KNOWLEDGE BIAS: IS THERE A LINK BETWEEN STUDENTS' FEEDBACK AND THE GRADES THEY EXPECT TO GET FROM THE LECTURERS THEY HAVE EVALUATED? A CASE STUDY OF ISRAELI COLLEGES
--Manuscript Draft--

Manuscript Number:	JKEC-D-13-00007R1
Full Title:	KNOWLEDGE BIAS: IS THERE A LINK BETWEEN STUDENTS' FEEDBACK AND THE GRADES THEY EXPECT TO GET FROM THE LECTURERS THEY HAVE EVALUATED? A CASE STUDY OF ISRAELI COLLEGES
Article Type:	Original Paper
Corresponding Author:	Yoav Gal, Ph.D. Tel – Hai College ISRAEL
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	Tel – Hai College
Corresponding Author's Secondary Institution:	
First Author:	Yoav Gal, Ph.D.
First Author Secondary Information:	
Order of Authors:	Yoav Gal, Ph.D.
	Adiv Gal, Ph.D.
Order of Authors Secondary Information:	
Abstract:	The article addresses two aspects of the same research question: (A) To what extent do feedback questionnaire scores reflect an objective basis for decision making? (B) Do the same mechanisms used in two different institutions produce the same results? The first section examines the question of whether the feedback is fair, while the second one examines whether it is a valid method. The principal findings show that a gender difference exists in the perception of the link between feedback ratings and expectations of personal grades that the students will receive for the course. The second section compared the way in which the administrations of two neighboring colleges with different policies relate to student feedback, resulting in different outcomes.
Response to Reviewers:	Cover Letter to the Editor: Dear Dr. Elias G. Carayannis, Editor-in-Chief, Journal of the Knowledge Economy. Revised manuscript. We have examined carefully the comments and there is no doubt that the reviewer well versed on the subject. We appreciate the reference to the exact details, and revised the article attached below. We believe that the detailed discussion of the comments reinforced the recognition of the importance of the subject. We feel that this is in a very important area and we hope the article is suitable now to get the research into print. We would like to thank the reviewer for the contribution to the discussion. Reviewer's comments: 1.Theory is missing in this article, which needs to be discussed at least in one or two paragraphs. The authors: Amendment Added: In the section of the Introduction and the amended text is highlighted in red. 2.More explanations in detail about the methods used, data collected and results, by giving more figures in this study which is lacking in the article, only two figures have been given.

The authors:
Because of limitations related to the technical aspects of the length of the article, it was not possible to extend the presentation of methods used, data collected and results. However, two more articles are now at the stage of the publishing process and are part of the overall study.

3.Please follow the same reference style throughout the text. Please see the line no. 36, 44 to 51 in page no. 3, which is different from line no. 6 to 8, 12, 13, 15, 17, 20, 22 in page no. 5.

The authors:
In the case of up to two authors, both were named. In the case of three or more authors the first author name specified + et al.. It is a common practice.

4.Please remove his/her words, and instead put any other word, as you have used many times, see line no. 60 in page no. 6 and line no. 1 to 3 in page no. 7.

The authors:
Amendment Added.

5.There is confusion, while in line no. 46 in page no. 5 you have written only economics and management department students of THC were taken for the study. Same way in line no. 11 and 13 in page no. 6 is also same. But in line no. 34 in page no. 8 you have written education along with the department of economics and management students were taken for study. Please check it again.

The authors:
Amendment Added: The sample included students studying in the Department of Economics and Management and those majoring in Education.

6.Please write out of 139 students, how many are from second and third year students, which is missing.

The authors:
Amendment Added: Of these, the number of novices was 63 and the number of seniors was 76.

7.Please explain the Two sample t-Test Assuming Equal Variances in the text given in the page no. 8 and also put the figure no. at the top of it.

The authors:
Amendment Added.

8.Explain what do you mean by Moodle Software given in line no. 29 in page no. 9.

The authors:
Amendment Added: Moodle (abbreviation for Modular Object-Oriented Dynamic Learning Environment) is a free source e-learning software platform, also known as a Learning Management System, or Virtual Learning Environment (VLE). As of October 2012 it had a user base of 70,793 registered and verified sites, serving 63,204,814 users in 6.7+ million courses with 1.2+ million teachers (http://en.wikipedia.org/wiki/Moodle).

9.What do you mean by 'Stations' given in line no. 18 in page no.10.

The authors:
Amendment Added: The term 'stations' was taken from the organizational culture of the college. The correct term is simply 'assignments' and should be treated in accordance.

10.Please show in the figure no.1, the grades for second and third year students, which are missing in figure 1, in page no. 12.

The authors:
Comparative analysis was carried out using Figure 1 between the two institutions. An analysis result by year is beyond the scope of this article.

11.There is a confusion and it is not clear, you have written 'out of 22 statements tested, use of negative consideration were made in 13 cases, and the same numbers of positive reasons were used, [How come same numbers of positive reasons were used, if total 22 statements were tested and 13 cases were negative, then remaining will be 9 positive reasons? See the line no. 49 and 51 in page no. 12.

The authors:
Amendment added and the amended text is highlighted in red: Positive/Negative considerations do not mean Positive/Negative statements. In each of the statements can appear a positive reference, and/or negative reference and/or neutral reference. A statement is an expression of an idea that is not a closed structured response. Therefore, the wording of the statement is free in order to express the perception and the attitudes of students rather than those of the researchers.
In addition, kindly note that feedback questionnaire is a tool for obtaining data from the environment to make decisions. Within the field of marketing in general, and especially in market research, for example, questionnaires are used for a variety of subjects. A

managerial and marketing orientation now permeates many disciplines. Many of the marketing tools utilized are quite old, and some are partly outdated. Nevertheless, college administrations frequently make use of the same tired marketing tools, especially student feedback questionnaires. Typically, solely for convenience, most of the questionnaires are based on closed questions. The questions usually elicit answers on a Likert scale, the most widely used approach to scaling responses. When responding to a Likert questionnaire item, respondents specify their level of agreement or disagreement on a symmetric agree/disagree scale for a series of statements. In this article, the subject of perceptions and attitudes is reviewed and therefore, it is difficult to capture the intensity of the feelings for a given item.

The comparison of the differences in wording of the statements is based on procedures in qualitative research, which is a broad umbrella term encompassing several specific methods and paradigms that rely on the collection, analysis and interpretation of non-statistical data. This is gathered principally through researcher-participant interaction and observation in real life settings. Qualitative research generally aims for depth rather than breadth in description and analysis, with researchers becoming closely acquainted with one particular community or study setting. These methods are particularly useful in accessing the lived day-to-day experience of the relevant population, allowing investigators to intimately explore and understand phenomena from a "subject's" point of view.

*Title Page

KNOWLEDGE BIAS: IS THERE A LINK BETWEEN STUDENTS' FEEDBACK AND THE GRADES THEY EXPECT TO GET FROM THE LECTURERS THEY HAVE EVALUATED?

A CASE STUDY OF ISRAELI COLLEGES

Abstract

The article addresses two aspects of the same research question: (A) To what extent do feedback questionnaire scores reflect an objective basis for decision making? (B) Do the same mechanisms used in two different institutions produce the same results? The first section examines the question of whether the feedback is fair, while the second one examines whether it is a valid method. The principal findings show that a gender difference exists in the perception of the link between feedback ratings and expectations of personal grades that the students will receive for the course. The second section compared the way in which the administrations of two neighboring colleges with different policies relate to student feedback, resulting in different outcomes.

Keywords*: Knowledge-Bias, Feedback, Likert-scale, Students'-grades, Decision-making*

Introduction

Decision-making processes in accelerated knowledge economics have become convoluted and difficult to manage. Feedback processes and external factors play central roles in managerial considerations and decisions at every level of organizations. This problem is also encountered in quintessential knowledge organizations like academic institutions, which also need to make managerial decisions based on data collected from organizational processes. This article shall explicate the direct link between students' ratings of teaching quality and their expectations of the personal grades they will receive in those courses. This article is the sequel to a previous one entitled, Knowledge Bias by Utilizing the Wording on Feedback Questionnaires (Gal and Gal, 2012). A large number of studies dealing with student feedback have been written, and the corpus of studies on the subject continues to expand. Therefore, from the outset, it cannot be argued that this article includes all aspects of the subject of student feedback. Although great efforts were made to find and identify as many sources as possible, in order to cover all of the germane theories, there is a great deal of interest in this field of study, and it is likely that some references will not be mentioned here.

The phenomenon of biased behavior of students is perceived as a world-wide problem (Bar-Yehuda, 2006). It can be found in a range of institutions, from the leading universities to the smallest community colleges, from the United States, through Europe, to the Far East. During the last 30 years, an increase in biased behavior of students has been observed and the most important factor that affects it has been the desire of students to improve their grades (Nath and Lovaglia, 2009).

Students tend to differentiate between intangible ethics and their real behavior. That is, if a student does not identify a situation as one has ethical characteristics, the student might, instead, use another consideration, i.e. relying on social norms or cost/benefit considerations (Eisenberg, 2004). One

reasons for the phenomenon are the deterioration in societal ethics and mores generally; and the competition between students striving for higher grades. Two types of factors in biased behavior of students can be distinguished: one relies on the students' behavior, and the other on the academic institution's culture and the messages it promulgates. The factors that were found to affect it that were associated with the students depended on such variables as: demographic variables; academic ability; psychological, ethnic, social and environmental variables. The factors associated with the institution's messages are characteristics of the institution, like the behavior of the academic staff, the institution's policies and the institution's perception of cheating and enforcement of its rules. Thinking bias in the direction of personal benefit leads a person to erroneous conclusions, according to which an act which is not in fact ethical is perceived as ethically legitimate. A widespread opinion in society holds that 'If everyone can then I can too, since I want to be like everyone else'. It can thus be said that students' values have changed over the years, and the desire to succeed at any cost and by any means has become the most important factor (Sharon et al., 2007).

Two types of motivation can be distinguished amongst students: internal and external. Internal motivation is what inspires study in order to acquire knowledge. External motivation prompts students to learn so as to prove their ability and display their diplomas. One view, therefore, holds that as early as elementary school, students can be distinguished by what motivates them to study. Furthermore, it is external motivation that is linked to dishonest behavior. A student who is motivated by performance and thus by external factors was found to have a higher probability of biased behavior. (Rettinger and Kramer, 2009).

This article addresses the students' perceptions of the relationship between the rating scores they give to the lecturer, and the personal grade (see definition below) that the student will receive later for the course. The motivation for examining the relationship between the ratings given on the feedback forms and the personal grades that the students receive for the course is related to an animated discussion on the subject of the perception of teaching and learning, as well as to the status of the lecturer and paths to advancement in academia. A great deal of criticism has been leveled against using feedback questionnaires to assess teaching, because it misses the principal goal, which is to continually improve the level of teaching. There are those who think that student feedback is not used to improve the quality of teaching, but in fact is used to improve/harm the personal status of the lecturer, particularly the conditions of the employment, salary and advancement. Conditioned by TV reality talent shows in which viewers "vote" for or against a candidate, students may be "voting" for or against a lecturer. Others think that the students' ratings of a lecturer have an effect on the personal grades that the lecturer gives the students, which might be higher than they deserve.

From the lecturers' point of view, evidence based on the literature shows that surveys of teaching constitute a source of anxiety, and might cause a negative counter-reaction to the students by some of the lecturers. These negative attitudes can stem from the lecturers' belief that the students' evaluations

are frequently biased, and are affected by things that are not necessarily germane and to the point, like the grade that the students are hoping to get. These beliefs in turn lead the lecturers to doubt the validity as well as fairness of the surveys, and to worry that the survey will be used against them in the future when the academic institution's administration makes its decisions about their continued employment, and about promotions. On the students' part, studies show that students perceive the process of filling out the feedback as something that is very important and they believe that the feedback reflects the quality of teaching. They also believe that they themselves are in a special position that enables them to evaluate aspects of teaching the course in a way that others, like colleagues or administrative supervisors, cannot. Nevertheless, a fair number of studies indicate that for the most part, the students do not believe that the administration even looks at their feedback questionnaires seriously, or places any importance whatsoever on their opinions.

A huge disparity exists in the way in which lecturers and students relate to feedback, and apparently, these disparities cause tension between the expectations and goals of the two groups. Nevertheless, several questions arise: based on that same tension between the lecturers and students, can the lecturers manipulate the feedback questionnaires to get higher scores from the students? And do the students use their power over the lecturers to receive higher personal grades? The relationship between the course personal grade that the student receives and the ratings on the feedback forms are problematic when they do not correlate with the quality of the student's learning. Therefore, the first part of the article once again addresses the question of whether the students perceive a relationship between the feedback scores and the personal grades they receive in the course. The analysis will be based on the question of gender differences in these perceptions. The study was conducted at Tel-Hai College (THC), located in the extreme north of Israel. Although THC is considered to be a peripheral college, it has an excellent reputation. The study was done as part of an effort undertaken in 2012 to improve the learning process, and was designed to identify potential problems.

The second part of the paper summarizes efforts to improve the learning systems that were conducted during the 2011-2012 academic year in two colleges: THC and Kinneret-College (KC). The two colleges are approximately 50 miles (80 km.) distant from each other, and attract the same target populations of students. They are both considered to be peripheral colleges, but are administered differently. At THC, great emphasis is placed on academic excellence and community involvement, while at KC, the emphasis is on providing opportunities for a broad range of students by setting acceptance requirements low. The different policies taken by the administrations of the two colleges result in different outcomes and different importance placed on student feedback as part of the decision-making process. Both colleges tested new, quite similar systems of learning during the 2011-2012 academic year. At THC, they emphasized excellence, and student feedback was only used as an aid in making decisions. By contrast, KC accorded primary importance to student feedback in making its academic and personnel decisions. The results were that at THC, the experiment continued to the second stage, with adjustments that took into account the students' feedback, while at KC, it was

decided to cease the initiative, since the students 'disliked' the entire process.

Literature review: Feedback questionnaire criticism

The feedback questionnaire constitutes a tool for acquiring data from the environment, to assist in making decisions. Managerial and marketing orientation now permeates many disciplines, even though many marketing tools utilized are quite old, and some are partly outdated. Nevertheless, college administrators frequently make use of the same tired marketing tools, especially student feedback questionnaires (Gal and Gal 2012). Student evaluations of teachers have been around for decades and are the most frequently used type of assessing faculty teaching performance. The information obtained from these questionnaires is used for a number of purposes; most obviously, they provide feedback to the lecturer to improve the teaching. As a result, student evaluations of teachers are the most common form of teacher evaluation used (Winchester and Winchester 2011).

Feedback forms utilizing Likert-type scales have been severely criticized in the literature. Among other reasons, students often believe that feedback will not improve teaching, and evidence exists that their cynicism hardens with experience (Huxham et al. 2008). Also, it was found (Gal and Gal 2012) that the way in which the questions on a feedback questionnaire are worded has a direct impact on the conclusions drawn by decision-makers, and thus on the decisions made. Different formulations of questionnaires yield different results and thus require more in-depth discussion of the validity of using those feedback questionnaires in making decisions regarding the quality of teaching. The use of feedback questionnaires has received criticism from all possible directions (Tennant et al. 2010; Davis 2009; Onwuegbuzie et al. 2007; Parayitam et al. 2007; McPherson 2006; Clayson and Sheffet 2006; Yao and Grady 2005; Felton et al. 2004; Martinson 2004; Dunegan and Hrivnak 2003; Becker and Watts 1999; Becker and Watts 1999; Hounsell 1999; Eiszler 2002; Marks 2000; Scriven 1995; Wetzstein et al. 1984; All as also quoted by Winchester and Winchester 2011). The following are some of the issues receiving attention:

- Feedback questionnaires are often the only measure of competence of a faculty member and are therefore useful for administrators. However, they can be of limited value to the lecturers.

- Since the students providing the feedback have finished the course of study, they cannot experience the positive outcome of these evaluations. As such, they are of limited benefit to the students completing them.

- The feedback questionnaires are designed to meet the needs of the organization rather than those of the lecturer, especially where they are used for performance and/or promotion decisions.

- Student feedback may influence on the difficulty of the course, by leveling down the course requirements, reducing the amount of homework, creating a leniency bias in grading or grade inflation to keep students happy.

- The lecturer's attractiveness or cosmetic factors may have an effect on the ratings given by students.

- Is the course a required or an elective one?

- The concern that students do not take feedback questionnaires seriously and see completing them as a chore.

- Feedback questionnaires may discourage lecturers from challenging students too much, for fear of students' retaliation at the end of the semester.

The widespread use of surveys of student perceptions of their teachers and their learning experiences has prompted an almost equally widespread examination of what these surveys are actually measuring. There are literally thousands of references to research on student ratings of teaching, most of them from the last two decades. The issue of the validity of student evaluations of teaching is a matter of some controversy, with evidence both supporting their continued use and evidence recommending their discontinuation. The controversy is unlikely to abate in the near future. It was found that cultural background significantly affects student evaluations. Other factors that have an influence on the average score include: year level; enrolment size; the quantitative nature of the subject; the gender of the student; fee-paying status by gender; course of study; the differences between the course mark and previous marks; the quality of workbooks; the quality of textbooks; and the score relative to those in other subjects taught at the same time. In addition, average scores for instructors who have taught in a mix of subjects are similar to those based on scores adjusted to account for subject and student characteristics (Davies et al. 2007, Davies et al. 2006). Student surveys have been used in administrative decision-making; for summative and formative purposes; as a basis for hiring, confirmation and promotion; or in matching teachers to appropriate courses. Surveys have also been conducted to fulfill the demands for the accountability of the institution and so that cross-institutional comparisons may be made. In some institutions, the results of such surveys have been made available to students, to inform their course selections. Deep controversy has arisen over how appropriate such use of rating data is for these purposes.

Several studies have confirmed the validity and reliability of student ratings in the U.S. (Bosshardt and Watts, 2001; Wachtel, 1998); in Australia (Marsh, 1987); in Hong Kong (Ting, 2000); and in Europe (Byrne and Flood 2003). Nevertheless, a great deal of debate is still ongoing about the usefulness of these ratings. Some argue that the precise numerical scores generated in student evaluations imply a level of measurement that simply does not exist (Wolfer & Johnson 2003, p. 117) and that it would be better if numerical scores were replaced with categories (d'Apollonia & Abrami, 1997), more global measures of assessment (Cashin & Downey, 1992) or a range of scores (Neumann, 2000). Others suggest that the fine measurements generated preclude differentiating

between good and bad teaching, and that utilizing cutoff points would be a more sensible approach (McKeachie, 1997) Still others claim that the use of student evaluations for administrative purposes is a misapplication of the data (Bedggood & Pollard, 1999; Sheehan, 1975).

With such criticism leveled against this tool, the question arises: if it was medicine would we use it? We all know the power of industry and the impact of big money on the introduction of tobacco products, dairy products, alcohol, etc. Thus, an important question is: What is the force behind the perpetuation of the use of feedback questionnaires, despite the harsh criticism of the method? Answering this interesting question is not within the purview of this article, but further examination of it is recommended. As noted above, this article addresses another variable regarding feedback questionnaires, by examining two aspects of the same research question, from two different angles: (A) To what extent do feedback questionnaire scores reflect an objective basis for decision-making? (B) Does the same mechanism used in two different institutions produce the same results? The first section examines the question of whether the feedback is fair, while the second one examines the question of whether it is a valid method.

Student evaluations of teachers by gender of students

The question examined is whether, in line with the students' perception, a link exists between the students' feedback scores on the questionnaires they are asked to complete and the final personal grades they receive for the course, amongst students of THC. The sample included students studying in the Department of Economics and Management and those majoring in Education. Despite the fact that this subject has been studied a number of times, the interest here derives from the desire to understand the source of the possible knowledge bias in the decision-making process regarding the evaluation of teaching.

The hypothesis was that a gender difference does, in fact, exist concerning the students' perception of the link between the rating and evaluation that the student gives the lecturer on the feedback questionnaire, and the grade that the student will receive for the course. Moreover, not only will a gender difference be found in the students' perception of the link between their feedback forms and the grades they will receive, but we hypothesize that the surety and smugness of male students about their opinions compared to those of female students will manifest vis-à-vis this link as well. That is, the male students will tend to believe that a high score on their feedback form will lead the lecturer to give them a higher grade for the course, as an ameliorating factor along with grades on exams and papers. This hypothesis relies on the assumption that 'accepted knowledge' exists, which is passed on to each succeeding cohort of students as part of the cultural tradition of the college, and enables students to identify lecturers who are 'easy graders'. The moment that the informal relationship of 'If you'll be OK with us, we'll be OK with you' is made, an unwritten agreement is created, that maintains the rules, i.e. a high feedback evaluation equals high student grades. Because most of the students majoring in Economics and Management are male, their opinions will be of greater importance and a *de facto* knowledge bias will result.

Definition of research variables and various reasons for deviation

The data for the purposes of this part of the study were collected from a sampling of students at THC. The emphasis was on an opinion survey that was meant to identify students' perceptions of the subject under study. The policies of THC's administration were not examined, nor were the considerations of lecturers, and no quantitative analysis of students' grades was done. The goal was to identify the students' opinion about the subject, in a way that would enable expression of the effect of existing knowledge on deviation. In fact, for the purposes of the study, it is unimportant whether the lecturer gave or did not give higher grades to students in the hopes that the students would give the lecturer a higher rating on their feedback questionnaires. Whether the administration of THC used or did not use the students' feedback in making its decision about promoting a lecturer is likewise unimportant. The only issue that is examined here is the reality of the students' perception concerning the subject. If, in the students' perception, a link exists between feedback ratings and personal grades, regardless of reality, then their decision-making process when completing the feedback questionnaires must be expressed. The variables examined are explicated below:

🔴 Reliability:

Can a student evaluate the quality of a lecturer's teaching appropriately? In accordance with the students' *weltanschauung* in the sample examined, they can and are entitled to evaluate their lecturers' teaching. The widespread argument was that if the student is the one who has to listen to the lecturer throughout the whole semester, she/he has the ability to say whether the lecturer/course contributed to her/his level of knowledge. In general, it can be said that students see themselves as mature adults with high levels of skills and knowledge, who can make considered decisions in their evaluation of the learning process.

🔴 Personal Grade:

This variable is defined in this study as the student's subjective estimation of the grade he/she will get for the course. The student determines this expected grade in accordance with the personal assessment of the level of learning and the investment of time and effort in the course, attendance, comprehensive learning including use of other sources, the interest in the subject, and the motivation.

🔴 Ratings given by students of the lecturer's teaching:

As with the variable of grade, this variable, too, is entirely subjective. The rating relies on the student's perception of teaching quality, and is determined solely by the student. No substantiation is required, neither quantitative nor according to the various items [on the forms]. In the event that the student determined that the lecturer was terrible/excellent, that

was sufficient to accept the rating as valid. The reason for this is that the aim of the study is to identify the factors contributing to deviation in the students' perceptions.

The question of whether the lecturer's gender is a factor in the students' ratings was not examined as part of this study. That is because it is assumed that during regular courses, there is no personal interaction between the lecturer and the student. Furthermore, a variety of reasons exist for bias in the answers. Comparisons are needed between various courses of different levels of difficulty, wherein differing workloads are assigned by different lecturers; whether the lecturer is a low-level staff member or a member of the senior faculty; differences in the number of students in each course; differing buildings, classrooms, and learning conditions; the number of weekly lecture hours; and all the ramifications of the foregoing. The reason for focusing the study derives from the assumption that if a knowledge bias exists within the more limited parameters, it would certainly expand in more comprehensive arrays.

Research process

The data was collected from a sample population, with emphasis on a few aspects:

- Characteristics

  The characteristics of the student providing the feedback were: gender, academic year, course of study and major. The students' names, I.D. numbers and other potentially identifying details were not collected, so as to reduce the opportunity to give biased answers.

- The heart of the study:

  The questions that were given to the students were constructed in a matrix. Some of them were worded positively and some negatively, so as to ensure reliability and consistency of the answers, and thus filter bias and operate controls in an unseen way. The questions in this part were closed questions, utilizing a Likert scale with four levels, thereby precluding the participant from avoiding expressing an opinion or expressing an intermediate one. Thus, the respondent was required to take a position in a certain direction. No emphasis was placed on the link between ratings and grades or on any other link; rather, it was characterized as an opinion survey. That is, the object of the study was not stated explicitly, although that could have been inferred generally from the questions themselves. Some of the questions were constructed in a general manner to the group of students as a whole, and others required individual responses. The aim of utilizing this structure was to reduce the creation of formulaic answers customarily deriving from stereotypes and prejudices.

● Open section

This section was meant to enable the participating student to express him/herself freely and raise new insights. Thus, subjects that were not covered in the closed section of the questionnaires could be addressed; the aim was to enable the student to add areas of knowledge that were not directly discussed.

Sample population

A total of 139 students from THC participated in this part of the study, of these, the number of novices was 63 and the number of seniors was 76. In preparation, a pilot study was conducted on 20 students who study at other colleges, to determine whether the instrument was understandable. The results of the pilot study were not included in the sample, since the process was only meant for oversight and corrections, and not for the purpose of drawing conclusions.

Results

Analysis of the closed questions was done by utilizing a t-test assuming unequal variances [$P(F<=f)$ = 0.02497, $\alpha<=0.05$], which examines the differences between two independent groups. The sample included students studying in the Department of Economics and Management and those majoring in Education. Analysis of the section of open questions relied on repeated cross expressions, use of words that denote power, and a summary of the significance of gender.

In dividing the sample by gender, a difference in the answers given by men and women was evident. It was found that men believe that there is a link between the ratings given to a lecturer on the feedback forms and the grades they expect for the course, to a greater extent than do women. This supported the hypothesis that states that gender is a factor influencing the results.

Table 1: Two-Sample t-Test Assuming Unequal Variances

	Men	Women
Mean	2.8	2.4
Variance	0.633	0.394
Observations	57	82
df	102	
t Stat	2.9711	
P(T<=t) one-tail	0.0019	
t Critical one-tail	1.6599	
P(T<=t) two-tail	0.0037	
t Critical two-tail	1.9835	

Thus, the hypothesis that a gender difference exists regarding the perception of a link between the rating given to a lecturer on the feedback questionnaires and the grade the student expect was supported. That is an especially interesting point, which constitutes another source of knowledge bias;

when no importance is attributed to gender, the un-weighted results are absorbed into averages.

Comparison of policies between two neighboring colleges

As noted above, during the 2011-2012 academic year, a controlled experiment was conducted, to introduce new learning systems to two neighboring colleges. THC and KC are located not far from each other and attract students from the same target population. The new systems were designed to solve a number of problems while enhancing the learning process:

- Solve the problem of copying on exams.

- Provide solutions for students with various types of learning disabilities.

- Change the focus of studying, from 'teaching to tests' to a continuous learning process throughout the semester.

Without coordinating between them, the two colleges independently chose to utilize the Moodle software, and likewise independently developed their processes of improving the quality of teaching using the range of possibilities that the Moodle program provides. Moodle (abbreviation for Modular Object-Oriented Dynamic Learning Environment) is a free source e-learning software platform, also known as a Learning Management System, or Virtual Learning Environment (VLE). As of October 2012 it had a user base of 70,793 registered and verified sites, serving 63,204,814 users in 6.7+ million courses with 1.2+ million teachers (http://en.wikipedia.org/wiki/Moodle). The opportunity offered by the introduction of the new systems was taken advantage of to conduct a pilot study that examined the changes in methodologies. The new system integrated class lessons with on-line forum discussions involving the community of students taking the course. As part of the experiment, six courses utilizing the new system of study were given at THC, and three at KC.

The use of technology raises questions regarding the learning expected to occur in the courses that used that methodology. What happens to students who are not self-motivated? How can subjects like advanced theories or professions related to them be taught using directed programs? How can academic standards be maintained? What will happen to the staff of teachers and lecturers? And those are only a few of the many possible questions raised. The professional literature is replete with articles addressing the trends that are likely to occur due to the greatly enhanced accessibility to high-quality academic contents. All of these trends are well known to anyone who deals with the subject in other fields as well. In this article, the new system and its processes will not be reviewed, nor will the rich literature dealing with the subject. What is interesting in the present discussion is that in both of the colleges that are apparently similar, similar experiments were conducted over a full academic year, with similar technological tools, in similar academic structures. Nevertheless, the outcomes were totally different. Furthermore, the material difference in the results primarily derives from the effect of student feedback.

Below is a summary of the important results from the comparison of the application of the new study

systems in the THC and KC.

A. At KC, by contrast with THC, the existence of two groups of students, a weaker and a stronger group, in the same course, is clearly obvious. At KC, the group of students who were unsuccessful in the course in 2012 was similar to the number of students who failed in the previous academic year. By contrast, a clear change in grades was obvious at THC, where the grades were higher as a result of the change in the teaching and learning system. Every group of students increased their grades; those showing the most improvement were the weaker students.

B. The change at KC, like the one at THC, was structural. The grades were determined at seven different assignments during the semester, with an increasing number of points given for each new assignment. The first assignment was worth ten points, and the last one was worth 30 points of the final grade. Thus, during the course of the semester, it was possible to acquire 100 points for the final grade. The new system thereby enables the students to improve their achievements if they make efforts throughout the semester, and not only on the midterm and final exams, as usually occurs. The new teaching system was accompanied by on-going discussions utilizing the Moodle program, which enabled the weaker students to participate continuously throughout the semester, raise questions, clarify points about which there were disagreements, learn and bring knowledge from other sources into the discussion.

C. At THC, the change in the system benefited the entire group of students, while at KC, the new system contributed value only to the group of stronger students. The new system significantly limited disciplinary issues, criticism about the validity and fairness of the exam, and in general created a positive learning atmosphere amongst the students who actively participated in the process. At the same time, this was not expressed in the students' formal feedback forms. That was because there was still a large group of weak students who compared the demands of the new system to other 'easier' courses. At KC, student feedback has a decisive impact on the administration's decisions and on the lecturers of other courses in an obvious way. Less obviously, the lecturers made the courses easier in order to forestall negative criticism from the students in their feedback

D. The relative groups of excellent students were similar in size and scope at the two colleges. The sub-group of students getting personal grades above 70 and above 80 at THC was similar in size for second- and third-year students.

E. Significant differences were also evident in the second-year students at THC compared with the second-year students at KC who had high grades, i.e. higher than 80, apparently due to differences in policies at the two colleges. While the findings obtained at KC support the existence of two different groups distributed around two different statistical poles, good and

poor, the distribution at THC displays a more normal Bell curve and shows bias toward higher grades. A comparison of the two colleges shows that more than 60% of the second-year students at THC are in the group that received grades of 71 to 89, while only 36% of second-year students at KC are in the group that received grades of 71 to 89.

F. Another interesting point of comparison is that, while at KC some 40% of the students in the course received grades below 70, there was a statistically significant difference in the number of students at THC who received such low grades. Only seven percent of second-year students at THC received grades below 70, and 18% of third-year students did. As noted, that is a manifestation of the different policies of the two colleges concerning the students and the demands made on them.

G. The great concern that the new system, which requires greater efforts from the students, will add to the burdens on the weaker students and drag them down has been proven false in THC. The exact opposite occurred: the entire cohort at THC improved their grades. That did not occur at KC, apparently because of the existence of two groups in the class and lack of structural homogeneity, as opposed to the situation at THC. This in turns shows the great importance that THC places on processes designed to improve the level of teaching and learning, so that it not only does not harm the sub-group of weaker students, but in fact helps them to improve their grades.

The following illustration presents a comparison of the two colleges vis-à-vis the new system introduced at both colleges during the 2011-2012 academic year.

Figure 1: Comparison of the Grades Received by the Students at TCH and KC
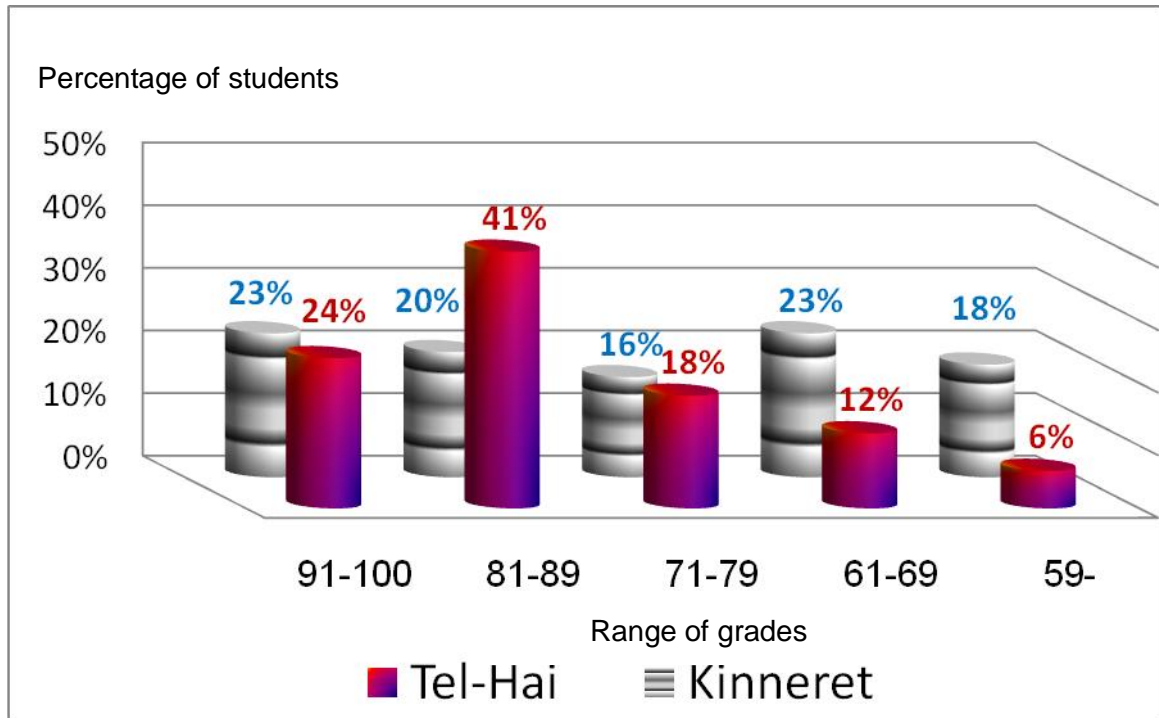
During the discussions in the on-line forum, which were open to all of those involved in the process, the statements below were the principal ones raised by the students concerning the new system. Those in italics are the central considerations in each statement that uses a negative reason, and those in bold use positive reasons, of course, according to the perceptions of the students. Positive/Negative considerations do not mean Positive/Negative statements. In each of the statements can appear a positive reference, and/or negative reference and/or neutral reference. Therefore, the wording of the statement is free in order to express the perception and the attitudes of students rather than those of the researchers. Attention should be paid to the fact that the italics / bold fonts do not signify pro or con the new system, but the type of substantiation and reasons given for the opinion. There are various dimensions to the discussion here: pro/con; considerations that hamper or considerations that support; use of positive/negative terms. This is a process of mapping the values of the students' perceptions, by reducing them to two dimensions, so as to draw conclusions from a wealth of data. An example of the statements are set forth below:

- I feel that the system contributes to my understanding the material better, than if I were to *study for the exam and forget* the material a few days later.

- We learned **how to study and how to think** in the course.

In the 22 statements tested, uses of negative considerations were made in 13 cases, and the same

numbers of positive reasons were used. A possible bias toward the positive could have been expected because of the fact that the respondent was identified by name, which was expected to result in flattery. A possible bias toward the negative could have been expected because it was an opportunity for the student to really make mincemeat out of the whole system, and could have been done without any consequences. Thus, a balanced system enables entering into the discussion with a certain degree of confidence that the statements reflect the true opinions of the students and not what seems politic only.

When only the key words that are in italics are examined, the following map of the statements results:

- Negative references to efforts: 4.

- Negative references to studying for the exam: 3.

- Negative references to the subject of copying: 2.

That is, the principal bothersome factors are the investment of time required by the new system, which is much greater, to a statistically significant extent, than the 'regular' amount of time invested [in other courses]. The fact is, learning for the exam is short-term learning only, and there is the issue of copying on the exam as well. A possible explanation can be postulated from what was stated, although it is not the only one: a 'vicious cycle' that stabilizes itself exists. The third-year students attribute a great deal of importance to the grades they receive in that year as a lever to raise their grade averages from the required courses they took in their first and second years. Thus, the mechanism of the 'regular' final exam enables the investment of a short concentrated burst of effort, aided by copying, in order to achieve the goal of a high grade obtained at any price, so as to improve one's personal grade average.

By contrast, when only the bold words are examined, for the most part, they refer to the positive contribution of the learning process. That is, if the pressure of the third year to improve one's grade average at any price is 'removed', the student feels that he/she benefited from the learning process, and didn't flinch from the much greater efforts and investment of time required. The determining factor is apparently the cost/benefit relationship in the goal of getting a better grade.

When the written opinions are added to complete interviews, the significance is:

A. At both THC and KC, the more the good student acquires experience and the further along in the academic career the student is, the higher the grades received, resulting in a concomitantly higher sense of inner value and increased level of satisfaction, as expressed in the feedback questionnaires.

B. At both colleges, what arises from the written comments on the feedback forms, and from the on-line forum discussions, is that the evaluation of the new system is made almost entirely in

comparison to other courses, and not of the course itself. The underlying claim that is repeated both on the feedback forms and in the on-line forums is that the course is harder than other courses, because the students have to make greater efforts to get high grades throughout the entire semester. As noted, in many other courses, the only effort required from the student is a relatively small one, that of studying for the final exam, without really having to invest in the process of on-going, long-term learning throughout the semester.

C. At THC, in a clear and statistically significant way, there are no disciplinary problems when using the new system, the problem of copying does not arise, and the academic level is high. At KC, despite the requirement that students attend classes, most of the weak students were absent from most of the classes.

D. At THC, no difficulties whatsoever exist for students with learning disabilities. The new system enables each student to carry out the assignments at the level of ability and character. At KC, because most of the students who are defined as weak were absent from classes, it was impossible to follow up on the impact of the new system on them.

What can be concluded?

Advantages

A. At THC, there were zero disciplinary problems. There were no discussions of any disciplinary committee, no question of proctoring exams, no problems with copying. There's no problem of 'sweeping the problem under the rug', since the problem does not exist at all. At KC, there was an unceasing fight with weak students, who refused to meet academic demands, and who received covert support from the academic secretary for their refusal.

B. The learning process continues at a high level throughout the entire semester, and not only in preparation for the final exam. The students are required to engage each other in the process of acquiring knowledge, and emphasis is on the process of acquiring knowledge, rather than preparing for a test. The result is that positive learning takes place in teams, motivated by internal and positive enjoyment, and not under the 'threat of the exam'. In both colleges, for the most part, the good students benefited from the new system. The difference between the two colleges was that in THC, the weaker students were given a boost upward, while at KC, with the active support of the academic secretary, the weak students waged a continuous battle against the new system, maintaining their rights to be absent from classes and avoid meeting the basic academic requirements. They utilized a mechanism of heavy pressure by writing letters of complaint and giving terrible feedback.

Difficulties:

A. In both colleges, it is clear that the students' feedback rely on 'relative value', that is, the ratings the students give about the course are strongly influenced by the way other courses are run.

B. The results from THC as of now show that the greatest beneficiaries of the new system are precisely the weaker students. The excellent students remained good students, but in general, there was an upward trend. The principal reason for this was the continuous discussions throughout the semester and the division of assignments into small ones that required the students to invest in the learning process instead of memorizing the material for the exam. By contrast, at KC, the group of weak students invested most of their efforts throughout the semester in organizational politics and complaints to the academic secretary about the 'unreasonable' difficulties of the new system. Although the grades of the excellent students in both colleges are similar, in THC the weaker students were 'dragged' up. That contrasts with the case at KC, in which the grades of the entire group of weak students were dragged down. They continually engaged the academic administration, which had to deal with their complaints and criticisms about the lecturer and the new system. These complaints were also expressed very strongly in the students' feedback at the end of the semester.

Discussion

The need to address the subject derives from two principle factors. First is the issue of knowledge bias in feedback questionnaires. The second is the existence of a broad, comprehensive academic conflict over the ability to introduce changes in the learning system without fear of political repercussions resulting from possible negative student feedback. The first part of this article reviewed a sample of students at THC regarding their perception of whether a link exists between the ratings they give a lecturer on their feedback forms and their grade expectations for the course. The findings show that the students in fact do think that such a link exists.

The second part of the article presents a comparison of an attempt to introduce a new learning system in two neighboring colleges. At KC, the group of weak students sowed an atmosphere of virtual terror through letters of complaint and extremely negative feedback, which led the academic administration to end the new system. That is, not only does a link exist between the ratings that the students give on their feedback and the grades they get, in accordance with the students' perception, they were able to manipulate their feedback and impose their will to prevent situations undesirable to them in academic improvement processes. Clearly, if the lecturers identified these mechanisms as threatening their source of income, although they would do everything they could to give the impression that they were fulfilling their academic roles, they would also do everything they could to prevent the students from 'going negative' against the lecturers.

Gender differences in the students' perception can be another key factor that explains the disparity between the various surveys on one hand, and the feeling of many lecturers on the other, including wide-ranging criticism of student feedback in the literature. However, this subject also needs a more

precise and broad examination, that is beyond the scope of this article. What is noteworthy is that, because more men than women usually major in Economics and Management, which can certainly manifest in the feedback forms; questionnaires that gather feedback from a similar number of male and female students can obscure this point. As noted, the study conducted supported the hypothesis that the students are of the opinion that a link exists between the ratings they give the lecturers on the feedback forms and the grades that they will receive for the course. That held true for the entire sample of students, as well as for division by gender. Quite possibly, the averages method used by other studies creates difficulties in identifying more precise patterns, and repeat studies are required. Moreover, low grades in a course do not necessarily stem only from the lecturer's ability to teach, but are affected by numerous other parameters, depending on the students, their personalities, and other factors.

The policies of KC is to provide as many students from the peripheral areas of Israel as possible with access to higher education. That statement creates difficulties, especially because the college's catchment area is larger than the distance to the closest metropolitan area. Furthermore, a declaration like this one creates an entire edifice of expectations by the students, who come to KC expecting that demands on them will be considerably lower than would be the case in other academic institutions. Such a policy isn't just left hanging, and the students do not ignore it. They absorb the message and function accordingly. As shown above, there is a group of good students, but there is also a group of very weak students that is no less important. The latter act energetically to lower the quality of teaching, reduce the demands made on them, make things easier for themselves and contemn efforts to improve academic processes. The end of the process at KC was the decision made by the academic administration that 'the satisfaction of the clients' is the foremost consideration, taking precedence over everything else. The complaints and criticisms of weaker students can damage the college's marketing ability, and thus it was decided to completely stop the new system.

At THC, by contrast, despite the many similarities in conditions, the policy is totally different. The improvement processes mentioned above were instituted along with direct oversight by the Dean of Faculty of the School of Humanities and the Faculty Committee on Excellence in Teaching. This latter committee carefully supervised the new system, formulating questions about the process and demanding that the courses involved in the new program meet appropriate academic standards. The Committee took the position that the format in and of itself was not the primary factor in creating learning. Rather, the type of challenges that the lecturers set for the students were what created learning. Therefore, the Committee wrote in the special report it presented to the college's administration, that the administration had to be tolerant of lower feedback ratings on some of the courses, in the hope that the evaluations would improve in the future as the program was stabilized and became more widely accepted. That is, with all due respect to the importance that the college's administration places on the students' feedback, the message is that academic excellence is not determined by popularity contests, but really reflects the fact that the accepted standards of

academic excellence are met.

The fundamental difference between the decision-making processes in KC and THC is that at KC, the administration insists on acting on popularity, as expressed in the ratings given by students on their feedback forms, which do not meet any academic test. By contrast, at THC, not only did the entire process receive backing from the moment it was decided to embark on the project, but it was tested according to measurements of improved teaching and not of 'client satisfaction'. When a process of pedagogic improvement is undertaken, and the most important factor in the decisions about it do not meet academic standards and are not objective, teaching may not be improved, and the whole process may be jeopardized. This has been well proven by the policies of the academic secretary of KC.

There are those who think that the problem is rooted in the fact that we live in a society in which status is very important, and that our attitude to others is determined according to that status. From this, a system of supposed mutual interests develops between the student and the lecturer, related to positions of power and ability to impose control. Accordingly, there is reason to believe that student feedback is not utilized as a means to improve the quality of teaching, but in fact serves as a means to improve the personal status of the lecturer and improve her/his employment conditions, i.e. promotion. Perhaps even worse is the excessive leniency used in grading the students who can be the ones to make future decisions in the fields of economics, society and government. On the other hand, the question arises regarding students utilizing their feedback as punishment of the lecturer. However, use of feedback surveys is continually spreading, despite the broadly based academic corpus that refutes their efficacy. In an up-to-date article that was published recently (Darling-Hammond et al., 2011), these deductions are validated:

> 'There is a widespread consensus among practitioners, researchers, and policy makers that current teacher evaluation systems do little to help teachers improve or to support personnel decision making.'

Clearly, the subject that needs to be raised in a serious discussion is not student feedback. Rather, such a discussion must focus on proven improvement in the level of academic teaching; on proven academic achievements; and on continuous improvement in the learning processes.

Summary and conclusions

The aim of this study was to attempt to understand in depth the viewpoint of students of the THC regarding the level of the link between student feedback about their lecturers' teaching and the grades that they expect to receive in those courses. Student feedback is a type of market study, in which a great deal of money is invested in order to reach the right operative decisions for the future, in everything concerning the quality of teaching. Despite that, apparently, student feedback is tainted with structural distortions and so a great deal of money is wasted. It is important to note that the students, as evaluators of teachers, are unaware of the various factors in deviation, and thus the control processes are carried out on levels other than decision-making processes regarding teaching.

The principal findings of this study show that the students perceive that a link exists between the ratings they give teachers on their feedback, and their expectations of the grades that they will receive for the course. This link is even expressed in a gender difference, wherein male students tend to think that this link is stronger than do female students. Moreover, not only do the students think that the link exists, they utilize the mechanism of the feedback as a lever to affect the academic decision-making of the institution of higher education. That was shown to be the case at KC. During the 2011-2012 academic year, a group of weaker students at KC expressed their opposition to new learning processes that were introduced, which required their active participation. Their opposition, in fact, brought about the cessation of the new processes. By contrast, at THC, which also utilized student feedback, the decision-making process there emphasized excellence. As a result, even the students' feedback change accordingly.

The principal significance that arises from the aforesaid, and from the corpus of articles that deal with criticism of the process of student feedback, is that the way in which feedback is obtained today, i.e. via feedback forms that rate the lecturers, must be fundamentally and materially changed. Student rating of lecturers functions, in large degree, as political ammunition, and not as a means of review so as to improve the quality of teaching. Proof of this conclusion is that the students see the feedback they give as something that can be manipulated, which they in fact do, so as to create more comfortable conditions for themselves, rather than improving the teaching quality. The moment that a tool of review becomes distorted and biased, use of the data obtained from it is also biased and unreliable, and leads to erroneous decision-making processes.

References

Bar-Yehuda, R. (2006), "The Purity of Examinations: Discussion in the Political Council for

Computer Science" (in Hebrew).

Becker W. E. and Watts M. (1999), "The state of economic education: How departments of

economics evaluate teaching", American Economic, 89:2, pp. 344-9.

Bedggood R. E. and Pollard R. J. (1999), "Uses and misuses of student opinion surveys in eight

Australian universities", Australian Journal of Education, 43:2, pp. 129-156.

Bosshardt W. and Watts M. (2001), "Comparing student and instructor evaluations of teaching",

Journal of Economic Education, 32:1, pp. 3-17.

Boysen G. A. (2008), "Revenge and student evaluations of teaching", Teaching of Psychology, 35, pp.

218-22.

Byrne M. and Flood B. (2003), "Assessing the teaching quality of accounting programs: An

evaluation of the course experience questionnaire", Assessment & Evaluation in Higher

Education, 28:2, pp. 135-145.

Cashin W. E. and Downey R. G. (1992), "Using global student ratings items for summative

evaluation", Journal of Educational Psychology, 84:4, pp. 563-572.

Clayson D. E. and Sheffet M. J. (2006), "Personality and the student evaluation of teaching", Journal

of Marketing Education, 28:2, pp. 149-60.

d'Apollonia S. and Abrami P. C. (1997), "Navigating student ratings of instruction", American

Psychologist, 52:11, pp. 1198-1208.

Davis B. G. (2009), "*Tools for Teaching*", San Francisco, CA: Jossey-Bass.

Davies M., Hirschberg J., Lye J., Johnston C. and McDonald I. (2007), "Systematic influences on

teaching evaluations: The case for caution", Australian Economic Papers, 46:1, pp. 18-38.

Davies W. M., Hirschberg J., Lye J. Johnston C. and McDonald I. (2006),"What Influences Teaching

Evaluations? Evidence from a Major Australian University", The Business Review,

Cambridge, 6:1, pp. 146-152.

Darling-Hammond L., Beardsley A., Haertel E. and Rothstein J. (2011), "Evaluating Teacher

Evaluation: What We Know about Value-Added Models and Other Methods", Stanford University School of Education, SUSE Open Archive, a copy of the peer-reviewed journal article at Phi Delta Kappan.

Dunegan K. J., and Hrivnak M. W. (2003), "Characteristics of mindless teaching evaluations and the moderating effects of image compatibility", Journal of Management Education, 27:3 pp. 280-303.

Eisenberg J., (2004), "To cheat or not to cheat: effects of moral perspective and situational variables on students' attitudes", Journal of moral education, 33:2, pp. 163-178

Eiszler C. F. (2002), "College students' evaluations of teaching and grade inflation", Research in Higher Education, 43:4, pp. 483-501.

Felton J., Mitchell J. and Stinso M. (2004), "Web-based student evaluations of professors: The relations between perceived quality, easiness and sexiness", Assessment & Evaluation in Higher Education, 29:1, pp. 91-108.

Gal Y. and Gal A. (2012), "Knowledge bias by utilizing the wording on feedback questionnaires: A case study of an Israeli college", Journal of Mixed Methods Research, Submitted.

Hounsell D. (1999), "The evaluation of teaching", In *A Handbook for Teaching and Learning in Higher Education: Enhancing Academic Practice,* H. Fry, S. Ketteridge and S. Marshall (ed.) 200-12. London: Kogan Page.

Huxhama M., Laybourna P., Cairncrossa S., Graya M., Browna N., Goldfincha J., Earl S. (2008), "Collecting student feedback: a comparison of questionnaires and other methods". Assessment & Evaluation in Higher Education, 33:6, December, pp. 675-686.

McPherson M. A. (2006), "Determinants of how students evaluate teachers". Journal of Economic Education, 37:1, pp. 3-20.

Marks R. B. (2000), "Determinants of student evaluations of global measures of instructor and course value", Journal of Marketing Education, 22:2, pp. 108-19.

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, pp. 253-388.

Martinson D. L. (2004), "A perhaps 'politically incorrect' solution to the very real problem of grade inflation", College Teaching, 52:2, pp. 47-51.

McPherson M. A. (2006), "Determinants of how students evaluate teachers", Journal of Economic Education, 37:1, pp. 3-20.

McKeachie W. J. (1997), "Student ratings: The validity of use" American Psychologist, 52:11, pp. 1218-1225.

Nath L. and Lovaglia M. (2009), "Cheating on Multiplechoice Exams: Monitoring", Assessment and an Optional Assignment. College Teaching, 57:1, pp. 3-8.

Neumann R. (2000), "Communicating student evaluation of teaching results: Rating interpretation guides (rigs)", Assessment & Evaluation in Higher Education, 25:2, pp. 121-134.

Onwuegbuzie A. J., Witcher A. E., Collins K. M. T. and Filler J.D. (2007), "Students' perceptions of characteristics of effective college teachers", American Educational Research Journal, 44:1, pp. 113-60.

Parayitam S., Desai K., and Phelps L. (2007), "The effect of teacher communication and course content on student satisfaction and effectiveness", Academy of Educational Leadership Journal, 11:3, pp. 91-105.

Rettinger D. A. and Kramer Y. (2009), "Situational and personal causes of student cheating. Research in Higher Education", 50:3, pp. 293-313.

Scriven M. (1995), "Students offer useful input to teacher evaluations", Practical Assessment, Research & Evaluation, 4:7, pp. 1-4.

Sharon, D., Bialik, H., Klishak, S., Yofi, L. H., Luz, S., Lior B., Tabak, N., (2007), "Promoting academic honesty and professionalism by means building and instilling a code of ethics for academic honesty in the Schoenbrun School of Nursing", Tel Aviv, Medicine and Law, 8:37 (in Hebrew).

Sheehan D. S. (1975), "On the invalidity of student ratings for administrative personnel decisions", Journal of Higher Education, 46:6, pp. 687-700.

Tennant M., Mcmullen C. and Kaczynski D. (2010), "Teaching, Learning and Research in Higher Education: A Critical Approach", Oxon: Routledge.

Ting K. (2000), "A multi-level perspective on student ratings of instruction", Research in Higher Education, 41:5, pp. 637-661.

Wachtel H. (1998), "Student evaluation of college teaching effectiveness: A brief review", Assessment & Evaluation in Higher Education, 23:2, pp. 191-211.

Wetzstein M. E., Broder J. M. and Wilson G. (1984), "Bayesian inference and student evaluations of teachers and courses", Journal of Economic Education, 14:1, pp. 40-45.

Winchester M. K., and Winchester T. M. (2011), "If you build it will they come? Exploring the student perspective of weekly student evaluations of teaching", Assessment & Evaluation in Higher Education, pp. 1-12.

Wolfer T. A. and Johnson M. M. (2003), "Re-evaluating student evaluation of teaching: The teaching evaluation form", Journal of Social Work Education, 39:1, pp. 111-121.

Yao Y. and Grady M. (2005), "How does faculty make formative use of student evaluation feedback? A multiple case study", Journal of Personnel Evaluation in Education, 18:2, pp. 107-26.