Student evaluations of instructors in higher education: A structural equation modeling

Hamzeh Dodeen
United Arab Emirates University

ABSTRACT

The study proposed a structural equation model (SEM) that hypothesizes the directional influence or causal relationships of six factors in the student's evaluations of higher education instructors. The proposed SEM (Model 1) considered teacher characteristics to be exogenous construct while course characteristics (organization/clarity, grading /fairness, workload/difficulty, and teaching performance) mediators, and students' perceived learning/interest to be dependent variable. Data for this study were collected from the students attending the United Arab Emirates University (UAEU) in two random samples with a total of 2298 students. The results of the five goodness of fit indices indicated that the data fit the developed model well. Teacher characteristics were mediated by course organization, teaching performance, and grading/fairness. However, the effect of course loading/difficulty was negative and small. This factor is neither influenced by the teacher characteristics nor it affects the perceived learning/interest. Thus, a modification was made on Model 1 by deleting the factor course loading/difficulty. Results indicated that the modified model (Model 2) fits the data well and important improvements were observed on the path coefficients between the remaining factors.

Keywords: students' evaluations, student evaluations of instructors, teaching effectiveness, teaching evaluation, structural equation model.

INTRODUCTION

The use of student evaluations of instructors has become common in higher education institutions (Davies, Hirschberg, Lye, Johnston, & McDonald, 2005; Feldman, 1997; Harrison, Douglas, & Burdsal, 2004; Neumann, 2000; Paswan & Young 2002; Rowley, 2003). These evaluations often are important and may be the only measure of teaching effectiveness (d'Apollonia & Abrami, 1997). The two main goals for using student evaluations of instructors in most universities are (1) to help administrators evaluate faculty members especially in reference to tenure, promotion, merit pay, salary increase, or contract renewal and (2) to provide faculty with feedback that has the potential to improve teaching effectiveness (Johnson, 2000; Ngware, 2005; Roberts, Irani, Telg & Lundy, 2005).

Students evaluations of instructors typically are multidimensional (Drago, Peltier, & Sorensen, 2002; Gage & Berliner, 1992; Gursoy & Umbreit, 2005; Huitt, 1995; Marsh, 2001; Paswan &Young, 2002). However, the number or the nature of the dimensions varies. For example, some researchers identified three dimensions namely, instructor delivery of course information, instructor-student interaction, and regulation of students' learning (Toland & Alyala, 2005). Others identified nine dimensions: learning/value, enthusiasm, organizing, group interaction, relationship with students, extent of coverage, examination and grading, assignments, and workload/difficulty (Marsh & Roche,1997). Still others identified six factors: relationship with students, course value, organization, grading, difficulty, and workload (Jackson et al., 1999); or course value, course difficulty, grading quality, course organization and design, work load and rapport with students (Harrison, Douglas, & Burdsal,2004). Some identified four factors: learning, instruction, organization, and workload (Gursoy & Umbreit, 2005). Some identified five factors: knowledge and performance in teaching, grading, overview of the course, requirements, and course outcomes (Badri, Abdulla, Kamali, & Dodeen, 2006). The commonalities of these studies seemingly focuses on four factors: organization, workload, grading, and learning.

Studies have examined interrelationships between the factors found in student evaluations. However, studies used statistical methods that examine only a single relationship at a time such as factor analysis, correlation, multiple regression, or multivariate analysis of variance (Paswan & Young, 2002). If the relationship between teaching and learning is direct, then we could easily conclude that good instruction results in good learning. Thus, teaching could be evaluated directly by the degree of learning. However, many intervening variables separate teaching and learning making this relationship more complicated. The use of structural equation modeling (SEM) provides a process that leads to a better understanding between teaching and learning (Stringer & Irwing, 1998). The use of SEM allows us to identify the relationships among factors included in the model and to establish direct and indirect effects of each factor (Rugutt & Chemosit, 2005). The main goal of this study is to utilize SEM to help explain interrelationships between factors found in student evaluations of instructors and to validate the proposed model through the use of empirical data.

LITREATURE REVIEW

Several studies used SEM to understand the underlying relationships between different factors found on student evaluations of instructor. Stringer & Irwing (1998) used SEM to examine relationships between teacher, course, and student characteristics among a sample of 1708 full-time undergraduate students taking health and social science courses. Students

completed a 25-item six factor teaching effectiveness questionnaire. The six factors were divided into four categories: overall evaluation of teaching (teaching quality); course (feedback, workload, and course integration); students' psychological responses (stimulation/learning); and the outcome measure. Teaching quality was predicted to influence course characteristics, which in turn were predicted to influence student stimulation/learning. The overall course evaluation was predicted to be determined by students' stimulation and learning. In a similar study, Marks (2000) used SEM with a sample of 2200 students enrolled in business courses who completed an instructor evaluation form consists of five factors: course organization, course workload, grading, liking instructor/ concern, and perceived learning. The developed model used workload/difficulty, liking instructor, and organization as exogenous factors and used grading, overall evaluation, and perceived learning as endogenous factors. Marks found that the instructor personality strongly influenced student rating and that organization has positive effect on learning.

Paswan & Young (2002) used SEM with a sample of 2059 students who completed a five factor instructor evaluation form: instructor involvement, student interest, student-instructor interaction, course demands, and course organization. In this SEM, three input factors (student-instructor interaction, course demands, and course organization) represented the course design input and influenced the two endogenous factors (instructor involvement and student interest). Results indicated that course organization and student-instructor interactions influence student interest and instructor involvement in a positive manner while course demand affects these endogenous factors in a negative manner.

Gursoy & Umbreit (2005) proposed a SEM in which organization, workload, and instruction were the exogenous constructs, and the learning construct was the dependent variable. The model was tested on data from a School of Hospitality Business Management. The results indicated that all three constructs have a significant positive impact on students' perception of learning. Recently, Mahrous & Kortam (2012) used SEM on a sample of students from three business schools. The students' evaluation instrument which was used in the study included 25 items representing six factors: organization of the course, workload difficulty, fairness of grading, instructor involvement, student-instructor interaction, and perceived learning. Four factors (organization of the course, workload difficulty, fairness of grading, and instructor involvement) were found to be significantly positively related to students' perception of learning while student-instructor interaction was not.

The results of these studies suggest that SEM methodology may be useful when examining interrelationships among factors on student evaluations of instructors. The purpose of this study is to propose a SEM to explain interrelationships between factors found in student evaluations of instructors and to validate it using empirical data.

METHODOLOGY

Samples

Data for this study were collected from the students attending the United Arab Emirates University (UAEU), a medium-sized four-year public university with an enrollment of approximately 15,000 students. First, eight focus groups, each with 10-15 randomly selected students, were organized to assist in generating items for the instrument that will be used to evaluate teaching effectiveness at the university level. Then data on this measure were collected on a random sample of 1096 students (Sample 1). Participating students represented both genders

and the eight colleges of the UAEU. Table 1 (Appendix) summarizes major demographic features of Sample 1. As can be observed, data included more females (80.7%) than males (19.3%) which reflected the actual ratio between males and females in the UAE University. All academic levels were represented; however, more participants were in their second or third level of study.

The data from Sample 1 were analyzed using exploratory factor analysis (EFA) to develop the final version of the instrument. Finally, another random sample of 1202 students (Sample 2) responded to the final version of the instrument and the collected data were used to validate the proposed SEM. Sample 2 also represented colleges, gender and all academic levels.

Procedure

The study used the following procedures. Assess university students' perceptions of effective teaching and use these perceptions to develop an initial instrument for evaluating instructors at the university level. Identify the instrument's underlying factors through exploratory factor analysis (EFA). Propose a structural equation model (SEM) to explain interrelationships between the factors of student evaluations of instructor, and finally, test the proposed SEM using empirical data.

Students perceptions about the characteristics of effective university instructors were assessed through conducting eight focus groups interviews (four for males and four for females). Focus groups participants were selected randomly to represent both genders and university colleges. Anonymity of participants and confidentially of the collected data were guaranteed. Participating students also discussed the main features of university courses generally and in relation to instructor evaluations in particular. Results of these interviews were collected and analyzed and a list of items or statements was developed. Then, related students' demographic variables (e.g., gender, school level, and college) were included on an initial version of the students' evaluation of instructor instrument.

Then, a panel of 10 faculty members from UAEU with background in education, educational measurement and evaluation, or educational psychology reviewed the initial version of the instrument. In this process, reviewers were asked to individually and carefully check item content, clarity, appropriateness, and any other related issue that may lead to improve the items or the instrument. All reviewers' comments and suggestions were reviewed and led to some item modification or deleting and the final draft of the instrument had 44 items. All were answered using 5-point Likert-type scale anchored from 1 *(strongly disagree)* to 5 *(strongly agree).*

RESULTS

Exploratory Factor Analysis (EFA)

An EFA using a principal components method with oblique rotation was applied on the 44 items using the data collected from a random sample of 1096 students (Sample 1). The purpose of this analysis was to determine the underlying dimensions or constructs measuring students' evaluation of effective instructors.

The appropriateness of the collected data to factor analysis was examined using two statistical methods: Kaiser-Meyer-Olkin measure of sampling adequacy and the Bartlett's test of sphericity. Kaiser-Meyer Olkin result indicated an acceptable level of .95., and the Bartlett's test

of sphericity was significant ($p < .001$). Oblique rotation methods, which assumes correlations between the extracted factors, were used because the components correlation matrix included many high (above .30) correlations. The results of the EFA identified eight components/factors in the instrument that together explain approximately 62% of the variance. Only components with at least three items with a loading score of equal to or greater than .40 were retained. Additionally, items were eliminated if their factor loadings were lower than .40 and items loaded on more than one dimension with a loading score of equal to or greater than .40 (Hattie, 1985; Chen & Hsu, 2001). This led to eliminate two factors because each one had less than three items with loading above .40. Additionally, eight items were eliminated because of low loadings or because of high (above .40) loadings on more than one component. Finally, six factors were clearly identified. Items highly loaded on each of these six factor and their corresponding loadings are presented in Table 2 (Appendix).

The first factor of five items (1, 2, 3, 4, and 5) discuses instructor characteristics and his/her relationships with students. The second factor of four items (6, 7, 8, and 9) discusses course clarity and organization. The third factor of five items (10, 11, 12, 13, and 14) discusses grades and exams. The fourth of five items (15, 16, 17, 18, and 19) discusses course workload and difficulty. The fifth factor of six items (20, 21, 22, 23, 24, and 25) discusses teaching methods. The sixth factor of seven items (26, 27, 28, 29, 30, 31, and 32) discusses students perceived learning and interest. More complete description of each of these six factors is presented in the next section.

Structural Equation Modeling (SEM)

The study developed a SEM to examine interrelations between the six factors described above based on previous studies in teaching evaluation and teaching effectiveness (e.g., Gursoy & Umbreit, 2005; Mahrous & Kortam, 2012; Marks, 2000; Paswan & Young, 2002; Stringer & Irwing, 1998). SEM establishes both direct and indirect effects of each factor in the model on the outcome or dependent variable. The proposed model (Figure 1) (Appendix) considered teacher characteristics to be exogenous construct while course characteristics (organization/clarity, grading /fairness, workload/difficulty, and teaching performance) mediators, and students' perceived learning/interest as the dependent variable.

The following describes each of the six factors followed by a justification or rational for including each in the model.

Teacher Characteristics. This factor focuses on personal characteristics of the instructor (e.g., knowledgeable, communicative, friendly, enthusiastic, caring, organized, approachable, treating students with respect, treat all students equally, and encouraging students). This factor may be the most important factor in that it impacts teaching performance as well as students learning (Polk, 2006; Murray, Rushton, & Paunonen,1990). Students ratings reflect the teacher's perceived personality more than his/her actual performance (Mango & Sembrano, 2007), and their ratings are influenced by an instructor's presentation style (Marks, 2000). Additionally, student ratings reflect their satisfaction and attitude toward instructors more than teaching effectiveness (Gursoy & Umbreit, 2005). In this model teacher characteristics is considered a latent variable.

Course Organization. This factor focuses on the course structure: whether it is well organized, concepts are presented systematically, and materials are presented at an appropriate

pace. Students considered course organization to reflect quality of teaching (Marks (2000). Organization and clarity of course structure contribute to students feeling comfortable, increasing their interest and perceived level of learning (Braskamp & Ory, 1994; Centra, 1993; Jackson et al. 1999; Paswan & Young, 2002; Marks, 2000).“Students see good teaching as presenting material at the right level at an appropriate pace within a clear, logical structure, and providing explanations that facilitate understanding and demonstrate both enthusiasm and empathy” (Gursoy & Umbreit, 2005, p. 103).

Grading/Fairness. This factor focuses on exams and evaluation methods: whether examinations were clear, fair and reasonable, assessed materials presented in the course, and the work was return in a reasonable time. Prior research identifies this factor as essential dimension of students’ evaluation of instructors (Braskamp & Ory, 1994; Centra, 1993; Jackson et al. 1999; Marsh, 1991; Paswan & Young, 2002; Marks, 2000).

Workload/Difficulty. This factor focuses on course workload: whether it included too much material, and the student’s ability to grasp its content. This factor appears commonly on student evaluations instruments (Burdsal & Bardo, 1986; Jackson et al., 1999; Marsh, 1991). Workload/difficulty influences students’ interest and their perception of learning (Gursoy & Umbreit, 2005).

Teaching Performance. This factor focuses on general teaching practices: whether materials were presented at an appropriately paced sequence, teaching aids and technology were used effectively, requirements were clear, difficult materials were clarified, methods of presenting materials were appropriate, and course presentations in the course were well prepared. Prior research indicated that perceived teacher performance is influenced by teaching methods and practices (Curits & Liying, 2001; Mango & Semnrano, 2007; Polk, 2006).

Student Perceived Learning/Interest. This factor focuses on student’s personal qualities: whether they were interested in learning course materials, were attentive in class, become more competent in this area, their found the course stimulated interest in area, was an effective learning experience, and increased their knowledge in this area. Prior research identified this factor as one of the most important for evaluating instructor effectiveness (Stringer & Irwing, 1998) and is included in most students’ evaluations of instructors (e.g., Braskamp & Ory, 1994; Centra, 1993; Jackson et al. 1999; Mark, 2000; Marsh, 1991; Paswan & Young, 2002).

The model (Figure 1) (Appendix) is developed on the premise that some perceptions of teaching effectiveness are developed before others and consequently affect final students rating (Gursoy & Umbreit, 2005; Mahrous & Kortam, 2012; Marks, 2000). In this model, teacher characteristics was considered the exogenous construct that influences course characteristics. This was assumed because teachers organize the course, set specific workload, set appropriate grading system, and exhibit the quality teaching (Stringer & Irwing, 1998). Additionally, students develop their opinion about their instructor early in the course. Later class and teaching-learning experiences do not materially change opinions (Marks, 2000; Stringer & Irwing, 1998). As noted previously, student evaluations of instructors usually are driven by perceptual factors such as instructor personality which is the most important factor on overall evaluation (Marks, 2000). Moreover, “students do not make a strict distinction between course evaluation and teacher evaluation. This is because student ratings of teaching effectiveness are primarily a

function of the instructor who teaches the course rather than the course that is being taught" (Marsh, 1987, p. 259).

The 32-item six factors described above form the final version of the students' evaluation instrument. This final version was used to collect data for assessing and validating the proposed structural equation model. Data were collected on a random sample of 1202 students (Sample 2). The second step of the analysis was to examine the internal reliability of each of the six factors as well as their inter-correlations. Cronbach's alpha reliability coefficient, mean, and standard deviation, and the correlation coefficients among these six factors were calculated and summarized in Table 3 (Appendix).

Correlations among five of the six factors in the model are high and significant. These are expected because these factors are subscales of one instrument designed to assess instructor teaching effectiveness. The exception, factor4, course workload/difficulty, has small and negative correlations with other factors. The internal reliability values of the six subscales as measured by Cronbach's alpha (the diagonal values) are high, ranging from .82 to .92. Among of the means of the six factors, the highest is for teacher characteristics while the lowest is for course workload/difficulty.

The next step of the analysis was testing the developed SEM which is called Model 1 Using EQS 6.1 software (Figure 2) (Appendix). Several fit indices commonly reported across the SEM studies were used to assess the fit of the developed model. Three commonly used fit statistics were the non-normed fit index (NNFI), the comparative fit index (CFI), and the goodness-of-fit index (GFI). Index values of .90 or higher indicate good fit for these three indices. The standardized root-mean square residual (SRMR) also was used to determine a residual estimate. A value of .06 or less of SRMR indicates a good fit. A fifth index used was the root mean square error of approximation (RMSEA), in which a value of .06 or less is considered to be adequate (Bentler & Bonnett, 1980; Hu & Bentler, 1995). Table 4 presents the selected goodness-of-fit statistics for the developed model. The results of the five goodness of fit indices indicated that the data fit the developed model well.

The results of SEM indicate that the standardized factor loadings or coefficients between the exogenous and the endogenous latent variables are large. These coefficients are the standardized regression coefficients that indicate the degree the dependent variable is influenced by the independent variable. Teacher characteristics are mediated highly by course organization (.91), teaching performance (.82), and grading/fairness (.80). All paths are statistically significant ($p<.05$). For example, the standardized coefficient value of .91, for example, for the path between teacher characteristics and course organization suggests that, while holding other variables constant, as teacher characteristics increases by one standard deviation, course organization is expected to increase by .91 standard deviations. In contrast, the effect of course loading/difficulty (-.12), although significant, is negative and small. This suggests small influence of teacher characteristics on course loading/difficulty.

Path coefficients also show how the mediators (organization/clarity, grading /fairness, workload/difficulty, and teaching performance) affect the dependent variable, perceived learning/interest. Results indicated that course organization (.17), teaching performance (.58), and grading/fairness (.21) exert a strong to moderate effect on the students' perceived learning/interest. In addition, the standardized residual variance of the dependent variable, perceived learning/interest, is .52, which indicates that approximately 50% of the variance in the perceived learning/interest is explained by the model.

In contrast, course loading/difficulty (-0.01) has no influence on perceived learning/interest. This factor, course loading/difficulty, is neither influenced by the teacher characteristics nor it affects the perceived learning/interest. Additionally, the errors or the residuals of this factor are very high (.99). This means that the influence on this factor from the model is very small as 99% of its variance is unexplained by the model. Thus, a modification was made on Model 1 by deleting the factor course loading/difficulty (Figure 3) (Appendix)

The modified model (Model 2) was tested to assess its goodness of fit to the data as well as to estimate its path coefficients. Results indicated that Model 2 also fits the data very well (Table 5) (Appendix).

The path coefficients of Model 2 provided evidence of important improvements through this modification. Increases were seen on teacher characteristics from course organization from .91 to .98, and on grading/fairness from .80 to .92, while the effect on teaching performance remains the same (.82) as in Model 1. In addition, the moderator variables now show a stronger impact on students perceived learning/interest as compared to the previous model. Increases were seen on course organization from .17 to .19, on teaching performance from .58 to .61. Grading/fairness effect remained the same (.21). The standardized residual variance of the dependent variable (perceived learning/interest) was reduced from .52 to .17. Thus, only 17% of the variance in the perceived learning/interest is not explained by the model.

DISCUSSION

This study proposed a model that hypothesize the directional influence or causal relationships of five factors in the students' evaluation of instructors in higher education. Specifically, the model shows the influence of teacher characteristics on other factors that, in turn, influence perceived students learning (the outcome of the teaching learning process).

Student learning constitutes the most widely utilized criterion of effective teaching (Burdsal & Harrison, 2008; Gursoy & Umbreit, 2005; Harrison, Douglas, & Burdsal, 2004; Marsh & Roche, 1997). Thus, attempts to evaluate teaching through its ultimate goal, learning, is logical. Therefore, students' evaluation of courses characteristics and instructor related-characteristics should focus on the perceived learning or interest (Clayson, 2009; Marks, 2000). Effective teaching evaluations need to focus on teacher-related behaviors as well as how they are translated into student/learning outcome (Ellet & Teddle, 2003; Gursoy & Umbreit, 2005; Ovando, 2001).

The availability of a model that links teaching with learning is very useful. Good teaching is to result in good learning. However, besides teaching, many variables intervene to influence students learning. Most of these variables are related to the course and its characteristics. Therefore, students were predicted to respond to course characteristics in terms of their perceived learning and interest. Students are unable to evaluate what they learned against what they should learn (Marks, 2000). Consequently, students rely on other factors (e.g., workload, organization, and grading) when evaluating quality of the learning.

The model indirectly relates teacher characteristics with student perceived learning through course characteristics. The model shows the significant effect of teacher characteristics on course organization, grading/fairness, and teaching performance, and emphasizes the importance of this factor in the evaluation of university instructor and teaching effectiveness. University instructor should not underestimate the effect of variables such as encouraging students to learn, meeting students outside the class, being enthusiastic, appearing

knowledgeable and organized, and showing sense of humor. These variables indirectly and strongly affect students learning and the evaluation of teaching effectiveness.

Results showed also how the mediator factors (course organization, grading/fairness, and teaching performance) strongly affect students' evaluation and their perceived learning. Students evaluate the quality of their learning through these factors. Knowledge of this relationship also is useful for instructors so they pay more attention to these variables in their teaching. Course workload/difficulty is not influenced by teacher characteristic. This suggests that the course load/difficulty and the instructor should be evaluated separately.

REFERENCES

Badri, M. A., Abdulla, M., Kamali, M., & Dodeen, H. (2006). Identifying potential biasing variables in student evaluation of teaching in a newly accredited business program in the UAE. *International Journal of Educational Management, 20*(1), 43-59.

Bentler, P. M., & Bonnett, D. G. (1980). Significant tests and goodness of fit in the analysis of covariance structure. *Psychological Bulletin, 88*, 588-606.

Braskamp, L. A., & Ory, J. C. (1994). *Assessing faculty work: Enhancing individual and institutional performance*. San Francisco: Jossey-Bass.

Burdsal, C., & Bardo, J. (1986). Measuring student's perceptions of teaching: Dimensions of evaluation. *Educational and Psychological Measurement, 56*(5), 63-79.

Burdsal, C., & Harrison, P. (2008). Further evidence supporting the validity of both a multidimensional profile and an overall evaluation of teaching effectiveness. *Assessment & Evaluation in Higher Education, 33*(5), 567–576.

Centra, J. A. (1993). *Reflective faculty evaluation: Enhancing teaching and determining faculty effectiveness*. San Francisco: Jossey-Bass.

Clayson, D. E. (2009). Student evaluations of teaching: Are they related to what students learn? A meta-analysis and review of the literature. *Journal of Marketing Education, 31*(1),16–30.

Curtis, A., & Liying, C. (2001). Teachers' self-evaluation of knowledge, skills and personality characteristics needed to manage change. *Asia-Pacific Journal of Teacher Education, 29*,139-143.

D'Apollonia, S., & Abrami, P. C. (1997). Navigating student ratings of instruction. *American Psychologist, 64*, 431-441.

Davies, M., Hirschberg, J., Lye, J., Johnston, C., & McDonald, I. (2005). Is it your fault? Influences on student evaluations of teaching in tertiary institutions. Unpublished manuscript. Retrieved from http://tlu.ecom.unimelb.edu.au/papers/Is_it_your_fault.pdf

Drago, W., Peltier, J., & Sorensen, D. (2002). Course content or the instructor: Which is more important in online teaching. *Management Research News*, 25(6/7), 69-83.

Ellet, C. D., & Teddle, C. (2003). Teacher evaluation, teacher effectiveness and school effectiveness: Perspective from the USA. *Journal of Personnel Evaluation in Education, 17*, 101-128.

Feldman, K. A. (1997). Identifying exemplary teachers and teaching: Evidence from student ratings. In R. P. Perry & J. C. Smart (Eds.), *Effective teaching in higher education: Research and practice* (pp. 368-395). Bronx, NY: Agathon.

Gage, N. L., & Berliner, D. C. (1992). *Educational psychology*, Boston, MA: Houghton, Mifflin Company.

Gursoy, D., & Umbreit, W. T. (2005). Exploring students' evaluations of teaching effectiveness: What factors are important? *Journal of Hospitality and Tourism Research, 29*(1), 91-109.

Harrison, P., Douglas, D., & Burdsal, C. (2004). The relative merits of different types of overall evaluations of teaching effectiveness. *Research in Higher Education, 45*(3), 311–323.

Huitt, W. (1995). A systems model of the teaching/learning process. *Educational Psychology Interactive.* Valdosta, GA: College of Education, Valdosta State University. Retrieved from http://www.edpsycinteractive.org/materials/tchlrnmd.html

Jackson, D. L., Teal, C. R., Raines, S. J., Nansel, T. R., Force, R. C., & Burdsal, C. A. (1999) The dimensions of students' perceptions of teaching effectiveness. *Educational and Psychological Measurement, 59*(4),580-596.

Johnson, R. (2000). The authority of the student evaluation questionnaire. *Teaching in Higher Education, 5*(4), 419-434.

Jöreskog, K. G., & Sörbom, D. (2004). *LISREL 8.71*. Chicago, IL: Scientific Software International, Inc.

King, A. S. (1983). How students size up prospective classes. *Journal of Business Education, 12*, 220-223.

Loveland, K. (2007). Student evaluation of teaching (SET) in web-based classes: Preliminary findings and a call for further research. *The Journal of Educators Online, 2*(4),1-18.

Mahrous, A., & Kortam, W. (2012). Students' evaluations and perceptions of learning within business schools in Egypt. *Journal of Marketing for Higher Education, 1*(22), 55–70.

Mango, C., & Sembrano, J. (2007). The role of teacher efficacy and characteristics on teaching effectiveness, performance, and use of learner-centered practices. *The Asia-Pacific Education Researcher, 16*(1),73-90.

Marks, R. B. (2000). Determinants of student evaluations of global measures of instructor and course value. *Journal of Marketing Education, 22*(2),108-119.

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

Marsh, H. W. (1991). Multidimensional students' evaluations of teaching effectiveness: A test of alternative higher-order structures. *Journal of Educational Psychology, 83*, 285-296.

Marsh, H. W. (2001). Distinguishing between good (useful) and bad workloads on students' evaluations of teaching. *American Educational Research Journal, 38*(1),183-212.

Marsh, H. W., & Roche, L. A. (1997). Making students' evaluations of teaching effectiveness effective. *American Psychologist, 52*,1187-1197.

Murray, H., Rushton, J., & Paunonen, S. (1990). Teacher personality traits and student instructional ratings in six types of university courses. *Journal of Educational Psychology, 82*, 250-61.

Neumann, R. (2000). A decade of student evaluation of teaching: Case study of an evolving system. *Journal of Institutional Research, 10*(1), 96-111.

Ngware, M. W. (2005). An improvement in instructional quality: Can evaluation of teaching effectiveness make a difference? *Quality Assurance in Education, 13*(3), 183-201.

Ovando, M. (2001). Teacher perceptions of a learner-centered evaluation system. *Journal of Personnel Evaluation in Education, 15,* 213-231.

Paswan, A., & Young, J. (2002). Student evaluation of instructor: A nomological investigation using structural equation modeling. *Journal of Marketing Education, 24*(3), 193-202.

Polk, J. A. (2006). Traits of effective teachers. *Arts Education Policy Review, 107*, 23-30.

Roberts, T. G., Irani, T. A., Telg, R. W., & Lundy, L. K. (2005). The development of an instrument to evaluate distance education courses using student attitudes. *American Journal of Distance Education, 19*(1), 51-64.

Rowley, J. (2003). Designing student feedback questionnaires. *Quality Assurance in Education, 11*(3), 142-149.

Rugutt, J. K., & Chemosit, C. C. (2005). A study of factors that influence college academic achievement: A structural equation modeling approach. *Journal of Educational Research & Policy Studies, 5*(1), 66-90.

Spooren, P., Mortelmans, D., & Denekens, J. (2007). Student evaluation of teaching quality in higher education: Development of an instrument based on 10 Likert-scales. *Assessment & Evaluation in Higher Education, 32*(6), 667-679.

Stringer, M., & Irwing, P. (1998). Students' evaluations of teaching effectiveness: A structural modeling approach. *British Journal of Educational Psychology, 68*(3), 409-426.

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

Wheeler, G. E., & Guerts, M. D. (1986). Student evaluation of faculty: A longitudinal study from one department in a business school. *Journal of Marketing Education, 8*, 24-31.

APPENDIX

Table 1
Demographic Characteristics of the Participating Students in Sample 1 and Sample 2

		Sample 1	Sample 2
Gender (n (%))	Males	211 (19.3)	245 (21.1%)*
	Females	885 (80.7)	771 (64.1%)*
Level (n (%))	First	77 (7.0)	104 (8.7%)
	Second	327 (29.8)	238 (19.8%)
	Third	414 (37.8)	340 (28.3%)
	Fourth	176 (16.1)	296 (24.6%)
	Fifth	76 (6.9)	119 (9.9%)
College (n (%))	Humanities	105 (9.6)	270 (22.5%)
	Sciences	209 (19.1)	201 (16.7%)
	Education	79 (7.2)	55 (4.6%)
	Business	222 (20.3)	191 (15.9%)
	Law	186 (17.0)	187 (15.6%)
	Engineering	106 (9.7)	94 (7.8%)
	Agriculture	84 (7.7)	54 (4.5%)
	IT	85 (7.8)	90 (7.5%)

*: Percentages do not make 100% due to missing data.

Table 2
Loadings of the Items on Six Extracted Factors from the Exploratory Factor Analysis

No.	Item	Component					
		1	2	3	4	5	6

1	The instructor encourages students to learn	.11	.01	.04	.02	.02	**.61**
2	The instructor is enthusiastic	.08	.00	.02	.01	.05	**.78**
3	The instructor appeared knowledgeable	.09	.06	.07	.03	.07	**.74**
4	The instructor is available outside the class	.05	.00	.15	.04	.12	**.58**
5	The instructor has sense of humor	.12	.08	.14	.10	.24	**.45**
6	The course is well organized	.14	.06	.07	.09	**.61**	.02
7	The course materials are updated	.10	.04	.06	.01	**.72**	.03
8	The course syllabus is clear	.01	.12	.05	.27	**.42**	.22
9	The course materials are related to real life	.06	.02	.06	.11	**.49**	.16
10	Exams' questions in this course were clear	.06	.00	**.77**	.01	.08	.05
11	Grading was fair in this course	.04	.09	**.70**	.05	.06	.07
12	Exams covered the materials taught in the course	.05	.02	**.73**	.06	.06	.03
13	Grades in this course were within my expectations	.02	.05	**.78**	.01	.00	.03
14	Valuable feedback was offered	.10	.01	**.48**	.27	.00	.21
15	Workload in this course was heavy	.05	**.70**	.03	.15	.05	.03
16	This course was very difficult	.02	**.83**	.13	.06	.04	.11
17	The main concepts of this course were difficult	.03	**.84**	.03	.01	.03	.03
18	More time is needed because of its difficulty level	.05	**.86**	.04	.02	.02	.04
19	Heavy workload	.01	**.74**	.02	.16	.01	.04
20	Variety of teachings methods were used	.02	.12	.02	**.63**	.12	.09
21	Appropriate paced sequence	.02	.00	.02	**.73**	.11	.02
22	Teaching aids were used effectively	.02	.01	.02	**.83**	.02	.07
23	Technology was used effectively	.04	.08	.04	**.79**	.06	.09
24	Difficult materials were clarified	.21	.01	.12	**.52**	.05	.09
25	Classes in the course were well prepared	.21	.02	.13	**.55**	.01	.00
26	I learned lots of new things in this course	**.70**	.01	.06	.02	.05	.13
27	I was attentive in this course	**.74**	.02	.03	.03	.07	.12
28	I became more competent in this subject	**.77**	.02	.00	.08	.05	.09
29	The course stimulated my interest in this subject	**.78**	.03	.02	.04	.10	.01
30	The course was a useful learning experience	**.77**	.08	.01	.09	.03	.03
31	The course increased my knowledge in this subject	**.76**	.02	.02	.04	.04	.07
32	The course enhanced my thinking skills	**.67**	.04	.03	.03	.03	.17

Table 3
Correlations among the Six Factors, Reliability Coefficients, Means, and Standard Deviations

Factor/Subscale	1	2	3	4	5	6	Mean	SD
1. Teacher Characteristics	.86	.71**	.60**	-.10**	.66**	.59**	4.19	.67
2. Course Organization		.82	.65**	-.08*	.69**	.68**	4.00	.74
3. Grading/Fairness			.83	-.06*	.63**	.62**	3.87	.72
4. Workload/ Difficulty				.92	-.00	-.06*	3.34	1.01
5. Teaching Performance					.87	.71**	4.00	.66
6. Perceived Learning						.92	4.00	.73

*: Significant at .05
**: Significant at .01

Table 4:
Goodness-of-Fit Indices of the Developed Structural Equation Modeling (Model 1)

Fit Index				
NNFI	CFI	GFI	SRMR	RMSEA
.91	.92	.88	.05	.06

Note. NNFI = Non-Normed Fit Index, CFI = Comparative Fit Index; GFI = Goodness-of-Fit Index; SRMR= Standardized Root Mean Square Residual; RMSEA = Root Mean Square Error of Approximation

Table 5:
Goodness-of-Fit Indices of the Developed Structural Equation Modeling (Model 2)

Fit Index				
NNFI	CFI	GFI	SRMR	RMSEA
.90	.91	.88	.05	.07

Note. NNFI = Non-Normed Fit Index, CFI = Comparative Fit Index; GFI = Goodness-of-Fit Index; SRMR= Standardized Root Mean Square Residual; RMSEA = Root Mean Square Error of Approximation

Figure 1

The Developed SEM of Relationships between the Factors of the Student Evaluations of Instructors

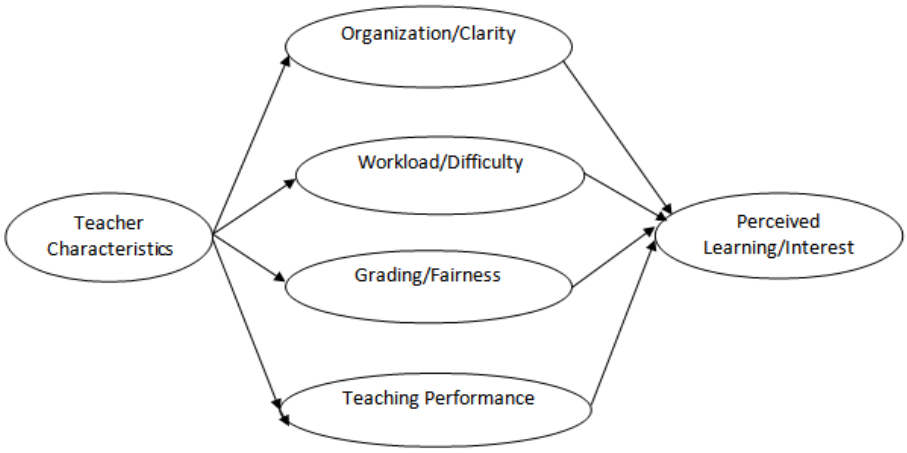
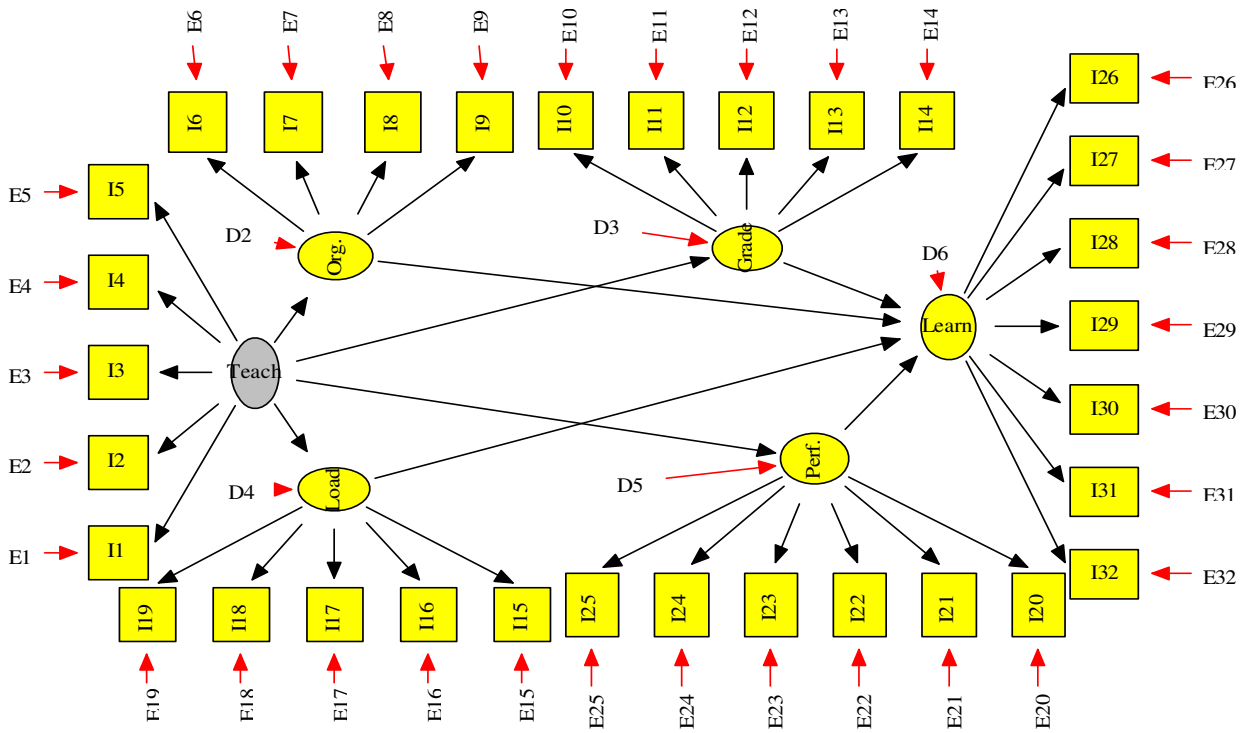
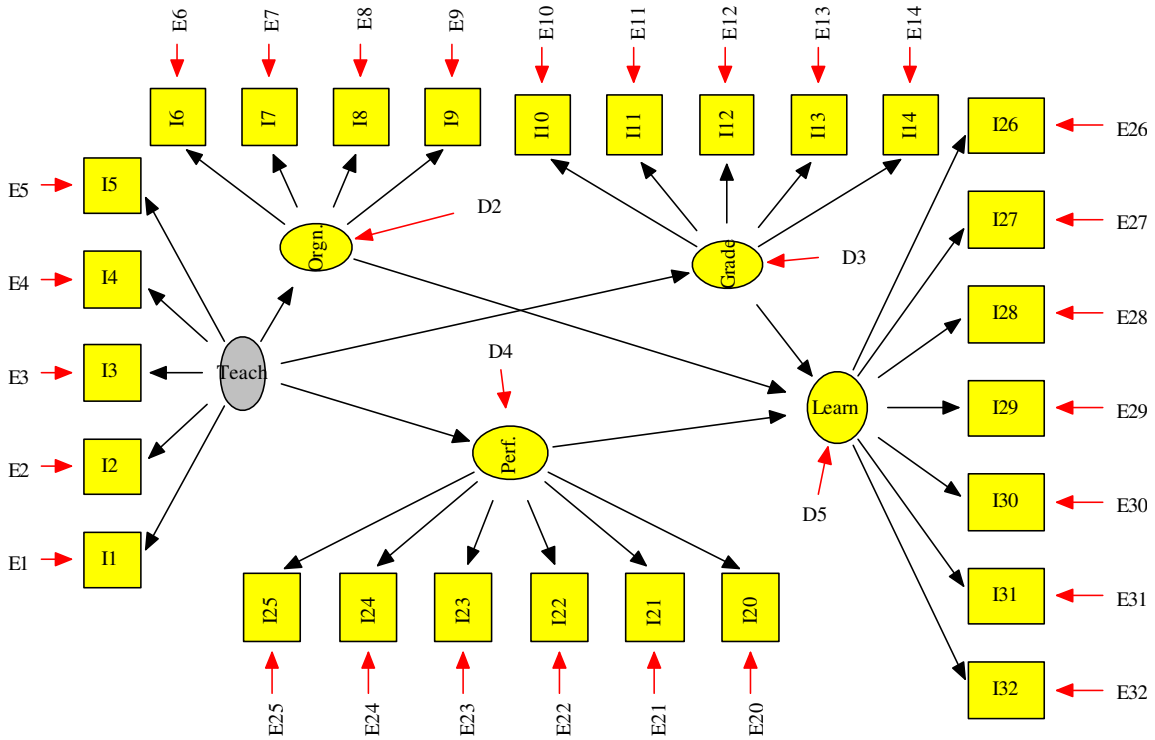


Figure 2
The Developed SEM (Model 1) Including the Course Workload/Difficulty



Note: Teach: Teacher Characteristic; Org.: Course Organization; Load: Course Workload/Difficulty; Grade: Grading/Fairness; Perf. Teaching Performance; Learn: Perceived Learning/Interest.

Figure 3
The Developed SEM (Model 2) Excluding the Course Workload/Difficulty



Note: Teach: Teacher Characteristics; Org.: Course Organization; Grade: Grading/Fairness; Perf. Teaching Performance; Learn: Perceived Learning/Interest.