Students' evaluation of Tutors in distance education: A Quasi -longitudinal study

Sotiris Kotsiantis[1,3], Nikolaos Tselios[2,3*] , Michalis Xenos[3]

[1]Department of Mathematics, University of Patras, University Campus, 26500 Rio, Patras, Greece

[2]Department of Educational Sciences and Early Childhood Education, University of Patras, University Campus, 26500 Rio, Patras, Greece

[3]Hellenic Open University, Parodos Aristotelous 18, 26335, Patras, Greece

Abstract: Methods used for evaluating teacher performance aim at providing the university with an in-depth understanding of its academic staff. Improving teacher performance and enhancing their training so as to enable them to improve their teaching methods, is a common goal of academic institutions. In the case of the Hellenic Open University the methodology used is mainly based on the evaluation of tutors by their students. The students are asked once a year to fill in an anonymous questionnaire and return it to the tutors' supervisor. Then all data is processed. The presented study used tutors' evaluation from 5 consecutive academic years**.** Analysis of the results showed that the tutors' evaluations remained relatively stable across the years. However, ratings were significantly higher for female tutors than for the male tutors. Students' participation in the evaluation, students' grades and students attendance in the face to face meetings were positively correlated with the students' ratings.

Keywords- Tutor evaluation; educational quality; distance education; higher education.

1 Introduction

An increasing number of higher education institutes offer courses using distance education approaches (Liu, 2012). An Open and Distance Learning (ODL) system enables students to control the learning process to a greater level. They are able to choose the place, time, the procedure and in some cases the material to study (Liu, 2012). On the contrary, in a distance education course the students often have a "feeling of uncertainty" (Green et al., 2004). Previous research has shown that the students come to distance education courses with varying expectations in terms of level of service and support they will receive from their tutors (Stevenson & Sander, 1998).

As a result, the role of a distance education tutor greatly differs from the role of a university professor assigned to deliver face to face instruction. Communication between the student and the tutor constitutes an important parameter for the success of a distance education program (Barron, 1999). Thus, additional domain-specific skills are a prerequisite for a tutor. They have to teach, evaluate and continuously support their students despite the distance between them. This includes the tutor's ability to organize and orientate study groups towards a specific learning objective and to reflect the theoretical approaches on practical experiences (Beijaard & Verloop, 1996). They should also align students, in order to focus on the most important concepts and emphasize upon contextual application skills (West & West, 2009) which are required to turn their theoretical knowledge into practice.

---

* Corresponding author. *E-mail addresses:* kotsiantis@gmail.com (S. Kotsiantis), nitse@ece.upatras.gr (N. Tselios), xenos@eap.gr (M. Xenos).

Communication with the students by post, telephone, e-mail, through the written assignments or at optional consulting meetings are the means provided to the tutor to fulfil this complex role. In any case, the tutor should promptly solve students' educational problems, discuss the issues that distract them in a friendly manner, instruct them during their studies, but most of all encourage them to continue their studies, understanding the difficulties they face and effectively supporting them (Rowntree, 1998).

1.1 Teaching effectiveness and Students' evaluations of teaching

The assessment of higher education has been a complex and extensively discussed issue in scientific literature (e.g. McKenzie et al., 1998; Wachtel, 1998; Mutch, 2003; Valimaa and Mollis, 2003; Carless, 2006; Emanuel and Adams, 2006; Weaver, 2006; Spooren, Brockx, & Mortelmans, 2013). McKenzie et al., (1998) stress the need to explore a variety of methods in the process of evaluation, so as to improve the communication process between tutor and student. In most cases, students evaluate teaching quality using carefully designed questionnaires. Despite the fact that the reliability and validity of such questionnaires has been extensively discussed, various issues remain to be tackled (Spooren, Brockx, & Mortelmans, 2013). For instance, factors such teacher's gender, age, ethnicity, class size, course type (compulsory/elective, undergraduate/postgraduate) and their possible influence on students' ratings (Marsh, 2007, Rubin, 1992) are examined and the results in many cases are controversial or non-conclusive. In addition, the differences -if any- in the obtained results, using such questionnaires in distance education courses related to conventional higher education environments, are not adequately examined (Liu, 2012). Thus, a need for examining rich, longitudinal datasets related to students' evaluation of teachers emerges.

Teaching effectiveness is the ability of an instructor to bring out the best in his or her students, to motivate them to succeed and to reach the point at which they can begin to appreciate the value and the aesthetic beauty of the material they are studying. Quality teaching in turn facilitates quality learning (Biggs, 1999).

Tutors evaluation in higher education is not a novel field. Research has shown that students are requested to evaluate many aspects of teaching quality, and their ratings are valid, multidimensional and quite reliable (Wachtel, 1998). It is frequently argued that increased quality assurance and evaluation is a solution to all problems of monitoring, control and accountability. However, strict procedures and processes have their downside since they encourage steeper institutional hierarchies (Kehm, 2007).

According to Rhoades and Sporn (2002) the quality assurance in Europe should adopt valid teaching evaluations of academics; in the U.S. tutors have been evaluated by students for nearly thirty years. In the 1970s, student evaluations of their tutors began to be widely used in colleges and universities, and were incorporated into the annual review process for faculty. There is an extensive literature on these evaluations (Feldman, 1976, 1977; Seldin, 1984), which are the principal criteria for evaluating faculty's instructional activity and are used by over 95% of department heads nationwide (Cashin, 1999).

1.2 Questionnaires to allow students to evaluate their tutors

The use of formal instruments to obtain student feedback in higher education is a field which has been substantially developed in the last decades (Marsh, 2007; Watkins et al., 1987; Donald, 1997; Brown and Frielick, 2000; Macpherson et al., 2000; McInnis et al., 2001;

Richardson, 2005). Thus, various questionnaires have been developed to allow students evaluate their tutors.

The *Student Evaluation of Educational Quality* (SEEQ) survey is the most widely used questionnaire in North America, Australia, New Zealand and Spain (Marsh, 2007; Watkins et al., 1987). It comprises 35 questions. The statements are intended to reflect nine aspects of effective teaching: learning/value, enthusiasm, organization, group interaction, individual rapport, breadth of coverage, examinations/grading, assignments and workload/difficulty. It allows students to rate a tutor against crucial areas of excellence. The SEEQ instrument has been exhaustively researched (Marsh, 2007). Evaluations applied for over 15 years (with responses from approximately 50,000 courses and almost 1 million students in a wide range of disciplines at both the undergraduate and graduate levels) have shown that SEEQ is both valid and reliable (Marsh, 2007).

The *Learning Improvement Strategies Questionnaire* (LISQ, Macpherson et al., 2000) was designed as a means of monitoring the quality of semester-length courses. The criteria for the five LISQ items were based on research on the quality of learning environments in North American universities (Donald, 1997). It was found in Donald's (1997) research that top faculties and departments use five common strategies to improve the quality of a course: motivate students to learn; teach in ways that deepen understanding; take teaching responsibilities very seriously; assess student learning in impartially; and provide environments which encourage learning. The LISQ is shown to be a relatively quick, reliable and valid method of monitoring courses. It is suggested that it could also assist with formative evaluation should teaching teams develop learning improvement strategies from response distributions (MacPherson and Frielick, 2001).

The *Students Course Experience Questionnaire* (SCEQ, Ginns, Prosser and Barrie, 2007) was designed for monitoring and improving the quality of curricula and now is a survey of current University of Sydney students. It uses the Course Experience Questionnaire CEQ factor scale items. Two trials using the CEQ were conducted in 1998 with another two in 1999. The version of the CEQ used in these trials has 25 items that are grouped into 6 scales. These are: the good teaching scale; clear goals and standards scale; appropriate workload scale; appropriate assessment scale; generic skills scale; and overall satisfaction item. It is available for formative purposes if desired by departments (Brown and Frielick, 2000). The CEQ is considered a valuable instrument for the purpose of improving the quality of teaching in universities and also for informing student choice, managing institutional performance and promoting accountability of the higher education sector (McInnis et al., 2001). The University of Sidney and the Institute for Teaching and Learning (ITL) adopted the SCEQ annually between 1993 and 2003, the SCEQ will be administered biannually (Ginns, Prosser, & Barrie, 2007).

The *Formative Feedback Questionnaire* (FFQ) was designed for staffs who want to tailor surveys to their individual needs and prefer to choose their own questions. An item bank is provided, but staff are also free to design their 20 own items for teaching contexts which may differ in some aspect. The item bank contains a number of items for specific settings e.g. seminar or laboratory based courses. The SEEQ and LISQ have a 9-point, the FFQ a 7- point, and the CEQ a 5-point Likert scale (Brown and Frielick, 2000).

Furthermore, there is a consensus that students' ratings of teaching effectiveness depends on a large number of dimensions. However, there is debate related to whether they can be subsumed under a single, more global dimension. Marsh, (2007) states that, although students' scores on the dimensions of the SEEQ were correlated with each other, they could not be

adequately captured by a single higher-order factor. On the other hand, D'Apollonia and Abrami (1997) proposed that students' evaluations of teaching were subsumed by a single overarching construct that they defined as 'general instructional skill' (Richardson, 2005).

As derived by the discussion presented above, the use of student evaluations of teaching quality as a performance measure is not new. Nowadays, such measures constitute a part of many universities promotion mechanisms. A simple review indicates that the majority of the literature related to the course evaluation is based on quantitative analyses of learners' assessments of a course (and its tutors) using closed-ended questions, with one or two opportunities for free text comments.

A common conclusion from such studies is that higher ratings are modestly correlated with higher achievement (Donald, 1997; D'Apollonia and Abrami, 1997; Brown and Frielick, 2000; Ginns, Prosser and Barrie, 2007; Macpherson et al., 2000; McInnis et al., 2001; Richardson, 2005). However, Onwuegbuzie et al. (2007) used a multistage mixed-methods analysis to assess the content-related validity and construct-related validity (i.e., substantive validity, structural validity, outcome validity, generalizability) of a teaching evaluation form by examining students' perceptions of characteristics of effective tutors. Their analysis showed that the content and construct validity of the evaluation scores are under dispute since some of the most prominent characteristics of an effective teacher are not represented in the questionnaire. Moreover, student characteristics, such as age and sex, in some cases influence their perceptions (Marsh, 2007). Finally, most authors agree that anonymous evaluation systems are preferred to open evaluation systems (Albanese, 2000). When evaluations are not anonymous, participants worry about the implications of their evaluations, since they might have to work with that faculty again in the future (Afonso et al., 2005).

In the present study the following research questions are attempted to answer:

1. Are there any differences in the students' ratings between the schools or/and the academic years?
2. Are there any differences between the ratings given by students about their tutor and the ratings given by the tutor's coordinator?
3. What is the correlation between student participation in the evaluation of the tutors and the students' evaluation score?
4. What is the correlation between student grade in the lesson of the tutors and the students' evaluation score of the tutor?
5. Is there any difference in the rating between male and female tutors?
6. Is there a difference in the students' evaluation score due to their participation in the face to face meetings?

The rest of the paper is organised as follows. First the method of the study is presented, followed by the results. Subsequently, the obtained findings are extensively discussed and compared to findings from other related studies. Finally, future research goals are presented.

2 Method of the study

2.1 Goal of the study

The goal of the study is to examine attributes of students' evaluation of tutors in the context of distance tertiary education. In many cases, a tutor's teaching and supporting abilities are measured by his/her own students (Kwan, 1999, Chen and Hoshower, 2003). However, various

issues related to the reliability and the validity of such an approach remain unanswered. Moreover, there is a lack of studies which contrast application of students' evaluation approaches in a distance education setting against conventional educational institutes (Liu, 2012). In the case of the Hellenic Open University (HOU) the methodology used is mainly based on the evaluation of tutors by their students. Initially, the students are asked once a year to fill in an anonymous questionnaire and return it to the tutors' supervisor. During the last years, the students complete online the questionnaire thus ensuring anonymity of the responses. Subsequently, all data were processed with respect to the legislation concerning private data protection. Tutors are only informed about the outcome of the evaluation, after the completion of the academic year, i.e. after everyone has submitted their final grades for the course.

2.2 Procedure and participants

The presented study examines HOU's tutors' evaluation data from 5 consecutive academic years (2003-2008). A formal evaluation based on students' opinion of their tutors is held using anonymous questionnaires with closed-ended (five-point rating or yes/no scales) and open-ended questions.

The students of the HOU study their material, which usually consists of printed textbooks supplemented by audio-visual or electronic material, from a distance. Furthermore, the students have the opportunity to attend 5 or 6, four hour face-to-face counselling meetings with their tutor. Students attending each module are divided into small groups of approximately 30 students, each supervised by one tutor. During these meetings, the solutions of the given assignments are discussed, as well as frequent/repeated mistakes in the solutions given by the students. In addition, the material under consideration during the timeframe of a given meeting is presented using slides, with a focus on solving representative examples or presenting case studies. Subsequently, the material under consideration in the timeframe in which the meeting is taking place, is presented. Finally, study problems experienced by the students are discussed and possible solutions are examined.

The tutors were trained and guided mainly by the coordinator of each course. In addition, various seminars were provided by the HOU related to distance education to better prepare the tutors. During the academic year, the tutor has a frequent communication with the students via (a) mail, (b) phone conversations (c) forum discussions. In particular, they provide detailed and personalised feedback related to proper approaches to solve the given assignments, answers to possible questions related to specific concepts, encourages the students to continue studying at a certain pace. Students have to take modules in order to get a degree or certificate. For an undergraduate degree they have to take at least 12 modules (each module requires an estimated effort of 400 to 450 hours in terms of study and involvement).

The research was conducted from 2003 to 2008 in all HOU's schools and include the students who participated in a specific counselling meeting. The questionnaire was given to the students of the Hellenic Open University during the fourth face-to-face meeting which usually takes place in mid-March. 9662 students completed the questionnaire for academic year 2003-4, 10878 for 2004-5, 14421 for 2005-6, 15447 for 2006-7, and 15593 for 2007-8. Presenting the results was delayed to ensure that all participated students had graduated, due to HOU regulations regarding unveiling personal data.

The questionnaire used in the HOU is based on the directions given by the Hellenic Quality Assurance and Accreditation Agency (HQA), which is a governmental agency that monitors all evaluation procedures, tailored to the distance education particularities of HOU. Before the

evaluation, all students were analytically informed about the way in which studies in the HOU are organized. The students have had received their tutor's comments on their written assignments as well as the necessary advisory letters. They had repeatedly communicated with their tutor and the administrative services of the HOU and they had already studied the educational material. The criteria used for the evaluation of the tutoring aspects were:

- The encouragement and the support given to the students (question example: Does your tutor support and encourage you during your studies?)
- The type and the quality of the communication with the students (question example: Are you satisfied with the communication with your tutor?)
- The quality and the effectiveness of communication during face to face optional meetings (counselling meetings) (question example: Are you satisfied with the communication with your tutor during face to face meetings?)
- The way of evaluating and commenting the written assignments (question example: the usefulness of the comments on the written assignments)
- The period of evaluating and commenting the written assignments (question example: Does your tutor evaluate the written as in time?)
- The general impression and satisfaction from tutors' educational efforts (results to this criterion are derived from the combination of the answers to a number of related questions).

The scales used to measure the abovementioned criteria were mainly ordinal scales such as the five point rating scale (1=totally disagree; 5=totally agree), and the Yes/No scale. A high level of internal reliability has been found for the adopted questionnaire (Cronbach's alpha = 0.93).

2.3 Equipment

The analysis of the questionnaires was initially carried out using Optical Mark Recognition (OMR) for the questionnaires handed to the students, and electronically for the following years, and is subject to the Greek legislation protecting 'personal' data. The obtained results, for this paper, were analysed using Excel 2013, R version 3.2.3 and SPSS v22.

3   Analysis of results

Ratings between the Schools across the academic years

HOU offers distance education in Greece and comprises four Schools (Human Studies, Social Sciences, Sciences and Technology, Applied Arts). Each School combines a number of courses (study programmes) and each programme has a number of modules. HOU employs a modular system, which consists of yearly modules. Each module is equal to 3 or 4 'conventional' university-level lessons, depending on the module's difficulty. In HOU, the tutor contributes to the lessons' design during the course module (organization of homework, clarifications related to the preparation and evaluation of the written assignments, information about the final exams). With regard to the tutors' evaluation score in the four different Schools, a recurrent rating of the schools was observed in each academic year. 1st has been ranked the school of Human studies, 2nd the school of Science & Technology, 3rd the school of Social Sciences and 4th the school of Applied Arts (see Figure 1 and Table 1). However, the observed

differences were not substantial. The tutors in all schools received a grade ranging from 4.45 to 4.67 on average. There were no statistical differences according to Levene's test in the tutors' ratings across the academic years (one-way ANOVA, $F_{(4, 780)}=.471$, $p=.757$, ns). The tutors' evaluation scores are summarised in Table 1.
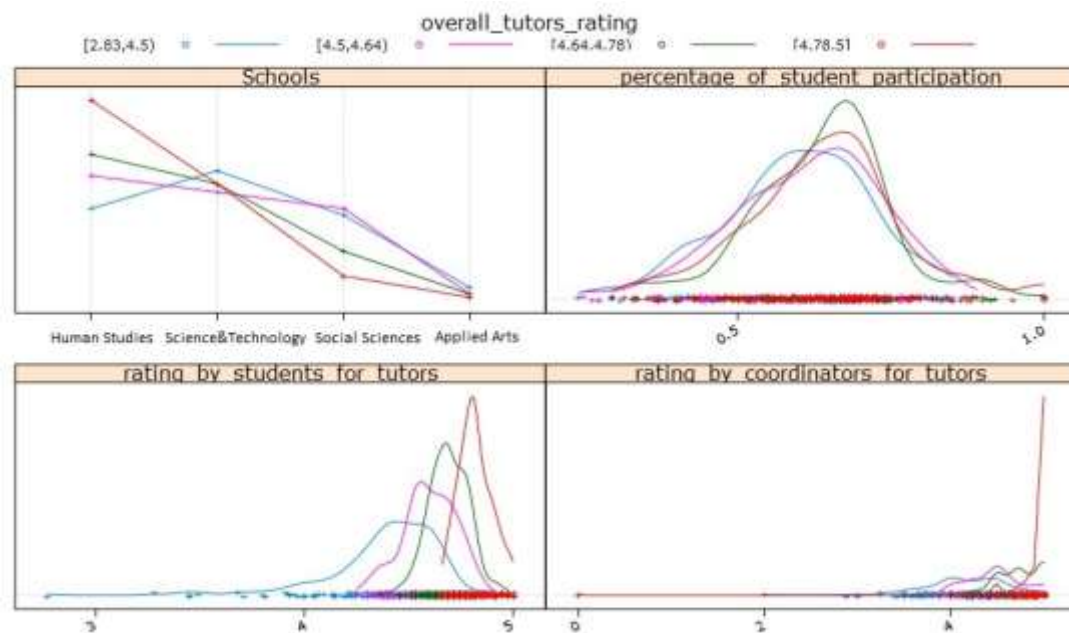
Figure 1 Graphical Representation of data

Table 1 Evaluation score of tutors in the four schools (Mean and Standard Deviation)

School/ Year	2003-4 (N=120)		2004-5 (N=139)		2005-6 (N=166)		2006-7 (N=180)		2007-8 (N=180)		TOTAL	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Applied Arts	n/a	n/a	4.64	.056	4.64	.16	4.16	.22	4.47	.15	4.45	.25
Human studies	**4.66**	.20	**4.66**	.22	**4.68**	.17	**4.68**	.16	**4.65**	.27	**4.67**	.21
Science &Technology	4.59	.18	4.57	.18	4.61	.21	4.59	.30	4.63	.20	4.60	.23
Social Sciences	4.58	.19	4.56	.16	4.51	.19	4.46	.29	4.51	.15	4.52	.20
Total	4.62	.19	4.62	.20	4.62	.20	4.59	.27	4.61	.23	4.61	.22

Differences between the ratings given by the students and by the tutors' coordinator

The tutor is responsible for communicating with each student in the group. The module's coordinator bears the overall responsibility of organizing the module, supervising the tutors and defining, in cooperation with the tutors, the educational methods and material as well as ways for supporting the students. In the HOU, all module coordinators belong to the University's permanent staff as opposed to tutors who sign / renew their contract with the HOU every year. The duties of module coordinators include the organization of the module and the provision of assistance to tutors in understanding the learning process and the particularities of ODL. The coordinator's duties include the responsibility of evaluating tutors based on their past activity and suggesting on contract renewing.

There were small but statistically significant differences between the tutors' ratings given by the students and the ratings given by the tutor's coordinator (paired samples test, $t(784)=3.085$, $p= 0.002$, s, Table 2). Moreover, there is a low but statistically significant correlation between the students' participation in the evaluation of the tutors and the students' evaluation score (Pearson's $r=.228$, $p=0.00$, s). Similar findings were obtained for each school (see Table 2).

Table 2 Summary of obtained data: Tutors' evaluation scores (* Correlation is significant at the 0.05 Level / (** Correlation is significant at the 0.01 Level)

School	students' evaluation score		coordinators' evaluation score		Students' evaluation score - coordinators' evaluation score correlation	Students' evaluation of tutors – students' performance	Students' participation in the evaluation- tutors' evaluation
	Mean	SD	Mean	SD	r	r	r
Applied Arts (N=14)	4.470	.21	4.532	.56	.008	.365**	.097
Human studies (N=339)	4.653	.24	4.701	.32	.218**	.243**	.048
Science &Technology (N=282)	4.615	.22	4.555	.51	.228**	.248**	.247**
Social Sciences (N=149)	4.606	.19	4.342	.41	.237**	.187**	.173*
Total (N=785)	4.627	.22	4.577	.44	.228**	.230**	.156**

Students' participation and correlation with the tutor's rating

Moreover, there is a low, but statistically significant, correlation between student participation in the evaluation of the tutors and the students' evaluation score (N=785, Pearson's $r=.156$, $p=0.00$, s). Thus, increased students' participation in the evaluation process is positively linked with the tutors' evaluation score. Figure 2 illustrates the percentage of

students' participation per curriculum. It should be stressed that student participation in the evaluation is not compulsory in the Hellenic Open University.

Figure 2 Box plot with the student participation in the tutors' evaluation for all academic years
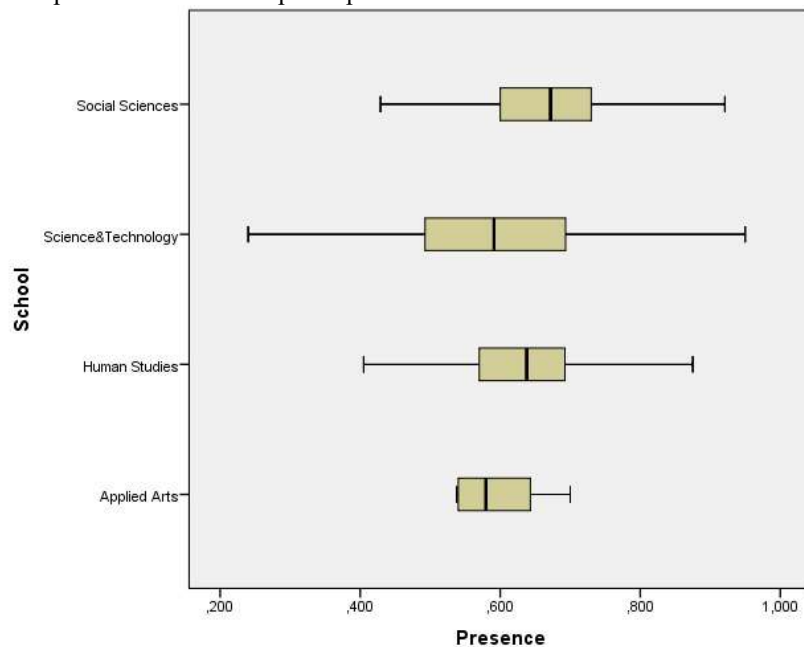
Students' grade and correlation with the tutor's rating

Data were obtained from 19308 students about their lesson performance. A significant modest correlation between the students' average group grade in the lessons' final examination and their tutor rating was observed (Pearson correlation .230, p=.000, s). The correlation was evident and statistically significant for the students of all four schools.

Differences on the ratings given by the students due to the tutors' gender

An independent-samples t-test indicated that scores were significantly higher for female tutors (mean=4.656, SD=.46) than for male tutors (mean=4.507, SD=3.57), t(1226) =5.049, p=.000, s. Levene's test indicated unequal variances (F=12.815, p=0.000), thus degrees of freedom were adjusted from 1254 to 1226.

Differences on the ratings given by the students according to their participation in the face to face optional meetings

Table 3 Evaluation score of tutors in relation to the attended face to face optional meetings

	Number of face to face optional meetings	N	Mean	SD
2003-4	>= 4	9913	4.291	.779
	< 4	2130	3.974	.957
2004-5	>= 4	11706	4.288	.789
	< 4	2426	3.958	.967
2005-6	>= 4	14026	4.293	.784
	< 4	3000	3.963	.970
2006-7	>= 4	14880	4.293	.782
	< 4	3198	3.967	.968
2007-8	>= 4	15189	4.292	.782
	< 4	3283	3.956	.971

On the contrary, there is a significant statistical difference between the students' evaluation if the number of face to face optional meetings at least 4 or less than 4 (see Table 3). The students that attended at least four face to face optional meetings rated significantly higher their tutors (mean=4.29, SD=.602) than the remaining students (mean= 3.96, SD= .504), independent samples t-test, $t(79749)= 17.231$, $p=0.000$, s. The difference is significant and stable in each academic year.

4 Discussion and Conclusion

In this paper a longitudinal study of students' evaluation of tutors in the context of a Distance Learning Higher Education institute is presented. From the data analysis, it arises that:

- There are no significant differences in the ratings between the academic years.
- In addition, with regard to the evaluation score of tutors in the four different schools, a recurrent pattern of schools rating of the schools was emerged in each academic year: 1st was the school of Human studies and 2nd was the school of Science & Technology, 3rd was the school of Social Sciences followed by the school of Applied Arts. Marsh (2007) also reports a weak tendency for higher ratings in Humanities.
- There were small but significant differences between the tutors' ratings given by the students and the ratings given by the tutor's coordinator.
- There is a low but significant correlation between the students' participation in the tutors' evaluation and the evaluation score.
- There is a medium, significant correlation between the students' grade in the final examination and their tutor's evaluation, a finding which is in line with Marsh (2007).
- Scores were significantly higher for the female tutors than for the male tutors. However, recent work from MacNell, Driscoll, and Hunt (2015) found significant differences in how

students rated the perceived male and female instructors, in the opposite direction. However, Liu (2012) found that students rate perceived male and female instructors the same.

- There is a significant statistical difference in the students' evaluation according to their participation in the face to face meetings >=4 or <4. The students who attended the majority (>=4) of the face to face optional meetings rated significantly higher the tutors in comparison to the students who did not attend a significant number of meetings (<4).

As the analysis of the obtained data shows, the aforementioned findings are verified each year, with very few exceptions due to small sample issues in a specific school. Therefore, they are recurrent phenomena, a fact which in turn is an indication of the stable performance of the HOU schools. High ratings provided by the students establish the credibility of online education. Tutor evaluation is a necessity for a lot of universities which offers distance education courses. To this end, students' perceptions of the teaching and the learning context, is a strong indication of quality of the learning outcome. Tutors can explicitly influence only some components of the learning process. From this perspective, information on how students perceive and respond to teaching is possibly of more importance (Lowenthal, Bauer, and Chen, 2015). Student perceptions of the learning context are vital and vulnerable. That is, because perceptions are influenced by many factors and in some cases are susceptible to manipulation and improvement (Marsh, 2007). At this point, it must be mentioned that the present research had as a concern that: "any comparisons among different courses should take into account the diversity of educational contexts and student populations" (Richardson, 2005).

Evaluation of tutors constitutes a crucial factor in personnel decisions. In the context of an open distance learning university (such as HOU) which employ a large number of external adjunct tutors/professors who have to renew their tutoring contracts with the institution every year, the reliability of the evaluation obtained has not been studied extensively. From the obtained findings is derived that there is a small but significant tutors' gender bias. Moreover, the link between students' grade and the tutors evaluation should be carefully examined since it is not clear whether higher grades represent "grading leniency, superior learning, or pre-existing differences" (Marsh, 2007). Perhaps, better grades given by the tutor (i.e. grade inflation) lead to higher evaluations. However, it is possible that the link unveiled demonstrates that a carefully designed learning process could both positively influence students' grades and tutors' ratings.

On the other hand, lack or major differences in the obtained results across the years is an indication of the credibility of the learning process. In addition, it is an indication that using students' evaluations in comparison to historical data as a means to renew or terminate a contract constitutes an adequate criterion. However, in order to avoid discriminations and superficial conclusions some of the obtained results should be interpreted with caution.

There is no doubt that the task of evaluating the competence and teaching effectiveness of its academic instructors is one of the most important challenges any university could face. Since the early 1970s, educational sciences and instructional psychology have paid a great deal of attention to the research on student ratings of instruction in higher education (Chonko et al., 2002). However, Kember et al. (2002) found no evidence for the hypothesis that the use of their Student Feedback Questionnaire improved teaching performance, at least as perceived by the students. Therefore, Ballantyne et al. (2000) state that when student evaluations are used to improve teaching quality, more attention must be given to personal follow-up and correct interpretation of such ratings.

Examination of the relation between the class size and the tutors' evaluation constitutes a notable future research goal (Liu, 2012). Moreover, investigation of the students' personality characteristics as measured with the Big Five Personality questionnaire and their tutors' rating as well as deeper examination of the link between students' grades and their tutors' evaluation are scheduled for future examination. In addition, the impact of interaction design aspects of the platform adopted to deliver the educational material on the students' satisfaction should be closely monitored and examined (Daskalakis, and Tselios, 2011; Kotsiantis, Tselios, Filippidi, and Komis, 2013; Katsanos, Tselios, and Avouris, 2008, Orfanou, Tselios, and Katsanos 2015). An open university such as the HOU is not a 'conventional' university which addresses to more homogenous student communities. It mainly addresses adults with special educational needs and incongruity (in terms their age, their professional and family obligations are concerned). As a result, investigation of such issues is of significant importance and factors related to students' evaluation of teacher need to be revisited by taking into account the distance education paradigm (Lowenthal, Bauer, and Chen, 2015).

References

Afonso, N. M., Cardoza, L. J., Mascarensas, O. A., Aranha, A. N., and Shah, C. (2005) 'Are anonymous evaluations a better assessment of faculty teaching performance? A comparative analysis of open and anonymous evaluation processes', *Family Medicine,* Vol. 37, pp. 43–47.

Albanese, M. A. (2000) 'Challenges in using rater judgments in medical education', *Journal of Evaluation in Clinical Practice,* Vol. 6, No. 3, pp. 305–319.

Ballantyne, R., Borthwick, J., and Packer, J. (2000) 'Beyond student evaluation of teaching: identifying and addressing academic staff development needs', *Assessment & Evaluation in Higher Education,* Vol. 25. No. 3, pp. 221–236.

Barron, A. (1999) *A teacher's guide to distance learning.* Florida: Florida Centre for Instructional Technology.

Beijaard, D., and Verloop, N. (1996) 'Assessing teachers' practical knowledge', *Studies in Educational Evaluation,* Vol. 22, No. 3, pp. 275-286.

Biggs, J. (1999) *Teaching for Quality Learning at University.* Buckingham: SRHE/Open University Press.

Brown, J., and Frielick S. (2000) Student evaluation of teaching at the University of Auckland. Proceedings of Teaching Evaluation Forum, Perth, Australia, pp. 17-22.

Cashin, W.E. (1999) 'Student ratings of teaching: Uses and misuses'. In P., Seldin. (ed.), *Changing Practices in Evaluating Teaching*. Bolton, MA: Anker Publishing Company.

Chen Y., and Hoshower, L. (2003) 'Student evaluation of teaching effectiveness: an assessment of student perception and motivation', *Assessment & Evaluation in Higher Education,* Vol. 28, No. 1, pp. 71-88.

Chonko, L. B., Tanner, J. F., and Davis, R. (2002) 'What are they thinking ? Students' expectations and self-assessments', *Journal of Education for Business,* Vol. 77, pp. 271–281.

Daskalakis, S., and Tselios, N. (2011) 'Evaluating e-learning initiatives: A literature review on methods and research frameworks'. *International Journal of Web-Based Learning and Teaching Technologies* (IJWLTT), Vol. 6, No. 1, pp. 35-51.

D'Apollonia, S., and Abrami, P. C. (1997) 'Navigating student rating of instruction', *American Psychologist,* Vol. 52, No. 11, pp. 1198–1208.

Donald, J. G. (1997) Improving the Environment for Learning: academic leaders talk about what works. San Francisco: Jossey-Bass.

Emanuel, R., and Adams, J.N. (2006) 'Assessing college student perceptions of instructor customer service via the Quality of Instructor Service to Students (QISS) Questionnaire', *Assessment & Evaluation in Higher Education*, Vol. 31, No. 5, pp. 535-549.

Feldman, K.A. (1976) 'Grades and colleges students' evaluations of their courses and teachers', *Research in Higher Education,* Vol. 4, pp. 69–111.

Feldman, K.A. (1977) 'Consistency and variability among college students in rating their teachers and courses: A review and analysis', *Research in Higher Education,* Vol. 6, pp. 233–274.

Ginns, P., Prosser, M., and Barrie, S. (2007) 'Students' perceptions of teaching quality in higher education: The perspective of currently enrolled students'. *Studies in Higher Education,* Vol 32, No 5,pp. 603-615.

Green, C.J., Van Gyn, G.H., Moehr, J.R. Lau, F.Y., and Coward, P.M. (2004) 'Introducing a technology-enabled problem-based learning approach into a health informatics curriculum', *International Journal of Medical Informatics,* Vol. 73, No. 2, 173-179.

Katsanos C., Tselios N., and Avouris N. (2008) 'AutoCardSorter: Designing the Information Architecture of a Web Site using Latent Semantic Analysis'. In Proceedings of the SIGCHI conference on Human factors in computing systems', CHI 2008, pp. 875-878, Florence, Italy:ACM Press, April 5-10, 2008.

Kember, D., Leung, D., and Kwan, K. (2002) 'Does the use of student feedback questionnaires improve the overall quality of teaching?', *Assessment & Evaluation in Higher Education,* Vol. 27, No. 5, pp. 412–425.

Kehm, B. (2007) Higher Education Research at the Turn of the Millennium: Dimensions and Issues. An Introduction. In: B. Kehm (Ed.) Looking Back to Look Forward (pp. 5-12). Kassel: INCHER- kassel.

Kotsiantis, S., Tselios, N., Filippidi, A., and Komis, V. (2013) 'Using Learning Analytics to identify successful learners in a blended learning course'. *Journal of Technology Enhanced Learning* (special issue on Learning Analytics), Vol. 5, No. 2, 133-150.

Kwan, K. P. (1999) 'How fair are student ratings in assessing the teaching performance of university teachers?', *Assessment & Evaluation in Higher Education,* Vol. 24, No. 2, pp.181-195.

Liu, O.L. (2012) 'Student evaluation of instruction: In the new paradigm of distance education. Research in Higher Education', Vol. 53, No. 4, pp. 471-486.

Lowenthal, P., Bauer, C., and Chen, K. Z. (2015) 'Student perceptions of online learning: An analysis of online course evaluations', *American Journal of Distance Education,* Vol. 29, No. 2, 85-97.

McInnis, C., Griffin, P., James, R., and Coates, H. (2001) 'Development of the course experience questionnaire (CEQ)'. *Melbourne: University of Melbourne, Department of Education, Training and Youth Affairs.*

MacPherson, R., and Frielick, S. (2001) The Learning Improvement Strategies Questionnaire. *ISEA,* Vol. 29, No. 3, pp. 40-50.

Macpherson, R. J. S., Pashiardis, P., and Frielick, S. (2000) 'The quality of courses at the University of Cyprus: a case study and comparative analysis', *Teaching in Higher Education,* Vol. 5, No. 2, pp. 219-232.

Marsh, H., and Roche, L. (1997) 'Making students' evaluations of teaching effectiveness effective: The critical issues of validity, bias, and utility', *American Psychologist,* Vol. 52, No. 11, pp. 1187-1197.

Marsh, H.W. (2007) Students' evaluations of university teaching: Dimensionality, reliability, validity, potential biases and usefulness. In *The scholarship of teaching and learning in higher education: An evidence-based perspective* (pp. 319-383). Netherlands: Springer.

McKenzie, J., Sheely, S., and Trigwell, K. (1998) 'Drawing on experience: an holistic approach to student evaluation of courses', *Assessment & Evaluation in Higher Education,* Vol. 23, No. 2, pp. 153-63.

MacNell, L., Driscoll, A., and Hunt, A.N. (2015) 'What's in a name: exposing gender bias in student ratings of teaching'. *Innovative Higher Education,* Vol. 40, Nr, 4, 291-303.

Mutch, A. (2003) 'Exploring the practice of feedback to students', *Active Learning in Higher Education,* Vol. 14, No. 1, pp. 24-38.

Onwuegbuzie, A., Witcher, A., Collins, K., Filer, J., Wiedmaier, C., and Moore, C. (2007) 'Students' Perceptions of Characteristics of Effective College Teachers: A Validity Study of a Teaching Evaluation Form Using a Mixed-Methods Analysis', *American Educational Research Journal,* Vol. 44, No. 1, pp. 113-160.

Orfanou, K., Tselios, N., and Katsanos, C. (2015) 'Perceived usability evaluation of learning management systems: empirical evaluation of the system usability scale'. *The International Review of Research in Open and Distance Learning* (IRRODL), Vol. 16, No. 2, 227-246.

Rhoades, G., and Sporn, B. (2002) 'Quality assurance in Europe and the U.S.: Professional and political economic framing of higher education policy', *Higher Education,* Vol. 43, No. 3, pp. 355-390.

Richardson, J. (2005) 'Instruments for obtaining student feedback: a review of the literature', *Assessment & Evaluation in Higher Education,* Vol. 30, No. 4, pp. 387-415.

Rubin, D.L. (1992) 'Nonlanguage factors affecting undergraduates' judgments of nonnative English-speaking teaching assistants'. *Research in Higher Education*, Vol. 33, No. 4, pp. 511-531.

Rowntree, D. (1998) *Exploring Open & Distance Learning.* Institute of Educational Technology, Open University. London: Kogan Page.

Seldin, P. (1984) *Changing Practices in Faculty Evaluations*. San Francisco: Jossey-Bass.

Spooren, P., Brockx, B., and Mortelmans, D. (2013). On the validity of student evaluation of teaching the state of the art. Review of Educational Research, Vol. 83, No. 4, pp. 598-642.

Stevenson, K., and Sander, P. (1998) 'How do Open University students expect to be taught at tutorials?', *Open Learning*, Vol. 13, No. 2, pp. 42-46.

Välimaa, J., and Mollis, M. (2003) 'Social Functions of Evaluation in Argentine and Finnish Higher Education', *Higher Education in Europe,* Vol. 29, No. 1, pp. 67-86.

Wachtel, H. K. (1998) 'Student evaluation of college teaching effectiveness: A brief review', *Assessment & Evaluation in Higher Education,* Vol. 23, No. 2, pp. 191-211.

Watkins, D., Marsh, H. W., and Young, D. (1987) 'Evaluating tertiary teaching: a New Zealand perspective', *Teaching & Teacher Education,* Vol. 2, pp. 41–53.

Weaver, M.R. (2006) 'Do students value feedback? Student perceptions of tutors' written responses', *Assessment & Evaluation in Higher Education,* Vol. 31, No. 3, pp. 379-394.