İşletme Fakültesi Dergisi, Cilt 17, Sayı 1, 2016, 65-87

Gönderme Tarihi: 1 Mart 2016
Kabul Tarihi: 24 Mayıs 2016

EVALUATING THE INSTRUCTOR EFFECTIVENESS: A TURKISH BUSINESS FACULTY EXAMPLE

Haeeun KIM[*], Melissa N. CAGLE[**], Ulaş ÇAKAR[***]

ABSTRACT

In light of recent developments in the educational system in Turkey, we aimed to investigate students' evaluation on teaching effectiveness in business school. We attempted to learn the characteristics students consider when they evaluate overall performance of instructors and the factors that influence the students when they appraise the instructors and their courses. In our study, we attempted to draw attention to a verified multidimensional measure of teaching evaluation. For this purpose, Students' Evaluation of Teaching Effectiveness Rating Scale (SETERS) was used in a business faculty in Turkey. Findings show that students consider "effective knowledge delivery" and "communication skill" as important aspects of a successful instructor.

Keywords: *Business Education, Student Evaluation, Teaching Effectiveness in Business School, Educational Research, Student Characteristics*

ÖĞRETİM ÜYESİ ETKİNLİĞİNİN DEĞERLENDİRİLMESİ ÜZERİNE BİR TÜRK İŞLETME FAKÜLTESİ ÖRNEĞİ

ÖZ

Türk eğitim sistemindeki son gelişmeler ışığında, "etkin öğretim" ve "öğrencilerin değerlendirmesi ile etkin öğretim" kavramlarını bir işletme fakültesi bağlamında araştırmayı amaçladık. Eğitimcilerin genel performansını değerlendirmek amacı ile öğrencilerin ne tür özellikleri dikkate aldığını; eğitimcillerin genel performansını ve derslerini değerlendirirken öğrencileri etkileyen faktörleri sorguladık. Çalışmamızda, "Student's Evaluation of Teaching Effectiveness Rating Scale" (SETERS), öğretim değerlendirme üzerine uluslararası boyutlu bir ölçek, Türkiye'deki bir üniversitede, işletme fakültesi altında ele alınmıştır. Ulaşılan bulgular, öğrencilerin "etkin bilgi aktarımı" ve "iletişim becerisi" özelliklerini etkin bir eğitimcide olması gereken önemli unsurlar olarak nitelendirdiklerini göstermektedir.

Anahtar Sözcükler: *İşletme Eğitimi, Öğrencilerin Öğretim Üyesi Değerlendirmesi, İşletme Yönetiminde Öğretim Etkinliği, Eğitim Araştırmaları, Öğrenci Karakteristikleri*

[*] Dokuz Eylül Üniversitesi, İşletme Fakültesi, Tınaztepe Kampüsü, Buca, İzmir, E-posta: haeeun09@gmail.com
[**] Dokuz Eylül Üniversitesi, İşletme Fakültesi, Tınaztepe Kampüsü, Buca, İzmir, E-posta: melissa.cagle@deu.edu.tr
[***] Dokuz Eylül Üniversitesi, İşletme Fakültesi, Tınaztepe Kampüsü, Buca, İzmir, E-posta: ulaş.çakar@deu.edu.tr

INTRODUCTION

Turkish higher educational system has a long history of modernization towards Western educational systems. After the Bologna process, which was enforced after 29 educational representatives of European countries signed on the Bologna declaration (Yüksek Öğretim Kurulu, 2016), this process has been accelerated and the adaptation of European educational standards resulted in a new era of Turkish educational system. The process includes revising the traditional Turkish educational curriculum into European Credit Transfer and Accumulation System (ECTS), and issuing Diploma Supplement (DS). The revisions consist of creating new or revising courses to fit into the criteria of the process, and changing the credit system and outcome based course evaluation. A major feature of these changes in Turkish universities is a fresh orientation to the responsibility of teaching. The act of teaching is no longer seen as the sole responsibility of the instructor, it became the responsibility of an entire teaching institution. Considering the current reformation, it is understandable that the interest on teaching effectiveness is increasing. Keeping up with the contemporary flow of educational reformation, Turkish business schools which bear increasing importance in society are also showing attention toward teaching effectiveness and students' evaluation.

The concept of teaching effectiveness in business education has been connected with student satisfaction and was discussed as a way to present further directions on which business institutions should take (DeShields Jr., Kara, & Kaynak, 2005; LeBlanc & Nguyen, 1997; Letcher & Neves, 2010). The development of this topic has followed a similar vein in Turkish literature as the improvements in educational institutions have increased the emphasis on teaching effectiveness. Karaca (2008) addressed the instructor capabilities and responsibilities in developing quality standards for education. The author suggested that two-way communication between student-instructor and contentious evaluation of the education quality are necessary. Açan and Saydan (2009) on the other hand, attempted to analyze the concept of quality under a university setting. They analyzed the satisfaction levels of 700 students' assessment of 'quality elements' of instructors and determined professional behavior, interest building, empathy and personal reputation to be significant characteristics. Gülcan, Kuştepeli, and Aldemir (2002) looked into the current student satisfaction levels within a business faculty. They suggested that further analysis and discussion are needed on factors leading to student dissatisfaction; issues such as two-way

communication with student-instructor, and assessment on student performance/expectation criteria. The authors also highlighted the importance of determining 'composition and quality' of efficient instructors.

The main purpose of the evaluation of teaching effectiveness is to provide feedback on the overall progress of the teacher/institution. Cohen (1980) stated that students' ratings hold three distinct purposes; for administrative purpose such as determining instructors' salary, promotion, and tenure, for instructors to improve quality of their lectures, and for students to decide which lectures/instructors they choose. Cohen and Herr (1979) stated that providing feedback to instructors on their within-class efforts allows instructors improve their teaching as the semester progresses, and also provides the benefit of motivating the instructor to feel as a part of the faculty development process. Santiago and Benavides (2009) stated that student evaluations can have the additional benefits of serving as an independent and objective assessment of the teachers' performance, generate awareness towards the consequences of the evaluation and form clear individual objectives with regards to all aspects of a teachers' performance.

Onwuegbuzie, Witcher, Collins, Filer, Wiedmaier, and Moore (2007) however, argued that scales used to measure teaching effectiveness do not fully represent important characteristics students consider when they evaluate instructors. The authors also pointed that the evaluation result may be influenced by different factors, which threatens content-related and construct-related validity of scales. Since higher educational institutions consider students' evaluation as trustable despite the lack of reliability and validity of measure, the authors suggested that it is an urgent matter to form a systematically trustable measure of teaching effectiveness. This assessment holds particularly true as the Turkish higher educational system continues its efforts to modernize towards Western educational systems. Tran (2015) also states that using Student Evaluation of Teaching (SET) in higher education is controversial because SETs lack theoretical framework and thus reliability and validity of the measures are not strongly constructed. We thus, propose to investigate the validity of a teaching evaluation form by examining students' perceptions of characteristics of instructors. Given the reformation environment and increasing concern on teaching effectiveness in Turkey, it is noteworthy that Turkish educational context especially Turkey's business education which has significant role in Turkish society needs to

be researched more in depth. In this study we emphasize on teaching effectiveness of business school in Turkey, and aim to provide insights on how to effectively manage instructors and students in business faculty.

In light of the potential benefits provided by the measurement of teaching effectiveness, the purpose of this research is to provide empirical evidence supporting the use of a multidimensional (Abrami, d'Apollonia, & Rosenfield, 1996; Dodeen, 2013; Feldman, 1976; Marsh, 1987) profile of student evaluation. We hope that the current study provides insight and understanding to future researchers about the importance and prominence of student evaluations of teaching efficiency in Turkey. However, as the study is not a representative of the Turkish students, we caution against generalizing our findings for Turkish universities.

In the first part of this study, we examine previous literature covering the usage of various teaching evaluation scales and demographic factors affecting SETERS. In the second part, we described measures of the study, data collection procedure, and quantitative techniques implemented. Statistical techniques to test our hypothesis are listed under data analysis. The third part of the study covers the results of our analysis and lastly, our findings are summarized under the discussion and conclusion presenting significance of the result on business faculty.

BACKGROUND LITERATURE

Different educational contexts and goals require the use of different teaching methods. Therefore, it is not easy to define a way of effective teaching into one single concept. Brophy and Good (1984) determined the instructors who present in an active manner, explain, illustrate, and reinforce concepts were considered to be more successful instructors than those who did not. Giovannelli (2003) studied 150 researches in order to develop a definition of effective teaching and found that, among the conducted research, there isn't a consensus regarding what constitutes as effective teaching. However, she observed a repeating pattern of several teaching behaviors; management of the classroom, instructional behavior, classroom organization, and the expectations from the teacher. Hattie and Learning (2009) examined over 800 studies ranging from early childhood through adult education and stated that providing formative evaluation to lecturers, teacher clarity, feedback, and self-reported grades have a positive impact on learning. Abrami et al. (1996) suggested product definition, process

definition and product-process definition of effective teaching. Product definition acknowledges students' outcome after teaching, process definition considers the whole process of teaching, and product-process definition links both product and process aspect of the teaching. Following the conceptual frame of Abrami et al, we perceive effective teaching as comprehensive idea which includes teachers' performance, behaviour, and attitude toward students, and interaction with students both during and after the teaching process.

Many scales have been developed to accurately investigate the teaching effectiveness. Feldman (1976) proposed that students' evaluations of teaching consists of instructor's stimulation of interest, knowledge of the subject, elocutionary skills, nature and value of the course material, and his/her intellectual expansiveness. One of the most consistently applied scales is SEEQ (Students' Evaluation of Educational Quality) of Marsh (1987). The scale consists of 9 dimensions such as Learning/Value, Instructor Enthusiasm, Organization/Clarity, Group Interaction, Individual Rapport, Breadth of Coverage, Examinations/Grading, Assignments/Readings, and Workload/Difficulty.

Abrami et al. (1996) stated evaluation on teaching effectiveness included three factors; instructor's role in delivering information, facilitating a social learning environment, and regulating student learning. Toland and De Ayala (2005) constructed the Student's Evaluation of Teaching Effectiveness Rating Scale (SETERS) based on the frame of Abrami et al. (1996) and d'Apollonia and Abrami (1997). The study included three factors, Instructor's Delivery of Course Information, Instructor's Role in Facilitating Instructor/Student Interactions, and Instructor's Role in Regulating Students Learning. Instructor's Delivery of Course Information measures the extent of how well an instructor presents course material and organizes a classroom. Instructor's Role in Facilitating Instructor/Student Interactions includes instructor's attitude and behaviour toward students and communication with them. Instructor's Role in Regulating Students Learning consists of instructor's interest on students' learning and feedback.

Additional studies conducted to identify Students Evaluation of Teaching (SET) factors have highlighted the importance of management of behavior and instructional presentation (Swartz, White, & Stuck 1990); respect, organization, and challenging students (Patrick & Smart, 1998). Martínez-Gómez, Sierra, Jabaloyes, and Zarzo (2011) developed a 'home-made' (Marsh & Roche, 1997) statistical analysis based on the Students'

Evaluation on Teaching (SET) scale of their institution. The scale included 19 items and the research found that 5 factors can account for approximately 78% of the teaching effectiveness. The study worked towards the validation of five dimensions for explaining the underlying structure of the SET. Onwuegbuzie et al. (2007) conducted a study to analyze the content-related and construct-related validity of a teaching evaluation form. The authors constructed the CARE-RESPECTED Model of Teaching Evaluation, which includes teacher's attributes that students considered when they evaluate teaching effectiveness (communicator, advocate, responsible, empowering) and nine themes (responsive, enthusiast, student centered, professional, expert, connector, transmitter, ethical, and director). Slate, LaPrairie, Schulte, and Onwuegbuzie (2011) examined the dominant traits of effective instructors and they identified 29 instructors' prevailing themes including knowledgeable, understanding, communication, caring, organized, and such. Thus we see that despite considerable research being conducted on the topic of teaching effectiveness, discussions on assessment issues have not still been resolved.

Most of the studies regarding SET survey were conducted in the North American context (Marsh, 2007). However, Watkins (1994) has expressed the need to test the reliability and validity of SEEQ in different countries such as India, Nepal, Nigeria, and Philippines. Through the cross-cultural comparison, Watkins concluded despite the low degree of generalizability of the scale to other cultures, SEEQ showed acceptable internal consistency. Marsh, Hau, Chung, and Siu's study (1998) on Chinese students has also shown that SEEQ was also valid in China.

When we examine the Turkish literature, we see that there are few researches conducted on teaching effectiveness in Turkish educational system. Therefore, conducting more research in Turkish educational context is required. Özgüngör (2010) measured the perception of teaching effectiveness of Turkish students with Course Experience Questionnaire (CEQ). The degree of student's self-efficacy was taken as an extra influential factor in the study. Özgüngör presented that self-efficacy plays a significant role in evaluation on teaching effectiveness. Taking this into consideration, our research aimed to measure business students' perception on teaching effectiveness with Turkish university students.

Thus, the main research question is *"Does the factor structure of SETERS show a similar structure with the Turkish students' sample?*". As mentioned before, it is not easy to define "effective teaching", as there

are multiple dimensions inherent under the concept. Thus, a single query of how satisfied the students were with the course is not adequate to accurately evaluate teaching effectiveness (Abrami et al., 1996; Feldman, 1976; Marsh, 1987) and would reduce the validity (Dodeen, 2013) of the analysis. It is imperative to analyze the structure of SETERS and determine its capability of successfully addressing the overall performance of the instructor. Therefore, our first hypothesis is, *Hypothesis 1: The factor structure of SETERS is influential in explaining the overall performance of the instructor.*

There has been a long standing debate on the effect of several potential influencing factors on the evaluation of teaching effectiveness. The potential factors include grading allowance by instructors and class size (Greenwald and Gillmore, 1997), gender of student, expected grade, and class size (Dodeen, 2013), age of students, gender of students and gender of instructor, teaching style and learning style, instructor's personality, locus of control (Sprinkle, 2009), pre-course interest (Francis, 2011), and students' self-efficacy (Ozgungor, 2010). However, numerous researchers (d'Apollonia & Abrami, 1996, 1997; Marsh, 2007; Renaud & Murray, 2005) also support the effectiveness and neutrality of these factors on teaching effectiveness evaluations. Thus, it is important to assess certain elements and determine if there is a strong relation with the SETERS measures. Bearing this position, we considered the influence of same-gender preferences, expected score and instructor's title on the scale.

When it comes to same-gender preferences, previous research findings are mixed. For example, Basow and Silberg (1987) stated that male students give female professors significantly poorer ratings than male professors. They also found that female students evaluate female professors less favorably than male professors. On the contrary, Centra and Gaubatz (2000) found that female instructors were rated higher by female students. They stated that the differences might stem from the teaching styles of the instructors. Taking these findings and our initial argument for the neutrality of evaluation measures into consideration, our second hypotheses are presented as below.

Hypothesis 2: There is no significant relationship between same-gender preferences of students and the evaluation of instructor's performance.

This hypothesis shall be tested through the measurement of Hypothesis 2a and 2b, by holding for the gender of the instructor and examining to see if there is a significant difference between male and female students'

perceived evaluation of instructors.

Hypothesis 2a: There is no significant relationship between students' gender and the evaluation of female instructor's performance.

Hypothesis 2b: There is no significant relationship between students' gender and the evaluation of male instructor's performance.

Another influencing variable for these measures is the student's end of year expected grade. Marsh (1987) and Feldman (1976) studied the effect of expected grades on evaluation and found a weak but positive association; as the expected level of grades of students increased, the overall performance of instructors were found to follow suit. However, Marsh and Dunkin (1992) and Howard and Maxwell (1982) proposed that the research on grading lacked support and its effect on SET was weakly correlated. Thus, our third hypothesis is as follow;

Hypothesis 3: There is no significant relationship between year-end expected score and evaluation of instructor's performance.

In regards to the instructors' title (level) within the organization, Nelson and Lynch (1984) supported that the instructors' rank had a significant effect on evaluations. On the contrary, Marsh and Bailey (1993) analyzed the instructor's academic degree (teaching assistants, visiting professors, tenure-track assistant professors or tenured professors) and found that it did not influence students' evaluations. In this study, we assume that instructor's title variable has little influence on teaching effectiveness. Therefore, our fourth hypothesis is presented as below.

Hypothesis 4: There is no significant relationship between the instructor's level in the organization and evaluation of instructor's performance.

Originally, Marsh and Roche (1997) argued that SET studies should consider the three aspects; theory, research, and practice of the measurements. The study of Toland and De Ayala (2005) has its foundation on the theory developed by Abrami et al. (1996) and d'Apollonia and Abrami (1997). Students Evaluation of Teaching Effectiveness Rating Scale was constructed through their research, and they proposed more empirical study to support the validity of the scale. Integrating the need to widen educational research area in Turkish educational system and the need to validate SETERS, the current study investigated how our samples of Turkish students perceive teaching effectiveness, by applying SETERS and testing our hypothesis. The study is expected to contribute to the increasing interest in educational effectiveness studies in Turkey and on the practical validity of SETERS.

RESEARCH DESIGN

Measures

The Students Evaluation of Teaching Effectiveness Rating Scale (SETERS) of Toland and De Ayala (2005) was employed under the analysis. Toland and De Ayala (2005) based their theoretical foundation on the studies of Abrami et al. (1996) and d'Apollonia and Abrami (1997). SETERS consisted of 34 items and 3 factors; Instructor's Delivery of Course Information (IDCI), Instructor's Role in Facilitating Instructor/Student Interactions (IRFI), and Instructor's Role in Regulating Students Learning (IRRL). Toland and De Ayala (2005) supported the position of process-product definition regarding effective teaching. Therefore, students' rating on teaching effectiveness bears two perspectives; *"what teachers do (process) and the impact teachers have on students (product)"*. d'Apollonia and Abrami (1997) also suggested students' measurement on teaching effectiveness includes a global component; delivering instruction, facilitating interactions, and evaluating student learning.

Procedure and Participants

The scale items from Toland and De Ayala (2005) were translated into Turkish using the forward and backwards translation method. In addition, two academic experts were asked to provide feedback on the translation for validity. Their feedback led to changes that helped to ensure the quality of the final questionnaire. The questionnaire was conducted on the students of Dokuz Eylül University, Faculty of Business, using convenience sampling.

The survey was conducted the second week after the midterms as students were aware of their midterm grades and the general teaching methods of the instructor (Özgüngör, 2010). 284 questionnaires were distributed and 282 were collected. The sample distribution between both genders (41% male and 59% female) was moderately equal in percentage. Out of the 282 observations, 48% of the coded instructors were male and 52% female (Professors, 30%; Associate Professors, 28%; Assistant Professors, 27%; and Lecturers, 13%).

The ages of the respondents were ranged from 18-30 (M=22.18), however 92% of the observations fell between the range of 20-24. Out of student responses 31% were Freshmen; 19% were Sophomore; 26% were Juniors and 22% were Seniors. From the

distribution of the respondents' expected end of year score, it could be said that the students were more willing to disclose their perceptions on instructors which they were expecting a higher end of year score; 25% of the respondents stated their expected score for the class was AA (90-100), 20% stated it to be BA (85-89), 12% as BB (80-84). Students stated that they viewed the overall performance of their instructor to be very successfully, 30%; somewhat successful, 33%; neutral, 25%; somewhat unsuccessful, 9%; and very unsuccessful, 2%.

Data Analysis

IBM SPSS version 20.0 was employed in interpreting the data. By analyzing the SETERS factor loadings that drive the overall rating of the instructor, the study attempted to provide information on the determinants of instructor ratings. As mentioned before, Toland and De Ayala (2005) proposed conducting further empirical study to support the validity of SETERS. By applying the scale to students studying in the Faculty of Business at Dokuz Eylül University, we investigate students' perception of teaching effectiveness and analyze the consistency of the original factor structure of SETERS by employing exploratory factor analysis. The main research question is *"Does the factor structure of SETERS show a similar structure with the Turkish students' sample?"* The constructs are developed and tested for reliability using the Cronbach's Alpha test.

In order to examine dimensions under the concept of effective teaching, it is imperative to analyze the structure of SETERS and determine its capability of successfully addressing the overall performance of the instructor. Therefore, the summated factor scores from the factor analysis are employed under the ordinal regression analysis to test the established hypothesis. The respondents were asked to answer on a 5-point scale (5=Very successful, 4=Somewhat successful, 3=Neutral, 2=Somewhat unsuccessful, 1=Very unsuccessful) depending on how satisfied they were with the instructor's performance in the classroom. The results are employed as a dependent variable under the ordinal regression for our first hypothesis;

Hypothesis 1: The factor structure of SETERS is influential in explaining the overall performance of the instructor.

F-test is employed and the detected relation was further analyzed to determine the same-gender relation of second hypothesis.

Hypothesis 2: There is no significant relationship between same-gender preferences of students and the evaluation of instructor's performance.

Hypothesis 2a: There is no significant relationship between students' gender and the evaluation of female instructor's performance.
Hypothesis 2b: There is no significant relationship between students' gender and the evaluation of male instructor's performance.

Assuming that variables have little influence on teaching evaluation, we tested our third and fourth hypothesis by use of correlation.

Hypothesis 3: There is no significant relationship between year-end expected score and evaluation of instructor's performance.
Hypothesis 4: There is no significant relationship between the instructor's level in the organization and evaluation of instructor's performance.

DISCUSSION OF FINDINGS

Data Screening Process

With a total sample size of 282, the ratio of observations to variables is approximately 8.3 and well above the minimum requirement of 8:1 (Hair, Black, Babin, & Anderson, 2010). An examination of all responses indicated missing data is below 10% for individual variables and of the overall data. No additional steps were taken to correct the discrepancy (Hair et al., 2010). Each variable was next screened for incorrectly entered information by use of frequency tables; any available responses were crosschecked via the coding manual. The data were further screened for outliers using the Mahalanobis Distance method ($\chi 2$ (34) = 65.247, $p < .05$). 12 variables that exceeded the critical value were detected and removed from the analysis, bringing the sample size down to 270.

Exploratory Factor Analysis

The Principal Component Analysis was employed under the study. Under the initial factor analysis, the KMO score was 96.6% accounting for a large amount of variance within the data and the Bartlett test was highly significant ($p<.001$). Also, the Pearson Correlation structure of the variables indicated one or two significant relations ($p<.001$) for each variable. Overall, the tests confirmed that the covariance structure is acceptable for a factor analysis. The Communalities were higher than 0.50 for all variables. SPSS listed three factors with Eigenvalues over 1 (20.354, 2.124 and 1.751 respectively). The three factors accounted for total variance of 71%. This coupled with a visual analysis of the Scree

Plot, supported the researchers' decision to separate the SETERS structure into three factors.

The Varimax rotation was employed as the results would further be used under an ordinal regression analysis. Five items were strongly cross-loaded. These were "IRFI" questions 8, 9, 10 and "IRRL" questions 1 and 2. The questionnaire items were removed and the analysis was re-run. Three factors were derived from the remaining 29 items. The new KMO score was 96.2% and the Bartlett test was highly significant ($p<.001$) as seen under Table 1.

Table 1: KMO and Bartlett's Test after deleting 5 items

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.962
Bartlett's Test of Sphericity	Approx. Chi-Square	7747.010
	df	406
	Sig.	.000

Table 2: Communalities after 5 item deletion

	Initial	Extraction
Instructor's Delivery of Course Information #1	1.000	.740
Instructor's Delivery of Course Information #2	1.000	.779
Instructor's Delivery of Course Information #3	1.000	.765
Instructor's Delivery of Course Information #4	1.000	.598
Instructor's Delivery of Course Information #5	1.000	.686
Instructor's Delivery of Course Information #6	1.000	.672
Instructor's Delivery of Course Information #7	1.000	.727
Instructor's Delivery of Course Information #8	1.000	.659
Instructor's Delivery of Course Information #9	1.000	.609
Instructor's Delivery of Course Information #10	1.000	.760
Instructor's Delivery of Course Information #11	1.000	.801
Instructor's Delivery of Course Information #12	1.000	.752
Facilitating Instructor/Student Interactions #1	1.000	.779
Facilitating Instructor/Student Interactions #2	1.000	.869
Facilitating Instructor/Student Interactions #3	1.000	.791
Facilitating Instructor/Student Interactions #4	1.000	.835
Facilitating Instructor/Student Interactions #5	1.000	.825
Facilitating Instructor/Student Interactions #6	1.000	.522
Facilitating Instructor/Student Interactions #7	1.000	.628
Instructor's Role in Regulating Students Learning #3	1.000	.598
Instructor's Role in Regulating Students Learning #4	1.000	.765
Instructor's Role in Regulating Students Learning #5	1.000	.774
Instructor's Role in Regulating Students Learning #6	1.000	.811
Instructor's Role in Regulating Students Learning #7	1.000	.540
Instructor's Role in Regulating Students Learning #8	1.000	.832
Instructor's Role in Regulating Students Learning #9	1.000	.785
Instructor's Role in Regulating Students Learning #10	1.000	.566
Instructor's Role in Regulating Students Learning #11	1.000	.797
Instructor's Role in Regulating Students Learning #12	1.000	.806

The new communalities were higher than 0.50 for all of the variables and the proportion of each variable's variance for the three factors increased to 73% as shown under Table 2 and 3, respectively.

The rotated component matrix of 29 items is presented under Table 4.

Table 4: Rotated Component Matrix: 29 items

	Component		
	1	**2**	**3**
Instructor's Delivery of Course Information #1	0.820		
Instructor's Delivery of Course Information #2	0.813		
Instructor's Delivery of Course Information #3	0.785		
Instructor's Delivery of Course Information #4	0.658		
Instructor's Delivery of Course Information #5	0.739		
Instructor's Delivery of Course Information #6	0.678		
Instructor's Delivery of Course Information #7	0.743		
Instructor's Delivery of Course Information #8	0.655		
Instructor's Delivery of Course Information #9	0.648		
Instructor's Delivery of Course Information #10	0.674		
Instructor's Delivery of Course Information #11	0.752		
Instructor's Delivery of Course Information #12	0.667	0.439	
Facilitating Instructor/Student Interactions #1			0.788
Facilitating Instructor/Student Interactions #2			0.875
Facilitating Instructor/Student Interactions #3			0.832
Facilitating Instructor/Student Interactions #4			0.791
Facilitating Instructor/Student Interactions #5			0.798
Facilitating Instructor/Student Interactions #6			0.435
Facilitating Instructor/Student Interactions #7			0.620
Instructor's Role in Regulating Students Learning #3		0.488	
Instructor's Role in Regulating Students Learning #4	0.493	0.644	
Instructor's Role in Regulating Students Learning #5		0.718	
Instructor's Role in Regulating Students Learning #6		0.747	
Instructor's Role in Regulating Students Learning #7		0.517	
Instructor's Role in Regulating Students Learning #8		0.824	
Instructor's Role in Regulating Students Learning #9		0.829	
Instructor's Role in Regulating Students Learning #10	0.509	0.546	
Instructor's Role in Regulating Students Learning #11		0.794	
Instructor's Role in Regulating Students Learning #12		0.808	
***Absolute value was limited as 0.40.**			

As it can be seen from Table 4, "IDCI" question 12 and " IRRL " questions 4 and 10 were cross-loaded. However, as the questionnaire items possessed strong theoretical foundation and were clearly listed under their primary loadings, they were retained under the analysis. Considering the first research question, *"Does the factor structure of*

SETERS show a similar structure with the Turkish students' sample?" the findings of exploratory factor analysis confirmed that the factor structure of SETERS is consistent with original structure of SETERS. The factor labels proposed by Toland and De Ayala (2005) suited the extracted factors and thus, were retained.

The reliability of the results was tested by use of the Cronbach's alpha test. The internal consistency for each of the three factors (.960; .951; .939) and the overall reliability (.976) of the SETERS was very high.

The inter-item correlation matrix showed that the alphas were very high. No substantial increases in alpha for any of the scales could have been achieved by eliminating more items, as the correlation values are above 0.40.

Ordinal Regression Analysis

The questionnaire items depicted under the three factors were summated using the original factor labels proposed by Toland and De Ayala (2005). These are IDCI, Factor 1; IRFI, Factor 2; and IRRL, Factor 3. The significant correlations between the factors are below 0.80 ($p<.001$). Our dependent variable is the overall performance of the instructor and the independent variables are the factors structured under SETERS. As previously depicted under Table 4, the dependent variable is ordinal in nature and thus, Ordinal Regression Analysis was employed to test the model.

The sample size remaining after the factor analysis was 270. The ratio of observations to variables is 68 to 1. The data was tested in accordance to the assumptions of the ordinal regression. The overall performance was positively correlated with IDCI (0.801), IRFI (0.612) and IRRL (0.700), however the variables aren't highly correlated, which would have otherwise lead to a multicollinearity problem within the analysis. This is provided under Table 5.

As seen on Table 6, the assumption of proportional odds was tested under the test of parallel lines. The model didn't present a significantly better fit to the data than the ordinal model 0.708 ($p>.05$), thus the regression assumption was not rejected. With a chi-square statistic of 273.518, ($p < .001$), the analysis indicated a significant improvement over the baseline intercept-only model as seen on Table 7.

Table 5: Correlations

		1	2	3	4
1. Overall Performance	Pearson Correlation	1	.801**	.612**	.700**
	Sig. (2-tailed)		.000	.000	.000
	N	270	270	270	270
2. Delivery	Pearson Correlation	.801**	1	.717**	.802**
	Sig. (2-tailed)	.000		.000	.000
	N	270	270	270	270
3. Facilitation	Pearson Correlation	.612**	.717**	1	.733**
	Sig. (2-tailed)	.000	.000		.000
	N	270	270	270	270
4. Regulation	Pearson Correlation	.700**	.802**	.733**	1
	Sig. (2-tailed)	.000	.000	.000	
	N	270	270	270	270
**. Correlation is significant at the 0.01 level (2-tailed).					

Table 6: Test of Parallel Lines

Model	-2Log Likelihood	Chi-Square	df	Sig.
Null Hypothesis	455.203			
General	448.886	6.317	9	.708
The null hypothesis states that the location parameters (slope coefficients) are the same across response categories.				
a. Link function: Logit.				

Table 7: Model Fitting Information

Model	-2Log Likelihood	Chi-Square	df	Sig.
Intercept Only	728.721			
Final	455.203	273.518	3	.000
Link function: Logit.				

The Nagelkerke fit of the model (.681) indicated an improvement over the null model with no predictors. We determined that the coefficient for IDCI and IRRL are significant ($p<.001$) and are influential in explaining the overall performance of the instructor. However, the significance of IRFL was above 0.05, thus, did not predict the dependent variable.

Based on the findings, the results to the first hypothesis *"Hypothesis 1: Factor structure of SETERS is influential in explaining the overall performance of the instructor"* is that IDCI and IRRL are significant predictors in explaining the instructor's performance whereas IRFI does not bear significance on the dependent variable.

F-Test and Correlation

We tested Hypothesis 2 through the measurement of Hypothesis 2a-b; by holding for the gender of the instructor and examining to see if there was a significant difference between male and female students' perceived evaluation of instructors. We see that there was no significant effect of student gender on the performance of female instructors; IDCI ($F_{1,140} = 0.947$, $p > .05$), IRFI ($F_{1,140} = 0.139$, $p > .05$) and IRRL ($F_{1,140} = 0.251$, $p > 0.05$). In regards to the relation between student gender and the performance of male instructors, we see that there was a slight trend towards significance on IDCI ($F_{1,125} = 8.264$, $p < .05$), however with IRFI ($F_{1,125} = 3.456$, $p > .05$) and IRRL ($F_{1,125} = 1.486$, $p > .05$) there are no significant effect of student gender on the performance of male instructors. As both Hypothesis 2a-b did not indicate meaningful significance between same-gender preferences of students on the perceived overall evaluation of instructors, thus, Hypothesis 2 is not rejected.

In regards to the relationship between the respondents' year-end expected scores and the SETERS measures, IDCI ($r = -0.29$ $p > .05$), IRFI ($r = -0.33$, $p > .05$) and IRRL ($r = -0.321$ $p > .05$) presented an insignificant, weak negative correlation. Also, the relationship between the instructor title and the SETERS measures, IDCI ($r = -0.30$ $p > .05$), IRFI ($r = -0.114$, $p > .05$) and IRRL ($r = -0.67$ $p > .05$) was found to be very weakly correlated and not significant. Thus, Hypothesis 3-4 are not rejected.

DISCUSSION AND CONCLUSION

Accompanying Bologna process, Turkish higher educational environment experiences reformation on overall aspect of teaching. Not only is it required to revise course format and course credits, but also the range of responsibility of teaching was extended from individual instructor to the entire educational institutions. Colleges and universities in Turkey, therefore, are increasingly interested in instructor's evaluation conducted by students. Business schools in Turkey are not exceptional in

this flow. Applying the traditional business concepts into business schools, improving the service quality based on customers' feedback should be always one of the most important attributes that business schools consider (LeBlanc & Nguyen, 1997). Previous researches on students' evaluation of teaching have been centered on constructing and confirming measures which are valid and reliable. Studies in Turkish educational context were also showing similar path, yet the researches are in the beginning stage.

Therefore, the purpose of this study was to provide empirical evidence supporting the use of teaching effectiveness measures in business faculty of a Turkish public university context. We hope that this study provides insights to academic members of business schools regarding how to satisfy customers, use the customer satisfaction results into better education quality, and further achieve competitive advantage among other business schools. The analysis confirmed that Toland and De Ayala (2005)'s three factor approach is appropriate in explaining the students' evaluation on teaching. The overall reliability of SETERS and the internal consistency for each of the constructed factors was supported by Cronbach's Alpha.

Addressing the hardship of defining "effective teaching", literature supports the argument that measures should be multidimensional (Abrami et al., 1996; Feldman, 1976; Marsh, 1987), in way of increasing the validity (Dodeen, 2013) of the study. Thus, the structure of SETERS was analyzed and it was determined that the sample students' perception on the Instructor's Delivery of Course Information and the Instructor's Role in Regulating Students Learning were predictors of the perceived overall performance of the instructor. Weimer (2007) suggested that a successful instructor is not only knowledgeable of the course content, but also aware of how to present this knowledge effectively to students. This is consistent with the process-product concept of effective teaching as suggested by Abrami et al. (1996). Instructor's regulating of learning focuses on consistent communication, feedback, encouragement of students' in learning. Allen (2008) proposed that giving positive feedback to students, and maintaining constructive relationship with students are key components to foster students' learning.

It was determined that same-gender preferences of students did not indicate meaningful significance on the evaluation of instructors' performance. Thus, the male (female) sample didn't present a tendency

to award higher ratings to male (female) instructors. Studies of Petridou and Sarri (2004) found no significant influence of students' gender and instructor's gender on evaluation ratings. Our studies' results are consistent with Petridou and Sarri, there was no indicator of same-gender preference found under the evaluation ratings. Therefore, faculties can consider this outcome when they set policy about employment of instructors.

Two other variables addressed under the study were the students' perception towards their expected year-end score and instructor's title. Both variables had no significant influence and didn't adversely affect the results of the SETERS measures for our sample. From the results for the students expected year-end score, it can be interpreted that students with low (high) expectations do not adversely punish (award) the instructor by awarding them lower (higher) grades. The results of the instructor's level suggest that Turkish students didn't consider academic hierarchy while evaluating instructor's performance. This is consistent with the findings of Özgüngör (2013). Turkish students do not perceive any difference on teaching performance whether the instructor is a lecturer, assistant professor, associate professor or full professor. Faculty should try to avoid prejudices such as full professors would receive highest ratings from students.

The application of student evaluation results should be handled carefully. Galbraith, Merrill, and Kline (2012) and Emery, Kramer, and Tian (2003) suggest inconsiderate usage of instructor effectiveness measures corrupted the teacher-student relationship. Also, Valsan and Sproule (2008) advised not to use the instructor effectiveness only for administrative purpose as the results could be distorted to maximize the evaluation score. Thus, the overall student evaluations of teaching efficiency measures should be understood as an aid in promoting students' engagement in learning and improving the quality of education.

There are several limitations of the study that should be noted. We collected data at a single business school with a relatively small number of students. We realize this sample is not a representative of the Turkish students and suggest caution against generalizing our findings for Turkish universities. A larger sample size could also offer an opportunity to study the influence of factors such as respondents' major, prior education (private/public), and attending preparatory class on the student's educational expectation.

It should be noted that research on evaluation of teaching efficiency in Turkey is currently in the beginning stage. Therefore, our

study was aimed to contribute to current area of educational research and attempted to provide insight and understanding to future researchers about the importance and prominence of evaluations of teaching efficiency in Turkey.

REFERENCES

Abrami, P. C., d'Apollonia, S., & Rosenfield, S. (1996). The Dimensionality of student ratings of instruction: what we know and what we do not. In J. C. Smart (Ed.), Higher education: handbook of theory and research (pp. 213-264). New York: Agathon.

Açan, B., & Saydan, R. (2009). Öğretim elemanlarının akademik kalite özelliklerinin değerlendirilmesi: Kafkas üniversitesi İİBF örneği. Atatürk Üniversitesi Sosyal Bilimler Enstitüsü Dergisi, 13(2).

Allen, S. (2008). 'Warming the climate for learning. Teaching Professor Report, 22 (2). Madison: Magna Publications.

Basow, S. A., & Silberg, N. T. (1987). Student evaluations of college professors: are female and male professors rated differently? Journal of Educational Psychology, 79(3), 308.

Brophy, J., & Good, T. L. (1984). Teacher behavior and student achievement. Occasional Paper (73).

Centra, J. A., & Gaubatz, N. B. (2000). Is there gender bias in student evaluations of teaching? Journal of Higher Education, 17-33.

Cohen, P. A., & Herr, G. (1979). Procedure for diagnostic instructional feedback-formative assessment of college-teaching (fact) model. Educational Technology, 19(12), 18-23.

Cohen, P. A. (1980). Effectiveness of student-rating feedback for improving college instruction: a meta-analysis of findings. Research in Higher Education, 13(4), 321-341.

d'Apollonia, S., & Abrami, P. C. (1997). Navigating student ratings of instruction. American Psychologist, 52(11), 1198-1208.

d'Apollonia, S., & Abrami, P. (1996, April). Variables moderating the validity of student ratings of instruction: a meta-analysis. Paper presented in 77th annual meeting of the American Educational Research Association, New York.

DeShields Jr, O. W., Kara, A., & Kaynak, E. (2005). Determinants of business student satisfaction and retention in higher education: applying Herzberg's two-factor theory. International Journal of Educational Management, 19(2), 128-139.

Dodeen, H. (2013). Validity, reliability, and potential bias of short forms of students' evaluation of teaching: the case of UAE university. Educational Assessment, 18(4), 235-250.

Emery, C. R., Kramer, T. R., & Tian, R. G. (2003). Return to academic standards: a critique of student evaluations of teaching effectiveness. Quality Assurance in Education, 11(1), 37-46.

Feldman, K. A. (1976). The superior college teacher from the students' view. Research in Higher Education, 5(3), 243-288.

Francis, C. A. (2011). Student Course Evaluations: Association with Pre-course Attitudes and Comparison of Business Courses in Social Science and Quantitative Topics. North American Journal of Psychology, 13(1).141-154.

Galbraith, C. S., Merrill, G. B., & Kline, D. M. (2012). Are student evaluations of teaching effectiveness valid for measuring student learning outcomes in business related classes? A neutral network and Bayesian analyses. Research in Higher Education, 53(3), 353-374.

Giovannelli, M. (2003). Relationship between reflective disposition toward teaching and effective teaching. The Journal of Educational Research, 96(5), 293-309.

Greenwald, A. G., and Gillmore, G. M. (1997). Grading leniency is a removable contaminant of student ratings. American psychologist, 52(11), 1209. (11): 1209-17.

Gülcan, Y., Kuştepeli, Y., & Aldemir, C. (2002). Yüksek öğretimde öğrenci doyumu: kuramsal bir çerçeve ve görgül bir araştırma. Süleyman Demirel Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi, 7(1), 99-114.

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). Multivariate data analysis. New Jersey: Pearson Education.

Hattie, J. A. C., & Learning, V. (2009). A synthesis of over 800 meta-analyses relating to achievements. New York.

Howard, G. S., & Maxwell, S. E. (1982). Do grades contaminate student evaluations of instruction?. Research in Higher Education, 16(2), 175-188.

Karaca, E. (2008). Eğitimde kalite arayışları ve eğitim fakültelerinin yeniden yapılandırılması. Dumlupınar Üniversitesi Sosyal Bilimler Dergisi, 21, 61-80.

LeBlanc, G., & Nguyen, N. (1997). Searching for excellence in business education: an exploratory study of customer impressions of service

quality. International Journal of Educational Management, 11(2), 72-79.

Letcher, D. W., & Neves, J. S. (2010). Determinants of undergraduate business student satisfaction. Research in Higher Education Journal, 6(1), 1-26.

Marsh, H. W. (1987). Students' evaluations of university teaching: research findings, methodological issues, and directions for future research. International Journal of Educational Research, 11(3), 253-388.

Marsh, H. W. (2007). Students' evaluations of university teaching: dimensionality, reliability, validity, potential biases and usefulness. In Perry, R. P., & Smart, J. C. (Ed.), The scholarship of teaching and learning in higher education: an evidence-based perspective (pp. 319-383). Netherlands: Springer.

Marsh, H. W., & Bailey, M. (1993). Multidimensional students' evaluations of teaching effectiveness: a profile analysis. Journal of Higher Education, 1-18.

Marsh, H. W., & Dunkin, M. J. (1992). Students' evaluations of university teaching: a multidimensional perspective. Higher education: handbook of theory and research, 8, 143-233.

Marsh, H. W., Hau, K. T., Chung, C. M., & Siu, T. L. (1998). Confirmatory factor analyses of Chinese students' evaluations of university teaching. Structural Equation Modelling: A Multidisciplinary Journal, 5(2), 143-164.

Marsh, H. W., & Roche, L. A. (1997). Making students' evaluations of teaching effectiveness effective: the critical issues of validity, bias, and utility. American Psychologist, 52(11), 1187.

Martínez-Gómez, M., Sierra, J. M. C., Jabaloyes, J., & Zarzo, M. (2011). A multivariate method for analyzing and improving the use of student evaluation of teaching questionnaires: a case study. Quality & Quantity, 45(6), 1415-1427.

Murray, H. G. (1983). Low-inference classroom teaching behaviors and student ratings of college teaching effectiveness. Journal of Educational Psychology 75: 138--149.

Nelson, J. P., & Lynch, K. A. (1984). Grade inflation, real income, simultaneity, and teaching evaluations. Journal of Economic Education, 21-37.

Onwuegbuzie, A. J., Witcher, A. E., Collins, K. M., Filer, J. D., Wiedmaier, C. D., & Moore, C. W. (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, 44(1), 113-160.

Özgüngör, S. (2010). Identifying dimensions of students' ratings that best predict students' self-efficacy, course value and satisfaction. Eurasian Journal of Educational Research (EJER), (38). 146-163.

Özgüngör, S. (2013). The relationship between instructor and course characteristics and students' perception of instructional quality. Procedia-Social and Behavioral Sciences, 93, 1324-1328.

Patrick, J., & Smart, R. M. (1998). An empirical evaluation of teacher effectiveness: the emergence of three critical factors. Assessment & Evaluation in Higher Education, 23(2), 165-178.

Petridou, E., & Sarri, K. (2004). Evaluation research in business schools: students' rating myth. The International Journal of Educational Management, 18(3), 152-159.

Renaud, R. D., & Murray, H. G. (2005). Factorial validity of student ratings of instruction. Research in Higher Education, 46(8), 929-953.

Santiago, P., & Benavides, F. (2009, December). Teacher evaluation: a conceptual framework and examples of country practices. Paper presented at OECD-Mexico Workshop 'Towards a Teacher Evaluation Framework in Mexico: International Practices, Criteria and Mechanisms', Mexico City.

Slate, J. R., LaPrairie, K. N., Schulte, D. P., & Onwuegbuzie, A. J. (2011). Views of effective college faculty: A mixed analysis. Assessment & Evaluation in Higher Education, 36(3), 331-346.

Sprinkle, J. E. (2009). Student Perceptions of Educator Effectiveness: A Follow-Up Study. College Student Journal, 43(4).1341-1356.

Swartz, C. W., White, K. P. & Stuck, G. B. (1990). The factorial structure of the North Carolina teacher performance appraisal instrument. Educational and Psychological Measurement, 50(1), 175–185

Tran, N. D. (2015). Reconceptualisation of approaches to teaching evaluation in higher education. Issues in Educational Research, 25(1), 50-61.

Toland, M. D., & De Ayala, R. J. (2005). A multilevel factor analysis of students' evaluations of teaching. Educational and Psychological Measurement, 65(2), 272-296.

Valsan, C., & Sproule, R. (2008). The invisible hands behind the student evaluation of teaching: the rise of the new managerial elite in the

governance of higher education. Journal of Economic Issues, 42(4), 939-958.

Watkins, D. (1994). Student evaluations of university teaching: a cross-cultural perspective. Research in Higher Education, 35(2), 251-266.

Weimer, M. (2007). Content knowledge: a barrier to teaching development in effective strategies for improving college teaching and learning. The Teaching Professor Report 22 (6). Madison: Magna Publications.

Yüksek Öğretim Kurulu. (2016). Bologna Process. Retrieved from http://www.yok.gov.tr/web/guest/anasayfa (09.05.16)