

고등학교 학생에 의한 영어수업 평가도구 개발을 위한 예비연구

권 현 자(아주대학교 박사과정)*
김 주 후(아주대학교교육대학원 교수)

《 요 약 》

본 연구의 목적은 고등학교 학생에 의한 영어수업 평가도구 개발을 위해 예비연구를 수행하는 것이었다. 연구 대상은 경기도 소재 일반계고등학교 영어교사 및 학생 각각 55명과 1,313명이었다. 전문가 집단은 영어수업 전문성 평가에 관한 교육과정평가원 보고서 등에 대한 이론적 검토를 거치면서 총 28개 평가문항을 추출하였다. 평가문항에 대한 학생 이해도 및 중요도를 점검하였으며, 예비조사를 실시하였다. 2010년 1학기말에 실시한 본 조사 결과, 평가도구의 신뢰도는 매우 높은 것으로 나타났다. 수업참여도가 지극히 낮은 학생들을 제외한 1,032명의 결과를 최종적으로 분석한 결과, 4개 요인들(학생수준을 고려한 수업실행, 평가 및 수업관리, 동기부여 및 유발, 효과적인 수업운영)이 추출되었으나, 이 4개 요인들이 영어수업의 질을 균형있게 설명하는데 어려움이 있었다. 첫 번째 요인 ‘학생수준을 고려한 수업실행’이 전체 변량의 47.38%로 나머지 3개 요인들의 설명력은 매우 제한적이었다. 그러나 이러한 결과를 통해 학생들은 학습자의 특성과 능력을 고려한 수업과 학습자의 다양한 학습요구에 부합하는 학습전략과 활동을 선호하는 것을 알 수 있다. 따라서 영어수업의 질 개선을 위한 영어수업평가에서 우선순위를 두어야 하는 것은 학습자측면이다. 요인부하량이 2개 이상의 요인에 걸쳐서 높게 나타난 문항들도 다수 있고, 첫 번째 요인에 대한 축소와 나머지 세 요인에 대한 수정 및 보완을 통하여 보다 조화롭고 타당한 평가도구를 위한 지속적인 연구가 필요하다. 마지막으로 본 연구가 갖는 제한점과 후속 연구에 대한 시사점을 논의하였다.

주제어 : 수업평가, 검사도구, 신뢰도, 타당도

* 제1저자 및 교신자, hyunjak@naver.com

I. 서론

수업의 질을 향상시키기 위한 방법의 하나로 학생이 참여하는 수업평가가 도입되어 활용되고 있다. 학생참여 수업평가는 학생이야말로 수업의 내용을 받아들이는 주체이자 교사와 가장 많은 시간을 보내는 당사자이기에 그 누구보다 타당한 평가자가 될 수 있다는 전제에 기초하고 있다(고영호·김병희, 2005). 가르친다는 것은 학생과 교사 사이에 친밀하고 분리할 수 없는 관계를 만들어내기 때문에 수업평가와 수업개선을 위해서 학생은 중요한 요소로 고려되어야한다(Francisco, 2007). 그러나 학생참여 평가의 타당성에 대한 공감대가 점진적으로 형성되어 왔음에도 불구하고 학생들이 수업을 평가하기 위해서 무엇을 고려해야 할 것인가에 대한 연구는 미진한 상태이다. 특히, 고등학교 수준에서 각 교과별 수업평가의 내용과 기준에 대한 구체적인 논의는 찾아보기 쉽지 않다.

학생이 참여하여 수업의 질을 평가한다는 것은 적어도 한국의 학교교육 문화에서 매우 낮은 경험이다. 전문성을 가진 교사의 수업을 학생이 평가하는 것에 대한 우려가 제기되고 있다. 그러나 학생참여 수업평가를 실시하는 것은 수업의 구성 요소 모두를 대상으로 하는 것이 아니라 피교육자가 더 잘 평가할 수 있는 부분이 있기 때문이다(김충행, 1997; 김원중 등, 1986). 교사와 학생들 간의 상호작용, 수업내용의 명료성, 교사의 열정이나 열의 등은 학생들이 평가할 수 있는 수업 영역이기에 학생이 수업 평가의 주체라 할 수 있다(Eisner, 1979). 학생들은 수업의 질에 대한 훌륭한 수업과 빈약한 수업에 대한 훌륭한 판단자이고, 수업 효과에 대한 타당한 자료를 제공할 수 있다(McKechie, 1990).

학생평가가 가능한 또 하나의 근거는 교사와 가장 많은 시간을 보내고 있는 당사자가 학생이라는 사실이다(안길훈, 2008). 학생은 교사와 가장 많은 시간을 보내며 교사의 교육을 가장 받아들이는 주체라는 점에서 수업평가의 핵심에 있다고 할 수 있다(권낙원, 1993). 교사직무의 대부분은 학생과의 대면 속에서 이루어지기 때문에 학생이 평가자로서의 능력을 보유하고 있는 것과는 별개로 교사를 가장 많이, 잘 알 수 있는 존재라는 사실을 부인하기 어렵다(고영호 등, 2005).

학생참여 수업평가의 효과를 다룬 몇몇 실증연구에서도 학생들에 의한 수업평가가 상당한 수준의 타당성이 있음을 보여주고 있다. 김종한(1998)의 연구에서는 학생에 의한 수업평가가 단지 학생의 수업권 차원뿐 아니라, 학생 스스로의 학업성취에 대한 동기가 유발되었다고 보고하였다. 임정원(2002)은 학생에 의한 수업평가를 통해 학생들의 수업의 효과성을 향상시킬 수 있었다고 주장하고 있다. 김충행(1983)은 초등학교 6학년을 대상으로 교사 수업행동에 대한 학생평정의 가능성을 발견하고, 교사의 노력에 따라 학생평정의 피드백의 효과가 있다고 주장하였다.

한편, 교과내용에 대한 전문성을 갖추고 있는 교사의 수업을 과연 피교육자가 타당하게 평가할 수 있을까에 대한 의문 역시 제기되어 왔다. 수업의 질을 평가하기 위해 학생이 참여하는 사례는 한국 상황에서 수집되고 분석된 자료가 충분치 않으나 국외 사례를 검토해보면 학생에 의한 수업평가의 효용성 증거 또한 상당 수준 축적되어 있다.

학생에 의한 수업평가는 고등교육에서 먼저 시작되었다. Marsh(1987)의 연구에서는 학생들의 수업평가가 상당히 신뢰롭고 타당성이 있으며, 비교적 편견이 없는 다양한 정보를 제공하는 것으로 보고되었다. FinKelstein(1995)은 학생들의 수업평가가 학생들의 학습이나 성취도, 수업의 효율성에 대한 동료평가와 긍정적인 상관을 보여주기 때문에 상당한 신뢰도가 있다고 주장한다. 특히 학생 수업평가에 대한 적절한 피드백을 통해서 수업에 대한 일반적인 영역보다는 특정한 교수행동 영역이 개선되는 결과를 초래하였으며(Murray & Smith, 1989), 학생들이 인식하는 수업의 문제점을 개선하는데도 도움이 되었다(Marsh & Roche, 1993). Cohen(1981)은 학생들의 수업평가 결과와 학업성취도간의 관계를 분석한 결과, 학생에 의한 수업평가가 수업효율성을 나타내는 타당한 준거가 된다고 주장한다. 특히, 학생들의 학습동기, 수업태도와 수업목표 측면에서는 상당히 유용한 정보를 제공한다(Marsh, 1984; McKeachie, 1997). 이외에도 학생에 의한 수업평가가 상당한 신뢰도와 타당도 증거를 확보하고 있음을 밝히는 실증적인 연구들이 있다(Braskamp, Brandenburg, & Ory, 1984; Marsh, 2007; McKeachie, 1979; Marsh & Overall, 1982).

이상의 논의에서 볼 수 있듯이 학생참여 수업평가의 타당성을 지지하는 증거들을 여러 연구를 통해서 발견할 수 있다. 그러나 중요한 것은 “학생에 의한 타당하고 신뢰로운 수업평가를 실시하기 위해 무엇을 평가해야하는가?”의 문제이다. 타당한 평가도구가 있어야 올바른 수업평가가 이루어지고, 올바른 수업평가를 통해 교사의 수업전문성 신장을 기대할 수가 있기 때문이다. 이런 맥락에서 수업의 질을 평가하기 위한 내용으로 Duke와 Stiggins(1990)은 교과에 대한 전문지식, 교수방법, 피드백 제공능력 등을 수업평가의 내용으로 제시하였다. 또한 Scriven(1990)은 수업평가 내용으로, 교과에 대한 전문적 지식, 평가방법에 대한 이해와 평가능력, 교수능력, 가르치는 것과 밀접하게 관련된 지적능력 등을 제시한 바 있다.

국내에서는 학생에 의한 수업평가를 일찍 시작한 서대전고등학교에서는 학습목표, 학습내용, 교수학습방법 및 운영, 평가, 학생자신에 대한 평가를 통해 학기당 두 차례씩 실시하고 있다(홍철표, 2005). 2009년 교원능력개발평가 선도학교나 2010학년도 1학기말에 교원능력개발평가를 실시한 학교에서는 교원에 의한 수업평가(수업준비, 수업실행, 평가 및 활용)와 학생에 의한 수업만족도 평가를 실시하였다. 교원이나 학생에 의한 수업평가는 교육부와 교육청 매뉴얼을 수정·보완하여 보건, 상담, 특수교과를 제외한 모든 교과에서 동일한 기준으로 실시되었다(경기도교육청, 2010; 교육부, 2009; 일반계 A 고등학교, 2010; 이상희, 2009).

영어수업을 평가하기 위해서는 일반 교육학 분야와 동일한 수업 관찰 도구를 사용하였으

나, 1970년대 중반 이후로 외국어 교육 분야만의 수업 관찰기준을 마련하는 연구가 활발해졌다(장경숙 등, 2007). 먼저, Fanslow(1977)의 Focus(Foci on Communication Used in Settings) 수업 관찰 도구에는 의사소통의 근원과 목표, 의사소통 목적, 의사소통 내용, 매체 종류, 매체 사용방법 등이 있다. TALOS(Target Language Observation Scheme)는 쉽게 관찰할 수 있는 저 유추(low inference)영역과 추론해야하는 고 유추(high inference)영역으로 구분하였다. 저 유추영역에는 교사측면(활동대상, 활동종류, 수업내용, 기술, 교수매체, 교수행위와 언어사용)과 학습자측면(활동대상, 발화형태, 질문형태, 언어사용)이 있고, 고 유추영역에는 교사와 학습자의 발화시간, 언어사용의 적절성과 다양성 영역이 있다(Ullmann & Geva, 1982). COLT(Communicative Orientation of Language Teaching)는 수업 관찰 영역을 두 가지로 나누었다. 첫 번째 영역에는 학습자 참여조성, 수업내용, 학습자 양상, 수업자료 영역이 있고, 두 번째 영역에서는 교사와 학습자 발화 상호작용을 다룬다(Fröhlich, Spada & Allen, 1984; Allen et al., 1990).

Nunan(1988)은 영어수업 평가도구를 교실관리, 과업분석, 상호작용 영역으로 나누고 영역별 평가요소를 제시하였으며, 1989년에는 교사측면에서 5가지 평가영역(수업계획, 실행, 관리, 담화, 학습자)과 평가요소를 제시하였다. Wajnryb(1992)도 학습자 측면과 교사측면으로 나누었다. Brown(2001)은 수업준비, 제시, 실행/방법, 교사의 개인적 특징, 교사와 학생 상호작용과 같은 다섯 가지 영역을 제시하였으며 각 영역별 평가요소를 제시하였다.

이러한 맥락에서 우리나라에서도 영어수업을 평가하는 준거에 관한 연구가 이루어져 왔다. 진경애 등(2002)은 영어 교수능력 신장 진단도구의 구성요소로 교육과정, 교사의 수업기술, 학습자 및 학습, 교사의 의사소통능력, 상호작용과 같은 다섯 가지 평가영역을 제안하였다. 또한 고경석 등(2008)은 교사의 영어의사소통능력, 교수내용에 대한 지식, 학생과의 의사소통능력, 교수방법의 효율성, 수업준비도 및 다양한 교재·교구의 활용 등을 제시하였다.

이상의 연구들은 동료교사나 전문가에 의한 영어수업 평가 및 진단 도구이다. 학생들이 참여하여 영어수업의 질을 평가할 수 있는 도구는 찾아보기 쉽지 않다. 동료교사에 의한 간접평가, 장학사나 외부 전문가들의 평가, 학생이나 학부모의 만족도 조사 등이 아닌 실제 교실상황에서 수업을 경험한 학생들이 참여하는 영어수업 평가도구가 필요하다. 본 연구의 목적은 일반계 고등학교 상황에서 영어수업 평가도구 개발을 위해 예비연구를 수행하는 것으로서, 학생이 평가할 수 있는 평가문항을 개발하여 그 타당성을 검증하고 후속연구를 위한 제안점을 제시하고자 하였다.

II. 연구 방법

1. 연구대상

본 연구의 대상은 경기도 소재 9개 고등학교 영어교사 및 학생 각각 55명과 1,313명이었다. 영어수업평가 선행연구 검토 및 문항 작성 단계에는 교육 총 경력 10년 이상의 영어교사 10명이 참여하였으며, 고등학교 1학년 65명의 학생들과 고등학교 3학년 65명의 학생들을 대상으로 각각 평가문항 이해도와 중요도를 검토하였다. 예비조사는 영어교사 2명과 고등학교 1학년 63명의 학생들을 대상으로 실시하였으며, 본 조사에서는 영어교사 43명과 고등학교 1,2학년 1,120명의 학생들이 참여하였다. 연구 단계별로 연구내용 및 연구대상, 검토방법은 <표 1>과 같다.

<표 1> 단계별 연구내용 및 연구대상, 검토방법

단계	연구내용	연구대상			검토방법
		교사	학생	표본 내용	
1	평가준거 점검	10		교육 총 경력 10년 이상 & 고등학교경력 5년 이상 영어교사 (연구자, 박사 1, 석사 3, 학사 5명)	서류검토 토의방식
2	평가문항 작성				
3	평가문항 이해도 검토		65	고등학교 1학년 2개 학급(별도표집)	5단계척도 설문지
4	평가내용 중요도 검토		65	고등학교 3학년 2개 학급(별도표집)	5단계척도 설문지
5	예비조사	2	63	고등학교 1학년 2개 학급	5단계척도 설문지
6	본 조사	43	1,120	고등학교 1학년 549명, 2학년 571명	5단계척도 설문지
계		55	1,313		

2. 평가도구 개발

이 연구의 목적은 학생에 의한 영어수업 평가도구 개발위한 예비연구를 수행하는 것이었다. 이를 위해 먼저, 선행연구 중에서 임찬빈 등(2004, 2005, 2006)의 영어교사전문성 평가기준을 기초자료로 삼고 이를 수정·보완하여 학생에 의한 영어수업 평가도구로 개발하였다. 그 이유는 임찬빈 등의 연구는 영어수업 관련 다양한 선행연구와 국내외 사례검토는 물론 국가수준의 대용량 설문조사 등을 통해 타당화된 기준을 제시하였기 때문이다. 그러나 이 기준은 동료교사나 전문가가 평가대상 영어교사에 대한 관찰이나 인터뷰를 통해 평가한다는

전제하에 개발된 것이기에 본 연구의 목적과 방향에 맞게 수정하였다. 또한 학생에 의한 영어수업 평가도구의 구체적인 개발과정은 모두 6단계로 나뉘어 진행되었다(<표 1> 참조).

가. 1단계: 평가준거 점검

첫 번째 단계에서는 영어수업 평가도구 전체 준거와 하위영역 점검이 이루어졌다. 전문가 집단은 교사 총 경력 10년 이상과 고등학교 근무 경력 5년 이상이면서 현재 일반계고등학교에서 근무하는 영어교사들로 구성되었다. 임찬빈 등(2006)은 영어교사 전문성 평가기준을 4개의 대 영역(지식, 계획, 실행, 전문성)과 6개의 중 영역(교과내용 및 방법지식, 학생이해, 수업설계, 학습환경 조성 및 학급운영, 수업실행, 전문성발달)으로 구분하였다. 그리고 중 영역을 세분화하여 22개 소 영역으로 세분화하고, 소 영역별로 1~2개의 수업전문성 기준이 제시되어 있으며, 이 기준을 평가하기 위한 질문과 지표가 소개되어 있는 영어과 수업평가 매뉴얼은 <표 2>과 같다.

<표 2> 임찬빈 등이 제시한 영어과 수업전문성기준

대 영역	전문성 영역	
	중 영역	소 영역
지식	교과내용 및 방법지식	교과내용
		언어능력
		내용교수법
	학생 이해	발달 특성 이해 개인차 이해
계획	수업 설계	내용선정 조직 및 학습목표 설정
		구조화된 수업 설계
		평가 계획
실행	학습환경 조성 및 학급운영	물리적 환경 조성
		상호작용과 존중
		집단운영
		학생관리
	수업실행	선행지식 활성화 와 동기유발
		다양하고 적절한 수업전략
		학습참여도 고취
		교사의 영어사용
		피드백 제공
		유연한 상황대처
전문성	수업설계	평가의 적절성
		수업반성
		동료교사 및 학부모와의 협력
		전문성 발달노력

평가준거 점검에서 검토자들은 평가의 실용성을 가장 우선적으로 고려하였으며, 전문가 집단 10명 전원이 참여하였다. 현장 교사들의 검토 결과, 학생들의 평가 가능성을 고려하여 대 영역 ‘전문성’은 삭제하였다. ‘전문성’을 평가하기 위한 자기수업반성, 동료교사 및 학부모와의 협력과 전문성 발달을 위해 노력하는 자세 등의 기준은 학생들이 평가하기에 한계가 있기 때문에 삭제하였다.

대 영역 ‘지식’에서 ‘내용 교수법’ 기준은 학생들이 평가하기에 한계가 있다는 점에 대해서는 모든 전문가들이 동의하였으나, 평가지표에 속하는 ‘영어를 통한 외국문화 소개 및 우리나라 문화 소개하기’에 대해서는 전문가들의 의견이 달랐으나, 다수의 전문가들이 평가항목에 포함시키는 것에 동의하였기 때문에 이 평가요소를 포함한 상태에서 다음 단계를 진행하였다.

나. 2단계: 평가문항 작성

다음 단계로, 소 영역을 평가하기 위한 수업전문성 기준 중에서 비슷한 개념은 통합하고, 두 가지 이상의 의미를 가진 질문은 단일개념을 묻는 것으로 수정·보완하여 학생들이 평가할 수 있는 적절한 진술문으로 작성하였다. 예를 들어, 표준자료에서 ‘수업이 효율적으로 이루어지도록 합리적인 절차와 규칙을 적용하고, 수업시간을 효율적으로 활용한다.’와 ‘일정한 행동기준을 정하여 공정하고 평등하게 학생을 지도하고 학생들이 야기하는 문제행동에 대해 적절히 대처한다.’는 평가요소는 ‘선생님은 영어수업 시간 50분간을 짜임새 있게 잘 활용하여 그날 배울 내용을 가르치시는가?’와 ‘선생님은 영어수업 시간에 학생들의 적극적인 참여를 유도하는 행동규칙을 정하여 공정하고 평등하게 적용하시는가?’로 진술하였다.

그리고 내용타당도 부분을 고려하여 고등학교 학생들이 쉽게 이해하고 평가할 수 있는 문장과 단어로 새롭게 수정하였다. 예를 들어, 표준자료의 ‘지식’ 영역에서 ‘영어의 언어학적 체계와 화용에 대한 충분한 지식’과 ‘영어로 의사소통하는 능력’에 관한 평가요소는 각각 ‘선생님의 영어 어휘, 문법, 발음, 독해에 대해 이해하기 쉽게 설명해주시는가?’와, ‘선생님의 영어 듣기, 읽기, 말하기, 쓰기를 학생들에게 시범을 보일 수 있는가?’로 수정하여 삽입하였다.

진술문 작성 단계에서는 연구자를 포함한 3명의 영어교사들이 주로 담당하였으며, 진술문 검토 마지막 단계에서는 전문가 집단 전원이 참여하였다. 이러한 과정을 통해 총 28개 문항을 진술하였다.

다. 3단계: 평가문항 이해도 검토

위 과정을 통해 제작한 진술문에 대해 학생 이해도를 점검하였다. 별도로 표집한 고등학교 1학년 65명의 학생들을 대상으로 실시하였으며, 각 문항에 대해 강한 부정에서 강한 긍

정까지 Likert식 5점 평정척도로 답하도록 하였다. ‘진술문을 어느 정도로 이해하는가?’에 대해 ‘전혀 그렇지 않다’ 1점, ‘그렇지 않다’ 2점, ‘보통이다’ 3점, ‘그렇다’ 4점과 ‘매우 그렇다’ 5점으로 답하도록 구성하였다. 불성실하게 답변한 학생들은 없었으며, 모든 문항 평균은 4.1~4.8로 나왔다.

라. 4단계: 평가내용 중요도 검토

다음으로 평가문항의 중요도를 점검하기 위해 진술된 문항에 대해 학생들이 생각하는 중요도를 검토하였다. 이 도구는 총 28개 문항에 대해 Likert식 5점 척도에서 평정하도록 하였으며, 별도로 표집된 고등학교 3학년 65명의 학생들을 대상으로 실시하였다. ‘전혀 중요하지 않다’ 1점에서 ‘중요하지 않다’ 2점, ‘보통이다’ 3점, ‘중요하다’ 4점과 ‘매우중요하다’에 5점으로 답하도록 구성하였다. 불성실하게 답변한 학생들은 없었으며, 2개 문항을 제외한 모든 문항들의 평균이 3.7이상으로 나타났다. ‘선생님은 영어수업 내용에 따라 짝 활동이나 그룹 활동 등으로 다양하게 구성하여 운영하는가?’에 대한 평균은 2.9로 낮은 수치를 보이지만, 전문가 논의를 참조하여 이 문항을 포함시키기로 하였다. 그리고 ‘선생님은 영어 시간을 통해 다양한 문화를 소개해 주는가?’는 평균 2.2로 낮은 수준이었으나, 예비 검사를 실시하고 나서 최종 삭제 여부를 결정하기로 하였다.

마. 5단계: 예비조사

위와 같은 이론적 검토와 문항 작성 및 검토 과정을 통하여 준비된 영어수업 평가도구는 본 조사에 앞서 예비조사를 실시하였다. 예비조사는 평가도구의 사전 현상분석을 통하여 문제점을 발견하고 문항의 신뢰도를 확보하기 위하여 이루어졌다. 예비조사 설문지는 고등학교 1학년 2학년 65명의 학생을 대상으로 내용 타당도 점검을 실시하였다. 평가대상 영어교사는 2명이었다. 응답이 제대로 이루어지지 않은 2부를 제외한 63부를 분석에 사용하였다. 문항의 내적일치도는 Cronbach's α 값이 .968로 매우 양호하게 나타났으며, 기술통계 결과를 통해 평균이 지나치게 높거나 낮은 문항 그리고 표준편차가 지나치게 크거나 작은 문항들에 검토가 이루어졌는데 이러한 패턴을 보인 문항은 발견되지 않았다.

바. 6단계: 본 조사

예비조사 과정을 거쳐서 최종적으로 선정된 28개 항목으로 구성된 설문지를 통해 본 조사를 실시하였다. 본 조사는 2010학년도 여름방학 이전, 1학기 기말고사가 결과가 나온 시점에서 실시하였으며, 연구대상 학생 및 교사의 특징을 <표 3>에 제시하였다.

본 조사 참여 대상은 경기도 7개 시에 소재한 9개 일반계 고등학교로부터 표집한 영어교

사 및 학생 각각 43명과 1,120명이었다. 교사들은, 1학년 담당 31명, 2학년 담당 12명으로 평균 연령은 36세였다. 학생들은 학년별로는 1학년 549명, 2학년 571명, 성별로는 남학생 577명, 여학생 543명으로 총 1,120명이었다.

학급을 대표할 수 있는 표집을 위해 학교별로 1학년 2개 학급, 2학년의 경우에는 인문과정 1개 학급과 자연과정 1개 학급을 선정하였으며, 남녀 분반의 경우에는 남학생 1개 학급과 여학생 1개 학급을 선정하였다. 연구자가 학교를 직접 방문하여 연구취지를 설명한 후 협조를 받을 수 있는 대표 교사에게 의뢰하여 실시하였다. 설문지 배부 및 회수는 현장 교사 25명이 실시하였는데, 평가의 객관성을 높이기 위해 해당 교과목의 수업을 담당하지 않는 교사가 실시하는 것을 원칙으로 하였다.

본 조사 대상 인원 총 1,239명 중 설문에 대해 부정확한 응답을 보이거나 패턴화된 답변 유형을 보인 119명(전체의 9.6%)명을 제외한 90.40%에 해당하는 1,120명을 1차 선정하였다. 그리고 설문조사에서 영어수업 참여도를 묻는 문항에 대해 ‘전혀 적극적이지 못함’으로 응답한 88명의 학생들을 제외한 1,032명을 최종 분석에 사용하여 평가도구의 구인타당도를 확인하였다.

<표 3> 연구대상 교사 및 학년별, 남녀별 학생 수

학년 \ 성별	학 생				교사		
	남	여	계	특징	남	여	계
1학년	315	234	549	수준별 수업	4	27	31
2학년	262	309	571	인문/자연과정	3	9	12
계	577	543	1,120		7	36	43

3. 자료 분석

본 조사에서 수집된 자료를 토대로 기술통계 분석을 위해 평균, 표준편차, 그리고 문항간 상관관계 계수를 계산하였다. 신뢰도 추정을 위해서는 내적 일관성 지수(Cronbach's α)를 산출하였다. 구인 타당도 증거를 확보하기 위해서는 요인분석을 실시하였는데 요인추출과 회전은 각각 주축분해(principal axis analysis)와 직교회전(varimax)을 활용하였으며, 모든 통계처리는 SPSSWIN 14.0 프로그램을 사용하였다. 평가척도의 신뢰도를 높이기 위해 학생 기초 조사 과정에서 수업참여도 문항에 대해 ‘전혀 적극적이지 못함’으로 응답한 학생들을 제외한 1,032명을 대상으로 요인분석을 실시하였으며, 그 결과를 다음 장에 제시하였다.

Ⅲ. 연구 결과

고등학교 수준에서 학생에 의한 영어수업 평가도구 개발을 위해 28개 문항을 작성하여 1,032명의 학생들로부터 수집된 자료를 분석하였다. 먼저, 기술통계량을 살펴보면 각 문항들의 평균은 2.39~3.31 사이의 분포를 보여 비교적 낮은 점수를 보고하였다. 높은 평균을 보인 문항들은 ‘선생님은 영어 듣기, 읽기, 말하기, 쓰기를 학생들에게 시험을 보일 수 있는가?(M=3.31)’, ‘선생님은 가르친 것을 적절하게 영어평가(정기조사나 수행평가)에 반영하는가?(M=3.14)’, ‘선생님은 영어수업시간 50분간을 짜임새 있게 잘 활용하여 그날 배울 내용을 가르치시는가?(M=3.13)’ 등이었다. 주로 교사의 교수 능력과 평가와 관련된 문항에서 높은 점수를 보였다. 각 문항들의 표준편차는 .941~1.265로 지나치게 높거나 낮은 경우는 발견되지 않았다. 문항간의 상관계수가 .187~.612로 나타났으며, Cronbach’s α 계수는 .951로 높은 신뢰도 수준을 보였다(<표 4> 참조).

<표 4> 문항내용, 문항별 평균 및 표준편차

(N=1,030)

문항	평가요소	평균	표준 편차
1	선생님은 영어 어휘, 문법, 발음, 독해에 대해 이해하기 쉽게 설명해주시는가?	3.04	1.063
2	선생님은 영어 듣기, 읽기, 말하기, 쓰기를 학생들에게 시험을 보일 수 있는가?	3.31	1.110
3	선생님은 영어수업을 통해 다양한 문화를 소개해 주는가?	2.64	1.148
4	선생님은 학생들의 학년 특성을 고려하여 영어수업 방법이나 자료를 준비하는가?	2.85	1.102
5	선생님은 학생의 개인차를 고려하여 영어수업을 진행하는가?	2.48	1.041
6	선생님은 학생들의 학년과 수준에 맞게 영어 학습목표를 제시하는가?	2.84	1.006
7	선생님의 수업내용 전개는 영어 학습목표를 달성하기에 적절한가?	2.89	.984
8	학생들의 수준과 상황에 비추어 영어수업의 양과 난이도가 적절한가?	2.95	.993
9	영어수업 내용의 앞뒤가 학습목표 달성과 잘 연결되어 있는가?	3.03	.947
10	선생님은 학생들의 상·중·하 수준에 적절한 영어수업 전략을 사용하고자 하는가?	2.68	1.061
11	선생님이 제시하는 수업 자료와 교구, 판서는 영어수업 내용 이해에 도움이 되는가?	3.00	1.005

12	선생님은 학습목표, 학생들의 수준, 수업내용을 고려하여 적절한 영어 평가기준을 제시하는가?	2.88	.941
13	선생님은 학생들이 영어수업에 집중할 수 있도록 교실 환경을 조성하는가?	2.83	1.061
14	선생님은 교사와 학생, 학생과 학생 사이에 활발하고 긍정적인 상호작용이 이루어지도록 영어 학습 분위기를 조성하는가?	3.00	1.121
15	선생님은 영어수업 시간 50분간을 짜임새 있게 잘 활용하여 그날 배운 내용을 가르치시는가?	3.13	1.083
16	선생님은 영어수업 내용에 따라 짝활동이나 그룹활동 등으로 다양하게 구성하여 운영하는가?	2.55	1.265
17	선생님은 영어수업 시간에 학생들의 적극적인 참여를 유도하는 행동규칙을 정하여 공정하고 평등하게 적용하시는가?	2.76	1.110
18	선생님은 새로운 영어수업 내용을 제시할 때, 이전에 배운 내용과 연계하여 설명하는가?	3.07	1.022
19	선생님은 새로운 영어수업 내용을 제시할 때, 학생들의 흥미와 학습동기를 유발하는가?	2.76	1.077
20	선생님은 가르치는 영어수업 내용에 따라 다양한 방법을 사용하는가?	2.64	1.089
21	영어수업 시간에 배운 내용을 실제로 활용해 보고 싶은 마음이 생기는가?	2.43	1.178
22	선생님은 학생들의 적극적인 영어수업 참여를 유도하고 있는가?	2.92	1.115
23	선생님은 수업진행에 필요한 영어를 학생들의 수준에 맞추어 적절하게 구사하는가?	2.88	1.010
24	선생님의 영어 사용이 학생들로 하여금 원어민과 영어로 대화하고 싶은 마음이 생기도록 하는가?	2.39	1.153
25	학생들의 질문이나 대답에 대한 선생님의 반응이 학생들의 영어 학습의욕과 효과를 향상시키는가?	2.66	1.058
26	선생님은 영어수업에서 학생들의 돌발적인 행동이나 예기치 못한 사태에 대해 잘 대처하는가?	2.99	1.045
27	선생님은 가르친 것을 적절하게 영어 평가(정기고사나 수행평가)에 반영하는가?	3.14	1.129
28	선생님은 평가 결과를 학생들의 영어 성적 향상이나 수업방법 개선을 위해 잘 활용하고 있는가?	2.84	.988

다음으로 구인타당도 점검을 위해 요인분석을 실시하였는데 그 결과를 <표 5>에 제시하였다. 고유값(eigen value) 1.0을 기준으로 하여 그 이상의 값을 갖는 요인만을 고려한 결과 4개의 요인들이 추출되었는데 전체 변량의 56.8%를 설명하였다. 그러나 요인부하량이 2개 이상의 요인에 걸쳐서 높게 나타난 문항들도 다수 있어 추출된 요인의 해석에 주의를 요하는 경우도 발견되었다.

<표 5> 문항별 회전된 요인행렬

문항 번호	평가내용	요인명	요인1	요인2	요인3	요인4
6	학생수준에 적절한 수업목표 제시	학생수준을 고려한 수업실행	.706	.201	.318	.102
5	학생개인차를 고려한 수업진행		.670	.047	.407	.193
4	학년특성을 고려한 수업방법		.663	.167	.194	.313
7	수업목표 달성을 위한 수업전개		.644	.323	.180	.154
8	학생수준에 적절한 수업양과 난이도		.605	.282	.149	.138
10	학생수준에 적절한 수업전략		.588	.068	.357	.204
12	학생수준과 수업내용을 고려한 평가기준		.582	.299	.181	.192
9	수업목표 달성을 위한 조직적인 진행		.578	.417	.124	.144
1	교사의 설명능력		.548	.433	.103	.253
11	효과적인 수업자료 및 판서	평가 및 수업관리	.531	.393	.181	.148
23	학생수준에 적절한 교사의 영어구사		.410	.309	.372	.235
27	수업내용과 평가의 연계성		.128	.690	.248	.077
26	돌발적인 상황이나 행동에 대한 대처		.213	.628	.333	.156
15	50분간 수업의 짜임새		.364	.601	.047	.322
18	수업내용의 전시학습과 연계성		.256	.566	.245	.089
2	교사의 시범능력		.428	.523	-.027	.269
28	평가 결과에 대한 피드백		.297	.499	.385	.124
21	실생활 영어 활용 동기부여	동기부여 및 유발	.209	.164	.714	.162
25	교사의 반응이 영어 학습의욕 유발		.287	.140	.669	.215
20	수업내용에 따른 다양한 수업방법		.254	.230	.641	.322
19	학생의 흥미와 동기유발		.178	.361	.533	.393
24	교사의 영어사용이 학생들의 동기유발		.293	.405	.512	.304
16	학습자 조직	효과적인 수업운영	.098	-.006	.286	.738
17	공정하고 적절한 학습자 행동관리		.192	.247	.319	.627
14	상호작용을 유도하는 수업분위기 조성		.242	.399	.111	.603
3	다양한 문화 소개		.420	.065	.170	.566
13	주의집중을 위한 교실환경 관리		.314	.380	.140	.503
22	학생들의 참여도 유도		.215	.375	.387	.496
고유치			12.238	1.473	1.172	1.027
설명분산			43.709	5.261	4.187	3.668
누적된 명분산			43.709	48.970	53.157	56.825

각 요인의 이름은 상대적으로 높은 요인부하량을 보인 문항들을 중심으로 하면서 선행연구 등을 참조하여 정하였다. 추출된 4개 하위 요인들을 살펴보면 다음과 같다.

첫 번째 요인은 11문항으로 전체 설명분산의 43.701%를 설명하는 주요인이다. 이 요인으로 묶인 주요 문항들을 보면 ‘선생님은 학생들의 학년과 수준에 맞게 영어 학습목표를 제시하는가?’, ‘선생님은 학생들의 학년 특성을 고려하여 영어수업 방법이나 수업자료를 준비하는가?’, ‘학생들의 수준과 상황에 비추어 영어수업의 양과 난이도가 적절한가?’ 등으로 첫 번째 요인의 저변에는 학생들의 수준과 특성을 고려한 학습목표 설정, 학습량, 학습전략 등에 관한 것이었다. 요인 2와 3에도 0.4 이상의 부하량을 보인 문항 1, 5와 9를 감안하여 ‘학생수준을 고려한 수업실행’으로 명명하였다.

두 번째 요인은 6개 문항으로 전체 설명분산의 5.261%를 설명하였다. 두 번째 요인에 묶인 문항들은 ‘선생님은 수업시간에 가르친 것을 적절하게 영어 평가에 반영하는가?’, ‘선생님은 영어수업에서 학생들의 돌발적인 행동이나 예기치 못한 사태에 대해 잘 대처하는가?’, ‘선생님은 영어수업 시간 50분간을 짜임새 있게 잘 활용하여 그날 배울 내용을 가르치시는가?’ 등이었고, 문항 2번은 첫 번째 요인과의 0.428의 요인부하량을 보였다. 주로 수업내용과 평가의 연계성, 상황에 대한 유연한 대처, 수업 내용의 구성 등에 관한 것으로 ‘평가 및 수업관리’로 명명하였다.

세 번째 요인은 5개 문항으로 전체 설명분산의 4.187%를 설명하였다. 문항들을 살펴보면 ‘영어수업 시간에 배운 내용을 실제로 활용해 보고 싶은 마음이 생기는가?’, ‘학생들의 질문이나 대답에 대한 선생님의 반응이 학생들의 영어 학습의욕과 효과를 향상시키는가?’, ‘선생님은 가르치는 내용에 따라 다양한 영어수업 방법을 사용하는가?’ 등으로 학생들의 실생활 활용 동기부여, 학생들의 영어 학습동기를 유발하는 교사의 반응이나 학습내용, 다양한 수업 방법 등에 관한 것이었다. 문항 24번은 두 번째 요인과의 0.405의 부하량을 보였는데 전체적으로 보아 ‘동기부여 및 유발’로 명명하였다.

네 번째 요인은 6개 문항으로 전체 설명분산의 3.668%를 설명하였다. 주된 문항은 ‘선생님은 영어수업 내용에 따라 짝활동이나 그룹활동 등으로 다양하게 구성하여 운영하는가?’, ‘선생님은 영어수업 시간에 학생들의 적극적인 참여를 유도하는 행동규칙을 정하여 공정하고 평등하게 적용하시는가?’, ‘선생님은 교사와 학생, 학생과 학생사이에 활발하고 긍정적인 상호작용이 이루어지도록 영어 학습분위기를 조성하는가?’ 등이었다. 주로 학생구성, 학습자 행동관리, 학습문화, 주의집중과 참여도 유도 등으로 ‘효과적인 수업운영’으로 명명하였다.

이상의 분석 결과를 종합해볼 때, 본 연구에서 개발한 영어수업 평가도구의 신뢰도는 매우 높은 것으로 나타났다. 그러나 구인타당도 증거를 살펴볼 때, 추출된 네 가지 요인들은 영어수업의 질을 균형있게 설명하는데 어려움을 보였다. 특별히 전체 변량의 43.7%를 설명한 첫 번째 요인은 나머지 요인들을 압도하는 주된 요인으로 밝혀졌다. 나머지 세 요인들의

경우 설명력이 적어 본 연구에서 설정한 구인을 이론적으로 해석하는데 제한점을 보였다.

IV. 논의

본 연구의 목적은 영어수업 평가도구를 개발을 위한 예비연구를 수행하는 것으로서, 국가 수준에서 개발한 영어교사전문성 평가기준의 내용을 토대로 학생들이 평가할 수 있는 문항을 개발하였다. 이 작업을 위해 10명의 영어교사로 구성된 전문가 집단과 43명의 영어교사 및 1,313명의 학생들이 참여하였다.

연구의 결과에 의하면, 새로 개발된 문항들은 높은 수준의 내적일관성을 보여 신뢰할만한 것으로 나타났다. 문항의 중복성은 높지 않았으며 고등학교 학생들이 이해하고 평정하기에 적절하였다. 예비조사 과정에서 일부 문항에 대해 중요도가 떨어진다는 반응도 있었으나 본 조사 과정을 통해 큰 문제가 없는 것으로 나타났다.

그러나 구인타당도 점검을 위한 요인분석 결과를 보면 본 연구에서 제시한 문항들은 네 가지 요소가 균형을 이루면서 이론적인 해석을 하는데 제한점을 보여 주었다. 특히, 첫 번째 요인인 ‘학생수준을 고려한 수업실행’은 전체 변량의 대부분을 설명하여 나머지 세 요인들의 설명력은 매우 제한적이었다.

이러한 결과는 우리나라 고등학교 영어수업 상황에서 학생들은 무엇보다 학생들의 수준과 특성을 감안한 수업진행이 매우 중요한 것으로 판단하고 있음을 보여준다. 이러한 결과는 케임브리지 ESOL(English for Speakers of Other Languages) 영어교사 수업평가 영역을 비롯하여(이병천, 2009), 영어수업의 실태에 대한 선행연구들(Danielson, 1996; Nunan, 1989; Brown, 2001; 신동일, 2003; 임찬빈 등, 2006)에서도 밝힌 것처럼 학생들의 입장에서 출발하고 그들을 배려하는 노력이 영어수업의 성패를 결정하는 요인임을 보여주고 있다. 또한 학생들의 영어수업 참여에 있어서 큰 편차가 있음을 간접적으로 보여주며 이 부분에 대한 고려 없이 효과적인 영어수업 진행의 어려움을 시사해 준다.

나머지 세 가지 요인들 ‘평가 및 수업관리’, ‘동기부여 및 유발’, ‘효과적인 수업운영’의 경우를 보면 비록 설명력이 떨어지기는 하나 영어수업의 질을 이루는 주요 요소임에 틀림없다. 이러한 요소들은 전문가나 동료교사에 의한 영어수업 평가 연구들(이병천, 2009; Danielson, 1996; Nunan, 1988; Wajnryb, 1992; Brown, 2001; 임찬빈, 2006)에서도 여러 차례 강조되고 있다. 즉, 효과적인 영어수업을 위해서는 평가내용의 적절성, 동기부여 및 효과적인 수업을 위한 영어수업 환경의 중요성을 시사해 주고 있다.

그런데 요인추출 과정에서 여러 개의 요인에 걸쳐 요인부하량이 높았던 ‘선생님은 영어의

어휘, 문법, 발음, 독해에 대해 이해하기 쉽게 설명해주시는가?’, ‘선생님은 영어 듣기, 읽기, 말하기, 쓰기를 학생들에게 시범을 보일 수 있는가?’, ‘영어수업 내용의 앞뒤가 학습 목표 달성과 잘 연결되어 있는가?’ 문항들의 경우, 학생들의 입장에서 보면 ‘학생수준을 고려한 수업 실행’과 ‘동기부여 및 유발’ 요인들이 뚜렷하게 구분되지 않을 수도 있음을 말해준다.

이러한 결과 및 논의를 토대로 향후 영어수업 평가도구의 타당화를 위한 제언을 하면 아래와 같다.

첫째, 영어수업 평가준거를 설정함에 있어서 보다 더 학생 수준 및 특성을 고려한 내용에 초점을 맞출 필요가 있다는 것이다. 기존의 평가도구들은 바람직한 영어수업의 특성을 감안하기는 했지만 다소 이론적인 분석에 치우친 측면이 있다. 이 연구 결과를 통해서 상당수의 고등학생들은 영어수업에 대해 학습자의 특성과 능력을 고려한 학습목표, 학습량과 난이도, 평가기준 설정을 원하고 학습자의 다양한 학습 요구에 부합하는 학습전략과 방법, 수업자료와 활동을 선호하는 것으로 나타났다. 학생들이 생각하는 보다 좋은 영어수업 개념을 조사하기 위해서는 제한된 질문에 학생들이 선택하는 유형이 아닌 자유질문지 형식을 통해 좋은 영어수업에 대한 학생들의 의견을 수렴할 필요가 있다.

둘째, 요인분석 결과 요인부하량이 2개의 요인에 걸쳐서 높게 나타난 문항들이 다수 있어 문항들이 독립적이지 못하고 상당 부분이 서로 관련되어 있는 것으로 판단된다. 학교수업의 성격 자체가 상호 관련된 일련의 활동으로 이루어진다는 점을 감안하더라도 문제점을 개선하기 위해서는 문항들 간에 상호 중첩되는 부분을 가급적 최소화시키는 것이 바람직하다고 본다.

셋째, 전체 설명분산의 43% 이상을 설명하는 첫 번째 요인은 문항의 내용 및 수를 압축하고 나머지 세 요인들은 문항을 수정하거나 추가할 필요가 있다. 새롭게 구성된 문항들을 활용하여 고등학교 학생들에 의한 영어수업 평가도구 타당화에 대한 지속적인 연구가 필요하다.

그러나 본 연구의 결과를 해석하는데 있어 몇 가지 제한점이 있다. 먼저, 본 연구의 전체적인 틀은 임찬빈 등(2006)의 연구결과를 기초로 수행되었다. 비록 이 연구가 국가수준의 연구를 토대로 진행되었기 때문에 기존의 개인수준에서 진행된 연구들에 비해 포괄적인 타당성을 보여준다고 할 수 있으나, 이 역시 학생입장의 평가를 전제로 한 연구는 아니었다. 따라서 학생에 의한 영어수업 평가를 위한 도구 개발을 위해서는 보다 학생의 입장과 상황을 고려한 이론적 틀을 바탕으로 접근할 필요가 있다. 특별히, 우리나라 고등학교 학생들이 영어를 공부하는 과정에서 나타나는 다양한 학습동기와 관련된 변인들을 고려하고, 교사와 학생 사이의 상호작용에 대한 다양한 측면들을 고려하여야 한다.

둘째, 본 연구의 자료수집 및 분석은 학생들이 스스로 자신들의 특성을 측정하는 일반적인 구인(예: 학업동기)과는 달리 매우 복잡한 구조를 보여 준다. 즉, 측정하고자 하는 구체적

인 구인이 영어수업의 질이기에 학생입장에서는 자신을 가르치는 특정 교사 수업의 질에 대한 평가만을 하게 된다. 연구 디자인상으로 보면 학생들의 응답이 교사들 안에 내재되는 (nested) 다층적인 구조를 보이기에 자료 분석에 있어서 보다 정교한 접근이 요구된다. 이번 연구에서는 학생에 의한 영어수업 평가도구 개발을 위한 첫 시도로서 측정문항의 선정 및 정교화 작업에 초점을 맞추었으나 학교 및 학생 표집 상의 제한점과 더불어 자료수집 및 분석상의 정교한 노력은 부족하였다.

세 번째, 위의 두 번째에 유사한 맥락에서 신뢰도 및 타당도 검증에 있어서 Cronbach's α 계수의 보고와 탐색적 요인분석 적용에 있어서도 한계가 있다. 측정상의 기술적인 문제해결 관점으로 보면 내재된 구조를 반영한 분석방법(예: 일반화가능도 이론, 위계적 회귀분석 등)이 고려될 필요가 있다. 또한 후속연구에서는 본 연구에서 얻어진 평가도구의 요인구조를 확인하는 보다 엄밀한 검증방법(예: 확인적 요인분석)이 이루어져야 할 것이다.

비록 연구결과 해석에 제한점이 있기는 했으나, 본 연구의 가장 큰 의의는 고등학교 학생 입장에서 평가할 수 있는 영어수업 평가도구 개발을 위한 탐색적 연구라는 데 있다. 즉, 학생에 의한 영어수업 평가도구 문항을 도출하고 이를 바탕으로 영어수업의 질을 평가하고자 하는데 의의가 있다. 이러한 평가도구를 활용하여 영어수업을 평가하면 외부자적 관점이 아니라, 학생들이 원하는 영어수업에 대한 실질적인 피드백을 받을 수 있고, 학생들의 영어수업 만족도를 높일 뿐만 아니라 영어수업의 질 개선을 기대할 수 있다.

참 고 문 헌

- 경기도교육청(2010). 교원 능력 개발 평가 매뉴얼(중등).
- 고경석, 김정렬, 김진석, 김해동, 류호열, 정숙경(2008). 중등학교 영어과 교사 자격기준 개발과 평가영역 상세화 및 평가도구 개발 연구. **한국교육과정평가원 연구보고서**.
- 고용호, 김병희(2005). 교사평가제의 쟁점연구, 경북대학교 사범대부속중등교육연구소, **중등교육연구**. 53(1), 1-28.
- 교육과학기술부(2009). 교원능력개발평가 매뉴얼: 초등·중등·교육청.
- 권기욱(1993). **대학교직원 평가: 교수·직원·행정가 평가**. 서울: 교육과학사.
- 권낙원(1993). 효과적인 수업을 위한 교사평가. **한국교원신문**. 1998.2.
- 김원중, 권오, 박운영, 이창원, 홍성애(1986). 대학생들에 의한 강의평가: 과학적 측정2 개발을 위한 노력. **경기대학교 학생생활연구**. 5(2), 26-59.
- 김이경·박상완(2005). 성장 중심 교사평가제 탐색: 교원의 인식을 중심으로. **교육학연구**. 43(3), 157-183.
- 김종한(1998). 고등학생의 교사수업평가에 관한 연구. 박사학위논문, 충남대학교 교육대학원.
- 김충행(1983). 교사효율성에 대한 학생평정과 그 피이드백의 효과에 관한 분석적 연구. 박사학위논문. 동국대학교 대학원.
- 김충행(1997). 수업방법 개선을 위한 학생에 의한 대학 강의 평가의 분석. 한국교육과정학회. **교육과정연구**. 15(1), 203-240.
- 신동일(2003). 비영어권 교사들의 영어수업 진행능력 평가영역 개발: ESP 관점에서. **영어교육**. 58(4), 61-98.
- 안길훈(2008). 학생의 교사평가에 대한 참여 가능성 연구: 현행 교원평가제도를 중심으로. **교육행정학연구**. 25(2). 153-180.
- 이병천, 김진석, 이명애, 이영주, 장우삼(2009). 영어로 진행하는 영어수업(TEE) 평가 도구 개발. **한국교육과정평가원 연구보고서**.
- 일반계 A고등학교(2010). 교원 능력 개발 평가 계획서.
- 이상희(2009). 교원능력개발 평가 선도학교 운영 실태 연구. 석사학위논문. 경인대학교 교육대학원.
- 임찬빈, 이화진, 박영순, 강대현, 박영석, 정유진(2004). 수업평가 기준 개발 연구(I): 일반 기준 및 교과(사회, 과학, 영어) 기준 개발. **한국교육과정평가원 연구보고서**.

- 임찬빈, 이화진, 서지영, 차우규(2005). 수업평가 기준 개발 연구(Ⅱ): 일반 및 교과(영어, 도덕, 체육) 기준상세화. **한국교육과정평가원 연구보고서**.
- 임찬빈, 이 윤, 어도선(2006). 수업평가 매뉴얼: 영어과 수업평가 기준. **한국교육과정평가원 연구보고서**.
- 임정원(2002). 중등학교에서의 학생에 의한 교사평가 가능성 연구. 석사학위논문. 부경대학교 교육대학원.
- 장경숙, 주형미, 김한주, 이복희(2008). 영어수업개선 연구대회. **한국교육과정평가원 연구보고서**.
- 진경애, 김진석, McNamara, T., Brown, A.(2002). 중등 영어 교수 능력 진단 프로그램 개발 연구. **한국교육과정평가원 연구보고서**.
- 홍철표(2005). 교사평가제도에 대한 학교구성원들의 인식비교, 전북대학교 교육대학원 석사학위논문.
- Allen, J. P., Swain, M., Harley, B., & Cummins, J. (1990). Aspects of classroom treatment: toward a more comprehensive view of second language education. In B. Harley, P. Allen, J. Cummins, & M, Swain(eds). *The development of second language proficiency*. Cambridge University Press. 58-81
- Braskamp, L. A., Brandenburg, D. C, & Ory, J. C. (1984). *Evaluating teaching effectiveness: a practical guide*. Beverly Hills: Sage Publications.
- Brown, H. D. (2001). *Teaching by principles: an interactive approach to language pedagogy*(2nd edition). Longman: NY.
- Cohen, P. A.(1981). Effectiveness of student ratings feedback and consultation for improving instruction in dental schools. *Journal of Dental Education*, 55(2), 145-150.
- Danielson, C. (1996). Enhancing professional practice: a framework for teaching. *Alexandria, VA: Association for Supervision and Curriculum Development*.
- Duke, D. L. & Stiggings, R. J. (1990). Beyond minimum competence: evaluation for professional development, In Millian, J. and Darling-Hammond, L(Eds.). *The new handbook of teacher evaluation: assessing elementary and secondary school teachers*. Newbury Park, CA: Sage, 116-132.
- Eisner, E. W. (1979). *The educational imagination: on the design and evaluation of school programs*. New York.: Macmillian Publishing Co.
- FinKelstein, M. J. (1995). College faculty as teachers. In H. Weeksler(Ed.), *The NEA 1995 Almanac of Higher Education*, 33-37. Washington, D.C.: National

- Education Association.
- Fanslow, J. (1977) Beyond rashomon: conceptualizing and describing the teaching act. *TESOL Quarterly*, 11(1), 17-41.
- Francisco, Z. (2007) The use and misuse of student evaluations of teaching. *Teaching in Higher Education*. 12(1), 55-76
- Fröhlich, M., Spada, N., & Allen, J. (1984). Differences in the communication orientation of L2 classrooms. *TESOL Quarterly*. 19, 51-62.
- Marsh, H. W. (1984) Marsh, H. W.(1984). Students' evaluations of university teaching: dimensionality, reliability, validity, potential biases and utility, *Journal of Educational Psychology*, 76(5), 707-754.
- Marsh, H. W. (1987) Students' evaluations of university teaching: research, findings, methodological issues, and directions for future research, *International Journal of Educational Research*, 11, 254-388.
- Marsh, H. W. (2007). Students' evaluations of university teaching: a multidimensional perspective", In R. P. Perry & J. C. Smart(Eds.). *The scholarship of teaching and learning in higher education: an evidence based perspective*. New York: Springer.
- Marsh, H. W. & Overall, J. U. (1982). Research current: students' evaluations of teaching, an update. *American Association for Higher Education Bulletin*. 35(4), 9-12.
- Marsh, H. W. & Roche, I. A. (1993). Making students' evaluation of teaching effectiveness. *American Psychology*, 52(11), 1187-1197.
- McKeachie, W. J. (1979). Student ratings of faculty: a reprise. *The American psychologist*. 52(11), 1218-1225.
- McKechie, W. J. (1990). Research on college teaching: the historical background. *Journal of Educational Research*. 82(2). 56-85.
- McKeachie, W. J. (1997). Student ratings: the validity of use. *American Psychologist*. 52(11), 1218-1225.
- Murray, H. G., & Smith, T. A. (1989). *Effectives of midterm behavioral feedback on end-of-term ratings of instructor effectiveness*. American Educational Research Association, San Francisco.
- Nunan, D. (1988). *The learner-centered curriculum*. Cambridge: Cambridge UP.
- Nunan, D. (1989). *Understanding language classrooms: a guide for teacher*

- initiated action*. London: Prentice-Hall.
- Scriven, M. (1990). Teacher selection. In Million, J. and Darling-Hammond, L.(Eds.), *The new handbook of teacher evaluation: assessing elementary and secondary school teachers*. Newbury Park, CA: Sage, 116-132.
- Ullmann, R. & Geva, E. (1982). *The target language observation scheme (TALOS)*. New York Board of Education, Core French Evaluation Project. Ontario Institute for Studies in Education, Toronto.
- Wajnryb, R. (1992). *Classroom observation tasks: a resource book for language teachers and trainers*. Cambridge University Press.

• 논문접수 : 2010년 9월 1일/ 수정본 접수 : 2011년 2월 25일/ 게재승인 : 2011년 3월 11일

ABSTRACT

A Preliminary Study on the Development of the Student Assessment Tool for Teaching English at High School Level

Hyunja Kwon(Doctor's course, Ajou University)

Juhu Kim(Professor, Ajou University)

The purpose of this preliminary study was to develop the student assessment tool for teaching English and examine the validity and reliability of it in high school settings. For this study, 53 English teachers and 1,313 students were sampled from 9 high schools in Kyeonggi-do. After the literature review was undertaken by the English teachers based on the KICE(Korea Institute for Curriculum and Evaluation) report which developed the evaluation instrument for teaching English, a total of 28 questions were utilized. The quality of 28 questions were ensured through examining the responses of the high school students.

The final survey was taken by the students after the first semester in 2010. With the deletion of the students who showed the strongly passive attitude in the English class, the final student sample was reduced to 1,032. On the basis of the survey results, we can conclude that student assessment tool in this study has high reliability(Cronbach's = .951). But we also conclude that the construct validity was found to be unsatisfiable levels. Because 56.8% of the entire variates falls on the first factor among the 4 confirmed factors(Teaching Considering Student Levels, Evaluation & Class Administration, Promotion of Motivation, Effective Class Management), which has meaningful implications. Student levels, needs, grades and characteristics need to be considered to improve English teaching. Finally, limitations and suggestions for future study were discussed.

Key Words : *Student assessment of teaching, Assessment tool, Validity, Reliability*