

# A Study Analyzing Current Courses at Seoul National University and Suggesting Directions for Course Improvement Based on Analysis of Student Course Evaluation

Hyeree Min\*

Center for Teaching and Learning, Seoul National University

*Abstract* — The purpose of this study was to determine directions for improvement of teaching based on analysis of student course evaluations. The data analyzed were students' ratings of courses at Seoul National University, Korea, from 2008 to 2011 (7 semesters). The contents of the research conducted were: 1) comparison of course evaluation items, 2) comparison of all, core and major courses, 3) comparison of core liberal arts, general liberal arts and foundation studies, 4) comparison within the teaching method category of all, core and major courses, 5) comparison within the teaching method categories of core liberal arts, general liberal arts, and foundation studies, and 6) correlation analysis (① correlations between individual questions on satisfaction and the average course satisfaction, ② correlations between students' grades and satisfaction with courses.) The results of this study were as follows. First, inadequacies of the course evaluation questionnaire showed the need for overall modifications to the questionnaire. Second, the clear decrease of satisfaction with core courses specifically indicated the need for improvement of this category. Third, in-depth analysis of the issues with core courses, foundation studies scored the lowest. Fourth, the necessity for improvements of the teaching method category was raised.

(Revised on 19 February, 2014)

## 1. Introduction

Course evaluation at Korean universities has caused various controversies ever since it was introduced in the early 2000s. Despite such differing opinions, it has been used as a crucial method to measure the quality of college education. In fact, course evaluation fits well into the current trend of concentrating efforts to reinforce the educational potential of universities, as it has been recently recognized as a tool to measure and evaluate the quality of education. As such, its role does not end at evaluation; efforts are being made to utilize

the evaluation results to provide information on courses, analyze the progress of education and present a path to improve education as a whole. Some enterprising foreign universities open the results of course evaluation to the public in order to help students choose their courses. Following this trend, an increasing number of Korean colleges are also making such results available to the public, and lively discussions on how to constructively utilize these results are increasingly becoming common.

Despite this environment, insufficient data have been accumulated in relation to the utilization of course evaluation. Previous studies on course evaluations

---

\*) Correspondence: Center for teaching and Learning, Seoul National University, #151-742, 61dong 306, 1 Gwanak-ro, Gwanak-gu, Seoul, Republic of Korea  
E-mail: hrmin82@snu.ac.kr

were mostly based on the development and use of questionnaires for these evaluations (Lee 2001; Ji et al. 2006; Kim and Kim 2008; Eyum 2008) and the elements that influenced the results of course evaluation and learner perspectives (Han et al. 2005; Kim 2006; Baek and Shin 2008; Oh 2009; Kim 2009; Lee 2010; Min and Lee 2011; Jang 2012) . Therefore, they were insufficient to fulfill the expected role of providing methods through which satisfaction analysis could bring improvements to courses.

Hence, an analysis of course evaluations with diverse and long-term perspectives is required to provide suggestions to improve the quality of courses according to each university’s individual characteristics. When this has been achieved, the initial goal of conducting course evaluations, improvement of both courses and the overall quality of education will be fulfilled.

## 2. Method of Study

### 2.1 Study Participants

This study is an analysis of the level of satisfaction of students of Seoul National University (SNU) with their courses during seven semesters, from the first semester of 2008 through the first semester of 2011. Table 1 shows the data collected from each semester. Overall, a greater amount was collected from the first semesters than the second semesters, and the amount of data increased as

the years passed.

### 2.2 Questionnaire for Course Evaluation

SNU’s questionnaire for the analysis of student satisfaction consists of 15 In this study, some of these items were classified into seven subcategories (Learner participation, Syllabus, Interaction, Motivation, Textbook assignments and assessments, Learner comprehension (difficulty) and e-Learning) under the broad category of ‘Methods.’ Table 2 outlines the questions that were used for analysis, as well as those not used in the categorization above.

### 2.3 Result analysis and methodology

In this study, we analyzed the level of student satisfaction with courses provided at SNU from the first semester of 2008 through the first semester of 2011. For this, we analyzed all courses as a whole, as well as the subsections of core courses and courses taught for majors. In addition, core courses were further categorized as follows: core liberal arts, general liberal arts, and foundation studies. The satisfaction with courses of individual colleges was also analyzed by category and question.

Learn the level of satisfaction of students about their courses for 5-point Likert scale. 1-point means ‘not at all satisfied,’ 5-point means extremely satisfied.’ To check the appropriateness of the questions of the satisfaction

Table 1. Number of courses and participating students

Semester	Core		Major		All	
	Number of courses	Number of participating students	Number of courses	Number of participating students	Number of courses	Number of participating students
2008-1 <sup>st</sup>	1,157	31,447	1,845	36,330	3,002	67,777
2008-2 <sup>nd</sup>	1,112	28,058	1,801	37,317	2,913	65,375
2009-1 <sup>st</sup>	1,130	33,612	1,875	42,584	3,005	76,196
2009-2 <sup>nd</sup>	1,096	30,971	1,787	42,218	2,883	73,189
2010-1 <sup>st</sup>	1,164	36,658	1,820	46,783	2,984	83,441
2010-2 <sup>nd</sup>	1,125	34,038	1,787	47,619	2,912	81,657
2011-1 <sup>st</sup>	1,163	37,930	1,863	52,215	3,026	90,145

**Table 2. Categorization of questions**

Category	Subcategory	Number*	Items
Content	Lecturer's qualification	7	The lecturer had the required knowledge and experience for the course.
	Course content	6	The course dealt with essential content for each topic.
	General	3	I would recommend the course to my juniors and friends since it was very informative.
		14	The course taught me the content that I had expected.
Method	Learner participation	1	How many hours on average did you study per week outside class hours?
		2	Did you participate in this course with enthusiasm?
	Syllabus	4	The syllabus helped me plan and decide to take this course.
		5	The course content adhered to the syllabus.
	Interaction	8	The lecturer encouraged student participation during classes.
	Motivation	9	The lecturer used appropriate methods to motivate and interest students towards the course topic.
	Textbook, assignments and assessments	10	Textbooks were used appropriately to aid comprehension.
		12	Feedback on assignments and assessments aided comprehension.
		13	The method of assessment was appropriate to evaluate student participation and academic achievement.
	Learner comprehension (difficulty)	11	The level of difficulty was appropriate and intellectually challenging.
	e-Learning	15	eTL was appropriately used.

\*Number: the number of actual survey items

**Table 3. Classification of core courses (based on the second semester of 2010)**

Category	Description	Subcategories of courses
Foundation studies	Subjects that help students achieve fundamental abilities required at university level	Academic interaction
		Communicating in a foreign language I (English)
		Communicating in a foreign language II (other languages)
		Mathematical analysis and deduction
		Scientific thought and experimentation
		Use of information technology
		Logical analysis and deduction
Core liberal arts	Subjects that help students form clear views and perspectives on important issues relating to human life, society, nature	Literature and arts
		History and philosophy
		Society and ideology
		Nature and technology
		Life and environment
General liberal arts	Introductory subjects that allow students to accumulate a broad spectrum of intellectual achievement and experience activities from diverse departments	Korean and composition
		Languages and foreign cultures
		Literature and art
		History and philosophy
		Society and ideology
		Understanding of nature
		Basic science
		Physical education and others
Special lectures for foundational education		

questionnaire, Pearson correlation coefficient analysis was conducted between the average score of all questions and each question, and also between individual students' grades and their satisfaction with the courses to analyze the correlations.

The main subject of analysis, core courses, can be classified into three categories as indicated above: foundation studies, core liberal arts, and general liberal arts. Examples of courses belonging to each category are outlined in Table 3.

### 3. Results

In this research, we compared the levels of

satisfaction with engineering courses and normal courses (all SNU courses except those of the College of Engineering). The contents of the research conducted were as follows:

- 1) Comparison by subject categories
- 2) Comparison of the method category
- 3) Correlation analysis (items: – satisfaction with courses; number of students – satisfaction with courses; students' grades – satisfaction with courses)

For the first and second items, research was conducted in the three separate categories of core courses, major courses and all courses (core and major courses included).

**Table 4. Comparison of satisfaction levels of core and major subjects (2008-2011)**

Average	Core (N=30,164)		Major (N=32,071)		t-value
	Mean	Standard Deviation	Mean	Standard Deviation	
2008-1 <sup>st</sup>	3.64	0.76	3.72	0.80	-12.362 ***
2008-2 <sup>nd</sup>	3.64	0.81	3.68	0.83	-6.155 ***
2009-1 <sup>st</sup>	3.67	0.78	3.72	0.81	-7.749 ***
2009-2 <sup>nd</sup>	3.69	0.80	3.70	0.84	-1.996
2010-1 <sup>st</sup>	3.68	0.81	3.75	0.84	-12.504 ***
2010-2 <sup>nd</sup>	3.72	0.82	3.76	0.84	-6.315 ***
2011-1 <sup>st</sup>	3.67	0.82	3.73	0.84	-9.535 ***
Total	3.67	0.80	3.73	0.83	-21.937 ***

\*p<.05, \*\*p<.01, \*\*\*p<.001

To determine the level of satisfaction of students with their courses, we asked them to rate each item on the questionnaire on a 5-point Likert scale with the following answers: 1, not at all satisfied; 2, slightly satisfied; 3, moderately satisfied; 4, very satisfied; 5, extremely satisfied.

**Table 5. Averages for foundation studies, core liberal arts and general liberal arts courses**

Semester	Foundation studies	Core liberal arts	General liberal arts
2008-1 <sup>st</sup> *	3.65	3.61	3.56
2008-2 <sup>nd</sup> *	3.56	3.84	3.65
2009-1 <sup>st</sup> *	3.59	3.85	3.71
2009-2 <sup>nd</sup> *	3.63	3.80	3.71
2010-1 <sup>st</sup> *	3.60	3.74	3.73
2010-2 <sup>nd</sup> *	3.65	3.82	3.76
2011-1 <sup>st</sup> *	3.61	3.70	3.74
Total *	3.60	3.79	3.71

\* significant differences between the semesters

### 3.1 Comparison by subject categories

#### 3.1.1 Comparison of all, core and major courses

For all courses, the average level of satisfaction was highest in the second semester of 2010 and lowest in the second semester of 2008. The average satisfaction score for the seven semesters was 3.70. In the case of the core courses, the average satisfaction was highest in the second semester of 2010 and the lowest in the first and second semesters of 2008, with the total average of the seven semesters being 3.67. The satisfaction with major courses was highest in the second semester of 2010 and lowest in the second semester of 2008, with an overall average of 3.73.

The detailed comparison shown in Table 4 reveals that the satisfaction with courses in the students' majors was higher than that with core courses in all semesters. Significant differences were found in the following items in particular:

- Textbooks were used appropriately to aid comprehension.
- The level of difficulty was appropriate and intellectually challenging.
- The method of assessment was appropriate to evaluate student participation and academic achievement.
- eTL was appropriately used (exceptionally for this, core courses achieved higher scores than those for majors did).

Item-by-item analysis revealed that No. 7 (the lecturer had the required knowledge and experience for the course) achieved the highest score, and No. 1 (how many hours on average did you study per week outside class hours?) had the lowest.

Items that showed lower satisfaction than the average were No. 1 (how many hours on average did you

study per week outside class hours?), 12 (feedback on assignments and assessments aided comprehension) and 15 (eTL was appropriately used).

#### 3.1.2 Comparison of core liberal arts, general liberal arts and foundation studies

The average levels of satisfaction for core liberal arts, general liberal arts and foundation studies courses are presented in Table 5.

The overall average was highest for core liberal arts courses, followed by general liberal arts and finally foundation studies courses. Significant differences were found for the following items:

- How many hours on average did you study per week outside class hours?
- I would recommend the course to my juniors and friends since it was very informative.
- The syllabus helped me plan and decide to take this course.
- The lecturer used appropriate methods to motivate and interest students towards the course topic, and
- eTL was appropriately used.

### 3.2 Analysis of the Method category

#### 3.2.1 Comparisons within the Method category of all courses, core courses and major courses

The Method category contains items regarding the educational method of the lecturer, which can help evaluate the lecturer's course management skills and use of appropriate educational methods. A comparison according to Table 6 revealed that the 'interaction' section scored highest for all the categories. Core courses showed high levels of satisfaction in 'motivation' On the other hand, 'learner participation' scored low in all

**Table 6. Comparisons by course category and method category**

Category	Highest scoring section	Lowest scoring section	Sections below average
All courses	Interaction	Learner participation	Learner participation, eTL
Core courses	Interaction, motivation	Learner participation	Learner participation, eTL
Major courses	Interaction	eTL	Learner participation, eTL

courses and core courses and ‘use of eTL’ scored the lowest in major courses. The sections that scored below average were ‘use of eTL’ and ‘learner participation.’

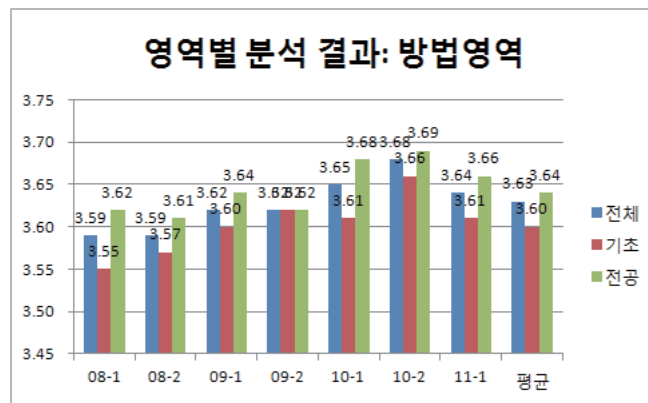
The satisfaction with all courses, core courses and major courses determined for the subcategories (sections)

of the Method category is presented in Table 7 and Figure 1 below. A detailed comparison of averages for core courses and major courses revealed that the level for the latter was higher than for the former except for the second semester of 2009. There were large disparities

**Table 7. Comparison of averages between core courses and major courses: Method category**

Semester	All	Basic	Major	Sections showing significant differences
2008-1 <sup>st</sup> *	3.59	3.55	3.62	Interaction Assignments and assessments Learner comprehension
2008-2 <sup>nd</sup> *	3.59	3.57	3.61	Interaction Assignments and assessments Learner comprehension
2009-1 <sup>st</sup> *	3.62	3.60	3.64	Motivation Learner comprehension eTL
2009-2 <sup>nd</sup> *	3.62	3.62	3.62	Learner participation (major > core) Assignments and assessments (major > core) eTL (core > major)
2010-1 <sup>st</sup> *	3.65	3.61	3.68	Interaction Assignment and assessments Learner comprehension
2010-2 <sup>nd</sup> *	3.68	3.66	3.69	Learner participation Assignments and assessments eTL (core > major)
2011-1 <sup>st</sup> *	3.64	3.61	3.66	Interaction Assignments and assessments Learner comprehension
Total *	3.63	3.60	3.64	Learner participation Assignments and assessments Learner comprehension

\* significant differences between the semesters



**Figure 1. Result of comparison by categories: Method category (left: All, middle: Core, right: Major)**

in ‘learner participation,’ ‘interaction,’ ‘assignments and assessments’ and ‘learner comprehension.’

In addition, Figures 2, 3, 4 and 5 presented below compare the scores of the questions that show large disparities (‘learner participation,’ ‘interaction,’ ‘assignments and assessments’ and ‘learner comprehension’). Through this it can be observed that

in this section the average of major courses was greater throughout all semesters and the differences in averages of the two categories of courses was large, indicating that core courses need improvements in these sections. In light of such results, lecturers of core courses should particularly focus on these categories to improve their course management skills.

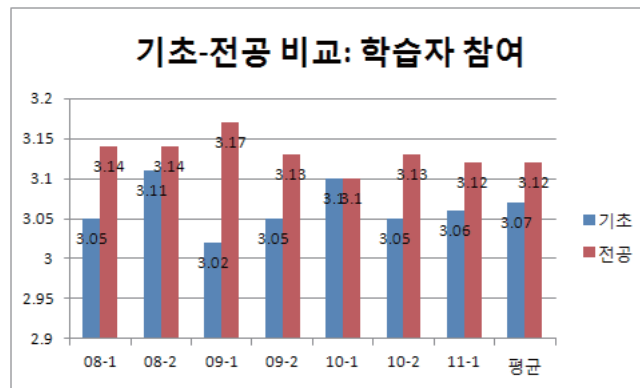


Figure 2. Comparison of satisfaction with core and major courses: learner participation (left: Core, right: Major)

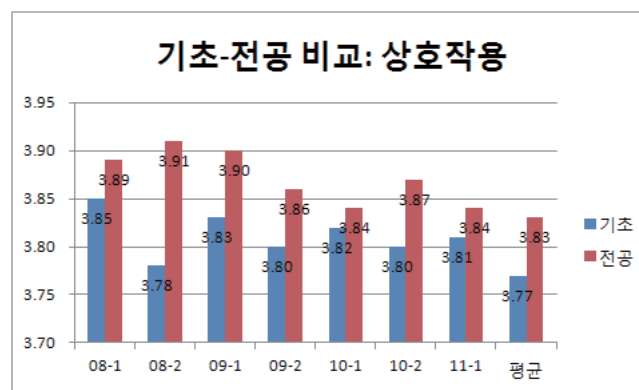


Figure 3. Comparison of satisfaction with core and major courses: interaction (left: Core, right: Major)

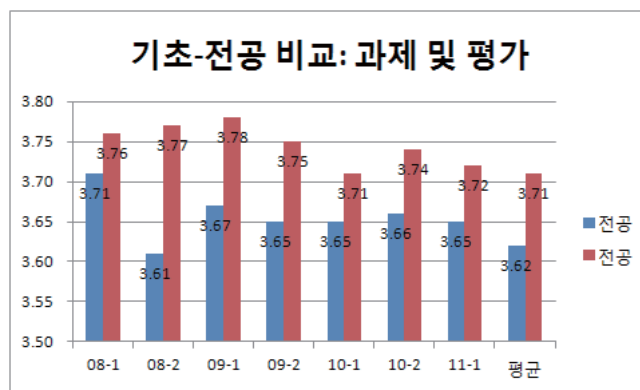


Figure 4. Comparison of satisfaction with core and major courses: assignments and assessments (left: Core, right: Major)

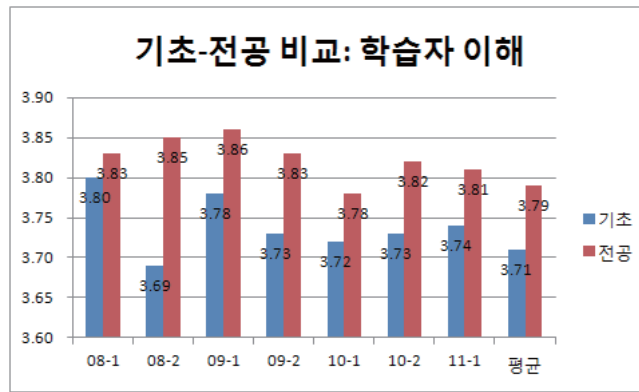


Figure 5. Comparison of satisfaction with core and major courses: learner comprehension (left: Basic, right: Major)

Table 8. Comparison of averages of core liberal arts, general liberal arts, foundation studies: Method category

Semester	Foundation studies	Core liberal arts	General liberal arts	Section with significantly different results
All *	3.52	3.72	3.65	<u>learners' participation</u> (core liberal arts > foundation studies > general liberal arts) <u>Motivation</u> (general liberal arts > core liberal arts > foundation studies) <u>eTL</u> (core liberal arts > general liberal arts > foundation studies)

\* significant differences between the semesters

Table 9. Pearson's correlation coefficient (r) between each question and average course satisfaction

Year -Semester	2008 -1 <sup>st</sup>	2008 -2 <sup>nd</sup>	2009 -1 <sup>st</sup>	2009 -2 <sup>nd</sup>	2010 -1 <sup>st</sup>	2010 -2 <sup>nd</sup>	2011 -1 <sup>st</sup>	Average
Question								
1	.403	.420	.396	.396	.401	.423	.411	.407
2	.777	.809	.799	.805	.807	.809	.802	.801
3	.862	.882	.870	.875	.874	.879	.874	.874
4	.806	.845	.823	.841	.844	.854	.850	.838
5	.810	.848	.828	.840	.833	.849	.836	.835
6	.844	.872	.852	.866	.860	.870	.861	.861
7	.791	.832	.807	.826	.820	.832	.821	.818
8	.819	.850	.833	.850	.844	.857	.845	.843
9	.861	.889	.873	.884	.882	.888	.883	.880
10	.794	.826	.804	.814	.816	.820	.819	.813
11	.857	.884	.868	.878	.876	.884	.876	.875
12	.818	.835	.825	.827	.834	.835	.838	.830
13	.848	.870	.853	.868	.867	.874	.868	.864
14	<b>.882</b>	<b>.903</b>	<b>.890</b>	<b>.900</b>	<b>.896</b>	<b>.904</b>	<b>.899</b>	<b>.896</b>
15	.547	.603	.600	.623	.641	.652	.650	.617



### 3.2.2 Comparison among averages of core liberal arts, general liberal arts, foundation studies

Table 8 below compares the averages of core liberal arts, general liberal arts and foundation studies. For all semesters the average satisfaction with core liberal arts and general liberal arts was higher than that with foundation studies. Satisfaction levels with core liberal arts and general liberal arts were quite similar, whereas for foundation studies the level fell short of the other two, showing significantly different results in the 'learner participation,' 'motivation' and 'use of eTL' sections.

### 3.3 Correlation Analysis

#### 3.3.1 Correlations between individual questions of the satisfaction survey questionnaire and the average course satisfaction

Item 14 (The course taught me the content that I had expected) was the question with the highest correlation with the average satisfaction (Table 9). On the other hand, Item 1 (How many hours on average did you study per week outside class hours?) and 15 (eTL was appropriately used) showed a relatively low correlation.

Admittedly, the result for Item 1 might be expected in the sense that hours spent for the course outside of classes cannot be considered as a factor contributing to satisfaction with the course. Item 15 shows that students do not consider the use of eTL to be an influential factor when evaluating a course, and as such the item should be deleted or amended appropriately. In particular, items such as No. 1 that focus on study hours dedicated to the class cannot be considered to be relevant with regard to satisfaction, and hence should be deleted.

#### 3.3.2 Correlation between students' grades and satisfaction with courses

The correlation between students' grades and course satisfaction was low (0.146 to 0.184) in all semesters. This suggested that, although the grades students received might have had a certain correlation with their satisfaction with the course, the extent was relatively low.

## 4. Results and Proposals

Analysis of SNU students' overall satisfaction with

their courses and comparison of satisfaction levels of major courses and core courses raised the following issues: there were problems with the course satisfaction questionnaire, lower satisfaction with core courses than major courses, and low satisfaction levels with foundation studies outside of core courses. Furthermore, meaningful results were obtained for the purposes of improving low satisfaction levels within the Method category. In this section, methods to improve SNU courses will be proposed based on a summary and analysis of these problems.

The first issue brought up by the research, inadequacies of the course evaluation questionnaire, showed the need for overall modifications to the questionnaire. The most problematic question was 'how many hours on average did you study per week outside class hours?' as it is difficult to identify the relationship between study hours and course satisfaction. In fact, the correlation for this item was 0.407, which was significantly lower than for the other items, which showed a correlation of at least 0.8. Therefore, its appropriateness in the satisfaction questionnaire should be reevaluated.

In addition, the item 'eTL was appropriately used' also showed a very low correlation compared to the other items. eTL is indeed an important tool for the smooth running of courses, but whether the extent of that management should be included to determine course satisfaction is questionable. The fact that the need or frequency of the use of eTL differs greatly based on the individual characteristics of each course strongly supports this view.

Second, the clear drop in satisfaction with core courses specifically called for the improvement of this category. Generally, the average satisfaction for courses in the students' majors was greater than that of core courses. The items that showed the greatest difference between core courses and those for majors related to the use of the teaching material, appropriate level of difficulty, method of assessment and appropriate use of eTL. Among these items, only the one relating to the appropriate use of eTL showed a greater satisfaction level in core courses than in major courses. This indicated that core courses should be highly emphasized in the efforts to improve course quality.

Third, to analyze issues with the core courses in depth, these courses were divided into three categories:

core liberal arts, general liberal arts and foundation studies. The results obtained by this analysis showed that satisfaction with both core and general liberal arts courses was higher than with foundation studies for every semester. According to the item-by-item analysis, students had complaints about the foundation studies syllabuses and the teachers' methods of motivating students. Analysis by category revealed that, for the Method category, great differences in satisfaction levels existed between the learner participation, motivation and use of eTL.

In contrast with core liberal arts and general liberal arts, foundation studies are subjects that help students achieve fundamental abilities required for learning (e.g. university level Korean, English, Calculus, Physics and Logic). These subjects are usually taken by freshmen, suggesting that they are significantly related to basic skills and academic abilities acquired through high school education. In addition, as some courses highly depend on lecturing rather than student presentations or discussion, it is hard to expect the participation or motivation levels to be high. Nevertheless, considering that these subjects are fundamental to students' further university education, improvement of these courses is particularly urgent.

Fourth, the necessity for improvements of the Method category was raised. Schematizing the results of course satisfaction for all categories, the Content category and the Method category, it was shown that in almost all colleges the average for all categories was situated somewhere between the results for the Content category and Method category. This indicated that students were generally more satisfied with the content, and that the method brought down the total average.

The result of comparing questions of the course satisfaction questionnaire regarding "Method" category showed that the responses for the 'interaction,' 'assignment and assessment' and 'learner comprehension' sections differed greatly, which suggested that greater attention should be given to these areas. From the learner-based education perspective currently emphasized in Korean education, 'interaction' and 'learner participation' are key methodologies for learner-based education. The low levels of satisfaction with these areas in particular revealed the lecture-centered course problem of Korean universities. Hence, more thought should be directed towards generating efficient teaching methods that can generate learner-based education and strengthen

interaction between students and lecturers.

Fifth, to ascertain the variables that affect satisfaction with courses, the correlation between grades and satisfaction was evaluated. The result revealed a very low correlation (0.146~0.184), which contradicted the widespread belief that lecturers who are generous with student grades will obtain better course evaluation scores. Arguably, this factor serves to enhance the credibility of course evaluations by students.

To conclude, proposals have been presented based on the satisfaction levels with courses of SNU from 2008 to 2011. If greater effort is dedicated to analyzing individual university's course evaluation results, and if data created through comparing such results from different universities are accumulated, we will be able to deduce further methods to increase the quality of education and courses at the university. This will allow course evaluations to fulfill their originally allocated role: actual use and application for improving the quality of university courses.

## References

- Chi Eunlim, Chang Jung-A (2006), "Validating a self-diagnosis instrument developed for university instructors and analyzing the related variables," *Journal of Educational Evaluation* **19**(1), 51-71
- Eunjoo Oh (2009), "How to improve course evaluation," *Journal of Educational methodology* **21**(2), 1-20
- Han Shin-il, Kim Hye-jeong, Lee Jung-yun (2005), "A Comprehensive Study of Korean Students' Evaluations of University Teaching," *The Journal of Educational Administration* **23**(3), 247-266
- Jang, YoungJoo (2012), "Analysis of the faculty and students' perspectives on student course evaluation," Dept. of Education Administrations & Higher Education, The Graduate School of Korea University
- Jin-Hwa Kim (2009), "The study of how influences in the environment of college education affects students' lecture evaluation : Based on data of Y university," *Department of Management Information Graduate School of Education*. Yeungnam University
- Kim, Hyunchul (2006), "Investigation of the determinants of teacher evaluation using ANOVA and categorical data analysis," *The Journal of Korean Education*

**33**(3), 121-144

- Kim, Sungsook, Kim Hakil (2008), "Developing and Validating Midsemester Students' Evaluations of College Teaching," *Journal of Educational Evaluation* **21**(1), 55-78
- LEE, Sun-A (2010), "A Study of Undergraduates' Perception on Evaluating Teacher Performance In Higher Education," Department of Education, The Graduate School Ewha Womans University
- Min, Hyeree, Lee, Heewon (2011), "The characteristics of good teaching based on the students' awareness : The comparative analysis of actual learning experiences between A university in Korea and oversea universities," *The Journal of Yeolin Education* **19**(1), 57-88
- Sun-Geun Baek, Hyo-Jeong Shin (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
- SungHeum Lee (2001), "A Study for Instrument Development of Course Evaluation in the University Based Upon the Theory of Instructional Design," *Journal of Educational Technology* **17**(1), 81-108
- Yum, Sichang (2008), "Validating Students' Ratings of Teaching Scale and Analyzing Multilevel Models," *Journal of Educational Evaluation* **21**(2), 25-52