영어어문교육 *제 20 권* 3 호 2014 년 가을

2015 학년도 수능 통합형 검사지 구성을 위한 영어영역 문항유형 타당성 분석

김 용 명*
(한국교육과정평가원)
박 지 선
(한국교육과정평가원)
황 필 아
(한국교육과정평가원)
이 병 천**
(한국교육과정평가원)

Kim, Yong-Myeong, Park, Jiseon, Hwang, Pil-Ah & Lee, Byeong-Cheon. (2014). The internal analysis of the validation on item-types of English Language Domain of the 2014 level-differentiated English CSAT for designing the integrated English tests of the 2015 CSAT. *English Language and Literature Teaching, 20*(3), 211-229.

This study investigated the validation on item-types of Foreign (English) Language Domain of the 2014 level-differentiated CSAT (College Scholastic Ability Test) for developing the integrated English tests of the 2015 CSAT. To this end, this study applied the testing criteria(authenticity, interactiveness, wash-back effectiveness, test-item development easiness, connection) which was used by the survey participants for evaluating each item-type based on five-point Likert scale, and revised the scoring procedures set by Kim(2010) which was designed for evaluating and analyzing the item-types of the Foreign (English) Language Domain on the basis of the 5 principles of item-selection. Following this criteria and procedures, it selected 60 test-evaluator and made them evaluate and analyze 18 item-types of the listening test and 20 item-types of the reading test of the 2014 level-differentiated CSAT. On the basis of the results of the evaluation and the analysis, it constructs an *item-types* for designing and developing the integrated tests of 2015 CSAT.

[College Scholastic Ability Test(CSAT)/Foreign (English) Domain/item-types/ validation, 대학수학능력시험(수능)/영어 영역/문항 유형/타당성]

* 제 1 저자 **교신저자

교보문고 KYOBO Book Centre

I. 연구의 필요성과 목적

본 연구는 정부의 '대입전형 간소화 및 대입제도 발전방안'(교육부, 2013.08.27.) 및 '2015 학년도 및 2016 학년도 대입제도 및 2015 학년도 대입전형 기본사항'(교육부·한국대학교육협의회, 2013.09.23.)에 의하여, 2014 학년도 수준별 영어 시험(A/B 형)이 2015 학년도(2014 년 11 월 시행)에는 폐지됨에 따라, 2015 학년도 '통합형' 수능 영어 영역의 검사지를 설계하는 데 그 목적이 있다.

대학수학능력시험(이하 수능)은 약 2 년간 7 회의 실험 평가를 거쳐 1993 년에 처음 시행되었다. 수능에서의 영어 능력 평가는 처음부터 독립된 영역으로 설정된 후, 수능 역사 20 년 동안 수능 전체 체제의 변화에 따른 변화 또는 외국어(영어) 영역 자체의 변화를 겪어 왔다. 특히 2013 년은 2014 학년도 수준별 수능 영어 영역 시험이 처음 시행되는 해이면서 동시에 폐지됨에 따라, 2015 학년도부터는 '수준별 시험' 대신 2013 학년도까지 유지되던 단일형 시험과 유사한 '통합형 시험'으로 변화하게 되었다. 그러나 2014 학년도 수준별 시험별 수능이 복잡한 대입 전형(수시 전형, 일반 전형, 특별 전형 등)과 맞물려 대입 전형이 더욱 복잡해졌다는 비판에 따라 2015 학년도부터는 통합형 수능이 도입되었다 (교육부, 2013.08.27.) 따라서 필연적으로 2015 학년도 '통합형' 수능 영어 영역의 검사지가 설계되어야 한다. 검사지 설계를 위해서는 먼저 읽기와 듣기 영역별 문항 유형이 결정되어야 하며, 이를 위해서는 기존의 문항 유형에 대한 타당성 분석 연구를 바탕으로 해야 한다. 문항 유형에 대한 타당성 분석 연구로는 강문구와 김용명(2013)이 있고, 그 이외에는 읽기 영역의 난이도 분석 연구(장경숙, 2004), 지문 친숙도 분석 연구(이경숙, 1999), 어휘 분석 연구(김낙복, 2008) 등과 같은 단편적인 연구만 있다. 대학수학능력시험이나 국가수준학업성취도평가와 같은 대단위 고부담 시험과 관련된 문항 유형에 대한 총체적 연구는 수준별 시험에 관한 강문구와 김용명(2013)을 참고할 수 있으나, 대학수학능력시험이 수준별 시험에서 통합형으로 변경됨에 따라 통합형 검사지에 대한 문항 유형 타당성 분석이 필요하다. 이와 같은 연구 목적을 구체화하기 위해 본 연구는 2014 수준별 외국어(영어) 영역의 문항 유형에 대한 타당성 분석, 그리고 문항 유형 타당성 분석 결과에 따른 2015 학년도 통합형 검사지 설계에 대한 제언이라는 두 가지의 연구 주제를 설정하였다.

이와 같은 연구 목적에 따른 연구 주제 탐색을 통하여 2014 학년도의 수준별 A/B 형을 통합형 검사지로 변경함에 있어 읽기와 듣기 영역별 문항 유형 타당성 분석 결과에 따라 유형이 결정되고, 이에 따른 검사지 구성 방안을 제안함으로써, 2015 학년도의 안정적 수능을 준비하는 데 본 연구의 의의가 있다.

교보문고 KYOBO Book Centre

Ⅱ. 이론적 배경

1. 문항 유형 타당성 평가 기준과 평가 요소[1]

문항 유형의 타당성을 평가하기 위해서는 문항 유형 평가에 대한 영역(기준)과 평가 요소를 설정해야 한다. 이를 위하여 본 연구에서는 강문구 와 김용명(2013)의 연구에서 개발한 문항 유형 타당성 평가 영역(기준)과 평가 요소를 통합형 수능 영어영역 시험에 적합하게 일부 수정하여 적용하였다. 먼저, Ⅱ장에서 논의한 문항 유형 결정 원리와 검사지 구성 원칙을 토대로, 문항 유형의 질적 타당도를 평가하기 위한 5 기준, 즉, 출제용이성, 연계성, 실제성, 상호작용성, 환류 효과성과 그 하위 기준 및 평가 요소를 추출하였다(표 1 참조).

표 1

문항 유형 타당성 평가 기준과 평가 요소(문항 유형 타당성 평가표)

영역	하위 영역	평가 요소(영역 명세화)	←매우 그렇지 않다 매우 그렇다→ 1	2	…	*n*
연계성	성취기준	문항 유형이 교육과정의 성취 기준과 부합하는가?			…	
실제성	TLU 영역	문항 유형의 특성이 TLU (Target Language Use)과업의 특성과 일치하는가?			…	
	필요불가결성	문항 유형이 TLU 영역에 비춰보아 필요불가결한 것인가?			…	
	이해 방식	문항 유형이 TLU 상황에서 듣기·읽기 이해 방식을 반영하는가?			…	
상호 작용성	상호작용성	문항 유형이 수험자의 다양한 언어능력과 상호작용하는가?			…	
	전략적 능력	문항 유형이 실시간 상으로 수험자의 전략적 능력과 관계하는가?			…	
출제 용이성	문항유형 타당성 지문 선정 용이성	문항 유형이 상업용 기출 문항과 중복되지 않고, 적절한 담화나 대화 그리고 지문을 선정하기에 용이한가?			…	
환류 효과성	교수 · 학습	문항 유형이 교수 및 수업 활동에 긍정적 영향을 주는가?			…	
	교실 평가	문항 유형이 교실 평가에 긍정적 영향을 주는가?			…	

[1] 본 연구에서의 이론적 배경 및 일부 선행 연구에 해당하는 타당성 평가영역 및 평가요소, 그리고 모듈형 문항 유형은 강문구와 김용명(2013) 및 김용명(2010)을 요약·정리하고, 용어의 정의나 설명에 해당하는 주요부분을 발췌하였기 때문에 부분적 인용 페이지 각각 개별적 표기하지 않음에 따라 이를 각주로 처리함을 밝혀 둔다.

교보문고 KYOBO Book Centre

2. 모듈형(modular) 평가 목표 분류표

2013 학년도까지의 단일형 검사지에서는 고정형 평가 목표 이원 분류표를 기준으로 하여 검사지를 구성하였다. 고정형 평가 목표 분류표는 단일 수준의 단일 시행 체제에서는 시험의 항상성과 동일성을 유지하기가 비교적 용이하고, 출제 과정도 비교적 안정적으로 운영할 수 있다는 장점이 있으나, 내용 영역 분류를 함에 있어 글의 종류에 따라 논설문, 설명문, 문학, 실용문, 기타로 구성하였다. 이처럼 글의 종류에 따라 내용 영역을 구성할 경우, 평가 이론상의 오류와 출제상의 문제점이 발생할 수 있다. 읽기 이론에 따르면, 글의 종류는 형식 스키마(formal schema)에 속하는데, 이는 읽기 능력의 주요 직접적 평가 요소에 속한다(김용명, 2010). 따라서 수능 외국어(영어) 영역 시험은 형식 스키마라는 직접 평가 요소를 내용 영역 기준으로 삼는 논리적 오류에 빠져 있으며, 그 결과 문항 개발 시, 형식 스키마의 이해 능력과 관련된 문체나 글의 구조에 관한 문항을 출제할 수 없다는 한계점이 있다. 또한 출제 과정에서 글의 종류에 따라 내용 타당도를 따져야 하므로 해당 지문의 상보성을 파악하기가 쉽지 않다. 보다 본질적인 문제는 글의 종류에 따른 분류는 상보성 원칙에 위배되는 경우가 많아서 수험자의 독서 성향이 시험 점수에 영향을 미칠 수 있다는 점이다. 이는 시험의 공정성 및 형평성의 문제를 야기할 수 있다(김용명, 권오량, 이완기, 김진석, 고현숙, 2011).

이 같은 한계점을 극복하기 위해, 2014 학년도 수준별 수능 영어 영역의 검사지 구성은 '모듈형(modular)' 평가 목표 분류표에 따랐다. 모듈형 평가 목표 분류표는 실제성 원리[2]에 따라 평가 목표 분류표의 각 영역과 그 하위 영역의 구성소를 가능한 한 세분하여 구성하지만, 상보성 원칙에 따라 각 영역과 그 하위 영역 간에 서로 상보적 분포를 이루도록 구성한 것을 말한다. 모듈형 평가 목표 분류표의 구성소는 내용 영역 체계, 행동 영역 체계, 문항 유형 특성 체계(김용명 외 4 인, 2011)로 구성된다.

또한 모듈형 평가 목표 분류표의 문항 유형 특성 체계는 검사지 구성의 4 원칙, 즉, 상보성(complementarity), 통합성(integration), 주축성(pivotality), 위계성(hierarchicality)으로 구성된다(김용명, 2010)[3]. 상보성(complementarity)은

[2] 실제성(authenticity)은 '시험 상황에서의 시험 과업의 특성이 TLU(Target Language Use, 이하 TLU) 영역에서의 TLU 과업의 자질과 일치하는 정도'로 정의한다(Bachman & Palmer, 1996, p. 23). 실제성은 시험 상황에서 시험 과업의 수행 결과(점수)를 비시험 상황으로 일반화하는 것과 관련된 개념이다. 다시 말해, 시험 상황에서 시험 과업의 특징이 TLU 영역에서의 TLU 과업의 자질과 일치할수록, 시험 상황에서 시험 결과(수행력)와 비시험 상황의 특정 TLU 영역에서의 실제 언어사용 능력과 일치할 가능성은 높아질 것이다. 실제성의 원리는 문항 유형 개발과 선별의 범위와 한계를 설정하는 역할을 한다. 따라서 실제성 원리는 수준별 영어시험(A/B형)의 평가 목표 분류표 상의 내용 영역의 주요 구성소를 설정하고, 이의 타당성을 검증하는 이론적 토대가 된다(김용명, 2010).

[3] 검사지 구성 원칙은 Bachman과 Palmer(1996)의 시험 유용성(test usefulness) 모

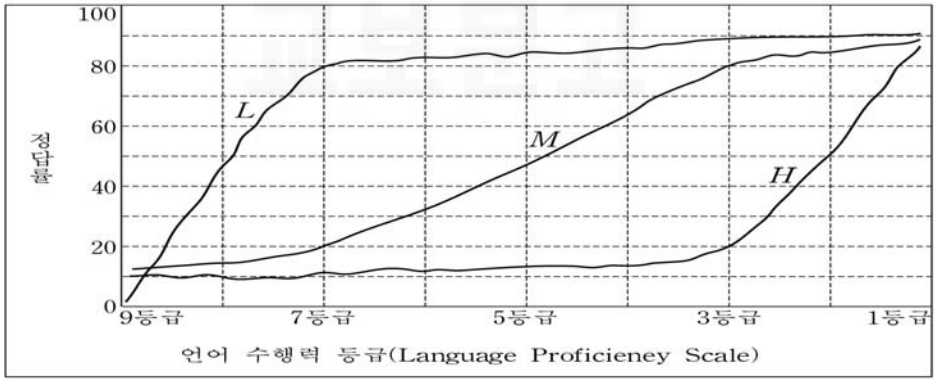
교보문고 KYOBO Book Centre

문항 또는 문항 유형이 언어능력, 내용 영역 및 행동 영역의 각 구성소가 서로 상보적 분포를 이룰 수 있도록 검사지를 구성해야 한다는 원칙이고, 통합성(integration)은 최근 교수 이론(Bachman & Palmer, 1996; Brown, 2007; Ellis, 2003)과 외국어과 교육과정(교과부, 2008)에서 강조되고 있는 4 기능의 통합 교수 원리를 토대로 구성한 개념이다. 주축성(pivotality) 원칙이란 하나의 검사지는 주축형 문항(pivot items)과 주변형 문항(peripheral items)으로 구성되는 원칙을 말한다. 일반적으로 주축형 문항은 평균 정답률을 기준으로 정답률의 편차가 상대적으로 작으며, 문항의 복잡도(어휘적, 언어적, 인지적 복잡도)가 높아지면, 이에 따라 난이도도 올라가는 경향을 보인다(문항의 복잡도와 난이도는 정비례 관계). 따라서 정답률의 편차가 상대적으로 작고, 학습자의 언어 능력(예, 문법적, 담화적 능력 등)에 민감성을 보인다는 사실에서 주축형 문항은 시험의 항상성 유지와 동등화에 기여할 수 있다는 것을 추론할 수 있다.

위계성(hierarchicality) 원칙은 한 검사지를 구성하는 각 문항의 난이도는 각 수험생별(또는 각 집단별) 수행 가능 단계와 일치할 수 있도록 위계화 되어야 한다는 원칙을 말한다. 수능 외국어(영어) 영역의 문항 분석(한국교육과정평가원, 2005, 2006, 2007, 2008, 2009, 2010) 결과에 따르면, 문항에 대한 수험자(학습자)의 반응에 따라 다양한 문항 반응 곡선들이 나타날 수 있지만, 그림 1 에서 보는 바와 같이, 문항 유형의 특성에 따라 이 다양한 곡선들을 (다소 추상화하여) 세 유형, 즉, 저난도(L)형, 중난도(M)형, 고난도(H)형으로 정형화할 수 있다[4].

그림 1

저난도(L)형, 중난도(M)형, 고난도(H)형 문항 유형 반응 곡선



델과 수능 외국어(영어) 영역의 문항 분석 자료(한국교육과정평가원, 2005, 2006, 2007, 2008, 2009, 2010)를 토대로 구성하였다.

[4] 문항 반응 곡선 저난도(L)형, 중난도(M)형, 고난도(H)형은 각각 low, mean, high의 첫 글자를 딴 것으로 하위 학습자, 중위 학습자, 상위 학습자를 보다 잘 변별한다는 의미를 담고 있다.

교보문고 KYOBO Book Centre

저난도(L)형 문항 반응 곡선은 하위 등급(예, 7·8·9 등급)에서는 등급이 올라감에 따라 정답률이 올라가지만, 일정 등급(예, 7 등급) 이상을 넘어서면 정답률(예, 80%대)이 고정되어 시험 고원을 형성하는 문항 유형을 말한다. 고난도(H)형의 문항 반응 곡선의 특성을 보이는 문항 유형에는 x, y, z 등의 유형이 있으며, 이들 문항 유형은 상위 등급 간의 난이도 및 변별도 조정에 결정적 역할을 한다(김용명, 2010, pp. 333-336).

이상에서 논의한 문항 유형 특성 체계의 4 구성소는 모듈형 평가 목표 이원분류표로부터 각 기능별 수준별 시험의 검사지를 구성할 때, 각 검사지의 질적 특성을 통제하는 역할을 한다. 이를 부연해서 설명하면, 문항 유형 특성 체계의 구성소, 즉, 통합성, 주축성, 위계성에 따른 각 문항 유형의 적정 구성 비율을 조정함으로써 시험의 항상성, 다양성, 동등성을 확보할 수 있다.

Ⅲ. 연구 방법

1. 설문 조사 대상

본 연구는 통합형 수능 영어 문항 유형의 타당성을 평가하기 위한 것이므로, 설문 조사 대상자를 2005 학년도부터 2013 학년도까지 수능 영어 영역 시험에 두 번 이상 출제 경험을 갖고 있는 출제 전문가 60 명을 문항 평가자로 선별하였다.[5] 구체적으로 살펴보면, 성별로는 남자 37 명(61.7%), 여자 23 명(38.3%)이었으며, 직종으로는 교수 32 명(53.3%), 교사 25 명(41.7%), 연구원 2 명(3.3%), 기타 1 명(1.7%)이었으며, 평균 나이는 41.9 세이다(표 2 참조).

표 2
설문 응답자 구성

분류	성별		직종			
	남	여	교수	교사	연구원	기타
인원(명)	37	23	32	25	2	1
(%)	61.7	38.3	53.3	41.7	3.3	1.7

2. 설문지 구성

문항 유형의 타당성을 평가하기 위해서는 문항 유형 평가에 대한 영역(기준)과 평가 요소를 설정해야 한다. 이를 위하여 본 연구에서는 다음의 과정을

[5] 본 연구의 특성상 출제경험을 가진 전문가에 한정하여 참여자의 구체적 인적 사항은 비공개로 하였으며, 연구참여자에게는 연구의 취지를 설명하고 동의를 득한 후 설문 시행하였다.

교보문고 KYOBO Book Centre

통해 문항 유형 타당성 평가 영역(기준)과 평가 요소 선정 연구(강문구 & 김용명, 2013)를 통합형 수능 영어영역 시험에 적합하게 일부 수정하여 적용하였다. Ⅱ장에서 논의한 문항 유형 결정 원리와 검사지 구성 원칙(김용명, 2010)을 토대로, 문항 유형의 질적 타당도를 평가하기 위한 5 기준, 즉, 출제용이성, 연계성, 실제성, 상호작용성, 환류 효과성과 그 하위 기준 및 평가 요소 및 이에 따른 문항 유형별 예시 문항을 중심으로 하여 설문지를 구성하였다.

3. 설문 조사 방법

현행 통합형 수능 영어 영역의 문항 유형 타당성 평가의 설문은 다음의 순서로 실시되었다(강문구 & 김용명, 2013).

(a) 문항 평가 영역(기준)과 평가 요소에 대해 이해도와 친숙도를 높이기 위해, 문항 평가자에게 약 2 시간에 걸쳐 문항 유형 타당성 평가와 관련된 평가 영역(연계성, 실제성, 상호작용성, 환류 효과성, 출제용이성) 및 평가 요소에 대해 설명하였다.
(b) 듣기 18 개 유형 및 읽기 20 개 유형 각각에 대해 예시 문항과 문항 유형 타당성 평가표로 구성된 문항 유형 타당성 평가지(표 3 참조)를 문항 평가자에게 각 한 부씩 나눠주었다.
(c) 문항 유형 타당성 평가표에 제시된 각 예시 문항 유형에 대해 문항 평가자에게 1-2 분에 걸쳐 PPT 로 각 문항 유형의 성립 요건, 출제 방법, 평균 정답률 등을 제시하였다.
(d) 문항 유형 타당성 평가표에 제시된 문항 유형 타당성과 관련된 기준인 출제용이성, 연계성, 실제성, 상호작용성, 환류효과성에 대해 문항 평가자에게 Likert 5 점 척도에 따라 1 점에서 5 점까지 점수를 부여하게 하였다.
(e) 이어서 평가자에게 약 3 분 동안에 걸쳐 예시 문항 유형과 평가표를 읽고, 문항 유형의 타당성을 평가·분석하여 이를 평가표에 표시하게 하였다.

4. 설문 분석 방법

문항 유형 타당도 평가 및 특성 분석의 점수화 방법은 다음과 같다[6].

(a) 통합형 영어 시험은 전국적으로 시행되는 고부담 시험이므로, 점수화 방법은 비보상 통합 점수(non-compensatory composite scores) 체계에

[6] 본 연구에서의 설문 분석은 (강문구 & 김용명, 2013)의 방법을 적용하였다.

교보문고 KYOBO Book Centre

따라 분할 점수(cut-off score)는 3 점으로 하였다.

(b) 평가자가 문항 유형의 평가 기준 및 평가 요소의 전부 또는 일부를 평가 하지 않았거나, 총 38 개 유형 중, 10 개 이상의 유형에 대해 평가 하지 않았을 경우, 평가 결과를 왜곡 시킬 수 있어 분석대상에서 제외 하였다.

Ⅳ. 연구 결과

1. 듣기 문항 유형 타당성 분석 결과

통합형 수능 영어 영역 듣기 유형의 18 개의 각 문항 유형별 타당성 평가 결과는 다음과 같다.

1) 대의 파악 모듈

대의 파악 모듈은 담화·대화를 듣고 전체적인 내용을 이해·추론하는 능력을 측정하는 것으로서, '담화·대화 목적', '담화·대화자 의견', '담화·대화 주제', '화자의 심정', '대화 장소', '대화자 관계' 등의 문항유형이 여기에 속한다.

첫째, '담화·대화 목적' 유형에 대한 타당성은 전체평균 3.85 점이었으며, 5 개 하위 영역 중 연계성에 대한 타당성이 4.18 점으로 가장 높게 나타났다. 다음으로 실제성(4.01 점), 환류효과성(3.81 점), 상호작용(3.68 점) 순이었으며, 출제용이성이 3.56 점으로 가장 낮게 나타났다.

둘째, '담화·대화자 의견' 유형에 대한 타당성은 전체평균 3.66 점이었며, 5 개 하위 영역 중 연계성에 대한 타당성이 3.91 점으로 가장 높게 나타났다. 다음으로 실제성(3.85 점), 환류효과성(3.8 점), 상호작용성(3.62 점) 순이었으며, 출제용이성이 3.14 점으로 가장 낮게 나타났다.

셋째, '담화·대화 주제' 유형에 대한 타당성은 전체평균 3.71 점이었으며, 5 개 하위 영역 중 연계성에 대한 타당성이 3.98 점으로 가장 높게 나타났다. 다음으로 환류효과성(3.85 점), 실제성(3.76 점), 상호작용성(3.56 점) 순이었으며, 출제용이성이 3.38 점으로 가장 낮게 나타났다.

넷째, '화자의 심정' 유형에 대한 타당성은 전체평균 3.18 점이었으며, 5 개 하위 영역 중 연계성에 대한 타당성이 3.45 점으로 가장 높게 나타났다. 다음으로 출제용이성(3.18 점), 상호작용성(3.13 점) 순이었으며, 실제성(3.09 점)과 환류효과성(3.03 점)이 가장 낮게 나타났다.

다섯째, '대화 장소' 유형에 대한 타당성은 전체평균 3.27 점이었으며, 5 개 하위 영역 중 출제용이성에 대한 타당성이 3.42 점으로 가장 높게 나타났다. 다음으로 연계성(3.35 점), 상호작용성(3.26 점), 실제성(3.17 점) 순으로 나타

교보문고 KYOBO Book Centre

났으며, 환류효과성은 2.79 점으로 가장 낮게 나타났다.

여섯째, '대화자 관계' 유형에 대한 타당성은 전체평균 3.27 점이었으며, 5 개 하위 영역 중 출제용이성에 대한 타당성이 3.45 점으로 가장 높게 나타났으며, 다음으로 상호작용성(3.3 점), 연계성(3.28 점), 실제성(3.21 점) 순이었으며, 환류효과성이 3.12 점으로 가장 낮게 나타났다.

2) 세부사항 파악 모듈

세부사항 파악 모듈은 담화·대화에 제시된 특정 정보를 사실적으로 파악하는 능력을 측정하는 것으로서, '할·한·부탁한 일', '숫자 파악', '5W1H 세부사항', '도표 내용 일치', '그림 대화 일치', '대화·담화 내용 일치', '대화·담화 언급 일치' 등의 문항유형이 여기에 속한다.

첫째, '할·한·부탁한 일' 유형에 대한 타당성은 전체평균 3.77 점이었으며, 5 개 하위 영역 중 실제성에 대한 타당성이 3.96 점으로 가장 높게 나타났다. 다음으로 연계성(3.93 점), 환류효과성(3.87 점), 상호작용성(3.72 점) 순이었으며, 출제용이성이 3.38 점으로 가장 낮게 나타났다.

둘째, '숫자 파악' 유형에 대한 타당성은 전체평균 3.68 점이었으며, 5 개 하위 영역 중 실제성에 대한 타당성이 3.84 점으로 가장 높게 나타났다. 다음으로 연계성(3.81 점), 상호작용성(3.77 점), 환류효과성(3.69 점) 순이었으며, 출제용이성이 3.3 점으로 가장 낮게 나타났다.

셋째, '5W1H 세부사항' 유형에 대한 타당성은 전체평균 3.69 점이었으며, 5 개 하위 영역 중 실제성에 대한 환류효과성이 3.37 점으로 가장 높게 나타났다. 다음으로 연계성(3.86 점), 실제성(3.79), 상호작용성(3.7 점) 순이었으며, 출제용이성이 3.22 점으로 가장 낮게 나타났다.

넷째, '도표 내용 일치' 유형에 대한 타당성은 전체평균 3.56 점이었으며, 5 개 하위 영역 중 상호작용성에 대한 타당성이 3.74 점으로 가장 높게 나타났다. 다음으로 연계성(3.68 점), 환류효과성(3.63 점), 실제성(3.59 점) 순이었으며, 출제용이성이 3.15 점으로 가장 낮게 나타났다.

다섯째, '그림 대화 일치' 유형에 대한 타당성은 전체평균 3.34 점이었으며, 5 개 하위 영역 중 연계성에 대한 타당성이 3.58 점으로 가장 높게 나타났다. 다음으로 환류효과성(3.33 점), 실제성(3.32 점), 상호작용성(3.26 점) 순이었으며, 출제용이성이 3.21 점으로 가장 낮게 나타났다.

여섯째, '대화·담화 내용 일치' 유형에 대한 타당성은 전체평균 3.62 점이었으며, 5 개 하위 영역 중 연계성에 대한 타당성이 4.01 점으로 가장 높게 나타났다. 다음으로 실제성(3.83 점), 환류효과성(3.81 점), 상호작용성(3.6 점) 순이었으며, 출제용이성이 3.36 점으로 가장 낮게 나타났다.

일곱째, '대화·담화 언급 일치' 유형에 대한 타당성은 전체평균 3.67 점이었으며, 5 개 하위 영역 중 연계성에 대한 타당성이 3.9 점으로 가장 높게

교보문고 KYOBO Book Centre

나타났다. 다음으로 환류효과성(3.8 점), 실제성(3.75 점), 상호작용성(3.69 점) 순이었으며, 출제용이성이 3.21 점으로 가장 낮게 나타났다.

3) 복합 모듈

복합 모듈은 상향, 하향, 상호작용 이해방식을 작동하여 보다 긴 담화·대화문의 대의 파악, 세부사항 파악 등을 측정하는 것으로서, '1 대화·담화문 2 문항' 유형이 여기에 속한다.

'1 대화·담화문 2 문항' 유형에 대한 타당성은 전체평균 3.64 점이었으며, 5 개 하위 영역 중 연계성에 대한 타당성이 3.98 점으로 가장 높게 나타났다. 다음으로 상호작용성(3.85 점)이었으며, 실제성과 환류효과성은 3.8 점으로 동일한 수치를 보였으며, 출제용이성이 2.75 점으로 가장 낮게 나타났다.

4) 간접 말하기 모듈

간접 말하기 모듈은 주어진 대화·담화문을 통해 이해한 바를 가상의 말하기에 적용할 수 있는 능력을 측정하는 것으로서, '짧은 대화 응답', '그림 자료(길찾기)', '대화에 응답', '담화에 응답' 등의 유형이 여기에 속한다.

첫째, '짧은 대화 응답' 유형에 대한 타당성은 전체평균 3.83 점이었으며, 5 개 하위 영역 중 연계성에 대한 타당성이 4.13 점으로 가장 높게 나타났다. 다음으로 실제성(3.99 점), 환류효과성(3.93 점), 상호작용성(3.7 점) 순이었으며, 출제용이성이 3.4 점으로 가장 낮게 나타났다.

둘째, '그림 자료(길찾기)' 유형에 대한 타당성은 전체평균 2.97 점으로 낮게 나타났다. 5 개 하위 영역 중 출제용이성에 대한 타당성이 3.25 점으로 가장 높게 나타났으며, 다음으로 연계성(3.1 점), 실제성(3 점), 환류효과성(2.79 점) 순이었다. 상호작용성이 2.71 점으로 가장 낮게 나타났다.

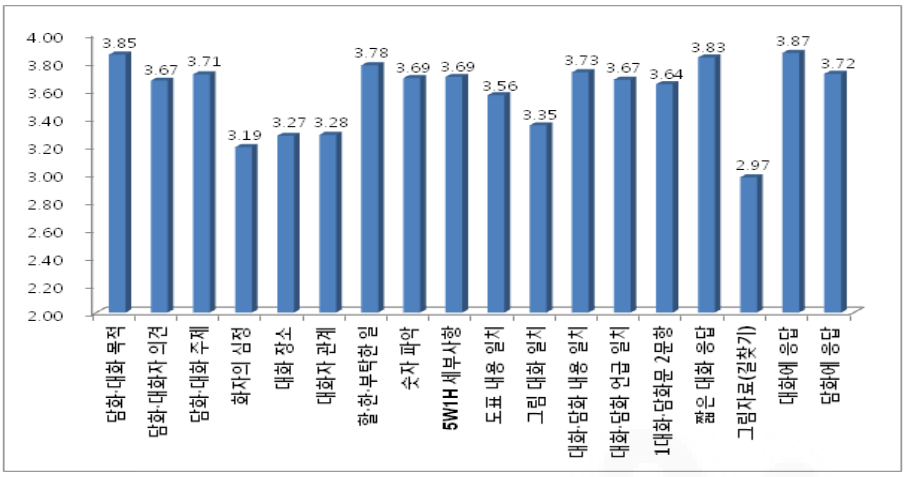
셋째, '대화에 응답' 유형에 대한 타당성은 전체평균 3.86 점이었으며, 5 개 하위 영역 중 환류효과성에 대한 타당성이 4.18 점으로 가장 높게 나타났다. 다음으로 연계성(4.15 점), 실제성(4.08 점), 상호작용성(3.94 점) 순이었으며, 출제용이성이 2.96 점으로 가장 낮게 나타났다.

넷째, '담화에 응답' 유형에 대한 타당성은 전체평균 3.71 점이었으며, 5 개 하위 영역 중 환류효과성에 대한 타당성이 3.95 점으로 가장 높게 나타났다. 다음으로 연계성(3.9 점), 실제성(3.86 점), 상호작용성(3.85 점) 순이었으며, 출제용이성이 3 점으로 가장 낮게 나타났다.

이상의 결과를 종합하여 듣기 유형 전체에 대한 타당성 평가 결과를 요약하면 그림 2 와 같다.

교보문고 KYOBO Book Centre

그림 2
듣기 유형에 대한 타당성 평가 결과(전체 평균 = 3.58 점)



구체적으로 살펴보면, 18 개 듣기 유형 타당성 평가의 전체 평균은 3.58 점으로 나타났으며 '그림 자료(길찾기)' 유형만이 3 점 미만으로 가장 낮게 나타났다. 구체적으로 살펴보면, 전체 평균보다 높게 나타난 유형은 대화에 응답(3.87 점), 담화·대화 목적(3.85 점), 짧은 대화 응답(3.83 점), 할·한·부탁한 일(3.78 점), 대화·담화 내용 일치(3.73 점), 담화에 응답(3.72 점), 담화·대화 주제(3.71 점), 숫자 파악(3.69 점), 5W1H 세부사항(3.69 점), 담화·대화자 의견(3.67 점), 대화·담화 언급 일치(3.67 점), 1 대화·담화문 2 문항(3.64 점) 등 12 개였다. 반면 전체 평균 보다 낮게 나타난 유형은 도표 내용 일치(3.56 점), 그림 대화 일치(3.35 점), 대화자 관계(3.28 점), 대화 장소(3.27 점), 화자의 심정(3.19 점), 그림 자료(길찾기)(2.97 점) 등 6 개였다.

2. 읽기 문항 유형 타당성 분석 결과

현행 통합형 수능 영어 영역의 20 개의 읽기 유형의 각 유형별 타당성 평가 결과를 요약하면 다음과 같다.

1) 문법·어휘 모듈

문법·어휘 모듈은 문단의 전체적 의미나 문장 간의 의미적 관련성을 고려하여 문법성을 판단할 수 있는 능력을 측정하는 것으로서, '어법', '어휘', '지칭추론(지시어)' 등의 유형이 여기에 속한다.

첫째, '어법' 유형에 대한 타당성은 전체평균 3.67 점이었으며, 5 개 하위

교보문고 KYOBO Book Centre

영역 중 연계성에 대한 타당성이 3.98 점으로 가장 높게 나타났다. 다음으로 환류효과성(3.94 점), 상호작용성(3.65 점), 실제성(3.61 점) 순이었으며, 출제용이성이 3.15 점으로 가장 낮게 나타났다.

둘째, '어휘' 유형에 대한 타당성은 평균 전체평균 3.79 점이었으며, 5 개 하위 영역 중 환류효과성에 대한 타당성이 4.12 점으로 가장 높게 나타났다. 다음으로 연계성(4.01 점), 상호작용성(3.93 점), 실제성(3.86 점) 순이었으며, 출제용이성이 3.03 점으로 가장 낮게 나타났다.

셋째, '지칭 추론(지시어)' 유형에 대한 타당성은 전체평균 3.08 점이었으며, 5 개 하위 영역 중 연계성에 대한 타당성이 3.4 점으로 가장 높게 나타났다. 다음으로 상호작용성(3.16 점), 환류효과성(3.08 점), 실제성(3.05 점) 순이었으며, 출제용이성이 2.72 점으로 가장 낮게 나타났다.

2) 대의 파악 모듈

대의 파악 모듈은 글을 읽고 전체적인 내용을 이해 · 추론하는 능력을 측정하는 것으로서, '글의 목적', '필자의 주장', '글의 주제', '글의 요지', '글의 제목', '글의 분위기·심경' 등의 문항유형이 여기에 속한다.

첫째, '글의 목적' 유형에 대한 타당성은 전체평균 3.66 점이었으며, 5 개 하위 영역 중 연계성에 대한 타당성이 3.96 점으로 가장 높게 나타났다. 다음으로 실제성(3.8 점), 환류효과성(3.7 점), 출제용이성(3.45 점) 순이었으며, 상호작용성이 3.39 점으로 가장 낮게 나타났다.

둘째, '필자의 주장' 유형에 대한 타당성은 전체평균 3.63 점이었으며, 5 개 하위 영역 중 연계성에 대한 타당성이 3.95 점으로 가장 높게 나타났다. 다음으로 환류효과성(3.71 점), 실제성(3.7 점), 상호작용성(3.56 점) 순이었으며, 출제용이성이 3.21 점으로 가장 낮게 나타났다.

셋째, '글의 주제' 유형에 대한 타당성은 전체평균 3.87 점이었으며, 5 개 하위 영역 중 연계성에 대한 타당성이 4.25 점으로 가장 높게 나타났다. 다음으로 환류효과성(4.05 점), 실제성(3.95 점), 상호작용성(3.76 점) 순이었으며, 출제용이성이 3.37 점으로 가장 낮게 나타났다.

넷째, '글의 요지' 유형에 대한 타당성은 전체평균 3.78 점이었으며, 5 개 하위 영역 중 연계성에 대한 타당성이 4.08 점으로 가장 높게 나타났다. 다음으로 실제성(3.94 점), 환류효과성(3.88 점), 상호작용성(3.65 점) 순이었으며, 출제용이성이 3.36 점으로 가장 낮게 나타났다.

다섯째, '글의 제목' 유형에 대한 타당성은 전체평균 3.7 점이었으며, 5 개 하위 영역 중 연계성에 대한 타당성이 4.05 점으로 가장 높게 나타났다. 다음으로 환류효과성(3.89 점), 상호작용성(3.72 점), 실제성(3.67 점) 순이었으며, 출제용이성이 3.18 점으로 가장 낮게 나타났다.

여섯째, '글의 분위기·심경' 유형에 대한 타당성은 전체평균 3.09 점이었으며,

교보문고 KYOBO Book Centre

5 개 하위 영역 중 연계성에 대한 타당성이 3.43 점으로 가장 높게 나타났다. 다음으로 실제성(3.08 점), 상호작용성(3.05 점), 환류효과성(3.02 점) 순이었으며, 출제용이성이 2.9 점으로 가장 낮게 나타났다.

3) 상호 작용 모듈

상호 작용 모듈은 상호 작용이해 방식을 작동하여 지문에 명시되지 않는 정보를 논리적으로 추론할 수 있는 능력을 측정하는 것으로서, '빈칸 추론', '연결사 추론' 등의 유형이 여기에 속한다.

첫째, '빈칸 추론' 유형에 대한 타당성은 전체평균 3.84 점이었으며, 5 개 하위 영역 중 상호작용성에 대한 타당성이 4.23 점으로 가장 높게 나타났다. 다음으로 환류효과성(4.14 점), 연계성(4.13 점), 실제성(3.97 점) 순이었으며, 출제용이성이 2.72 점으로 가장 낮게 나타났다.

둘째, '연결사 추론' 유형에 대한 타당성은 전체평균 3.66 점이었으며, 5 개 하위 영역 중 연계성에 대한 타당성이 3.9 점으로 가장 높게 나타났다. 다음으로 환류효과성(3.77 점), 상호작용성(3.75 점), 실제성(3.59 점) 순이었으며, 출제용이성이 3.3 점으로 가장 낮게 나타났다.

4) 세부사항 파악 모듈

세부사항 파악 모듈은 글에 제시된 특정 정보를 사실적으로 이해하는 능력을 측정하는 것으로서, '실용 자료 내용 일치', '도표 내용 일치', '내용 일치·불일치' 등의 유형이 여기에 속한다.

첫째, '실용 자료 내용 일치' 유형에 대한 타당성은 전체평균 3.55 점이었으며, 5 개 하위 영역 중 연계성에 대한 타당성이 3.83 점으로 가장 높게 나타났다. 다음으로 실제성(3.81 점), 환류효과성(3.59 점), 상호작용성(3.45 점) 순이었으며, 출제용이성이 3.05 점으로 가장 낮게 나타났다.

둘째, '도표 내용 일치' 유형에 대한 타당성은 전체평균 3.29 점이었으며, 5 개 하위 영역 중 연계성에 대한 타당성이 3.56 점으로 가장 높게 나타났다. 다음으로 실제성(3.43 점), 상호작용성(3.31 점), 환류효과성(3.27 점) 순이었으며, 출제용이성이 2.89 점으로 가장 낮게 나타났다.

셋째, '내용 일치·불일치' 유형에 대한 타당성은 전체평균 3.46 점이었으며, 5 개 하위 영역 중 연계성에 대한 타당성이 3.66 점으로 가장 높게 나타났다. 다음으로 환류효과성(3.5 점), 실제성(3.44 점), 상호작용성(3.4 점) 순이었으며, 출제용이성이 3.3 점으로 가장 낮게 나타났다.

교보문고 KYOBO Book Centre

5) 복합 모듈

복합 모듈은 상향, 하향, 상호작용 이해방식을 작동하여 보다 긴 지문의 대의 파악, 세부사항 파악, 논리적 추론, 쓰기 능력 등을 측정하는 것으로서, '1 지문 2 문항'과 '1 지문 3 문항'의 유형이 여기에 속한다.

첫째, '1 지문 2 문항' 유형에 대한 타당성은 전체평균 3.76 점이었으며, 5 개 하위 영역 중 상호작용성에 대한 타당성이 4.05 점으로 가장 높게 나타났다. 다음으로 환류효과성(4.04 점), 연계성(4.03 점), 실제성(3.85 점) 순이었으며, 출제용이성이 2.81 점으로 가장 낮게 나타났다.

둘째, '1 지문 3 문항' 유형에 대한 타당성은 전체평균 3.68 점이었으며, 5 개 하위 영역 중 연계성에 대한 타당성이 4 점으로 가장 높게 나타났다. 다음으로 상호작용성(3.95 점), 환류효과성(3.93 점), 실제성(3.76 점) 순이었으며, 출제용이성이 2.75 점으로 가장 낮게 나타났다.

6) 간접 쓰기 모듈

간접 쓰기 모듈은 주어진 지문을 통해 이해한 바를 가상의 쓰기에 적용할 수 있는 능력을 측정하는 것으로서, '흐름에 무관한 문장', '글의 순서 찾기', '주어진 문장 삽입', '문단 요약' 등의 유형이 여기에 속한다.

첫째, '흐름에 무관한 문장' 유형의 타당성은 전체평균 3.72 점이었으며, 5 개 하위 영역 중 연계성에 대한 타당성이 3.98 점으로 가장 높게 나타났다. 다음으로 상호작용성과 환류효과성이 3.9 점으로 동일하게 나타났으며, 연이어 실제성(3.68 점) 순이었다. 출제용이성이 3.12 점으로 가장 낮게 나타났다.

둘째, '글의 순서 찾기' 유형에 대한 타당성은 전체평균 3.78 점이었으며, 5 개 하위 영역 중 연계성에 대한 타당성이 4.11 점으로 가장 높게 나타났다. 다음으로 상호작용성(4.1 점), 환류효과성(3.99 점), 실제성(3.77 점) 순이었으며, 출제용이성이 2.92 점으로 가장 낮게 나타났다.

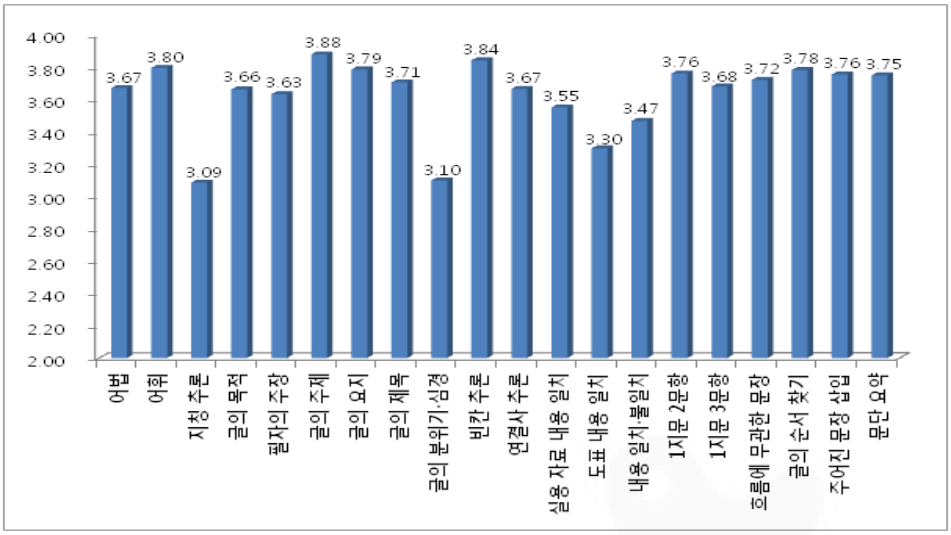
셋째, '주어진 문장 삽입' 유형에 대한 타당성은 전체평균 3.75 점이었으며, 5 개 하위 영역 중 환류효과성에 대한 타당성이 4.05 점으로 가장 높게 나타났다. 다음으로 연계성(4.03 점), 상호작용성(4 점), 실제성(3.74 점) 순이었으며, 출제용이성이 2.95 점으로 가장 낮게 나타났다.

넷째, '문단 요약' 유형에 대한 타당성은 전체평균 3.74 점이었으며, 5 개 하위 영역 중 연계성에 대한 타당성이 4.11 점으로 가장 높게 나타났다. 다음으로 상호작용성(4.05 점), 환류효과성(4 점), 실제성(3.98 점) 순이었으며, 출제용이성이 2.58 점으로 가장 낮게 나타났다.

이상의 결과를 종합하여 읽기 유형 전체에 대한 타당성 평가 결과를 요약하면 그림 3 과 같다.

교보문고 KYOBO Book Centre

그림 3
읽기 문항 유형에 대한 타당성 평가 결과(전체 평균 = 3.63 점)



구체적으로 살펴보면, 20 개 읽기 유형 타당성 평가의 전체 평균은 3.63 점으로 나타났다. 전체 평균보다 높게 나타난 유형은 글의 주제(3.88 점), 빈칸 추론(3.84 점), 어휘(3.80 점), 글의 요지(3.79 점), 글의 순서 찾기(3.78 점), 1 지문 2 문항(3.76 점), 주어진 문장 삽입(3.76 점), 문단 요약(3.75 점), 흐름에 무관한 문장(3.72 점), 글의 제목(3.71 점), 1 지문 3 문항(3.68 점), 어법(3.67 점), 연결사 추론(3.67 점), 글의 목적(3.66 점), 필자의 주장(3.63 점) 등 15 개였다. 반면 전체 평균보다 낮게 나타난 유형은 실용 자료 내용 일치(3.55 점), 내용 일치·불일치(3.47 점), 도표 내용 일치(3.30 점), 글의 분위기·심경(3.10 점), 지칭 추론(3.09 점) 등 5 개였다.

V. 결론 및 제언

2015 학년도 수능 영어영역 검사지를 구성하기 위하여, 본 연구에서는 2014 학년도 수준별 외국어(영어) 영역의 문항 유형에 대한 내적 타당성을 분석하여, 이를 통합형 영어 시험 검사지 구성을 위하여 제언하고자 한다.

이 같은 연구 목적을 구현하기 위해, 본 연구에서는 수능 참여 경험을 갖고 있는 교수 및 교사를 대상으로 문항 유형 타당성 평가의 기준과 평가 요소에 따라 문항 평가자에게 현행 듣기와 읽기 문항 유형에 대한 타당성을 Likert 5 점 척도로 평가하게 하였다. 내적 타당성 평가를 분석한 결과는 다음과

교보문고 KYOBO Book Centre

같았다. 2015 학년도 수능 영어 듣기 영역 문항 유형별 타당성 설문조사 분석결과, 대의 파악과 세부사항 파악, 그리고 복합 모듈에 해당하는 문항 유형은 모두 3.0 이상의 평가를 나타내었으나, 간접 말하기 모듈의 그림 자료(길찾기) 문항 유형은 3.0 에 미달하여 해당 유형만 문항 유형 은행에서 배제하는 것이 바람직한 것으로 나타났다.

한편 수능 영어 읽기 영역 문항 유형별 타당성 설문조사 분석결과, 문법. 어휘, 대의 파악, 세부사항 파악, 상호작용(빈칸), 복합 모듈, 그리고 간접 쓰기 모듈의 전체 문항 유형이 모두 3.0 이상의 평가를 나타내었고, 타당성 분석 결과 3.0 에 미달하는 문항 유형은 없으므로, 전체 읽기 모듈의 문항 유형은 2015 학년도 통합형 영어 읽기 영역에서 모두 사용하는 것이 바람직한 것으로 나타났다. 문항 유형별 타당성 평가 결과를 바탕으로 하고, 검사지 구성 원칙 및 정부의 쉬운 수능 정책을 적용하여 2015 수능 통합형 영어 시험의 검사지 구성을 위한 제언은 다음과 같다.

첫째, 2015 학년도 통합형 수능 영어 영역의 듣기 평가 문항 수는 현행 수준별 수능보다 5 개가 줄어든 17 문항이며, 읽기 평가의 문항 수는 현행보다 5 개 늘어난 28 문항이다. 이에 따라 듣기 영역과 읽기 영역에서 어떤 모듈에서 어떤 문항 유형을 줄이고, 또 늘릴 것인가를 결정해야 한다.

둘째, 2015 학년도 통합형 수능 영어 영역의 검사지도 2014 학년도 수준별 수능에서와 마찬가지로 모듈형 검사지 구성 방식에 따라 구성하는 것이 타당하다. 왜냐하면, 2015 학년도 통합형 수능의 총 문항 수가 45 개이므로 수능의 평가 요소와 교육과정의 시험 과목의 성취 기준과 연계할 경우, 부득이하게 수능의 평가 요소와 성취 기준이 1:1 로 대응될 수 없기 때문이다. 이를 구체적으로 설명하면, 과거 수능처럼 고정형 검사지 구성 방식에 따라 검사지를 구성할 경우, 예를 들어, 대의 파악 모듈 중, 특정 3 유형, 가령, 목적.주제·요지 파악 문항 유형만 검사지에 포함되고, 나머지 유형, 가령, 주장·제목·시사점 파악 유형은 검사지에서 배제될 것이며, 이는 교실 수업에서도 이 유형을 학습할 기회도 배제될 것이기 때문이다. 그러나 모듈형 검사지 구성 방식에 따라 검사지를 구성할 경우, 대의 파악 모듈에 속한 문항 유형 중에서 매 시험마다 그 구성을 달리함으로써 교실 수업에서 특정 대의 파악 유형을 배제할 수 없을 것이기 때문이다.

셋째, 듣기 영역의 문항 수를 축소할 경우, 상보성 원칙에 따라 현행 수준별 수능 B 형의 듣기 영역에서 동일 유형이나 유사한 유형이 두 개 이상 있는 경우, 그 중 한 개를 삭제할 필요가 있다. 읽기 영역에서 문항 수를 확대할 경우, 역시 상보성 원칙에 따라 현행 A 형에는 있지만, B 형에는 없는 문항 유형을 새로 추가할 필요가 있다.

넷째, 통합성 원칙에 따라 단일형 문항(순수 듣기 및 읽기 문항)과 통합형

교보문고 KYOBO Book Centre

문항(간접 말하기 및 쓰기 문항)의 적정 구성 비율을 설정할 경우, 통합형 문항의 비율은 현행 수준별 수능에서의 비율을 그대로 유지하는 것이 타당하다. 왜냐하면, 수능 영어 영역 시험이 갖고 있는 본질적인 문제점은 말하기와 쓰기 능력을 직접 평가할 수 없다는 것이다. 이런 상황을 고려한다면, 비록 간접 평가라 하더라도 말하기와 쓰기 모듈 관련 문항의 비율은 현행과 동일하게 유지하거나 좀 더 늘리는 방안을 검토해야 할 것이다.

다섯째, 주축성 원칙에 따라 주축형 문항과 주변형 문항의 적정 구성 비율을 설정할 경우, 현행 A 형을 대비해서 학습해 온 수험생이 주로 예체능계 및 중하위 등급임을 고려하여, 주변형 문항 유형(예, 대의 파악 유형, 내용 일치 유형)의 비율을 현행보다 늘릴 필요가 있다. 따라서 새로 추가되는 문항은 주변형 문항 유형으로 대체하는 것이 바람직하다.

여섯째, 위계성 원칙에 따라 L 형, M 형, H 형 문항의 적정 구성 비율을 설정할 경우, 현행 B 형의 난도를 통합형 수능에서도 그대로 유지할 필요가 있으므로 H 형 문항의 수는 현행대로 유지하지만, 현행 A 형 응시생의 대부분이 중하위 등급에 속하는 것을 고려하여, L 형과 M 형 문항의 비율은 현행보다 늘릴 필요가 있다. 따라서 읽기 영역에서 새로 추가되는 문항은 L 형과 M 형 문항 유형으로 대체할 필요가 있다.

일곱째, 본 연구는 출제전문가를 대상으로 하여 제한된 설문 참여자의 의견을 중심으로 하였으나, 우리나라에서 교육에서 수능이 차지하는 비중을 고려할 때 후속연구는 설문참여자의 폭을 확대하고 다양화하여 장기적 관점에서 연구할 필요가 있다.

참고문헌

강문구, 김용명. (2013). 2014 수능 수준별 영어 시험의 검사지 구성을 위한 외국어 영역 문항 유형에 대한 내외적 타당성 분석. *영어교과교육, 12*(2), 1-35.

교육과학기술부. (2008a). *외국어과 교육과정(I)*: 교육과학기술부 고시 제 2008-160 호 [별책 14]. 서울: 교육과학기술부.

교육부. (2013). 대입전형 간소화 및 대입제도 발전 방안. 보도자료 (2013. 08.27.)

교육부 & 한국대학교육협의회. (2013, 09.23.). 2015 학년도 및 2016 학년도 대입제도 및 2015 학년도 대입전형 기본사항. 보도자료.

김낙복. (2008). 대학수학능력시험 외국어(영어) 영역의 코퍼스 언어학적 어휘 비교 분석. *영어어문교육, 14*(4), 201-221.

교보문고 KYOBO Book Centre

김용명. (2010). 국가영어능력평가시험(NEAT)의 검사지 구성의 원칙과 절차: 문항 유형 확정 모델. *영어어문교육, 16*(4), 153-184.

김용명, 권오량, 이완기, 김진석, 고현숙. (2011). *수능 외국어(영어) 영역 평가 방법 개선 후속 연구*. 한국교육과정평가원, 연구보고 CAT 2011-2.

이경숙. (1999). 문제 수, 지문 길이, 지문 친숙도가 영어 청해와 독해시험에 미치는 영향. *영어교육, 54*(4), 327-349.

장경숙. (2004). 대학수학능력시험 외국어(영어)영역 읽기 난이도 예측 모형 개발. *외국어교육, 11*(1), 111-130.

한국교육과정평가원. (2005-2010). *수능 외국어(영어) 영역 문항 분석* (비공개 자료집).

Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice: Designing and developing useful language tests*. Oxford: Oxford University Press.

Brown, H. D. (2007). *Teaching by principles: An interactive approach to language pedagogy*. New York: Longman.

Ellis, R. (2003). *Task-based language learning and teaching*. Oxford: Oxford University Press.

예시언어(Examples in): Korean
적용가능 언어(Applicable languages): English
적용가능수준(Applicable Levels): Secondary/College

김용명
한국교육과정평가원 수능본부 출제연구실
100-784 서울시 중구 정동 15-5 정동빌딩
Tel: 02-3704-3533
Email: mencius@kice.re.kr

이병천
한국교육과정평가원 수능본부 출제연구실
100-784 서울시 중구 정동 15-5 정동빌딩
Tel: 02-3704-3956
Email: bclee@kice.re.kr

교보문고 KYOBO Book Centre

박지선
한국교육과정평가원 수능본부 출제연구실
100-784 서울시 중구 정동 15-5 정동빌딩
Tel: 02-3704-3607
Email: jspark@kice.re.kr

황필아
한국교육과정평가원 교육평가본부
100-784 서울시 중구 정동 15-5 정동빌딩
Tel: 02-3704-3698
Email: pilahh@kice.re.kr

Received in July 15th, 2014
Reviewed in August 26th, 2014
Revised version received in September 5th, 2014

KYOBO 교보문고
교보문고 KYOBO Book Centre