

影響學生教學評鑑結果之教師因素

賴宜弘* 黃芬芬** 楊雪華***

摘要

本研究主要探討哪些因素會影響學生教學評鑑結果。本研究所要探討的因素包含學術價值、教師之教學熱誠、教學明晰度、團體互動、師生關係、課程廣度、成績評量方法、作業、班級人數、教師年齡、教師性別與教師職別。本研究主要以階層線性模式，分析 172 班 5178 名學生評量教師教學績效的資料，以瞭解個人因素與班級因素的相對重要性。由研究結果可以發現，學術價值、教師之教學熱誠、教學明晰度、師生關係、課程廣度與作業對學生教學評鑑結果皆有顯著正向影響。女性教師與高階教師可以在學生教學評鑑獲取較高的分數。教師的年紀與級職對於教學明晰度與學生教學評鑑結果的交互作用有正向影響；班級人數對團體互動與學生教學評鑑結果的交互作用有負向影響。授課老師應於課堂上讓學生充份瞭解該科目的學術價值、展現教學熱誠、組織調理的講授內容、友善的師生關係、適度的課程廣度與良好規劃的家庭作業皆可提升生教學評鑑結果。資深教師與高階老師對於教學明晰度與學生教學評鑑結果的交互作用表現較佳；較多班級人數對團體互動與學生教學評鑑結果的交互作用有負向影響。

關鍵字：教學評鑑、學生回饋教學、教學績效

* 亞東技術學院 醫務管理系 助理教授
** 亞東技術學院 醫務管理系 副教授
*** 亞東技術學院 醫務管理系 教授



Exploring the Teacher's Factors Influencing the Results of Students' Evaluation of University Teaching

Yi-Horng Lai * Fen-Fen Huang ** Hsieh-Hua Yang ***

Abstract

The purposes of this study was to investigate the relationship of academic value, enthusiasm for teaching, clarity, interaction, relationship, curriculum breadth, methods of assessment, assignments, class size, the age of teacher, the gender of teacher, the rank of teacher, and students' evaluation of teaching quality. Student rating for teaching was analyzed in relation to individual-class and class-level factors in a sample of 5178 students in 172 classes in order to survey the relative importance of the factors of class-level and student-level in the students' evaluation of teaching quality. With hierarchical linear modeling, it could be find that individual measures accounted for most of the variance. With the individual-level, more academic value, enthusiasm for teaching, clarity, the relationship of teacher and student, course breadth, and good in grading method, would upgrade the score of students' evaluation of teaching quality. With the class-level, the score of students' evaluation of teaching quality of female teacher was higher than male teacher, and higher rank of teachers' score of students' evaluation of teaching quality was higher than lower rank of teachers. Besides, the old teacher or higher rank of teacher was helpful in lecture clearly for getting high score of students' evaluation of teaching quality. Larger class was not helpful in relationship of teachers and students for higher students' evaluation of teaching quality. Teachers should make students understanding the academic value of the course with a clear way and enthusiasm for teaching. Experienced teachers or high rank teacher could get higher score of students' evaluation of teaching quality than other teachers. Too many students in one class were not helpful in the interaction for students' evaluation of teaching quality.

Keywords: teaching evaluation, student feedback to teaching, teaching performance

* Assistant Professor, Department of Health Care Administration, Oriental Institute of Technology, the corresponding author

** Associate Professor, Department of Health Care Administration, Oriental Institute of Technology

*** Professor, Department of Health Care Administration, Oriental Institute of Technology



INTRODUCTION

The measuring of the teaching quality was an important topic in education. For the Ministry of Education policy in Taiwan, the establishment of educational quality control mechanisms had become the urgent task for the universities. Students' evaluation of teaching effectiveness (SETE) was one of most popular teaching evaluation approach in Taiwan.

Most studies (McKeachie, 1990; Marsh & Roche, 1997) pointed out class size play an important role in students' evaluation of teaching quality. But some studies (Fernández, Mateo, & Munizñ, 1998) didn't support this point. Kaschak's study (1978) found that university students' evaluation scores of female teachers were lower than male teachers, but Elmore & LaPointe's study (1975) found that the student evaluation in male and female was similar.

More and more studies suggest that hierarchical linear modeling was better than linear regression analysis in the sample with multi subgroup (such as class, school, department, and team). In order to survey the relative importance of the factors of class-level and student-level in the students' evaluation of teaching quality, the main object of this study focus on university students, and explored the effect of feedback in different background of students and different class with hierarchical linear modeling.

The Questionnaire for SETE

SETE has generally been implemented in universities in Taiwan. But the education sector and academia has been controversial in reasonableness and applicability of SETE (d'Apollonia & Abrami, 1997; Marsh & Roche, 1997; Mckeachie, 1990; Schellhase, 2010). Some Taiwanese teachers don't like SETE for mental and campus ethics. Some teachers believe that SETE would infringe on the autonomy of teachers teaching, contrary to the spirit of respecting, cause mutual suspicion among teachers, and lower requirements of teachers to students. Some people that agree with SETE point out that the myth of the teachers' authority could be broken by SETE, and teaching quality can be improved.

SETE has been promoted for a long time, and the Students' Evaluation of Educational Quality (SEEQ) is the most popular questionnaire for SETE. SEEQ was developed by Herbert Marsh, and it is usually used for the assessment of teaching quality and individual curriculum model. The reliability and validity of SEEQ was recognized in many studies (Marsh & Roche, 1992; Coffey & Gibbs, 2000; Morley, 2011; Ibrahim, 2012; Lidice & Saglam, 2013). There are 35 closed quizzes for measuring the teach efficiency from nine faces (Marsh, 1982). The nine faces of SEEQ were



academic value, enthusiasm for teaching, clarity, interaction, relationship, curriculum breadth, methods of assessment, assignments, and overall. The SEEQ was usually executed by a person other than teachers in the last week of the course. SEEQ commonly used in universities in North America, now.

Class size and SETE

Class size was one of the important factors in class-level variables in SETE research. It is logical to speculate that students in larger class were given less opportunity to interact with the instructor. It is also quite likely that the instructor requires less homework and spends more time lecturing and less time in discussion in a large class (McKeachie, 1990). March and Roche (1997) point out that class size is moderately correlated with group interaction and individual link in SEEQ (Marsh, 1982) scales, and class size is nearly uncorrelated with other SEEQ scales.

Fernández, Mateo, and Munizñ (1998) stated that the relationship between class size and student ratings of teaching quality was weak and nonlinear. Although there were seem to be a relationship between class size and students' ratings, the magnitude of class size effect is small.

For medical learning by problem-based learning (PBL), small group format was better in academic value, tutor factors, group interaction, e-learning support, organizational factors, personal study time/hours, and No. of learning resources used. (Roberts, Lawson, Newble, Self, & Chan, 2005)

The instructor and SETE

It is inconsistent in different study in the relationship of SETE and gender. Kaschak's study (1978) found that university students' evaluation scores of female teachers were lower than male teachers. Elmore & LaPointe's study (1975) found that the different of the student evaluation in male and female was not significant. Feldman (1983) found that student satisfaction of teaching was positive with teachers' rank. After Centra (1981) analyzed 8000 different students, he point out that students' evaluation was close to teaching experience. In Civian & Brennan's study (1996), the relationship of teaching experience and students' evaluation was curve relationship, and the best pried is 3-years to 12-year. Feldman (1983) found that student satisfaction of teaching was negative with the age of teaching.



METHODOLOGY

The conceptual framework of this study is shown on Figure 1. In order to investigate the hypotheses of this study and measured the multi-level model for the students' evaluation of teaching quality. The student-level was built with Students' Evaluation of Educational Quality (SEEQ) (Marsh, 1982), and academic value, enthusiasm for teaching, clarity, interaction, relationship, curriculum breadth, methods of assessment, and assignments were evaluated in this level. The class-level (class size, the age of teacher, the gender of teacher, and the rank of teacher) factors as well as their cross-level interaction on the students' evaluation of teaching quality are included in the model.

Hypotheses of this study were as Figure 1. The Hypothesis 1 was the student-level factors that measuring with SEEQ (such as academic value, enthusiasm for teaching, clarity, interaction, relationship, curriculum breadth, methods of assessment, and assignments) are positive with students' evaluation of teaching quality. The Hypothesis 2 was the class-level factors (such as class size, the age of teacher, the gender of teacher, and the rank of teacher) are positive with students' evaluation of teaching quality. The Hypothesis 3 was the class-level factors are positive with the relation of students-level factors and students' evaluation of teaching quality.

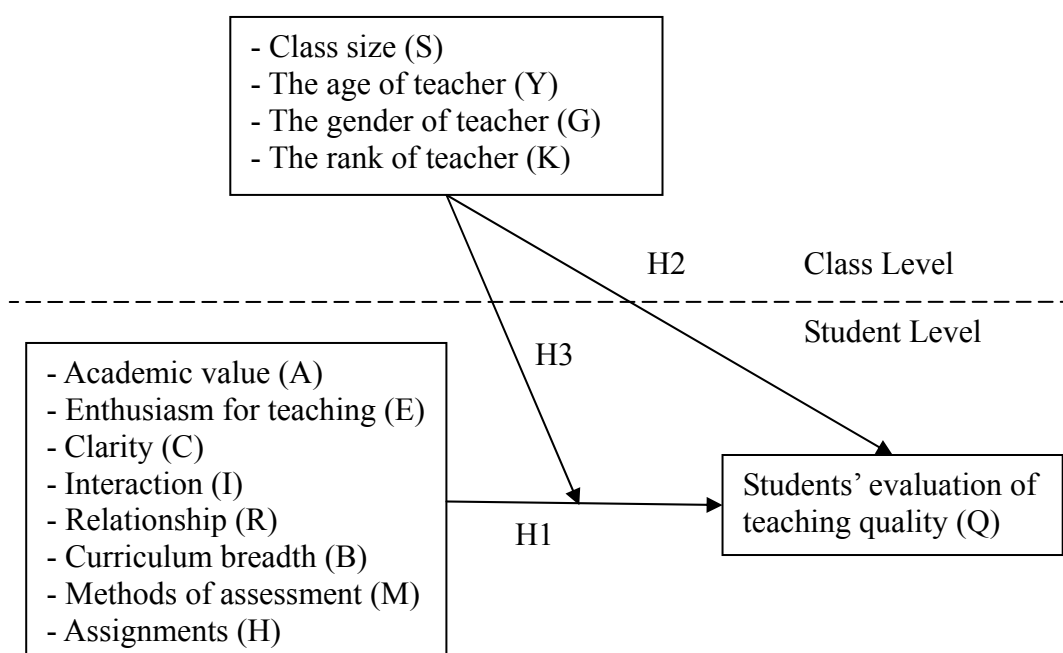


Figure 1: Research Framework



Research Data

The research data in this study was collected by C. J. Huang for the project: Can students evaluate the teaching quality of teacher? This project was finished on October, 31, 2011. The data could be obtained from Survey Research Data Archive (SRDA) (Huang, 2003). In Huang's survey (2003), there were 5574 feedbacks been collected in a public university of education from August 1, 2001 to July 1, 2002. This study focus on Huang's research data that without missing value. A total of 5178 university students' feedbacks were collected (Table 1) in 172 classes (Table 2). There were 2167 male students (41.83%) and 3011 female students (58.17%). Most of them were college of education students (26.07%), and most of them were first grade students (31.87%).

Table 1: Data Summarize of Students Level

Geographic		Frequency	Percent (%)
Gender	Male	2167	41.83
	Female	3011	58.17
College	Arts	1172	22.63
	Education	1350	26.07
	Science	1229	23.74
	Technology and Vocational Education	1251	24.16
	Business	82	1.58
	Social Science and Physical Education	11	0.21
	Management	19	0.37
	Engineering	64	1.24
	Grade	First grade	1650
	Second grade	1580	30.51
	Third grade	986	19.04
	Fourth grade	565	10.91
	Education courses	76	1.47
	Master	315	6.08
	Ph.D.	6	0.12
Total		5178	100.0

In the 172 classes, 91 classes were male lecturer (52.91%), and 81 classes were female lecturer (47.09%). Most of them were associate professor (43.60%). The code of female instructors was 0, and the code of male instructors is 1. The codes of lecturer, assistant professor, associate professor, professor are 1, 2, 3, and 4.



Table 2: Data Summarize of Class Level

Geographic		Frequency	Percent (%)
The gender of teacher	Male (1)	91	52.91
	Female (0)	81	47.09
The rank of teacher	Professor (4)	63	36.63
	Associate professor(3)	75	43.60
	Assistant professor (2)	3	1.75
	Lecturer (1)	31	18.02
Total		172	100.0

Research Tools

A questionnaire was administered to evaluate the students' perceptions of the quality of teaching based on a modification of the SEEQ (Marsh, 1982). Psychometric characteristics of SEEQ have been well established. It has a high degree of internal consistency (0.917 to 0.779) and it has reasonable levels of validity in that scale scores correlate significantly with a wide range of measures of learning outcome (Coffey & Gibbs, 2001). There are 8 parts in the research questionnaire with 9 scales. All variables exhibit a high level of reliability with the Cronbach's alpha values (as Table 3) exceeding the recommended 0.6 (Nunnally, 1978).

Table 3: The reliability of the questionnaire in this study

	N of Items	Cronbach's α
Academic value (A)	4	0.863
Enthusiasm for teaching (E)	4	0.876
Clarity (C)	4	0.882
Interaction (I)	4	0.917
Relationship (R)	4	0.860
Curriculum breadth (B)	4	0.847
Methods of assessment (M)	3	0.779
Assignments (H)	2	0.881

Data Analysis

For exploring the factors of class-level and student-level in the students' evaluation of teaching quality, the data analysis methodology in this study was hierarchical linear modeling (HLM) with restricted maximum likelihood with SSI HLM 6.08 computer software in this study.

HLM or multi-level modeling analysis was used to analyze the relationships between an individual analysis and group analysis. This analysis was using many variables to adjust the regression of base level dependent variables on base level independent variables. The objective of this technique was to analyze many variables by using the hierarchical structure (Hofman, 1997). It



would be adopted the HLM method as described by Bryk and Raudenbus (1992) and tested the model in four steps. First, it would be estimated a null model that had no predictors from either individual level or the group level. Null model was required to separate the students' evaluation of teaching quality variance in to within-team and between-team components. In the Second step, which was level 1 analysis of individual students' evaluation of teaching quality, students' evaluation of teaching quality was regressed on grand-mean-centered individual level predictor. In third step or the level 2 analysis, it be used the intercept estimates obtained from level 1 as outcome variables and regressed these on the group level predictors. The last step is corresponding hypothesis, hence the cross level interactions were tested.

RESULTS

For testing the hypotheses in this study (Figure 1), HLM been used for data analysis in this study. This study explored the relationship between the cross-level variables with HLM.

Null Model

This study explored the relationship of students' personal opinion factors and class factors, and detecting the effect of cross-level with HLM method. The result of data analysis was as Table 4 and Table 5. The different between each class with null model were significant.

With the result of the null model in this study, it could be found that between group components (τ_{00}) was significant ($\chi^2=3181.711$; $df=170$; $P\text{-value}<0.001$), and $ICC(1)$ was 0.3825. It mean 38.25% of variance was between group, and 61.75% of variance was within group. $ICC(2)$ was 0.938. It could be found that the $ICC(1)$ of this model was larger than 0.138 and the $ICC(2)$ of this model was larger than 0.7 (Cohen, 1988). The different of students' evaluation of teaching quality between each class was significant, and the other factor could be detected with next step.

Random Coefficient Model

Random coefficient model was the model with all individual level variables. It mean a regression model with all individual level variables in individual level and null model in group level. The relationship of dependent variables and independent variables in individual level could be detected with random coefficient model.

The result of random coefficient model was as Table 4 and Table 5. It could be found that the



relationship of academic value (A) (t-ratio=6.940; df=170; P-value<0.001), enthusiasm for teaching (E) (t-ratio=16.145; df=170; P-value<0.001), clarity (C) (t-ratio=5.221; df=170; P-value<0.001), relationship (R) (t-ratio=6.937; df=170; P-value<0.001), curriculum breadth (B) (t-ratio=2.869; df=170; P-value=0.005), and assignments (H) (t-ratio=2.643; df=170; P-value=0.009) with students' evaluation of teaching quality were significant, but interaction (I) (t-ratio=0.631; df=170; P-value=0.529), and methods of assessment (M) (t-ratio=1.353; df=170; P-value=0.178).

With the result, more academic value, enthusiasm for teaching, clarity, teacher-student relationship, course breadth, and good in grading method, would upgrade the score of students' evaluation of teaching quality.

Contextual Model

The percent of the direct effects that can be explained by intercept variance could be detected with contextual model.

The result of contextual model was as Table 4 and Table 5. It could be found that the score of students' evaluation of teaching quality could be effected by the gender of teacher (G) (t-ratio=-2.723; df=166; P-value=0.008), and the rank of teacher (K) (t-ratio=3.478; df=166; P-value=0.001) significantly, and class size (S) (t-ratio=-0.626; df=166; P-value=0.532), and the age of teacher (Y) (t-ratio=-1.269; df=166; P-value=0.206) were not significant.

With the result, the score of students' evaluation of teaching quality of female teacher was higher than male teacher, and higher rank of teachers' score of students' evaluation of teaching quality was higher than lower rank of teachers.

Full Model

Table 4: The Result of Hierarchical Linear Model - Fixed Effect

		Null Model	Random Coefficient Model	Contextual Model	Full Model
I	γ_{00}	6.701 (0.081)*	6.702* (0.081)	6.935* (0.114)	6.942* (0.117)
S	γ_{01}	-	-	-0.004 (0.005)	-0.003 (0.005)
Y	γ_{02}	-	-	-0.013 (0.010)	-0.019 (0.011)
G	γ_{03}	-	-	-0.442* (0.162)	-0.456* (0.169)
K	γ_{04}	-	-	0.255* (0.073)	0.254* (0.077)
A	γ_{10}	-	0.122* (0.018)	0.123* (0.018)	0.146* (0.028)
A	γ_{11}	-	-	-	-0.001 (0.001)
A	γ_{12}	-	-	-	0.001 (0.002)
A	γ_{13}	-	-	-	-0.037 (0.040)
A	γ_{14}	-	-	-	-0.003 (0.017)
E	γ_{20}	-	0.312* (0.019)	0.313* (0.019)	0.320* (0.031)



E	γ_{21}	-	-	-	-0.001 (0.001)
E	γ_{22}	-	-	-	0.001 (0.003)
E	γ_{23}	-	-	-	-0.010 (0.044)
E	γ_{24}	-	-	-	-0.023 (0.019)
C	γ_{30}	-	0.114* (0.022)	0.113* (0.022)	0.145* (0.033)
C	γ_{31}	-	-	-	0.001 (0.002)
C	γ_{32}	-	-	-	0.006* (0.003)
C	γ_{33}	-	-	-	-0.054 (0.048)
C	γ_{34}	-	-	-	0.055* (0.022)
I	γ_{40}	-	0.010 (0.016)	0.010 (0.016)	0.021 (0.026)
I	γ_{41}	-	-	-	-0.001 (0.001)
I	γ_{42}	-	-	-	-0.001 (0.002)
I	γ_{43}	-	-	-	-0.021 (0.036)
I	γ_{44}	-	-	-	-0.005 (0.016)
R	γ_{50}	-	0.146* (0.021)	0.145* (0.021)	0.130* (0.032)
R	γ_{51}	-	-	-	-0.003* (0.002)
R	γ_{52}	-	-	-	-0.003 (0.003)
R	γ_{53}	-	-	-	0.049 (0.046)
R	γ_{54}	-	-	-	-0.009 (0.020)
B	γ_{60}	-	0.051* (0.018)	0.054* (0.018)	0.004 (0.027)
B	γ_{61}	-	-	-	0.002 (0.001)
B	γ_{62}	-	-	-	0.001 (0.002)
B	γ_{63}	-	-	-	0.079 (0.038)
B	γ_{64}	-	-	-	-0.011 (0.017)
M	γ_{70}	-	0.020 (0.015)	0.021 (0.015)	0.024 (0.023)
M	γ_{71}	-	-	-	-0.001 (0.001)
M	γ_{72}	-	-	-	-0.001 (0.002)
M	γ_{73}	-	-	-	0.008 (0.033)
M	γ_{74}	-	-	-	-0.034 (0.015)
H	γ_{80}	-	0.037* (0.014)	0.037* (0.014)	0.053* (0.022)
H	γ_{81}	-	-	-	0.003* (0.001)
H	γ_{82}	-	-	-	0.001 (0.002)
H	γ_{83}	-	-	-	-0.056 (0.032)
H	γ_{84}	-	-	-	-0.008 (0.014)

1: * p-value<0.05. 2: () is Standard deviation.

The result of full model was as Table 4 and Table 5. It could be found that the age of teacher was positive with the correction of clarity (C) and students' evaluation of teaching quality (γ_{32}) (t-ratio=2.048; df=166; P-value=0.042), the rank of teacher (K) was positive with the correction of clarity (C) and students' evaluation of teaching quality (γ_{34}) (t-ratio= 2.531; df=166; P-value=0.013), and class size (S) was negative with the correction of relationship of teachers and students and students' evaluation of teaching quality (γ_{51}) (t-ratio=-2.041; df=166; P-value=0.043).

With the result, the old teacher or higher rank of teacher was helpful in lecture clearly for getting high score of students' evaluation of teaching quality. Larger class was not helpful in relationship of teachers and students for higher students' evaluation of teaching quality.



Table 5: The Result of Hierarchical Linear Model - Random Effect

		Null Model	Random Coefficient Model	Contextual Model	Full Model
	τ_{00}	1.057*	1.092*	0.955*	0.955*
A	τ_{11}	-	0.009	0.010	0.010
E	τ_{22}	-	0.007	0.007	0.008
C	τ_{33}	-	0.026*	0.025*	0.024*
I	τ_{44}	-	0.006	0.006	0.007
R	τ_{55}	-	0.026*	0.025*	0.023*
B	τ_{66}	-	0.013	0.013	0.011
M	τ_{77}	-	0.005	0.005	0.005
H	τ_{88}	-	0.006	0.006	0.006
	σ^2	1.708	0.927	0.927	0.926
-2LL		17960.865	15239.387	15237.175	15448.300

1: * p-value<0.05.

DISCUSSION AND CONCLUSION

The main purpose of this study was to adopt the multilevel methodology to join both class-level and student-level variables in the analysis to resolve the problem. Based on the result of this study, the students' evaluation of teaching quality of each class was different with different class-level and individual-level variables.

With the student-level, more academic value, enthusiasm for teaching, clarity, teacher-student relationship, course breadth, and good in grading method, would upgrade students' evaluation of teaching quality.

With the class-level, the score of students' evaluation of teaching of female teacher was higher than male teacher, and higher rank of teachers' score of students' evaluation of teaching quality was higher than lower rank of teachers. Besides, the experienced teacher or higher rank of teacher was helpful in lecture clearly for getting high score of students' evaluation of teaching quality. Larger class was not helpful in relationship of teachers and students for higher students' evaluation of teaching quality, and this result the same as Marsh (1982) and Marsha, Muthén, Asparouhov, Lüdtke, Robitzsche, Morin, & Trautwein, (2009).

ACKNOWLEDGEMENT

This study is based in part on data from the Survey Research Data Archive (SRDA) provided by the Academia Sinica. The interpretation and conclusions contained herein do not represent those of Survey Research Data Archive (SRDA) or Academia Sinica.



REFERENCES

- Bryk, A. S. & Raudenbus, S. W. (1992). *Hierarchical linear models*. Newbury Park: CA:Sage.
- Coffey, M. & Gibbs, G. (2000). The evaluation of the student evaluation of educational quality (SEEQ) questionnaire in UK higher education. *Assessment and Evaluation in Higher Education*, 26(1), 89-93.
- Coffey, M. & Gibbs, G. (2001) Reliable student feedback: a replication of the factor structure of the SEEQ. *Assessment and Evaluation in Higher Education*, 26, 89-93.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed). Hillsdale, NJ: Erlbaum.
- d'Apollonia, S. & Abrami, P. C. (1997). Navigating student ratings of instruction. *American Psychologist*, 52, 1198-1208.
- Fernández, J., Mateo, M. A., & Munizñ, J. (1998). Is there a relationship between class size and student ratings of teaching quality? *Educational and Psychological Measurement*, 58, 596-604.
- Hofman, D. A. (1997). An overview of the logic and rationale of hierarchical linear models. *Journal of Management*, 23(6), 723-744.
- Huang, C.J. (2003). Can students evaluate the teaching quality of teacher? *Survey Research Data Archive*. Retrieved July 18, 2012 from <https://srda.sinica.edu.tw/search/advsciitem/858>.
- Ibrahim, M. (2012). A Psychometric Evaluation of Two Teaching Effectiveness Scales. *Journal of Sustainable Development*, 5(7), 91-103.
- Lidice, A., & Saglam, G. (2013). Using Students' Evaluations to Measure Educational Quality. *Procedia - Social and Behavioral Sciences*, 70, 1009 - 1015.
- Marsh, H.W. (1982). SEEQ: a reliable and useful instrument for collecting student evaluations of university teaching. *British Journal of Educational Psychology*, 52, 77-95.
- Marsha, H.W., Muthén, B., Asparouhov, T., Lüdtke, O., Robitzsche, A., Morin, A.J.S., & Trautwein, U. (2009). Exploratory Structural Equation Modeling, Integrating CFA and EFA: Application to Students' Evaluations of University Teaching. *Structural Equation Modeling: A Multidisciplinary Journal*, 16(3), 439-476.
- Marsh, H.W., & Roche, L.A. (1992). The use of student evaluations of university teaching in different settings: the applicability paradigm. *Australian Journal of Education*, 36, 278-300.
- Marsh, H.W., & Roche, L.A. (1997). Making students' evaluations of teaching effectiveness effective: The critical issues of validity, bias, and utility. *American Psychologist*, 11,



1187-1197.

McKeachie, W. J. (1990). Research on college teaching: The historical background. *Journal of Educational Psychology*, 82, 189-200.

Morley, D.D. (2011). Claims about the reliability of student evaluations of instruction: The ecological fallacy rides again. *Studies in Educational Evaluation*, 38(1), 15-20.

Nunnally, J.C. (1978). *Psychometric Theory*. New York: McGraw Hill.

Roberts, C., Lawson, M., Newble, D., Self, A., & Chan, P. (2005). The introduction of large class problem-based learning into an undergraduate medical curriculum: an evaluation. *Medical Teacher*, 27(6), 527–533.

Schellhase, K.C. (2010). The Relationship Between Student Evaluation of Instruction Scores and Faculty Formal Educational Coursework. *Athletic Training Education Journal*, 5(4), 156-164.

