THE EFFECT OF PERCEIVED INSTRUCTIONAL EFFECTIVENESS ON STUDENT LOYALTY: A MULTILEVEL STRUCTURAL EQUATION MODEL

ALGILANAN ÖĞRETİMSEL ETKİLİLİĞİN ÖĞRENCİ SADAKATİNE ETKİSİ: ÇOK AŞAMALI YAPISAL EŞİTLİK MODELİ

Gülhayat GÖLBAŞI ŞİMŞEK[*], Fatma NOYAN[**]

ABSTRACT: Social sciences research often entails the analysis of data with a multilevel structure. An example of multilevel data is containing measurement on university students nested within instructors. This paper concentrate on multilevel analysis of structural equation modeling with educational data. Data used in this study were gathered from 17647 university students in Turkey taking course from 202 instructors during the first term of 2004 academic year. The main topic of this paper is to investigate the effect of Perceived Instructional Effectiveness (PIE) on Student Loyalty (SL). From the both of within and between model results, it was supported that student loyalty is positively affected by perceived instructional effectiveness. The total variation of SL explained by PIE in within model was 57%; on the other hand total variation of SL explained by PIE and instructor's academic status in between model was 92%. The effects of the other background variables were also considered.

Keywords: multilevel structural equation modeling, perceived instructional effectiveness, student loyalty.

ÖZET: Eğitim araştırmalarında öğrencilerin, sınıfların ve okulların çok aşamalı yapıda olması gibi, aynı öğretim üyesinden ders alan öğrenciler de çok aşamalı veri yapısına bir örnektir. Bu çalışmada Algılanan Öğretim Üyesi Etkililiğinin (AÖÜE) Öğrenci Sadakatine (ÖS) etkisinin araştırılması için, çok aşamalı yapısal eşitlik modelinden yararlanılmıştır. 2004 akademik yılında 202 öğretim üyesinden (ikinci aşama örnekleme birimi) ders alan 17647 öğrenci (birinci aşama örnekleme birimi) araştırmanın örneklemini oluşturmuştur. Hem öğretim üyesi içi, hem de öğretim üyeleri arası yapısal eşitlik modelleri AÖÜE'nin ÖS'yi olumlu etkilediğini doğrulamıştır. ÖS'deki değişkenliğin AÖÜE tarafından açıklanan kısmı, öğretim üyesi içi modelde %57 iken, öğretim üyeleri arası modelde %92 bulunmuştur. Ayrıca çalışmada ikinci aşama değişkenlerinin öğretim üyeleri arası modelde ÖS'ye etkisi de araştırılmıştır.

Anahtar sözcükler: çok aşamalı yapısal eşitlik modeli, algılanan öğretim üyesi etkililiği, öğrenci sadakati.

1. INTRODUCTION

Many colleges and universities are using some form of student evaluation to measure teaching quality and effectiveness. These instruments include Endeavor Instrument, Student Instructional Rating System (SIRS), Instructor and Course Evaluation System (ICES), Student Description of Teaching (SDT), Students' Evaluations of Educational Quality (SEEQ), and Instructional Development and Effectiveness Assessment (IDEA). These evaluation instruments are usually administered at the end of a semester or quarter. "How well does student evaluation indicate teaching quality?" is a question many educators have asked in recent years. Although these instruments are widely used in American universities, the question of whether student evaluation indicates teaching quality remains largely unanswered (Chan, 2001).

Student ratings or student evaluation of teaching effectiveness continue to be a controversial topic. Some instructors view students as an obvious and convenient choice as raters. (e.g. Jackson et al., 1999), whereas others are convinced that students are essentially unqualified to provide a valid assessment of teaching quality (e.g., Bonetti, 1994). Despite these conflicting views, colleges and universities continue to collect data on teaching quality through the use of questionnaires administered to their students (Nasser and Hagtvet, 2006; Ekinci and Burgaz, 2007).

While only a few instructors argue strongly against the usefulness of student ratings in providing feedback about instructional effectiveness, many of them continue to challenge the use of student ratings for personal decisions. Many researchers have questioned the use of student ratings as a

[*] Assistant Prof. Dr., Department of Statistics, Yıldız Technical University, gulhayat@yildiz.edu.tr
[**] Research Assistant, Department of Statistics, Yıldız Technical University, fnoyan@yildiz.edu.tr

measure of teaching effectiveness (e.g., Chau,1997; Marsh, Overall and Kesler, 1979) on grounds that such ratings are biased by variables unrelated to teaching effectiveness. Other researchers consider these variables as a source of relevant and legitimate variance in student ratings (Marsh, 1987; Marsh and Rochi, 1997,2000).

In recent years there has been some discussion on the validity of using student evaluation instruments to measure instructional effectiveness and quality. One side of the argument is that student satisfaction is indicative of teaching quality. The counter-argument is that student satisfaction measurements are too customer-oriented. A satisfied student doesn't always imply that the instructor is being effective or the course is relevant (Chan, 2001).

1.1.Instructional Effectiveness and Student Loyalty

The assessment of instructional effectiveness to student satisfaction relationship is advocated (Guolla, 1999). Rather than abandoning the use of student evaluation instruments, Guolla evaluates the impact of multiple teaching quality factors on course satisfaction and instructor satisfaction as perceived by students. Guolla applied established theory from customer satisfaction and educational psychology research to a sample of MBA and undergraduate students and found that there was a strong correlation between learning and course satisfaction; strong correlation was also found between instructor enthusiasm and instructor satisfaction. Guolla believes that measuring satisfaction at the end of a course raises two questions: "Does the course in its entirety constitute a single, yet extended, consumption experience? Or, do all class sessions represent separate episodes resulting in cumulative satisfaction?". Guolla discusses that student satisfaction is a suitable criterion variable that can be predicted by six teaching quality factors: learning, enthusiasm, organization, interaction, rapport, assignments, and material. His rationale is that students are customers while being clients, producers, and products at the same time. Their satisfaction level should allow instructors to analyze how their teaching performance can be improved.

From the marketing research theories and empirical studies, perceived quality and satisfaction have been found to be major antecedents of customer loyalty (Beardona and Teel, 1983; Innis, 1991; Oliver, 1980, 1981; Roest and Pieters, 1997). In the conceptual model developed there:

Perceived Quality $\rightarrow$ Customer Satisfaction $\rightarrow$ Customer loyalty

Perceived quality is the customer's overall assessment of the standard of the service delivery process. Customer satisfaction is the degree of overall pleasure or contentment felt by the customer, resulting from the ability of the service to fulfill the customer's desires, expectations and needs in relation to the service. Customer loyalty is the individual's judgment about buying again a designated service from the same company, taking into account his or her current situation and likely circumstances (Hellier, Geursen, Carr and Rickard, 2003).

In this content, student evaluations or ratings of instructional effectiveness (perceived instructional effectiveness) can be regarded as perceived quality not customer satisfaction, and student's general evaluation of instructor/course can be considered as customer satisfaction. Perceived quality is the fundamental requirement for customer loyalty. On the other hand, in educational studies, it is difficult to measure satisfaction because students are not considered as customers exactly. Under this situation, it may be useful to examine the direct effects of perceived instructional effectiveness on student loyalty.

In previous researches on the relationship between perceived instructional effectiveness and/or student satisfaction, data were either disaggregated to the individual (e.g., student) level or aggregated to group (e.g., instructor) level. Neither approach is adequate for a proper understanding of the actual structure of clustered data. Consequently, this state of affairs suggests the need for a methodology that accounts for the dependency in the multilevel data. Multilevel regression methods (e.g.,Roundenbush and Willms, 1991) have contributed significantly to the understanding of clustered data. However, these methods do not readily allow for the specification of systems of structural equations within and between levels of the clustered data (Kaplan and Elliott, 1997). Multilevel structural equation

modeling (MSEM) provides solutions to these shortcomings and recent advances by Muthén and his colleagues (Muthén , 1989, 1991, 1994; Muthén and Sattora, 1989) allow researchers to combine full structural equation models (e.g., Jöreskog, 1977) with multilevel models. This advanced method was used in the current study to address some of the shortcomings that characterized conventional analysis methods frequently used in previous researches. More specifically, the purpose of the current study was to examine the extent to which perceived instructional effectiveness by students affects student loyalty to instructor through MSEM. This powerful method allows for simultaneous estimation of the within and between instructor variation.

2. METHODOLOGY

2.1. Multilevel Structural Equation Modeling

Social science often studies systems that possess a hierarchical structure. Naturally, such systems can be observed at different hierarchical levels. Familiar examples are the educational system, with its hierarchy of pupils within instructors and/or classes within schools, families, with family members within families, and other social structures where individuals are grouped in larger organizational or geographical groups. As a consequence the data can be regarded as multistage or cluster sample from different hierarchical levels. In the most general case, there are not only variables at separate levels, but there may be different sets of variables at the separate levels (Hox, 1993).

Structural equation modeling (SEM) is used when the variables of interest cannot be measured perfectly. Instead, there are either sets of items reflecting a hypothetical construct (e.g. depression) or fallible measurements of a variable (e.g. caloryintake) using different instruments. The latent variables, or factors, are interpreted as constructs, traits or 'true' variables, underlying the measured item sand inducing dependence among them. The measurement model is some times of interest in its own right, but relations among the factors or between factors and observed variables (the structural part of the model) are often the focus of investigation. (Rabe-Hesketh, Skrondal and Zheng, 2007). Similar to the applications of the hierarchical linear model to regression in the context of the multilevel model, Multilevel SEM (MSEM) is a direct generalization of SEM in the context of the multilevel model.(Cheung and Au, 2005)

The popularity of multilevel modeling and SEM is a striking feature of quantitative research in the medical, behavioral and social sciences. Although developed separately and for different purposes, SEM and multilevel modeling have important communalities since both approaches include latent variables or random effects to induce, and therefore explain, correlations among responses. (Rabe-Hesketh, Skrondal and Zheng, 2007).

MSEM, a synthesis of multilevel and structural equation modeling, is required for valid statistical inference when the units of observation form a hierarchy of nested clusters and some variables of interest are measured by a set of items or fallible instruments. Multilevel structural equation modeling also enables researchers to investigate exciting research questions which could not otherwise be validly addressed. (Rabe-Hesketh, Skrondal and Zheng, 2007). For instance, in this paper we will consider an important question in education: does student loyalty (single item latent variable) depend on student evaluation of instructor effectiveness (a latent variable) at both of student and instructor level?

Although in principal, special formulas and software could be developed for MSEM maximum likelihood (ML) estimation, Muthén (1989,1990) showed that multiple group SEM software can be modified for MSEM-ML analysis. In line with this idea, Muthén proposed a simpler multiple group SEM software. This estimator uses the customary between and pooled within sample covariance matrices. In the balanced case, the estimator is consisted and, despite the fact that it uses less information than ML, has given similar results in the analyses to date. The true ML procedure will be referred to as FIML (Full information ML) and the simpler estimator as MUML (Muthén's ML based estimator) (Muthén,1994).

As pointed Muthén (1989), MSEM is a complex analysis, which needs to follow a sound strategy. The actual MSEM should, in a typical case, be preceded by four important analysis steps: conventional SEM of S_T (Total Sample Covariance Matrix), estimation of between variation, estimation of within structure, and estimation of between structure (Muthén,1994). S_T, S_{PW} (Pooled Within Sample Covaraince Matrix) and S_B (Between Sample Covariance Matrix) can be calculated using standard statistical packages or SEM software.

2.2. Data Source and Sample

The study was conducted at the Yıldız Technical University which is the one of the largest universities in Turkey. Course/instructor evaluation was initiated at this university in 2003. Data used in this study were gathered from 17647 students taking course from 202 instructors during the first term of 2004 academic year. Student evaluations of one course taught by each instructor were included in the data. Through simple random selection of instructors with p=0.10, the instructor evaluation questionnaire was applied to all students of these instructors. The distribution of the instructors' sample: (a) gender: 68.4% were men, and 31.6% were women; (b) status: not professor or associated professor = 45.4%, professor or associated professor= 54.6%. The students were not asked for any personal information due to avoid biased responses.

Student ratings of instructor effectiveness with 8 items measuring perceived instructional effectiveness and student loyalty with 1 item were collected using an evaluation questionnaire (see items in Appendix 1). Students responded to each item using a 5-point Likert-type scale. Responses to the perceived instructional effectiveness items yielded high internal consistency (Cronbach's alpha=0.9048).

Instructor data were obtained from university records. Variables included at the between group only as follows:

Course_n : Number of distinct courses given by the instructor in the related academic term ; Course_h: Total course hours in a week given by the instructor in the related academic term ; Status: Instructor's academic status (0= not professor or associated professor, 1= professor or associated professor); Public_r: Ratio of number of instructor's publications to his/her colleagues in his/her department; Gender: Instructor's gender: 0=male, 1=female.

2.2. Theory-Based Research Model

Our research hypothesis was that these data should fit a MSEM, with two latent variables on which same type linear relationships with different estimates are hold at both levels. Perceived instructional effectiveness (exogenous latent variable) affects student loyalty to instructor (endogenous latent variable) positively. We have no additional variables in student level other than exogenous and endogenous latent variables with corresponding indicator variables. In instructor level (between group), the positive relationship between perceived instructional effectiveness and student loyalty is conceptually sensible and theoretically persuasive respect to marketing research theory. On the other hand, the instructor level part through MSEM should also be considered due to student responses to these 9 items not only reflect individual variation but also systematic instructor differences in beliefs/attitudes about instructional effectiveness and loyalty. We have also background variables to be used for explaining instructor level variation in the latent variables. Instructor level exogenous latent variable, perceived instructional effectiveness is proposed to be affected from course_n and course_h negatively; while from public_r positively. Instructor's status and gender are assumed to affect between groups endogenous latent variable student loyalty.

3. FINDINGS

In this study, all analyses were performed using SPSS 13.0, LISREL 8.80 and Mplus 3.0. Muthén's first step is performed in LISREL, robust ML estimation was performed using conventional asymptotic covariance matrix and Weighted Least Squares Estimation was performed using

conventional polychoric correlation matrix and corresponding asymptotic covariance matrix computed all in LISREL in addition to conventional ML estimation for S_T. In Muthén's second step ICCs are computed through Mplus directly and SPSS indirectly. Muthén's third step was carried out in LISREL using S_{PW}. In addition to conventional estimation, Robust ML estimation was performed using asymptotic covariance matrix. Muthén's fourth step was carried out analyzing both of between group sample covariance matrix and estimated between group covariance matrix performed in LISREL, In this step, asymptotic covariance matrix was also used for comparison purpose. And the last step, MSEM, was performed both of Mplus's default MUML estimation and LISREL's two-group ML estimation with estimated between group and sample pooled within group covariance matrices.

Multiple indices were referenced to determine model fit. First, model fit was assessed using the chi-square goodness-of-fit statistics. A well-fitting model would be expected to have small (relative to degrees of freedom), non-significant value of chi-square. However, when sample sizes are large, the chi-square statistic may be statistically significant even though the model is substantially correct. Thus, we also used the Comparative Fit Index (CFI), the Root Mean Square Error of Approximation (RMSEA), and the Standardized Root Mean Square Residual (SRMR) as guides in assessing fit. For lack of any standard protocol with MSEM, we followed Hu and Bentler's (1999) recommendations in regards to fit indices. We looked for values of CFI in the mid 0.90's or higher, and RMSEA and SRMR of 0.05 or less (Dyer, Hanges and Hall, 2005).

The intraclass correlation coefficients (ICCs) were calculated for each indicator and the ICC(1) values for these 9 indicators ranged from 0.118 to 0.204, with an average ICC of 0.173. B. O. Muthén's (1994) experience with survey data suggested that the common values of ICC ranged from 0.00 to 0.50. Moreover, the ICC(1) may underestimate the true intraclass correlation because individual level measurement error contributes to the within variance. Given our relatively high ICC values, we concluded that there was sufficient between group variation offering MSEM for making valid statistical inferences.

We investigate the between structure using both covariance matrices of $\mathbf{S_B}$ and $\hat{\mathbf{\Sigma}}_\mathbf{B}$ computed in Mplus. It is tempting to use $\mathbf{S_B}$ to explore the between structure. Note, however, that $\mathbf{S_B}$ is not an unbiased or consistent estimator of $\mathbf{\Sigma_B}$. The $\mathbf{\Sigma_B}$ estimator is also a function of $\mathbf{S_{PW}}$. In other words, any simple structure expected to hold for $\mathbf{\Sigma_B}$ does not necessarily hold for $\mathbf{S_B}$, but it should hold within sampling error for the ML estimate of $\mathbf{\Sigma_B}$.(Muthén, 1994). On the contrary to this usual situation, our $\hat{\mathbf{\Sigma}}_\mathbf{B}$ is positive definite, while $\mathbf{S_B}$ is not positive definite. Then we conducted the analyses in LISREL using both of covariance matrices and ridge option was taken analyzing $\mathbf{S_B}$.In this step, Robust ML estimation with asymptotic covariance matrix computed through LISREL. The effects of background variables to latent variables also considered. The effects of course_n, course_h and public_r on perceived instructional effectiveness of instructor level were found theoretically true signs but statistically insignificant. On the other hand, effect of status variable on student satisfaction of instructor level was found statistically significant (structural coefficient estimate= 0.11 with standard error=0.028 and t=3.94) and theoretically reasonable while gender effect was found statistically insignificant.

MSEM was conducted in both of Mplus and LISREL's two-group settings. Specifying the MSEM in Mpus is relatively straightforward than LISREL. The estimates, standard errors and t values of these analyses were almost same except model fits. In LISREL, covariance matrices of $\hat{\mathbf{\Sigma}}_\mathbf{B}$ and $\mathbf{S_{PW}}$ were fitted to proposed models simultaneously. Table 1 shows parameter estimates, standard errors of parameter estimates and t values obtained from the Mplus's MUML estimation and, model fit indices and chi-squares obtained from both of Mplus's MUML estimation and LISREL's two-group ML estimation.

Table 1: MSEM Results of Proposed Model

Item	Within Model (Student Level)				Between Model (Instructor Level)			
	Estimate	SE of Estimate	t	Reliability	Estimate	SE of Estimate	t	Reliability
x1	1	-	-	0.37	1	-	-	0.68
x2	1.20	0.016	74.49	0.49	1.47	0.094	15.67	0.84
x3	1.11	0.014	78.36	0.46	1.27	0.080	15.87	0.84
x4	1.48	0.019	80.07	0.60	1.83	0.112	16.30	0.87
x5	1.56	0.019	82.96	0.67	1.87	0.108	17.20	0.93
x6	1.28	0.017	74.08	0.48	1.42	0.096	14.78	0.78
x7	1.46	0.019	76.42	0.53	1.80	0.109	16.49	0.89
x8	1.18	0.018	64.062	0.33	1.36	0.122	11.11	0.52
	Structural Path							
	Estimate	SE of Estimate	t	R^2	Estimate	SE of Estimate	t	R^2
SL on PIE	1.60	0.020	78.60	0.57	2.05	0.121	16.96	0.92
SL_B on Status	-	-	-		0.11	0.03	3.64	
	Model Fit							
	MUML				Two-Group ML			
Chi-square	1948.004				1997.93			
df	58				68			
CFI	0.98				0.98			
TLI	0.97				0.97			
RMSEA	0.043				0.057			
SRMR	Within=0.022 Between=0.047				Within= 0.023 Between=0.0498			

As shown in Table 1, the MSEM fitted the data reasonably well, especially RMSEA=0.043 . The results from the proposed model suggest good fit at the within level, and adequate, but little worse fit at the between level. The total variation of SL (Student Loyalty) explained by PIE (Perceived Instructional Effectiveness) in within model was 57%; on the other hand total variation of SL explained by PIE and STATUS in between model was 92% owing to the absence of individual (student) level measurement errors in the indicator variables.

Construct reliabilities for PIE in within and between models were found as 88.43% and 96.90% respectively. %96.90 of construct reliability of PIE in between model was higher than the within model one because between model indicators should be treated in suspect owing to individual level measurement errors free. On the other hand, even 88.43% of construct reliability of PIE in within model was higher than 86.84% of construct reliability value calculated from the analyzing the Total Sample Covariance Matrix (S_T). As a result, MSEM offers higher and possibly more accurate

reliabilities than conventional SEM ones. In addition to these results; all reliabilities are above of the suggested value of 70% for conventional SEM.

Variance explained calculated for PIE were 59.77%, 80.19% and 45.50% for within model, between model and ignoring nested structure respectively. MSEM rendered variance explained very close to 60%. As a result, variance explained of PIE was adequate to interpret findings.

4. DISCUSSION AND CONCLUSIONS

Student ratings or student evaluation of instructional effectiveness continue to be a controversial topic while many colleges and universities are still using hardcopy and/or online questionnaires of student evaluation to measure teaching quality and effectiveness. Misleading inferences have been made due to conceptual confusion. In this questionnaires the concepts such as teaching effectiveness, instructional effectiveness, perceived instructional/teaching effectiveness, student satisfaction and teaching quality are misused. In this paper, we considered perceived instructional effectiveness and student loyalty. Then it is needed to clarify the underlying concepts. What was perceived instructional effectiveness? Was it teaching quality, was it instructor performance (publications, teaching hours etc.)? After reviewing educational literature, it is concluded that perceived instructional effectiveness is student evaluation of instructional quality and ability. What is student loyalty? We utilized from marketing research terms and borrowed the term of customer loyalty from them. Customer loyalty is the individual's judgment about buying again a designated service from the same company, taking into account his or her current situation and likely circumstances. In this content, student evaluations or ratings of instructional effectiveness (perceived instructional effectiveness) can be regarded as perceived quality not customer satisfaction, and concept like customer repurchase intention (I would like to take another course from the same instructor) may be considered as loyalty for students.

When we carefully consider the problem of analyzing data arising from organizations, it is clear that neither SEM nor Multilevel modeling alone give a complete picture of the problem under investigations. Indeed use of either methodology separately could result indifferent but perhaps equally serous specification errors. Specially, utilization SEM modeling alone would ignore the sampling schemes that often used to generate educational data and would result in biased structural regression coefficient (Muthén and Satorra, 1989). The use of multilevel modeling alone would preclude the analyst from studying complex indirect and simultaneous effects within and across levels of the system. What is required, therefore, is a method that combines the best of both methodologies (Kaplan. 2000).

In this paper we examined the relationship between perceived instructional effectiveness and student loyalty. It was hypothesized that perceived instructional effectiveness affects student loyalty positively. This hypothesis was not rejected for which both of within and between model. The second that student loyalty of between model are affected from the status of instructor. The higher the status, the higher the student satisfaction is. This hypothesis was not rejected. The effects of the other background variables (number of distinct courses, total course hours, publication ratio and instructor's gender) were found statistically insignificant. Construct reliability of perceived instructional effectiveness for within model and between model was adequate, and variance explained of them for underlying concept were marginally well. 88.43% of construct reliability of perceived instructional effectiveness in student level indicate that student evaluation of instructional effectiveness is reliable and it can be used for various purposes.

REFERENCES

Bearden, W.O., & Teel, J.E. (1983). Selected determinants of consumer satisfaction and complaint reports. *Journal of Marketing Research, 2,* 21-28.

Bonetti, S. (1994). On the use of student questionnaires. *Higher Education Review, 26*(3), 57- 64.

Chan, K.C. (2001). *Does Instructor Witness and Character Matter in Student Perception of Faculty Performance and Course Quality.* Paper presented at the meeting of The Christian Business Faculty Association 17th Annual Conference. Olivet Nazarene University.

Chau, C.T. (1997). A bootstrap experiment on the statistical properties pf students' ratings of teaching effectiveness. *Research in Higher Education, 38*(4), 497-517.

Cheung, M_KE. W-L., & Au, K. (2005). Applications of multilevel structural equation modeling to cross-cultural research. *Structural Equation Modeling, 12*(4), 598-619.

Dyer, G.N., Hanges, P.J., & Hall Rosalie, J. (2005). Applying multilevel confirmatory factor analysis techniques to the study of leadership. *The Leadership Quarterly, 16*, 149-167.

Ekinci C. E., & Burgaz B. (2007). Hacettepe Üniversitesi öğrencilerinin bazı akademik hizmetlere ilişkin beklenti ve memnuniyet düzeyleri. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi, 33*, 120-134.

Guolla, M. (1999). Assessing the teaching quality to student satisfaction relationship: Applied customer satisfaction research in the classroom. *Journal of Marketing Theory and Practice, 16,* 87-97.

Heck, R.H., & Thomas, S.L. (2000). *An introduction to multilevel modeling techniques* Mahwah, New Jersey: Lawrence Erlbaum Associates.

Hellier, P.K., Geursen, G.M., Carr, R.A., & Rickard, J.A. (2003). Customer repurchase intention: A general structural equation model. *European Journal of Marketing, 37*(11/12), 1762-1800.

Hofmann, D.A. (1997). An overview of the logic and rationale hierarchical linear models. *Journal of Management, 23*, 723-744.

Hox, J.J. (1993). Factor analysis of multilevel data: Gauging the Muthén model. In J.H.L. Oud and R.A.W. van Blokland-Vogelesang (Ed.), Advances in longitudinal and multivariate analysis in the behavioral sciences (pp. 141-156). Nijmegen: NL:ITS.

Hu, L., & Bentler, P.M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*, 1-55.

Innis, D.E. (1991). Customer service, repurchase intentions, market orientation and firm performance in the channel. Unpublished doctoral dissertation, University of Ohio State.

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.

Jöreskog, K.G. (1977). *Structural Equation modeling in the social sciences: Specification, estimation and testing.* In P.R. Krishnaiah (Eds.), Application of Statistics (pp. 265- 287). North-Holland, Amsterdam.

Kaplan, D. (2000). *Structural Equation Modeling: Foundations and Extensions*. Thousand Oaks: Sage Publications.

Kaplan, D., & Elliot, R.P. (1997). A didactic example of multilevel structural equation modeling applicable to study of organizations. *Structural Equation Modeling, 4*(1), 1-4.

Kozlowski, S.W., & Klein, K.J. (2000). *A multilevel approach to theory and research in organizations: Contextual, temporal, and emergent properties*. San Francisco: Jossey Bass.

Lee, S.Y. (1990). Multilevel analysis of structural equation models. *Biometrika, 77,* 763-772.

Marsh, H.V., & Rochi, L.A. (1997). Making student's evaluations of teaching effectiveness effective. *American Psychologist, 52*(11), 1187-1197.

Marsh, H.W. (1987). Student's evaluations of university teaching: Research findings, methodological issues, and directions for future research. *International Journal of Educational Research, 11*(3), 263-388.

Marsh, H.W., Overall, U.J., & Kesler, P.S. (1979). Validity of student evaluation of instructional effectiveness: A comparison of faculty self-evaluations and evaluation by their students. *Journal of Educational Psychology, 71*(2), 149-160.

Marsh, H.W., & Rochi, L.A. (2000). Effects of granting leniency and low workload on students' evaluations of popular myth, bias, validity, or innocent bystanders. *Journal of Educational Psychology, 71*(2), 149-160.

Muthén, B.O. (1989). Latent variable modeling in heterogeneous populations. *Psychometrika, 54*, 557-585.

Muthén, B.O. (1990). *Mean and covariance structure analysis of hierarchical data*. UCLA, Statistics Series.

Muthén, B.O. (1991). Multilevel factor analysis of class and student achievement components. *Journal of Educational Measurement, 28*(4), 338-354.

Muthén, B.O. (1994). Multilevel covariance structure analysis. *Sociological Methods and Research, 22*, 376-398.

Muthén, B.O., & Sattora, A. (1989). Multilevel aspects of varying parameters in structural models. In R.D. Bock (Ed.), *Multilevel analysis of educational data*. San Diego: Academic.

Nasser, F., & Hagtvet, K.A. (2006). Multilevel analysis of the effects of student and instructor/course characteristics on student ratings. *Research in Higher Education,47*(5), 559-590.

Oliver, R.L. (1980). A cognitive model of the antecedents and consequences of satisfaction decisions. *Journal of Marketing Research, 17*, 460-469.

Oliver, R.L. (1981). Measurement and evaluation of satisfaction processes in retail settings. *Journal of Retailing, 57*(3), 35-48.

Rabe-Hesketh, S., Skrondal, A., & Zheng, X. (2007). Multilevel structural equation modeling. *Handbook of Latent Variable and Related Models* (pp. 209-227). Amsterdam: Elsevier.

Roest, H., & Pieters, R. (1997). The nomological net of perceived service quality. *International Journal of Service Industry Management, 8*(4), 336-351.

Roundenbush, S.W., & Bryk, A.S. (2002). *Hierarchical Linear models: Applications and data analysis methods.* Thousand Oaks, CA:Sage.

Roundenbush, S.W., & Willams, J.D. (Eds.). (1991). *Schools, classrooms, and pupils: International studies of schooling from a multilevel perspective.* Orlando: Academic press.

Stapleton, L.M. (2006). Using multilevel structural equation modeling techniques with complex sample data. *Structural Equation Modeling: A second course,* 345-383.

Appendix 1-Items On The Evaluation Questionnaire

Part I. Perceived Instructional Effectiveness
X1.The instructor taught the course according to the syllabus.
X2.The knowledge and ability of instructor was sufficient.
X3.Preparation of course was done by instructor before the courses.
X4.The instructor had a good communication with students.
X5.The instructor provided clear understanding of the course.
X6.The instructor used course and teaching equipment effectively.
X7.The instructor encouraged student participation and/or discussion.
X8.The instructor encouraged the students to research (library, internet etc.)
Part II. Student Loyalty
Y. I would like to take another course from the same instructor.

GENİŞLETİLMİŞ ÖZET

Özellikle son yıllarda üniversitelerdeki eğitim ve öğretim faaliyetlerinin geliştirilmesi kapsamında, öğretim üyesi etkililiğinin ve kalitesinin ölçülmesi için öğrenci değerlendirme formları kullanılmaktadır. Bununla beraber bu formların geçerliliği hakkındaki tartışmalar da halen devam etmektedir. Bu konuda ileri sürülen fikirlerden biri, öğrenci memnuniyetinin öğretme kalitesini ifade ettiği şeklindedir. Bir diğeri de öğrenci memnuniyetinin ve öğretme etkililiğinin farklı kavramlar olduğudur. Buna göre öğrencinin öğretim üyesinden memnun olması, öğretim üyesinin de etkili olduğu anlamına gelmemektedir. Öğretim üyesi etkililiğinin öğrenci memnuniyetine etkisinin araştırıldığı çalışmalarda, öğretim üyesi etkililiğinin öğrenci memnuniyeti üzerinde olumlu etkisinin olduğu sonucuna da varılmıştır.

Öğretim üyesi etkililiği, öğrenci memnuniyeti ve öğrenci sadakati kavramları temellerini pazarlama araştırmalarından almaktadırlar. Pazarlama araştırmaları alanında yapılan deneysel çalışmalarda müşteri memnuniyeti, algılanan ürün/hizmet kalitesi, müşteri sadakati, ürünü/hizmeti yeniden satın alma niyeti vb. kavramlar hemen her çalışmada kullanılmakta ve bu kavramlar ayrı ayrı ele alınmaktadırlar. Pazarlama teorisine göre algılanan kalite ve müşteri memnuniyeti, müşteri sadakatinin belirleyicileri olmaktadırlar. Algılanan kalite müşterinin ürün veya hizmet sağlayıcının standardı hakkındaki genel değerlendirmesidir. Müşteri memnuniyeti ise, ürün/hizmet sağlayıcının, müşterinin istek, beklenti ve ihtiyaçlarını karşılaması bağlamında müşteri tarafından hissedilen genel memnuniyet ve iç huzurdur. Müşteri sadakati ise en genel tanımıyla müşterinin ürünü/hizmeti aynı sağlayıcıdan almaya devam etme isteğidir. Bu bağlamda, bu kavramlar eğitim alanında ele alındıklarında algılanan kalite, algılanan öğretim üyesi etkililiğine, müşteri sadakati ise öğretim üyesine karşı öğrenci sadakatine karşılık gelmekte ve pazarlama araştırmalarında olduğu gibi algılanan öğretim üyesi etkililiğinin öğrenci sadakatini pozitif olarak etkilediği düşünülmektedir. Burada tekrar üzerinde durulması gereken nokta algılanan öğretim üyesi etkililiğinin pazarlama literatüründeki ürün/hizmet kalitesine denk geldiği ve öğrenci memnuniyeti olmadığıdır.

Eğitim alanında daha önce yapılmış olan çalışmalarda, yukarıda verilen kavramlar arasındaki ilişkilerin incelenmesinde, öğrencilerin öğretim üyelerinde gruplandığı dikkate alınmayarak, yalnızca öğrencilerin örnekleme birimler oldukları veya yalnızca öğretim üyelerinin ele alınmasıyla, her öğretim üyesi için öğrencilerinden elde edilen skorları kullanılması şeklinde iki tür veri kullanılmıştır.

Fakat bu iki yaklaşımın da bazı dezavantajları bulunmaktadır. Bu yaklaşımları ilki olan ve dağıtma adı verilebilen yaklaşımda, öğrencilerin öğretim üyeleri üzerinde yuvalanmış olduğu, aşamalı veri yapısı göz ardı edilmektedir. Böylece istatistiksel analizlerin en temel varsayımı olan birimlerin bağımsızlığı varsayımı en başta ihlal edilmiş olmaktadır. Çünkü aynı öğretim üyesinden ders alan öğrencilerin, bu öğretim üyesine verdikleri puanlar arasında ilişki olacağı ve böylece de gözlemlerin bağımsızlığı varsayımının bozulacağı açıktır. Bu varsayım ihlali halinde tahminler sistematik hatalı olacak ve birinci tür hata da artacaktır. Veri analizinde kullanılan ikinci yaklaşım her bir öğretim üyesi için öğrenciler tarafından verilen puanların toplanmasıdır ve bu yaklaşıma da kısaca toplanma yaklaşımı denilebilmektedir. Bu durumda da örnek büyüklüğü yalnızca öğretim üyesi sayısı kadar olacak ve böylece büyük miktarda bilgi kaybı ortaya çıkacaktır. Ayrıca analizler sonucunda bulunacak sonuçların yorumlanmasında da sorunlar yaşanacaktır. Çünkü dağıtma ve toplanma yaklaşımlarından elde edilen veri aynı anlamda değildir. Verinin bu şekilde yuvalanmış olduğu durumlarda, çok aşamalı verideki bağımlılık yapısını telafi etmek için, çok aşamalı modellerden yararlanılmaktadır.

Bu makalenin asıl konusu algılanan öğretim üyesi etkililiğinin öğrenci sadakati üstündeki etkisinin araştırılmasıdır. Sözü edilen bu iki kavram latent (örtük) değişkenler olduklarından ve örtük değişkenler arasındaki ilişkiler incelenmek istendiğinden, Yapısal Eşitlik Modellerinin (YEM) kullanılması uygun olmaktadır. Bununla beraber, verinin aşamalı yapısı da hesaba katıldığında, modelleme için kullanılabilecek en uygun yaklaşımın Çok Aşamalı Yapısal Eşitlik Modelleri (ÇAYEM) olduğu açıktır. ÇAYEM gözlem birimlerinin hiyerarşik veya kümelenmiş olması durumunda geçerli istatistiksel çıkarsama yapılmasına olanak sağlayan, çok aşamalı modellerin ve YEM'in sentezi olmaktadır. Bu güçlü yöntem, öğretim üyeleri arası ve öğretim üyeleri içi değişkenliği eş anlı olarak fakat ayrı ayrı ele almaktadır.

Bu çalışmada kullanılan veriler, Türkiye'deki en büyük üniversitelerden biri olan Yıldız Teknik Üniversitesi'nin öğrenci ve öğretim üyelerinden elde edilmiştir. Ders/öğretim üyesi değerlendirme formlarının kullanılması üniversitede 2003 yılında başlamış olup hala devam etmektedir. Tesadüfi olarak seçilen 202 öğretim üyesinde ders alan toplam 17647 öğrenciye ait değerlendirme formlarından elde edilen veriler, bu çalışmanın verilerini oluşturmaktadır. Örnekteki öğretim üyelerinin %68.4'ü erkek, %31.6'sı kadındır; %54.6'sı doçent veya profesör, %45.4'ü öğretim görevlisi veya yardımcı doçenttir. Sistematik hatalı cevaplardan kaçınmak amacıyla, öğrenci değerlendirme formlarında öğrenci bilgileri sorulmamıştır.

Algılanan öğretim üyesi etkililiğini ölçen sekiz madde ve öğrenci sadakatini ölçen tek madde Ek1'de verilmektedir. Soruların cevaplandırılmasında 5 noktalı Likert ölçek kullanılmıştır. Algılanan öğretim üyesi etkililiği için güvenilirlik katsayısı da oldukça yüksektir (Cronbach's alpha=0.9048).

Bu çalışmadaki araştırma hipotezimiz, algılanan öğretim üyesi etkililiğinin, öğrenci sadakatini olumlu olarak etkilediği şeklindedir. Bu hipotez hem öğretim üyeleri arası modelde hem de öğretim üyeleri içi modelde reddedilmemektedir. Yani algılanan öğretim üyesi etkililiği arttıkça öğrencinin öğretim üyesinden tekrar ders alma isteği yani öğrenci sadakati de artmaktadır. Öğretim üyeleri içi modelde öğrenci sadakatinin algılanan öğretim üyesi etkililiği tarafından açıklanan varyansı %57 iken, öğretim üyeleri arası modelde bu oran %92 olmaktadır. Bunun asıl sebebi öğretim üyeleri arası modelde ölçme hatalarının etkisinin yok denecek kadar az olmasıdır. Çalışmadaki ikincil amacımız, öğretim üyeleri arası modelde, öğrenci sadakatinin öğretim üyesinin akademik ünvanından, öğretim üyesinin algılanan etkililiğinin de verilen ders sayısı, toplam ders saati, öğretim üyesinin yayınlarının ilgili bölüm yayınlarına oranı olan yayın oranından, ve öğretim üyesinin cinsiyetinden etkilendiği şeklindedir. Bu hipotezlerden yalnızca ilki kabul edilirken, diğerleri reddedilmiştir. Yani öğretim üyeleri arası modelde, öğretim üyesinin ünvanı ve algılanan öğretim üyesi etkililiği arttıkça öğrenci sadakati de artmaktadır. Diğer taraftan öğretim üyesinin göreli yayın sayısının, verdiği farklı ders sayısı ve toplam ders saatinin ve cinsiyetinin algılanan öğretim üyesi etkililiği üzerinde anlamlı bir etkisi bulunamamıştır. Sonuç olarak öğretim üyesinin akademik performansı öğrenciler tarafından algılanan öğretimsel etkililiğini etkilememekte, algılanan öğretim üyesi etkililiği ve öğretim üyesinin akademik ünvanı ise öğrencilerin öğretim üyesine sadakatini arttırmaktadır.