</i>	<i>Insult</i>	<i>Revenge</i>	<i>Unreliable</i>	<i>Irrelevant</i>	<i>Small</i>	<i>Excessive</i>	<i>Frustration</i>
<i>Complain</i>	-								
<i>Attend</i>	.26***	-							
<i>Insult</i>	-.01	-.04	-						
<i>Revenge</i>	.03	.28***	-.05	-					
<i>Unreliable</i>	-.02	.04	-.04	.06	-				
<i>Irrelevant</i>	-.03	-.05	.28***	-.01	.27***	-			
<i>Small</i>	.36***	.28***	-.08	.09	.21**	.01	-		
<i>Excessive</i>	.32***	.13	-.04	.05	.27***	.001	.57***	-	
<i>Frustration</i>	-.09	.15*	-.03	.42***	.05	.20**	.08	.01	-
Mean	.07	.10	.09	.12	.18	.10	.15	.10	.09
SD	.26	.30	.28	.32	.38	.31	.36	.31	.29

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

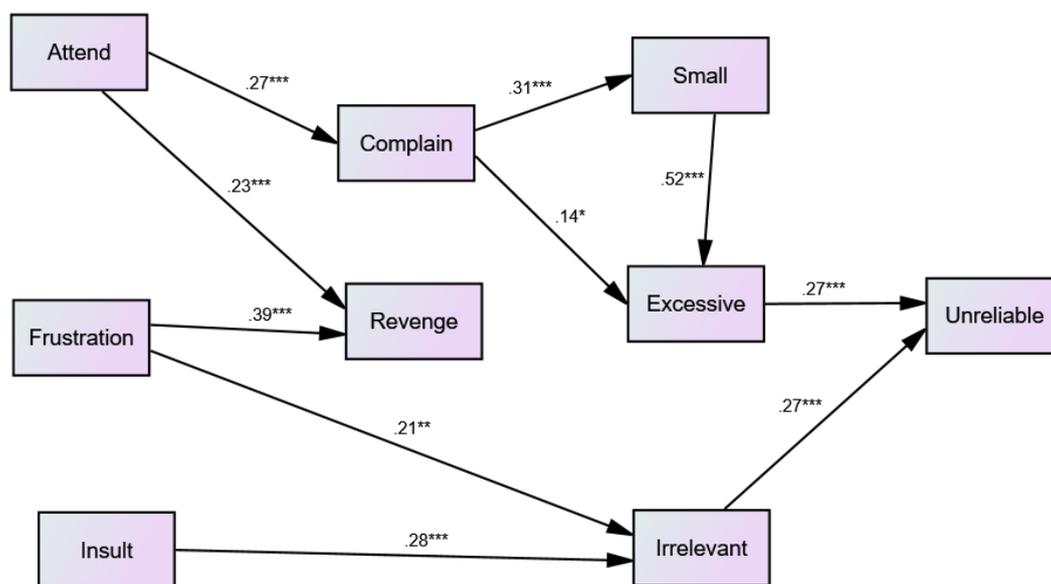


Figure 1. Model results

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

The hypothesized model shows a very good fit: CMIN/DF = 0.85 ( $p > .05$ ), CFI = 1, TLI=1.03, RMSEA = 0. All hypotheses were supported. Attend positively affected Complain (H1a) and Revenge (H1b), Frustration positively affected Revenge (H2a) and Irrelevant (H2b), Insult positively affected Irrelevant (H3), Complain positively affected Small (H4b) and Excessive (H4a), Excessive positively affected Unreliable (H5), Irrelevant affected Unreliable (H6), and Small positively affected Excessive (H7).

In the next step, age, gender, and academic institution (college or university) were added as control variables, i.e., a relationship between age, gender, and academic institution and all other variables was added. Results still showed a good fit with the model, CMIN/DF = .94 ( $p > .05$ ), CFI = 1, TLI=1.02, RMSEA = 0. All previous relationships remained significant. With regard to age, gender, and academic institution, only the following relationships were significant: Age significantly affected only Revenge ( $B = .26, p < .001$ ), which means that older teachers expressed more complaints that the teaching survey is used by students as revenge against the

lecturer. Gender significantly affected only Insult ( $B = -.21, p < .01$ ), meaning that women complain about offensive remarks more than do men. Institution type affected only Frustration ( $B = -.16, p < .05$ ), meaning that faculty from colleges complain more than faculty from universities that the teaching survey is used to express frustration.

## 6. Conclusion

In this study, mixed empirical and qualitative methods were used, to examine lecturers' perceptions regarding the damage caused by teaching surveys and the associations between perceived damage and background variables. The research findings show that faculty do not consider teaching surveys to be a reliable measure of their performance or as a tool that can be utilized to improve teaching.

The background variables of faculty members affect their perception of the damage caused by teaching surveys. We checked whether and to what degree age, sex, and academic institution (college or university) constitute factors that affect perceived damage as seen by faculty members, in various categories. The findings indicate that these factors are indeed significant: Age significantly affects perceptions that evaluations are used as a tool of as revenge by students: Older faculty member subscribed to this view more than younger faculty members. Gender affects perceptions that teaching surveys are used to insult faculty members. That is to say, women complain about students' offense comments more than do men. Additionally, compared to faculty members at universities, faculty members at colleges were found to more strongly hold the opinion that students used teaching surveys as an opportunity to express their frustration.

Assuming that students' voices and their satisfaction with teaching is important and a significant resource for academic institutions, the Council for Higher Education, and the faculty members, the lecturers themselves – who wish to perform their job well – the question is how should these teaching surveys be treated: what do they indicate and who do they represent? Do they constitute a trustworthy management tool that can be utilized as a foundation for improving teaching – or should other tools be developed, unrelated to students' opinions?

Teaching surveys completed by students hold significance in the capitalist academic world, in the sense that they reflect the customers' opinion of a service that they consume. It is nonetheless necessary to critically judge the meaning of teaching surveys in the current era, particularly when students are accustomed to express their opinion on digital media with little thought or planning.

Many ways of evaluating the quality of teaching in academia are used worldwide (<https://cft.vanderbilt.edu/guides-sub-pages/teaching-portfolios>; see Table 3). Each method has its benefits and weaknesses, but it appears that a combination of methods may facilitate more reliable and balanced information on the quality of teaching and not only on student satisfaction.

Table 3. Strengths and weaknesses of accepted methods of teaching evaluation

<b>Evaluation source</b>	<b>Method</b>	<b>Benefits</b>	<b>Weaknesses</b>
By students	Teaching surveys	A relatively inexpensive large-scale measure of student satisfaction with teaching	Some aspects of a lecturer's teaching practice cannot be properly measured by the survey, such as the quality of the syllabus (relevance of the topics and literature) and others. Validity and reliability must be confirmed.
By superiors / academic colleagues / academic teaching experts	Peer evaluation	Requires consensus on the purpose of the process and manner of performance. The process also requires training observers for valid and reliable performance of the process. Requires at least two evaluators and time.	Subjective – The evaluations might be biased, in both directions, for reasons of internal politics or due to acquaintance and personal relationships between the lecturer and the person evaluated. The lecturer might prepare specifically for the evaluation.
By the lecturers	Self-evaluation of one's own educational initiatives	Allows lecturers to address activities that are typically not evaluated in teaching surveys (e.g., unique use of technology, development of a new and innovative course, development of unique study materials and/or teaching methods)	
By measuring students' learning outcomes	Checking grades, projects, final papers, publications, students' participation in conferences		
By graduates	Retrospective questionnaires	Retrospective view	Expensive and complicated. Difficult to locate former students or to motivate them to rate lecturers in retrospect.

One strategy to assess lecturers' quality of teaching is the use of individual teaching portfolios. A portfolio of one's teaching is a platform for faculty members to group together documents that describe their work. It can include feedback/observations by peers, student ratings and comments, assignments given in class, evaluation of the course materials and corrections made accordingly, as well as varied methods of teaching documentation. A portfolio of one's teaching reflects faculty members' perceptions of the teaching, learning, and evaluation. A portfolio can be a critical factor for devising strong evidence indicating effective teaching. The teaching portfolio, or parts of it, can be used to evaluate the teaching efficacy of faculty members. This is particularly true when those making decisions concerning faculty members have only limited knowledge of teaching in general and of teaching in the faculty member's discipline in particular. For this reason, in the case of faculty members for whom procedures for reaching decisions about tenure or promotion have been initiated, the portfolio also constitutes an important tool for those engaged in faculty development, when working with faculty members at different stages of their professional development.

Furthermore, the use of portfolios as a frame for quality assessment extends the assessment process, transforming it into a process involving all stakeholders – the lecturer, her students, and her superiors. There appears to be a causal relationship between reflecting on teaching and seeking ways of improving it. As researchers, faculty members engage in examination, evaluation, and analysis of research papers in their field of expertise. If faculty members invest time in examining, evaluating, and analyzing their teaching, it is only logical that as a result they will develop a variety of theories on how to improve their teaching methods and students'

learning outcomes. This strategy gives faculty members an opportunity to reflect on their teaching and to emphasize the association between their teaching goals and students' documented learning. Reflection on their work lets faculty members transcend details of everyday teaching and analyze the success of their teaching endeavors. They must provide answers to questions, such as what can be changed in order to improve; does evidence of student learning support claims of teaching efficacy. Preparing a portfolio promotes reflections on their teaching and linking components of the teaching approach with course materials, peer evaluation, and student outcomes. Furthermore, new faculty members can be instructed to gather materials for the portfolio from the beginning of their work as teachers, to reflect on the goals of their teaching, and when they develop courses – to document the pedagogical strategies utilized.

Academic teaching is undergoing deep changes around the world and therefore it is important to properly and fairly examine assessment methods of the quality of teaching in academia. Evaluation by means of portfolios produced by faculty members is only one of these – and it is capable of forming a balance to the voice of students and providing a platform for the voice of lecturers, faculty members, as well. Further research would do well to explore the damage caused by teaching surveys, perceived by faculty members as factors that block their process of promotion, requiring an essential balance for quality assessments of teaching.

### References

- Anderson, K. H., & Siegfried, J. J. (1997) Gender differences in rating the teaching of economics. *Eastern Economic Journal*, 23(3), 347-357. <https://doi.org/10.1057/ej.2008.50>
- Avdor, S. (2006). For whom and for what? Bias factors in evaluating teaching and courses at a college of education and the question of the procedure's contribution to those evaluated and to the organization. *Dapim*, 41, 10-37. [Hebrew]
- Ballantyne, R., Borthwick, J., & Packer, J. (2000). Beyond student evaluation of teaching: identifying and addressing academic staff development. *Assessment & Evaluation in Higher Education*, 25(3), 221-236. <https://doi.org/10.1080/713611430>
- Chandler, T. (1978). The questionable status of student evaluations of teaching. *Teaching of Psychology*, 5(3), 150-152. [https://doi.org/10.1207/s15328023top0503\\_12](https://doi.org/10.1207/s15328023top0503_12)
- 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.
- Davidovitch, N., & Notzer, D. (2004). Students' attendance and scores of teachers' evaluations. *Al Hagova*, 3, 38-41. [Hebrew]
- Davidovitch, N., & Eckhaus, E. (2019). Teaching students to think : Faculty recommendations for teaching evaluations employing automated content analysis. *International Journal of Higher Education*, 8(3), 83-93. <https://doi.org/10.5430/ijhe.v8n3p83>
- Eckhaus, E. (2019). Happiness in fashion. In J. I. Kantola, S. Nazir, and T. Barath (Eds.), *Advances in human factors, business management and society. Proceedings of the AHFE 2018 International Conference on Human Factors, Business Management and Society*: (pp. 15-25), Cham, Switzerland: Springer International Publishing.
- Eckhaus, E., & Ben-Hador, B. (2017). 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. <https://doi.org/10.12973/eu-jer.7.3.445>
- 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>
- Ghasempour, Z., Bakar, M. N., & Jahanshahloo, G. R. (2014). Mix-method design in educational research: Strengths and challenges. *International Journal of Pedagogical Innovations*, 2(02), 83-90.

<https://doi.org/10.12785/ijpi/020204>

- Eckhaus, E. (2019). Happiness in fashion. In J. Kantola, S. Nazir, & T. Barath (Eds.), *Advances in Human Factors, Business Management and Society. AHFE 2018. Advances in Intelligent Systems and Computing* (Vol. 783, pp. 15-25). Cham: Springer International Publishing.
- 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.  
<https://doi.org/10.1023/b:rihe.0000019592.78752.da>
- Hativa, N. (2003a). Are teachers in the humanities and arts better than their colleagues in management and engineering? *Al Hagova*, 2, 38-40. [Hebrew]
- Hativa, N. (2003b). What teachers should know and be able to do - reform in teaching standards. *Hed Hachinuch*, 6, 14-19. [Hebrew]
- Hativa, N. (2008). Myths and facts about evaluation surveys by students. *Al Hagova*, 7, 13-14. [Hebrew]
- Hativa, N., Many, A., & Dayagi, R. (2010). The whys and wherefores of teacher evaluation by their students. *Al Hagova*, 9, 30. [Hebrew]
- Hinz, A., Sander, C., Glaesmer, H., Brähler, E., Zenger, M., Hilbert, A., & Kocalevent, R.-D. (2017). Optimism and pessimism in the general population: psychometric properties of the Life Orientation Test (LOT-R). *International Journal of Clinical and Health Psychology*, 17(2), 161-170.  
<https://doi.org/10.1016/j.ijchp.2017.02.003>
- 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.  
<https://doi.org/10.1080/10705519909540118>
- Kulik, J. A. (2001). Student ratings: validity, utility, and controversy. *New Directions for Institutional Research*, 2001(109), 9-25. <https://doi.org/10.1002/ir.1>
- Marsh, H. W., & Roche, L. A. (1993). The use of students' evaluations of university teaching to improve teaching effectiveness. *American Educational Research Journal*, 30(1), 217-251.  
<https://doi.org/10.3102%2F00028312030001217>
- Mizrachi, Y. (1994). *History of the tension between general liberal higher education and disciplinary specialist higher education and inspecting the option of a balancing curricular program in the regional colleges* (Doctoral dissertation). School of Educational Leadership, Jerusalem. [Hebrew]
- Rinderman, H., & Schofield, N. (2001). Generalizability of multidimensional student ratings of university instruction across courses and teachers. *Research in Higher Education*, 42(4), 377-399.  
<https://doi.org/10.1023/a:1011050724796>
- 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.  
<https://doi.org/10.1080/02619760802000248>
- Tomas, J. R. (2017). The effects of advertising alcohol on young people. *Journal of Social Sciences & Humanities Research*, 3(1), 1-7.
- Wachtel, H.K. (1998) Student evaluation of college teaching effectiveness: a brief review, *Assessment and Evaluation in Higher Education*, 23(2), 191-211.
- Wu, K., Chen, S., & Yuan, Y. (2018, october). *Research on the customer loyalty of bicycle-sharing company based on PLS-SEM model* Paper presented at the 2nd International Conference on Management Engineering, Software Engineering and Service Sciences. <https://doi.org/10.1145/3180374.3181358>

## Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).