Asia Pacific Education Review
2008, Vol. 9, No.4, 475-486.

Multilevel Analysis of the Effects of Student and Course Characteristics on Satisfaction in Undergraduate Liberal Arts Courses

Sun-Geun Baek **Hyo-Jeong Shin**
Seoul National University
Korea

The purpose of this study is to analyze the effects of student and course characteristics on student satisfaction in courses by using a two-level hierarchical linear model (HLM). Based on literature reviews, 13 research variables (① student gender, ② student academic year, ③ student major, ④ student reason for taking the course, ⑤ student level of course participation, ⑥ student expected grade, ⑦ student achieved grade, ⑧ faculty gender, ⑨ faculty age, ⑩ faculty status, ⑪ academic field of the course, ⑫ class size, and ⑬ course type) were selected and specified as fixed effects in the analysis model. Data were 57,216 ratings of 1,481 undergraduate liberal arts courses at Seoul National University in 2006. The result of unconditional model analysis revealed that student characteristics' effects on student satisfaction in courses (within-course effects, 88.1%) were much larger than course characteristics' effects (between-courses effects, 11.9%). The result of conditional model analysis specifying student and course level predictors revealed that those 12 research variables, with the exception of student gender, had statistically significant effects on student satisfaction in courses. The explained variance was 22.0% in student level, 65.8% in course level, and 27.2% of the total variance.

Key words: hierarchical linear model (HLM), student satisfaction in courses, student characteristics, course characteristics, teaching competence

Introduction

Since the late 1990's, colleges and universities in Korea have been conducting surveys measuring student satisfaction in courses (or student ratings of instruction) in order to increase institutional accountability and improve lecturers' teaching competencies (Baek, 2006; Choi, 2002; Shin, 2005; Yum, 2008). In recent years, about 82% of colleges and universities in Korea use measurements of student satisfaction in courses for the purpose of improving the quality of instructions and about 58% for the purposes of helping faculty reflect upon their teaching, in deciding on their retention, as well as the granting of tenure, and promotion (Han, 2005a).

In order to survey students' responses, some questionnaires on student satisfaction in courses have been developed and widely utilized in the United States. For instance, the SEEQ (Student Evaluation of Educational Quality) developed by Marsh (1987, 1991) has measured student ratings of instruction in 9 categories: ① learning and value, ② instructor enthusiasm, ③ organization clarity, ④

Sun-Geun Baek, Department of Education, Seoul National University, Korea; Hyo-Jeong Shin, Depaerment of Education, Seoul National Uniersity, Korea.

An earlier version of this paper was presented at the 8th International Conference on Education Research, Seoul National University, Seoul, Korea, October 23-25, 2007.

Correspondence concerning this article should be addressed to Sun-Geun Baek, Department of Education, Seoul National University, Gwanak-gu, Seoul, Korea. E-mail: dr100@snu.ac.kr

group interaction, ⑤ individual rapport, ⑥ breadth of coverage, ⑦ examinations/grading, ⑧ assignments/readings, and ⑨ workload. SRS (Student Rating System) developed in Kansas State University (IDEA, 2007) and SIR II (Student Instructional Report II) developed by Education Testing Service (ETS, 2006) are widely used to evaluate the instructional quality from the students' perspectives. Additionally, colleges and universities in Korea have tended to develop their own student course evaluation scales (Baek & Shin, 2008; Han, 2005b; Han, et al., 2005; Kim, et al., 2001). For instance, Seoul National University developed its own SCSS (Student Course Satisfaction Scale) and has been using it since 2003 (Byun & Kim, 2003).

Despite the current popularity of and demands for student ratings of instruction, there still remain some issues to be examined empirically. The main stream of research so far has focused on finding out what student, lecturer or course characteristics may affect student satisfaction in courses (Cohen, 1981; Kulik, 2001; Marsh & Roche, 1997; McKeachie, 1997). According to Centra's (1993) and Feldman's (1998) defining of "potential bias", student ratings are likely to depend on non-instruction factors such as lower workload, courses with lenient grading, or instructor showmanship, which are irrelevant to teaching competence per se. At this point, we must determine the validity of the results of measurements of student satisfaction in courses, which are basically evaluations of instructional quality or the instructor's teaching competence.

As for research methods, many empirical research exercises have been conducted through ANOVA or regression analysis, and thus these exercises have limitations in terms of their sometimes inappropriate application of units of analysis (Ethington, 1997; Feldman, 1998). In other words, many exercises used either individual data or the mean score of each class. As a matter of fact, a clustered student group of each course as well as lecturers is multi-level data, and thus the class context or lecturer characteristics for the course affecting students' responses should be taken into account. In this respect, if individual student response data were used as the unit of the analysis, the independence assumption would be violated. Moreover, its variance is inflated. On the other hand, if the mean score of each class were used, an aggregation bias appears. In order to solve these problems, a hierarchical linear model (HLM), which analyzes not only student-level data but also course-level data, might be more appropriate (Chin, 2007; Civian & Brennan, 1996; Nasser & Hagtvet, 2006; Raudenbush & Bryk, 2002; Tabachnick & Fidell, 2007; Ting, 2000; Umbach & Porter, 2002). The HLM is a data analysis technique for research designs where the data for participants is organized at more than one level. For example, there may be a measure of student achievement at the student level, and a measure of teacher's teaching skills at the classroom level. If a HLM were used, the data analysis can be protected from inflation related Type I errors, and variables can be specified in different levels at the same time (Kang, 1998; Kim, 2007; Hong, 2007; Raudenbush & Bryk, 2002).

In summary, a HLM appears to be the appropriate model to analyze the effects of student and course characteristics on student satisfaction in courses. In this respect, this study attempts to analyze the effects of student and course characteristics on student satisfaction in courses by using a two-level HLM.

Research Questions

In this study, the following research questions were proposed and investigated.

1. How much of student satisfaction in courses can be explained by student characteristics and course characteristics respectively?
2. What kind of student characteristics affects student satisfaction in courses?
3. What kind of course characteristics affects student satisfaction in courses?

Methodology

Data

The data collected for this research was derived from

undergraduate students' ratings of liberal arts education courses at Seoul National University in 2006. S/U (Successful/Unsuccessful) graded courses, 1 or 2 credit courses, and very small size classes (less than 5 students) were excluded from the analysis data. Therefore, 57,216 ratings for 1,481 courses were analyzed in this study.

Instrument and Variables

Student satisfaction in courses were collected using the SCSS (Student Course Satisfaction Scale) developed by and utilized in Seoul National University since 2003 (Byun & Kim, 2003). The questionnaire consists of 15 items measuring instructional quality for all of the undergraduate liberal arts courses using a 5-point Likert type scale as follows; ① strongly disagree, ② disagree, ③ neutral, ④ agree, and ⑤ strongly agree. Responses to the 15 items yielded high internal consistency (Cronbach's alpha = 0.970), and student ratings as a dependent variable was acquired in average student's responses to the 15 items (see Table 1).

13 variables used in this research were selected on the basis of literature reviews and some data were obtained from the responses of students in questionnaire (i.e., the reason for taking the course) as well as university's course records (i.e., student achieved grades). The variables included in research models were coded as follows:

Analysis Models

In this study, 4 research models were analyzed by using HLM: ① unconditional model, ② conditional model specifying student level predictors, ③ conditional model specifying course level predictors, and ④ conditional model specifying student and course level predictors. All of these predictors were specified as fixed effects, meaning that all of

Table 1
The Items of SCSS

No.	Items
1	I was able to check the syllabus or be informed of the course before the course begins.
2	The course was performed as it was outlined in the syllabus or better.
3	The contents of the course were well-organized.
4	The instructor appeared to be knowledgeable and experienced enough to teach the course.
5	The course was delivered in a challenging way which held my interest.
6	The instructor monitored student's process and corresponded with students sincerely.
7	The textbooks and reading materials were properly used and helpful.
8	The assignments were appropriate and useful for learning.
9	I was able to get feedback before or after the submission of the assignments.
10	Every student could actively participate in the class activities.
11	Grading was properly assigned based on the student's participation and achievement.
12	The teaching assistant was helpful at facilitating learning and other facilities were satisfactory.
13	The course level of difficulty and the coverage were appropriate.
14	The course was useful in helpful me to be competent in this academic field.
15	I would like to recommend this course to other students.

the variables have the same effects over the courses. Additionally, student's academic year, the degree of course participation, expected grade, achieved grade, faculty's age, and class size were grand-mean centered. Each research model is as follows:

Table 2
Research Variables

Levels	Variables		Codes
Level 1: Student	Gender	(SGENDER)	0 = male, 1 = female
	Academic Year	(SYEAR)	1 = freshman, 2 = sophomore, 3 = junior, 4 = senior
	Major	- (SOCIAL) (NATURAL) (ARTS)	1: humanities(reference group) 2: social science 3: natural science and engineering 4: arts and physical education
	Reason for Taking the Course	- (CAREER) (REQUIRE) (FOUNDATION) (COMPETENCE)	1: for getting high grades(reference group) 2: for preparing better jobs 3: for the requirement course 4: for the basic foundations of the major 5: for improving overall competence in this area
	Course Participation	(PARTICI)	5-point scale from 1 = very low to 5 = very high
	Expected Grade	(EGRADE)	0 = F, 1 = D, 2 = C, 3 = B, 4 = A
	Achieved Grade	(AGRADE)	0.0=F, 0.7=D-, 1.0=D0, 1.3=D+, 1.7=C-, 2.0=C0, 2.3=C+, 2.7=B-, 3.0=B0, 3.3=B+, 3.7=A-, 4.0=A0, 4.3=A+
Level 2: Course	Faculty Gender	(FGENDER)	0 = male, 1 = female
	Faculty Age	(FAGE)	instructor's age
	Faculty Status	(STATUS)	0 = tenured, 1 = non-tenured
	Academic Field of the Course	- (CSOCIAL) (CNATURAL)	1: humanities (reference group) 2: social science 3: natural science and engineering
	Class Size	(SIZE)	the number of students who registered the course
	Course Type	- (CORE) (GENERAL)	1: foundations of learning (reference group), 2: core courses 3: general education courses

Table 3
Analysis Models

Analysis Model	Equations
① Unconditional Model	$Y_{ij} = \gamma_{00} + u_{0j} + r_{ij}$
② Conditional Model Specifying Student Level Predictors	$Y_{ij} = \gamma_{00} + \gamma_{10}(SGENDER) + \gamma_{20}(SYEAR - \overline{SYEAR..})$ $+ \gamma_{30}(SOCIAL) + \gamma_{40}(NATURAL) + \gamma_{50}(ARTS)$ $+ \gamma_{60}(CAREER) + \gamma_{70}(REQUIRE) + \gamma_{80}(FOUNDATION)$ $+ \gamma_{90}(COMPETENCE) + \gamma_{100}(PARTICI - \overline{PARTICI..})$ $+ \gamma_{110}(EGRADE - \overline{EGRADE..})$ $+ \gamma_{120}(AGRADE - \overline{AGRADE..}) + u_0 + r$
③ Conditional Model Specifying Course Level Predictors	$Y_{ij} = \gamma_{00} + \gamma_{01}(FGENDER) + \gamma_{02}(FAGE - \overline{FAGE..})$ $+ \gamma_{03}(STATUS) + \gamma_{04}(CSOCIAL) + \gamma_{05}(CNATURAL)$ $+ \gamma_{06}(SIZE - \overline{SIZE..}) + \gamma_{07}(CORE) + \gamma_{08}(GENERAL) + u_0 + r$
④ Conditional Model Specifying Student and Course Level Predictors	$Y_{ij} = \gamma_{00} + \gamma_{10}(SGENDER) + \gamma_{20}(SYEAR - \overline{SYEAR..})$ $+ \gamma_{30}(SOCIAL) + \gamma_{40}(NATURAL) + \gamma_{50}(ARTS)$ $+ \gamma_{60}(CAREER) + \gamma_{70}(REQUIRE) + \gamma_{80}(FOUNDATION)$ $+ \gamma_{90}(COMPETENCE) + \gamma_{100}(PARTICI - \overline{PARTICI..})$ $+ \gamma_{110}(EGRADE - \overline{EGRADE..})$ $+ \gamma_{120}(AGRADE - \overline{AGRADE..})$ $+ \gamma_{01}(FGENDER) + \gamma_{02}(FAGE - \overline{FAGE..})$ $+ \gamma_{03}(STATUS) + \gamma_{04}(CSOCIAL) + \gamma_{05}(CNATURAL)$ $+ \gamma_{06}(SIZE - \overline{SIZE..}) + \gamma_{07}(CORE) + \gamma_{08}(GENERAL) + u_0 + r$

Results

Descriptive Statistics

Table 4 shows the summary of 57,216 ratings of students. As shown in Table 4, students who are female, in higher academic years, majoring in humanities or arts and physical education, taking the course for improving their overall competence in this area, and expecting higher grades tend to show higher satisfaction in courses. In addition, the correlation between student satisfaction in courses and students' course participation was 0.479 (p < .001). The average course participation was 3.80.

Table 5 shows the descriptive statistics of 1,481 undergraduate liberal arts courses. As shown in Table 5, female lecturers, non-tenured faculty, and humanities courses tend to show higher student satisfaction in courses. In addition, faculty's age was from 27 to 71, and average was 43.5 years old. Class size was from 5 to 261, and average was 40.3 people.

Multilevel Results

In this study, 4 research models were analyzed using HLM. Table 6 shows the multilevel results of each analysis model.

Table 4

Descriptive Statistics of 57,216 Ratings of Students

		No. of Cases	(%)	Student Satisfaction in courses	
				Mean	SD
Student Gender	Male	36,714	64.2%	3.77	0.86
	Female	20,502	35.8%	3.88	0.72
Student Academic Year	Freshman	24,344	42.5%	3.63	0.80
	Sophomore	11,344	19.8%	3.83	0.80
	Junior	8,175	14.3%	3.97	0.78
	Senior	13,353	23.3%	4.02	0.80
Student Major	Humanities	7,043	12.3%	3.93	0.75
	Social	15,792	27.6%	3.89	0.76
	Natural	30,939	54.1%	3.72	0.85
	Arts	3,442	6.0%	3.96	0.75
Reason for Taking the Course	High grades	2,552	4.5%	3.01	1.23
	Career	502	0.9%	3.47	1.01
	Requirement	19,282	33.7%	3.63	0.79
	Foundation	5,555	9.7%	3.78	0.76
	Competence	29,325	51.3%	4.00	0.72
Expected Grade	F	456	0.8%	3.28	1.13
	D	1,007	1.8%	3.39	0.91
	C	7,353	12.9%	3.44	0.79
	B	22,637	39.6%	3.73	0.74
	A	25,763	45.0%	4.00	0.82
Total		57,216	100.0%	3.81	0.82

Table 5

Descriptive Statistics of 1,481 Courses

		No. of Cases	(%)	Student Satisfaction in courses	
				Mean	SD
Faculty Gender	Male	937	63.3%	3.84	0.32
	Female	544	36.7%	3.96	0.30
Faculty Status	Tenured	462	31.2%	3.82	0.30
	Non-tenured	1,019	68.8%	3.92	0.33
Academic Field of the Course	Humanities	972	65.6%	3.97	0.29
	Social	165	11.1%	3.86	0.27
	Natural	344	23.2%	3.66	0.32
Total		1,481	100.0%	3.89	0.32

Table 6

Multilevel Results of 4 Analysis Models

Models		①Unconditional		②Student		③Course		④Student + Course	
		coefficient	S.E	coefficient	S.E	coefficient	S.E	coefficient	S.E
Student Satisfaction		3.868***	(0.008)	3.494***	(0.027)	3.945***	(0.023)	3.568***	(0.031)
Student Level Variables									
Student Gender				0.007	(0.006)			0.006	(0.006)
Year				0.071***	(0.003)			0.071***	(0.003)
Major	Social			-0.038**	(0.011)			-0.032**	(0.011)
	Natural			-0.068***	(0.011)			-0.054***	(0.011)
	Arts			0.045**	(0.015)			0.049**	(0.015)
Reason for Taking	Career			0.208***	(0.045)			0.206***	(0.045)
	Requirement			0.302***	(0.026)			0.300***	(0.026)
	Foundation			0.411***	(0.027)			0.416***	(0.027)
	Competence			0.466***	(0.025)			0.467***	(0.025)
Course Participation				0.300***	(0.005)			0.298***	(0.005)
Expected Grade				0.028***	(0.006)			0.029***	(0.006)
Achieved Grade				0.019**	(0.005)			0.018**	(0.005)
Course Level Variables									
Faculty Gender						0.052**	(0.016)	0.028*	(0.012)
Faculty Age						-0.003**	(0.001)	-0.004***	(0.001)
Status						-0.083***	(0.022)	-0.070***	(0.016)
Academic Field	Social					-0.086**	(0.024)	-0.034	(0.019)
	Natural					-0.249***	(0.023)	-0.090***	(0.017)
Class Size						-0.002***	(0.000)	-0.001***	(0.000)
Course Type	Core					0.045	(0.029)	-0.007	(0.021)
	General					0.115***	(0.018)	-0.036**	(0.015)
Variances									
Course Level		0.079		0.031		0.055		0.027	
Student Level		0.587		0.458		0.587		0.458	
Total		0.666		0.489		0.642		0.485	
ICC		0.119		0.063		0.086		0.056	
Explained Variance									
Course Level				0.608		0.304		0.658	
Student Level				0.220		0.000		0.220	
Total				0.266		0.036		0.272	

*p < .05; **p < .01; ***p < .001

① Unconditional Model

In the unconditional model, no research variables are specified. The use of an unconditional model as a basic step for in-depth analysis shows decomposed variances in both

student level (within group effects) and course level (between group effects) with intra-class correlation. Intra-class correlation means the proportion of variance in the dependent variable that is explained by group membership (Luke, 2004; Raudenbush & Bryk, 2002). In this study, the result of the unconditional model analysis showed that the intra-class correlation coefficient is 0.119, which means the proportion of explained variance by course membership is 11.9%. The remaining 88.1% of the total variance can be explained by student level predictors. Figure 1 shows the box-whisker plot of student ratings of 30 courses as random. As shown in Figure 1, student satisfaction in courses is much more different across individual students rather than the average student ratings over the courses.

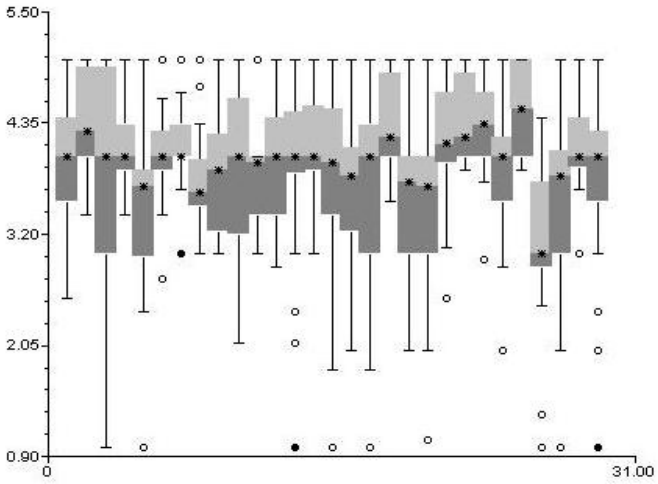

Figure 1. Box-whisker plot of student satisfaction in random 30 courses

② Conditional Model Specifying Student Level Predicators

The result of the conditional model analysis specifying student level predictors revealed that student's academic year, major, the reason for taking the course, the degree of course participation, expecting grade and achieved grade had statistically significant effects on student satisfaction in courses. In this model, the explained variance was 22.0% in student level, 60.8% in course level, and 26.6% of total variance.

In particlar, the relationship of student expected grade and course satisfaction needs to be more thoroughly examined. Most of the research on this issue agrees that the higher grades student expect, the higher ratings of instruction appear. However, this relationship can be interpreted as either student's preference for faculty or student's active course participation (Baek et al., 2005; Marsh & Roche, 2000; Nasser & Hagtvet, 2006).

From the perspective of faculty who were seen as being lenient in grading, the data in this study showed that the correlation between student satisfaction in courses and percentage of 'A' grades (A-, A0, A+) in course is fairly high (0.454, $p < .01$, see Figure 2). This indicates that students might assume that such faculty already issued grades before they take the course or respond to the student course evaluation scale.

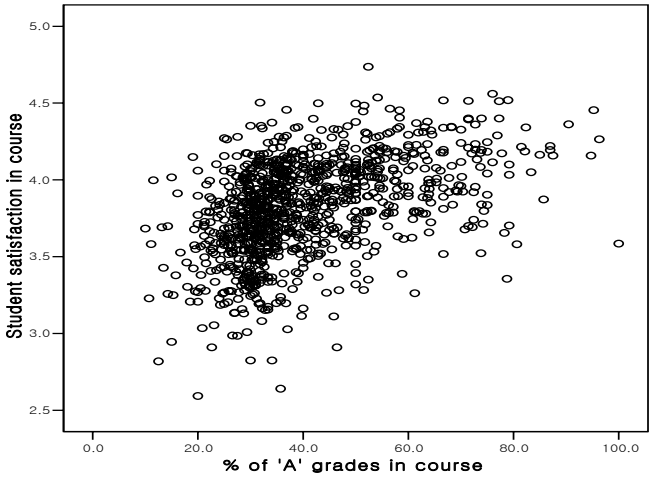

Figure 2. Relationship between % of 'A' grades in course and student satisfaction

On the other hand, from the perspective of student's active participation in the course, the data in this study showed that the correlation between student's course participation and expected grade is fairly high (0.461, $p < .01$). This indicates that student's active participation might have a number of positive effects on not only expected grade but also student satisfaction in courses.

In summary, student's course participation and the idea of faculty who were seen as being lenient in grading may influence student satisfaction in courses in a rather complicated way.

③ Conditional Model Specifying Course Level Predictors

The results of the analysis of the conditional model

specifying course level predictors revealed that faculty gender, faculty age, faculty status, academic field of the course, class size, and course type (basic, core, general) had statistically significant effects on student satisfaction in courses. In this model, the explained variance was 0.0% in student level, 30.4% in course level, and 3.6% of total variance.

In particular, the relationship between class size and course satisfaction needs to be more thoroughly examined. Researchers have suggested various ideas about the relationship between class size and course satisfaction (Baek et al., 2005; Civian & Brennan, 1996; Han, 2001, 2002; Ting, 2000). For the data in this study, the correlation between class size and student satisfaction in courses was -0.286 ($p < .01$, see Figure 3). This indicates that the more students who register in one class, the lower the level of student satisfaction in courses.

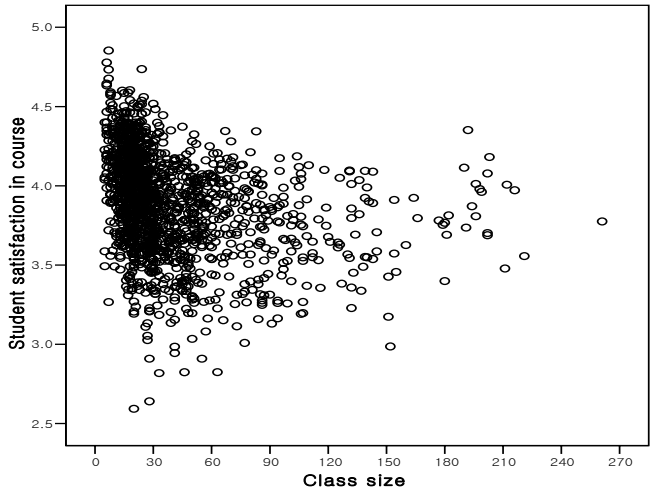

Figure 3. Relationship between class size and student satisfaction

In addition, this negative relationship between class size and student satisfaction in courses saw further variations in terms of the academic field of the course. Table 7 shows the correlation between class size and student ratings upon each academic field of the course for 1,216 courses consisting of 5-60 students. This implies that class size of natural sciences and engineering courses need to be reduced with a higher level of priority compared to humanities courses in order to improve students' satisfaction in courses.

In summary, student satisfaction in courses tends to be high when the class size is small, and this relationship may vary upon the academic field of the course.

Table 7
Correlations of Class Size and Student Satisfaction upon the Academic Field of the Course

Academic Field of the Course	No. of Courses	Correlation Coefficient
Humanities	886	-0.252**
Social sciences	79	-0.294**
Natural sciences and Engineering	251	-0.376**

④ Conditional Model Specifying Student and Course Level Predictors

The results of the conditional model analysis specifying student and course level predictors revealed that almost all of the variables had statistically significant effects on student satisfaction in courses. In this model, the explained variance was 22.0% in terms of student level, 65.8% in terms of course level, and 27.2% of the total variance.

In summary, if the other conditions are controlled, student satisfaction in courses may show higher ratings when:

- Students are in a higher academic year
- Students major in arts and physical education or humanities rather than social sciences or natural sciences and engineering
- Students take the course in order to improve their competence in that area or wish to acquire basic concepts of the major rather than taking the course to get high grades or to prepare for employment, or to satisfy other requirements
- Students participate the course more actively
- Students expect higher grades
- Students achieve higher grades
- Faculty's gender is female and the faculty is relatively young
- Faculty's status is tenured
- The course being taken is a humanities subject rather than a social science or, more particularly, a natural science and engineering course.
- The class size is small

Summary and Discussion

The purpose of this study was to analyze the effects of student and course characteristics on student satisfaction in courses by using a two-level hierarchical linear model (HLM). Based on literature reviews, 13 research variables (① student gender, ② student academic year, ③ student major, ④ student reason for taking the course, ⑤ student level of course participation, ⑥ student expected grade, ⑦ student achieved grade, ⑧ faculty gender, ⑨ faculty age, ⑩ faculty status, ⑪ academic field of the course, ⑫ class size, and ⑬ course type) were selected and specified as fixed effects in the analysis models. Data were 57,216 ratings of 1,481 undergraduate liberal arts courses at Seoul National University in 2006. The results of this study are summarized as follows.

First, the result of the unconditional model analysis showed that the intra-class correlation coefficient is 0.119, which means that the proportion of explained variance by course membership is 11.9%. The remaining 88.1% of the total variance can be explained by student level predictors.

Second, the results of the conditional model analysis specifying student level predictors revealed that student's academic year, major, the reason for taking the course, the degree of course participation, expecting grade and achieved grade had statistically significant effects on student satisfaction in courses. In this model, the explained variance was 22.0% in terms of student level, 60.8% in terms of course level, and 26.6% of the total variance.

Third, the result of the conditional model analysis specifying course level predictors revealed that faculty gender, faculty age, faculty status, academic field of the course, class size, and course type (basic, core, general) had statistically significant effects on student satisfaction in courses. In regards to this model, the explained variance was 0.0% in terms of student level, 30.4% in terms of course level, and 3.6% of the total variance.

Fourth, the result of the conditional model analysis specifying student and course level predictors revealed that all of the research variables, except student gender, had statistically significant effects on student satisfaction in courses. In this model, the explained variance was 22.0% in student level, 65.8% in course level, and 27.2% of total variance.

In conclusion, student characteristics had much greater effects on student satisfaction in courses (within-course effects, 88.1%) than course characteristics (between-courses effects, 11.9%). Therefore, there are many restrictions on the use of student ratings as indicators of the course quality or teaching competence.

However, there are some limitations in generalizing this study's results because the data was gathered from only one university. To generalize and confirm the results, more university cases should be analyzed in the future. In addition, insincere responses from students should be controlled to improve the quality of data. Almost 40% of the analyzed data had exactly the same responses across all of the items. Such potentially insincere responses can be obstacles to gathering more reliable student ratings of instruction (Dey, 1997; Johnson, 2003; Kim, 2005; Kim, 2006). Empirical research, done to filter such insincere responses are required. Lastly, additional student and course characteristics that might have influential effects upon student satisfaction in courses should be explored, and various kinds of statistical analysis models, including random effect models should be analyzed to confirm and improve the results of this study.

References

Baek, S. (2006). A review of course evaluation in colleges and universities. *Educational Development, 157*, 22-24.

Baek, S., Cho, Y., Jung, S., & Kim, J. (2005). *An evaluation study on the general education in Seoul National University*. Faculty of Liberal Education in Seoul National University.

Baek, S., & Shin, H. (2008). Multilevel analysis of the effects of student and course characteristics on student course evaluation: Focused on the undergraduate liberal education program. *Journal of Educational Evaluation, 21*(2), 1-24.

Byun, C., & Kim, J. (2003). *Research for developing student course satisfaction scale in Seoul National University*. Seoul National University.

Centra, J. A. (1993). *Reflective faculty education*. San Francisco: Jossey-Bass.

Chin, T. (2007). Comparison of three growth modeling techniques in the multilevel analysis of longitudinal academic achievement scores: Latent growth modeling, Hierarchical linear modeling, and longitudinal profile

analysis via multidimensional scaling. *Asia Pacific Education Review, 8*(2), 262-275.

Civian, J. T., & Brennan, R. T. (1996). Student and course factors predicting satisfaction in undergraduate courses at Harvard University. paper presented at the annual meeting of the American Educational Research Association.

Choi, G. (2002). Study on the evaluation system for institutional accreditation. Unpublished doctoral dissertation, Yonsei University, Seoul, Korea.

Cohen, P. A. (1981). Student rating of instruction and student achievement: A meta-analysis multi-section validity studies. *Review of Educational Resesarch, 51*(3), 281-309.

Dey, E. L. (1997). Working with low survey response rates: The efficacy of weighting adjustments. *Research in Higher Education, 38*(2), 215-227.

Ethington, C. A. (1997). A hierarchical linear modeling approach to studying college effects. In J. C. Smart(Ed.), *Higher Education: Handbook of theory and research* (Vol. 12, pp. 165-194). NY: Agathon Press.

ETS (2006). *The Student instructional report Ⅱ: Its development, uses and supporting research.* Retrieved June 1, 2006, from http://www.ets.org/.

Feldman, K. A. (1998). Reflections on the study of effective college teaching and student ratings: one continuing quest and two unresolved issues. In J. C. Smart(Ed.), *Higher Education: Handbook of theory and research*(Vol. 13, pp. 35-74). NY: Agathon Press.

Han, S. (2005a). Research on the current student ratings in course and its efficient way of operation, paper presented at the annual meeting of Council of the Dean of Academic Affairs.

Han, S. (2005b). Research on the items of student ratings in college courses, paper presented at the annual meeting of Council of the Dean of Academic Affairs.

Han, S. (2001). Research on student ratings of university teaching: Analysis of determinants related to professor, student and class. *The Journal of Educational Administration, 19*(4), 247-266.

Han, S. (2002). A study of the effect of class size on student ratings of university teaching. *Korean Journal of Higher Education, 13*(1), 155-176.

Han, S., Kim, H., & Lee, J. (2005). A comprehensive study of Korean students' evaluations of university teaching. *The Journal of Educational Administration, 23*(3), 379-403.

Hong, S. (2007). *Hierarchical and longitudinal data analysis*. Department of Social Welfare in Yonsei University.

IDEA (2007). *Instructional development and effectiveness assessment student ratings of instruction.* Retrieved April 15, 2007, from http://www.idea.ksu.edu.

Johnson, T. D. (2003). Online Student Ratings: Will students respond? *New Directions for Teaching and Learning, 96*, 49-59.

Kang, S. (1998). Analytical comparisons between classical linear models and multilevel models as educational and social research methods. *Journal of Educational Evaluation, 11*(1), 207-258.

Kim, K. (2007). *Theories and Practices of Multilevel Analysis*. The 6th research methods workshop of the Korean society of educational psychology in 2007.

Kim, M. (2005). Validity and reliability of lecture evaluation. *Asian Journal of Education, 6*(3), 1-24.

Kim, S. (2006). Optimizing measurement conditions for students' evaluation of e-learning system using generalizability study. *Journal of Educational Evaluation, 19*(1), 305-322.

Kim, S., Park, J., & Kim, J. (2001). *A Study on the Development of Faculty Evaluation Model Focused on Teaching by Course Evaluations*. Research report on policy of Ministry of Education in Korea.

Kulik, J. A. (2001). Student ratings: Validity, utility, and controversy, *New Directions for Institutional Research, 109*, 9-25.

Luke, D. A. (2004). *Multilevel modeling*. CA: Sage Publications.

Marsh, H. W. (1987). Student's evaluation of university teaching: Research findings, methodological issues, and directions for future research. *International Journal of Educational Research, 11*(3), 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., & Roche, L. A. (1997). Making students' evaluations of teaching effectiveness effective: The critical issues of validity, bias, and utility. *American Psychologist, 52*(11), 1187-1197.

Marsh, H. W., & Roche, L. A. (2000). Effects of grading leniency and low workload on students' evaluations of teaching: Popular myth, bias, validity, or innocent bystanders? *Journal of Educational Psychology, 92*(1), 202-228.

McKeachie, W. J. (1997). Student ratings: The validity of use. *American Psychologist, 52*(11), 1218-1225.

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.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods*. Advanced quantitative techniques in the social sciences series 1, 2nd ed., CA: SAGE Publications. Shin, H. (2005). Hangook godung kyoyook gaehyuk jungchaek [*Korean Policy of Higher Education Reform]*, Seoul: Hakjisa.

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). NY: Pearson Education, Inc.

Ting, K. (2000). A multilevel perspective on student ratings of instruction: lessons from the Chinese experience. *Research in Higher Education, 41*(5), 637-661.

Umbach, P. D., & Porter, S. R. (2002). How do academic departments impact student satisfaction?: Understanding the contextual effects of departments. *Research in Higher Education, 43*(2), 209-234.

Yum, S. (2008). Validating students' ratings of teaching scale and analyzing multilevel models. *Journal of Educational Evaluation, 21*(2), 25-52.

Received November 1, 2007
Revision received May 29, 2008
Accepted June 21, 2008