영어어문교육 18권 2호 2012년 여름

대학수학능력시험(영어시험)의 문항개선에 대한 연구*

전 성 애
(전북대학교)

Jeon, Sung-Ae. (2012). A study on the improvement of the test items in Korean scholastic ability test (English test). *English Language & Literature Teaching, 18*(2), 189-211.

The purpose of the study was to explore ways to improve the test items on the Korean scholastic ability test. More specifically, the researchers investigated whether use of the target language in test items would make a difference in total scores, discriminatory power, and item difficulty. A total of 288 high school seniors participated in the study. The subjects were divided into the experimental group (N=145) and the control group (N=143). A 25-item test resembling the Korean scholastic ability test was administered to both groups. The experimental group was given items whose questions and alternatives were all presented in English, whereas the control group was given items whose questions and alternatives were presented in Korean only. Statistical analyses revealed that use of English vs. Korean in the questions and alternatives made a significant difference in total scores, item discrimination, and item difficulty level. The findings strongly suggest that use of English is one way to improve the quality of the Korean scholastic ability test by enhancing item discrimination and face validity. Considering that the test in question is a high-stakes exam in Korea, further research on how to improve the Korean scholastic ability test is urgently called for.

[Language testing/test format/Korean scholastic ability test/item discrimination/ item difficulty, 언어평가/시험형태/대학수학능력시험/문항변별도/문항난이도]

I. 서론

1. 연구의 필요성 및 목적

우리나라에서 시험이 가지고 있는 의미나 시험의 결과는 학습자들의 학습 발달을 위해 매우 중요하고 그 영향이 지대하다. 시험은 교육 목표의 달성

* 이 논문은 저자의 학위 논문 일부를 수정 보완한 것임.

교보문고 KYOBO Book Centre

정도를 점검하는 과정으로 교육과정에 포함되어 있다. 초·중등학교의 올바른 공교육이 이루어지기 위해서는 국가가 제시하고 있는 교육과정의 틀 안에서 교수.학습의 효과와 효율성이 극대화되어야 한다. 특히, 교육 목표의 달성 여부를 확인하는 작업으로 평가(feedback)가 이루어져야 한다.

고등학교 영어과 교수.학습 목표의 근거는 교육과학기술부가 고시한 국가 수준의 교육과정에서 찾을 수 있다. 고등학교 교육과정에서 의도한 성취 목표의 도달 여부와 대학에서 수학할 수 있는 능력에 대한 공식적인 최종 평가는 대학수학능력시험으로 이루어진다. 대학수학능력시험은 Hughes(1989, p. 1)가 제시하고 있는 역류 효과(backwash effect)가 지나치게 커서 그것의 유형과 내용에 따라 고등학교 영어교육 전반에 걸쳐 영향을 주는 것이 현실이다(전병만, 2004).

대학수학능력시험은 국립평가원의 주관 하에 고등학교 영어교육의 정상화를 위한 시도로서 1990 년 12 월에 1 차 실험평가가 실시된 이래 1992 년 11 월까지 7 차례의 실험평가를 거쳐 새로운 평가 모형으로 자리를 잡게 되었다. 실제로 1993 년도에 처음으로 실시된 대학수학능력시험은 과거 학력고사에 비해 우리나라 교육 역사상 처음으로 음성 언어인 듣기를 포함시켜 영어교육의 정상화를 이룰 수 있는 획기적인 전기를 마련하였다. 1993 년에 처음 도입된 대학수학능력시험은 현재까지 문항 수, 문항 유형, 배점 등 약간의 변화는 있었지만 기존의 모형을 대체로 유지해오고 있다.

대학수학능력시험이 처음 실시된 1993 년도부터 현재까지 교육과정의 내용은 큰 변화를 보였다. 이 시기 동안에 국가교육과정은 5 차 교육과정(1987-1992), 6 차 교육과정(1992-1997), 7 차 교육과정(1997-2007), 개정 7 차 교육과정(2008), 재개정 7 차 교육과정(2011 이후)으로 발전해왔다. 교육과정의 변화는 교육 목표, 교육 내용, 교수·학습 방법, 평가가 달라진다는 것을 의미한다. 그러나 교육과정의 변화에 따라 평가가 달라져야 함에도 불구하고 기존의 방법을 그대로 유지함으로써 영어교육의 정상화를 기대할 수 없었다. 실제 우리나라의 영어교육은 바람직한 교수·학습 방법을 수용하고 교육현장에 적용하여 영어교육의 목적을 달성하려는 노력 보다는 오히려 대학입학 시험의 준비 교육으로만 존재하고 있는 상황에서 학생들의 영어능력의 질적 향상이나 기능적인 면의 개발은 기대하기 어려운 실정이다(전병만, 2004).

이 연구는 지금 시행되고 있는 현행 대학수학능력시험(영어시험)의 모형을 중심으로 하고 있다. 현 시점에서 정부가 2016 년에 시행 계획으로 밝히고 있는 국가영어능력시험(1 급, 2 급, 3 급)은 iBT(internet-Based Test)로서 언어의 네 가지 기능을 직접 평가한다는 점에서 매우 바람직하며 영어교육의 정상화를 이룰 수 있어 긍정적이다. 1 급은 대학생 이상 성인을 대상으로 한 영어 능력 시험이며, 2 급이나 3 급은 고등학교 학생을 대상으로 한 대학입학 시험으로서 활용될 예정이다. 그러나 우리나라 현실을 고려할 때 2 급이나

교보문고 KYOBO Book Centre

3 급의 시험은 영어 능력 자격 시험으로는 가능하나 현행 입학시험과 같이 당락을 결정짓는 평가 도구로 사용될 경우에 많은 문제점을 야기할 수 있다는 점은 자명하다. 예를 들면, 표현 기능의 평가에 있어서 채점자 신뢰도, 동형 문항간의 난이도, 컴퓨터 등 기자재의 오작동 등이 여기에 해당된다.

최근에 정부는 국가영어능력시험 개발에 많은 관심을 기울이고 있지만, 대학수학능력시험의 문제점 분석과 질적 향상을 위한 문항 개선에 대한 연구는 지금까지 소홀히 해왔다. 특히, 질문과 답지의 목표 언어 사용을 통한 문항 개선에 대해서는 체계적인 관심을 기울이지 못했다. 예를 들면, 영어 숙달도를 측정하는 세계적인 표준화 시험-TOEFL(Test of English as a Foreign Language), IELTS(International English Language Testing System) 등-의 문항은 질문과 답지가 모두 영어로 되어 있다. 또한 중국, 일본, 러시아, 인도네시아 등의 나라에서 시행하고 있는 대학입학 영어시험의 질문과 답지가 모두 영어로 되어 있다(김용명, 이완기, 김진석, 고현숙, 김상재, 2010). 그러나 우리나라의 대학수학능력 시험(영어시험)에서는 영어시험 문항의 모든 질문은 우리말로, 답지는 영어나 우리말로 구성되어 있다. 따라서 이 부분과 관련된 영어 시험 문항의 개선이 요구되며 이에 대한 연구가 필요하다.

이러한 점을 고려하여, 이 연구는 대학수학능력시험(영어시험)의 문항 개선에 대한 연구로서 문항 질문과 답지를 영어로 제시하는 새로운 모형을 개발하고 그 효과를 검증함을 목적으로 한다. 시험 문항의 분석에서 가장 중요한 요소인 문항 변별도나 문항 난이도를 통계적으로 분석하고 검증하고 대학수학능력시험(영어시험)의 바람직한 방향을 제시함을 목적으로 한다.

2. 연구의 문제

이 연구는 다음 세 가지의 문제를 다룬다.

1) 실험집단(문항을 영어로 질문하고 답지도 영어로 제시하는 경우)과 통제 집단(문항을 우리말로 질문하고 답지도 우리말로 제시하는 경우) 사이에 시험 결과는 차이가 있는가?
2) 실험집단과 통제집단 사이에 문항 변별도는 차이가 있는가?
3) 실험집단과 통제집단의 정답율은 차이가 있는가?

II. 이론적 배경

이 장에서는 이 연구에 대한 이론적 배경으로 언어능력 평가 이론의 틀 안에서 대학수학능력시험(영어시험)에 대한 전반적인 검토를 한다. 우선 언어 능력 평가 이론에 대한 검토로서 시험, 평가, 교수에 대한 개념과 정의를 살펴

교보문고 KYOBO Book Centre

본다. 언어능력 평가 이론을 기초로 대학수학능력시험(영어시험)의 성격과 방향, 평가 목표와 내용을 다룬다.

1. 시험, 평가, 교수의 개념과 정의

평가는 교육과정의 한 영역으로 교육 현장에서 아주 중요한 의미를 지닌다. Brown(2004)은 시험(test)을 "a method of measuring a person's knowledge or performance in a given domain"(p. 3)으로 정의하고 있다. 이 정의에 따르면 시험은 수험자에게 특정한 영역의 지식이나 수행을 요구하는 기법, 절차, 문항으로 하나의 '방법'(method)을 지칭한다.

이완기(2003)는 "평가 분야에서는 시험, 테스트, 검사 등 여러 용어가 평가라는 말과 비슷하게 쓰인다. 영어에서도 테스트(test), 테스팅(testing), 이그재미네이션(examination), 메저먼트(measurement), 어세스먼트(assessment), 이밸류에이션(evaluation) 등의 용어들이 명확한 구분 없이 혼용되는 경우가 많다"(p. 22)고 지적하고 있다. 용어의 불분명함이나 혼용이 시험의 특성에 따라 달리 사용되고 있음에도 불구하고 우리말 용어가 합당하지 않아 필자는 'testing'과 'assessment'를 '평가'라는 같은 개념으로 사용한다. Hughes (2003, p. 5)는 이 두 용어의 차이점을 "It (Testing) is just one form of assessment."라고 제시하고 있다. 따라서 assessment 가 testing 보다 더 넓은 개념으로 사용된다. Brown(2004)은 평가가 시험보다 훨씬 넓은 영역에서 사용되며 일회성이 아니라 계속해서 평가가 계속되는 속성을 지니고 있음을 다음과 같이 제시하고 있다(p. 4).

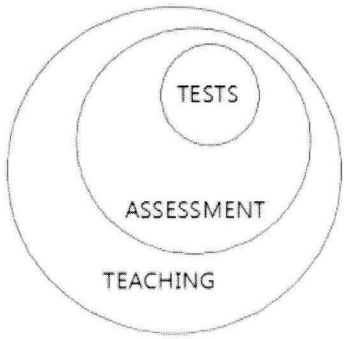
> Assessment, on the other hand, is an ongoing process that encompasses a much wider domain .... Tests, then, are a subset of assessment; they are certainly not the only form of assessment that a teacher can make. Tests can be useful devices, but they are only one among many procedures and tasks that teachers can ultimately use to assess students(Brown, 2004, p. 4).

Brown(2004)은 교육 내용을 교수, 학습하는 가운데 평가가 지속적으로 이루어질 수 있으며 언어 평가에서 시험(test), 평가(testing, assessment), 교수(teaching)와의 관계를 그림 1 과 같이 제시하고 있다

교보문고 KYOBO Book Centre

그림 1

Tests, assessment, and teaching (Brown, 2004, p. 5)

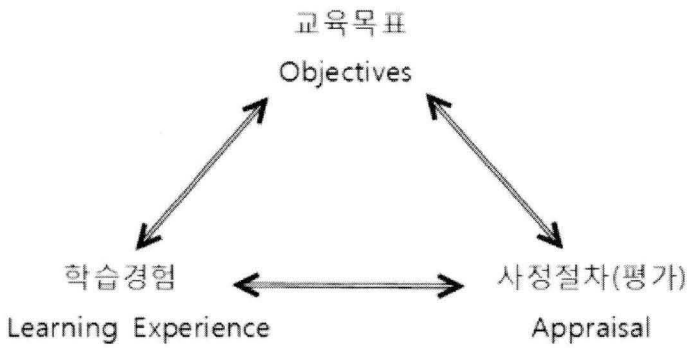


Tyler(1950)의 교육과정 모형은 다음 그림 2 에 제시된 바와 같다. 다시 말하면, 이 모형은 교육과정에서 교육목표와 학습경험과 평가의 유기적인 관계를 자세히 보여주고 있다.

그림 2

교육과정의 형태

(Tyler, 1950; 남명호, 김성숙, 지은림, 2000, p. 10 재인용)



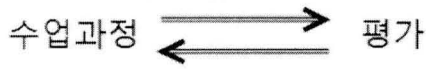
황정규(1998, p. 56)는 Tyler(1950)의 모형에 대하여 교육과정을 "첫째, 교육목표의 선정 및 세분화, 둘째, 학습경험의 선정과 조직, 셋째, 교수-학습 과정의 투입, 넷째 교육평가의 과정"으로 설명하고 있다. Brown(2004)이나 Tyler(1950)의 모형은 교수-학습과 평가는 아주 밀접한 상호 유기적인 관계에 있다는 점을 강조한다. 임병빈(2005, pp. 29-30)은 "바람직한 'backwash'는 평가와 교수 과정이 상호간에 영향을 미쳐서 평가가 교수

교보문고 KYOBO Book Centre

학습 과정에 도움을 줄 수 있다"고 언급하면서 평가와 교수 과정 사이의 상호 의존 관계를 그림 3과 같이 제시하고 있다.

그림 3

backwash의 유형 (임병빈, 2005, p. 30)



Bachman 과 Palmer(1996, p. 30)는 역류 효과를 "수험자 개개인에 대한 평가의 직접적인 영향"(direct impact of testing on individuals)이라고 말하고 있다. 평가는 교육 목표, 교육 내용, 교수-학습 방법에 지대한 영향을 미치고 있다(Hughes, 2003). 따라서 실제 교육 현장에서 바람직한 교육이 이루어질 수 있도록 신뢰성, 타당성, 실용성이 있는 평가가 무엇보다도 중요하다.

2. 대학수학능력 시험의 성격과 방향

대학수학능력시험은 1993 년도에 처음 실시된 이후 지금까지 대학입학 시험으로 과거에 실시해오던 학력고사와는 전혀 다른 시험의 형태와 내용을 포함하고 있다. 교육과정평가원(2005)은 대학수학능력시험의 외국어(영어) 영역에 대해 "의사소통능력의 달성 정도를 중점적으로 측정하는 시험"(p. 183)으로 정의하고 있다. 이 시험은 고등학교 교육과정의 틀 안에서 언어의 네 가지 기능 -듣기, 말하기, 읽기, 쓰기- 의 균형적인 교육과 더불어 언어의 실제적인 사용을 중시하는 의사소통 능력을 측정할 수 있는 바탕을 마련하고 있다. 이는 그 동안 소홀히 다루어왔던 음성 언어의 중요성과 지나치게 문법 중심의 암기식 교육의 배제가 그 중심이 되었으며 통합교과적인 내용을 소재로 학생들의 고등 사고능력을 측정하기 위하여 고안된 시험이다.

김임득(1995)은 대학수학능력시험(영어시험)에 대해 "고등학교 교육과정(교육부 고시 1992-19 호)에서 규정하고 있는 영어 사용 능력을 측정하는 다양한 문항을 측정함으로써 고등학교 영어교육이 정상화되도록 하고, 대학 수학에 필요한 독해를 비롯한 여러 언어 기능에 걸쳐 의사소통 능력을 함양하는 영어 교수법을 유도할 수 있는 문항이 출제되도록 하고 있다"(p. 52)고 제시하였다. 이는 교육과정의 내용과 수준을 출제의 원칙으로 하고 있어 영어교육 현장에서 바람직한 영어교육이 이루어질 수 있다는 점을 시사하고 있다. 특히, 대학수학능력시험(영어시험)이 영어교육 현장에서 역류 효과로서 긍정적인 효과를 가져왔다는 공감대가 꾸준히 있어왔다(김임득, 1995; 이재희, 2001; 전병만, 2004).

교보문고 KYOBO Book Centre

대학수학능력시험(영어시험)에 있어서 2001 년 이후 듣기 문항은 2011 년 현재까지 17 개 문항이 출제되고, 읽기 문항은 5 개가 줄어든 33 개의 문항으로 총 50 개의 문항이 출제되었다. 대학수학능력시험의 성격과 출제 방향은 도입기 때보다 더 구체적으로 제시되어 있으며 그 내용은 다음과 같다.

1) 대학 수학에 필요한 영어사용능력의 기본이 되어 있는 가를 측정한다.
2) 고등학교 영어교육의 목표를 달성한 정도를 측정하고, 고등학교 영어과 교육 과정이 제시하고 있는 영어사용능력을 함양시킬 수 있는 수업이 되도록 출제한다.
3) 독해력에 많은 비중을 두고, 듣기는 직접 방법으로, 말하기와 쓰기는 간접 방법으로 측정한다.
4) 숙어를 포함한 어휘력은 문장이나 단락 속에서 의미를 파악하는 능력을 측정하고, 문법성 판단력은 복잡한 구조를 이해하고 그 속에 담긴 정보를 이해하는 능력을 측정하는 수준으로 출제한다.
5) 독해는 단락 속에 담긴 내용을 사실적 이해, 추론적 이해, 감상이나 판단과 같은 종합적 이해, 적용의 능력을 중심으로 출제한다.
6) 청해는 대화나 담화 속에 담긴 내용을 사실적 이해, 추론적 이해, 종합적 이해를 하였는지를 측정하는 수준으로 출제한다.
7) 말하기와 쓰기는 간접적으로 출제하되 의사소통능력을 측정하는 데 필요한 수준으로 출제한다.

(교육과정평가원, 2005, p. 205)

듣기 시험의 경우 1993 년도 시험에서는 듣기 문항이 8 문항으로 시작하여 최근에는 17 문항으로 증가되었으며, 읽기 문항은 일 지문 일 문항 원칙에서 벗어나 2005 학년도 영어시험에서는 300 단어 이상이 되는 한 지문에 3 개의 문항이 출제되어 짧은 지문보다는 긴 지문을 활용하는 것이 바람직한 현상으로 자리하게 되었다.

3. 대학수학능력시험 평가 목표와 내용

대학수학능력시험은 넓은 의미로 두 가지의 평가 목적을 가지고 있다. 첫째는 대학 수학에 필요한 영어사용 능력을 평가하며, 둘째로 고등학교 영어과 교육과정에 제시된 영어교육 목표의 달성 정도를 평가한다. 고등학교 영어과 교육과정에 기초한 일반적인 평가 목표는 교육과정평가원(2002)에 나타난 바와 같다.

첫째, 외국어(영어)영역의 시험은 일상생활에서 영어를 의사소통의 도구로 활용할 수 있는 능력을 측정한다.

교보문고 KYOBO Book Centre

둘째, 일상생활과 일반적인 화제에 관한 말이나 글의 의미를 이해하고 자연스럽게 표현하는 능력을 측정한다.
셋째, 영어를 통하여 다양한 정보를 받아들이고 활용하는 능력을 측정한다.
넷째, 외국문화를 이해하고 바르게 수용하며, 쉬운 말이나 글로 우리의 문화를 외국인에게 소개하는 능력을 측정한다.
다섯째, 대학 수학 능력에 필요한 학문적인 영어사용 능력을 측정한다.
여섯째, 실용적인 글을 이해하고 활용하는 능력을 측정한다.

(교육과정평가원, 2002, p. 12)

일반적인 평가 목표에 따른 구체적인 평가 내용은 고등학교 영어과 교육과정의 성취 기준에 나타나 있다.

김임득(1995)은 평가 내용에 대해 음성 언어(듣기와 말하기)와 문자 언어(읽기와 쓰기)의 평가 내용으로서 기초 언어능력과 언어사용 능력으로 구분하여 구체적인 예를 제시하고 있다. 음성 언어의 경우, 기초 언어능력으로 "(1) 문장 재성 능력 (2) 어휘-의사소통 상황에서 빈번하게 쓰이는 쉬운 어휘 (3) 음운론 규칙, 담화 규칙 (4) 회화체 상용 표현법(언어기능, 상황별 상용 표현법)" 등의 내용을 포함하고 있다(김임득, 1995, p. 53). 김임득(1995)은 언어사용 능력으로 "음성으로 전달되는 정보를 듣고 그 내용을 사실적으로 이해하거나 추론하고, 종합적으로 판단할 수 있는 청해력을 측정하고 있다"(p. 53)고 제시하고 있다.

문자 언어의 경우, 기초 언어 능력으로 어휘의 의미 파악, 문장 구조, 담화 구조, 사회언어학적 지식 등의 내용을 제시하고 있다. 언어사용 능력으로 "인쇄 매체로 전달되는 정보를 얻는 목적과 종류에 따라 scanning, skimming, extensive reading, intensive reading 을 하면서 얻은 핵심적 내용을 사실적으로 이해하느냐, 아니면 추론하여 연역적으로 생각을 하여 결론을 예측할 것인가, 또는 필자가 주는 정보의 내용을 학습자의 인지능력을 근거로 하여 종합적으로 판단하고 감상하는 능력을 측정하고 있다" (김임득, 1995, p. 60). 다시 말하면, 글의 사실적 이해, 추론적 이해, 종합적 이해와 적용을 중심으로 하고 있다.

대학수학능력시험에 있어서 2004 년도에 실시한 2005 대학수학능력시험은 이전의 시험과 달리 출제범위가 7 차 교육과정이 적용되는 시기로 시험의 범위가 공통영어 수준에서 영어 1(2,300 단어 수준)의 수준과 내용을 평가 목표로 하는 변화가 이루어졌다. 또한 읽기 시험에서 일 지문 다 문항의 문항이 등장하게 되었다. 2005 학년도 대학수학능력시험 이후 현재까지 시험의 형태나 내용은 큰 변동 없이 유지되고 있다.

대학수학능력시험의 영어시험이 포함하고 있는 내용은 다음 표 1 의 외국어 (영어) 영역 이원분류표에 나타나 있다.

교보문고 KYOBO Book Centre

표 1

외국어(영어)영역 이원분류표

행동영역 / 내용영역		어휘	문법성판단력	이해			적용	문항수	비율(%)
				사실적	추론적	종합적			
듣기	대화								
	담화								
말하기	대화								
	담화								
쓰기	문장								
	문단								
읽기	논설문								
	설명문								
	문학								
	실용문								
	기타								
문항수								50	
비율(100%)									100

표 1 에 제시된 대학수학능력시험 외국어(영어) 영역 이원분류표는 언어 평가에 있어서 내용 영역과 행동 영역을 구분하여 측정함을 보여준다. 내용 영역으로는 언어의 네 가지 기능 -듣기, 말하기, 읽기, 쓰기- 을 포함하며 듣기와 말하기 기능의 음성 언어는 대화와 담화로 구분하여 평가하고 있다. 쓰기 기능은 문장과 문단으로 나누며, 읽기는 논설문, 설명문, 문학, 실용문, 기타의 글의 내용을 중심으로 평가한다.

행동 영역에 포함되어야 할 사항은 어휘, 문법성 판단력, 사실적 이해, 추론적 이해, 종합적 이해, 적용이며, 이 요소가 음성 언어와 문자 언어에 적용된다. 음성 언어나 문자 언어에 대한 평가는 기초 언어능력을 토대로 하고 있다. 기초 언어능력은 어휘 능력, 문장의 구조, 담화 구조, 사회언어학적 지식을 포함하고 있다. 이를 바탕으로 언어 사용 기술로서 사실적 이해, 추론적 이해, 종합적 이해를 중점적으로 평가한다(김임득, 1995; 교육과정평가원, 2005).

교보문고 KYOBO Book Centre

III. 연구 방법

이 장에서는 이 연구의 가설을 설정하고 가설을 검증하기 위한 실험 대상자 선정, 실험을 위한 자료(평가 문항) 개발, 개발된 시험의 실시, 그리고 자료 분석 방법에 대해 기술한다.

1. 가설의 설정

연구자는 서론의 2 에서 제기된 문제점을 토대로 하여 다음 세 가지의 가설을 설정하였다.

가설 1: 실험집단(문항을 영어로 질문하고 답지도 영어로 제시하는 경우)과 통제집단(문항을 우리말로 질문하고 답지도 우리말로 제시하는 경우) 사이에 시험 결과는 차이가 없다.
가설 2: 실험집단이 통제집단보다 문항 변별도가 높을 것이다.
가설 3: 실험집단이 통제집단보다 정답율이 낮을 것이다.

2. 실험 대상자

이 연구를 위해 전라북도 도청 소재지인 전주시에 위치한 인문계 S 고등학교에 재학 중인 3 학년 남녀 학생 288 명을 선정하였으며, 이들을 통제집단(143 명)과 실험집단(145 명)으로 나누었다. 연구 대상자들은 대부분 대학 진학을 원하는 학생들로 2011 년 11 월에 실시될 대학수학능력시험을 준비하고 있기 때문에 이 연구에 가장 적합한 실험 대상이 되었다.

실험집단과 통제집단은 2011 년 6 월에 전국적으로 실시한 2011 년 대학수학능력 모의시험의 결과를 토대로 동질집단이 되게 하였다. 이 시험한 참여한 대상자는 실험집단(152 명)과 통제집단(151 명)으로 총 303 명이었다. 이들 중에서 본 실험에 결시한 학생 15 명을 제외한 288 명이 본 연구의 실험 대상자로 최종 선정되었다(표 2 참조).

표 2
실험 대상자의 구분

집단 / 성별	실험집단	통제집단
남	71	71
여	74	72
소계	145	143
총계	288	

(단위: 명)

교보문고 KYOBO Book Centre

실험 대상자의 선정은 문과 4 개반과 이과 4 개반, 남학생 4 개반과 여학생 4 개반의 요소를 고려하여 표집이 균형을 이룰 수 있도록 하였다. 다시 말하면, 각 집단에는 문과 2 개반(남여 1 개반 포함)과 이과 2 개반(남여 1 개반 포함)으로 실험 대상자가 골고루 분포되게 하였다.

실험대상자의 동질성 검증에 있어서 실험집단과 통제집단은 2011 년 6 월 2 일에 교육과정평가원이 주관하여 실시한 대학수학능력시험 모의시험의 결과를 토대로 동질성 t-검증을 하였다. 두 집단의 시험 결과는 다음 표 3 에 나타나 있다.

표 3

실험 대상자의 동질성 검증

항목 / 집단	대상자수	평균	분산	공동분산	t-값	p
실험집단	152	59.96	380.80	385.79	0.19	0.84
통제집단	151	60.39	390.81			

($p > 0.05$)

표 3 에서 보는 바와 같이 두 집단의 평균이 아주 비슷하며 t-값이 0.19, p 값이 0.84 로 유의 수준 0.05 보다 크므로 귀무가설을 채택한다. 따라서 두 집단 간에는 차이가 없어 두 집단이 동질집단임이 증명되었다.

3. 실험 자료 개발

통제집단과 실험집단에 사용할 평가 내용이 같은 두 가지 다른 형태의 시험 문항을 개발하였다. 통제집단에게는 현재 시행하고 있는 시험의 형태로 문항의 질문과 답지가 우리말로 된, 실험집단에게는 새로 개발된 질문과 답지가 모두 영어로 된 시험 문항을 개발하였다.

두 집단에 사용되는 문항은 듣기 12 문항, 읽기 13 문항으로 총 25 문항을 사용하였다. 본 연구의 문항 수가 실제 대학입학시험의 문항 수보다 적은 이유로는 읽기 시험에 있어서 빈칸 채우기나 어법 등에 관한 내용은 영어로 답지를 제시해야 하며, 우리말로 답지를 제시할 수 없는 제약이 있기 때문이다. 두 집단에 적용한 문항은 전성애(2012)의 논문 부록 1 과 부록 2 에 제시되어 있다. 다음 예 1 과 예 2 는 사용된 듣기 문항과 읽기 문항의 대표적인 예이다.

교보문고 KYOBO Book Centre

예 1
두 집단의 듣기 문항 비교

<실험집단에 사용된 문항>

7. Listen to the dialog and choose what the man is going to do next.

① cleaning the kitchen
② doing the laundry
③ going to his friend's house
④ giving his friend a call
⑤ mowing the lawn

<통제집단에 사용된 문항>

7. 대화를 듣고, 남자가 다음에 할 일로 가장 적절한 것을 고르시오.

① 부엌 청소하기
② 빨래하기
③ 친구 집에 가기
④ 친구에게 전화하기
⑤ 잔디 깎기

W: Thomas, have you finished your chores?
M: Almost. I cleaned the kitchen and did the laundry.
W: Did you mow the lawn, too?
M: No, not yet.
W: When are you planning on doing that?
M: I'll just do it tomorrow. I'm going over to Brian's house right now.
W: I don't think so. The last time you waited until Sunday, you never finished your chores.
M: I'll definitely do it tomorrow. I promise!
W: No. You can't go out until you finish your work.
M: But Brian is waiting for me.
W: Well, you'd better call and let him know you can't make it.
M: All right. I guess I can see him tomorrow

예 2
두 집단의 읽기 문항 비교

<실험집단에 사용된 문항>

20. Which one best describes the main point in this passage?

교보문고 KYOBO Book Centre

My father took me for a walk after my baseball game. I expected him to congratulate me for winning, but he didn't seem cheerful. "You and your team won played well today. Normally, that is something people would respect, but not with your team's behavior after the game." I thought back to what happened after the game, when a few players on my team had pointed and laughed at the other team, and even I had participated in chanting "Fifteen to zero!"—the final score of the game—over and over. "We were just excited, Dad," I tried to explain. "I know you were, but that's not an excuse to be disrespectful. A good competition is based on mutual respect, and if you don't give that to your opponents, you won't receive it from others."

① It is more important to act on your will than on that of others.
② A honest loss is worthier than a shameful victory.
③ It is necessary to respect others in order to be respected*
④ Doing your best is more important than the outcome.
⑤ It is not helpful for you to focus on the past.

<통제집단에 사용된 문항>
20. 다음 글의 요지로 가장 적절한 것을 고르시오.
① 다수의 행동에 휩쓸리지 않는 주체성이 필요하다.
② 비겁한 승리보다는 정정당당한 패배가 더 가치 있다.
③ 존중 받기 위해서는 남을 먼저 존중해야 한다.
④ 최선을 다했다면 결과는 그다지 중요하지 않다.
⑤ 이미 지난 과거 일에 연연하는 것은 좋지 않다.

연구자는 우리나라 ○○출판사로부터 영어 지문의 사용 허가를 받아 평가 문항의 기초 자료를 마련하고, 대학수학능력시험에서 여러 차례 출제 위원으로 참여한 교수의 도움을 받아 시험 문항을 개발하였다. 또한 듣기 시험의 경우는 실제 대학수학능력시험과 매우 유사한 환경(원어민 녹음, 발화 속도, 문항간 휴지 등)을 구현하도록 노력하였다. 읽기의 경우도 두 집단에 문항당 1.5 분이 소요되도록 대학수학능력시험과 거의 같은 시간을 배정하였다.

4. 시험 실시

사전시험으로 한국교육과정평가원의 주관 하에 2011 년 6 월 2 일에 실시된 대학수학능력 모의시험의 자료를 활용하였다. 이 연구를 위해 개발된 시험은 실험 대상자에게 2011 년 8 월 20 일 약 한 시간에 걸쳐 실시하였다.

교보문고 KYOBO Book Centre

두 집단의 모든 학생이 동시에 시험에 참여하였다. 듣기는 안내 방송을 포함하여 약 14 분, 읽기는 19 분이 소요되도록 하였다. 이 시험은 시험의 공정성을 위해 감독관의 엄격한 관리 하에 실시하였다.

5. 시험 결과 처리

시험의 채점은 답안지 OMR 카드를 활용하여 컴퓨터 상에서 자동 처리하도록 하였으며 Excel 통계 프로그램을 활용하여 두 집단의 성적 결과, t-검증, 문항 난이도, 문항 변별도, 문항난이도로서 답지 반응율 등을 분석하였다.

Ⅳ. 결과 분석 및 논의

제 4 장에서는 서론에서 제기된 문제점을 중심으로 실험집단과 통제집단의 시험 결과를 통계적으로 분석하며, 대학수학능력시험(영어시험)의 문제점을 밝히고, 이에 대한 개선 방법을 논의한다.

1. 가설 1 의 검증과 논의

가설 1: 실험집단과 통제집단 사이에 시험 결과는 차이가 없다.

가설 1 의 검증을 위해 같은 내용의 시험을 질문과 답지가 다른 두 가지 유형으로 나누어 각각 실험집단과 통제집단에 실시하였다. 한 유형은 문항을 영어로 질문하고 답지도 영어로 제시하는 시험으로 실험집단에 적용하였고, 다른 또 하나의 유형은 문항을 우리말로 질문하고 답지도 우리말로 제시하는 시험으로 통제집단에 적용하였다. 두 집단의 시험 결과는 다음 표 4 에 나타나 있다.

표 4

실험집단과 통제집단의 시험 결과

항목 / 집단	대상자	문항수	평균	표준편차	표준오차	범위
실험집단	145	25	14.22	4.92	0.40	20 (23-3)
통제집단	143	25	16.52	4.73	0.39	19 (24-5)

교보문고 KYOBO Book Centre

표 4 에서 보는 바와 같이 통제 집단이 실험 집단보다 평균이 2.3 점 높게 나타나고 있다. 이 점수를 100 점으로 환산하면 9.2 점으로 상당한 차이가 나타나고 있다. 이는 통제집단의 경우, 영어를 모르더라도 우리말을 통해서 답의 단서를 유추할 수 있는 가능성이 있는 것으로 나타나고 있다. 또한 두 집단의 점수 범위는 실험집단(최소 3 문항에서 최대 23 문항)과 통제집단(최소 5 문항에서 최대 24 문항) 사이에 거의 차이가 없지만 실험집단에 적용한 문항이 어렵다는 사실을 알 수 있다.

가설 1 의 검증은 t-검증으로 하였으며 그 결과는 다음 표 5 에 나타나 있다.

표 5

실험집단과 통제집단의 동질성 검증

항목 / 집단	대상자수	평균	분산	공동분산	t-값	p
실험집단	145	14.22	24.24	23.35	4.04	***
통제집단	143	16.52	22.46			

($p < 0.001$)

표 5 에서 나타난 바와 같이 t-검증 결과, t-값이 4.04 로 $p<0.001$ 이므로 귀무가설을 기각한다. 따라서 두 집단은 차이가 있다는 것이 증명되었다. 표 4 와 표 5 의 시험 결과를 종합하면, 두 집단 사이에는 통계적으로 유의한 수준으로 평균 점수의 차이를 보임을 알 수 있다.

2. 가설 2 의 검증과 논의

가설 2: 실험집단이 통제집단보다 문항 변별도가 높을 것이다.

문항 분석의 중요한 분석 방법으로서 문항 난이도(item difficulty)와 문항 변별도(item discrimination)가 사용된다. 문항 난이도는 보통 정답율로 나타내며 문항 변별도는 각 문항에 있어서 상위 집단(상위 25%)과 하위 집단(하위 25%)의 능력을 변별해내는 정도를 나타내는 지수이다.

가설 2 에 대한 검증으로 다음 표 6 은 두 집단의 문항 변별도와 문항 정답율을 나타내고 있다.

교보문고 KYOBO Book Centre

표 6

두 집단의 평균 문항 변별도와 정답율

항목 / 집단	문항 변별도	정답율(%)
실험집단	0.52	56.88
통제집단	0.46	66.09

표 6 에서 보는 바와 같이, 실험집단이 통제 집단보다 변별도가 높으며 정답율에 있어서는 실험집단이 통제집단보다 낮게 나타나고 있다. 이는 질문과 답지가 우리말로 제시되기 때문에 수험자들에게 영어로 질문하고 답지도 영어로 되어 있는 시험보다 더 쉬운 시험으로 정답율이 높게 나타나고 있다. 특히, 실험집단에 적용한 영어 질문과 영어 답지로 되어 있는 시험이 문항 변별도에 있어서 더 좋은 결과를 갖게 되어 가설 2 가 증명되었다.

교육과정평가원의 주관 하에 시행된 최근 5 년간(2005-2009 학년도)의 대학수학능력시험의 영어시험의 변별도는 평균 0.44 로 나타나고 있다(김용명, 이완기, 김진석, 고현숙, 김상재, 2010). 그러나 이 연구에서 시험 문항의 개선의 일환으로 개발하여 현장 실험집단에 적용한 시험 모형의 결과는 변별도가 평균 0.52 로 높게 나타나고 있다. 따라서 '각 문항에 있어서 영어로 질문하고 답지도 영어로 제시하는' 문항으로 개선할 필요가 있다는 근거를 제시할 수 있다.

이 연구에 사용된 25 문항(듣기 12 문항과 읽기 13 문항)을 각 문항별 문항 변별도, 문항 정답율, 문항 답지 반응에 대한 결과는 실험집단의 경우는 표 7, 통제집단의 경우는 표 8 에 나타나 있다.

표 7

각 문항의 변별도, 정답율, 문항 반응(실험집단)

문항 번호	변별도	정답율 (%)	답지반응(%)				
			1 번	2 번	3 번	4 번	5 번
1	0.156	18.6	9.7	1.4	14.5	18.6	55.9
2	0.533	43.4	43.4	13.8	13.8	8.3	20.7
3	0.656	70.3	8.3	1.4	70.3	15.2	4.8
4	0.344	56.6	0.7	7.6	32.4	56.6	2.8
5	0.672	55.2	9.7	55.2	3.4	6.9	24.8
6	0.633	71.7	2.8	71.7	22.8	2.1	0.7
7	0.456	73.1	2.1	5.5	11.0	73.1	8.3

교보문고 KYOBO Book Centre

8	0.097	20.0	7.6	11.7	42.1	18.6	20.0
9	0.341	43.4	4.1	43.4	24.8	13.1	14.5
10	0.679	55.2	55.2	8.3	20.0	6.2	10.3
11	0.749	71.7	4.1	6.9	11.0	71.7	6.2
12	0.672	55.9	6.9	14.5	12.4	10.3	55.9
13	0.482	80.0	80.0	3.4	10.3	2.8	3.4
14	0.649	74.5	7.6	74.5	13.1	3.4	1.4
15	0.400	89.7	0.7	1.4	89.7	6.9	1.4
16	0.513	55.9	12.4	10.3	12.4	55.9	9.0
17	0.449	77.9	1.4	2.1	15.2	3.4	77.9
18	0.646	47.6	10.3	12.4	15.9	47.6	13.8
19	0.036	10.3	64.8	10.3	9.0	3.4	12.4
20	0.690	66.9	14.5	6.2	66.9	8.3	4.1
21	0.700	71.7	4.8	11.7	10.3	71.7	1.4
22	0.700	66.9	5.5	7.6	66.9	13.1	6.9
23	0.500	39.3	11.0	39.3	14.5	15.9	19.3
24	0.779	43.4	43.4	11.0	20.0	12.4	13.1
25	0.700	62.8	0.7	7.6	17.9	11.0	62.8

표 8

각 문항의 변별도, 정답율, 문항 반응(통제집단)

문항 번호	변별도	정답율 (%)	답지반응(%)				
			1 번	2 번	3 번	4 번	5 번
1	-0.028	49.0	9.1	0.0	11.2	49.0	30.8
2	0.660	55.9	55.9	7.0	13.3	16.1	7.7
3	0.493	69.2	13.3	4.9	69.2	9.1	3.5
4	0.597	59.4	2.8	4.2	31.5	59.4	2.1
5	0.625	65.7	8.4	65.7	4.2	4.9	16.8
6	0.264	84.6	1.4	84.6	11.9	2.1	0.0
7	0.438	65.0	5.6	7.0	4.9	65.0	17.5
8	0.535	33.6	9.1	11.9	29.4	16.1	33.6
9	0.549	46.9	3.5	46.9	23.1	21.7	4.9
10	0.632	54.5	54.5	1.4	14.7	6.3	23.1
11	0.340	87.4	4.2	3.5	2.8	87.4	2.1
12	0.396	84.6	0.0	4.9	4.9	5.6	84.6
13	0.368	79.7	79.7	3.5	14.0	0.7	2.1

교보문고 KYOBO Book Centre

14	0.222	92.3	0.7	92.3	4.2	1.4	1.4
15	0.313	87.4	3.5	4.2	87.4	3.5	1.4
16	0.458	58.7	20.3	1.4	13.3	58.7	6.3
17	0.674	70.6	1.4	4.2	21.7	2.1	70.6
18	0.743	60.1	7.0	9.8	14.0	60.1	9.1
19	-0.125	11.2	80.4	11.2	2.1	2.8	3.5
20	0.569	79.0	1.4	7.7	79.0	9.8	2.1
21	0.396	79.7	4.9	8.4	2.8	79.7	4.2
22	0.667	65.0	15.4	6.3	65.0	9.1	4.2
23	0.507	75.5	7.0	75.5	4.2	7.0	6.3
24	0.674	62.2	62.2	4.9	10.5	9.1	13.3
25	0.611	74.8	2.8	9.8	7.7	4.9	74.8

표 7 의 실험집단의 문항 분석에서 발견되는 특이한 예는 문항 변별도가 0.30 이하의 문항이 3 문항으로 1 번(0.16), 8 번(0.10), 19 번(0.04)이 이에 해당한다. 이는 세 문항의 정답율이 각각 18.6%, 20.0%, 10.3%로 문항이 어려운 경우이다. 문항이 아주 어려워 실험 대상자들이 무리하게 추측하여 답지를 골랐다는 뜻이다. 이와 대조적으로, 문항이 아주 쉬워 변별도에 영향을 미치는 예는 표 8 의 통제집단의 문항 분석표에 나타나 있다. 6 번(0.26)과 14 번(0.22)의 경우 문항 변별도는 0.30 이하이며, 정답율이 각각 84.6%와 92.3%로 문항이 쉬운 경우이다. 문항 변별도가 낮은 문항은 문항 개선을 위해 답지 반응 등을 다시 검토할 필요가 있다.

답지 반응은 수험자의 답지에 대한 반응을 나타내며 각 문항에 있어서 매력적인 적절한 오답지가 되었는지에 대한 검토의 실마리를 제공해 준다. 표 7 에 나타난 1 번과 12 번의 문항에서 답지 반응을 하지 않은 경우가 1 번 문항 오답지 2 번과 12 번 문항 오답지 1 번에 나타나 있다. 이 시험에 있어서 수험자가 특정한 오답지에 반응하지 않음으로써 이 시험은 결국 5 지선다형 시험이라기보다 4 지선다형 시험이라고 말할 수 있다. 이러한 오답지는 다시 수정하고 보완하여 새로운 매력적인 오답지가 되도록 해야 한다. 따라서 문항 반응에 대한 검토는 문항 개선을 위해 반드시 필요하다.

대학수학능력시험의 영어시험은 언어의 네 가지 기능(듣기, 말하기, 읽기, 쓰기)을 측정하고 있지만 엄밀한 의미에서 본다면 음성 언어로서 듣기와 문자 언어로서 읽기 기능의 측정으로 이루어져 있다. 이 연구에서 듣기와 읽기 기능의 변별도는 다음 표 9 에 나타나 있다.

교보문고 KYOBO Book Centre

표 9

듣기 문항과 읽기 문항의 변별도(평균)

영역 / 집단	듣기	읽기
실험집단	0.50	0.56
통제집단	0.45	0.47

표 9 에서 보는 바와 같이 실험집단이 통제집단보다 듣기와 읽기의 두 기능 모두에서 문항 변별도가 높음을 나타나고 있다. 또한 읽기가 듣기보다 더 변별력이 높게 나타나고 있다. 따라서 영어시험에서 질문과 답지가 영어로 되어 있는 문항이 문항 변별도가 높아 문항 개선의 방법으로서 이의 활용이 필요하다.

3. 가설 3의 검증

가설 3: 실험집단이 통제집단보다 정답율이 낮을 것이다.

가설 3 의 검증은 문항 정답율의 분석에 의하여 이루어진다. 두 집단에 나타나는 문항 정답율은 다음 표 10에 나타나 있는 바와 같다.

표 10

두 집단의 평균 문항 정답율(평균)

항목 / 집단	정답율 (%)
실험집단	56.88
통제집단	66.09

표 10 에서 보는 바와 같이 실험집단의 정답율은 56.88 로 통제집단의 정답율 66.09 보다 정답율에 있어서 훨씬 낮다. 이는 영어 질문과 영어 답지로 구성된 시험 문항이 우리말로 제시되는 문항보다 훨씬 어렵다는 것을 의미하고 있다. 질문과 답지가 우리말로 제시되는 경우, 영어를 모르더라도 우리말에 의해 문제를 해결할 수 있는 단서를 제공하기 때문이다. 또한 각 항별 정답율은 표 7 과 표 8 에 나타나 있으며 이에 대한 논의는 IV-2 에 제시하였다.

김용명, 이완기, 김진석, 고현숙, 김상재(2010)는 다음 표 11 에서 처럼 주요 국가의 국가 주도형 영어 시험의 양상 분석을 제시하고 있다. 표 11 을 통해서 보는 바와 같이 질문과 선택지(답지)가 모국어(한국어)로 제시되어 있는 경우는 우리나라 밖에 없다.

교보문고 KYOBO Book Centre

표 11

주요 국가의 국가 주도형 영어 시험의 시험지 양상 분석

(김용명, 이완기, 김진석, 고현숙, 김상재, 2010, p. 33)

양상 / 국가	지시문	질문	선택지	비고
한국	한국어	한국어	한국어/영어	▪지시문, 질문 한국어 ▪선택지 한국어·영어 혼용
중국	중국어	영어	영어	
일본	일본어	영어	영어	
독일	-	영어	-	▪주관식 3문항
러시아	러시아어	영어	영어	
인도네시아	영어	영어	영어	

현재 사용되고 있는 대학수학능력시험(영어시험) 문항은 질문과 답지가 우리말로 되어 있어 안면 타당도(face validity)의 결여가 지적된다. 또한 질문과 답지를 우리말로 제공함으로써 문제를 해결하는데 단서를 제공할 뿐만 아니라 문항 자체가 쉬어지는 경향을 갖게 될 수 있다.

이상에서 논의한 바와 같이 대학수학능력시험의 영어시험은 분리평가에 입각한 객관식 선다형 시험으로 이해 기능을 적절히 평가할 수 있으나 표현 기능을 직접적으로 평가할 수 없는 단점을 지니고 있다. 이 연구에서 문항 변별도와 문항 난이도를 통하여 문항 분석을 하였으며 이를 토대로 가설을 검증하였다. 따라서 문항 개선의 방법으로 질문과 답지를 영어로 제시하는 시험 문항으로의 전환이 필요하다.

V. 결론 및 제언

1. 결론

우리나라에서 대학입학시험으로서 대학수학능력시험은 국가 단위로 시행 되는 고등학교 공교육의 마지막 단계의 시험으로 그 중요성은 매우 크다. 시험은 교과의 목표, 내용, 교수.학습 방법에 지대한 영향을 주는 역류 효과 가 있기 때문에 공교육의 바람직한 교수.학습을 위해 평가 방법이 타당해야 한다.

이 연구는 대학수학능력시험(영어시험)의 문항 개선을 위한 실험연구로서 시험 문항의 형태를 질문과 답지의 제시 방법의 개선을 통하여 올바른 영어 교육이 교육 현장에 이루어질 수 있는 바탕을 마련하는데 그 목적이 있다. 다시 말하면, 새로운 영어시험 문항으로서 '질문과 답지를 영어로 제시하는 문항'을 개발하여 그 효과를 검증함을 목적으로 하고 있다.

교보문고 KYOBO Book Centre

이러한 목적을 달성하기 위해 새로운 평가 문항 모형을 개발하였으며, 이 모형을 고등학교 3 학년 남녀 학생 총 288 명을 대상으로 실험집단과 통제집단으로 나누어 실험하였다. 새로 개발된 모형은 실험집단에 기존에 사용된 모형은 통제집단에 적용하여 세 가지의 가설을 검증하였다.

가설의 검증 결과, 실험집단과 통제집단 사이에는 평균 점수, 변별도 및 정답율에 있어서 큰 차이가 있는 것으로 드러났다. 이는 질문과 답지를 영어로 사용하는 시험이 기존의 우리말 질문과 우리말 답지를 사용하는 시험보다 더 타당한 것으로 확인되었다. 또한 이 실험의 결과 분석과 논의를 통하여 실험집단에 적용한 새로 개발된 모형은 다음 세 가지의 긍정적인 면을 지니고 있어 문항 개선의 여지가 있음을 보여주고 있다. 첫째, 문항 변별도가 높다. 둘째, 영어시험다운 안면 타당도를 가진다. 셋째, 영어 사용 능력의 활성화로 인해 긍정적인 역류 효과를 낼 수 있다. 이상에서 살펴본 바와 같이 대학수학능력시험의 영어시험 문항은 반드시 필요하며 V-2 에서 밝힌 바와 같이 문항 개선을 위한 후속 연구가 필요하다.

2. 제언

대학수학능력시험(영어시험)은 국가가 주도하는 중요한 시험으로 영어교육에 지대한 영향을 미치는 것이 사실이다. 영어시험의 타당성에 대한 연구나 문항 개선에 대한 지속적인 연구가 이루어져야 한다. 후속 연구로서 다음 세 가지를 제안한다.

첫째, 이 연구는 대학수학능력시험(영어시험)의 문항 개선을 위한 연구로서 이 연구의 제한점에서 밝힌 바와 같이 특정한 지역, 적은 수의 대상자로 실험하였기 때문에 실험 대상자들의 특성(상급 집단, 중급 집단, 하급 집단)이나 지역(대도시, 중도시, 소도시) 등의 요소를 고려하여 다각적으로 연구할 필요가 있다.

둘째, 이 연구에서 다루지 않은 문항 지시문을 영어로 제시하거나 듣기 문항에 있어서 질문이나 답지를 문자로 제시하지 않고 음성으로 제시하는 연구가 필요하다.

셋째, 올바른 영어교육을 위해 이 시험의 단점으로 지적되는 분리 시험에서 탈피하여 통합 시험으로서의 전환이 요구된다. 특히, 표현 기능의 평가는 분리식 선다형 시험으로는 직접적인 측정을 할 수 없다는 단점이 있다. 현재 교육과정평가원이 개발하고 있는 국가공인영어시험이 대안으로 제시될 수 있다. 그러나 국가공인영어시험이 대학 선발을 위한 시험으로 시행된다면 여러 가지 문제점을 일으킬 소지가 있다. 국가공인영어시험이 신뢰성, 타당성, 실용성을 가진 시험으로 정착하여 우리나라 영어교육의 정상화가 이루어질 수 있도록 다각적인 연구가 지속되어야 한다.

교보문고 KYOBO Book Centre

참고문헌

교육부. (1992). 제 6 차 고등학교 교육과정. 교육부 고시 제 1992-19. 307-363.
교육부. (1997). 제 7 차 영어과 교육과정. 서울: 교육부.
교육인적자원부. (2008). 교육인적자원부 고시 제 2006-75 호 및 제 2007-79 호에 따른 중학교 교육과정 해설(V) 외국어(영어). 서울: 대한교과서 주식회사.
교육과학기술부. (2009). 교육과학기술부 고시 제 2008-160 호에 따른 초등학교 교육과정 해설. 서울: 한솔기획.
교육과학기술부. (2011). 영어과 교육과정. 교육과학기술부 고시 2011-361 호.
한국교육과정평가원. (2002). 영어 수준별 평가의 한국적용 가능성. 연구자료 RRC 2002-11.
한국교육과정평가원. (2002). 2005 학년도 대학수학능력시험 세부 시행 방안 세미나. 연구자료 ORM 2002-13.
한국교육과정평가원. (2005). *대학수학능력시험 10 년사 I.* 서울: 한국교육과정평가원.
한국교육과정평가원. (2005). *대학수학능력시험 10 년사 II.* 서울: 한국교육과정평가원.
한국교육과정평가원. (2006). 영어교육 혁신방안 공개 세미나 자료집. 연구자료 ORM 2006-23.
김용명, 이완기, 김진석, 고현숙, 김상재. (2010). 수능 외국어(영어)영역 개선 연구. 한국교육과정평가원. 대수능 CAT 2010-11.
김임득. (1995). 대학수학능력 시험과 영어교육의 방향. *Foreign Languages Education, Vol. 1*, 39-70.
남명호, 김성숙, 지은림. (2000). *수행평가 -이해와 적용.* 서울: 문음사.
우종옥 외. (1993). 대학수학능력시험의 영역별 출제모형정립 및 모형활용 방안연구. 한국교육평가원.
이영식. (1997). 최근 언어평가의 연구와 이론적 배경. *영어교육, 52(1)*, 221-246.
이완기. (2003). *영어 평가 방법론.* 서울: 문진미디어.
이재희. (2001) 대학수학능력시험 영어 영역의 총체적 개선 방안. *영어교육, 56(2)*, 333-364.
임병빈. (2005). *영어교육 평가 방법.* 서울: 경문사.
전병만. (1985). *언어능력평가.* 서울: 한신문화사.
_____. (1988). New Directions of Language Testing. *영어영문학, 제 11 집.* 영어영문학회 전북지부, 한국영어영문학회, 237-272.
_____. (2004). 한국, 중국, 일본의 대학입학시험 문항 비교 연구. *영어어문교육, 10(2)*, 231-256.

교보문고 KYOBO Book Centre

전성애. (2012). 대학수학능력시험(영어시험)의 문항개선에 대한 연구. 전북대학교 대학원 석사학위 논문.

황정규. (1998). *학교학습과 교육평가(개정판)*. 서울: 교육과학사.

Bachman, L. F., and Palmer, A. S. (1996). *Language Testing in Practice*. Oxford: Oxford University Press.

Brown, H. Douglas. (1980). *Principles of Language Learning and Teaching*. (2th Ed.). Englegood Cliffs, N.J.: Prentice-Hall.

_____. (2004). *Language Assessment: Principles and Classroom Practices*. Pearson Education, Inc.

Corder, S. (1973). *Introducing Applied Linguistics*. Harmondsworth, UK: Penguin Books.

Harris, David P. (1969). *Testing English as a second Language*. New York: McGraw-Hill.

Hughes, Arthur. (2003). *Testing for Language Teachers*. (2nd Ed.). Cambridge: Cambridge University Press.

Lado, R. (1961). Language Testing. New York: McGraw-Hill.

Oller, John W. Jr. (1979). *Language Tests at School*. London: Longman.

Valette, Rebecca. (1977). *Modern Laguage Testing*. (2nd Ed.). New York: Harcourt, Brace, and Jovanovich.

Weir, C. J. (1990). *Communicative Language Testing*. New York: Prentice-Hall.

예시언어(Examples in): English
적용가능언어(Applicable Languages): English
적용가능수준(Applicable Levels): Secondary

전성애
전북대학교 언어교육원
(561-756) 전북 전주시 덕진구 덕진동 1가 664-14
Tel: 063-270-2250
E-mail: sungaejeon@gmail.com

Received in March, 2012
Reviewed in May, 2012
Revised version received in June, 2012

교보문고 KYOBO Book Centre