

---

---

# 이공계 강의평가 결과의 실증적 분석을 통한 강의평가제도 개선방안

김학일\*, 김성숙\*\*, 권오양\*\*\*, 이천\*\*\*\*, 노경호\*\*\*\*\*

인하대학교 정보통신공학부 교수\*, SGInstitute Inc. 대표\*\*,  
인하대학교 기계공학부 교수\*\*\*, 인하대학교 전자전기공학부 교수\*\*\*\*,  
인하대학교 생명화학공학부 교수\*\*\*\*\*

## Improving Course Evaluation System of Engineering Education

Hakil Kim\*, Sungsook Kim\*\*, Oh-Yang Kwon\*\*\*, Cheon Lee\*\*\*\* and Kyung-Ho Row\*\*\*\*\*  
College of Engineering, Inha University\*\*\*\*\*,\*\*  
SGInstitute, Inc.\*\*

### 국문요약

대학의 강의평가는 수업효과성에 대한 진단적 피드백을 제공하고 교수에 대한 의사결정에 도움을 주며 학생이 수강 시 활용할 수 있는 정보를 제공하는 등 광범위하게 사용되고 있다. 그러나 활용범위가 확대될 수록 강의평가도구의 타당성과 신뢰성이 제기되고 있으며, 강의평가 제도 자체에 대한 측정학적 연구가 요구되고 있다. 본 연구의 목적은 이공계 대학의 강의평가 결과에 대한 실증적 분석을 통하여 강의평가제도의 순기능을 극대화하기 위한 개선방안을 제안하는 것이다. 평가도구의 타당도와 신뢰도, 평가결과에 대한 문항반응 유형 등을 통계적으로 분석하기 위해 국내 대학교 이공계 개설강의 중 471개 과목의 385명 교수에 대한 총 49,127 명 학생의 강의평가 결과자료를 사용하였다. 분석 결과, 첫째, 강의평가의 유형별 차이를 줄이기 위하여 전공별 또는 수업유형별로 평가문항 군에서 문항을 선택하도록 개선할 것을 제안하고, 둘째, 평가결과점수 상하위 10% 해당하는 강의의 특성을 분석하여 결과활용의 방향을 제시하였다.

### Abstract

An evaluation system of teaching is one of authentic assessment tools for improving the quality of higher education. The purpose of this study is to cultivate the class evaluation system in the college of engineering based on the empirical analysis of the results of the class evaluation. Especially, this study investigates the validity evidence using the confirmatory factor analysis of the class evaluation. The data used in this study were acquired from 49,127 student's evaluation

responses of 471 courses offered in colleges of natural science and engineering at a university in Korea. The reliabilities are quite good for every construct by producing an index value from 0.92 to 0.98. The results provides a guideline for an appropriate measurement model to report the information, to clarify quality and appropriateness of instrument items, to make recommendations for which items should be left or merged in the revised instrument. A special suggestion for improving student's evaluation of each course is to prepare a well-designed instruction for students explaining why and how to evaluate the course in order to produce reliable and valid results.

주제어: 강의평가, 평가도구, 타당도, 신뢰도

Keywords: Evaluation of teaching, Evaluation instrument, Validity, Reliability

## I. 서론

대학에서 시행하고 있는 강의평가의 기본 목적은 교수의 수업효과성을 평가하기 위함이다. 특히 강의평가 결과가 수업효과성에 대한 진단적 피드백을 제공하고 교수에 대한 의사결정에 도움을 주며 학생이 수강 시 활용할 수 있는 정보를 제공하는 등, 대학교육의 전반적인 질적 제고를 위한 긍정적인 역할을 담당해야 한다. 우리나라의 경우 각 강좌에 대한 학생의 강의평가는 1990년대 초반 몇몇 대학에서 시범적으로 실시하였다. 그러나 1997년 이후 공식적인 대학종합평가의 평가항목 중 강의평가제도 적용여부가 포함되면서 거의 모든 대학에서 의무적으로 시행하게 되었다. 따라서 강의평가제도 도입 초반에는 평가도구의 신뢰도, 타당도, 관찰의 일반화가능도, 점수산출 방법, 평가시행방법, 결과의 활용방안 등의 연구도 동반되었지만(김성숙, 1993; 송미섭·지은림, 1994), 대학내 강의평가가 정착된 이후 오히려 관련 연구도 소홀하게 전개되는 실정이다. 그동안 강의평가에 대하여 과연 (1)누가 평가를 받으며 누가 평가 하는가 (2)무엇을 평가해야하는가 (3)어떤 평가도구를 어떤 방법으로 시행해야하는가 (4)어떤 기준으로 평가해야하는가 (5)평가 결과는 어떻게 처리되며 어떻게 활용 되는가 (6)평가결과는 타당하고 신뢰할만한가 등, 다양한 문제를 제기해 왔다(Hoy & Miskel, 2005). 즉 평가가 지향하는 목적에 따라 평가도구의 구인과 문항내용, 평가방법과 기준 등이 다르게 구성되는 물론 평가결과를 어떻게 다르게 활용하는지도 차별되기 때문이다.

대학교육의 효과성을 보다 객관적으로 측정하기 위하여 강의에 대한 평가는 학생평가를 기본으로 교수의 자기평가, 포트폴리오, 동료평가, 전문가의 모니터링 등 다면적 평가를 제안하고 있다(Ory, 2001). 대부분 대학이 강의평가를 하는 목적은 강의의 질 개선을 최우선적으로 두고 있으나 최근 교수 업적평가에 항목을 포함함은 물론 우수교육 교수를 선발하는 근거로 활용하는 대학도 증가하고 있다(김성열 외, 2001). 그러나 이렇게 활용범위가 확대될수록 강의평가도구의 타당성과 신뢰성에 대한 문제점을 제기하며 측정학적 연구는 물론 강의평가 제도 자체에 대한 심층적 연구를 요구하고 있다(김성숙, 2006).

강의평가 시행에서 나타나는 특성은 학생에 의한 강의평가가 교실에서 무기명 온라인 방식으로 거의 전환하고 있다는 점이다. 자료수집과 관리 그리고 결과 처리의 편이성을 고려하여 초기 비용과 노력을 감안하더라도 온라인 평가로 급속히 증가하고 있다. 특히 우리나라의 경우 대학평가의

지표로서 강의평가제도의 적용은 물론 학생의 참여수준도 포함되어야 하므로 대부분 대학에서 성적열람의 선행 또는 수강신청의 선행조건으로 온라인 평가를 의무화하는 추세이다. 그러나 온라인 방식에서는 모든 학생이 수강하는 모든 강의에 대하여 성적열람 이전 또는 수강신청 이전에 수행하므로 학기가 지나갈수록 무성의한 답변으로 일관하고 있는 현상을 간과할 수 없다.

이와 같이 최근 대두되고 있는 강의평가제도의 문제점과 구체적으로 연계되는 연구는 첫째, 강의평가도구의 타당성과 신뢰성, 평가문항의 부적절성, 강의유형별 또는 규모별 문항의 무차별성 등 평가도구와 문항 개선의 필요성, 둘째, 점수 계량화 등 평가결과 점수 산출방법 개선의 필요성, 셋째, 강의평가 시기의 부적절성과 작성도구의 일방성 등 시행방법 개선의 필요성, 마지막으로 결과 활용에 대한 불신과 부적절한 활용 등 결과의 활용측면 개선의 필요성 등으로 요약할 수 있다.

이렇게 다양하게 제기되는 필요성에 근거하여 이 연구의 목적은 이공계 대학의 강의평가 결과를 효과적으로 활용하기 위하여 평가결과를 실증적으로 분석함으로써 현행 강의평가제도의 문제점을 정량적으로 분석하고, 강의평가제도에 대한 개선방법을 제안하는 것이다. 구체적인 연구범위는 첫째, 사용되고 있는 강의평가 결과의 신뢰도와 타당도를 점검하고, 둘째, 평가도구의 유형별 특성과 평가결과를 분석함으로써 강의유형과의 관계를 점검하며, 셋째, 강의평가점수에 따른 집단별 강의에 대한 문항반응유형을 분류함으로써 평가제도에 대한 시사점을 도출하는 것이다.

## II. 이론적 배경

최근 우리나라 대학의 강의평가 실태를 조사한 한신일·김혜정·이정연(2005) 연구는 대학의 강의평가에 대한 국내외 동향을 몇 가지 특성으로 요약하였다. 첫째, 교수의 책무성 차원에서 강의평가에 대한 인식이 확대되었으며 대학의 질을 결정하는 요인으로 교수능력과 책무성을 강조하면서 질적 관리에 부응하는 평가의 유용성이 중요하게 부각되고 있음을 설명하였다. 즉 강의평가는 단순한 수업개선이나 교수의 강의능력 향상 평가의 준거로서의 역할 뿐 아니라 대학교육의 총체적 질 관리 도구의 하나로 기능하고 있음을 의미하는 것이다. 둘째, 강의평가의 목적을 충실히 수행하려면 현재처럼 평가의 주체를 학생에 한정하지 말고 동료와 대학 내 전문기관 등을 통하여 다면적 평가를 수행해야 한다는 점이다. 왜냐하면 학생의 강의평가를 통해 교수의 수업개선의 목적은 달성할 수 있을지 몰라도 강의 효과나 결과 활용에 대한 거시적 의사결정을 위한 강의평가로는 한계가 있기 때문이다(Hoyt & Pallett, 1999).

국내 대학의 경우, 전체대학의 약 65% 정도가 단일 평가 설문지를 사용하고 있으며 수업특성별 평가설문지를 2개 유형에서부터 12개 유형까지 다양하게 나타나고 있다. 특히 대단위 강의식 수업에서부터 실험 실습 등 다양한 형태의 수업을 진행하고 있는 이공계 대학의 강의평가는 평가설문지 구성항목이 차별화되어야 한다. 그러나 평가도구 구성 항목을 보면 대부분 유사한 영역으로 교수자 관련 특성과 학습자 관련 성취문항, 그리고 수업관련 계획, 내용, 방법, 평가, 상호작용 등과 전체적 총평으로 구성되어 있다고 지적하였다(한신일·김혜정·이정연, 2005). 국내 대부분 대학이 수업의 특성이나 전공분야 교수법의 특수성을 고려하지 않은 채 단일한 유형을 사용하고 있는 것들 가장 큰 문제점으로 지적하였다.

국의 대학의 경우, 최근 퍼듀 대학교나 캔사스 주립대학교 등 미국에서 모범적으로 강의 평가를 실시하고 있는 대학의 경우 문항개발에서부터 관리와 결과활용에 이르기까지 전문평가기구에서 전담하므로 피드백이 원활하게 이루어지며 수업개선에 미치는 영향도 긍정적으로 나타나고 있다. McKee 와 Dowaliby(1985) 와 Feldman(1997) 의 연구에서도 각각 학생들의 강의에 대한 피드백이 교수의 수업개선에 반영된다는 신뢰가 형성되어야 긍정적인 참여를 유도할 수 있다고 지적하였다. 또한 강의평가가 일차적으로 수업개선이라는 본질적 목적을 달성해야한다는 입장과 함께 강의평가의 타당성과 신뢰성을 입증하려는 노력에도 불구하고 학생의 강의평가만으로 강의에 대한 전반적 개선을 위한 합리적 판단을 하기에는 부적절하다는 논의가 확산되고 있는 실정이다(Theall, Abrami, & Mets, 2001; Mckeachie, 1996).

미국의 경우 대부분의 강의평가를 대학운영의 필수과정으로 사용하고 있고 캐나다의 경우도 94% 이상 강의평가를 시행하고 있다고 보고하였다(Saroyan & Amundsen, 2001). 대부분 관찰적으로 사용한 설문지이며 설문의 내용 역시 대체로 강의의 구조와 내용, 교수와 학생의 상호작용, 의사전달의 명확성, 과목의 난이도, 과제 분량과 평가방법, 강의에 대한 충실도, 지식과 이해정도 등을 포함하고 있다(Marsh, 1994; Feldman, 1997). 한편, Bukalski 와 Zirpola(1993)는 미국에서 강의평가 제도를 도입한지 70여년이 지났음에도 불구하고 과연 근본적인 목적을 달성하고 있는지에 대한 비판과 함께 이 제도의 유용성과 무용성에 대한 상반된 결과를 지적하였다.

그동안 강의평가에 대한 피드백을 받은 집단과 받지 않은 집단을 비교할 때 학생들의 수업만족도와 학업성취수준이 향상되었다는 연구는 꾸준히 수행되었다 (Marsh & Roche, 1993; Wilson, 1986; Brinko, 1990; Murray, 1985; L'Hommendieu, Menges, & Brinko, 1990). 이러한 연구결과는 실제 교수학습 상황에서 연구된 것이 아니라 단기적 실험적 상황에서 수행된 것이므로 교수의 수업행동변화가 실제적으로 확인되었다고 보기는 어렵다. 따라서 최근 연구에서 강의평가가 실질적으로 교수의 수업의 질을 개선하는데 얼마나 영향력이 있는지 체계적인 검증이 요구된다고 지적하였다(Spooren & Mortelmans, 2006; Algozzine, et al., 2004; Grimes, Millea, & Woodruff, 2004; Shevlin, et. al., 2000).

한편, Broder와 Dorfman (1994)에 의하면 미국에서 대학 교수들에 대한 강의평가의 출발이 단순 교육적 목적으로 비롯된 것이 아니라 학생 등록의 감소, 학부수업 질에 대한 불만, 교수들의 승진심사를 위한 평가자료 등과 같이 교육외적인 요소로 비롯되었다고 하였다. 많은 연구에서 대학 강의에 대한 학생평가를 네가지 목적으로 요약하고 있다. 첫째, 교수자 스스로 자신의 수업의 장단점을 파악하여 단점을 개선함으로써 수업의 효율성을 향상시키고, 둘째, 교수의 승진 심사 등 각종 인사 관련 결정을 위한 자료로 사용되며, 셋째, 학생들이 수강과목을 선택하는데 참고자료로 제공하며, 마지막으로 교수학습과 관련된 연구와 교육과정 및 평가 전문가들에게 필요한 자료를 제공함에 있다(Marsh & Boche, 1993; Cook, 1989; Seldin, 1993; Smith & Cranton, 1992).

우리나라의 경우 강의평가의 제도적 측면에 대한 기초연구나 결과활용에 대한 실증적 근거가 부족한 상황에서 강의평가 자체만 반복 시행되고 있는 실정이다. 강의평가도구 또는 결과의 타당성에 대한 연구는 일관된 양상을 보이지 않고 있는 상황이므로 단정적인 결론을 내리기 어렵다(이종승, 1993). 지은림과 장정아(2006)는 대학의 수업평가에 관한 선행 연구들의 검토를 통하여 강의평가가 타당성이 있다고 지적하였으나, 이에 관한 실증적인 연구는 매우 부족하다(길양숙·김동진, 2003). 이렇게 기존 연구를 검토해 볼 때 현재 시행되고 있는 강의평가결과의 타당성과 신뢰성, 그

리고 효과성 등에 관한 연구는 지속적으로 병행되어야 강의평가제도의 근본취지를 달성할 수 있으며 대학교육제도 안에서 긍정적 장치로 자리매김 할 수 있을 것이다.

### Ⅲ. 연구 방법

평가결과에 대하여 실질적으로 활용하려면 실제 평가자료에 대한 심층적 분석이 요구된다. 이 연구에서 평가도구의 타당도와 신뢰도, 평가결과에 대한 문항반응유형 등을 통계적으로 분석하기 위하여 실제 강의평가 결과를 사용하였다.

#### 1. 연구 대상

이 연구를 위해 2005학년도 1학기 강의평가 결과를 분석하였고 해당 학부와 과목, 그리고 교수와 학생자료 구성은 <표 1>과 같다. 분석에 실제 사용된 자료는 국내 I 대학교 공과대학과 자연과대학에서 개설된 학부과목 총 471개 과목, 385명의 강사에 대해 49,127명 학생들이 평가한 결과이다.

<표 1> 이공계 대학 2005학년도 1학기 강의평가 결과 자료

단과대학	학부(과)	과목 수	교수 수	학생 수
공과 대학	기계공학부	74	64	8,141
	생명화학공학부	34	23	2,962
	나노시스템공학부	24	13	1,177
	전자전기공학부	43	37	4,962
	컴퓨터공학부	36	37	5,265
	정보통신공학부	19	14	1,722
	신소재공학부	25	17	1,767
	환경토목공학부	59	35	3,068
	건축학부	36	42	2,697
전체	350	282	31,761	
자연 과학 대학	수학통계학부	44	40	6,132
	물리화학부	39	42	9,009
	생명해양과학부	38	21	2,225
	전체	121	103	17,366
합계		471	385	49,127

#### 2. 강의평가 도구

I 대학의 경우 최근 사용되고 있는 강의평가 유형은 강의특성, 교수방법, 그리고 수업환경에 따라 A형에서 K형까지 12 개의 유형으로 분류되어 있다(표 2). 객관식 문항은 21개 문항으로 유형에 따라 몇 가지씩 차이가 있으며 주관식 문항은 유형에 상관없이 제시된다. 객관식 문항은 교수방

<표 2> I 대학교 강의평가 유형별 특성

유형	특성
A형	인문사회 일반강좌, 중단위 이하의 수업 규모
B형	이공계 일반 강좌, 이론 중심의 수업
C형	토론 중심 수업
D형	이공계 전공 강좌, 소단위 이하의 전공 수업
E형	실습 중심의 강좌, 소단위 수업
F형	중단위 이하 강좌 중 강의중심의 수업, 교양선택 강좌
G형	중단위 이상의 교양선택 강좌, 강의중심 수업
H형	강의, 조별활동과 같이 다양한 수업진행 병행 강좌
I형	실습중심 강좌의 실습시간을 LA나 실습현장의 책임자가 담당
J형	외국인 강사가 담당하는 “대학영어” 과목
K형	E-Learning 강좌
L형	원어강의

<표 3> 이공계 일반강좌에 대한 강의평가도구(B형)

평가영역	평가 문항
I. 교수방법에 관한 평가	1. 강의노트(유인물 등)가 학습에 도움을 준 정도는
	2. 강의내용 이해를 위한 교수님의 설명 방법은
	3. 이론전개에 있어서 개념들 간의 연관성 제시는
	4. 학습을 위한 교수님의 예시 제공은
	5. 강의시간 중 질문할 기회는
	6. 질문에 대한 교수님의 답변은
	7. 시험 및 과제 제출 시 교수님의 결과처리는
	8. 교수님의 강의에 대한 열의는
	9. 학생들의 학습곤란에 대처하는 교수님의 방법은
	10. 교수님의 강의시간외의 질문을 위한 면담기회의 제공은
	11. 강의시간의 효율적인 사용은
	12. 학생학습에 대한 교수님의 관심은
II. 학습성취에 관한 평가	1. 본 강좌를 통해 배운 지식의 성취도는
	2. 학습내용의 난이도는
	3. 본 강좌의 평가방법은
	4. 보고서를 포함한 과제물이 학습에 도움을 준 정도는
	5. 시험문제(중간, 퀴즈 등)와 강의 내용과의 연관성
III. 강의 전반에 관한 평가	1. 본 강좌는 전반적으로
	2. 본 강좌에서 다루는 내용은
	3. 강의교재 선택의 적절성은
	4. 교수님의 강의 방법은

법에 대한 평가 12개 문항, 학습성취 관련 평가 5개 문항, 강의 전반에 대한 평가 4개 문항으로 구성되어 있으며 주관식 문항은 5개 문항으로 수업전체에 대한 의견과 이유를 직접 기술하도록 되어 있다. 강의평가 도구 유형을 살펴보면 A형은 인문사회 일반강좌, B형은 이공계 일반강좌, C형은

토론중심 수업, D형은 이공계 전공강좌, E형은 실습중심 수업, F형은 중단위 이하 강좌중 강의중심 수업, G형은 중단위 이상 강좌중 강의중심 수업, H형은 조별 활동 활용수업, I형은 실험실습 중심 수업(실습시간시 조교등 담당자가 다른 경우), J형은 영어로 진행되는 영어강의 그리고 K형은 온라인으로 진행되는 이러닝 수업 등으로 분류되며, 교수자가 수업특성에 따라 선택하여 평가받도록 되어있다.

각 문항에 대한 척도는 5단계 Likert 척도에 따라 ‘매우 긍정적’, ‘긍정적’, ‘보통’, ‘부정적’, ‘매우 부정적’ 등으로 동일하게 구성되어 있고, 해당강좌에 없거나 부적절한 문항은 ‘해당없음’란에 표시하도록 하였다. 이공계 대학에서 주로 사용하고 있는 B형의 강의평가 도구의 예를 보면 <표 3>과 같다.

### 3. 평가점수 산출 방법

강의평가 유형은 강의규모(수강인원), 전공영역(단과대학·학부), 강의유형(이론/실습), 과목분류(전공/교양), 수업방법(강의/토론 등), 강사직위(교수/강사 등) 등의 특성에 따라 선택적으로 사용할 것을 권하고 있고, 실제 학생들의 평가 결과를 보면 강사에 의해 결정되는 수업내용의 질을 판단하기 이전에 강의유형 자체에서 이미 평가점수의 차이가 크게 나타나고 있다. 이러한 강의평가의 유형에 따른 점수 차이를 보완해 주기 위해 표준점수를 산출하여 강사에 대한 평가를 실시하고 있으며 표준점수 산출방법은 다음과 같다.

$$\left( \frac{\sum_{\text{강좌}} \left( \frac{\text{점수} - \text{해당집단 I 평균}}{\text{해당집단 I 표준편차}} \times 0.4 + \frac{\text{점수} - \text{해당집단 II 평균}}{\text{해당집단 II 표준편차}} \times 0.6 \right)}{\text{강좌수}} \right) \times 5 + 70 = \text{표준점수}$$

여기서 집단 I 은 수강인원에 의한 분류로 20명 이하, 21~40명, 41~80명, 81~120명, 121명 이상의 5집단으로 구분되고, 집단 II는 단과대학에 의한 분류로 과목이 개설된 단과대학의 집단 특성을 반영하게 된다. 즉 수강인원과 단과대학에 따른 강의평가 결과를 반영하여 점수를 조정해주는 것이다.

### 4. 자료 분석 방법

이 연구에서 분석하고자 하는 범위는 첫째, 사용되고 있는 강의평가 결과의 신뢰도와 타당도를 점검하고, 둘째, 평가도구의 유형별 특성과 평가결과를 분석함으로써 강의유형과의 관계를 점검하고, 셋째, 강의평가점수에 따른 집단별 강의에 대한 문항반응유형을 분류하여 특성을 구분하고자 한다. 측정영역에서 검사의 타당도란 측정과 검사의 충실성과 적합성을 의미하여, 신뢰도는 측정과 검사의 정확성과 일관성을 말한다(APA, AERA, NCME, 1985, 1999). 검사결과가 바르게 사용되려면 신뢰도와 타당도, 즉 측정과 검사의 양호도를 보장하여야 한다. APA, AERA, NCME (1985, 1999)에서 검사도구의 타당도의 종류를 내용타당도, 준거관련타당도, 구인타당도로 분류하면서 각각의 개념과 추정방법을 정리하였다. 강의평가도구의 타당도는 검사지가 측정하려고 하는 행동특성, 즉 구인(construct)을 얼마나 의미있게 반영하고 있는지 검증하는 것이 구인타당도이다. 즉 도구를 제작하는 과정에서 가장 중요하게 다루는 ‘구인’이 어느 정도 이론적 구조에 부합되는가를 점

검하고자 하였다. I 대학의 강의평가도구에서 제시하고 있는 세 가지 영역에 대하여 각 문항이 어떻게 설명되고 있는지 살펴보기 위해 확인적 요인분석(Confirmatory Factor Analysis) 방법을 적용하였다.

신뢰도 추정은 충분한 사례수가 있으므로 각 영역별 그리고 전체평가도구에 대하여 Cronbach  $\alpha$  (성태제, 2002) 를 산출하였다. Cronbach  $\alpha$  는 내적일관성 신뢰도의 하나로 피험자가 검사에 답하는 일관성정도를 파악하기 위해 검사-재검사법 신뢰도나 동형검사 신뢰도와 달리 검사를 나누지 않고, 한 검사 내에서 피험자 반응에 대한 문항 간 유사성을 계산하는 것이다. Cronbach  $\alpha$  는 이 분점수와 연속점수 모두 적용할 수 있고 통계패키지프로그램에 대부분 포함되어 있으므로 신뢰도 계수 산출시 대표적으로 적용되고 있다 .

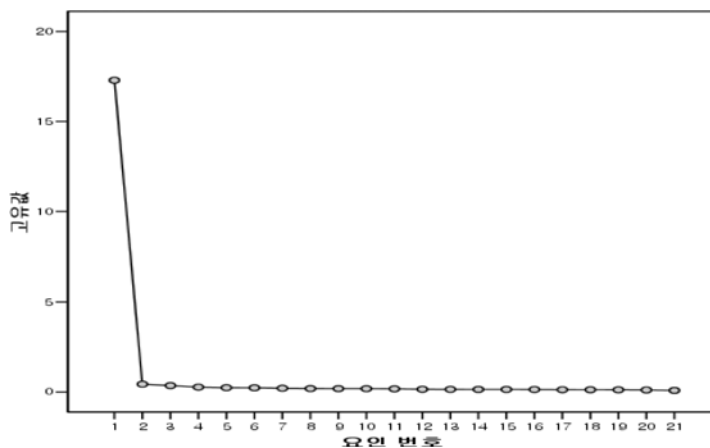
또한, 강의유형별 결과에 대한 분석과 답지반응 유형분석 등은 단순한 기술통계를 계산하여 비교함으로써 실제 다른 유형의 평가도구와 학생반응에 대한 시사점을 도출하고자 하였다. 그리고 평가점수가 상위 10% 와 하위 10% 에 해당하는 강의의 특성을 분석하여 학생과의 반응과 비교하였다. 분석에 사용한 프로그램은 SPSS/PC 12.0 버전이다.

## IV. 연구 결과

### 1. 평가도구의 양호도 검증 결과

강의평가를 위한 설문은 유형에 따라 다소 문항진술이 틀리지만 위에서 제사한 바와 같이 교수 방법에 관한 평가 12문항, 학업성취에 관한 평가 5문항, 강의전반에 관한 평가 4문항으로, 총 3개 영역 21개의 문항으로 구성되어 있다. 각 문항의 응답은 5단계 Likert 척도로 주어지며, 강의평가 설문의 타당도를 검증하기 위하여 요인분석을 실시하였고, 21개의 문항으로 이루어진 강의평가 내용이 교수방법, 학업성취, 강의전반의 3개의 요인으로 구분되는지 살펴보았다.

[그림 1]의 Scree 도표에서 확인할 수 있는 바와 같이, 아이겐 값이 1보다 크게 나타난 것을 기



[그림 1] Scree 도표



<표 4> 요인분석을 통한 구인타당도 검증 결과

내용영역	문항번호	1요인
교수방법	1	.881
	2	.910
	3	.919
	4	.916
	5	.880
	6	.910
	7	.898
	8	.889
	9	.920
	10	.891
	11	.909
	12	.917
학습성취	13	.925
	14	.889
	15	.919
	16	.901
	17	.901
강의전반	18	.944
	19	.927
	20	.897
	21	.935
고유값		17.335
설명된 총분산		82.548

<표 5> 신뢰도 검증 결과 (Cronbach  $\alpha$ )

하위 요인	문항 수	Cronbach $\alpha$
교수방법	12	.981
학습성취	5	.963
강의전반	4	.966
전체	21	.989

준으로 할 때 21개 문항은 하나의 요인으로 설명된다는 것이다. 따라서 강의평가 설문내용을 현재와 같이 교수방법, 학습성취, 강의전반의 3개 요인으로 구분하는 것은 타당하지 않다고 볼 수 있다. 요인분석을 통하여 이공계대학에서 가장 많이 사용되었던 B형 강의평가지의 구인타당도를 검증한 결과는 <표 4>와 같다. 요인분석 결과, 하나의 요인만 추출되었고, 요인1의 고유값은 17.335, 설명된 총분산은 82.548로 매우 높게 나타났다. 또한, 강의평가의 신뢰도를 검증하기 위하여, 문항내적일관성 신뢰도인 Cronbach  $\alpha$  값을 산출한 결과는 <표 5>와 같고, 신뢰도 검증결과 Cronbach  $\alpha$  값은 하위요인과 전체 모두 각각 .963~.989로 매우 높게 나타났다. 평가결과에 대한 신뢰도가 이렇게 높게 산출된 것은 학생들의 문항에 대한 반응이 일관적이라는 의미이다.

## 2. 강의평가의 유형별 분석 결과

강의평가 설문은 개설되는 강의의 특성에 따라 총 12개의 유형으로 제작되었으며, 강사가 그 가운데 자신의 강의에 가장 적합한 강의평가 유형을 선택하여 학생들에게 제공하도록 되어 있다. 이공계대학의 경우, 전공영역과 과목분류상 B형(이공계 일반 강좌)이 가장 많으며, 이론과목의 경우는 D형(이공계 전공 강좌)이, 실습과목인 경우는 I형(조교가 실습을 담당하는 강좌)이 그 다음으로 많이 사용된다.

각 학부별로 2005학년도 1학기에 강의평가 유형에 따라 응답한 학생 수와 %는 <표 6>과 같다. 학생들이 강의평가에 참여한 결과를 살펴보면, 이공계 일반 강좌에 해당하는 B형에 응답한 학생이 가장 많았고, I형과 D형의 순으로 나타났다.

<표 6> 강의평가지 유형별 응답 학생 수 (%)

단과대학	학부(과)	유형								전체
		A	B	D	E	F	G	H	I	
공과대학	기계공학부	1274 (15.6)	<b>3760</b> <b>(46.2)</b>	1138 (14.0)	868 (10.7)	-	124 (1.5)	197 (2.4)	780 (9.6)	8141 (100.0)
	생명화학공학부	269 (9.1)	<b>2074</b> <b>(70.0)</b>	117 (4.0)	55 (1.9)	233 (7.9)	-	69 (2.3)	145 (4.9)	2962 (100.0)
	나노시스템공학부	-	<b>943</b> <b>(80.1)</b>	173 (14.7)	-	-	-	-	61 (5.2)	1177 (100.0)
	전자전기공학부	393 (7.9)	<b>3008</b> <b>(60.6)</b>	365 (7.4)	201 (4.1)	-	84 (1.7)	-	911 (18.4)	4962 (100.0)
	컴퓨터공학부	239 (4.5)	<b>2674</b> <b>(50.8)</b>	18 (.3)	403 (7.7)	-	-	-	1931 (36.7)	5265 (100.0)
	정보통신공학부	227 (13.2)	<b>711</b> <b>(41.3)</b>	461 (26.8)	183 (10.6)	-	-	-	140 (8.1)	1722 (100.0)
	신소재공학부	-	<b>1390</b> <b>(78.7)</b>	169 (9.6)	16 (.9)	-	-	-	192 (10.9)	1767 (100.0)
	환경토목공학부	68 (2.2)	<b>1361</b> <b>(44.4)</b>	1294 (42.2)	126 (4.1)	-	-	-	219 (7.1)	3068 (100.0)
	건축학부	-	<b>1937</b> <b>(71.8)</b>	143 (5.3)	528 (19.6)	-	89 (3.3)	-	-	2697 (100.0)
	전체	2470 (7.8)	<b>17858</b> <b>(56.2)</b>	3878 (12.2)	2380 (7.5)	233 (.7)	297 (.9)	266 (.8)	4379 (13.8)	31761 (100.0)
자연과학대학	수학통계학부	554 (9.0)	<b>1775</b> <b>(28.9)</b>	343 (5.6)	313 (5.1)	2918 (47.6)	-	-	229 (3.7)	6132 (100.0)
	물리화학부	1084 (12.0)	<b>3311</b> <b>(36.8)</b>	952 (10.6)	35 (.4)	-	-	-	3627 (40.3)	9009 (100.0)
	생명해양과학부	560 (25.2)	<b>397</b> <b>(17.8)</b>	670 (30.1)	89 (4.0)	-	171 (7.7)	-	338 (15.2)	2225 (100.0)
	전체	2198 (12.7)	<b>5483</b> <b>(31.6)</b>	1965 (11.3)	437 (2.5)	2918 (16.8)	171 (1.0)	-	4194 (24.2)	17366 (100.0)

강의평가 유형별로 각 문항과 총점의 평균은 <표 7>과 같다. 이공계 대학의 강의평가 유형별로 살펴보면, D형(이공계 전공 강좌 - 소단위 이하의 전공)과 H형(강의, 조별활동 등 다양한 수업방식 병행 강좌)의 평균이 다른 유형에 비해 높게 나타난 반면, A형(인문사회 일반 강좌)과 I형(조교가 실습을 담당하는 강좌)의 평균은 다른 유형보다 낮게 나타났다. 즉, 학생들은 교양보다는 전공과 관련된 과목에, 대규모보다는 소규모의 수업에, 조교가 담당하는 수업보다는 교수나 강사가 직접 관리하는 수업에 보다 높은 점수를 부여하고 있음을 알 수 있다.

하지만 이 표준점수로는 수업규모와 단과대학의 특성 외에는 강의평가 유형에 따른 점수 차이를 제대로 보정해줄 수 없다. 예를 들어 교양과목에 비해 전공과목에 대한 강의평가 점수가 높게 나타나지만 현 표준점수 산출방법으로는 이 차이를 반영하여 평가 결과를 조정해줄 수가 없다. 또한 수업규모와 단과대학에 따라 가중치를 달리 주고 있으나 실제 경험적인 자료 분석에 의한 결과가 아니므로, 강의평가 결과를 유형에 상관없이 비교하거나 교수를 평가하기 위한 행정적인 목적으로 사용하기 위해서는 최근의 강의평가 결과를 기초로 정교한 검토가 필요하다. 따라서 강의평가 유형 선택 자체가 강의평가 결과에 미치는 영향을 고려하여, 강의평가 유형 구분에 대한 타당성을 입증할 수 있는 지속적인 연구가 필요하며, 강의평가 유형별로 문항의 내용과 수준을 점검하고 비교

<표 7> 강의평가 유형별 문항과 총점 평균

문항번호	유형								
	A	B	D	E	F	G	H	I	합계
1	3.90	3.99	<b>4.13</b>	3.94	3.88	3.98	4.13	3.89	3.97
2	3.88	3.99	<b>4.13</b>	4.03	3.96	4.06	4.12	4.01	4.00
3	3.88	4.00	<b>4.13</b>	3.96	3.88	3.94	4.14	3.90	3.98
4	3.89	4.01	<b>4.12</b>	3.94	3.96	3.96	4.11	3.83	3.97
5	3.84	3.95	<b>4.12</b>	3.96	3.96	4.03	4.04	3.91	3.96
6	3.87	4.00	<b>4.14</b>	3.92	3.94	4.04	4.09	3.93	3.98
7	3.93	3.94	<b>4.09</b>	3.89	3.79	3.99	4.09	3.94	3.95
8	3.87	4.07	<b>4.21</b>	3.90	3.87	4.09	4.19	3.90	4.02
9	3.79	3.93	<b>4.10</b>	3.80	4.05	3.99	4.11	3.77	3.91
10	3.82	3.89	<b>4.04</b>	3.90	3.86	3.95	4.08	3.99	3.92
11	3.83	3.97	<b>4.18</b>	3.96	3.99	4.10	4.11	3.88	3.97
12	3.78	3.98	<b>4.15</b>	3.94	3.92	3.99	4.14	3.93	3.97
13	3.83	3.96	<b>4.12</b>	3.94	3.91	3.97	4.15	3.91	3.96
14	3.83	3.87	<b>4.02</b>	3.98	3.94	3.99	4.13	3.89	3.90
15	3.84	3.93	<b>4.07</b>	3.91	3.89	3.98	4.07	3.85	3.92
16	3.84	3.93	<b>4.07</b>	3.92	3.85	3.94	4.09	3.90	3.93
17	3.95	3.98	<b>4.11</b>	3.94	3.95	4.03	4.09	3.88	3.97
18	3.87	4.00	<b>4.15</b>	3.95	3.97	4.01	4.15	3.93	3.99
19	3.84	4.00	<b>4.13</b>	3.97	3.95	4.02	4.10	3.92	3.98
20	3.69	3.95	<b>4.08</b>	3.88	3.89	3.88	4.09	3.84	3.91
21	3.74	3.98	<b>4.14</b>	3.92	3.96	4.00	4.10	3.88	3.96
총점	80.71	83.33	<b>86.45</b>	82.55	82.37	83.95	86.33	81.88	83.17

해 보아야 한다. 또한 강의평가 유형에 따른 결과 차이를 최대한 보정할 수 있는 점수 산출방법에 대한 수정이 필요하고, 유형이 다른 강의평가 결과의 활용에 있어서 주의를 기울여야 할 것이다.

### 3. 강의평가의 응답 반응유형 분석결과

강의평가 결과를 교수-학습 개선과 강사에 대한 평가 자료로 활용할 때, 가장 문제가 되는 것은 강의평가 결과를 얼마나 신뢰할 수 있는가하는 점이다. 일반적으로 강의평가의 참여율은 '성적조회 제한'을 하기 때문에 90% 이상으로 높게 나타나지만, 강제적인 참여로 인해 학생들의 불성실한 응답 또한 높기 때문이다. 강의평가결과 총점 평균점수가 상위 10%에 해당하는 강의와 하위 10%에 해당하는 강의에서 불성실하게 응답한 학생 수와 비율은 <표 8>, <표 9>와 같다. 학생들의 불성실한 응답에 대한 조작적 정의는 동일한 척도로 21개의 문항 모두에 응답한 경우를 의미한다.

총점평균점수 상위 10%와 하위 10%에 해당하는 강의에서 불성실하게 응답한 학생의 비율은 각각 20.00 ~ 100.00%, 9.76 ~ 90.48%로 높게 나타났다. 불성실하게 응답한 학생의 비율이 이렇게 높게 나타남에도 불구하고 강의평가 결과 신뢰도가 높게 나타나는 이유는 학생들이 강의 전체에 대한 평가를 반영하여 한 번호로 답하기 때문이다. 즉, 해당 강의에 매우 만족한 학생은 강의평가 문항에 상관없이 1번부터 21번까지 모두 '매우 긍정적'에 해당하는 5번으로 답하고, 강의에 대해 매우 불만이 많았던 학생은 21개의 문항 모두를 '매우 부정적'에 해당하는 1번으로 답하는 경향이 있었다. 상위 10%에 해당하는 강의에서는 4번과 5번으로 모두 응답한 학생이 많았고, 하위 10%에 해당하는 강의에서는 1번과 2번으로 모든 문항에 대해 응답한 학생이 많았다. 또한 강의평가 결과 상·하위와 상관없이 3번으로 모든 문항에 응답한 학생도 많았다. <표 8>에서 볼 수 있듯이 총점평균 상위 10%에 해당하는 98개 강의 중 불성실하게 응답한 학생이 70%가 넘는 강의가 무려 39개 (약 40%) 나 된다는 것은 강의 평가 결과를 어떻게 해석해야하는가 하는 문제를 제기하고 있다. 도리어 총점평균 하위 10%에 해당하는 98개 강의는 평균점수는 낮다하더라도 불성실하게 응답한 학생이 70% 넘는 강의가 7 개(약 7.2%) 정도로 나타났다 (<표 9>참조). 강의평가 결과를 활용하기에 앞서 학생들이 어떻게 강의평가에 임하고 있는가를 파악해야한다는 것이다.

### 4. 강의평가 결과 상·하위 10%에 해당하는 강의 분석

강의평가는 궁극적으로 평가 결과를 반영하여 강의의 질을 개선하고 이후의 교수-학습에 도움을 주고자 하는 것을 목적으로 한다. 이를 위해서 특히 높은 평가를 받은 강의와 낮은 평가를 받은 강의에 주목해 볼 필요가 있다. 높은 강의평가를 받은 강의는 벤치마킹하여 어떠한 점에서 학생들이 만족하고 있는지, 특별한 강사의 노하우가 있는지, 다른 강의에 적용할 수 있는 교수방법상 특징이 있는지 등 그 강의에 한정해서 뿐만 아니라 다른 강의의 질을 개선하는 목적으로도 확대하여 활용할 수 있다. 또한 낮은 강의평가를 받은 강의는 무엇이 가장 문제인지, 강사에게 부족한 교수적 능력은 무엇인지, 해당 강의를 통해서 학생들이 가장 어려움을 느끼는 부분은 무엇인지를 파악하여 강사에게 도움을 줄 수 있는 프로그램을 제공하거나, 이와 유사한 강의에서 발생할 수 있는 문제를 사전에 파악하여 대처하는 데 활용할 수 있다. 강의평가 결과 상·하위 10%에 해당하는 강의의 특성을 요약하면 <표 10>과 같다.

강의평가 결과, 공과대학이 자연과학대학보다 상위 10%에 해당하는 강의의 비율이 약간 높고

<표 8> 상위 10% 강의에서 불성실하게 응답한 학생 비율

과목 이수번호	강의평가 유형	전체 학생수	불성실하게 응답한 학생 수		과목 이수번호	강의평가 유형	전체 학생수	불성실하게 응답한 학생 수	
			명	%				명	%
IC305-01	4	10	2	20.00	IA211-03	2	62	42	67.74
CL104-02	4	96	28	29.17	IT121-01	2	56	38	67.86
CK001-04	4	100	36	36.00	IX210-01	4	25	17	68.00
CL104-01	4	92	34	36.96	BS357-02	2	57	39	68.42
CL104-03	4	88	35	39.77	BS255-01	4	51	35	68.63
BS213-08	6	64	29	45.31	IA311-01	1	96	66	68.75
YN137-02	6	49	23	46.94	BT453-01	4	29	20	68.97
BS213-22	6	80	40	50.00	IR215-01	2	71	49	69.01
BS213-24	1	174	93	53.45	IN318-02	2	69	48	69.57
IW212-01	5	31	17	54.84	IB212-01	4	79	55	69.62
IA411-03	7	124	68	54.84	IV178-01	4	10	7	70.00
IW201-01	5	9	5	55.56	BS203-02	2	50	35	70.00
IL309-01	4	36	20	55.56	IL310-02	4	67	47	70.15
BS213-19	6	61	34	55.74	IB304-02	2	87	62	71.26
BP334-01	4	16	9	56.25	IA213-02	2	115	82	71.30
IN322-01	2	23	13	56.52	IM471-01	4	7	5	71.43
BO343-01	4	28	16	57.14	IV162-03	5	14	10	71.43
IN423-01	2	81	47	58.02	IJ107-01	2	21	15	71.43
BS203-01	2	65	38	58.46	IP411-01	2	35	25	71.43
BP331-01	4	17	10	58.82	IL324-01	5	21	15	71.43
BO241-01	4	22	13	59.09	IU213-01	2	53	38	71.70
BS213-09	6	66	39	59.09	IL203-02	2	78	56	71.79
BS255-02	4	42	25	59.52	IU203-02	2	32	23	71.88
BS213-10	6	52	31	59.62	IB218-03	2	94	69	73.40
BS213-18	6	63	38	60.32	IE352-01	2	65	48	73.85
IA224-03	2	65	40	61.54	IR312-01	2	54	40	74.07
IL409-03	1	34	21	61.76	IU307-01	2	24	18	75.00
IJ302-01	4	21	13	61.90	IU306-01	4	52	39	75.00
IL310-01	4	29	18	62.07	IU203-01	2	66	50	75.76
IN318-01	2	90	56	62.22	IU214-02	4	50	38	76.00
IX201-01	4	16	10	62.50	IB406-02	9	47	36	76.60
IA216-02	2	96	60	62.50	BS323-02	2	52	40	76.92
CK503-02	4	30	19	63.33	IV183-01	4	9	7	77.78
IF606-01	2	60	38	63.33	IU214-01	4	42	33	78.57
IL207-03	9	14	9	64.29	BS341-01	4	38	30	78.95
BS357-01	2	40	26	65.00	IJ401-01	4	38	30	78.95
IE343-01	4	43	28	65.12	IU213-02	2	67	53	79.10
IA224-08	2	66	43	65.15	IR315-01	2	29	23	79.31
IB314-01	8	26	17	65.38	IE345-01	1	25	20	80.00
IL203-01	2	87	57	65.52	BS371-01	2	36	30	83.33
IJ304-01	2	12	8	66.67	IU215-02	5	21	18	85.71
IB316-05	5	27	18	66.67	IB218-04	2	53	46	86.79
IN408-02	2	36	24	66.67	IU215-01	5	39	34	87.18
IR413-01	4	15	10	66.67	EP226-01	2	8	7	87.50
IA224-04	2	69	46	66.67	IE243-01	2	22	20	90.91
IJ404-01	2	27	18	66.67	IM459-01	4	23	21	91.30
BS341-02	4	61	41	67.21	IV162-01	5	18	17	94.44
IN408-01	2	83	56	67.47	IU308-01	2	6	6	100.00
IA216-05	2	80	54	67.50	IQ414-01	4	8	8	100.00

<표 9> 하위 10% 강의에서 불성실하게 응답한 학생 비율

과목 이수번호	강의평가 유형	전체 학생수	불성실하게 응답한 학생 수		과목 이수번호	강의평가 유형	전체 학생수	불성실하게 응답한 학생 수	
			명	%				명	%
BM100-01	2	41	4	9.76	IL202-02	1	53	26	49.06
BM101-11	2	67	15	22.39	IN103-02	2	79	39	49.37
BM101-05	2	57	14	24.56	BS101-36	1	295	147	49.83
BM101-30	2	66	17	25.76	IN303-01	2	60	30	50.00
IO101-02	5	35	11	31.43	IN103-04	9	82	41	50.00
BM101-22	2	66	21	31.82	IN303-02	2	60	30	50.00
IX179-01	5	12	4	33.33	YN137-06	6	47	24	51.06
BM101-25	2	62	21	33.87	IA211-05	1	263	137	52.09
IN105-01	2	41	14	34.15	IA111-01	5	42	22	52.38
IN103-03	2	89	31	34.83	IA226-04	5	49	26	53.06
BS201-01	1	20	7	35.00	IN201-02	9	26	14	53.85
BS101-34	6	50	18	36.00	IT120-06	5	13	7	53.85
IN101-06	9	58	21	36.21	IA226-02	5	52	28	53.85
BM101-08	2	57	22	38.60	IA226-01	5	50	27	54.00
BM101-20	2	62	24	38.71	YC008-02	5	80	44	55.00
BN105-04	2	74	29	39.19	IT119-06	4	27	15	55.56
BM101-10	2	76	30	39.47	IK103-01	2	97	54	55.67
IA226-03	5	43	17	39.53	IL332-01	2	57	32	56.14
IO101-01	5	35	14	40.00	IN101-05	9	53	30	56.60
BN105-05	2	64	26	40.63	IO207-07	9	26	15	57.69
BN105-01	2	68	28	41.18	YC007-05	9	78	46	58.97
BN105-07	2	63	26	41.27	BN211-01	4	69	41	59.42
BR101-01	2	41	17	41.46	IX191-01	4	13	8	61.54
YC008-04	9	76	32	42.11	OE076-01	1	108	67	62.04
BN105-03	2	73	31	42.47	IO203-01	1	87	54	62.07
IA111-05	5	40	17	42.50	IQ311-01	2	55	35	63.64
BN105-06	2	63	27	42.86	BN101-04	2	69	44	63.77
YC008-03	5	86	37	43.02	IL202-03	2	92	59	64.13
BN105-02	2	69	30	43.48	YN137-04	6	51	33	64.71
YC006-01	9	79	35	44.30	IK103-02	2	80	52	65.00
IN105-02	2	88	39	44.32	IV164-02	5	20	13	65.00
BS101-21	6	48	22	45.83	BS485-01	2	42	28	66.67
BN101-09	2	65	30	46.15	BS421-01	2	9	6	66.67
IN303-03	2	56	26	46.43	IK405-01	2	31	21	67.74
EP212-01	2	15	7	46.67	IL332-02	2	60	42	70.00
YC007-04	9	85	40	47.06	IN419-01	4	18	13	72.22
IN101-03	9	57	27	47.37	IF408-01	5	22	16	72.73
IN101-04	9	86	41	47.67	BS323-01	2	40	30	75.00
BN101-12	2	69	33	47.83	BS451-01	2	19	15	78.95
BS101-11	6	48	23	47.92	IV164-01	5	21	19	90.48
IN305-01	2	48	23	47.92	IV164-01	5	21	19	90.48
IO203-02	2	85	41	48.24					

<표 10> 강의평가 결과 상·하위 10%에 해당하는 강의의 특성

구분		상위 10%		하위 10%		전체	
		명	%	명	%	명	%
단과대학	공과대학	3,520	11.1	2,829	8.9	31,761	100.0
	자연과학대학	1,277	7.4	2,170	12.5	13,919	100.0
강사 직위	교수	3,299	9.6	1,914	5.6	34,272	100.0
	강사	1,083	7.8	3,085	22.3	13,853	100.0
	조교	485	10.1	0	0.0	1,002	100.0
강의평가 유형	A형	407	8.7	826	17.7	4,668	100.0
	B형	2,289	9.8	2,475	10.6	23,341	100.0
	D형	1,290	22.1	127	2.2	5,843	100.0
	E형	165	5.9	621	22.0	2,817	100.0
	F형	435	13.8	244	7.7	3,151	100.0
	G형	124	26.5	0	0.0	468	100.0
	H형	26	9.8	0	0.0	266	100.0
	I형	61	0.7	8.2	14.1	8,573	100.0

하위 10%에 해당하는 강의의 비율이 낮게 나타났으나 큰 차이는 아니라고 볼 수 있다. 반대로 강사 직위에 있어서는 하위 10%에 해당하는 강의에, 교수는 전체 교수의 5.6% 밖에 되지 않는 반면, 강사는 전체 강사의 22.3%나 속해, 학생들의 강사에 대한 불만이 상대적으로 크다는 것을 알 수 있다. 또한 강의평가 유형별로 볼 때도 상위 10%에서는 D형과 G형의 강의평가 설문을 사용한 강의가, 하위 10%에서는 A형과 E형의 강의평가 설문을 사용한 강의의 비율이 상대적으로 높게 나타났다. 이러한 결과는 강사보다는 교수의 강의가 경험을 바탕으로 교수방법과 강의전반에 있어서 높은 평가를 받고 있다는 것을 시사하며, D형의 강의평가 점수 평균이 높고 A형의 강의평가 점수 평균이 낮은 결과와 유사하게 상·하위 10%에 해당하는 비율에 차이가 있는 점으로 비추어 볼 때 강의평가 유형에 따른 점수 차이를 현재 표준점수가 보정해 주고 있지 못함을 반영하고 있다.

## V. 결론 및 제언

대학의 강의평가가 수업의 질을 향상시키기 위함이라는 것은 모두 동의하는 바이다. 그러나 이미 강의평가의 활용범위가 확대되고 있는 시점에서 다시 강의평가제도의 문제점을 제기하며 강의평가 자체 개선을 위한 실증적 분석을 시도하였다. 연구결과를 요약하면 I 대학의 강의평가 설문내용을 3개 요인으로 구분하는 것은 통계적으로 타당하지 않았으며 21개 문항은 하나의 요인을 설명하고 있었다. 따라서 문항의 수를 조정하거나 재배열 또는 문항내용을 수정해야 할 필요가 있다. 신뢰도 경우 Cronbach  $\alpha$  값이 .963~.989로 높게 산출되었으나 학생들의 강의평가에 대한 태도와도 연결하여 해석해야 할 것이다.

강의평가 유형별 분석결과에서 지적하였듯이 소단위 이하의 전공강의와 조별활동 등을 포함한 다양한 수업방식을 병행한 과목에 대한 평가결과가 다른 유형에 비해 높게 나타났으며 인문사회

일반강의와 조교가 담당하는 실습강의에 대한 평균은 다른 유형보다 낮게 나타났다. 이것은 학생들이 교양보다는 전공관련 과목에, 대규모보다는 소규모 수업에, 그리고 조교가 담당하는 수업보다는 교수가 직접 관리하는 실습강의를 더 좋게 평가하고 있음을 보여줌으로써 최근 연구(Spooren & Mortelmans, 2006; Algozzine, et al., 2004; Grimes, Millea, & Woodruff, 2004; Shevlin, et. al., 2000)를 실증적으로 뒷받침한다. 또한 사용하고 있는 표준점수는 수업규모와 단과대학 특성 외 강의평가유형에 따른 점수 차이를 조정할 수 없다는 문제점도 지적하였다. 따라서 강의평가 결과를 유형에 상관없이 비교하거나 교수를 평가하기 위한 행정적인 목적으로 사용하기 위해서는 최근의 강의평가 결과를 기초로 정교한 검토키트가 필요하며 점수산출 공식을 어떻게 수정해야 하는지 연구해야 할 것이다.

연구 결과 가장 큰 문제점으로 나타난 것은 학생들이 강의평가를 불성실하게 응답하는 경향이 있다는 점이다. 총점평균점수 상위 10%와 하위 10%에 해당하는 강의에서 불성실하게 응답한 학생의 비율은 각각 20.00 ~ 100.00%, 9.76 ~ 90.48%로 높게 나타났다. 이렇게 불성실하게 응답한 학생의 비율이 높게 나타남에도 불구하고 강의평가 결과 신뢰도가 높게 나타나는 이유는 학생들이 강의 전체에 대한 평가를 반영하여 동일한 번호로 답하기 때문인 것으로 검증되었다. 따라서 강의평가 결과를 강의와 강사에 대한 평가 자료로 활용할 수는 있으나, 그 활용에 있어서 주의를 기울여야 한다. 즉, 교수-학습 개선을 위해 결과를 활용할 경우는 점수 자체보다는 수업방법과 학생들이 어려움을 겪는 내용을 보완하는 것을 목적으로 하여야 하며, 강사에 대한 평가 자료로 활용할 경우에는 강의평가 결과를 참고하되 강의평가 결과만을 반영하여 행정적인 의사결정을 하는 것은 지양하여야 한다.

한편 강의평가 결과 높은 점수를 받은 강의와 낮은 점수를 받은 강의의 특성을 살펴본 결과, 강의평가 유형에 상관없이 상위 10%에 해당하는 강의에서 높은 점수를 받은 문항은 강의 전반에 대한 만족도와 강의 내용의 적절성, 교수의 열의에 대한 내용이었다. 또한 강의평가 유형별로 강의방법과 관련된 문항에서도 높은 점수를 얻고 있음을 확인할 수 있다. 이와는 달리 하위 10%에 해당하는 강의에서 상대적으로 더 낮은 점수를 받은 문항의 내용은 보다 구체적으로 학습방법과 학생 지도에 대한 내용이 많았으며, 이러한 부분이 전반적인 평가 결과에도 좋지 않은 영향을 주었다. 따라서, 이러한 구체적인 사례는 강사와 학생 모두에게 실질적인 도움을 줄 수 방법으로 강의평가 결과를 적극적으로 활용하여야 한다. 강의평가 결과 상·하위 10%에 해당하는 강의 내용을 질적으로 분석하여 장단점을 보다 체계적으로 구조화할 필요가 있으며, 그 결과를 다시 교수-학습 개선을 위한 자료로 강사들에게 제공해 주어야 한다. 특히, 상위 10%에 해당하는 강의는 우수 강의 사례로 강의평가 결과와 수업 내용을 공개하여 다른 강사와 학생들에게 제공해주는 방식으로 활용하고, 하위 10%에 해당하는 강의는 새로운 교수방법에 관련된 내용이나 문제로 지적된 점을 수정하는데 실질적 도움을 줄 수 있는 프로그램을 개발하는 방식으로 지원한다면 긍정적 효과를 얻을 수 있을 것이다.

이 연구 결과를 토대로 다음과 같은 결론을 제안하고자 한다. 첫째, 평가대상은 교수 개인에 대한 평가에 국한하지 말고 교수와 강의 그리고 학생으로 구분하여 평가하여야 한다. 즉 수업유형으로 평가도구를 단순 분류하는 것보다 교수 개인에 대한 평가, 강의자체에 대한 평가, 그리고 학생의 성취나 태도에 대한 평가 등으로 분류되어야 결과 활용에 있어 즉각적인 피드백이 이루어질 것이다.



둘째, 평가시기는 한학기 1회를 원칙으로 하되 실제로 강의 개선을 위한 피이드백을 수합하기 위하여 강의가 진행되는 학기 중에 한번 더 시행하는 것을 제안하고자 한다. 더욱 중요한 것은 우리나라 대부분 대학에서 성적열람 이전에 온라인 평가방법을 사용함으로써 발생하는 문제점을 시정해야 할 것이다. 즉 학생들의 무성의한 답변을 유도할 뿐 아니라 그 결과에 대한 신뢰도에 치명적인 영향을 미치고 있다. 따라서 외국의 경우처럼 마지막 수업 중 OCR 카드를 사용하는 것이 무리라면 온라인 평가를 유지하되 종강 이전 1-2 주전 강의평가에 대한 정확한 홍보와 가이드를 주며 수업의 연장선 상에서 학생들의 신중한 평가를 유도해야 한다.

셋째, 평가도구는 유형의 종류가 다양한 것보다 평가도구의 하부 요인에 포함하는 문항을 어떻게 구성할 것인가가 더욱 중요하다. 강의평가 문항 군에서 전공별 또는 수업유형별로 문항을 선택하도록 하는 것이 바람직하며 문항의 수도 15개 안팎이 적절하다고 본다. 특히 평가도구 유형은 다양한 것도 중요하지만 좀 더 타당한 문항으로 수정할 것을 제한한다.

넷째, 평가 결과를 강의의 질 자체를 개선하는데 활용하기보다 교수의 교수업적 평가나 강사의 강의부여 판정등 행정적 의사결정으로 확대되고 있다는 또 다른 문제점을 발견하였다. 따라서 강의평가 결과 활용이 본질에 충실할 수 있도록 학기 중 또는 다음 학기 시작 직후 실질적인 수업개선을 위해 강의평가 결과를 어떻게 반영하였는지 교수와 학생 간의 상호 피이드백이 필요하다. 이러한 강의평가결과에 대한 순기능적 절차는 교수와 학생간의 상호 신뢰를 높일 뿐 아니라 강의평가에 대한 학생의 태도를 결정적으로 개선시킬 수 있다.

다섯째, 평가 제도의 측정학적 문제를 해결하여야 한다. 이미 앞에서 지적한 문항내용의 타당성 문제를 포함하여 학생들의 성실한 참여를 기대함으로써 평가결과의 신뢰도를 보장하여야 한다. 강의평가 참여율은 매우 높으나 학생들의 불성실한 응답 역시 매우 높으므로 높은 신뢰도는 동일한 번호를 응답했기 때문에 수치상 높게 나타난 것이라 무의미하다. 따라서 평가문항의 개선과 결과 활용의 순기능을 보강함에 앞서 가장 중요한 것은 강의평가에 임하는 학생들의 태도를 변화시켜야 한다.

여섯째, 강의평가 결과를 교수 개인적 차원에서 수업개선을 위해 사용할 뿐 아니라 우수수업 모형을 개발함으로써 대학 차원에서 교육의 질 향상을 위한 교육방법을 개발하고 교육공학적 자료를 체계적으로 지원하도록 한다. 특히 상하위 10% 에 해당하는 강의에 대한 집중적 분석을 통해 높은 강의평가 결과를 받은 강의의 특성과 장점을 제시하고 모형화 함으로써 우수강의의 방법을 적용할 수 있는 프로그램을 개발할 수 있다. 한편 낮은 강의평가를 받은 경우의 문제점과 학생들의 반응도 유형화하여 실질적인 개선책을 제공할 수 있다.

마지막으로, 이공계 대학 뿐 아니라 대학 전체에서 강의평가제도가 보다 긍정적 역할로 자리매김하려면 대학 내 관련 전문기구를 활용하여 지속적인 연구가 진행되어야 한다. 학기별, 교수별 또는 강의별 평가결과점수를 제공하는 것으로 그치지 말고 각 대학 내 교수학습 관련 전문기구 또는 교육평가 전문가에게 강의평가 결과가 개인적 차원은 물론 대학전체 차원에서 활용 될 수 있는 연구지원이 반드시 우선되어야 한다. 앞에서 언급한 바와 같이 평가대상, 시기, 도구, 문항, 그리고 결과활용 등 강의평가제도의 전반적 개선이 요구되며 제안된 구체적 방안이 실천으로 옮겨질 시점인 것이다.

[ 감사 의 글 ]

본 연구는 2005년 학술진흥재단의 학술연구비에 의하여 지원되었음. (KRF-2005-080-D00006)

[ 참 고 문 헌 ]

- 길양숙, 김동진(2003). 수업평가 결과와 쟁점. 한국교육학회. 교육학연구, 41(3), 303-323.
- 김성숙(1993). 관찰을 통한 교수 평가체계에 대한 측정의 일반화가능도 연구, 교육학 연구, 31(1), 23-40.
- 김성숙(2006). e-learning 강의평가 도구의 일반화가능도와 평가활용의 최적화 조건, 교육평가연구. 19(1), 305-322.
- 김성열, 박재완, 김종철, 강현석(2001). 대학 학사과정 강의평가제 실태분석을 통한 교육업적 평가모형 개발연구. 교육인적 자원부 정책연구, 2001-일-30.
- 성태제(2002). 현대 교육평가. 학지사.
- 송미섭, 지은림(1994). 강의평가 설문지에 관한 문항분석 연구. 교육평가연구, 7(2), 263-283.
- 이종승(1993). 교수의 강의평가에 관한 연구. 충남대학교 교육연구소. 교육발전논총, 14(1), 83-95.
- 지은림, 장정아(2006). 대학 교수자를 위한 강의 자가진단도구 개발의 타당화 및 관련 변인 탐색. 교육평가연구, 19(1), 51-71.
- 한신일, 김혜정, 이정연(2005). 한국 대학의 강의 평가 실태 분석. 교육행정학 연구. 23(3), pp. 379-403.
- AERA, APA, NCME (1985, 1999). *Standards for educational and psychological testing*, Washington D.C.
- Algozzine, B., et. al. (2004). Student evaluation of college teaching: A practice in search of principles. *College Teaching*, 52(4), 134-141.
- Brinko, K. T. (1990). Instructional consultation with feedback in higher education. *J. of Higher Education*, 61, 65-83.
- Broder, J. M., & Dorfman, J. H. (1994). Determinants of teaching quality: What's important to students?. *Research in Higher Education*, 35, 235-249.
- Bukalski, P. J., & Zirpola, D. J. (1993). *Guide for nontenured faculty members: Annual evaluation, promotion, and tenure*. University Film and Association of Georgia State University.
- Cook, S. S. (1989). Improving the quality of student ratings of instruction: A look at two strategies. *Research in Higher Education*, 30, 31-45.
- Feldman, K. A. (1997). Identifying exemplary teachers and teaching : Evidence from student ratings. In R. P. Perry, & J.C. Smart(eds). *Effective teaching in Higher Education* :

*Research and Practice*, New York: Agathon.

- Grimes, P. W., Millea, M. J., & Woodruff, T. W. (2004). Grades-Who's to blame? Student evaluation of teaching and locus of control. *J. of Economic Education*, 35(2) 129-147.
- Hoy, W. & Miskel, C. (2005). *Educational Administration: Theory, Research, and Practice (7th ed.)*. New York: Mcgraw Hill.
- Hoyt, D.P. & Palltett, W. H. (1999). Appraising teaching and effectiveness: Beyond student ratings, *IDEA paper*, 36, IDEA center, Kansas State University, 1-8.
- L'Hommedieu, R., Mengers, R. J., & Brinko, K. T. (1990). Methodological explanations for the modest effects of feedback from student ratings. *J. of Educational Psychology*, 82, 232-241.
- Marsh, H.W.(1994). Weighting for the right criteria in the instructional development and effectiveness assessment(IDEA) system: global and specific ratings of teaching effectiveness and their relation to course objectives. *J. of educational psychology*, 86(4), pp. 631-648.
- Marsh, H.W., & Roche, L. (1993). The use of students' evaluation and an individually structured intervention to enhance university teaching effectiveness. *American Educational Research Journal*, 30, 217-251.
- McGreal, T. L. (1990). The use of rating scales in teacher evaluation: Concerns and recommendations. *J. of Personnel Evaluation in Education*, 4, 41-58.
- McKeachie, W.J. (1996). Research on college : The historical background. *J. of Educational Psychology*, 82, 189-200.
- McKee, B. G., & Dowaliby, F. J. (1985). The relationship between student, course, and instructor characteristics and hearing-impaired students' ratings of instruction. *Volta-Review*, 87, 77-86.
- Murray, H. G. (1985). Classroom teaching behaviors related to college teaching effectiveness. *New Directions for Teaching and Learning*, 23, 21-34.
- Ory, J. C. (2001). Faculty thoughts and concerns about students ratings. *New Directions for Teaching and Learning*. 87, 3-15.
- Saroyan, A. & Amundsen, C. (2001). Evaluating University Teaching: time to take stock. *Assessment & Evaluation on Higher Education*, 36(4), 341-353.
- Seldin, P. (1993). Rating student ratings of professors. *The Education Digest*, 70-72.
- Shevlin, M., et. al. (2000). The validity of student evaluation of teaching in higher education: Love me, love my lectures. *Assessment & Evaluation in Higher Education*, 25(4), 397-405.
- Smith, R. A., & Cranton, P. A. (1992). Students' perceptions of teaching skills and overall effectiveness across instructional settings. *Research in Higher Education*, 33, 747-764.
- Spooren, P., Mortelmans, D. (2006). Teacher professionalism and student evaluation of teaching: Will better teachers receive higher ratings and will better students give

higher ratings?. *Educational Studies*, 32(2), 201-214.

Theall, M., Abrami, P.C. & Mets, L.A. (Eds). (2001). *The Student Ratings Debate: Are They Valid? How can we best use them?* San Francisco: Jossey-Bass.

Wilson, R. C. (1986). Improving faculty teaching—Effective use of student evaluations and consultants. *Journal of Higher Education*, 57, 196-211.