

의사소통 능력 측정 도구로서의 공인 영어 표준화 시험 타당도

이 현 우*

(인하대학교)

이 소 영**

(인하대학교)

Lee, Hyunoo & Lee, Soyoung. (2004) Validity of TOEFL, TOEIC and TEPS as a measure of communicative competence. *English Language & Literature Teaching*, 10(3), 191-210.

The scores of TOEFL, TOEIC and TEPS have been increasingly used for many purposes in Korea. In particular, these test scores are being used as a measure of the test-takers' overall communicative competence. Nonetheless, few studies have been conducted that investigate the validity of the test scores used for this purpose. As a preliminary step to explore the validity of the test scores, we conducted a small scale study by comparing 30 university students' scores in TOEFL, TOEIC and TEPS with their class performances in English conversation and composition courses. The correlations between the test scores and the grade point averages of the courses show that the test scores are much harder to use as a valid measure of test-takers' overall communicative competence than usually thought to be and that the score in TOEFL is, nonetheless, a more reliable measure than the ones in the other tests. Although this study has a few limitations such as the small number of participants, the homogeneity of the sample as a group, etc., it provides some insight into the use of the three tests for measurement of overall communicative proficiency and suggests need for conducting further validation studies in these areas.

[validity/standardized English test, 타당도/영어 표준화 시험]

* 제1저자 표시임

** 제2저자 표시임

I. 서론

한국에서 지난 20여 년간 영어는 어쩌면 모국어보다도 더욱 중요하게 여겨져 왔는지도 모른다 국제간 정치, 문화, 예술 등의 교류 확대와 국제 무역 증대는 한국인에게 영어의 필요성을 각인시키는 계기가 되었고, 이에 따라 영어 교육은 각급 학교에서뿐만 아니라 사회의 거의 전 영역과 전 계층에서 그 어느 때보다도 더욱 강조되어 왔다. 영어는 이제 학교에서 배우는 교과목의 하나로 인식되기보다는 학교 졸업 후 사회생활을 하는 동안에도 능숙하게 구사할 수 있을 정도로 배워야하는 학습의 대상으로 여겨지고 있다.

영어가 학습의 도구 언어로서 기능하는 데 그치지 않고 실생활의 언어 도구로까지 기능해야 한다는 요구가 커짐에 따라, 한국의 영어 교육은 영어로의 의사소통 능력 향상을 궁극적 목표로 삼고 있다 제7차 중등 교육과정의 목표와 대학 영어 강좌의 강의 목표가 이 점을 잘 보여주고 있다 영어 교육의 목표가 영어로의 의사소통을 강조함에 따라, 각종 영어 평가 시험 또한 전반적인 영어 의사소통 능력 측정에 중점을 두고 있다. 각급 학교의 영어 시험, 대학수학능력시험, 각종 임용 및 취직 또는 승진 시험이 모두 이러한 평가 방향을 지향하고 있다고 해도 과언이 아니다. 이점에 있어서 공인 영어 표준화 시험도 마찬가지이다.

한국에서 가장 일반적으로 실시되고 있는 공인 영어 표준화 시험으로는 미국 Educational Testing Service (ETS)에서 개발하고 시행하는 Test of English as a Foreign Language (TOEFL), ETS에서 개발한 Test of English for International Communication (TOEIC) 그리고 서울대 언어교육원에서 주관하고 시행하는 Test of English Proficiency developed by Seoul National University (TEPS)가 있다. 이들 시험은 각각의 세부 목적이 다소 다르지만, 수험생의 영어 의사소통 능력을 측정한다는 점에서는 동일하다. 비록 TOEFL (특히, computer-based test인 경우) 시험의 경우에는 쓰기 능력을 제한적으로 측정하고 있지만, 이들 시험은 표준화 시험으로서 어해 능력 측정에 중점을 둔 시험이라고 할 수 있다. 이러한 시험의 구조적 특성, 즉 듣기와 읽기 능력 측정이 시험의 대부분을 차지한다는 점 때문에 과연 이들 시험이 수험생의 영어로의 의사소통 능력을 제대로 측정할 수 있는가에 대한 의문이 항상 제기되어 왔다.

다음 장에서 보듯이, TOEFL과 TOEIC의 경우에는 영어로의 의사소통 능력을 측정하는 도구로서의 타당도를 검증하기 위한 연구가 ETS를 중심으로 수년간 이루어져 왔고 꽤 많은 연구 결과가 축적되었다. 반면에 TEPS는 본격적으로 개발돼 시행된 지 채 10년이 되지 않은 관계로, 이 시험에 대한 타당도 연구가 매우 빈약한 형편이다. 더군다나, 세 시험에 대한 기존의 타당도 연구가 ESL 상황에서 영어를 배우는 학습자

또는 일본인 영어 학습자를 대상으로 했기 때문에, 이들 시험이 한국의 EFL 상황에서 영어를 학습한 수험생을 대상으로 할 때에도 타당한 의사소통 능력의 측정 도구로 기능하는지는 경험적으로 그 진위를 밝혀야만 한다. 이런 짐을 염두에 두고, 본 연구는 대학생의 공인 영어 표준화 시험 점수와 이들의 영어 회화 및 영어 작문 과목의 평점과의 상관관계를 조사하고 그 결과를 기반으로 이들 공인 영어 표준화 시험의 타당도에 대한 시사점을 얻고자 한다. 동일한 시기에 동일한 표본으로부터 얻어진 세 공인 영어 표준화 시험의 점수를 이들 표본이 두 학기에 걸쳐 취득한 영어 기능 과목의 평점과 비교하는 것은 이들 평점이 표본의 영어 구사 능력을 가장 객관적으로 보여준다는 점과 이 상관관계 연구가 하나 이상의 공인 영어 표준화 시험의 타당도를 동시에 동일 기준에 대해 비교한다는 점에서 그 의미가 있다고 여겨진다.

본 논문은 다음과 같이 구성되어 있다. II장에서는 본 연구의 주 대상인 TOEFL, TOEIC, TEPS를 간략히 소개하고 이들 시험의 타당도와 관련된 선행 연구를 검토할 것이다. III장에서는 본 연구에 참여한 피험자의 소개와 연구 실행 절차 및 방법을 기술할 것이다. IV장에서는 연구의 결과를 제시하고 그에 대한 논의를 다룰 것이다. 마지막으로 V장에서는 본 연구의 결론을 도출하고 이것이 함축하는 바와 후속 연구에 대한 제언을 기술할 것이다.

II. 선행 연구

공인 영어 표준화 시험의 타당도 관련 연구는 주로 각 공인 영어 표준화 시험에 초점을 두고, 이것과 다른 공인영어 시험간의 상관관계를 구하는 양상으로 수행되어 왔다. GRE 혹은 GMAT 등과 같은 적성 검사와 함께 타당도 검증을 한 연구(Wilson, 1982)가 있기는 하나, 본 연구의 주제와 관련되어 있는 TOEFL, TOEIC, TEPS 시험 중 두 가지 이상을 가지고 함께 그 타당도를 조사한 연구는 찾아보기 어렵다. 따라서 본 연구의 주제와 관련되어 있는 선행 연구를 살펴볼 이 장에서는 TOEFL, TOEIC, TEPS 각각의 시험에 대해 이루어진 타당도 연구를 시험별로 나누어 살펴볼 것이다. 그러나 이에 앞서, 독자의 편의를 위해 이들 시험의 성격 및 특징을 간략히 개관하는 것이 필요할 것이다.

1 TOEFL, TOEIC, TEPS 시험의 구성

TOEFL은 모국어가 영어가 아닌 사람이 북미 지역의 대학에서 학업을 수행하는 데 필요요한 영어 구사력을 어느 정도 갖추었는지 측정하려는 목적에서 만들어졌다. 본

연구의 자료로서 활용되는 TOEFL은 paper-based 테스트로서 다음과 같이 세 개의 영역으로 구성되어 있다

Section 1 Listening Comprehension	
Part A Short Conversations	30문항
Part B Long Conversations	8문항
Part C Short Talks	12문항
Section 2 Structure and Written Expression	
Structure Incomplete Sentences	15문항
Written Expression Error Recognition	25문항
Section 3 Reading Comprehension	
Reading Passages	50문항

TOEFL과 마찬가지로 TOEIC 또한 모국어가 영어가 아닌 사람의 영어 구사력을 측정하지만, TOEIC 점수는 직장에서 다른 사람과 영어로 얼마나 잘 의사소통을 하는지 보여준다 TOEIC은 크게 청해와 독해(문법 포함)에 관한 두 개의 영역으로 나뉘어져 있으며, 각 영역의 하위 구성은 다음과 같다¹⁾

Section 1 Listening Comprehension	
Part 1 Photographs	20문항
Part 2 Questions and Responses	30문항
Part 3 Short Conversations	30문항
Part 4 Short Talks	20문항
Section 2 Reading	
Part 5 Incomplete Sentences	40문항
Part 6 Error Recognition	20문항
Part 7 Reading Comprehension	40문항

마지막으로, TEPS는 한국인의 영어로의 의사소통 능력을 측정할 목적으로 서울대학교에서 개발한 시험으로 크게 네 개의 영역, 즉 청해, 문법, 어휘, 그리고 독해로 구성되어 있다

1) TOEIC의 Part 6은 문법에 관한 문항으로 이루어졌고, Part 5는 문법 및 어휘에 관한 문항으로 이루어졌다

Domain 1' Listening Comprehension	
Part 1. Questions and Responses	15문항
Part 2 Incomplete Short Conversations	15문항
Part 3' Short Conversations	15문항
Part 4' Short Talks	15문항
Domain 2. Grammar	
Part 5 Incomplete Short Conversations	20문항
Part 6: Incomplete Sentences	20문항
Part 7 Error Recognition. Short Conversations	5문항
Part 8. Error Recognition' Short Paragraphs	5문항
Domain 3' Vocabulary	
Part 9 Incomplete Short Conversations	25문항
Part 10' Incomplete Sentences	25문항
Domain 4 Reading Comprehension	
Part 11' Incomplete Paragraphs	16문항
Part 12' Reading Passages	21문항
Part 13. Unity of Paragraphs	3문항

2 TOEFL의 타당도

TOEFL은 1960년대부터 지금까지 계속 여러 종류의 영어 숙달도 검사와의 상관관계를 통해 그 타당도에 대한 검증을 해 오고 있다 (TOEFL Test and Score Manual, 1997 Edition; TOEFL Score User Guide, 2000-2001 Edition) 100명의 수험자를 대상으로 하여 Michigan Test of English Language Proficiency 점수와 TOEFL 점수 간에 .89라는 높은 상관관계 지수를 보고한 Upshur (1966)의 연구, TOEFL 점수와 구두 면접 및 작문 점수와의 관계를 살펴본 후 59에서부터 87에 이르는 비교적 높은 상관관계 지수를 보고한 Pike (1979)의 연구 등이 그 예다 이와 같은 준거 지향적 타당도 (criterion-referenced validity)에 대한 검증은 1990년대에 들어서도 계속 이루어졌고 (Henning & Cascallar, 1992), 이와 더불어 구인 타당도에 대한 검증이 다양한 연구 설계 및 통계 분석을 통해 이루어져 오고 있다 (Oltman, Stricker, & Barrows, 1988) 그러나 TOEFL Score User Guide, 2000-2001 Edition에서 지적하고 있는 바와 같이, TOEFL의 타당도를 검증한 많은 연구가 대부분 대규모로 이루어져 온 것은 단점으로 작용할 수도 있는 바, 이는 다음의 인용에서도 확인할 수 있다

One way to approach validity is to conduct a local study that establishes a relationship between TOEFL scores and the linguistic and curricular demands faced by students. The TOEFL program has undertaken a large-scale study that assesses the linguistic needs of international students in a variety of discrete academic settings. Future studies will examine the relationship between scores and placements at various colleges and universities.

(TOEFL Score User Guide, 2000-2001 Edition, p. 29)

즉 TOEFL 점수가 입학 및 배치와 같은 목적의 준거로 사용되기 위해서는 특정 TOEFL 점수가 그 목적의 효과적 달성에 기여하는 정도를 TOEFL 점수를 준거로 채택한 학교 및 기업 등과 같은 구체적 장에서 확인해야 하는 것이다. 이럴 경우, TOEFL을 주관하고 있는 기관인 ETS에서 수행하는 연구와는 달리 특정한 장을 배경으로 하는 소규모적인 연구가 되어 샘플의 크기 등과 같은 요소가 연구 결과의 일반화에 걸림돌이 될 수도 있을 것이다. 그렇지만 문항 제시 방법의 영향을 알아보기 위해서는 “실제” 수험자를 대상으로 하여 확인할 필요성이 있다고 강조한 Chalhoub-Deville과 Deville (1999)의 시각처럼, 우리나라에서 사용되는 TOEFL 점수의 타당성 검증은 우리나라에서 TOEFL 점수가 사용되는 목적이 무엇인지 알아보는 것에서부터 출발하여 이루어져야 할 필요가 있다고 여겨진다.

3 TOEIC의 타당도

TOEIC은 일본 기업의 지원 아래 ETS에서 개발한 시험으로서 주로 일본과 한국에서 사용되고 있다. 그러나 TOEIC의 타당도와 신뢰도에 대한 체계적인 연구는 TOEFL에 비하면 상대적으로 적은 편이다. 이런 상황에서 TOEIC의 타당도에 관해 살펴볼 주된 문헌은 역시 TOEIC을 개발한 기관에서 출간된 자료일 수밖에 없다. TOEIC의 타당도에 관해 부분적으로 기술해 놓은 TOEIC Technical Manual에 따르면, TOEIC은 여러 영어 숙달도 검사와의 상관관계를 통해 행해진 준거 타당도 검증에서 상당히 높은 상관관계 지수를 보여줌으로써 타당도가 높은 검사라고 한다. 이를 좀 더 구체적으로 살펴보면 다음과 같다. 다양한 모국어 배경을 가진 393명의 TOEIC 수험자의 점수와 FSI (Foreign Service Institute)에서 개발한 말하기 검사인 LPI (Language Proficiency Interview) 점수간의 상관관계 지수가 청해 영역에서 74, 독해 영역에서 68, 평균 74로 나타나 TOEIC이 구두 언어 평가 검사와 마찬가지로 구어의 이해 및 사용 능력을 효과적으로 측정하는 타당한 검사라는 것이다. 이러한 타당도 검증은 말하기 검사뿐만 아니라 여러 가지의 듣기, 읽기, 쓰기 검사에서 적게는 30명부터 많게는 306명의 수험자를 대상으로 하여 이루어졌고, 그 결과 낮게는 67부터 높게는 92에 이

르는 상관관계 지수가 보고되었다 (TOEIC과 개별 검사간의 상관관계 지수에 대한 구체적 내용은 TOEIC Technical Manual의 III-1에서 III-5를 참조하라)

그러나 각각의 언어 기능을 다루는 검사와의 상관관계에서 이처럼 높은 지수를 보여준 TOEIC이 영어 교수-학습의 장인 교실에서 나타나는 학습자의 영어 수행 능력과의 상관관계에서는 상당히 편차가 큰 지수를 보여주고 있다 학습자의 영어 숙달도에 따라 네 등급으로 구분하여 문법, 듣기, 읽기, 말하기, 쓰기 영역별 수행 능력과 TOEIC 점수 간에 구해진 상관관계 지수는 03에서부터 .82에 이르기까지 다양하게 나타나고 있다 (TOEIC Technical Manual, III-6) TOEIC Technical Manual이 네 등급 중에서 상대적으로 일관적이고 높은 상관관계 지수를 보여주고 있다고 언급한 중급 학습자 집단에서조차 36에서 52 정도에 그치는 상관관계 지수가 보고되어 있다 비교적 높고 일관된 상관관계 지수를 보여주었던 개별 언어 기능 검사와의 관계와 달리 교실 내 학습자의 영어 수행 능력에서는 변이가 큰 상관관계 지수가 나온 것은 이에 대해 심도 있는 분석 및 해석이 수반되어야 함을 함축한다고 여겨진다 또한 이와 유사한 맥락 하에서 후속 연구가 이루어져 선행 연구 결과와의 비교, 검토가 이루어져야 할 필요성을 제기하는 것이라고 여겨진다

4 TEPS의 타당도

TOEFL과 TOEIC에 비해 뒤늦게 개발되었기에 TEPS의 타당도에 관한 연구는 상대적으로 다른 두 시험에 비해 매우 적은 편이다 또한 "1999년 이후 실제 시행된 문제와 답안 등의 자료를 공개한 적이 없기 때문에 그러한 자료들을 바탕으로 한 연구가 거의 없다는 것이 대단히 아쉬운 점이다" (이완기, p 218, 2003) 여기에서는 TEPS 개발에서 중심적 역할을 한 최인철의 두 연구를 살펴볼 것이다 최인철 (1999)의 "Test fairness and validity of the TEPS" 연구에서는 56명의 연구 참여자들을 대상으로 하여 당시 서울대학교 언어연구원에서 개발한 구두 면접 시험인 Test of Oral Proficiency (TOP) 점수와 TEPS 점수와의 상관관계를 구했다 그 결과 TEPS의 하위 영역 중 일부인 어휘와 독해를 제외하고는 전반적으로 TEPS는 TOP와 높은 상관관계 지수를 지닌 것으로 나와 있다. 특히 TEPS의 청해 점수와 TOP는 6 이상의 높은 상관관계를 보여주고 있다 또한, 이 연구에서는 TOEFL이나 TOEIC의 시험과는 다른 검사 방법 (test method)을 지닌 TEPS에 대해서 314명의 수험자에게 설문지 조사를 통해 1지문 1문형 원칙등과 같은 TEPS 고유의 검사 방법이 타당함을 보여주기도 했다 하지만 TOP는 수험자의 구두 능력을 간접적으로 평가하는 simulated oral proficiency interview (SOPI)이어서 직접적인 의사소통 능력을 측정하는 검사인지에 관해서는 논란의 여지가 남아 있다 (SOPI에 관한 자세한 내용은 Stansfield & Kenyon (1992)을

참조하라)

한편, 최인철 (1995)의 또 다른 연구는 TEPS의 초기 버전인 SNUCREPT와 TOEIC 간의 비교에 중점을 둔 것으로서, 42명의 수험자를 대상으로 해서 이 두 시험의 상호간 상관관계뿐만 아니라 회화 과목에서 실시한 일련의 검사 점수로서 나타내어진 의사소통 능력과 이 두 시험의 하위 요소들간의 상관관계도 보고하고 있다 그 결과, 의사소통 능력을 측정하는 도구로서 TEPS가 TOEIC보다 더 타당함을 주장하고 있다 그러나 의사소통 능력의 지표로서 사용된 회화 과목의 일련의 검사가 무엇이었는지 명확하게 나타나 있지 않아 TOEIC에 비해 TEPS의 타당도가 높다고 한 결론이 후속 연구에 의해 검토될 필요가 있다고 본다

III. 연구 방법

1 연구 참여자

2002년 데이터 수집 당시 경인 지역에 위치하고 있는 한 대학교의 영어 관련 학과에 재학 중이던 3학년 남녀 학생 30명이 본 연구에 참여하였다 이들의 입학 당시 대학 입학 수학 능력 시험 점수는 상위 5~10% 정도이며, 본 연구의 데이터 수집을 위해 이들이 치른 TOEFL, TOEIC, 그리고 TEPS 시험의 평균 점수는 각각 520, 665, 720점으로 중급의 영어 수준을 지닌 집단이라고 할 수 있다 영어 구사력에 대한 이러한 해석은 아래 제시된 표 1, 2, 3을 토대로 이루어진 것이다

표 1은 1996년 7월부터 1997년 6월까지 실시된 TOEFL 시험에 응시한 수험생 794, 989명을 대상으로 작성된 점수 분포표이다 이 표의 내용과 본 연구 참여자들이 얻은 TOEFL 평균 점수인 청해 50, 문법 52, 독해 54, 그리고 총점 520을 비교해 볼 때, 이들은 상위 55%에 해당하는 집단으로 해석할 수 있다 즉, 본 연구 참여자들의 영어 수준은 중급 수준이라고 할 수 있다

표 1
TOEFL 점수에 따른 퍼센타일 등위

Listening Comprehension	Structure & Written Expression	Reading Comprehension	Total	Percentile Rank
63	55	50	540	57
52	45	39	520	45
40	35	29	500	33

(TOEFL Test and Score Data Summary, 1997-1998 Edition, p 2)

본 연구 참여자들의 영어 수준에 대한 이러한 해석은 TOEIC User's Guide에 나와 있는 TOEIC 점수 해석 지침과도 비슷하다. FSI가 개발한 FSI/TOEIC 점수 등급은 모두 5등급으로 (3, 2+, 2, 1+, 1의 순서)²⁾ 되어 있는데, TOEIC 총점 평균 720을 얻은 본 연구의 참여자들은 중간 수준인 2에 해당한다고 볼 수 있다 FSI가 규정하고 있는 2 등급의 영어 구사력은 다음의 표 2에 자세히 제시되어 있다

표 2
FSI의 TOEIC 점수 해석

FSI/ TOEIC	Communication	Comprehension	Expression
2 605-780	Capable of holding a conversation with assurance but a limited mastery of vocabulary	Comprehension of general meaning of non-complex conversations in known subjects	Can provide professional and personal information Can use simple constructions and circumlocution

(TOEIC User's Guide, p 14)

한편, TEPS 역시 연구 참여자의 영어 수준에 대해서는 TOEFL 및 TOEIC과 거의 동일한 결과를 보여준다 본 연구 참여자들의 TEPS 총점 평균은 665, 청해 258, 독해 282, 문법 63, 어휘 61인데, 이는 총 10급의 (1+, 1에서부터 5까지) 수준으로 나누어진 TEPS 등급에서 중상위급인 2급에 해당하는 것이다 2급이 나타내는 영어구사력은 다음의 표 3에 자세히 나타나 있다.

표 3
TEPS 등급표: 2급

601-700	전반 외국인으로서 중상급 수준의 의사소통능력 증장기간 집중 교육을 받으면 일반분야 업무를 큰 어려움 없이 수행할 수 있음
2급 241-280	청해 일반적 상황에 보통수준의 내용을 대체로 이해하는 정도의 청해력 독해 일반적 소재에 보통수준의 내용을 대체로 이해하는 정도의 독해력
61-70	문법 일반적인 구문을 대체로 이해하는 정도의 문법 능력 어휘 일반적인 표현을 대체로 이해하는 정도의 어휘력

2) 실제 TOEIC User's Guide에는 0+와 0등급까지 있으나, 연구의 목적상 이 두 등급은 의미가 없어 제외한다

2 데이터 수집 및 분석

본 연구를 위해 수집한 데이터는 다음과 같이 크게 두 가지로 나누어 기술할 수 있다. 첫째, 연구 참여자들이 재학 중이던 학과의 교과과정상 필수 과목으로 지정되어 있던 총 네 개의 영어 기능 영역 과목들, 즉 회화 1과 2, 작문 1과 2를 수강한 후에 받은 점수이다. 연구 참여자들이 소속되어 있던 학과에서는 이 과목들을 1, 2학년용으로 개설하고 있었다. 이 과목들은 모두 원어민 교수자가 담당하였는데, 이들이 본 연구의 데이터로 사용된 연구 참여자들의 학점을 어떤 평가 기준으로 부여하였는지 등에 관한 정보를 포함하고 있는 교수 요목은 부록 1에 제시되어 있다.

둘째, 연구 참여자들의 동의 및 자발적인 참여 하에 2002년 4월부터 5월에 걸쳐 이들이 치른 TOEFL, TOEIC, 그리고 TEPS 시험의 점수이다. TOEFL 점수는 현재 실시되고 있는 CBT 형태의 TOEFL 대신 PBT 형태의 Institutional Testing Program (ITP) TOEFL을 연구 참여자들이 2002년 4월 27일에 치르고 난 후 ITP TOEFL 시험을 관장하는 한미교육위원단으로부터 개별 통보받은 것을 수집하였다. TOEIC과 TEPS 점수의 경우, 연구 참여자들이 각각 해당 정기 시험에 응시하여, 2002년 5월 26일에는 TOEIC 시험을, 2002년 5월 19일에는 TEPS 시험을 치른 후 관장 기관으로부터 개별 통보받은 것을 수집하였다 (연구 참여자별 공인 영어 표준화 시험 점수는 부록 2에 제시되어 있다). 이와 같이 수집한 세 가지의 공인 영어 표준화 시험 점수와 영어 기능 영역 과목 점수를 토대로 기술 통계 및 상관관계 지수를 구하였다.

IV. 결과 및 논의

표 4는 공인 영어 표준화 시험과 영어 기능 과목의 점수에 대한 기술 통계를 보여주고 있다. 표 4에서 볼 수 있듯이, TOEFL의 독해 영역을 제외하고는 모든 데이터의 척도 지수와 왜도 지수가 -2에서 +2 내에 있어 공인 영어 표준화 시험과 영어 기능 과목 점수 모두 정규 분포를 갖는다고 볼 수 있다. -1.94에서 -0.39의 왜도 지수를 보여준 TOEFL의 경우, 문법이나 청해 영역에 비해 독해 영역이 쉬웠음을 알 수 있다. 하지만 -0.97에서 -0.32의 왜도 지수를 보여준 TOEIC에서는 청해에 비해 독해가 어려웠음을 알 수 있다. TEPS의 경우, 왜도 지수가 -0.69에서 0.35에 걸쳐 있는데 어휘가 다른 영역에 비해 어려웠음을 알 수 있다. 영어 기능 과목의 학점 스케일 (A+, A0, D+, D0, F)을 (4.5, 4, 1.5, 1, 0)로 전환한 표 4의 자료는 영어 회화 과목이 영작문 과목보다 조금 쉬웠음을 나타내고 있다.

표 4
공인 영어 표준화 시험 점수와 영어 기능 과목 점수에 대한 기술 통계

	TOEFL				TOEIC		
	총점	청해	문법	독해	총점	청해	독해
평균	520	50	52	54	720	369	351
표준오차	6.43	71	64	1.06	19.46	10.63	11.56
중앙값	530	51	53	55	718	365	353
최빈값	547	49	55	58	830	420	350
표준편차	35.21	3.89	3.53	5.81	106.60	58.20	63.32
분산	1239.52	15.14	12.46	33.77	11362.90	3387.39	4009.31
첨도	1.87	-24	-72	4.80	17	1.27	-95
왜도	-1.28	-62	-39	-1.94	-67	-97	-32
최소값	410	41	46	33	450	215	230
최대값	563	56	58	61	870	455	455

표 4 (계속)

	TEPS				영어 기능 과목		
	총점	청해	문법	어휘	독해	회화	작문
평균	665	258	63	61	282	3.18	2.98
표준오차	17.68	7.73	1.62	1.40	8.80	14	11
중앙값	670	257	67	61	278	3.25	3
최빈값	693	265	68	66	245	3.25	2.5
표준편차	96.84	42.35	8.88	7.69	48.18	78	60
분산	9377.90	1793.58	78.8	59.14	2321.65	61	36
첨도	-80	-63	-56	-04	-76	-46	-96
왜도	-11	-09	-69	35	-02	-36	14
최소값	456	169	45	46	181	1.5	2
최대값	815	330	76	79	358	4.5	4

다음의 표 5는 세 가지 공인 영어 표준화 시험의 총점 및 각 시험의 하위 영역 점수

와 영어 기능 과목의 평점 사이의 상관관계를 보여주고 있다 이들 상관 계수의 검토 및 비교를 통해 얻은 결과는 다음과 같다

표 5
공인 영어 표준화 시험 점수와 영어 기능 과목 점수간의 상관관계

	토플 총점	토플 성취	토플 문법	토플 독해	토익 총점	토익 성취	토익 독해	텡스 총점	텡스 성취	텡스 문법	텡스 어휘	텡스 독해	회화	작문
토플 총점	1													
토플 성취	.76*	1												
토플 문법	.66*	.21	1											
토플 독해	.91*	.58*	.44	1										
토익 총점	.76*	.68*	.49*	.63*	1									
토익 성취	.69*	.72*	.34	.56*	.87*	1								
토익 독해	.65*	.48*	.50*	.55*	.89*	.54*	1							
텡스 총점	.74*	.73*	.44	.60*	.84*	.70*	.77*	1						
텡스 성취	.64*	.72*	.31	.49*	.80*	.77*	.65*	.94*	1					
텡스 문법	.64*	.53*	.47*	.53*	.80*	.66*	.74*	.69*	.55*	1				
텡스 어휘	.58*	.50*	.47*	.43	.70*	.53*	.70*	.64*	.56*	.67*	1			
텡스 독해	.72*	.64*	.45	.61*	.72*	.53*	.73*	.95*	.82*	.61*	.51*	1		
회화	.48*	.40	.26	.44	.28	.34	.16	.11	.05	.27	.13	.11	1	
작문	.43	.48*	.21	.34	.21	.28	.09	.23	.21	.26	.10	.21	.62*	1

* $p < .01$

연구 참여자의 공인 영어 표준화 시험 점수간의 상관 계수는 각 시험의 총점을 기준으로 했을 때, TOEFL과 TOEIC 간에는 .76, TOEFL과 TEPS 간에는 .74, 그리고 TOEIC과 TEP 간에는 .84로 나타나고 있다 반면 회화 과목 점수와 세 공인 영어 표준화 시험의 총점 간에는 TOEFL의 경우 .48, TOEIC의 경우 .28, 그리고 TEPS의 경우

11의 낮은 상관 계수를 보여주고 있다 이러한 낮은 상관 계수는 작문 과목 점수와의 비교에서도 비슷한 현상으로 나타나고 있는 바, TOEFL의 경우 43, TOEIC의 경우 21, 그리고 TEPS의 경우 23의 상관 계수를 확인할 수 있다 공인 영어 표준화 시험 점수 상호간에 높은 상관관계가 있음에도 불구하고, 이 시험 점수가 교실 내 학습자의 영어 수행 능력, 특히 말하기와 쓰기라는 표현 능력을 나타내는 지표인 기능 과목의 점수와는 낮은 상관관계를 보여주고 있는 점은 II장에서 언급했던 선행 연구의(TOEIC Technical Manual) 결과와 같은 것이어서 주목할 만하다 TOEIC 점수와 교실 내 영어 수행 능력 점수간의 상관관계를 조사한 이 연구(표 6 참조)에 따르면, TOEIC 총점 및 하위 영역 점수와 기능 과목 점수 간의 상관 계수가 학습자의 언어 수준에 따라 매우 큰 편차(03~ 82)를 보여주고 있다.

표 6
TOEIC 점수와 교실 내 영어 수행 능력간의 관계

	Grade	Listen- ing	Read- ing	Total		Grade	Listen- ing	Read- ing	Total
Advanced Beginner	Gram- mar	03	33	18	Advanced Intermediate	Gram- mar	48	38	45
	Listen- ing	19	40	31		Listen- ing	70	59	74
	Read- ing	10	32	21		Read- ing	49	37	50
	Speak- ing	20	32	28		Speak- ing	42	05	29
	Writ- ing	17	29	24		Writ- ing	26	25	29
Intermediate	Gram- mar	35	57	52	Advanced	Gram- mar	11	74	57
	Listen- ing	.40	21	.36		Listen- ing	75	63	72
	Read- ing	28	51	45		Read- ing	66	82	78
	Speak- ing	.35	48	47		Speak- ing	68	69	73
	Writ- ing	.26	54	45		Writ- ing	50	65	61

(TOEIC Technical Manual, III-6)

표 6에서 주목할 것은 본 연구의 참여자와 동일한 수준의 영어 구사력을 갖추었다고 여겨지는 Intermediate 집단의 상관 계수이다 이들 계수에 비해, 연구 참여자의

TOEIC 점수와 영어 기능 과목 점수간의 상관 계수는 상당히 낮다 연구 참여자의 TOEIC 총점과 회화 및 작문 과목의 평점과의 상관 계수는 각각 28과 .21로 표 6에서 이에 상응하는 계수 36 또는 47과 45보다 낮다 특히, 연구 참여자의 TOEIC 독해 점수와 작문 과목의 평점과의 상관 계수 .09는 표 6의 상응하는 계수 54보다 매우 현저히 낮다

TOEFL이나 TEPS 점수와 교실 내 영어 수행 능력 점수간의 상관관계를 조사한 연구가 없는 관계로 본 연구의 결과를 일반화하는 데에 다소의 난점이 있으나, 표 6에 제시된 TOEIC의 결과가 TOEFL과 TEPS에서도 대동소이하게 나타난다고 가정한다면, 위의 결과는 학습자의 교실 내 영어 수행 능력을 나타내는 지표로서 공인 영어 표준화 시험 점수를 사용하는 것은 조심스럽게 이루어져야 함을 함축한다 더 나아가 본 연구의 영어 기능 과목 점수는 회화, 작문 각각 2학기라는 비교적 장기간에 걸쳐 연구 참여자들의 영어 구사력을 평가한 결과이고 따라서 일회적인 시험의 결과보다는 좀 더 충실하게 학습자의 영어 구사력을 드러낸 지표임을 고려해 볼 때, 공인 영어 표준화 시험 점수와 이 과목의 점수 간의 낮은 상관 계수는 공인 영어 표준화 시험이 지향하고 있는 구인에 대한 재고의 필요성을 드러내는 것이라고 여겨진다

앞서 살펴본 바와 같이 영어 기능 과목의 점수와 공인 영어 표준화 시험 점수간의 상관 계수가 기존의 연구보다 낮게 나왔지만, 그래도 이 세 공인 영어 표준화 시험 가운데에서는 TOEFL이 TOEIC이나 TEPS보다는 상대적으로 좀 더 높은 상관 계수를 보여주고 있다 TOEFL은 총점을 기준으로 했을 때, 회화와 작문 과목의 평점과 각각 48과 (01 수준에서 통계적으로 유의한 차이를 나타냄) 43의 상관 계수를 가지고 있는 반면 TOEIC은 28과 21, 그리고 TEPS는 11과 23의 상관 계수를 나타내고 있다. 즉, 회화와 작문 각각의 점수와의 상관관계에 있어서 TOEFL이 다른 두 시험에 비해 상대적으로 우위에 있다

다른 두 공인 영어 시험에 비해 TOEFL이 타당도 면에서 우월하다는 것은 보다 거시적인 측면에서도 확인된다 영어 회화와 작문 과목에서의 평점을 합산한 수치가 연구 참여자의 영어 구사력을 가장 정확하게 계량화한 것이라고 할 때, 아래 표 7의 결과는 이 점을 보다 분명하게 보여준다

표 7

공인 영어 표준화 시험 점수와 회화·작문 과목 점수(합산)간의 상관관계

	토플 총점	토플 청해	토플 문법	토플 독해	토플 총점	토플 청해	토플 독해	토플 총점	토플 청해	토플 문법	토플 어휘	토플 독해	회화	작문
회화 +	51*	48*	26	44	28	35	14	18	13	30	13	17	92*	87*
작문														

* $p < .01$

세 공인 영어 표준화 시험의 총점과 회화와 작문 점수의 합산과의 상관 계수가 TOEFL은 51, TOEIC은 28 그리고 TEPS는 18로 나타나고 있다 이러한 차이는 영어 구사력의 척도로서 TOEFL 점수가 다른 두 시험의 점수에 비해 상대적으로 우위에 있음을 확인해 주는 것이라고 할 수 있다³⁾

영어 구사력의 측정 도구로서 TOEFL이 TOEIC이나 TEPS보다 좀 더 신뢰할 만하다는 사실은 이들 시험의 하위 영역의 점수와 이들 영역에 해당하거나 밀접한 관계가 있는 기능 과목의 성적과의 상관관계를 비교함으로써 한 번 더 확인할 수 있다 표 5에서 보듯이, 세 공인 영어 표준화 시험의 청해 점수와 회화 과목의 평점과의 상관관계를 보면, TOEFL은 40, TOEIC은 34 그리고 TEPS는 05이다 반면에 공인 시험의 독해 점수와 영작문 과목의 평점과의 상관관계를 보면, TOEFL은 34, TOEIC은 09 그리고 TEPS는 21이다 이러한 비교는 TEPS의 청해가 다른 두 시험의 청해에 비해 문제가 있으며 TOEIC의 독해가 다른 두 시험의 독해에 비해 문제가 있음을 시사한다고 여겨진다 특히, 영작문 평점과의 상관관계에 있어서, TOEIC의 청해 점수가 독해 점수보다도 지나치게 높게 나온 것은 TOEIC의 독해 시험이 안고 있는 한계를 보여주는 것이라고 여겨진다.

V. 결론 및 제언

본 연구는 장기간에 걸쳐 수강하는 영어 회화 과목과 영작문 과목에서의 성적이 연구 참여자의 영어 의사소통 능력을 가장 잘 반영한다는 가정 아래, 연구 참여자의 기능 과목의 성적과 TOEFL, TOEIC, TEPS의 총점 및 하위 영역에서의 점수 사이의 상관관계를 조사하였다 연구의 결과를 요약하면 다음과 같다 첫째, 세 공인 영어 표준화 시험의 총점 또는 그 하위 영역 점수와 영어 회화 및 작문의 수업 수행 능력 사이의 상관 계수는 기존의 연구에서 밝혀진 것보다 상당히 더 낮게 나왔다 둘째, 비록 본 연구의 상관 계수들은 기존 연구의 결과보다 낮게 나왔지만, 일정한 양상을 보여주고 있다 즉, 전반적인 영어 구사력을 측정하는 도구로서 사용된다고 할 때, TOEFL이 TOEIC이나 TEPS보다 전술한 기존의 모든 면에서 구인 타당도가 더 높게 나타나고 있다

첫 번째 결과와 관련하여, 영어 기능 과목에서의 평가가 수강생의 영어 구사력을 얼

3) 표 7에서 보듯이, TOEIC은 회화와 작문의 평점을 합산한 것과의 상관관계에 있어서 총점은 28, 청해는 35 그리고 독해는 14이다. 청해 문항보다 독해 문항의 구인 타당도가 더욱 문제시된다고 말할 수 있다 TOEIC보다도 결과가 더 좋지 않은 TEPS는 문법 영역이 다른 영역에 비해 구인 타당도가 그래도 좀 괜찮은 것으로 여길 수 있다

마나 객관적 반영할 수 있겠느냐고 의의를 제기할 수도 있겠지만, 이와 비슷한 의의는 일회성 지필 고사로 영어 구사력을 측정한다는 공인 영어 표준화 시험에도 얼마든지 제기할 수 있다 영어 회화 및 작문 과목에서의 수업 수행 능력을 2학기에 걸쳐 평가해 얻은 성적의 합이 한 번 치러 얻은 세 공인 시험의 점수보다 연구 참여자의 영어 구사력을 보다 정확하게 반영하고 있다고 본다면, 두 점수 사이의 낮은 상관 계수는 공인 영어 표준화 시험을 통해 수험자의 영어 의사소통 능력을 추론하는 것에 무리가 있음을 보여준다고도 할 수 있겠다

TOEFL 점수가 TOEIC이나 TEPS 점수보다 좀 더 신뢰할 만한 영어 구사력의 척도 역할을 한다는 두 번째 결과는 대학 수학 능력과 공인 영어 표준화 시험간의 상관관계를 조사한 이현우와 이소영(2003)의 연구 결과와도 부합한다 이 연구에 따르면, TOEFL 점수가 다른 두 시험의 점수보다도 더욱 신뢰할 만한 대학 수학 능력의 척도가 된다 이런 결과들은 TOEFL이 다른 두 시험에 비해 상대적으로 더 오랜 기간동안 다수의 수험자를 대상으로 실시해 왔고, 실시 과정 및 결과에 대한 지속적인 연구 및 개선 노력을 기울여왔기 때문인 것으로 보인다

글을 맺으며, 본 연구에서 얻은 결과 및 이에 대한 논의는 사례 연구의 한계성을 내포하고 있음을 밝혀둔다 즉, 본 연구는 제한된 수의 인원만이 연구의 실험에 참여한 것을 바탕으로 결과를 도출하고 있는 점, 이들 연구 참여자들이 특정 대학의 특정 전공을 이수하고 있던 동질 집단이라는 점, 그리고 세 공인 영어 시험 점수의 비교 대상이었던 영어 회화 및 작문 점수가 의사소통능력을 가장 잘 보여주는 지표이기는 하나 숙달도 평가를 지향하는 공인 영어 시험과는 달리 성취도 평가의 속성도 가지고 있다는 점 등과 같은 제한점을 가지고 있기도 하다 따라서 이런 제한점에 극복한 후속 연구를 통해 본 연구 결과의 검증이 이루어질 필요가 있다 이런 제한점에도 불구하고 본 연구는 세 주요 공인 영어 표준화 시험을 국내의 EFL 상황 속에서 동일한 대상에 대하여 동일한 시기에 실험함으로써 전반적인 의사소통 능력의 측정 도구로서의 공인 영어 표준화 시험의 사용에 대해서 몇몇 시사점을 던져주었고, 이와 관련한 구인 타당도의 연구가 더욱 활발히 이뤄져야 한다는 필요성을 확인시켜 주었다고 본다

참 고 문 헌

- 이현우, 이소영 (2003) TOEFL, TOEIC, TEPS 시험 점수와 대학 수학 능력과의 연관성 연구 *영어어문교육*, 9(1), 153-171
- 이완기 (2003). *영어 평가 방법론*. 서울 문진미디어
- 최인철 (1995) A comparability study on SNUCREPT and TOEIC *Language*

- Research*, 31(2), 357-386
- 최인철 (1999) Test fairness and validity of the TEPS *Language Research*, 35(4), 357-386.
- Chalhoub-Deville, M, & Deville, C (1999) Computer adaptive testing in second language contexts *Annual Review of Applied Linguistics*, 19, 273-299
- Henning, G and Cascallar, E. (1992). *A preliminary study of the nature of communicative competence* (TOEFL Research Report 36). Princeton, NJ: Educational Testing Service
- Oltman, P K, Stricker, L J, and Barrows, T. (1988) *Native language, English proficiency, and the structure of the Test of English as a foreign Language* (TOEFL Research Report 27) Princeton, NJ Educational Testing Service
- Pike, L. (1979). *An evaluation of alternative item formats for testing English as a foreign language* (TOEFL Research Report 2) Princeton, NJ. Educational Testing Service
- TEPS 시험안내 (2003) Retrieved October, 7, 2003, from the World Wide Web. http://www.teps.or.kr/html/info/teps_3.html
- TOEFL Test and Score Data Summary, 1997-1998 edition (1997) Retrieved October, 7, 2003, from the World Wide Web <http://www.toefl.org/pubs/pubsindx.html>
- TOEFL Test and Score Manual, 1997 edition (1997) Retrieved October, 7, 2003, from the World Wide Web <http://www.toefl.org/pubs/pubsindx.html>
- TOEFL Score User Guide, 2000-2001 edition (2000). Retrieved October, 7, 2003, from the World Wide Web <http://www.toefl.org/pubs/pubsindx.html>
- TOEIC Technical Manual (2003). Retrieved October, 7, 2003, from the World Wide Web http://www.toeic.com/2_4info/library.htm
- TOEIC User's Guide (2003) Retrieved October, 7, 2003, from the World Wide Web http://www.toeic.com/2_4info/library.htm
- Upshur, J A (1966) *Comparison of performance on "test of English as a foreign language" and "Michigan test of English proficiency"* Unpublished manuscript University of Michigan
- Wilson, T. (1982) *GMAT and GRE aptitude test performance in relation to primary language and scores on TOEFL* (TOEFL Research Report 12). Princeton, NJ Educational Testing Service

부록 1

영어 회화 및 영작문 과목의 교수 요목

A 「영어 회화 1-2」 교수 요목

-강의 요지

- 1 To practice vocabulary, structures, and functions common in everyday conversations in English
- 2 To increase active vocabulary, vocabulary common in everyday conversations
- 3 To discuss and solve problems in English, collect and examine information, then report/share results with others
- 4 To be creative and at ease when speaking English with others.
- 5 To study and practice the following structures in conversational situations statements & responses, tag questions & responses, simple past tense, past continuous tense, when, while & during, present perfect tense, simple past vs present perfect; *for, since & ago, most vs most of, both, either & neither, used to, be used to + -ing*
- 6 To manage and participate in small group discussions The lead of a small group discussion for one hour ENTIRELY in English

평가 방법

- | | |
|----------------------|----------------------|
| 1 Attendance/Tardy 5 | 2. Homework 5 |
| 3. Use of English 10 | 4 (Oral) Quizzes 20 |
| 5. Group Work 10 | 6 Midterm-written 10 |
| 7 Midterm-oral 10 | 8 Final-written 10 |
| 9 Final-oral 20 | |

B. 「영작문 1-2」 교수 요목

-강의 요지

The main goal of this course is to become more familiar with PROCESS WRITING in paragraphs and essays We will learn about the writing process, grammar and style Students will map their progress by writing regularly in a JOURNAL, and will have the opportunity to write an ESSAY as their final exam We will write about various topics, and focus specifically on TECHNIQUES for good writing.

-평가 방법

- | | |
|--------------------------|-------------------------|
| 1. Attendance: 5 | 2 Participation 5 |
| 3 Written Assignments 20 | 4. Journal 30 |
| 5 Midterm Exam 20 | 6 Final Exam/Project 20 |

부록 2

연구 참여자의 공인 영어 표준화 시험 점수

이름	TOEFL 총점	TOEFL 영역 점수			TOEIC 총점	TOEIC 영역 점수		TEPS 총점	TEPS 영역 점수			
		청해	문법	독해		청해	독해		청해	문법	어휘	독해
1	563	55	53	61	715	365	350	698	270	62	71	295
2	557	52	57	58	830	420	41	677	265	67	68	277
3	557	53	58	56	830	395	435	772	280	76	65	351
4	553	51	57	58	805	420	385	801	319	76	64	342
5	550	52	55	58	840	425	415	808	301	71	77	353
6	550	50	55	60	735	360	435	759	280	68	62	349
7	547	52	55	57	870	415	455	813	330	67	67	349
8	547	55	55	54	845	435	410	770	322	69	71	308
9	547	53	55	56	730	350	380	693	252	68	79	254
10	543	49	55	59	710	360	350	652	253	59	61	279
11	543	53	53	57	860	440	420	765	317	66	60	321
12	540	56	51	55	825	440	385	815	319	72	66	358
13	537	50	55	56	695	360	335	629	244	60	56	269
14	537	53	49	59	800	390	410	728	260	73	55	340
15	533	51	53	56	675	335	340	570	213	64	57	236
16	527	49	51	58	675	405	270	525	193	68	56	208
17	523	51	53	53	655	385	270	693	265	68	60	300
18	517	52	48	55	810	455	355	693	292	70	66	265
19	517	43	58	54	625	320	305	576	207	64	58	247
20	507	45	53	54	680	330	350	642	249	52	58	283
21	507	49	52	51	640	345	295	564	215	54	50	245
22	503	51	46	54	650	340	310	689	274	59	56	300
23	500	49	46	55	750	390	360	638	269	66	64	239
24	497	46	48	55	605	305	300	547	201	47	54	245
25	490	46	54	47	765	365	400	662	247	73	67	275
26	487	49	50	47	570	335	235	605	252	47	52	254
27	483	41	53	51	720	325	395	596	213	68	66	249
28	470	47	51	43	680	420	260	566	250	51	50	210
29	453	42	48	46	450	220	230	547	211	45	46	245
30	410	44	46	33	495	215	280	456	169	52	54	181

예시언어(Examples in) English
적용가능 언어(Applicable Language) English
적용가능 수준(Applicable Levels) College/Tertiary

이현우

인하대학교 사범대학 영어교육과
402-751 인천시 남구 용현동 253
Tel 032-860-7857
Fax 032-865-3857
Email hylee@inha.ac.kr

이소영

인하대학교 사범대학 영어교육과
402-751 인천시 남구 용현동 253
Tel 032-860-7856
Fax 032-865-3857
Email soyoun@inha.ac.kr

Revised in Oct, 2004

Reviewed by Nov, 2004

Revised version received in Dec, 2004