

영어교육 74권 2호 2019년 여름
DOI: 10.15858/engtea.74.2.201906.3

대학 수능 영어시험 절대평가 실시: 정치적 맥락에 의해 좌우된 시험의 사례*

이영식
(한남대학교)

Lee, Young Shik. (2019). Implementing a policy of absolute grading for CSAT English in Korea: A case of politically embedded test. *English Teaching*, 74(2), 3-25.

The College Scholastic Ability Test (CSAT) has focused on relative ranks among students since it started from 1993. In December 2014 The Ministry of Education (MOE) announced that it would adopt an absolute grading system (AGS) for English scores in the CSAT, starting from November 2017, and discard the relative grading. The MOE hoped that the AGS of CSAT English would eventually serve to normalize English education by having teachers place more emphasis on improving students' communication skills rather than on solving CSAT questions. This paper describes the rationale behind the implementation of AGS for CSAT English by the MOE and how the criteria and grades of AGS have been developed and set. Then current problems arising from the AGS for CSAT English are discussed by touching on its theoretical, practical and political issues based on relevant literatures, public hearings and news media reports, and informal interviews with contemporary teachers of English.

Key words: 대학수학능력시험(College Scholastic Ability Test: CSAT), 절대평가 (absolute grading system: AGS), 준거참조평가(criterion-referenced test), 정치적 맥락(political dimension)

1. 서론

교육부는 보도자료(2014년 12월 26일)를 통하여 2018학년도 대학수학 능력시험(이하 수능)부터 영어 영역에 절대평가를 도입한다고 발표하였다. 그러한 보도자료의 내용은 다음과 같다.

* 본 논문은 2016년도 한남대학교 학술연구조성비 지원에 의하여 연구되었음.

수능시험의 목적은 학교 교육과정에 따른 학습수준과 학습량을 이수하였는지 평가하여 학생의 수학능력을 측정하기 위함이나, 학생간의 상대적 서열을 중시하는 상대평가 체제의 현행 수능 영어 평가 방식은, 성적향상을 위한 무한경쟁을 초래하여 교육과정의 범위와 수준을 넘는 과잉학습이 유발되는 문제가 있었다. 또한 학교 현장에서 학생들의 영어능력을 실질적으로 향상시키기 위한 수업보다 수능 대비를 위한 문제풀이 위주의 수업이 이루어지기 때문에 균형 있는 영어능력 향상에 한계가 있었으며, 학생을 변별하기 위해 난이도가 높은 문제를 출제하는 경향이 나타나, 불필요한 학습 부담과 사교육비 부담이 초래된다는 지적도 많았다. 수능영어의 평가방식을 절대평가로 전환하는 취지는 단순히 높은 수능 점수를 받기 위한 학생과 학교현장의 무의미한 경쟁과 학습 부담을 경감시킴으로써, 의사소통 중심의 수업 활성화 등 학생들의 실제 영어 능력을 향상시키는 방향으로 학교 영어교육이 정상화되는 계기를 마련하기 위함이다.

그러나 2017년 11월 및 2018년 11월에 실시된 수능 영어시험뿐만 아니라, 앞으로 본론에서 언급될 수능영어시험에 관련된 당사자들(stake-holders)의 논의와 언론보도를 살펴보면, 본래 교육부가 의도하였던 수능영어 절대평가가 여러 문제가 있는 것을 알 수 있다. 이에 따라 본 논문에서는 최근 실시된 수능영어 절대평가에 대하여 우선 이론적 논의를 진행하고, 수능영어 절대평가의 공청회 논의와 토론 자료, 언론보도 내용 및 기타 통계 자료 등을 참고하여 실제적인 문제를 논의하여 현재 수능영어 절대평가가 안고 있는 문제를 파악하고자 한다.

2. 수능 영어 절대평가의 이론과 현실

2.1. 수능 영어 절대평가의 목표 성취 수준의 서술

현재 실시되는 수능영어 절대평가에 대한 여러 문제를 논하기 이전에, 우선 수능영어 절대평가의 바탕이 되는 준거참조평가를 기존의 상대평가 방식이었던 규준참조평가와 구분하여 논할 필요가 있다. 규준참조평가(norm-referenced testing)는 흔히 상대평가라고 일컫기도 하는데, 평가의 결과를 수험자들끼리 상호 비교하여 순위나 등수를 내는 방식의 평가이다. 상당히 큰 표본 집단으로 정규분포를 가정하고 있으며, 수험자 전체 집단의 평균점수를 기준(norm)으로 각 수험자의 점수를 비교, 분석하고 해석한다(한국교육학회 교육평가연구회, 1995; Bae, 2014). 반면에 준거참조평가(criterion-referenced testing)는 한 수험자의 지식이나 수행능력을 사전에 구체화된 수행준거 또는 목표에 의해서 직접 해석

할 수 있는 측정치를 제공할 수 있도록 제작된 평가이다. 수험생의 성취수준을 일정한 준거에 따라 판단하는 평가 방법으로, 수험자의 위치를 어떤 수행 표준 또는 준거에 비추어 확인하도록 제작된다. 준거참조평가의 목적은 한 수험자가 무엇을 할 수 있는가를 알려주는 데에 있으며, 다른 수험자와의 비교에 있지 않으므로 준거참조평가는 정규분포(normal distribution)를 가정하지 않는다(한국 교육학회 교육평가연구회, 1995; Brown & Hudson, 2002; Davidson, 2012; Hudson, 2014).

준거참조평가의 예로써 유럽공통참조기준(Common European Framework of Reference: CEFR)과 American Council on the Teaching of Foreign Languages(ACTFL)를 들 수 있다. 이러한 CEFR이나 ACTFL의 준거참조평가 모형은 다음과 같은 특징을 지니고 있는데, 상당히 오랜 기간 동안(30년 이상)의 많은 연구와 타당도 검증을 통하여 이룩하였으며, 그러한 결과로 각 등급에 맞는 능력에 대한 서술이 자세하게 되어 있어서 언어 학습자들이 성취하고자 하는 목표 수준을 미리 파악할 수 있게 배려하였다. 우선 CEFR 영역 중 읽기(reading) 등급을 살펴보면 표 1과 같다(Council of Europe, 2001, p. 69).

표 1
유럽공통참조기준의 독해력 준거

Level	Description
C2	Can understand and interpret critically virtually all forms of the written language including abstract, structurally complex, or highly colloquial literary and non-literary writings. Can understand a wider range of long and complex texts, appreciating subtle distinctions of style and implicit as well as explicit meaning.
C1	Can understand in detail lengthy, complex texts, whether or not they relate to his/her own area of speciality, provided he/she can reread difficult sections.
B2	Can read with a large degree of independence, adapting style and speed of reading to different texts and purposes, and using appropriate reference sources selectively. Has a broad active reading vocabulary, but may experience some difficulty with low frequency idioms.
B1	Can read straightforward factual texts on subjects related to his/her field and interest with a satisfactory level of comprehension.
A2	Can understand short, simple texts on familiar matters of a concrete type which consist of high frequency everyday or job-related language.
A2	Can understand short, simple texts containing the high frequency vocabulary, including a proportion of shared international vocabulary items.
A1	Can understand very short, simple texts a single phrase at a time, picking up familiar names, words and basic phrases and rereading as required.

반면에 ACTFL은 원래 Novice, Intermediate, Advanced 및 Superior로 크게 4단계로 구분하고, 각 단계를 각각 high, mid, low로 세분하여 총 10등급으로 설정하여 놓았으며, 2012년부터는 최상위에 Distinguished 등급을 추가로 설정하여 총 11등급이 되었다(ACTFL, 2012). 이러한 ACTFL 지침은 각 등급에 맞는 4기

능(speaking, writing, listening, reading)의 능력을 상세하게 기술하여 놓았다(자세한 내용은 웹사이트 참고: www.actfl.org). 또한 ACTFL은 세 등급(Novice, Intermediate, Advanced)의 수행 수준 척도(performance descriptors)를 상세하게 기술하고 각 수준의 예를 들어주어 언어 학습자로 하여금 자기에 맞는 언어 능력 수준을 파악하여 성취할 수 있도록 배려하였다(ACTFL, 2015).

위와 같이 수능 절대평가는 절대기준평가로서 규준참조평가가 아닌, 준거참조평가를 의미하므로, 사전에 학생들이 성취하여야 할 기준-준거(criterion) 또는 표준(standard)-을 미리 구체적으로 제공함으로써 학생들은 그 기준을 성취할 수 있도록 배려하여야 한다. 또한 절대평가는 평가 기준에 대해 오랜 기간 동안의 심층적인 연구와 검증을 하여야 하며, 특히 수능영어 절대평가는 고등학교 교육과정의 성취도 시험으로 대학을 입학한 후의 목표언어사용(target language use: TLU)을 고려하여야 한다(Bachman & Palmer, 1996, 2010). 그러나 수능영어 절대평가가 2014년 12월에 발표되고, 아주 짧은 기간이 지난 2017년 11월부터 실제 수능영어 절대평가를 실시하다 보니, 결국 현재 수능영어 시험이 정부(교육부) 주도하에 단기간 동안 이루어진 하향식 절대평가 체제가 되고 말았다. 또한 수능영어 절대평가의 기준에 필요한 영어 교육과정의 성취도와 대학 입학 후 목표언어사용의 구체적인 준거에 대한 이론적 및 실제적 규명이 없다 보니, 수능영어 절대평가가 추구하는 각 등급의 서술이 전혀 없게 되었다. 우리의 수능영어 절대평가는 위에서 언급한 CEFR이나 ACTFL과는 달리 학생들이 성취하여야 하는 목표 수준의 영어능력에 대한 정의가 되어 있지 않으며, 이러한 결과로 학생들은 자기가 성취하여야 하는 영어 능력에 대한 실제적 윤곽도 제대로 파악하지 못한 채로 수동적으로 수능영어 시험을 치를 수밖에 없는 현실이 되었다.

2.2. 절대 기준 및 등급을 명확히 설정하지 못한 채로 시험 실시

대학입학을 위한 수능영어시험은 그 시험이 수험자의 일생을 좌우할 수 있는 고부담 시험(high-stakes exam)이므로 점수기준이 정확 하여야 하며, 아주 엄정한 절대기준 등급체제를 확립하여야 한다. 특히 수능영어 시험이 본래의 절대평가의 취지를 갖기 위해서는 수험생들에게 시험의 결과로 성취 수준(등급)이 부여 되어야 하며, 수험생들에게 등급이 부여되기 위해서는 각 수험생들의 점수에 따라 등급을 구분할 수 있는 분할점수가 정해져야 된다. 한국교육과정평가원(2014)에서는 공청회를 통하여 수능영어 절대평가에 대한 점수체제의 기초연구(강태중 등, 2014)를 진행하였다고 하지만, 그러한 점수체제에 대한 연구는 탐색적인 논의로 그쳤다. 박찬호(2014)는 수능 영어영역 절대평가 점수체제로 (1) 분할점수를 미리 정하고 그에 맞게 출제하는 고정분할점수 방식, (2) 전문가의 내용 분석에 의거하여 준거를 설정하는 내용 분석에 의한 준거설정 방식, (3) 고정분할점수 방식이나 내용 분석에 의한 준거설정을 기본으로 하면서 단점을

보완하는 혼합 방식을 탐색하여 제시하였지만, 실제로 수능영어 절대평가에 맞는 구체적인 점수체제 방안을 제시하였다고 보기는 어렵다(김준식 등, 2015; 한국교육과정평가원, 2014). 이렇게 공청회를 통한 제시나 논의가 수능영어 절대평가에 맞는 현장 연구와 검증을 이룬 것은 결코 아니어서 수능시험 영어영역의 절대평가에 맞는 타당한 연구를 수행하였다고 말할 수 없다!

등급 설정에 대해서도, (1) 5등급 고정 분할 안, (2) 9등급 고정 분할 안, (3) 5등급 준거설정 안, (4) 9등급 혼합방식 안이 논의되었다가, 한국사 절대평가 점수체제와 일관성을 유지하여 최소화하며 기존 수능시험과 같이 수능영어 절대평가를 실시한다는 측면에서 결국 9등급 고정분할 방식이 제안되었다(김준식 등, 2015). 이렇게 우여곡절 속에 수능영어 절대평가의 등급이 결국 현재의 9등급으로 설정되었는데, 본래 수능영어 시험 절대평가에 필요한 각 등급의 기준(준거 또는 표준)이 제대로 확립되지 않았고, 또한 각 등급에 필요한 분할 점수도 제대로 규명이 안된 절대평가가 되고 말았다. 막연한 기준으로 설정된 9등급 체계로 수능영어 절대평가는 2017년 11월과 2018년 11월에 시행되게 되었는데, 그러한 9등급 체계가 종래의 상대평가 9등급 체계와 다른 점은 거의 없고 단지 고득점자만 많이 배출한 결과만 초래하였다(한국교육과정평가원, 2017, 12월 12일, 2018, 12월 5일).

이와 같이 수능영어 시험의 기준과 등급 체제에 대한 제안과 논의(강태중, 2014; 강태훈, 2014; 박찬호, 2014)는 있었으나, 수능영어 절대평가의 기준에 필요한 영어 교육과정의 성취도와 대학입학 후 목표언어사용의 구체적인 준거뿐만 아니라 영어능력의 등급 체계 및 분할 점수에 대한 적합한 연구와 타당한 검증이 있었다고 볼 수 없다. 현재의 수능영어 시험은 절대기준 설정이 없이 무늬만 절대평가가 될 수 있다는 지적(김준식 등, 2015; 박찬호, 2014)도 있었다. 이러한 수능영어 평가 방식은 절대평가가 아니라 쉬운 수능 영어 비율 등급제로 결국 상대평가라고 할 수 있으며(송옥진, 2018, 11월 19일; 이영식, 2018; 정혜선, 2019), 결국 현재까지 두 차례에 실시된 이러한 수능영어 절대평가 실시는 많은 혼란과 문제를 야기할 수밖에 없는데, 이러한 논의를 다음 장에서 구체적으로 하기로 한다.

2.3. 대부분의 주요 당사자에게 중요하지 않은 영어시험

항상 대규모 표준화 시험은 그 시험에 대하여 당사자를 가지고 있다. 그러한 대규모 시험에 대해서는 학생, 교사, 학부모, 입학전형을 실시하는 대학 및 교육부 등의 정치 및 사회적 기관이 당사자가 될 수 있으며, 특히 수능영어시험

¹ 박찬호(2014)는 ‘절대평가의 도입만으로 모든 문제가 해결될 것이라는 낙관론이나 아직 확정되지 않은 구체적인 방안을 전제로 하는 비관론 모두 경계의 대상이다’라고 말하였다.

은 수험자의 대학 입학전형 목적으로 하므로 수능영어시험에 대해서는 수험자와 입학전형을 실시하는 대학이 주요 당사자가 된다. 그러나 2018학년도(2017년 11월 실시) 절대평가의 수능영어시험 결과는 주요 당사자인 상당히 많은 대학에게는 무의미한 영어시험이 되었다고 볼 수 있다(박세인, 2017, 12월 11일). 2019학년도 수능영어시험(2018년 11월 실시)도 지난해에 이어 절대평가로 치러졌으나, 1등급(원점수 90점 이상) 비율이 5.3%로 지난해 1등급 비율(10.03%)의 절반이 되어 지난 해보다 어려운 시험이라고 할 수 있다. 그럼에도 불구하고 영어 시험이 국어 및 수학 보다는 여전히 상급 등급의 비율이 많은 시험이 되어서, 수험자와 대학의 입장에서는 영어보다는 변별력이 많은 다른 과목에 더 많은 비중을 줄 수밖에 없었다(박세미, 2018, 12월 5일; 신혜정, 2018, 12월 4일). 결국 쉬운 수능영어시험의 결과가 대학입학 전형자료로서는 합당하지 않으므로, 상당히 많은 대학이 수험자들을 전형할 때 수능 영어시험 결과 보다는 오히려 타 과목의 수능시험 결과를 치중하여 고려하게 되었다. 특히 상위권 대학에서 수능 영어 성적을 입학 전형자료로 충실하게 사용할 수 없게 되니, 이러한 수능영어 절대평가 시험 결과는 국어 및 수학과는 달리 상위권 학생들과 상당히 많은 대학에게는 결국 그리 중요하지 않은 시험이 되었다(박세인, 2017, 12월 11일; 한국영어관련학술단체 협의회, 2018; 한국영어학학회, 2018).

2.4. 불안정한 수능영어 시험 난이도의 변동

절대평가에서는 등급 기준과 그에 따른 분할 점수가 엄정해야 할 뿐만 아니라, 각 시험마다 변하지 않는 일정한 난이도가 더욱 중요하다. 그러나 수능영어 절대평가에 대한 기준 설정이 분명하지 않고, 각 등급 점수도 10점대로 임의적이다 보니, 표 2에서 보는 바와 같이 지난 2년간 치렀던 수능영어 시험의 난이도가 일정하지 않게 되었다. 2018학년도 수능영어 시험에서는 1등급이 10.03%, 2등급이 19.65%로 상위 1-2등급이 29.68%(약 30%)이며, 2019학년도 수능영어 시험에서는 1등급이 5.30%, 2등급이 14.34%로 상위 1-2등급이 19.64%(약 20%)이다. 수능영어시험이 절대평가를 지향한다면 2018학년도 난이도와 2019학년도 난이도가 동일하거나 거의 비슷하여야 하는데, 2019학년도 상위 1-2등급이 2018학년도 상위 1-2등급의 2/3에 불과하였다.

이와 같이 일정하지 않은 난이도에 대한 언론보도도 살펴볼 필요가 있는데, 2018학년도 수능 영어시험이 기대이상으로 쉬워짐에 따라 많은 수험생들이 혼란을 겪었고(김기홍, 2017, 12월 12일), 2019학년도 수능 영어시험은 어려워져서 1등급 비율이 전년도에 비해 절반에 그치다 보니 수험자의 혼란은 가중되었다(박세미, 2018, 12월 5일). 특히 “절대평가 도입의 취지를 살리려면 일정한 난도가 핵심인데, 학생들 입장에선 시험이 너무 쉬웠다가 너무 어려웠다가 하며 갈피를 못 잡으니 혼란만 가중되고 있다”라는 언론보도(송옥진, 2018, 11월 19일)도 있었다.

표 2

영어 영역 등급 구분 원점수, 등급별 인원 및 비율(2018학년도 및 2019학년도)

등급	등급 구분 점수	2018학년도		2019학년도	
		인원(명)	비율(%)	인원(명)	비율(%)
1	90	52,983	10.03	27,942	5.30
2	80	103,756	19.65	75,565	14.34
3	70	134,275	25.43	97,577	18.51
4	60	94,871	17.97	110,176	20.91
5	50	55,391	10.49	87,117	16.53
6	40	35,576	6.74	56,219	10.67
7	30	25,050	4.74	39,136	7.43
8	20	18,441	3.49	24,179	4.59
9	20미만	7,721	1.46	9,114	1.73

출처: 한국교육과정평가원, 2017, 12월 12일; 한국교육과정평가원, 2018, 12월 5일

이상에서 보는 바와 같이, 절대평가에서는 여러 시험에 걸쳐 변하지 않은 일정한 난이도가 중요한데, 2회에 걸쳐 실시된 수능영어 절대평가에서도 난이도의 변화는 심하였다. 이와 같이 절대평가라고 일컫는 수능영어 시험의 난이도가 일정하지 않고 심하게 변화하다 보니, 과연 현재 수능영어 절대평가가 상대평가와 다르다고 할 수 있는지 많은 수능영어 시험 당사자들이 의문을 품지 않을 수밖에 없고, 본래의 절대평가 도입 취지가 무색 해졌다고 할 수 있다(송옥진, 2018, 11월 19일; 이영식, 2018).

3. 수능 영어 절대평가의 갈등

3.1. 영어 절대평가의 부정적 환류 효과

3.1.1. 주요 과목 중 덜 중요시되는 영어 교육

대규모 표준화 시험으로서 수능영어 절대평가가 엄정하지 않은 채로 진행되다 보니 심한 후유증을 야기하고 있다고 하여도 과언이 아니다. 한국영어관련 학술단체협의회(이하 한국영단협)에 의하면, 현재 수능 영어 시험을 쉽게 출제하여 고득점자를 많이 하는 수능영어 절대평가 정책으로 인하여 영어 학력 저하에 대해 많은 영어교사들과 영어 교수들이 심각한 우려를 표명하고 있다(한국영단협, 2018). 특히 어느 현직 영어교사는 다음과 같은 우려를 표명하였다(한국교육과정평가원, 2018, 12월 19일).

대입 수능시험에서 영어과목만 절대평가로 실시되니 상당히 많은 학생들이 수능영어 시험이 쉬울 것이라고 생각하면서 영어 공부를 제대로 하지 않고, 반면에 국어나 수학 같은 과목은 더 어렵게 출제될 것이라는 기대감으로 국어와 수학에 더 치중하여 공부하는 것이 문제라고 할 수 있습니다. 이런 상황에서는 대입 수능시험에서 영어 과목만 절대평가를 실시하지 않고 국어 및 수학에서도 절대평가를 실시하든지, 또는 수능영어 시험을 종전처럼 상대평가로 환원하여야 한다고 생각합니다.

비록 2019학년도 수능영어시험의 1등급 비율이 적어서 이전의 시험보다 어려웠지만, 여전히 국어 및 수학 보다는 쉬운 시험이라고 볼 수 있다. 위와 같은 내용을 종합하여 볼 때, 현재 고등학교 학생들이 주요 과목 중에서 영어 과목을 이전 보다 공부를 많이 하지 않을 뿐만 아니라, 결국 영어 학력 저하를 가져올 수 있는 현실에 대해 상당히 많은 영어 교육자들이 우려하고 있다고 볼 수 있다(정지윤, 2018, 8월 8일).

3.1.2. 고등학교의 영어 수업 주당 시수 감소

표 3에서 보는 바와 같이 지난 2년 동안 고등학교에서의 영어 수업 시수가 국어 및 수학 과목에 비해 감소하였다.

표 3
인문계 고등학교 주당 평균 영어 수업시간 변화²

	1학년	2학년	3학년
2017학년도 이전	4시간	5-6시간	6-8시간
2018학년도 이후	4시간	4시간 + 국어 영어 수학 학교 별 또는 학생별 선택 과목으로 과목간 2-3 시간 차이가 있을 수 있음	4시간 + 국어 영어 수학 학교 별 또는 학생별 선택 과목으로 과목간 2-3 시간 차이가 있을 수 있음

2017학년도 이전에는 고등학교 영어 수업 주당 시수가 1학년에서는 4시간, 2학년에서는 5~6시간, 3학년에서는 6~8시간이었는데, 2018학년도 이후에는 고등학교 영어 수업 시수가 1,2,3학년 모두 평균적으로 4시간으로 줄어들었다. 물론 각 학교 별로 또는 학생 별로 고등학교 2~3학년에서는 국어, 영어, 수학 중에서 선택 과목을 2~3시간 늘려서 실시하는 것을 고려한다고 하여도, 고등학교 2~3

² 표의 통계는 학교알리미(www.schoolinfo.go.kr) 홈페이지에서 무작위로 전국 고등학교의 수업 시수를 검색하고, 또한 주위의 현직교사들을 면담하여 확인한 결과이다.

학년에서의 영어 수업 시수는 국어 및 수학 수업 시수보다 감소된 것은 분명하다. 결국 이러한 통계는 공교육에서 영어 공부를 중요시하지 않고, 오히려 다른 주요 과목인 국어와 수학을 더 중요하게 여긴다는 현실을 반영하는 것을 이라고 볼 수 있다. 그 결과로 많은 현직 영어교사 및 교수들은 수능영어 절대평가의 부정적 효과에 대해 우려를 표명하고 있다(한국영단협, 2018).

3.1.3. 타 주요 과목과의 갈등

진술한 바와 같이, 대학 수능시험에서 영어가 변별력이 없으므로, 국어와 수학에서 변별력을 높여야 한다고 많은 사람들이 우려하고 있다(정혜선, 2019; 한국영단협, 2018). 실제 그러한 우려를 반영하듯이, 두 번 치른 수능시험에서 국어 및 수학의 난이도와 변별력이 영어보다 높은 결과가 나왔으며, 이러한 우려가 주요 일간지에 보도되었다(박세미, 2017, 12월 11일; 박세인, 2017, 12월 11일; 송옥진, 2018, 11월 19일; 신혜정, 2018, 12월 4일). 본 연구자는 현직 고등학교 교사와 면담할 기회를 가졌는데, 실제로 고등학교에서 교육을 담당하는 교육자들은 영어교육을 경시하고, 오히려 다른 과목(국어와 수학)을 중시하게 되었고, 국어와 수학이 영어 대신 변별력을 높여야 한다는 부담감을 안고 있는 현실이다(이영식, 2018, 7월 31일; 정혜선, 2019). 이러한 현실은 지난 몇 년 동안의 기초과목 교사들의 채용 현황을 살펴보면 알 수 있다(표 4 참조).

표 4

영어 절대평가가 결정된 2014년 이후 공립 중등 영어교사 및 타 교과 교사 채용 인원: 2014학년도 대비율

	2014년	2015년	2016년	2017년	2018년	2019년
영어	672명 (100%)	528명 (76.8%)	420명 (62.5%)	311명 (46.3%)	263명 (39.1%)	268명 (39.9%)
국어	583명 (100%)	516명 (88.5%)	426명 (73.1%)	375명 (64.3%)	362명 (62.1%)	322명 (55.2%)
수학	567명 (100%)	485명 (85.5%)	421명 (74.3%)	375명 (66.1%)	334명 (58.9%)	291명 (51.3%)

주. 각 시도교육청 공립 중등학교 교사 임용후보자 선정경쟁시험 시행계획 공고 참조.

2014년부터 2018년까지 영어 과목 교사들뿐만 아니라 국어 및 수학 과목 교사들도 감소하였는데, 이러한 교사들의 전반적인 감소는 학령인구의 감소로 기인한 것으로 볼 수 있다. 그러나 위 통계를 구체적으로 살펴보면, 국어 및 수학 과목 교사와 비교하여 영어 과목 교사의 채용이 가장 많이 감소하였다. 즉, 중등학교에서는 학생들의 대학 수학능력 시험 준비를 위하여 영어 교사보다는 국어 및 수학 교사의 채용에 더 많은 관심을 두었다고 볼 수 있다. 결국 이러한 통계는 고등학교의 주요 과목 교육에서 영어보다는 국어 및 수학에 더 많은 중

점을 두었으며, 상대적으로 영어 과목 교육은 소홀히 한다는 것을 의미한다.

아울러 우리는 2018학년도 및 2019학년도 수능능력시험의 영어, 국어 및 수학의 채점 결과를 살펴볼 필요가 있다(한국교육과정평가원, 2017, 12월 12일; 한국교육과정평가원, 2018, 12월 5일).

그림 1
2018학년도 및 2019학년도 수능시험 과목별 1-2등급 비율

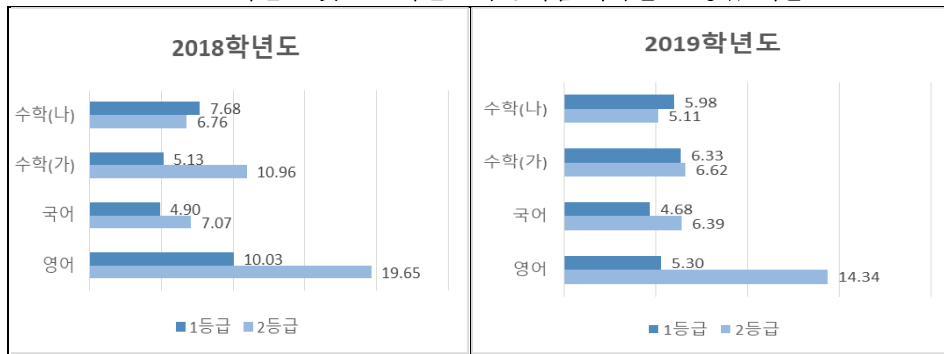


그림 1에서 보는 바와 같이, 국어에서 1등급이 4.90%, 2등급이 7.07%를 차지하였고(1-2등급 11.97%), 수학(가)에서 1등급이 5.13%, 2등급이 10.96%를 차지하였으며(1-2등급 16.09%), 수학(나)에서 1등급이 7.68%, 2등급이 6.76%를 차지하였고(1-2등급 14.44%), 영어에서 1등급이 10.03%, 2등급이 19.65%를 차지하였다(1-2등급 29.68%). 2019학년도 수능시험에서는 국어에서 1등급이 4.68%, 2등급이 6.39%를 차지하였고(1-2등급 11.07%), 수학(가)에서 1등급이 6.33%, 2등급이 6.62%를 차지하였으며(1-2등급 12.95%), 수학(나)에서 1등급이 5.98%, 2등급이 5.11%를 차지하였고(1-2등급 11.07%), 영어에서 1등급이 5.30%, 2등급이 14.34%를 차지하였다(1-2등급 19.64%).

2018학년도 및 2019학년도 수능시험 주요 과목의 상위 1등급 및 2등급의 비율을 살펴볼 때, 영어의 상위 1-2등급의 비율이 국어 및 수학의 상위 1-2등급 비율보다 훨씬 높아서 수능시험 영어 영역은 지난 2년에 걸쳐서 국어 및 수학 영역 보다 훨씬 쉬운 것으로 드러났다. 이러한 수능시험의 영어 영역과 국어 및 수학 영역 간의 등급 비율의 불균형은 결국 영어 영역만 절대평가를 실시하여 야기된 것이라고 볼 수 있다. 그 뿐만 아니라, 절대평가를 치른 영어 영역에서도 지난 2년간 상위 1등급 및 2등급 비율이 차이가 있는 것을 알 수 있다. 대학입학 수능시험에서 영어 영역이 절대평가가 실시 되지 불과 2년밖에 안되었지만, 대학입학 수능시험에서 타 기초과목 영역과 영어 영역간의 상위 등급 비율의 큰 차이뿐만 아니라, 2년간 절대평가로 치른 수능영어 시험만의 현저한 난이도 차이가 현재 수능영어 절대평가가 안고 있는 문제라고 할 수 있으며, 결국 이러한 수능영어 절대평가의 타당성 및 유용성에 대해서도 심각한 의문을

제기할 수 밖에 없다(송옥진, 2018, 11월 19일; 양선아, 2018, 12월 4일; 한국영단협, 2018, 6월 22일).

3.2. 감소되었다고 볼 수 없는 사교육비

교육부 및 통계청은 초등·중·고등학교 수준의 1,484개 학교에서 40,000명 학부모를 대상으로 2017년 초·중·고 사교육비를 조사한 통계를 발표하였으며, 이러한 통계에 근거하여 한겨레신문 등의 주요 언론들은 표 5와 같이 학생 1인당 월평균 사교육비를 보도하였다(김미향, 2018, 3월 16일; 이도경, 2018, 3월 16일).

표 5
학생 1인당 월평균 사교육비*

년도	액수(단위: 천원)	년도	액수(단위: 천원)
2007	222	2013	239
2008	233	2014	242
2009	242	2015	244
2010	240	2016	256
2011	240	2017	271
2012	236		

주: 교육부 및 통계청(2018.3.15) '2017년 초·중·고 사교육비 조사'

또한 학생 1인당 월평균 사교육비를 초등학교, 중학교, 고등학교로 구분하여 표 6과 같은 통계 결과를 제시하였다.

표 6
초·중·고 학생 1인당 월평균 사교육비(단위: 천원)

년도	2013	2014	2015	2016	2017
평균	239	242 (+1.1%)	244 (+1.0%)	256 (+4.8%)	271 (+5.9%)
초등학교	232	232 (+0.0%)	231 (-0.4%)	241 (+4.5%)	253 (+4.8%)
중학교	267	270 (+1.2)	275 (+1.9%)	275 (-0.1%)	291 (+5.7%)
고등학교	223	230 (+2.9%)	236 (+2.9%)	262 (+10.9%)	284 (+8.4%)

표 6을 살펴 보면, 우선 중학생 사교육비가 가장 많은 것으로 밝혀졌고(291,000원), 2017년 고등학교 학생 1인당 월평균 사교육비는 284,000원으로, 전년도 2016년 대비 8.4% 증가하였다. 이러한 언론보도에 의하면, 수능시험의 학생 1인당 월평균 사교육비는 198,000원이며, 전년도(2016년) 대비 3.4% 증가하였다. 수능시험 과목 사교육비는 국어 14.2% 증가, 사회탐구 및 과학탐구 8.5%

증가, 수학 3.3% 증가, 영어 0.5% 증가로 나타났다. 수능시험에서 영어 영역이 절대평가로 전환되자 국어와 사회 및 과학의 다른 과목으로 ‘풍선 효과’가 나타난 것(이도경, 2018, 3월 16일)으로 풀이할 수 있지만, 수능영어 절대평가 실시로 교육부가 의도한대로 사교육비(특히 사교육비 총량)가 감소되었다고는 볼 수 없다. 특히 영어 사교육비는 유아 및 초등학생들이 중·고등 학생들보다 사교육비가 많은 것을 고려한다면(한국영단협, 2018), 과연 수능영어 사교육비 0.5%증가가 수능영어 절대평가에 의한 소폭 증가로도 단정하기에는 어려울 수 있다. 또한 수능영어시험을 쉽게 출제함으로써 영어로 인한 학생들 간의 경쟁은 감소하였다고 일부 주장할 수 있으나, 오히려 타 과목에서 경쟁을 부추긴 결과가 되었으며 오히려 1인당 사교육비 총액은 증가한 것으로 밝혀졌다. 따라서 영어영역만 절대평가를 실시함으로써 학생들의 사교육비를 줄이고자 하는 정부의 의도는 상당히 현실성이 없었다고 볼 수 있다.

3.3. 수능 영어 절대평가의 윤리성

AERA(American Educational Research Association), APA(American Psychological Association) 및 NCME(National Council on Measurement in Education)는 수험자의 권리와 책임에 대한 표준(standards)을 다음과 같이 규정하였다(2014, pp. 133-135).

8. Standards for Test Takers' Right and Responsibilities

Cluster 1. Test Takers' Rights to Information Prior to Testing

Standard 8.2: Test takers should be provided in advance with as much information about the test, the testing process, the intended test use, test scoring criteria, testing policy, availability of accommodations, and confidentiality protection as is consistent with obtaining valid responses and making appropriate interpretations of test scores.

위와 같은 표준에 따르면, 대규모 표준화 시험은 수험자들에게 알아야 할 정보를 미리 구체적으로 제공함으로써 수험자의 권리를 지켜주도록 최선을 다하여야 한다. 특히 절대기준 평가(즉, 준거참조평가)는 수험자들에게 미리 평가기준을 정확히 알려주어서, 수험자들이 그 기준을 성취할 수 있도록 배려하여야 한다. 그러나 현재 수능영어 절대평가는 너무 정치적인 일정에 의해 진행되다 보니 절대평가에 대한 타당한 연구 및 검증이 없었고, 시험에 대한 다양한 정보뿐만 아니라 절대평가 기준을 수험자에게 사전에 전혀 제공하지 않았으며, 결국 이러한 문제는 미국 AERA, APA 및 NCME에서 공동으로 규정한 평가 표준을 어겼다고 볼 수 있다. 이렇게 미국 AERA, APA 및 NCME에서 공동으로 규정한 수험자의 권리와 책임 기준을 현재 수능영어 절대평가가 지키지 못하였다고 판단된다면, 결국 수능영어 절대평가는 수험자의 권리와 책임에 대한 윤

리성에도 저촉될 소지가 있다.

3.4. 수능 영어 결과 타당도(consequential validity)

2014년 12월 26일 교육부의 보도 자료에 의하면, 수능영어 절대평가를 실시하는 본래 교육부의 의도는 (1) 사교육 억제, (2) 학교 영어교육 정상화, (3) 학생 상호간의 불필요한 경쟁 해소라고 볼 수 있다. 그러나 2018학년도 및 2019학년도 수능영어 절대평가의 결과를 살펴보면, 원래 교육부가 의도한 바와 다른 결과가 나왔고, 본래 절대평가의 취지가 무색하였다고 볼 수 있다(송옥진, 2018, 11월 19일). 우선 사교육비가 감소되지 않았으며, 수능영어 절대평가에 의하여 학교 영어 교육 정상화는 이루어지지 않았다는 것을 쉽게 알 수 있다(정혜선, 2019). 결국 정부에 의해 주도된 교육정책 집행이나 단순히 시험의 변화를 통해서만 학교 교육의 정상화를 이룩하고자 하는 것은 현실적으로 타당하지 않다는 것을 다시 한번 확인할 수 있게 되었다(Meier & Wood, 2006). 특히 영어 과목만 절대평가를 실시하고, 국어 및 수학과 같은 과목에서는 상대 평가를 실시하다 보니 국어 및 수학 과목과 영어 과목간의 불균형으로 인하여 학생 상호간의 경쟁은 더욱이 해소되지 않았을 뿐만 아니라, 수능 영어영역의 절대평가도 학생 상호간의 경쟁을 완전히 해소하였다고 볼 수 없다.

지난 세기 말에 평가에 대한 결과 타당도라는 개념이 대두되었는데, 이러한 개념은 본래 의도된 평가 기준을 측정하는 정확성, 그러한 정확성이 수험자의 시험 준비에 미치는 효과, 시험의 해석 및 사용에 대한 사회성 및 윤리성 등을 모두 포괄하는 시험 결과를 의미한다(Messick, 1998; Weir, 2005). 전술한 바와 같이, 수능영어 절대평가의 결과를 살펴보면, 수능영어 절대평가의 도입 및 실시의 본래 취지와 타당도—특히 결과타당도—에 대해서 많은 의문의 여지가 있게 되었다.

지금까지 수능영어 절대평가에 대해 진행된 논의를 살펴보면, 절대기준평가의 이론적 타당성 및 현실성이 부족하여 많은 모순을 지니고 있으며, 무늬만 절대평가인 쉬운 상대평가로 그치고 말았다. 이러한 수능영어 절대평가는 교육부에 의해 주도되어 제대로 된 이론적 연구 및 타당도 검증 없이 하향적으로 진행된 대규모 표준화 시험(large-scale standardized test)의 대학입학시험에 불과하게 되었다. 이러한 상황에서 우리는 McNamara와 Roever(2006, p. 224)의 말을 유념할 필요가 있다.

Test constructs are the site of struggle over values, a struggle that was not resolvable by validity research, but through normal political process.

이러한 맥락에서 볼 때, 수능영어 절대평가는 타당하고 합리적인 연구에 의해서 진행된 것이 아니라, 정치적인 고려에 의해 진행되었다고 볼 수 있다. 그

결과로 수능영어 절대평가는 여러 갈등을 안게 되었고, 그러한 수능영어 절대평가의 갈등도 타당한 연구에 의해 해결되는 것으로 기대하기 보다는 정치적인 맥락에서 해결될 수 있는 논란의 소지가 있게 되었다. 결국 현재의 수능영어 절대평가는 정치적 맥락에 좌우된 대규모 시험에 불과하게 되었다.

4. 제언 및 결론

4.1. 제언

교육부에 의해 2014년 12월에 보도되어 2017년 11월 및 2018년 11월에 실시된 수능영어 절대평가를 살펴보면, 전술한 바와 같이 많은 문제를 야기하였다고 볼 수 있다. 이에 따라 현시점에서 우리는 다음과 같은 두 가지 대안을 고려하여 볼 수 있다.

- (1) 수능의 영어뿐만 아니라 주요 타 과목에 대해서도 절대 평가 실시
- (2) 수능 영어 절대평가 폐지 및 상대평가로의 환원

특히 (1)안에 대해서는 한국교육과정평가원이 주최한 공청회에서 언급되었기 때문에 상당히 가능성 있는 제안이 될 수 있었다(강태중 등, 2014; 양도웅, 2018, 8월 20일; 한국교육과정평가원, 2014, 10월 20일). 그러나 수능 영어에 대한 위 두 가지 방안 모두 사실상 불가능하게 되었다. 교육부(보도자료, 2018, 8월 17일)는 “2022학년도 대학입학제도 개편방안 및 고교교육 혁신 방향 발표”라는 제목으로 다음과 같이 발표하였다.

수능 평가방법의 경우, 국가교육회의 권고안을 존중하여 국어·수학·탐구는 현행 방식을 유지하되, 영어, 한국사와 과목 출립 문제가 있는 ‘제2외국어/한문’은 절대평가로 변경한다.

결국 수능영어 절대평가가 돌이킬 수 없는 현실이라면, 우리 영어교육자들은 수능영어 타당성 제고를 위한 이론적인 노력을 다음과 같이 기울일 필요가 있다고 본다.

4.1.1. 수능영어시험이 수험생의 영어구사력을 진정으로 평가하고 있는지 검증할 필요

우선 수능영어시험이 수험생의 진정한 영어능력을 제대로 평가하고 있는지 검증할 필요가 있다. 아래와 같은 도식에서 보는 것처럼, 본래 시험점수는 수험

자의 진정한 능력(true ability), 시험방법의 국면(test method facet: TMF/test-wiseness), 수험자의 특성(test-taker's characteristics: TC) 및 오류(error)의 결과라고 할 수 있다.

$$\text{Score} = \text{TA} + \text{TMF} + \text{TC} + \text{Error}$$

여기에서 수험자 특성이라는 것은 성별의 차이(남성 또는 여성), 장애인과 같은 특성, 수험자들의 사회경제적 배경(socioeconomic status: SES) 등을 의미하는데, 이러한 수험자 특성이 시험 결과(즉, 성적)에 영향을 미치는 것은 이미 알려져 있다(Abedi, 2008; Kunnan, 1995; Taylor, 2012; Weir, 2005). 한편 시험방법의 효과라는 것은 어떤 수험자의 시험 성적이 본래 수험자의 능력만을 반영하는 것이 아니라, 시험을 치르는 요령을 반영하는 것을 의미한다(Bachman, 1990, 1991). 특히 자기 일생의 진로를 결정할 수 있는 고부담 시험에서는 수험자가 자신의 본래 능력에 맞게 시험을 치르는 것도 중요하지만, 반면에 특정한 시험 문제를 그 시험의 요령에 의거하여 잘 풀거나 또는 그 시험 문항에 실수하지 않는 요령을 발휘하는 것도 중요하게 된다고 말할 수 있다.

현재 우리는 수능영어 시험에 대해서는 수험자 특성보다는 일단 시험방법의 효과에 대해 더욱 관심을 가질 필요가 있다고 본 연구자는 생각한다. 물론 시험 방법에 대처하는 요령이나 전략은 반드시 부정적인 것은 아니라고 볼 수 있다(Cohen, 1987, 2012). 그러나 이제 수능영어시험이 1993년부터 시행되어 지난 25년간 치러지면서 대부분 1지문 1문항 형식으로 출제되어 왔으며, 특히 한국에서만 실시되는 특유한 영어시험 문항도 있다. 현재 수능영어 시험은 각 문제마다 다른 유형을 지니고 있고 또한 지난 세월 동안 특별한 개선 없이 거의 똑같은 유형의 문항을 출제하여 오다 보니 그 동안 드러나지 않은 문제들이 잠재할 수도 있다. 따라서 우리는 4반세기 동안의 수능영어의 특유한 영어 문항들이 과연 학생들의 진정한 영어능력만을 타당하게 반영하고 있는지, 아니면 각 문항에 따른 방법의 효과(즉, 문제 푸는 요령)를 필요 이상으로 과도하게 반영하고 있지 않는지를 검증할 필요가 있다. 특히 수능영어 시험에서 시험 문제를 잘 해결하는 능력 못지 않게 각 문제에 대해 실수를 하지 않는 요령이 더 중요하게 수험자들이 느끼지 않는지를 검증할 필요가 있다고 본다.

현재 수능영어 입시지도로 받는 학생의 입장에서 보건대, 학생들은 본래 영어능력의 향상보다는 문제 푸는 요령에 더욱 관심을 가질 우려가 있다. 이러한 수능 영어시험의 요령 효과에 대해서는 현재까지 구체적이고 심층적인 연구는 없었지만, 주위의 현직 영어 교사나 학생들에게 문의하면 수능영어시험의 요령 효과에 대해서는 상당히 감지할 수 있다. 그 결과로 학생들의 입장에서는 학교 교육을 통하여 스스로 자기 영어능력을 향상시키는 자기 주도 학습을 하기 보다는, 수능영어시험을 잘 풀게 하는 학교 영어 교사의 지도나 학원의 사교육에 의존할 수밖에 없는 현실이 될 수 있다. 특히 고부담 시험인 수능영어 대비를

위하여 시험 요령 위주로 지도하는 사교육이 더욱 효과적이라고 많은 고등학교 학생들에게 판단된다면, 수능영어를 대비하는 학생들의 입장에서는 오히려 더욱 영어 사교육에 의존할 수 없는 현실이 될 수밖에 없다.

따라서 수능영어 시험의 각 문항들이 수험자의 영어 능력을 제대로 반영하고 있는지, 또는 수험자의 본래 영어 능력뿐만 아니라 각 시험 문제의 요령 효과(Bachman, 1990, 1991; Cohen, 1987, 2012)를 상당히 반영하고 있는지를 일단 체계적으로 검증하여 타당한 수능영어시험이 잘 실시되도록 유념할 필요가 있다. 이러한 수능영어의 시험 요령 효과에 대한 규명은 각 수험자의 시험 성적 데이터가 주어지면 수능 영어시험의 신뢰도 측면에서 이제는 더욱 정교하게 발달한 측정 도구(psychometric tool)를 이용하여 정량적 연구로 규명할 수 있고, 타당도 측면에서는 수험자 및 수능영어 지도 교사들을 면담하여 정성적인 연구로 실행할 수 있다. 결국 수능영어시험에 대한 방법/요령의 효과 규명은 수능 영어 자체에 대한 타당도 검증인데, 한편 이미 국제 공인 영어시험에서는 타당도 검증에 대한 많은 연구가 이루어졌다(Chapelle & Voss, 2014; Geisinger, 2016; Xi, 2010). 그러나 현재 이러한 연구는 본 논문의 범위를 넘으므로, 일단 수능영어의 요령의 효과에 대해서 우려만 표명한다.

4.1.2. 문제 유형의 단순화 및 개선

만약 수능 영어시험 문항이 수험자(고등학교 학생)의 진정한 영어능력을 벗어나 다른 요인들을 반영한다고 판단되고 또한 수능영어 시험이 절대평가에 포함되도록 출제되기 위해서는 현재 수능영어 시험 문항들을 개선 및 보완할 필요가 있게 된다. 수능영어의 독해 문항 방식을 1지문 1문항으로 하여 많은 유형의 문항을 복잡하게 출제하기 보다는, 여러 개별적인 문항이 합쳐져 있는 문항군(testlet)으로 개발하여 전체 시험문제를 소수의 독해 문항군으로 출제할 필요가 있다고 본다. 그러한 문항군에 속한 몇 개의 문제(예를 들면, 한 독해 문항군에 속한 4-6개의 시험 문제)를 출제하면서 문항의 영어 수준은 높이면서도, 개별적으로 출제되어 1문항으로만 구성된 문제 유형의 수는 대폭 줄이는 방안도 고려하여 볼 필요가 있다. 그럼으로써 많은 유형의 문항에 대한 등급의 준거를 세우는 불편을 없애고, 소수의 문항군에 맞는 등급 준거를 세울 수 있도록 보장할 수 있다(Jiao & Chen, 2014; Wainer, 2007). 현재 수능영어 독해 문제를 살펴보면 대부분 한 문항이 하나의 시험 문제유형으로 되어 있다. 반면에 대부분의 국제적인 영어 구사력 시험의 독해 부분은 한 문항군에 여러 문제가 속하여 있으며, TOEFL, IELTS, 중국 CET(College English Test), 대만 GEPT(General English Proficiency Test) 등의 독해 문제 예를 들 수 있다. 만약 우리 영어시험도 이렇게 소수의 독해 문항군으로 출제하게 되면, 수험자의 진정한 영어능력이 가능한 한 많이 반영되면서도, 오히려 현재의 수능처럼 각 유형의 문제를 푸는 요령의 효과는 감소할 수 있다고 예상할 수 있다.

문항군으로 구성된 수능 영어시험이 수험자의 진정한 영어능력을 제대로 반영하면서도 시험을 치르는 요령의 효과가 적다고 판명되면, 학생들은 본래 영어능력의 향상을 위해 노력할 것으로 본 연구자는 예상하여 본다. 전술한 바와 같이, 수험자가 자신의 영어 독해 능력의 향상을 통하여 본래 독해 문항의 문제 해결에 집중하고, 시험에만 대처하는 요령에 대한 수험자의 관심과 노력은 감소되는 수능 영어시험의 출제가 궁극적으로는 학생의 자기주도 학습을 유도하고, 사교육에 지나치게 의존할 필요를 감소시킬 수 있다고 본다. 또한 그러한 문항군에 의한 시험을 통해 수능영어 절대기준평가에 맞는 토대를 현재 수능 영어시험 체제보다는 잘 마련할 수 있다고 기대할 수 있다.

본 연구자는 수능 영어시험을 문항군으로 출제하는 것을 주장하고 있지만, 한편 그러한 문항군 출제가 현실적 문제가 없다고는 할 수 없다. 현재의 수능 영어시험은 각 출제위원들이 약 4주 동안 모처에 모여 각 문항을 출제하여 수합하는 체제이다. 수능 영어시험의 45문항을 과연 현재의 출제 체제 하에서 소수의 문항군으로 모아서 출제하기에는 어려움이 있으리라는 우려가 든다. 이러한 어려움은 결국 문제은행 방식으로 문제가 개발되어 수능 문제가 출제되어야 하는 것을 의미할 수 있고, 결국 다량의 문제은행 방식으로써 수능영어 절대평가가 가능할 수 있다는 주장을 뒷받침할 수 있다. 또한 수능 영어시험이 문항군으로 출제한다고 하더라도 과연 시험 문제에 대처하는 방법/요령의 효과가 전혀 없다고는 볼 수 없다.

그러므로, 본 연구자는 수능 영어시험을 가능하면 절대평가 취지에 맞도록 문항을 개선하여야 한다고 본다. 현재처럼 각 문항이 하나의 문제 유형인 영어 시험으로 출제되고 또한 지난 세월 동안 수능영어 문항의 개선이나 수능영어 문제 유형의 개혁이 전혀 이루어지지 않은 것을 고려한다면 이제는 수능 영어 시험의 문항 개선 및 보완뿐만 아니라 부단한 타당도 검증에 노력하여야 한다. 또한 현재 수능 영어의 듣기 및 읽기 능력은 타당하면서도 진정한 영어 능력으로서 목표언어사용-대학에서 영어원서를 어려움 없이 읽거나 또는 영어로 정보를 빠르게 파악하는 능력 등-을 배려하여야 하는데(Bachman & Palmer, 1996, 2010), 과연 현재의 수능 영어시험이 대학에서의 목표언어사용을 잘 반영하고 있는지도 재고할 필요가 있다.

4.1.3. 영어공교육 정상화를 위한 영어 표현력(말하기 및 쓰기) 평가 실시

현재 수능영어시험이 절대평가라고 하지만, 아직도 이해력(듣기 및 읽기) 위주로 시행되고 있다. 현재 수능영어 듣기 및 읽기 시험은 대규모 시험으로서(Fox, 2014) 현재의 절대평가 방식으로 치르기에는 여러 가지 문제가 노출되고 있다. 그러나 우리의 영어 공교육이 정상화되려면 영어교육과정과 영어평가가 일치되어야 하며, 공교육에서의 영어평가가 이해력 위주의 시험에서 탈피하도록 노력하여야 한다. 특히 영어시험의 절대평가는 영어 듣기 또는 독해보다는

오히려 영어 말하기 또는 쓰기에서 준거만 잘 설정하면 더욱 적절할 것으로 사료된다.

이러한 상황에서 영어 표현력(말하기와 쓰기)은 교육과정에 맞추어서 최소 수준만 통과하면 되는 절대평가를 실시하는 것도 바람직하다고 주장할 수 있다. 영어 표현력 평가는 대규모 표준화 시험으로서 컴퓨터를 통하여 시행할 수도 있지만, 현재 영어교사의 능력은 의사소통적 영어수업을 할 수 있는 능력이 충분히 갖추어져 있다고 판명되었으므로(Lee, 2015), 영어 교사가 영어 표현력 시험을 교실 수업에서 시행하는 것도 고려할 수 있다. 이러한 영어 표현력 시험에 대해 국가는 영어 표현 능력에 대한 척도(최저 수준의 등급)를 잘 제공하고, 영어 교사들의 영어 표현력 채점 및 신뢰도에 대한 안목을 도모하여 영어 표현력에 대한 평가가 잘 진행될 수 있도록 배려할 수 있다.

현재 전세계적으로 언어교사의 교실수업에서의 언어평가 문식력(language assessment literacy: LAL)에 대한 논의가 활발히 진행되었고, 언어평가 전문 학술지 *Language Testing* 30권 3호(2013)는 언어평가 문식력에 대해 집중적으로 다루었다. 그러나 국내에서는 언어평가 문식력에 대한 논문이 미비하고, 국내에서 실시되는 영어평가 문식력에 대한 연수 프로그램이 많은 현직 교사들을 만족시키지 못하고 있다(Chung & Nam, 2018). 현직 영어교사들에게 언어평가 문식력을 갖추게 하는 직전 교육(pre-service education) 및 현직 연수(in-service development) 뿐만 아니라, 온라인 및 자기 연수를 실시할 수 있도록 배려한다면, 교실 수업을 통한 기본 수준의 영어 표현력(특히 말하기)에 대한 절대평가는 상당히 가능한 것으로 예상할 수 있다. 이러한 영어표현력(특히 영어 말하기 능력)의 절대 평가는 마치 운전 면허시험처럼 최소 능력 수준만 갖추면 통과하게 함으로써, 거의 모든 학생들이 영어표현력에 대한 최저 수준 만이라도 갖추도록 유도하고, 결국 공교육의 영어교육과정과 영어평가를 일치시킬 수 있다. 이렇게 함으로써 그동안 수능영어시험뿐만 아니라 공교육의 영어교육이 이해력 위주로만 진행되어 야기된 파행(교육과정과 평가의 불일치)을 방지할 수 있다고 본다.

4.2. 결론

우리의 대학입시를 위한 영어 평가가 너무나도 급속도로 정치적 맥락에 의해 좌우되다 보니, 본 연구자로서도 체계적인 연구 조사를 통한 연구 결과를 제시하지 못하고, 다분히 추측적인(speculative) 논리를 제시하게 된 것이 본 연구의 제한점이라고 할 수 있겠다. 이미 논의한 바와 같이, 현재 수능영어 절대평가는 기준과 점수체제가 심층적인 연구와 검증을 통하여 제대로 규명이 안 된 채로 막연한 등급을 설정하다 보니, 결국 수능영어 절대평가의 이론적 결함 및 타당한 연구 부족으로 인한 문제가 드러나게 되었다. 절대 기준에 대한 이론적 규명이 없고, 수능영어 절대평가—특히 수능영어 시험의 절대 기준과 등급—에 대한 충실한 검증 연구가 진행되지 않은 상태에서 절대평가를 추진하다 보니,

그 결과로 현재 실시되고 있는 수능영어 평가 방식은 절대평가가 아니라 쉬운 수능 영어 비율 등급제로 고득점자가 많아진 상대평가가 되었다. 결국 교육부에 의해 주도된 수능영어 절대평가의 본래 취지가 무색하게 되었으며, 또한 교육부 정책적인 의도(사교육 억제, 학생들간의 과열 경쟁 방지 및 공교육 정상화)가 효과적으로 반영되었다고 보기도 어렵다. 따라서 수능영어 이해력(듣기 및 읽기)에 대한 절대평가는 사실상 타당하지 않은 것으로서 교육부의 하향식 정책에 의해 무리하게 집행되었다고 하여도 과언이 아니다.

국어나 수학과 같은 다른 주요 과목도 절대평가로 실시하지 않고 또한 수능 영어 절대평가가 이전의 상대평가로 돌아가지 않는 현실에서 현재 절대평가라는 명목 하에 실시되고 있는 수능영어 시험이라도 비교적 엄정하게 실시될 수 있도록 절대평가의 기본이 되는 준거참조평가의 실증적인 이론적 연구, 꾸준한 타당도 검증 및 영어 교육현장 의견을 반영한 실행 연구를 통하여 보완하는 노력이 절실히 요구된다고 할 수 있다. 그러나 그와 아울러 영어교육자들의 중지가 모아져서 교육부에서 하향식으로 진행되어온 영어 교육 및 평가 정책에 대해 영어교육 및 평가 당사자들이 상향식으로 대처하여 상호 작용적이고 타당한 영어 교육 및 평가 정책이 확립되게 하는 노력도 중요할 수 있다고 본다.

참고문헌

- 강태중, 김용명, 김진구, 이병천, 김성혜, 박지선, 황필아, 주현우, 이용상, 김인숙, 임황규, 이완기, 박찬호, 최창완. (2014). *대학수학능력시험 영어영역 절대평가 도입 방안* (연구보고 CAT 2014-36). 서울: 한국교육과정평가원.
- 강태중. (2014). 대학수학능력시험 발전 방향 모색. *제26회 KICE 교육과정·평가 정책포럼: 수능 영어영역 절대평가 도입 방안 탐색* (수능 CAT 2014-13) (pp. 3-13). 서울: 한국교육과정평가원.
- 강태훈. (2014). 수능 영어 절대평가 점수체제 탐색. *제26회 KICE 교육과정·평가 정책포럼: 수능 영어영역 절대평가 도입 방안 탐색* (수능 CAT 2014-13) (pp. 87-100). 서울: 한국교육과정평가원.
- 교육부. (2014, 12월 26일). 보도자료. 대학수학능력시험 영어영역 절대평가 도입: 현재 중학교 3학년 학생들이 대학에 진학하는 2018학년도 수능부터 적용.
- 교육부. (2018, 8월 17일). 보도자료. 2022학년도 대학입학제도 개편방안 및 고교 교육 혁신 방향 발표.
- 김기홍. (2017, 12월 12일). 첫 절대평가 영어, 1등급이 10%. *조선일보*, p. A16.
- 김미향. (2018, 3월 16일). 1인당 사교육비 월27만원1천원... 7년제 최고액 경신. *한겨레신문*, p. 10.
- 김준식, 이용상, 박지선, 민호기, 박상욱, 박용호, 황필아, 박찬호, 신상근, 신동광. (2015). *대학수학능력시험 영어영역 절대평가 시험체제 및 점수체제*

- 방안 (연구보고서 CAT2015-20). 서울: 한국교육과정평가원.
- 박세미. (2018, 12월 5일). 국어 1등급 10점, 2등급 9점 떨어져... 영어 1등급은 반 토막. *조선일보*. 월드와이드웹: http://news.chosun.com/site/data/html_dir/2018/12/05/2018120500205.html에서 2018년 12월 5일에 검색했음.
- 박세인. (2017, 12월 11일). 절대평가 영어 난이도 조절실패... 정시 지원 혼란 예고: 10명중 1명이 1등급. *한국일보*, p. 10.
- 박찬호. (2014). 수능 영어 절대평가 점수체제 탐색에 대한 토론. *제26회 KICE 교육과정·평가 정책포럼: 수능 영어영역 절대평가 도입 방안 탐색* (수능 CAT 2014-13) (pp. 111-116). 서울: 한국교육과정평가원.
- 송옥진. (2018, 11월 19일). 영어 널뛰기 난도에 ‘상대평가와 다르게 뭐냐’. *한국일보*, p. 8.
- 신혜정. (2018, 12월 4일). ‘14년만에 최고 난도’ 국어가 정시 단락 좌우. *한국일보*, p. 6.
- 양도웅. (2018, 8월 20일). 한국영단협, “영어에도 다른 기초과목과 동일한 평가 방법 적용해야”. *교수신문*. 월드와이드웹: <http://www.kyosu.net/news/articleView.html?idxno=42491>에서 2019년 5월 28일에 검색했음.
- 양선아. (2018, 12월 4일). 2019학년도 수능 ‘불수능’으로 확인... 평가원 “난도 조절 실패 송구” 사과. *한겨레신문*, p. 10.
- 이도경. (2018, 3월 16일). 1인당 사교육비 역대 최고... 손놓은 정부. *국민일보*, p. 10.
- 이영식. (2018, 7월 31일). 수능영어 절대평가의 문제점. *헤럴드경제*. 월드와이드웹: http://biz.heraldcorp.com/view.php?ud=20180731000374&ACE_SEARCH=1에서 2018년 8월 5일에 검색했음.
- 이영식. (2018, 12월). 대학 수능 영어시험 절대평가 실시: 정치적 맥락에 의해 변화된 시험 사례. *2018년 한국영어학학회 추계 학술대회 발표논문집: 수능 영어 절대평가와 영어공교육이 대학영어교육에 미치는 영향*(pp. 109-115). 숭실대학교 벤처중소기업센터 311호.
- 정지윤. (2018, 8월 8일). 2022 수능개편 영어도 다른 과목과 동일한 평가방법 도입해야. *경향신문*, p. 10.
- 정혜선. (2019). *대학 수학능력 시험 영어영역 절대평가에 대한 고등학교 교사와 학생들의 인식*. 박사학위논문, 한남대학교, 대전.
- 한국교육과정평가원. (2014). *수능 영어영역 절대평가 도입 방안 탐색*. 제26회 KICE 교육과정 평가 정책포럼 (수능 CAT 2014-13). 한국교육과정평가원 대회의실.
- 한국교육과정평가원. (2014). *수능 영어영역 절대평가 도입 방안 공청회(1차) 자료집* (수능 CAT 2014-29-1). 한국교육과정평가원 대회의실.
- 한국교육과정평가원. (2017, 12월 12일). 보도자료. 2018학년도 대학수학능력시험 채점결과. 월드와이드웹: <http://www.kice.re.kr/boardCnts/view.do?boardID=10024&boardSeq=5021730&lev=0&m=050102&searchType=S&statusYN=W&p>

- age=7&s=kice에서 2018년 1월 8일에 검색했음.
- 한국교육과정평가원. (2018, 12월 5일). 보도자료. 2019학년도 대학수학능력시험 채점결과. 월드와이드웹: <http://www.kice.re.kr/boardCnts/view.do?boardID=10024&boardSeq=5033660&lev=0&m=050102&searchType=S&statusYN=W&page=1&s=kice>에서 2018년 1월 8일에 검색했음.
- 한국교육과정평가원. (2018, 12월 5일). 보도자료. 2019학년도 대학수학능력시험 채점결과.
- 한국교육과정평가원. (2018, 12월 19일). *학교 영어교육 내실화 방안 연구 설명회* (연구자료 ORM 2018-122). 진천: 한국교육과정평가원.
- 한국교육학회 교육평가연구회. (1995). *교육 측정·평가·연구·통계 용어 사전*. 서울: 중앙교육진흥연구소.
- 한국영어관련학술단체협의회. (2018). *한국 영어교육의 위기: 학교 영어교육과 국가 경쟁력 발표집*. 한국영어관련학회 공동 심포지엄. 한국프레스센터 20층.
- 한국영어관련학술단체협의회. (2018, 6월 22일). 보도자료. 대입개편 공론화 결과 발표에 관한 한국영어관련학술단체협의회 입장문. *한국 영어교육의 위기: 학교 영어교육과 국가 경쟁력 발표집*, p. 5.
- 한국영어학학회. (2018). *수능영어 절대평가와 영어공교육이 대학영어교육에 미치는 영향*. 2018년 추계 학술대회 발표논문집. 숭실대학교 벤처중소기업 센터 311호.
- Abedi, J. (2008). Utilizing accommodations in assessment. In E. Shohamy & N. H. Hornberger (Eds.), *Language testing and assessment* (pp. 331-348). New York: Springer.
- American Council on the Teaching of Foreign Languages. (2012). *ACTFL proficiency guidelines 2012*. Retrieved on December 3, 2018, from the World Wide Web: <https://www.actfl.org/publications/guidelines-and-manuals/actfl-proficiency-guidelines-2012>
- American Council on the Teaching of Foreign Languages. (2015). *ACTFL performance descriptors for language learners*. Retrieved on December 3, 2018, from the World Wide Web: <https://www.actfl.org/sites/default/files/pdfs/ACTFLPerformance-Descriptors.pdf>
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.
- Bachman, L. F. (1991). What does language testing have to offer? *TESOL Quarterly*, 25(4),

671-704.

- Bachman, L. F., & Palmer, A. (1996). *Language testing in practice*. Oxford: Oxford University Press.
- Bachman, L. F., & Palmer, A. (2010). *Language assessment in practice*. Oxford: Oxford University Press.
- Bae, J. (2014). Norm-referenced approach to language assessment. In A. Kunnan (Ed.), *The companion to language assessment* (pp. 542-560). Chichester, UK: John Wiley & Sons.
- Brown, J. D., & Hudson, T. (2002). *Criterion-referenced language testing*. Cambridge: Cambridge University Press.
- Chapelle, C. A., & Voss, E. (2014). Evaluation of language tests through validation research. In A. Kunnan (Ed.), *The companion to language assessment* (pp. 1081-1097). Chichester, UK: John Wiley & Sons.
- Cohen, A. (1987). Study in learner strategies: Feedback on composition. *PASAA*, 17(2), 29-38.
- Cohen, A. (2012). Test taking strategies. In C. Coombe, P. Davidson, B. O'Sullivan, & S. Stoyhoff (Eds.), *The Cambridge guide to second language assessment* (pp. 96-104). Cambridge: Cambridge University Press.
- Chung, S. J., & Nam, Y. (2018). Language assessment literacy of Korean EFL teachers: An investigation of their training experiences and needs. *Modern English Education*, 19(1), 38-48.
- Council of Europe. (2001). *Common European framework of reference for languages: Learning, teaching, assessment*. Cambridge: Cambridge University Press.
- Davidson, F. (2012). Test specifications and criterion referenced assessment. In G. Fulcher & F. Davidson (Eds.), *The Routledge handbook of language testing* (pp. 197-207). New York: Routledge.
- Fox, J. (2014). Large-scale assessment. In A. Kunnan (Ed.), *The companion to language assessment* (pp. 533-541). Chichester, UK: John Wiley & Sons.
- Geisinger, K. F. (2016). Test evaluation. In S. Lane, M. R. Raymond, & T. M. Haladyna (Eds.), *Handbook of test development*. (pp. 624-638). New York: Routledge.
- Hudson, T. (2014). Criterion-referenced approach to language assessment. In A. Kunnan (Ed.), *The companion to language assessment* (pp. 561-577). Chichester, UK: John Wiley & Sons.
- Jiao, H., & Chen, Y-F. (2014). Differential item and testlet functioning analysis. In A. Kunnan (Ed.), *The companion to language assessment* (pp. 1282-1300). Chichester, UK: John Wiley & Sons.
- Kunnan, A. (1995). *Test taker characteristics and test performance: A structural equation modeling*. Cambridge: Cambridge University Press.

- Lee, Y. S. (2015). Innovating secondary English education in Korea. In B. Spolsky & K. Sung (Eds.), *Secondary school English education in Asia* (pp. 47-64). New York: Routledge.
- McNamara, T., & Roever, C. (2006). *Language testing: The social dimension*. Oxford: Blackwell.
- Meier, D., & Wood, G. (2006). *Many children left behind: How the No Child Left Behind Act is damaging our children and our schools*. Boston, MA: Beacon Press.
- Messick, S. (1998). Test validity: A matter of consequence. *Social Indicators Research*, 45, 35-44.
- Taylor, L. (2012). Accommodation in language testing. In C. Coombe, P. Davidson, B. O'Sullivan, & S. Stoyloff (Eds.), *The Cambridge guide to second language assessment* (pp. 105-112). Cambridge: Cambridge University Press.
- Wainer, H. (2007). *Testlet response theory and its applications*. Cambridge: Cambridge University Press.
- Weir, C. J. (2005). *Language testing and validation: An evidence-based approach*. New York: Palgrave Macmillan.
- Xi, X. (2010). Methods of test validation. In E. Shohamy & N. H. Hornberger (Eds.), *Language testing and assessment* (pp. 177-196). New York: Springer.

적용가능 수준(Applicable levels): Secondary, tertiary

이영식

교수

한남대학교 사범대학 영어교육과

34430 대전광역시 대덕구 한남로 70

Email: youngslee@hnu.kr

Received on March 15, 2019

Reviewed on May 26, 2019

Revised version received on May 31, 2019