

## 대학영어 프로그램을 위한 배치고사 연구: 수능과 TOEIC의 비교를 중심으로

조수경\*, 박찬호†

승실대학교, 한국교육과정평가원

**Cho, Sookyung & Park, Chanho. (2011). Investigations of the placement test for a college English program: With a comparison of CSAT and TOEIC. *Modern English Education*, 12(3), 368-383.**

This study investigates which of two tests—a college entrance exam or an English proficiency test—better places college freshmen into the courses appropriate for their levels. The participants are 2,474 freshmen, enrolled in a 4-year university in Seoul. Most students in the college English program, in which students were placed in courses according to their English subtest scores on the College Scholastic Ability Test (CSAT), also took the Test of English for International Communication (TOEIC). First, students who would have been placed into a lower-level course if the TOEIC scores had been used as a criterion for placement were identified. Later, discovering that a large portion of the students had been "misplaced", the lecturers of the college English program were interviewed about the levels of their students. The lecturers generally stated that the students should have been placed into a lower-level course. It is thus argued that TOEIC, the English proficiency test, would be a better measure for placement purposes. However, the results should be used with a caution because some lecturers did not find that the students with lower TOEIC scores were problematic in their classes. (188)

[placement tests/college English program/TOEIC/CSAT/  
배치고사/대학영어프로그램/토익/수능]

### I. 서론

현재 많은 대학에서 영어교육을 모든 재학생들에게로 확대 실시하고 있는 추세이다. 글로벌 시대에 맞는 인재로 재학생들을 키워내기 위한 일환으로 명실공히 세계 공용어로 자리잡은 영어로 자신의 의견을 말이나 글로 표현해

---

\* 제1저자

† 교신저자

내는 능력을 키워내기 위함이다(성명희, 현태덕, 이지연, 장형지, 2011). 그러나 영어교육의 전면적인 확대가 실효성을 거두기 위해서는 학생들의 수준에 맞는 영어교육이 필수적이라고 하겠다. 다시 말해서 학생들의 현재 영어 수준을 고려하여 이에 알맞은 교육을 받도록 배려해야 한다는 것이다.

학생들의 수준에 맞는 교육을 하기 위해서는 우선 학생들을 적절한 수준에 맞도록 배치할 필요가 있다. Linn과 Gronlund(2000)는 배치평가(placement assessment)를 교수학습이 시작되기 전 학생들의 출발행동을 분석하는 평가로 정의하였다. Bachman(1990)이나 Brown(2005)은 학생들을 그들에게 적합한 수준의 반에 배치하는 것이 얼마나 중요한지를 역설한 바 있다. 수준이 너무 높은 반으로 가면 수업을 따라가는 데 어려움이 있고 반대로 너무 낮은 반으로 갈 경우에는 수업에 흥미를 잃고 학습 능력이 저하될 수 있다는 것이다. 이러한 문제를 최소화하기 위해서 반편성에 앞서 배치고사(placement test)를 실행할 필요성이 있게 된다. Wall, Clapham과 Alderson(1994)과 Fulcher(1997)에 따르면, 배치고사의 궁극적 목적은 어학 수준이 낮아서 학교 생활을 하는 데 어려움이 있거나 실패할 확률이 높은 학생들을 알아내어 그들을 돕고자 하는 데에 있기 때문이다.

그러나 안타깝게도 이 부분에 있어서 현재 한국 대학의 상황은 그리 낙관적이지 않다. Kyung-Ah Yu, Chae-Young Moon, Sang-Keun Shin(2007)의 연구에 의하면 총 다섯 개의 대학과 네 개의 대학 연계 어학원을 조사한 결과 단 두 개의 대학과 네 개의 어학원에서만 배치고사를 실시하고 있는 것으로 나타났으며 이들 대부분에서조차 수업 목표와 배치고사 내용이 일치하는 경우는 드물었다. 교수의 답변을 분석한 결과도 이와 별반 다르지 않음을 알 수 있다. 53%의 교수가 동일하지 않은 집단을 가르친 것으로 답했으며 33% 이상이 자신들이 가르친 반에 수준에 맞지 않는 학생이 있었다고 답했던 것이다. 이러한 결과에 기반하여 한국 대학의 영어 프로그램에서 배치고사를 좀 더 적극적으로 활용하는 동시에 수업의 내용과 동일한 종류의 지식이나 능력을 측정하는 방향으로 나아갈 것을 권장했다.

그러나 배치고사의 중요성을 인지하더라도 배치고사를 실행하기에 앞서 여러 가지 현실적인 제반 상황과의 타협을 피하기 힘들다. 예를 들어, Wesche, Paribakht와 Ready(1993)는 학생들의 수준별 분반이 이루어지지 않더라도 다른 종류의 시험과 달리 배치고사는 이에 대한 후속 조치가 가능하기 때문에 정확하진 않더라도 초기에 어느 정도의 분류라도 이루어진다면 다행이라고 한 바 있다. 따라서 영어 프로그램의 운영자나 대학의 입장에서는 자신들의 영어 프로그램에 가장 적합한 배치고사가 무엇인지(“What is the best possible test (method) for this program?” p. 3) 보다는 기존의 시험들 중 어느 것이 배치고사로서 자신들의 프로그램에 가장 알맞은가(“Which of the possible tests (methods) is best for this program?” p. 3)를 고민하게 된다고 했다. 본 연구의 대상이 되는 서울에 위치한 4년제 대학의 경우도 바로 Wesche 등이 지적한

경우에 해당된다. 학생들에게는 수준별 수업을 제공하고 교수에게는 원활한 수업 진행을 위해서 일찍부터 배치고사의 중요성을 인식해 왔지만 프로그램의 내용과 부합하는 배치고사를 자체적으로 제작할 만한 여건은 되기 않기에 현재 이용 가능한 시험들 중 수능의 외국어 영역 점수를 이용하여 분반을 하고 있는 실정이다. 3월 개강 며칠 전이나 최종 신입생 명단이 확인되는 현재 상황에서 3,000여 명에 가까운 신입생들을 3월 1일 개강에 맞춰 분반하는 것은 사실상 불가능 하기 때문이다. 현재 개강 전의 이 짧은 기간 동안 신입생들에게 온라인으로 토익을 보도록 유도하고 있긴 하지만 그 참여율이 약 85% 정도에 미치고 있어 배치고사로서의 활용은 아직 미흡한 단계이다. 따라서 본 연구에서는 한 4년제 대학의 실례를 통하여, 전체 신입생을 대상으로 한 두 시험 중 어느 시험이 해당 대학의 영어 프로그램에 가장 적합한 배치고사인지를 알아보기 위하여 수능과 토익의 기능을 비교·분석하겠다.

## II. 이론적 배경

배치고사에 대한 그 간의 연구는 특정 시험에 대한 타당도(validity) 검사로 이루어져 왔다. 다시 말해서 배치고사가 실제로 검사하고자 하는 바를 검사하고 있는가에 관심을 가져왔다. Wall 등(1994)은 배치고사의 타당도를 검사할 때에 다음과 같은 네 가지 질문을 염두에 두어야 한다고 했다.

1. 배치고사가 영어를 가장 필요로 하고 이러한 능력을 가르치는 수업을 필요로 하는 학생들을 정확하게 분별해내는가?
2. 배치고사를 보는 학생들은 자신들의 언어 능력이 정확하게 측정된다고 생각하는가?
3. 배치고사의 내용이 그 용도에 부합하는가?
4. 배치고사가 신뢰할 만한가? (p.321)

이러한 질문에 대답하기 위하여 대부분의 연구들은 여러 가지 타당도 검사를 수행하였다. 예를 들어, Wall 등(1994)은 네 가지 타당도 검사를 통하여 Lancaster 대학에서 실시하고 있는 배치고사가 의도한 대로 기능하고 있는가를 살펴보았다. 학생들이 시험에 대해서 어떻게 생각하고 있는가를 물어봄으로써 액면 타당도(face validity)를, 배치고사의 내용과 프로그램의 내용을 살펴봄으로써 내용 타당도(content validity)를, 시험의 각 영역이 실제로 서로 다른 능력을 검사하고 있느냐에 관한 구인 타당도(construct validity), 마지막으로 배치고사와 다른 시험과의 연관성을 보는 동시 타당도(concurrent validity)를 조사했다. Fulcher(1997) 역시 Surrey 대학의 배치고사 타당도를 각 시험 영역의 상관계수와 토플과의 동시 타당도, 수업 내용과의 내용 타당도 등을 통해

측정하였다.

그러나 최근의 추세는 특정 시험에 대한 단순한 타당도 검증을 넘어서 배치고사가 학업 성취도를 얼마나 정확하게 예측할 수 있느냐에 초점이 맞추어지고 있다. 예를 들어, Young-Ju Lee 와 Greene(2007)은 양적 연구와 더불어 질적 연구를 혼합한 혼합 방식(mixed methods)을 이용하여 외국인 대학원생을 대상으로 영어 배치고사가 그 학생들의 향후 대학원에서의 학업 성취도를 예측하는 데 도움이 되는가를 살펴보았다. 학생들의 배치고사 점수와 그들의 학점(GPA), 자기 평가, 그리고 그들에 대한 교수의 평가를 비교해 본 결과, 학업 성취도는 복잡한 요소를 포괄하고 있기 때문에 단순히 영어 배치고사 점수만으로 이를 예측하는 것이 힘들다는 것을 보여주었다.

한국에서도 학생들의 대학 수학능력을 측정하겠다는 목표로 1994 학년도부터 수능 시험이 실시된 이후, 많은 연구자들이 수능 시험이 대학에서의 학업 성취도를 얼마나 잘 예측하는지를 연구해왔다. 비교적 복합적인 결론을 제시한 연구로는 김현철(2004)의 연구가 있다. 1999 학년도에 서울에 소재한 한 대학에 입학한 학생들을 성별에 따라 나눈 후 4 년 간 얻은 평점평균에 대한 수능 성적과 고교 내신 성적의 예측력을 비교한 결과 성별에 따라 차이가 있음을 관찰하였다. 즉, 여학생의 경우 내신 성적이 수능 성적보다 대학에서의 학업 성취도에 더 큰 영향을 미쳤지만 남학생의 경우 반대로 내신 성적의 영향력이 더 컸다. 성별과 같은 학생들의 특성에 따라 수능 성적과 고교 내신 성적이 대학에서의 학업 성취도에 미치는 영향이 달라질 수 있다는 것이다.

반면, 대부분의 연구는 수능 성적보다는 내신 성적의 예측력이 더 좋다는 일관된 결론을 내리고 있다. 김상호(1995)는 한 대학의 신입생들을 대상으로 수능 성적, 고교 내신 성적, 그리고 대학에서의 학업 성취도 지표로서 첫 학기 평점평균을 수집하여 그 변인들 간 상관관계를 분석하였다. 학생들이 속한 계열에 따라 결과에 차이가 있기는 하였지만 대체로 고교 내신 성적이 수능 점수보다 대학에서의 학업 성취도와와의 상관계수가 더 높았다. 수능 성적과 대학 학업 성취도 간 상관 계수는 비록 통계적으로 유의미하기는 하였으나 .04 정도의 매우 미미한 수준이었다. 윤명희, 안도희(2001)도 이와 비슷한 결론을 내렸다. 윤명희, 안도희는 1995 학년도부터 5 개 학년도에 걸쳐 4 학년 성적까지 포함한 대학 학업 성취도 결과와 수능 및 고교 내신 성적을 이용하여 단순 상관분석과 중다회귀분석을 실시하였다. 이 연구의 결과에서도 역시 고교 내신 성적이 수능 성적보다 대학에서의 학업 성취도를 일관되게 더 잘 예측하는 것으로 나타났다. 이와 유사한 결과는 이후의 여러 연구에서도 지속적으로 보고되고 있다(박성수, 2007; 부재율, 2007; 신정철, 정지선, 신택수, 2008).

이와 같이 수능과 대학에서의 학업 성취도의 관계는 그리 높지 않음이 밝혀져 있는 현재, 수능의 외국어 영역 점수를 영어회화 프로그램에 대한 배치고사로 활용하고 있는 서울에 위치한 한 4 년제 대학의 상황을 진단해 보고자 하는 노력에서 본 연구가 비롯되었다. 본 연구는 특히 앞서 설명한

것처럼, 대안으로 제시되고 있는 토익 시험을 배치고사로 활용했을 때의 가상 분석을 통하여 두 시험을 비교, 대조하여 다음과 같은 질문에 답하고자 한다.

1. 수능과 토익이 레벨에 맞는 적당한 수업에 학생들을 배치하고 있는가?
2. 서로 다른 배치 고사 시험 방식이 실제 수업에 어떤 영향을 미치는가?

### III. 연구 방법

#### 1. 연구 대상

영어회화 수업은 본 연구의 대상이 되는 대학에 입학한 신입생 모두가 들어야 하는 교양 필수 과목이다. 이 대학에 입학한 총 3 천 여명의 학생들은 전공에 따라 9개의 군(A, B, C, D, E, F, G, H, I)으로 나뉘어, 약 160 여 개의 반에 배치된다. 분반 배치 기준은 수능의 외국어 영역 점수로서 점수에 따라 같은 군에 속한 학생들을 22 명 혹은 23 명씩 한 반으로 배정한다. 대상은 2011학년도 1학기 총 3,077명의 신입생 중, 수능과 모의 토익 점수가 둘 다 있는 2,474명으로 한정하였다.

#### 2. 연구 도구 및 절차

수능은 2010년도 11월 12일에 치러진 수능의 외국어(영어) 영역 점수를 사용하였다. 토익의 경우는 3월 개강 직전에 온라인으로 치러진 모의 토익 점수를 사용하였다. 이 대학에 입학이 확정된 총 3,077명의 신입생들에게 배치고사 목적으로 3월 개강 직전에 모의 토익을 보도록 권하였으나, 실제로는 80%에 해당하는 2,474명 만이 토익 시험에 응시하였다. 정식 토익은 청취 100문제와 독해 100문제로 구성되어 총 120분이 소요되나, 이 대학에서 사용한 모의 토익의 경우는 청취 47문제, 독해 48문제로 총 한 시간이 소요되었다.

시간상의 제약과 응시율 저조로 신입생들은 결국 수능의 외국어(영어) 영역 성적에 따라 영어회화 수업에 분반 배치되었다. 모든 영어회화 수업은 동일한 커리큘럼을 따르고 있으며, 학생들의 자신감을 북돋아 영어로 효과적인 의사소통을 할 수 있도록 돕는다는 목표를 내세우고 있다. 특히 언어의 네 영역 중 표현기능에 해당하는 말하기와 글쓰기에 중점을 두고 있다. 학생들은 16주 동안 대화에 대한 기초적인 요소로 시작하여, 가족이나 친구에 대해 이야기하기, 직업이나 공부, 오락 및 여가 활용, 길 물어보기, 미래의 계획 등등과 같은 다양한 주제에 대해 말하고 쓰는 것을 연습했다.

### 3. 분석 방식

본 연구는 양적 연구를 주로 활용하였으며 이에 대한 결과를 설문조사를 통하여 보완하였다. Young-Ju Lee와 Greene(2007)은 배치고사와 학업 성취도와의 관계와 같이 복잡한 현상을 연구할 경우 다양한 방법을 혼합한 방식의 이점에 대해서 역설한 바 있다. 다양한 방식을 통합적으로 사용할 경우, 하나의 결과를 다른 결과와 대조함으로써 추론에 대한 타당도를 높이며 또한 어떤 결과를 사용하여 다른 결과를 설명함으로써 가장 완벽하면서도 균형 잡힌 유용한 결과물을 나올 수 있다고 했다(p. 386). 어느 분석 방법도 그 자체로 완벽할 수 없기에 본 연구에서는 이러한 혼합 연구를 활용하여 양적 연구 결과와 설문 조사 결과를 서로 보완하고자 하였다.

우선, 일차적으로 수능과 토익의 배치고사로서의 기능을 살펴보기 위하여 신입생들의 수능과 모의 토익 점수를 이용한 양적 분석을 시행하였다. 수능 점수<sup>1</sup>를 이용한 분반 시와 모의 토익 점수<sup>2</sup>를 이용한 분반 시를 비교하여 오배치된 학생들(misplaced students)을 찾아내어 보았다. 본 연구에서는 토익 성적을 기준으로 할 때의 분반이 수능 성적을 기준으로 할 때의 분반과 달라지는 학생들을 오배치로 지칭한다. 구체적으로는 모의 토익 성적을 이용하여 같은 신입생들을 각 그룹별로 가상 분반을 하였다. 이 때 그룹별로 각 반의 커트라인을 찾아내어 실제 분반에서 이 커트라인보다 낮은 학생들을 찾은 뒤, 수능 점수를 기준으로 분반하였을 때와 비교하였다. 예를 들어, A 그룹의 경우 영문학과나 일문학과 등과 같은 각종 인문학 계열의 학생들이 소속되어 있는데, 실제로는 수능의 외국어(영어) 영역 점수에 따라 총 16개의 분반으로 나뉘었다. 이 중 가장 높은 수준인 분반 1의 경우 총 23명의 학생이 속해 있지만, 이 중 19명만이 모의 토익을 치렀다. 모의 토익 성적을 가지고 재분반을 할 경우 같은 인문학 계열 학생 그룹 중 상위 19명만을 분반 1로 재분반을 했다. 이 때 이 가상 분반의 커트라인은 토익 860점에 해당했는데, 실제로 분반 1의 학생 중 이 커트라인보다 낮은 학생들은 오배치된 것으로 간주하였다. 각 그룹별, 분반별 학생들의 오배치 여부를 계산하고 학생들의 수능 및 토익 성적이 오배치에 미치는 영향을 파악하기 위해 그룹별로 로지스틱 회귀분석을 실시하였다. 양적 분석에는 컴퓨터 언어와 통계 패키지의 기능을 모두 갖춘 R 시스템(R Development Core Team, 2010)을 이용하였다.

또한 수능과 모의 토익 점수 비교를 통해 오배치된 학생들이 실제로 오배치되었는가를 확인하기 위하여 2차적으로 설문 연구를 시행하였다. 총 128개의 분반을 가르친 24명의 원어민 교수에게 이메일로 설문 조사를

<sup>1</sup> 수능 성적으로는 외국어(영어) 영역의 백분위 점수를 이용하였다.

<sup>2</sup> 토익 성적은 청해(listening comprehension)와 독해(reading comprehension) 성적을 합한 총점을 이용하였다.

진행하였으며 설문 조사는 크게 두 가지 질문으로 구성되었다(부록 참조). 첫째, 오배치된 것으로 확인된 학생들 중 각 분반별로 커트라인보다 낮은 학생들 2~3명을 추려 학기가 끝난 뒤 각 교수에게 이 학생들에 대한 영어 구사 능력에 대한 평가를 수집하였다. 이 때, 분반의 커트라인과 비교하여 성적 차이가 지나치게 큰 학생의 경우는 문의 대상 학생에서 제외하였다. 앞서 예로 들었던 그룹 A, 분반 1의 경우 모의 토익 점수로 가상 분반 시 커트라인이 860점이었는데 실제 분반 1에 배치된 두 학생의 경우 모의 토익에서 65점과 30점을 받았다. 이들의 경우 수능의 외국어 영역 성적이 각각 99, 98 백분위수로 최상위권이었는데 토익 점수는 지나치게 낮아 모의 토익 점수가 이들의 실제 영어 구사 능력을 제대로 반영하지 못했을 가능성이 크다. 따라서 이러한 이상값(outlier)을 제외하고 그 외에 분반의 커트라인과 비교적 많은 차이를 보이는 학생들을 골라서 교수의 의견을 물었다.

다음으로는 분반 전체에 대한 원어민 교수들의 의견을 알아보았다. 특히 오배치된 학생 수와 그들의 모의 토익 점수를 기준으로 영어 구사 능력 면에서 동질성(homogeneity)이 뛰어난 경우와 그렇지 않은 경우(heterogeneity)를 선별하여 교수들의 개인적인 의견을 물어보았다. 동일성이 뛰어난 경우는 오배치된 학생 수가 소수에 그치면서(3~4명) 동시에 그들의 모의 토익 점수가 반 평균과 크게 차이를 보이지 않는 경우를 선별했다. 그 반대의 경우에 대해서는, 오배치된 학생의 수가 반 전체 인원의 1/2 이상을 넘는 동시에 그들의 모의 토익 점수가 반 평균에 비해 확연히 낮은 경우로 선별했다.

## IV. 결과

### 1. 수능 및 토익 성적과 오배치 비율

표 1은 전체 학생들의 수능 백분위 성적과 토익 점수에 대한 셋째와 넷째 중심 적률(central moments) 정보인 왜도(skewness)와 첨도(kurtosis)를 보여준다.

표 1  
학생들의 수능 및 토익 성적의 왜도(skewness)와 첨도(kurtosis)

구분	왜도	첨도
수능 백분위	-2.00	7.63
토익 총점	-0.26	2.47

표 1에서 보듯이 수능 백분위 성적은 왜도가 -2.00, 첨도가 7.63이므로 심한 부적 편포를 이루고 있음을 알 수 있다. 반면, 토익 성적은 왜도와 첨도가 각각

-0.26, 2.47이어서 정규분포와 가까운 형태를 보이고 있다. 즉, 수능 영어 성적을 기준으로 할 때에는 학생들 대부분이 상위 10% 안에 몰려있는 분포를 보이지만 토익 성적은 상대적으로 골고루 퍼져있음을 알 수 있다.

또한 앞에서 설명한 바와 같이 모의 토익 점수를 기준으로 한 가상 분반을 만들고 이를 수능 기준의 분반과 비교했을 때, 절반에 가까운 대상이 오배치된 것으로 나타났다. 조사 대상 총 2,474명 중 40.99%에 해당하는 1,014명이 수능 점수와 토익 점수 사이에 차이가 있는 것으로 조사되었다(표 2 참조). 특히 각 그룹별로 상위 레벨일수록 오배치된 학생 수가 하위 레벨에 비해 더 많은 것으로 나타났다. 그룹별로 반의 수와 학생의 수가 다르기 때문에 그룹 내 상위 1/3까지의 학생들만으로 구성된 반을 상, 그 다음 상위 2/3까지의 학생들만으로 구성된 반을 중, 나머지를 하로 하여 상, 중 하 각각에 속한 반의 오배치 비율 평균을 계산한 결과 D그룹에서 상에 속하는 반의 오배치 비율 평균이 73.18로 가장 높았고, H그룹이 가장 낮은 53.47이었다. 반면, D그룹에서 하에 속하는 반의 오배치 비율 평균이 24.08로 가장 높았고, F그룹이 가장 낮은 10.52였다. 중에 속하는 반의 오배치 비율 평균은 36.62부터 49.62의 범위에 있었다. 즉 모든 그룹에서 상위 1/3 이내에 속하는 반의 오배치 비율이 50% 이상으로 가장 높았고 중에 속하는 반의 오배치 비율 평균이 그 다음이었으며 하가 평균적으로 가장 낮았다. 이는 하위 레벨의 경우, 대부분이 모의 토익 컷트 라인이 100점 대나 그 이하였으므로 이 보다 더 낮은 반에 배치되었어야 하는 경우가 극히 드물거나 전무하기 때문이다.

표 2

그룹별 분반별 오배치된 학생 수와 학급 인원 대비 비율

A 그룹				B 그룹				C 그룹			
반	오 배치	학급 인원	비율 (%)	반	오 배치	학급 인원	비율 (%)	반	오 배치	학급 인원	비율 (%)
1	8	19	42.1	1	18	20	90.0	1	12	19	63.2
2	16	22	72.7	2	17	22	77.3	2	14	19	73.7
3	16	23	69.6	3	13	20	65.0	3	12	20	60.0
4	14	23	60.9	4	10	22	45.5	4	14	20	70.0
5	13	22	59.1	5	16	22	72.7	5	14	22	63.6
6	10	21	47.6	6	8	21	38.1	6	10	20	50.0
7	14	22	63.6	7	8	22	36.4	7	11	21	52.4
8	7	20	35.0	8	8	21	38.1	8	11	17	64.7
9	5	21	23.8	9	5	20	25.0	9	4	23	17.4
10	5	22	22.7	10	10	22	45.5	10	7	18	38.9



11	3	22	13.6	11	6	21	28.6	11	3	23	13.0
12	7	23	30.4	12	4	19	21.1	12	4	23	17.4
13	4	19	21.1	13	7	22	31.8	13	7	22	31.8
14	4	18	22.2	14	4	21	19.0	14	2	14	14.3
15	1	13	7.7	15	3	23	13.0	15	0	13	0.0
16	0	5	0.0	16	0	8	0.0				
계	127	315	40.3	계	137	326	42.0	계	125	294	42.5
D 그룹				E 그룹				F 그룹			
반	오 배치	학급 인원	비율 (%)	반	오 배치	학급 인원	비율 (%)	반	오 배치	학급 인원	비율 (%)
1	12	14	85.7	1	13	21	61.9	1	12	19	63.2
2	13	18	72.2	2	17	18	94.4	2	13	15	86.7
3	13	14	92.9	3	13	20	65.0	3	10	21	47.6
4	9	16	56.3	4	15	19	78.9	4	10	18	55.6
5	10	17	58.8	5	11	19	57.9	5	9	20	45.0
6	5	12	41.7	6	10	16	62.5	6	10	23	43.5
7	12	20	60.0	7	7	19	36.8	7	8	19	42.1
8	4	14	28.6	8	9	23	39.1	8	6	21	28.6
9	8	16	50.0	9	9	21	42.9	9	3	21	14.3
10	1	16	6.3	10	6	20	30.0	10	1	15	6.7
11	5	19	26.3	11	6	19	31.6	11	6	19	31.6
12	4	17	23.5	12	5	21	23.8	12	0	16	0.0
13	6	16	37.5	13	5	21	23.8	13	0	9	0.0
14	6	14	42.9	14	3	20	15.0				
15	2	14	14.3	15	0	15	0.0				
16	0	13	0.0	16	0	3	0.0				
계	110	250	44.0	계	129	295	43.7	계	88	236	37.3
G 그룹				H 그룹				I 그룹			
반	오 배치	학급 인원	비율 (%)	반	오 배치	학급 인원	비율 (%)	반	오 배치	학급 인원	비율 (%)
1	12	21	57.1	1	11	20	55.0	1	13	20	65.0
2	21	25	84.0	2	10	23	43.5	2	12	17	70.6

3	11	21	52.4	3	13	21	61.9	3	16	19	84.2
4	11	22	50.0	4	13	22	59.1	4	13	19	68.4
5	12	22	54.5	5	7	19	36.8	5	11	21	52.4
6	10	21	47.6	6	9	21	42.9	6	8	21	38.1
7	8	24	33.3	7	9	19	47.4	7	7	20	35.0
8	9	20	45.0	8	3	19	15.8	8	6	19	31.6
9	7	23	30.4	9	6	21	28.6	9	8	25	32.0
10	6	25	24.0	10	4	21	19.0	10	4	21	19.0
11	3	20	15.0	11	0	11	0.0	11	2	21	9.5
12	0	10	0.0					12	3	13	23.1
								13	0	15	0.0
계	110	254	43.3	계	85	217	39.2	계	103	251	41.0

## 2. 수능 및 토익 성적에 따른 오배치

학생들의 분반은 그룹별로 수능 성적을 바탕으로 이루어졌기 때문에 수능 성적이 높을수록 오배치될 확률이 높음은 이미 표 1의 결과를 보아 알 수 있다. 이러한 경향이 통계적으로 유의미한지, 특히 토익 성적을 통제된 상태에서도 유의미한지 파악하기 위하여 토익 성적과 수능 성적을 독립변인, 오배치 여부(0: 오배치 아님, 1: 오배치)를 종속변인으로 하는 로지스틱 회귀분석을 실시하였으며, 그 결과는 표 3과 같다.

표 3  
그룹별 로지스틱 회귀분석 결과

그룹		B 추정값	표준오차	Wald	유의확률	Exp(B)
A	Intercept	-6.66	1.60	17.28	0.00	0.00
	TOEIC	-0.02	0.00	73.21	0.00	0.98
	CSAT	0.22	0.03	54.11	0.00	1.24
B	Intercept	-11.38	1.89	36.26	0.00	0.00
	TOEIC	-0.03	0.00	61.40	0.00	0.97
	CSAT	0.33	0.04	59.45	0.00	1.39
C	Intercept	-5.94	1.25	22.52	0.00	0.00
	TOEIC	-0.02	0.00	68.99	0.00	0.98
	CSAT	0.20	0.03	61.07	0.00	1.22

D	Intercept	-0.18	0.72	0.06	0.80	0.83
	TOEIC	-0.02	0.00	56.95	0.00	0.98
	CSAT	0.14	0.02	56.27	0.00	1.16
E	Intercept	-29.86	4.81	38.58	0.00	0.00
	TOEIC	-0.03	0.00	57.12	0.00	0.97
	CSAT	0.54	0.08	50.14	0.00	1.71
F	Intercept	-11.84	2.83	17.53	0.00	0.00
	TOEIC	-0.01	0.00	48.77	0.00	0.99
	CSAT	0.23	0.04	32.71	0.00	1.26
G	Intercept	-14.88	2.95	25.44	0.00	0.00
	TOEIC	-0.03	0.00	55.68	0.00	0.97
	CSAT	0.35	0.05	42.73	0.00	1.43
H	Intercept	-10.80	2.29	22.22	0.00	0.00
	TOEIC	-0.02	0.00	49.99	0.00	0.98
	CSAT	0.28	0.04	41.18	0.00	1.32
I	Intercept	-6.24	1.68	13.74	0.00	0.00
	TOEIC	-0.02	0.00	50.96	0.00	0.98
	CSAT	0.20	0.03	43.59	0.00	1.22

표에 나타난 결과를 보면, 결과 수능 성적이 높을수록 오배치될 확률은 유의미하게( $p < .01$ ) 높았다. 즉, 수능 외국어(영어) 영역 표준점수의 백분율이 1 올라갈수록 오배치될 오즈(odds)는 0.14(그룹 D)에서 0.54(그룹 E)만큼 올라갔다. 반면 토익 성적이 1점 올라갈수록 오배치될 오즈는 0.01(그룹 F)에서 0.03(그룹 B, E, G)만큼 낮아져 토익 성적이 높을수록 오배치될 확률이 유의미하게( $p < .01$ ) 낮아지는 결과를 보였다. 오배치의 기준이 토익 성적이었으므로 이러한 경향성은 어느 정도 예측 가능한 것이지만 이것이 통계적으로 유의미하며, 특히 서로의 성적을 통제한 상태에서도 유의미한 결과를 보였다는 점에서 흥미로운 결과라 할 수 있다. 이러한 로지스틱 회귀분석의 결과는 모든 그룹에서 비슷한 양상을 보였다.

### 3. 설문 조사

수능이 아닌 토익 성적을 기준으로 할 경우 상당한 수의 학생들이 오배치되었으며 수능 성적이 높을수록 오배치될 확률이 높은 반면 토익 성적이 높을수록 오배치될 확률은 낮아진다는 양적 분석의 결과를 얻었다. 이를 토대로 원어민 교수들에게 설문 조사를 실시한 결과, 교수들도 상당수의

학생들이 오배치되었음을 인정하였다. 설문 조사를 했던 총 24명의 교수 중 15명이 오배치된 학생들에 대한 개인적인 의견을 보내왔다. 설문에서 개인적인 의견을 물었던 총 168명의 오배치된 학생들에 대해서 약 39%에 해당하는 65명 정도의 학생들에 대해서 수준이 더 낮은 반으로 배치되었어야 한다고 답했다. 이 65명은 더 낮은 수준으로 분반되었어야 한다고 교수가 명시적으로 답한 경우이다. 이 외에도 대체로 각 반의 평균적인 영어 구사 능력 보다는 떨어지지만 더 낮은 레벨로 분반되어야 한다고 생각하지 않는 경우(N=7)의 학생까지 포함시켰을 경우, 양적 연구에서 오배치된 것으로 확인된 168명의 학생들 중 약 43%의 학생에 대해서 교수들 또한 동일한 의견을 제시한 것으로 볼 수 있다. 뿐만 아니라, 이들 학생들 외에 18명의 학생들은 영어 구사 능력 면에서는 같은 반에 있는 다른 학생들에 비해 뒤지지 않지만 자세 면에서 문제가 있는 것으로 지적되었다. 예를 들어서, 수업을 자주 빠지거나 중간에 수강을 취소한 경우도 있었고, 수업에 들어오더라도 적극적으로 수업에 참여하지 않은 경우이다. 이러한 학생들까지 포함시켰을 경우, 조사 대상 168명 중 무려 54%에 해당하는 90명이 한 학기 동안 이들을 지도했던 교수에 의해서도 배치에 문제가 있는 것으로 평가되었다고 볼 수 있다.

영어 구사 능력 면에서 동일한 집단인지 아닌지를 묻는 두 번째 질문에 대해서는 개인적인 편차를 많이 보였다. 양적 조사 결과 오배치된 학생들의 수가 많을 뿐만 아니라 이들 학생들의 토익 점수가 반 평균과 많은 차이를 보이는 경우도, 영어 구사 능력 면에서 동일하다고 평가하는 경우가 있는가 하면, 이들에 비해서 상대적으로 오배치된 학생 수가 적고 토익 점수 또한 그 편차가 크게 나지 않는 경우에 오히려 영어 구사 능력 면에서 동일하지 않다고 평가하는 경우도 있었다. 예를 들어, 한 교수의 경우, 오배치된 학생이 정원의 60%에 이르고(20명 중 12명) 이 오배치된 학생들의 평균 토익 점수가 분반 커트라인(710점)보다 약 150점 정도가 낮은(562점) 경우에도 다음과 같이 영어 구사 능력에 있어서 차이가 없다고 하였다.

In regards to class levels, I felt that [this class] was almost the same across the board. With the exception of one or two students they were very competent, similar, and very enthusiastic.

반면에 같은 교수는 오배치된 학생의 수가 거의 비슷한(N=11) 다른 분반에 대해서는 영어 구사 능력면에서 동일하지 않다고 지적했다. 이 분반의 경우 분반 커트라인이 650점이지만 오배치된 사람들의 토익 평균은 이보다 200점 정도 차이가 나는 446점이었다. 이 분반의 경우는 레벨뿐만 아니라 수업에 임하는 자세 또한 문제가 있었던 것으로 보인다.

[This class] was a very low level class with levels ranging from low to very poor. They

could have been restructured and placed in a new class. Also the manner and rudeness of the students was quite prevalent.

이처럼 개인적인 편차를 많이 보이고 있지만, 전반적으로는 레벨이 아주 높거나 낮은 경우 오배치된 학생이 있더라도 같은 반 학생들과 점수 차이가 많이 나지 않는다면 교수가 언어 구사 능력 면에서 크게 차이가 나지 않는다고 느낄 수도 있음을 시사한다. 마찬가지로 레벨이 아주 낮은 경우에도 오배치자가 많더라도 점수 차이가 많이 나지 않기 때문에 동일 집단으로 평가하는 것으로 보인다. 다만 이러한 오배치자가 많으면서 토익 점수가 200점 이상으로 현격히 차이를 드러낼 경우, 그리고 수업에 임하는 전반적인 학생들의 태도가 부정적인 경우는 영어 구사 능력 면에서 동일 집단이 아니라는 평가에 이를 가능성이 있는 것으로 보인다.

## V. 논의 및 결론

이상과 같은 양적 연구와 설문 조사 연구를 통해 수능과 토익의 배치고사로서의 기능을 살펴본 결과, 두 시험은 다르게 기능하고 있는 것으로 나타났다. 수능으로 배치한 분반과 모의 토익 점수를 기준으로 배치한 분반의 경우, 40.99%의 학생들이 일치하지 않는 것으로 나타났다. 그리고 이들 중 일부에 대해 교수에게 설문 조사를 한 결과 50% 이상에 대해 문제가 있거나 오배치된 것으로 답하였다. 이러한 결과는 두 시험의 성격을 비교했을 때도 예상 가능하다. 수능의 경우는 학생들의 대학 수학 능력(college scholastic ability)을 알아보기 위한 시험이며, 토익은 영어의 능숙도(English proficiency)를 측정하는 시험이기 때문이다. 뿐만 아니라 Wesche 등(1993)도 이와 비슷한 결론에 도달한 바 있다. 회화 수업에서 네 가지 시험 유형이 배치고사로서 어떻게 기능하고 있는가를 살펴본 결과, 학생들에게 유창성을 요구하는 시험들, 읽기와 듣기와 같은 시험이 배치고사로서 가장 잘 기능하고 있음을 발견했다. 그들은 회화라는 수업의 성격과 이들 시험이 가장 밀접하게 연관되어 있기 때문인 것으로 추정했다. 마찬가지로 본 대학의 회화 수업 또한 말하기와 쓰기를 비롯한 영어 능력 향상에 중점을 두고 있기 때문에 비즈니스 영어에 주안점을 두고 있긴 하지만 전반적인 영어 능력을 측정하는 토익 시험이 수능에 비해서 배치고사로 더 효과적으로 기능했다고 볼 수 있다.

그러나 이러한 결론에 대해서 좀 더 주의 깊은 해석 또한 요구된다. 첫째, 본 연구에서 사용한 학생들의 토익 점수를 얼마나 신뢰할 수 있는가이다. 앞서 언급했듯이, 신입생들에게 입학 전에 온라인 모의 토익을 보도록 독려하기는 했지만 의무화하지는 않았기 때문에 영어 회화 수업을 들은 3,077명의 학생 중 오직 2,474명만이 시험에 응시를 했다. 뿐만 아니라 이들이 얼마나 성실하게

토익 시험을 보았는가 하는 문제도 남는다. 예를 들어, 분반의 토익 커트라인이 780인 반에서 각각 590과 460점을 받은 두 학생의 경우, 그들을 가르쳤던 교수는 두 학생이 다른 학생들에 비해 뛰어나다고 평가하면서 아마도 이 학생들이 토익 시험에서 최선을 다하지 않았을 것이라고 언급했다.

둘째로 학생들의 영어 수준 이외에도 수업에 영향을 미치는 다양한 요소가 존재하기 때문에 오배치 자체가 반드시 문제가 된다고 할 수 없다. 수업을 진행하는 교수의 경우 오배치된 학생이 수업에 성실히 임하는 모습을 보일 경우 레벨에 상관없이 대체로 관대함을 지니고 있었다. 게다가 학기 초와 학기 말 성취도의 차이를 비교 했을 때 낮은 레벨에도 불구하고 이러한 발전을 이루었다 해서 오히려 좋은 인상을 주는 경우도 있었다. 다만 교수들은 레벨의 차이와 더불어 수업에 대한 그릇된 태도가 나타날 때, 어려움을 호소하고 있는 것으로 나타났다. 이러한 수업에 대한 부정적 태도 증가가 동료 학생들의 탈동기화를 가속화하는 것은 이미 김태영, 김은주, 김혜영(2011)에서도 확인된 바 있다.

앞에서도 언급했지만 대부분의 학자들이 가장 이상적인 배치고사는 가르치는 내용을 검사할 수 있는 자체적인 시험이라고 했다. 그러나 이러한 자체적인 개발에 필요한 물리적, 경제적 자원을 감안하면, 현재 많은 대학에서는 자체 개발 보다는 현존하는 시험을 이용할 수밖에 없는 상황이다. 이러한 시험들 중, 본 연구에서와 같은 타당도 검사를 통하여 각 대학의 영어 프로그램의 내용 및 목적과 가장 부합하는 동시에 오배치의 가능성을 최대한 배제할 수 있는 시험을 선택하는 것이 필요하다. 또한 본 연구에서와 같이 학생들의 영어 수준에 대한 교수들의 평가뿐만 아니라 실제 수업에서의 성적까지도 연구에 포함시켜 배치고사 결과와 비교함으로써, 배치고사 결과가 학생들의 대학에서의 학업 성취도와 어떤 연관 관계를 가지고 있는가를 살펴보는 후속 연구도 필요할 것이다.

#### 참고문헌

- 김상호. (1995). 대학수학능력시험 및 고등학교 내신성적과 대학학업성적과의 상관관계 연구. *한국교육문제연구소논문집*, 10, 91-111.
- 김태영, 김은주, 김혜영. (2011). 온-오프라인 결합 교양영 수강 대학생들의 학습동기 유형과 변화. *현대영어교육*, 12(1), 126-156.
- 김현철. (2004). 대학생의 학업성취(I): 대입 전형요소와 학습동기의 예측력 비교. *교육학연구*, 42(4), 343-378.
- 박성수. (2007). 대학입학 전형요소와 대학 학업성취도와의 상관관계에 관한 연구. *한국교육*, 34(4), 27-56.

- 부재율. (2007). 2008 학년도 대학입시 제도에서 전형자료들의 당락에 대한 영향력과 학업성취도 예측력 변화에 대한 분석. *교육평가연구*, 20(3), 57-79.
- 성명희, 현대덕, 이지연, 장형지. (2011). 대학 영어 선이수를 위한 표준교육과정 개발 방향 제안. *현대영어교육*, 12(2), 140-157.
- 신정철, 신태수, 정지선. (2008). 대학생의 학업성취도와 그 영향요인들 간의 인과관계 분석. *교육행정학연구*, 26(1), 287-313.
- 윤명희, 안도희. (2001). 새입시 제도하에서의 지방사립대학 입시방안 연구: 동의대학교를 중심으로. *교육평가연구*, 14(1), 151-170.
- Bachman, L. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.
- Brown, J. D. (2005). *Testing in language programs* (2nd ed.). New York; McGraw-Hill.
- Fulcher, G. (1997). An English language placement test: Issues in reliability and validity. *Language Testing*, 14, 113-139.
- Lee, Young-Ju, & Greene, J. (2007). The predictive validity of an ESL placement test: A mixed methods approach. *Journal of Mixed Methods Research*, 1, 366-389.
- Linn, R. L., & Gronlund, N. E. (2000). *Measurement and assessment in teaching* (8th ed.). Upper Saddle River, NJ: Prentice Hall.
- R Development Core Team (2010). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. Retrieved May 29, 2011, from the World Wide Web: <http://www.R-project.org/>.
- Wall, D., Clapham, C., & Alderson, J. C. (1994). Evaluating a placement test. *Language Testing*, 11, 321-344.
- Wesche, M., Paribakht, S.T., & Ready, D. (1993, August). A comparative study of four placement instruments. Paper presented at the Annual Language Testing Research Colloquium, Cambridge, England.
- Yu, Kyung-Ah, Moon, Chae-Young, & Shin, Sang-Keun. (2007). Placement practices in university English language programs in Korea. *Foreign Languages Education*, 14(4), 27-48.

부록  
설문조사

Dear \* \* \*,

I am interested in seeing whether CSAT (Korean SAT) or TOEIC is a more reliable tool to place students in classes. Through a primary analysis of the scores of the two tests, I found that the following students in your classes have a big discrepancy between their CSAT and TOEIC scores.

Class	Student ID	Student Names
A1	2011****	***
A2	2011****	***
B2	2011****	***
B3	2011****	***
.	.	.
.	.	.

So I would like to hear your opinions about their proficiency levels: what you think of their proficiency, whether they were appropriately placed in your classes, or whether they should have been placed in lower-level classes. I am also curious about your opinions of one of your classes (Group XX, Class YY), whether you think you taught a homogeneous group of students in the class in terms of English proficiency.

Feel free to add additional comments you would like to make on the students. Also, your ideas about how to improve our program are welcome.

조수경  
 숭실대학교 베어드학부대학  
 156-743 서울 동작구 상도로 369  
 Tel: (02) 828-7211  
 Email: sookjungcho@ssu.ac.kr

박찬호  
 한국교육과정평가원  
 100-784 서울 중구 정동길 21-15  
 Tel: (02) 3704-3613  
 Email: cpark@kice.re.kr

Received 15 July 2011  
 Revised 20 August 2011  
 Accepted 3 September 2011