

국어과 국가수준 학업성취도 평가의 문항 개선 방안 연구

—2008~2010년 검사지의 타당도 검토를 중심으로—

김영란 · 김인숙*

<차 례>

- I. 머리말
- II. 선행 연구 검토
- III. 연구 방법
- IV. 연구 결과 및 논의
- V. 맺음말

I. 머리말

우리나라는 국가가 주도적으로 학교교육의 질을 관리해 온 전통이 깊다. 정부 수립 이후 줄곧 국가가 교육과정을 제정·보급하여 교육 내용의 준거를 제공하였고, 교육평가 관련 정책도 국가가 주도적으로 수립하여 학교 현장으로 보급·확산함으로써 학교교육의 질을 확보하고자 하였다. 이러한 맥락 속에서 국가수준 학업성취도 평가(이하 학업성취도 평가)도 국가 차원에서 학교교육의 성과를 점검하고 각종 교육 정책 및 교육 방안 수립에 필요한 기초 자료를 산출하기 위해 시행되고 있다. 학업성취도 평

* 김영란(한국교육과정평가원, 제1저자) · 김인숙(한국교육과정평가원, 교신저자)

가는 체계적으로 이루어지지 못하다가 2003년에 와서야 일정한 체계를 갖추어(정구향 외, 2004 : 1-3) 일부 학교를 대상으로 시행되었고, 2008년부터는 전 학교를 대상으로 시행되어 올해 5년째를 맞고 있다.

이렇게 학업성취도 평가가 대상 학년의 전체 학생이 치르는 시험(해당 과목)인 만큼 검사지의 타당도를 검토할 필요가 있다. 그리고 전수 대상 시행 이후에는 학업성취도 평가 결과가 교육 현장에서 중요한 지표로 활용되는 등, 그 중요성에 비추어 볼 때에도 검사지의 타당도를 점검하는 것은 일차적으로 요구되는 사안이다. 더불어 2008년 이후에는 우리나라 학생 전체의 학업성취도를 보여주는 평가 결과가 축적되고 있는데, 이 데이터는 교육적으로 유의미한 연구를 위해 활용될 필요가 있다. 그러나 아직까지는 학업성취도 평가에 대한 기본 연구 보고서 외에 국어과 학업성취도 평가 데이터를 바탕으로 한 연구가 많이 이루어지지 못하였다.

이에 본 연구는 학업성취도 평가 결과 데이터를 바탕으로 검사지의 타당도를 검토하고, 그 결과를 바탕으로 평가 문항의 개선점을 제안하려고 한다. 전수 대상 시행 이후의 국어과 학업성취도 평가 문항은 매년 평가 결과 보고서(김도남 외, 2011 ; 남민우 외, 2010 ; 남민우 외, 2009)에서 문항별 정답률, 변별도 등을 바탕으로 그 양호도가 기입되고 있다. 그러나 이들 보고서에는 문항별 평가 결과에 대한 일차적인 분석 기술이 선행되어야 하다 보니, 검사지 전체에 대한 심층적인 논의를 게재하기 어려운 형편에 있다. 또한 학업성취도 평가의 문항이 타당성 있는 것이 되도록 많은 사람들이 노력을 기울이고 있지만,¹⁾ 동시에 평가 결과라는 객관적인 데이터를 바탕으로 시험의 타당도를 다각도로 점검하는 작업도 필요하다. 이를 통해서 좀더 신뢰롭고 타당한 검사지가 될 수 있는 개선점을 찾음으로써 학업성취도 평가의 신뢰성과 타당성을 보다 더 높일 수 있을 것이다.

이를 위하여 본 연구는 전수 시행 이후인 2008~2010년까지의 국어과

1) 이 연구 보고서들을 보면, 여러 전문가들이 문항을 작성, 검토, 수정, 보완하는 절차를 거치면서 학업성취도 평가 문항의 신뢰도와 타당도를 높이는 작업을 하는 것을 알 수 있다. 특히 2008년의 경우는 예비 검사 시행을 통해 평가 도구의 완성도를 높이는 과정을 거쳤음을 알 수 있다(남민우 외, 2009 : 37-39).

학업성취도 평가 결과(초, 중, 고)를 바탕으로 각 검사지의 신뢰도와 타당도를 점검하고, 그 결과에 따라서 문항 개선점을 제안하고자 한다. 각 검사지의 신뢰도와 타당도를 점검하기 위해서 검사지의 신뢰도 계수 산출, 문항별로 신뢰도 계수에 영향을 주는 정도 점검, 요인분석을 통한 요인적재량 검토 등을 수행하고, 정답률과 변별도 등 문항 분석 결과를 함께 검토한다.

아래 2장에서는 본 연구와 관련된 선행 연구를 검토하고, 3장에서는 연구 방법을 자세히 서술하며, 4장에서는 분석 결과와 그에 관한 논의를 펼치고 5장에서는 논의에 따른 제언을 하면서 마무리한다.

II. 선행 연구 검토

국가 차원의 학업성취도 평가가 오랫동안 시행된 것에 비해 국어과 학업성취도 평가 관련 연구는 많이 이루어지지 않았다. 또한 그 동안에 수집된 국어과 학업성취도 평가 관련 데이터에 비해, 한국교육과정평가원의 연구 보고서를 제외하고는 소수의 연구만이 데이터를 기반으로 이루어진 실정이다. 국어과 학업성취도 평가 관련 연구는 대부분이 전문가의 정성적인 평가를 바탕으로 평가의 타당도에 대한 비판적 검토가 주를 이루고 있다.

우선 전수 시행 이후, 국어과 학업성취도 평가의 전반적인 측면(김창원, 2011)과 시행 측면(남민우·남가영, 2011)을 비롯하여, 듣기(임철성, 2011), 읽기(김혜정, 2011), 쓰기(박영민, 2011), 문법(주세형, 2011), 문항 영역(최지현, 2011)에 대한 검토를 바탕으로 발전 방향이 제안된 바 있다. 이들 논의는 국어과 학업성취도 평가의 전반 또는 영역별 검토를 바탕으로, 보다 거시적인 관점에서 현재의 문제점을 지적하고 보다 타당성 있는 평가가 되는 데 필요한 점들을 제안하였다.

그리고 이들 논의보다는 미시적으로 영역별 문항에 대한 검토를 바탕

으로 평가의 내용 타당도를 점검한 연구들이 있다. 남민우(2011)는 2004-2009년 초6 문학 문항과 평가 결과를 바탕으로, 시 평가 틀이 지식과 수용 범주에 치우쳐 있고 창작 범주를 포괄하지 못하고, 특정 성취기준이 중복적으로 설정되어 있으며, 시적 능력을 측정하는 데에 부족한 수의 평가 요소가 설정되어 있음을 지적하였다. 그리고 초등 시교육이 특정 교육 내용에 치우쳐 있고, 교육과정과 교과서 등에서 학습 용어가 엄밀하게 사용되지 못한 점을 지적하였다(김영란 외, 2011 : 10에서 재인용). 문법 영역을 살핀 주세형(2009)는 2005-2008년 문법 문항에 대한 검토를 바탕으로, 문법 문항에서 ‘텍스트’가 활용되는 방식이 문법 교육적으로 유의미하게 다루어지지 못하고 있음을 비판하였다. 또한 이은희(2011)는 2004-2009년까지의 문법 문항에 대한 검토를 바탕으로, 지속적으로 낮은 정답률을 나타낸 성취기준에 대한 원인 점검, 교수·학습 대책 마련, 문항 구성상의 문제점 보완을 촉구하였다. 또한 텍스트를 활용한 문법 문항이 보다 실질적인 의미를 가질 수 있도록 구성되어야 하고, 수업 시간에 학습한 문법 용어가 평가 장면에서 배제되지 않아야 한다고 역설하였다.

또한 최근에는 평가를 점검하는 이론적 틀에 따라 국어과 학업성취도 평가를 점검한 논의가 있었는데(조재운, 2011 ; 김영란 외, 2011), 조재운(2011)은 평가의 적합성을 검토하는 틀로서 타당도, 신뢰도, 유용도, 투명도를 살펴보는 VRUT²⁾ 모형을 제안하면서 국어과 학업성취도 평가를 검토하였다. 연구자는 이 틀에 따라 학업성취도 평가가 인지적 능력, 정의적 능력, 인지적·정의적 능력의 사용으로 구성되어야 하는 종합적인 국어능력을 측정하는 도구로서는 부적합하여 구인 타당도를 확보하지 못하고, 적은 문항수로 인해 내용 타당도 확보도 어려우며, 서답형 채점의 신뢰도도 확보하지 못하고, 평가 결과표에 매우 제한적인 정보가 제공됨으로써 평가의 청자(학생, 학부모, 교사)가 얻는 유용성이 낮다고 하였다. 또한 평가 결과 공개로 인해 학교(지역) 간 경쟁 유발, 사교육 심화, 성적 우수 지역으로 학생 쏠림 현상, 문제풀이식 교육과정 운영 등의 부정적 결과가 있고, 평가

2) VRUT은 타당도(Validity), 신뢰도(Reliability), 유용도(Usefulness), 투명도(Transparency)의 영문 첫 글자를 조합한 것이다.

의 청자들이 평가 내용, 채점 과정 등에 대해 투명하게 알 수 없다고 비판하였다. 이 연구는 평가를 점검하는 이론적 틀에 따라 학업성취도 평가를 검토한 의의가 있고, 타당한 비판을 제기하였다. 그러나 학업성취도 평가의 타당도, 신뢰도, 유용도, 투명도에 대해 인상적인 평가를 하고, 연구자의 비판을 뒷받침할 객관적인 증거 제시가 부족하다는 점에서 아쉬움이 있다.

한편, 평가의 내용 타당도를 보다 집단적인 방법으로 점검한 논의가 있었다. 천경록(2009, 2010, 2011)은 평가가 교육과정(또는 성취기준)과 밀접한 관련성을 가져야 한다는 정합성(alignment)의 논리에 따라 국어과 평가를 검토할 필요가 있다고 보고, 이에 따라 학업성취도 평가의 틀이 되는 성취기준을 바탕으로, 2006년 초6, 중3, 고1 문항에 대해 검토하였다. 이들 논의에서는 Webb의 정합성 판정 모형에 따라 계량적, 전문가 집단적 방법을 통해 학업성취도 평가의 내용 타당도를 점검하였다.³⁾ 그 결과 2006년 초6, 고1 평가 문항은 성취기준과 문항이 측정하는 지식의 깊이가 일치하지 않는 경우가 있고, 해당 영역에 적합하지 않은 문항이 일부 있는 것으로 나타났다(천경록, 2010, 2011). 그리고 중3 문항의 경우는 지식의 깊이가 적합하지 않은 것이 일부 있는 것으로 나타났다(천경록, 2011). 이들 논의는 평가 문항의 내용 타당도를 전문가의 판단에 의존하면서도 집단적으로 점검하고, 그 결과로 학업성취도 평가 문항이 성취기준을 만족스럽게 구현해 내지 못하는 부분에 대해 지적인 점에서 의의가 있다. 그러나 감사지

3) Webb의 정합성 판정 모형은 평가의 영역 일치도, 지식의 범위 일치도, 표집의 균형성 정도, 지식의 깊이 일치도를 바탕으로 한다. 영역 일치도는 평가가 측정하고자 하는 교과 영역을 고려하여 적절한 수의 문항으로 구성되었는지를 판단하는 것이다. 하위 평가 영역별로 최소 여섯 문항 이상이 출제되어야 한다는 기준을 적용한다. 지식의 범위 일치도는, 평가가 측정하는 지식의 범위가 성취기준이 요구하는 지식의 범위와 유사한지를 판정하는 것이다. 이 기준을 만족하려면 성취기준 중에서 평가 문항으로 개발된 성취기준이 적어도 50% 이상이 되어야 한다. 표집의 균형성 정도는 어떤 영역에 속한 각 성취기준들이 문항 개발에서 고르게 반영되는지를 판정하는 것이다. 이는 특정 성취기준이 과도하게 반복되어 출제되는 것을 막기 위한 것이다. 지식의 깊이 일치도는, 평가 문항이 성취기준이 의도하는 정도의 지식이나 기능의 인지적 복잡성을 묻는지 여부를 판정하는 것이다(천경록, 2010 : 39-41, 김영란 외, 2011 : 9에서 재인용).

가 성취기준을 완전히 구현해 낼 수 없는 형식적 제약(선다형 위주이고 서답형 문항에서는 열린 반응을 묻기 어려우며 시간·지면 제약이 따르는 지필 고사라는 점)을 받는다는 점을 고려하지 않고, 지필 평가의 특성을 고려하지 않고 작성된 성취기준이라는 근본적인 한계가 있음을 간과하고 있어서 아쉬움이 있다.

이상에서 전문가의 판단에 근거하여, 타당한 국어과 학업성취도 평가를 위한 논의들이 있었음을 살펴보았다. 이들은 타당한 평가가 되기 위해서 필요한 점들을 제안하고 있기는 하지만 검사지에 대한 점검은 부족해서 현재의 국어과 학업성취도 평가 문항이 구체적으로 어떻게 개선되어야 하는지에 대해서는 시사점을 얻기 어렵다.

다른 한편으로 국어과 학업성취도 평가의 결과 데이터를 바탕으로 한 보고서가 있는데(김도남 외, 2011 ; 남민우 외, 2010 ; 남민우 외, 2009), 이들은 성별, 지역별, 학교유형별 등의 변인에 따른 평가 결과의 기초적 분석 내용을 주로 다루고 있어서 매년 평가 결과에 대한 정보를 얻을 수 있다. 그러나 평가 결과와 관련된 심층적인 탐구가 부족하고, 평가 문항에 대한 점검도 정답률, 변별도를 위주로 하는 것이어서 새로운 각도의 점검이 요구된다. 전수 대상 시행 이후의 학업성취도 평가의 중요성은 증대되었고, 그만큼 검사지도 더 높은 신뢰도와 타당도를 갖출 필요가 있다는 점을 감안하면 검사지에 대한 심층 연구가 필요한 상황이다.

평가 결과를 바탕으로 한 심층 연구의 필요성에 부응하고자 김영란 외(2011)은, 2009-2010년 평가 결과를 바탕으로 평가를 점검하는 이론적 틀인 CURRV 모형⁴⁾에 따라 국어과 학업성취도 평가의 전반에 대하여 검토하였다. 이 연구는 교사 대상의 설문 조사와 심층 면담을 통해서 학업성취도 평가가 현장에 미치는 결과, 유용성, 평가 관계자들의 역할과 책임에 대해서 조사하였고, 평가 결과 데이터를 바탕으로 통계적 분석을 하여

4) CURRV 모형은 어떤 평가를 점검할 때 그 평가가 현장에 미치는 결과(Consequences), 유용성(Usefulness), 평가 관련자들의 역할과 책임(Roles and Responsibilities), 신뢰도(Reliability), 타당도(Validity)를 점검할 필요가 있다고 보는 이론적 틀이다. 보다 자세한 내용은 김영란 외(2011)와 Afflerbach(2007, 조병영 외 역, 2010)을 참조할 수 있다.

검사지의 신뢰도와 타당도를 점검하였다. 이 연구는 조재윤(2011)의 접근과 마찬가지로 교육 현장에 영향력이 큰 평가를 이론적 틀에 따라 점검한 의의가 있고, 특히 그러한 점검을 데이터를 바탕으로 한 의의가 있다. 그러나 이 연구는 하나의 연구에서 국어과 학업성취도 평가의 결과, 유용성, 평가 관계자의 역할과 책임, 신뢰도와 타당도의 다섯 가지 측면을 두루 살펴본다 보니, 각각의 측면을 깊이 있게 점검하지 못한 한계를 지닌다.

이에 본 연구는 김영란 외(2011)에서 살핀, 국어과 학업성취도 평가 검사지의 신뢰도와 타당도 측면을 보다 면밀히 점검하여 이 논의를 좀더 확장하고자 한다. 또한 타당한 국어과 학업성취도 평가를 위한 전문가의 질적인 판단에 근거하는 것이 아니라, 평가 결과를 바탕으로 통계적 기법을 활용하여 국어과 학업성취도 평가의 신뢰도와 타당도를 점검하고자 한다. 국어과 학업성취도 평가의 타당도에 대한 전문가의 검토는 매우 의미 있는 것임에는 틀림없지만 평가 관련 데이터가 축적되고 있는 현실을 고려할 때, 이러한 데이터를 바탕으로 양적, 질적 연구가 활성화될 필요가 있다.

III. 연구 방법

본 연구에서는 전수 대상 시행 이후의 국어과 학업성취도 평가 검사지에 대해서 신뢰도와 타당도를 점검하고, 그 결과에 따라 검사지의 신뢰도와 타당도를 다소 떨어뜨리는 문항에 대해 그 특징을 분석하였다. 이를 위해서 <표 1>과 같이 2008~2010년⁵⁾ 국어과 학업성취도 평가 9개의 검사지에 대한 학생들의 평가 결과에 기초하여 각 문항별 변별도, 검사지 신뢰도, 각 문항 제거 시 신뢰도, 요인분석을 통한 각 문항의 요인적재량/loading)을 살펴보았다.

타당도에는 여러 가지가 있지만 앞에서 살펴본 바와 같이 국어과 학

5) 2011년 평가 결과는 현재 미발표된 상황임.

업성취도 평가의 내용 타당도에 대한 논의는 일부 이루어진 바 있다. 그리고 김영란 외(2011)에서 요인분석을 통한 구인타당도를 점검했는데, 이 결과는 국어과 학업성취도 평가 문항들이 대체로 하나의 요인으로 구성됨을 보여준 바 있다⁶⁾. 더불어 현재의 국어과 학업성취도 평가의 틀에 관한 연구인 정은영 외(2010 : 149)는 평가 목표를 ‘국어능력’을 측정하는 것으로 설정하고 그 아래에 듣기, 읽기, 쓰기, 문법, 문학의 영역을 두고, 어휘를 부수적인 영역으로 두고 있다. 이로부터 국어과 학업성취도 평가 검사지는 ‘국어능력’을 측정하는 내용으로 구성되어야 한다는 가정을 도출할 수 있다. 따라서 본 연구는 국어과 학업성취도 평가 검사지는 ‘국어능력’이라는 하나의 요인으로 구성된다는 가설을 바탕으로 요인분석을 통해 요인적 재량을 점검하였다.

〈표 1〉 분석 대상 검사지별 문항수와 인원수⁷⁾

연도 학교급	2008년		2009년		2010년	
	문항수	인원수	문항수	인원수	문항수	인원수
초등학교	40	23,601	30	634,106	29	605,102
중학교	40	30,034	40	647,623	34	660,218
고등학교	40	29,631	40	636,767	36	631,352

6) 김영란 외(2011)은 검사지의 구인타당도를 점검하기 위해 요인분석을 했는데, 국어과 학업성취도 평가 검사지를 몇 가지의 요인으로 구분하기 위한 이론적 근거가 미약한 채로 요인분석을 실시한 약점이 있다. 따라서 요인분석의 결과로 구분된 요인 1, 2, 3, 4의 특성이 무엇인지 제대로 규명하지 못한 한계를 지닌다. 그러나 이 요인분석의 결과에서는 국어과 학업성취도 평가 문항이 극소수의 문항을 제외하고는 대체적으로 하나의 요인으로 묶이는 것을 알 수 있었다. 본 연구는 이러한 발견을 바탕으로 국어과 학업성취도 평가 검사지가 하나의 요인으로 구성된다는 가설을 설정하였다.

7) 학업성취도 평가는 2009년까지는 고등학교 1학년을 대상으로 하다가 2010년에는 고등학교 2학년을 대상으로 하였고, 해를 거듭하면서 학생들의 부담을 줄이는 방향에서 문항수도 축소되었다. 표 안의 문항수는 서답형 문항을 포함한 것이고 인원수는 유효한 분석 대상 인원수이다. 2008년은 표집 데이터이다.

1. 신뢰도

신뢰도는 검사자가 측정하려는 것을 얼마나 안정적으로 일관성 있게 측정하는지를 나타내는 것으로, 본 연구에서는 크론바흐 알파(Cronbach's α)를 이용한 내적 일관성(internal consistency) 신뢰도로 살펴보았다. 이는 검사를 구성하고 있는 문항 간의 일관성의 정도를 나타내는 것으로, 개별 문항을 하나하나의 검사점으로 하여 상관관계를 구한다. 그런 뒤 이들의 평균 상호 상관관계를 신뢰도의 추정치로 삼는 방법으로 평균 상호 상관은 모든 가능한 반복법을 사용하여 구한 신뢰도 계수의 평균을 의미한다. 대체로 신뢰도(α)는 0.6 보다 작으면 내적 일치도가 결여된 것으로 판단하고, 0.6~0.7 이면 수용할 만한 것으로, 0.8~0.9 이상이면 높은 것으로 여겨진다(성태제, 2002).

2. 문항별 변별도

각 문항과 모든 문항의 총점 간의 상관으로 판별하는 변별도는 -1.00 ~ +1.00에 분포한다. 변별도가 양수(+)이면 정적 변별력을 갖고, 0에 가까우면 변별력이 거의 없다는 것을 의미하며, 음수(-)인 경우는 하위집단이 상위집단보다 더 높은 점수를 얻은 경우로써 이는 문항이 문제가 있음을 의미한다. 학업성취도 평가 결과 보고서(김도남 외, 2011)에서는 변별도 지수에 따라서 변별도 정도를 다음과 같이 구분하고 있다.

변별도	변별 정도
0.40 이상	높음
0.30 ~ 0.39	있음
0.20 ~ 0.29	낮음
0.10 ~ 0.19	매우 낮음
0.1 미만	없음

* 김도남 외(2011 : 63)의 표를 수정하여 제시함.

따라서 검사지 각 문항의 변별도를 점검하면 각 문항이 검사지에서 문항 본래의 기능을 하는지를 일차적으로 알 수 있다. 검사지 문항별 변별도는 김도남 외(2011), 남민우 외(2010, 2009)에서 보고된 바 있는데, 이를 참조하였다.

3. 문항 제거 시 신뢰도

문항 제거 시 신뢰도는 해당 문항이 제거되었을 때의 신뢰도를 나타내는 것으로 전체 신뢰도 계수보다 높게 나타나면, 해당 문항을 제거함으로써 검사지의 신뢰도가 올라가는 것을 의미하고, 전체 신뢰도 계수보다 낮은 문항은 제거하면 검사지의 신뢰도가 낮아지는 것을 의미한다. 따라서 각 문항을 제외했을 때의 신뢰도 계수를 검사지 전체 신뢰도와 비교해서, 각 문항이 검사지의 신뢰도를 낮추는 것인지를 점검하였다. 가령, 2010년 고등 문항은 검사지의 신뢰도가 0.903으로 매우 높은 신뢰도를 나타낸다. 그런데 문항 28번을 제외하고 신뢰도 계수를 산출해 보면 검사지의 신뢰도가 0.905로 높아지게 된다. 즉, 문항 28번은 이 검사지의 신뢰도를 떨어뜨리는 문항인 것이다. 이런 식으로 문항별로 검사지의 신뢰도에 어떤 영향을 주는지를 살펴보았다.

4. 요인분석을 통한 요인적재량

구인타당도는 연구자가 측정하고 있다고 믿는 가설적인 구인(국어능력)을 검사점수가 제대로 측정하고 있는지의 여부를 나타내는데, 이는 요인분석을 통해 점검할 수 있다. 요인분석은 변수들 간의 상호관계를 분석하여, 상관이 높은 변수들을 모아 요인으로 규명하고 그 요인의 의미를 부여하는 통계분석기법이다. 요인분석의 기본 원리는 항목들 간의 상관관계가 높은 것끼리 하나의 요인으로 묶어내며, 하나의 요인으로 묶여진 측정

항목들은 하나의 개념을 측정하는 것으로 간주할 수 있다. 이를 통해 검사점수가 검사의 사용 목적에 얼마나 부합하는가를 나타내는, 검사지 목적의 적합성을 알 수 있다. 즉, 검사점수가 조작적 정의에서 규명한 심리적 구인들을 제대로 측정하였는가를 검증할 수 있다.

요인적재량(factor loading)은 각 변수와 요인 간의 상관관계 정도를 나타낸다. 즉 요인적재량이 클수록 해당요인에 대한 설명력이 크고 각 변수들은 요인적재량이 가장 큰 요인에 속하게 된다.

본 연구는 앞서 밝힌 바와 같이 요인의 수를 ‘국어능력’이라는 하나로 설정하여 분석하였다. 이순목(1994)은 요인적재량을 0.3~0.4 이상인 기준을 제시하면서 이러한 기준을 너무 경직되게 적용시킬 필요가 없음을 설명한 바 있다. 본 연구는 이러한 논의를 토대로 하고, 국어과 검사지의 특성⁸⁾을 감안하여 0.3보다는 완화된 기준인 0.2를 기준으로 ‘국어능력’을 측정하는 요인으로 구분하였다.

IV. 연구 결과 및 논의

1. 신뢰도와 문항별 변별도

<표 2>와 같이 전수 대상 시행 이후의 국어과 학업성취도 평가 검사지의 신뢰도 계수는 0.791~0.914로 신뢰도가 높은 것으로 나타났다. 앞에서 언급한 바와 같이, 신뢰도 계수가 0.8 이상이면 신뢰도가 높은 것인데, 국어과 검사지들이 모두 신뢰로운 평가 도구였음을 알 수 있다.

8) ‘국어능력’에는 ‘독해력’과 ‘문법 지식’과 같이 다소 거리가 있는 특성이 함께 포함되어 있는 것을 가정한다.

〈표 2〉 검사지별 신뢰도 계수

연도 학교급	2008년	2009년	2010년
초등학교	0.863	0.824	0.791
중학교	0.912	0.914	0.873
고등학교	0.881	0.906	0.903

여기에서 나아가 세부적으로 각 문항 제거 시 신뢰도를 살펴본 결과, [부록 1]에서와 같이 2008년 초등, 2009년 고등, 2010년 중등 검사지의 경우에는 검사지 전체의 신뢰도를 낮추는 문항이 하나도 없었다. 즉, 모든 문항이 검사지의 신뢰도에 기여했음을 알 수 있다. 그러나 나머지 검사지에서는 다음 표와 같이 극히 일부분인 한두 개의 문항이 검사지의 신뢰도를 다소 낮추는 것임을 발견하였다. 이들 문항의 특징에 대해서는 타당도 분석 결과와 함께 아래에서 자세히 검토한다.

〈표 3〉 검사지의 신뢰도를 낮추는 문항(번호)

연도 학교급	2008년	2009년	2010년
초등학교	-	서답4	20
중학교	3	17	-
고등학교	4, 20	-	24, 28

더불어 각 검사지 문항별 변별도를 살펴보았는데, 검사지의 문항 대부분이 변별도가 있거나 높은 것(0.2 이상)으로 나타났다. 다만, 5개의 검사지에서 변별도가 매우 낮은 문항(0.19 미만)이 1~3개 포함된 것을 알 수 있었다. 검사지의 신뢰도를 낮추는 문항이 동시에 변별도도 매우 낮은 문항인 경우가 많은데, 두 경우가 반드시 일치하는 것은 아니었다(부록 1 참조).

2. 요인분석 결과

<표 4>의 설명분산은 각 검사지의 설명된 분산을 나타내는 것으로 ‘국어능력’ 요인이 전체분산을 16.48~23.99% 선에서 설명하는 것으로 나타났다. 즉, 가령 2009년 중등 검사 결과는 국어능력이 전체분산의 23.44%를 설명하는 것으로 풀이할 수 있다. 이는 국어과 학업성취도 평가 결과가 학생의 국어능력의 일부분을 보여주는 자료로 활용될 수는 있지만, 국어능력의 전체로 간주될 수 없다는 것을 시사한다.⁹⁾

〈표 4〉 검사지별 요인분석 결과

연도 학교급	2008년		2009년		2010년	
	0.2미만문항	설명분산(%)	0.2미만문항	설명분산(%)	0.2미만문항	설명분산(%)
초	8	16.48	19	17.11	20	17.26
중	3	22.83	17, 8	23.44	28	21.25
고	25, 9, 4, 20	17.42	20, 27	21.31	28	23.99

그리고 요인분석 결과, 요인적재량 0.2를 기준으로 각 검사지 대부분의 문항이 요인 1(국어능력)로 묶이는 것으로 나타났다. 즉, 대부분의 문항

- 9) 김영란 외(2011)에서 활용한 2009년, 2010년 데이터는 본 연구의 데이터와 동일한 것인데, 본 연구에서는 선다형과 서답형을 포함한 전체 문항을 대상으로 요인분석을 하여 <표 4>와 같은 결과를 얻었다. 반면, 김영란 외(2011)에서는 서답형 문항을 별도로 요인분석한 바 있는데, 위의 결과와 아래 표의 서답형 문항의 설명분산(%)을 비교해 볼 수 있다.

연도 학교급	2009년	2010년
초등학교	19.27	39.71
중학교	44.67	47.41
고등학교	44.59	51.30

*김영란 외(2011 : 51)의 표를 수정하여 제시함.

위 표와 같이 서답형 문항만의 설명분산이 검사지 전체의 설명분산보다 높고, 2010년 고등 검사지의 경우 그 설명력이 51.30%나 된다는 것을 알 수 있다. 즉, 검사지 전체의 결과보다는 서답형 문항에서 추출된 국어능력 요인이 전체분산에 대한 설명력이 더 높다는 것을 알 수 있다.

이 국어능력을 측정하는 특성을 가진 문항으로 구분되었다. 다만 검사지 별로 극히 일부 문항이 요인 1로 묶이지 않는 것을 볼 수 있었다. <표 4>와 같이, 여섯 개 검사지에서는 이러한 문항이 단 1개, 두 개의 검사지에서는 2개, 한 개의 검사지에서는 4개가 있는 것으로 나타났다(각 검사지 문항별 요인적재량은 [부록 2] 참조).

요인 1로 구분되지 않는 문항들은 국어능력과는 거리가 있는 특성을 측정하는 문항으로, 국어능력 측정 기능이 낮은 문항이라고 할 수 있다. 이들은 공통적으로 해당 검사지 내에서는 변별도가 상대적으로 낮은 문항인 것임을 알 수 있다.

3. 논의

이러한 분석의 결과, 9개 검사지에서 신뢰도를 낮추거나 국어능력 측정 기능이 낮은 문항 16개를 추출해냈다.

검사지 신뢰도를 낮추는 6개 문항(<표 3> 참조) 중에서 2010년 고등 24번 문항이하 ‘10-고-24’로 표기함과 ‘09-초-서답4’는 신뢰도를 떨어뜨리는 문항이지만 국어능력 측정 기능은 있는 문항으로 구분되었다. 이를 제외한 나머지 문항들은 검사지의 신뢰도를 낮추는 문항이면서 국어능력 측정 기능이 낮은 문항들이었다. 이 문항들은 모두 해당 검사지 내에서 변별도가 상대적으로 낮은 문항이라는 공통점이 있었다. ‘09-초-서답4’는 변별도 0.38로 변별도가 있는 문항이지만 해당 검사지의 서답형 문항들 중에서는 상대적으로 낮은 변별도를 보인다. ‘09-고-20’과 ‘09-고-27’도 변별도 0.27로 변별도가 낮지 않지만 해당 검사지 내에서는 변별도가 가장 낮은 문항이었다.

그러면 검사지의 신뢰도와 타당도를 낮추는 문항들에는 어떤 공통된 특징이 있을까? 이들 문항에서 한 가지 공통점을 추출할 수 있었다. 바로 학습자들이 정답을 확정하는 데에 장애가 되는 요소가 포함된 문항이라는 것이다.

그리고 이러한 장애를 주는 원인을 크게 네 가지 유형으로 구분할 수 있었다. 첫째, 문항을 풀이하는 데 필요한 정보가 불충분하여 학습자가 진위 여부를 확정하기 어려운 경우였다. 둘째, 문항 내에 불명확한 진술이나 혼동을 일으키는 내용이 포함되어서 학습자가 진위 여부를 확정하기 어려운 경우였다. 셋째, 학습자의 수준을 상회하는 어려운 요소가 포함되어서 학습자가 가지고 있는 지식으로는 정답을 확정하기 어려운 경우였다. 그리고 이들 유형의 요소가 복합적으로 포함된 경우도 있었다. 넷째, 문항이 언어 양식(음성 언어) 처리 방식에 적합하지 않은 내용으로 구성되어서 학습자가 정답을 확정하는 데 무리가 따른 경우였다. 이상의 유형에 대해 문항을 중심으로 살펴보자.¹⁰⁾

1) 불충분한 정보가 제공된 문항

첫째, 문항을 풀이하는 데 충분하지 못한 정보가 제공되어 학습자가 진위 여부를 확정하기 어려운 문항으로, ‘08-초-8’를 예로 들 수 있다. 이 문항 <자료>의 그림만을 보면 이것이 어떤 상황인지 짐작하기 어렵다. 다만, 답지의 진술을 바탕으로 그림에 제시된 상황을 역으로 짐작해 볼 때 정전이 된 상황임을 유추해 볼 수 있다. 답지①, ②, ③은 각각 1% 미만대의 답지 반응률을 보이는 것처럼(아래 표 참조), 정답이 되기에는 불충분한 진술들이다. 학습자들은 답지 ④와 ⑤ 중에서 정답을 가려야 하는데, 세 번째 컷의 그림으로는 이 가족이 서로 대화를 하는 것인지, 아빠가 이야기를 하는 것인지, 서로 쳐다보고 웃는 것인지 알기 어렵다. 많은 학습자들이 ⑤번을 선택하기는 했지만 ④번을 선택한 경우도 있는데, 이 선택이 충분한 정보를 바탕으로 한 것이 아닐 가능성이 크다. 즉, 국어능력이 충분히 있는 학습자들도 마지막 그림을 ‘대화를 하게 되었다’로 해석할 충분한 정보가 없다고 판단하고, 그림에서 명백히 유추할 수 있는 상황(마지막 그림에서 가족은 서로를 쳐다보고 있고, 정전이 되었으므로 텔레비전은 꺼져 있

10) 지면 관계상 해당 문항은 [부록 3]을 참조하기 바람.

을 것을 짐작할 수 있음인 답지④를 선택했을 가능성이 있다. 다시 말해, 이 문제는 국어능력에 따라서 반응 결과가 나누어진 것이 아니라 주어진 <자료>에서 제공한 정보 자체가 불충분했기 때문에 판단이 엇갈렸던 것으로 볼 수 있다.

문항	정답률(%)	변별도	답지 반응 분포				
			①	②	③	④	⑤
08-초-8	90.95	0.23	0.45	0.79	0.83	6.83	90.95

2) 불명확한 진술이나 혼동을 일으키는 내용이 포함된 문항

둘째, 불명확한 진술이나 혼동을 일으키는 내용이 포함되어서 학생들이 진위 여부를 확정하기 어려운 문항이 있었는데, ‘08-고-9’, ‘09-중-17’, ‘08-고-4’를 들 수 있다.

‘08-고-9’는 검사지의 신뢰도를 낮추지는 않지만 국어능력 측정 기능이 낮은 문항이다. 우선 이 문항의 <자료>는 상당히 전문적인 내용으로 고1 수준을 상회하는 것이고, 답지의 진술이 분명하지 않아서 혼동을 일으키는 내용들이 있다. 답지①과 ②는 일부 어구를 수정하거나 삭제하는 것에 대한 판단으로, 많은 학습자가 적절성 여부를 판단할 수 있었던 것으로 보인다. 그리고 정답지③도 지문의 ‘건강 관련 전문학술지에 게재’한다는 조건에 주목했다면 틀린 진술이라는 것을 파악할 수 있었을 것이다. 그러나 답지 반응에서 보듯이 학생들이 이 부분을 많이 놓치고 있다. 그리고 문제는 답지④인데 (나)의 내용을 고려할 때 어렵지 않게 판단할 수 있었을 것이다. 그러나 남민우 외(2009)의 지적처럼 4개의 답지에서 ‘-해야겠어’의 진술로 되어 있는데, 답지④만이 ‘-도 -좋겠어’라고 하여 “다른 답지들과 달리 필수적인 수정 내용을 언급하는 것처럼 읽히지(180쪽)” 않는다. 즉 수정 여부를 확정하지 않는 진술로 읽힘으로써 오답으로 판정해 버렸을 가능성이 크다는 것이다. 이는 답지 진술 상에서 학습자의 오해를 불러일으킬 수 있는 요소가 들어 있고, 그로 인해 학습자가 정답을 확정

하는 데 장애를 일으켰던 것으로 보인다. 그리고 답지 ⑤는 마치 결론에 ‘예방의 중요성’이라는 내용을 새롭게 추가하는 설명인 것으로 읽힐 수 있는데, 10명 중 거의 2명에 가까운 학생들이 ⑤를 선택한 것을 보아도 오해의 여지가 있었던 것으로 판단된다.

‘09-중-17’도 쓰기 문항으로, 답지의 진술에서 학생들이 충분히 혼동할 여지가 있는 내용이 포함되어 있다. 답지③은 (가)를 보완한다고 할 때에 적절한 내용임에는 틀림이 없다. 그러나 답지④는 만약 (가)를 새로운 내용으로 교체해서 보완한다면, 틀린 내용이 될 수 없는 진술이다. 10명 중 3명에 가까운 학생들이 ④를 선택한 것으로 보아도, 이 진술은 학생들에게 상당한 혼동을 초래한 것으로 판단된다.

문항	정답률(%)	변별도	답지 반응 분포				
			①	②	③	④	⑤
08-고-9	20.18	0.18	7.69	14.91	20.18	39.97	17.02
09-중-17	39.88	0.24	10.47	6.47	39.88	27.96	15.06

한편, ‘08-고-4’는 검사지의 신뢰도를 낮추고 국어능력 측정 기능도 낮은 문항으로, 강연을 듣고 신뢰성을 높이는 방안을 선택하는 문항이다. 강연은 신뢰성이 크게 떨어지는 내용은 아닌데, 학습자는 강연을 한 번 듣고 신뢰성을 판단하거나 답지를 통해 역추적해야 한다. 듣기 자료의 끝부분에서 ‘숲을 가까이 하자’는 제안을 했으므로 학습자들은 답지②와 같이 숲 생태계에 대한 구체적인 정보를 제시함으로써 보다 신뢰성있는 강연이 될 수 있을 것이라고 판단했을 수 있다. 또한 강연은 생태맹의 의미를 상당히 추상적으로 설명하고 있는데, 이것에 대한 설명이 더 필요하다고 판단한 학습자들은 답지④를 선택했을 개연성이 크다. 또 학습자들은 자신의 판단을 들은 내용에 기대야 하니 더 어려움이 있었을 것이다. 이 문항은 다소 혼동을 일으키는 답지 구성으로 인해 학습자들에게 곤란을 준 것으로 판단된다.

문항	정답률(%)	변별도	답지 반응 분포				
			①	②	③	④	⑤
08-고-4	38.19	0.18	2.87	17.95	38.19	29.71	11.01

3) 학습자의 수준을 상회하는 어려운 요소가 포함된 문항

셋째, 해당 학년 학습자의 수준에서는 해결하기 어려운 내용이 들어 있어서 학습자들이 정답을 확정하기 어려운 문항들이 있었다. 여기에는 문장 구조에 관한 문항인 ‘08-고-20’과 ‘09-고-20’, 단어 분류 문항인 ‘09-초-서답4’, 어법 문항인 ‘09-중-8’, 부사어의 호응에 관한 문항인 ‘10-초-20’, 사회 방언에 관한 문항인 ‘10-고-28’, 그리고 어려운 표현이 포함된 읽기 문항인 ‘09-초-19’가 해당된다. 여섯 개가 문법 문항이고, 읽기 문항이 한 문항 포함되었다.

먼저 ‘08-고-20’과 ‘09-고-20’을 살펴보자. 두 문항은 모두 국어능력 측정 기능이 낮은 문항이고, ‘08-고-20’은 해당 검사지의 신뢰도도 떨어뜨리는 문항인 것으로 나타났다. 두 문항은 ‘문장 구조를 파악할 수 있다’는 성취기준을 구현한 것인데, 반응 결과는 다음과 같다.

문항	변별도	답지 반응 분포					전체 정답률	우수학력 정답률	보통학력 정답률	기초학력 정답률	기초미달 정답률
		①	②	③	④	⑤					
08-20	0.14	19.08	12.01	26.63	22.76	19.11	26.63	34.89	26.41	22.18	15.90
09-20	0.27	7.72	25.64	23.17	23.77	19.35	25.64	41.19	19.41	16.75	14.81

‘08-고-20’에서 주어진 문장은, 주어(콩고는)를 수식하는 관형절(아프리카 중부에 위치한)이 안겨 있고 서술부에는 ‘나라’를 수식하는 관형절(골탄이 많이 생산되는)이 안겨 있는, 두 개의 관형절을 안고 있는 복잡한 구조의 문장이다. 학생들은 이 문항을 해결하기 위해 많은 문법 지식을 동원해야 하고, ‘적절하지 않은’ 것을 선택해야 하기 때문에 모든 진술의 진위를 확정해야 한다. 우선 답지①에서는 주성분의 개념을 알고, 문장에 적용할 수 있어야 한다. 둘째, 답지③에서는 관형절과 부사절이 무엇인지를 알고, 문

장에 적용할 수 있어야 한다. 셋째, 답지②, ④, ⑤에서는 문장 성분 각각의 기능을 파악하고 문장 구조를 분석할 수 있어야 한다. 이러한 지식들을 한꺼번에 요구하는 이 문항은, 20%대의 정답률로 볼 때 학습자에게 상당히 어려웠던 것을 알 수 있다. 그리고 답지별 반응 분포를 보아도 모든 답지의 진위를 파악하는 데에 상당히 곤란을 겪었던 것으로 나타났다. 또한 이러한 문법 지식은 우수학력 학생들도 30-40% 정도만이 알고 있는 것으로, 보통학력, 기초학력, 기초미달 집단이 모두 잘 알지 못하는 것으로 나타났다.

‘09-고-20’도 ‘08-고-20’과 성취기준과 문항 형태가 동일하고, 같은 문제를 안고 있다. 이 두 문항에 대해서는 이은희(2011: 7)에서도 “평가로서의 기능을 거의 하지 못”하는 문항이라고 지적한 바 있는데, 두 문항은 해당 학년 학습자의 국어능력을 변별해 주지 못한다. 두 문항이 국어능력 측정 기능이 낮다는 것에서도 알 수 있듯이, 이 문항이 측정하는 문법 지식은 국어능력을 가늠하는 지표가 되지 못한다. 여기에는 하나의 문항이 여러 가지 문법 지식을 복잡하게 묻고, 분석 대상 문장이 학습자의 수준을 상회하는 어려운 것이라는 것에서 원인을 찾을 수 있다. 2010년 고등검사지에는 동일한 성취기준에 대한 문항이 있는데, 다음과 같이 비교적 쉬운 문장을 제시하고 답지 구성도 보다 간략하게 제시했는데, 그 반응 결과가 양호했다.

㉠ 국어는 우리 겨레가 살아온 모습을 반영한다.

10-고-8. ㉠에 대한 다음 설명으로 적절한 것은?

- ① 서술어가 두 번 나타난다.
- ② 문장 전체의 주어는 ‘우리 겨레가’이다.
- ③ 안은 문장과 안긴 문장의 주어는 같다.
- ④ 안긴 문장의 주성분은 ‘주어, 목적어, 서술어’이다.
- ⑤ 관형절로 안긴 문장과 인용절로 안긴 문장이 있다.

‘10-고-8’은 본 연구의 분석 결과, 변별도(0.30)가 있는 문항이면서 해당 검사지의 신뢰도에도 기여하고 국어능력 측정 기능도 있는 문항인 것으로

나타났다. 이렇게 ‘08-고-20’, ‘09-고-20’와 ‘10-고-8’에 대한 분석 결과를 대비해 보면 이 성취기준문장 구조 분석을 구현한 문항이, 학습자의 능력으로 분석할 만한 수준의 문장을 제시하고 답지의 내용도 비교적 손쉽게 구성된다면 얼마든지 변별력을 가지면서 문항 본래의 기능을 할 수 있다는 것을 보여준다. 반면에 학습자의 수준을 상회하는, 지나치게 복잡하거나 어렵게 구성된 문항은 오히려 문항으로서의 제 기능을 하지 못하는 것을 확인할 수 있었다.

한편, ‘09-초-서답4’는 국어능력을 측정하는 기능은 있으나 서답형 문항 중에서는 유일하게 검사지의 신뢰도를 떨어뜨리는 문항으로 판별되었다. 이 문항은 해당 검사지 내 서답형 문항 중에서는 변별도가 가장 낮은 것으로, ‘외래어’의 개념을 알고 제시된 단어들 중에서 고유어를 고를 것을 요구한다. 학생들은 ‘도시, 시골, 거실’ 중에서 고유어를 하나 선택해야 하는데, 이 단어들 중에서 고유어와 한자어를 구별하는 것이 까다로웠던 것으로 나타났다.

문항	정답률(%)	변별도	부분 점수별 비율		
			0	1	2
09-초-서답4	58.94	0.38	28.89	24.34	46.77

‘09-초-19’는 국어능력 측정 기능이 낮은 문항으로 나타났다. 답지②의 ‘먹을 감다’, 답지③의 ‘소를 먹이다’의 표현이 학생들의 반응을 엇갈리게 한 요소였던 것으로 판단된다. 즉 두 가지 표현의 의미를 정확히 모르는 상태이다 보니 답지 진술의 진위를 확정하기 어려웠던 것으로 보인다.

문항	정답률(%)	변별도	답지 반응 분포				
			①	②	③	④	⑤
09-초-19	34.35	0.30	34.35	26.27	17.10	10.27	10.73

‘09-중-8’도 국어능력 측정 기능이 낮은 문항으로 어법에 대해 묻고 있다. 학생들이 답지②의 ‘남어’와 ‘너머’, 답지③의 ‘-던지’와 ‘-든지’를

정확하게 구분하는 데 어려움이 있다는 것을 알 수 있다. 이는 국어능력이 좋은 중3 학생들 중에서도 이 표현들에 대해 정확히 모르는 이가 많은 것을 시사한다.

문항	정답률(%)	변별도	답지 반응 분포				
			①	②	③	④	⑤
09-중-8	47.11	0.26	1.81	16.40	26.59	7.88	47.11

‘10-초-20’은 국어능력 측정 기능도 낮고 검사지의 신뢰도도 떨어뜨리는 문항으로, 부사가 적절하게 쓰이지 않은 경우를 찾는 것이다. 학습자들은 답지①을 정답만큼이나 많이 선택하였다. 초6 학생들이 ‘-크지라도’와 호응을 이루는 ‘비록’에 대한 이해가 충분치 않기 때문에 반응이 엇갈렸던 것으로 판단된다.

문항	정답률(%)	변별도	답지 반응 분포				
			①	②	③	④	⑤
10-초-20	47.0	0.13	39.1	6.8	47.0	5.2	1.5

‘10-고-28’은 검사지의 신뢰도도 떨어뜨리고 국어능력 측정 기능도 낮은 문항으로, 사회 방언에 따라 달라지는 어휘의 의미에 대해 묻고 있다. 학생들의 반응이 답지③, ④, ⑤로 엇갈리는데, 이는 학생들이 물건 구매 상황에서 판매원 발화의 이면적 의미, 군대 언어의 의미와 특징에 대해 잘 알지 못한다는 것을 보여주는 것이다.

문항	정답률(%)	변별도	답지 반응 분포				
			①	②	③	④	⑤
10-고-28	29.0	0.14	10.5	29.0	18.5	22.0	19.5

마지막으로 ‘10-중-28’은 국어능력 측정 기능이 낮은 문항이다. 이 문항은 부정문에서만 쓰이는 부사어를 찾는 문항인데, 답지에 제시된 부사

가 모두 부정어와 호응하고 있다. 즉 제시된 부사는 모두 부정문에서 쓰이는 것으로, 학생들은 이 중에서 긍정문에서도 쓰이는 것을 찾으면 된다. 그런데 학생들은 답지⑤를 정답보다 더 많이 선택하였다. 이것은 학생들이 ‘않다’와 ‘없다’에 비해 ‘아니다’를 일견 부정어로 인식하지 못했기 때문인 것으로 추정해 볼 수 있다. 또한 ‘아직’이 ‘우리에게는 아직 희망이 남아 있다’와 같이 긍정문에서도 쓰인다는 것을 놓치고 있기 때문인 것으로 판단된다.

문항	정답률(%)	변별도	답지 반응 분포				
			①	②	③	④	⑤
10-중-28	21.3	0.13	21.3	9.5	12.8	10.9	45.1

지금까지 살펴 본 문항에서는 학습자가 잘 몰라서 정오(正誤)를 판단하기 어려운 점이 있음을 살펴보았다. 이 중에서 국어능력 측정 기능이 낮은 문항(‘09-초서답4’를 제외한 나머지 문항들)은, 문항 속에서 학습자들이 곤란을 느낀 내용(지식)이 해당 학년 학습자의 국어능력을 가늠하는 지표가 되지 못한다는 것을 시사한다. 이는 국어능력이 좋은 학습자도 해당 지식의 습득이 충분치 않음을 의미한다. 이러한 지식에 대해서는 교수·학습 측면에서는 지도를 강화해야 하고, 평가 측면에서는, ‘문장 구조 분석’ 문항의 예가 시사하듯, 학생들이 처리할 수 있는 수준의 자료를 바탕으로 보다 쉽게 문항을 구성해야 할 것이다.

지금까지 학습자가 정답을 확정하는 데 장애를 주는 원인으로 세 가지 유형을 살펴보았는데, 이러한 원인이 복합적으로 포함된 문항들도 있었다. 여기에는 ‘08-고-25’와 ‘09-고-27’을 예로 들 수 있는데, 모두 문학 문항이다.

먼저 ‘08-고-25’의 경우를 보자. 이 문항은 국어능력 측정 기능이 낮은 것으로, <오발탄> 전체를 읽은 학생이나 그렇지 않은 학생 모두에게 어려웠을 것으로 짐작된다. 답지①은 <자료>를 통해서 한국 전쟁 때문에

‘철호’ 가족이 궁핍하게 산다는 것을 유추할 수 있다. 답지②의 경우도 <자료>에서 월남 난민에 대해 언급한 점, 지문에서 ‘가자 가자’를 외치는 점 등을 통해서 철호 어머니가 월남 난민이라는 점을 짐작해 볼 수 있다. 정답인 ③의 해방촌은, 지문의 내용을 바탕으로 궁핍한 생활을 하는 사람들이 사는 동네라는 것을 짐작할 수 있다. 짐작컨대 전쟁 직후에 그것도 70년대의 산물인 도시 개발에 의해 형성된 동네라고 추론하기 어렵다. 그러나 답지④는 주어진 지문의 내용이나 <자료>의 정보로는 확인하기 어렵고, 작품 전체의 내용을 통해서도 가족이 남북통일을 간절히 소망한다는 점을 명백히 확정하기 어렵기 때문에, ‘적절히 않은 것’으로 상당수의 학생들이 선택하였다. 답지⑤의 경우도 시대에 대한 배경지식이 부족한 학습자는 1950년대에 중학교를 진학하는 것이 유복한 것인지 아닌지를 판단하기 어려울 수 있다. 즉 이 문항은 문항 풀이에 필요한 정보를 충분히 제공하지도 않고, 문항이 요구하는 시대 배경에 대한 지식이 학습자의 수준을 상회하기 때문에 학습자가 정답을 확정하기 곤란했던 것으로 판단된다.

문항	정답률(%)	변별도	답지 반응 분포				
			①	②	③	④	⑤
08-고-25	38.38	0.23	6.47	9.07	38.38	26.93	18.40

‘09-고-27’도 국어능력 측정 기능이 낮은 문항인데, 시대상을 추론하는 문항이다. 작품과 1930년대의 시대상에 대한 배경지식이 부족한 학생들은 지문에서 충분한 정보를 제공받지 못하기 때문에 곤란을 겪었던 것으로 판단된다. 정답인 ③을 제외하고, 다른 답지에 대한 반응률이 대동소이한 것을 볼 수 있다. 답지①은 인쇄소 과장의 말을 통해서 비교적 쉽게 유추할 수 있는 것으로 보인다. 답지②의 ‘열차’는 ‘경성역’을 통해 유추할 수 있는데, ‘기차’가 익숙한 학생들에게는 ‘열차’가 다소 생경했을 수 있다. 답지④는 지식인인 P가 자신의 아들을 인쇄소 과장에게 부탁하는 것에 주목한 학생이라면, 지식인의 처지가 좀더 비참했다고 오판할 수 있다. 또한

‘지식인과 노동자 모두 생활이 힘들었었군’과 같이 지문에서 보다 쉽게 유추할 수 있게 답지가 진술되었더라면 오판의 여지를 줄일 수 있지 않았을까 한다. 그리고 월급도 필요 없다는데 어린 아이를 흔쾌히 받아주지 않는 인쇄소 과장의 간간한 태도에 주목한 학생들이라면 답지⑤도 정답으로 오판할 수 있다. 이 문항은 학습자가 진위를 판정하는 데에 지문이 충분한 정보를 제공하지 못하고, 답지에 다소 모호한 진술도 포함되어 있어서 학습자의 반응이 엇갈렸던 것으로 판단된다.

문항	정답률(%)	변별도	답지 반응 분포				
			①	②	③	④	⑤
09-고-27	35.05	0.27	13.44	18.97	35.05	16.92	15.25

‘10-고-24’는 검사지의 신뢰도를 낮추지만 국어능력 측정 기능은 있는 문항으로, 두 시조의 특성을 비교하는 문항이다. 이 문항은 다소 어려운 내용이 포함되고, 그 진술도 간명하지 않아서 학생들에게 혼동을 준 것으로 판단된다. 상당수의 학습자들은 답지④를 선택했는데, 문학 작품은 드러내고자 하는 의미를 형상화 과정을 통해 표현하는 것임에 주목한다면, 두 작품은 모두 삶의 태도를 간접적으로 표현하는 것이라고 할 수 있다. 그리고 인물의 모습을 직접적으로 묘사하고 있는 작품 (나)에 주목했다면 ‘인물의 삶의 태도를 간접적으로 드러내고 있다’를 틀렸다고 오판할 수도 있었을 것이다. 인물의 모습을 묘사하는 것, 시적 화자의 삶의 태도를 직접적으로 드러내는 것, 인물의 삶의 태도를 인물 행위 서술을 통해 간접적으로 드러내는 것 등이 의미하는 바를 정확하게 이해하는 것이 학습자에게는 까다로웠던 것으로 나타났고, 그러한 내용이 진술된 방식도 간명하지 못해서 혼동을 준 측면이 있는 문항으로 판단된다.

문항	정답률(%)	변별도	답지 반응 분포				
			①	②	③	④	⑤
10-고-24	39.8	0.23	9.5	10.1	39.8	27.8	12.4

4) 언어 양식(음성 언어)의 특성에 적합지 않은 내용으로 구성된 문항

넷째, 언어 양식의 특성에 적합지 않은 내용이 포함되어 있어서 학습자가 정답을 확정하는 데 무리가 따른 문항이 있었다. 여기에는 듣기 문항인 ‘08-중-3’이 해당되는데, 검사지의 신뢰도를 낮추고 국어능력 측정 기능도 낮은 것으로 나타났다. 이 문항은 세 출연자가 발언한 내용을 모두 종합할 수 있는 의미로 정리해야 하는 문항이다. 학습자들은 세 출연자의 발언을 듣고, 발언의 내용을 기억하면서 세 개의 발언에서 공통된 의미를 추출해야 한다. 문자 언어를 대상으로 세 발언에서 공통된 의미를 추출하는 것도 쉽지 않은데, 순간적으로 흘러가는 발언(음성 언어)에서 공통적인 의미를 추출한다는 것은 듣기 문항의 요구로는 상당히 무리한 것으로 판단된다. 이런 이유로 학생들이 답지 중에서 가장 그럴듯한 진술인 정답지③을 선택해 버렸을 가능성을 배제하기 어렵다.

문항	정답률(%)	변별도	답지 반응 분포				
			①	②	③	④	⑤
08-중-3	69.15	0.18	5.00	6.85	69.15	5.87	12.96

이상으로 본 연구에서 검토한 16개 문항 중에는 다행히 이후에 개발된 검사지에서 개선된 경우도 있다. ‘08-초-8’(그림을 제공하고, 그것을 언어로 표현하게 한 문항)은 2009년 검사지부터는 등장하지 않고, ‘09-초-서답4’(외래어, 한자어, 고유어를 구분하는 초등 문항)은 2010년부터는 질문의 의도가 보다 잘 드러나는 형태의 문항으로 바뀌었다. 그리고 앞에서 언급한 바와 같이, 문장 구조 분석에 대한 고등학교의 문항은 2010년의 검사지부터는 분석 대상 문장의 수준을 낮추고 답지 구성을 상대적으로 수월하게 구성하여 문항이 제 기능을 하도록 개선된 바 있다. ‘맺음말’에서는 이러한 논의를 토대로 향후 문항 개발 시에 개선점을 제언하면서 마무리하도록 한다.

V. 맺음말

본 연구는 전수 대상 시행 이후의 9개 국어과 국가수준 학업성취도 평가 검사지에 대해서 신뢰도와 타당도를 검토하였다. 그 결과 9개 검사지가 모두 높은 신뢰도와 타당도를 확보하고 있다는 것을 확인하였다. 다만, 극히 일부의 문항이 다소 미흡한 것을 알 수 있었다. 이들 문항은 학습자가 정답을 확정하는 데에 장애를 주는 요소를 지닌 것으로, 향후의 검사지 구성 시에는 이러한 문항이 없도록 해야 할 것이다.

이를 위해서 첫째, 문항에서는 문항 풀이에 필요한 정보를 충분히 제공할 필요가 있다. 정답을 확정하는 것은 일차적으로는 학습자의 국어능력에 달려있는 것이다. 그러나 능력과 상관없이 문항이 풀이에 필요한 정보를 불충분하게 제공해서 학습자들이 곤란을 겪게 해서는 안 된다. 충분한 정보를 제공해서 학습자들이 문항 풀이를 할 수 있는 여건을 마련해 주어야 한다.

둘째, 학습자가 진위를 확정할 수 없는 다소 모호한 진술이나 혼동을 주는(misleading) 내용이 포함되게 해서는 안 된다. 분석 결과에서 살펴본 듯이 이러한 문항은 문항 본래의 기능을 하지 못한다. 출제 과정에서 난도를 높이고자 진술을 다소 모호하게 하거나 어렵게 만드는 경우가 있는데, 이는 문항이 제 기능을 하는 데 장애가 될 뿐이다. 때문에 문항은 반드시 학습자가 질문의 의도와 답지의 의미를 명쾌하게 이해할 수 있는 언어로 진술되어야 한다.

셋째, 문항은 학습자의 수준에 적합한 내용으로 구성되어야 한다. 매우 어려운 자료를 활용하는 것이 아니라 학습자의 국어능력을 의미있게 변별해 낼 수 있는 언어 자료를 바탕으로 묻는 문항을 구성해야 할 것이다. 또한 하나의 문항에서 여러 가지 지식을 물어서 문항을 어렵게 하는 것을 지양해야 한다. 분석 결과에서 보았듯이, 해당 학년 학습자들의 수준을 상회하는 어려운 요소를 포함하고 있는 문항은, 문항의 본래 기능을 다하지 못하고 검사지의 신뢰도와 타당도 확보에도 지장을 준다.

한편으로 분석 결과를 통해서 우리나라의 학생들이 어떤 부분에 대한 지식이 부족한지를 알 수 있었는데, 이러한 부분에 대해서는 향후에 해당 교육 내용에 대한 지도가 강화되어야 할 것이다. 더불어 해당 교육 내용의 학년 위계 설정에도 이 같은 결과가 송환(feedback)될 필요가 있다.

넷째, 분석 결과에서 듣기 문항이 음성 언어 처리 방식으로 해결하기 곤란한 경우에는 제 기능을 하지 못한다는 것을 알 수 있었다. 듣기 문항은 들은 음성 정보를 바탕으로 해당 학년 학습자가 순간적으로 정보 확인, 추론, 비판이 가능한 범위 내에서 구성되어야 한다. 수능에서 듣기 문항이 없어진 상황을 감안한다면, 학업성취도 평가에서 비용이 많이 드는 듣기 문항이 존속할 수 있을지 예견하기 어렵다. 이러한 때에 학업성취도 평가의 듣기 문항이 본래적 기능을 구현하는 것으로 구성하는 일은 시급한 일이다.

본 연구의 결과에서 보았듯이, 학업성취도 평가 검사지는 상당히 높은 신뢰도와 타당도를 확보한 검사이다. 그러나 향후에 위에서 언급한 네 가지 지점을 보완한다면 보다 더 신뢰롭고 타당한 검사지가 될 수 있을 것이다.*

* 본 논문은 2012. 6. 30. 투고되었으며, 2012. 7. 10. 심사가 시작되어 2012. 7. 31. 심사가 종료되었음.

▣ 참고문헌

- 김도남, 김영란, 김미경, 권택일(2011), “2010년 국가수준 학업성취도 평가 결과 분석—국어—”, 연구보고 RRE 2011-3-2, 한국교육과정평가원.
- 김영란, 조병영(2011), “CURRV 모형에 기초한 국어과 국가수준 학업성취도 평가 고찰”, 연구보고 RRE 2011-14, 한국교육과정평가원.
- 김창원(2011), “국가수준 학업성취도 평가의 본질과 제 문제”, 『국어교육』 134, 1~33.
- 김혜정(2011), “국가수준 학업성취도 평가 개선을 위한 ‘읽기’ 평가의 비판적 검토”, 『국어교육』 134, 61~86.
- 남민우(2011), “초등학생들의 시적 능력과 시교육 평가의 개선 방안—국어과 학업성취도 평가 결과를 바탕으로—”, 『국어교육학연구』 제41집, 379-401.
- 남민우 · 김도남 · 최은정 · 정은영(2009), “2008년 국가수준 학업성취도 평가 연구—국어—”, 연구보고 RRE 2009-9-1, 한국교육과정평가원.
- 남민우 · 남가영(2011), “국가수준 국어과 학업성취도 평가의 시행 체제 개선 방안”, 『국어교육』 134, 113~130.
- 남민우 · 남가영 · 이창훈 · 정은영 · 최은정(2010), “2009년 국가수준 학업성취도 평가 연구—국어—”, 연구보고 RRE 2010-6-2, 한국교육과정평가원.
- 박영민(2011), “국가수준 학업성취도 평가 개선을 위한 ‘쓰기’ 평가 이론의 정립”, 『국어교육』 134, 87~112.
- 성태제(2002), 『타당도와 신뢰도』, 서울 : 양서원.
- 이순묵(1994), “요인분석의 관행과 문제점”, 『한국심리학회지 : 산업 및 조직』 7(1), 1~27.
- 이은희(2011), “문법 평가의 현황과 과제—국가 수준 학업 성취도 평가를 중심으로—”, 『문법교육』 제15집, 1~29.
- 임칠성(2011), “국가 수준 학업성취도 평가 개선을 위한 ‘듣기’ 평가 이론의 정립”, 『국어교육』 134, 35~60.
- 정구향 · 김경희 · 김재철 · 반재천 · 민경석 · 오희숙 · 권재기(2004), “2003년 국가수준 학업성취도 평가 연구—총론—”, 연구보고 RRE 2004-1-1, 한국교육과정평가원.
- 정은영 · 남민우 · 김도남 · 김혜숙 · 박가나 · 이봉주 · 권점례 · 최원호 · 이인호 · 조보경 · 송민영 · 최인봉 · 김희경 · 김소영(2010), 국가수준 학업성취도 평가의 교과별 평가 틀 개발 연구, 연구보고 CRE 2010-7, 한국교육과정평가원.
- 조재운(2011), “국가수준 국어과 학업성취도 평가의 적합성 검토 연구 : 국어과 평가

- 의 적합성 검토 VRUT 모형을 사용하여”, 『한국초등국어교육』 제46집, 349~377.
- 주세형(2009), “국가수준 학업 성취도 평가에서의 소위 ‘텍스트 중심 원리’에 대한 비판—2005-2008년 문법 영역을 중심으로—”, 『국어교육학연구』 제35집, 481~506.
- 주세형(2011), “국어과 문법 평가 이론의 발전 방향”, 『국어교육』 135, 39~66.
- 천경록(2009), “국어과 평가의 정합성 분석의 기초”, 『국어교육』 129, 113~139.
- 천경록(2010), “초6 국어과 성취도 평가의 정합성 판정”, 『한국초등국어교육』 제43집, 35~60.
- 천경록(2011), “국가수준 국어과 평가의 정합성 분석—중3, 고1 학업성취도 평가를 중심으로—”, 『국어교육』 136, 297~324.
- 최지현(2011), “국가 수준 학업 성취도 평가 개선을 위한 ‘문학’ 평가 이론의 정립”, 『국어교육』 135, 1~35.
- Afflerbach, P. (2007), *Understanding and using reading assessment : K-12*. 조병영 · 김소현 · 조재윤 · 서수현 · 김종윤 옮김(2010), 『독서 평가의 이해와 사용』, 서울 : 한국문화사.

<초록>

국어과 국가수준 학업성취도 평가의 문항 개선 방안 연구

—2008~2010년 검사지의 타당도 검토를 중심으로—

김영란 · 김인숙

본 연구는 전수 대상 시행 이후의 국어과 국가수준 학업성취도 평가 검사지의 신뢰도와 타당도를 검토하고, 그 결과에 따라 검사지의 신뢰도와 타당도를 떨어뜨리는 문항의 특징을 분석하여 문항 개선점을 제시하는 것을 목적으로 한다. 이를 위하여 2008-2010년의 초, 중, 고 9개 국어과 검사 결과를 이용하여 신뢰도 계수 산출, 문항 제거 시 신뢰도, 문항 변별도, 요인분석을 통한 요인적재량을 검토하였다. 신뢰도 분석 결과, 9개 검사지의 신뢰도가 모두 높은 것으로 나타났다. 그리고 6개 검사지 중 8개 문항이 검사지의 신뢰도를 다소 떨어뜨리는 문항인 것으로 나타났다. 그리고 구인 타당도 분석 결과, 14개의 문항이 국어능력을 측정하는 것과 는 거리가 있는 것으로 나타났다.

이렇게 검사지의 신뢰도를 떨어뜨리는 문항과 국어능력 측정 기능이 낮은 문항은, 공통적으로 해당 검사지의 문항들 중에서 변별도가 상대적으로 낮은 문항이라는 점을 발견하였다. 그리고 이러한 문항들은 모두 학습자가 정답을 확정하는데 장애가 되는 요소가 있다는 특징을 추출하였다. 또한 정답 확정에 장애를 주는 원인을 네 가지 유형으로 분류할 수 있었다. 첫째, 문항에서 불충분한 정보가 제공되어서 학생들이 진위 여부를 확정하기 어려운 경우였다. 둘째, 불명확한 진술이나 혼동을 일으키는 내용이 포함되어서 학생들이 진위 여부를 확정하기 어려운 경우였다. 셋째, 학습자 수준을 상회하는 어려운 요소 때문에 학생들이 정답을 확정하기 어려운 경우였다. 넷째, 언어 양식(음성 언어)의 처리 방식에 적합지 않은 내용으로 구성되어서, 학생들이 정답을 확정하는 데 무리가 따른 문항들이었다.

【핵심어】 국어과 국가수준 학업성취도 평가, 타당도, 신뢰도, 구인타당도, 요인분석

<Abstract>

A Study on the Improvement of Test Items in the Korean NAEA

—Using Data from the 2008~2010 NAEA Korean Tests—

Kim, Young-ran · Kim, In-suk

This study examines the reliability and the validity of the Korean National Assessment of Educational Achievement (NAEA) tests since they began to be administered nationwide. In addition, it suggests some improvements to item development by analyzing those characteristics of the items that lower the reliability and the validity of the test. Using the data from 2008~2010 elementary, middle, and high school Korean NAEA tests, the reliability coefficient (Cronbach's alpha), Cronbach's alpha if item deleted, item discrimination, and factor loadings by factor analysis were examined. The results showed that the internal consistency of all the Korean NAEA tests was high : eight items from six of the nine tests lower the Cronbach's alpha. Also, fourteen items from the nine tests had low factor loadings; that is, the correlation coefficients between the items and the factor (Korean ability). These items had both relatively low item discrimination in common and factors which distracted students from answering correctly. These items can be categorized into four types : first, they have limited information ; second, they have imprecise statements and ambiguous content ; third, they have inappropriate difficulty levels for the students to whom they are administered ; and fourth, they have improper factors for the language style ; any and all of these may mislead students into answering incorrectly.

【Key words】 Korean National Assessment of Educational Achievement (NAEA) test, Reliability, Validity, Construct Validity, Factor Analysis

【부록 1】 문항 제거 시 신뢰도 및 문항별 변별도

문항	2008						2009						2010					
	초		중		고		초		중		고		초		중		고	
	신뢰	변별	신뢰	변별	신뢰	변별	신뢰	변별	신뢰	변별	신뢰	변별	신뢰	변별	신뢰	변별	신뢰	변별
1	.860	.42	.912	.27	.880	.33	.819	.45	.913	.37	.904	.44	.783	.38	.870	.32	.902	.36
2	.862	.33	.910	.47	.880	.34	.820	.40	.913	.40	.904	.39	.790	.19	.870	.33	.901	.41
3	.860	.43	.913	.18	.880	.34	.822	.35	.914	.30	.904	.40	.789	.22	.869	.39	.902	.31
4	.861	.36	.910	.48	.883	.18	.817	.51	.912	.53	.904	.39	.786	.34	.871	.27	.902	.30
5	.861	.39	.910	.46	.877	.49	.824	.31	.913	.44	.904	.48	.789	.27	.872	.25	.901	.43
6	.860	.42	.909	.55	.880	.38	.820	.43	.912	.51	.904	.46	.784	.32	.871	.29	.902	.35
7	.862	.25	.911	.31	.881	.28	.821	.42	.912	.48	.904	.43	.788	.28	.870	.34	.901	.38
8	.862	.23	.910	.49	.878	.45	.822	.36	.914	.26	.903	.55	.785	.36	.868	.46	.902	.30
9	.861	.35	.910	.45	.882	.18	.820	.44	.914	.32	.905	.37	.786	.32	.868	.44	.903	.26
10	.859	.47	.911	.38	.879	.42	.821	.39	.913	.40	.903	.51	.786	.29	.869	.43	.902	.34
11	.860	.43	.911	.33	.879	.41	.817	.52	.913	.43	.904	.44	.784	.33	.871	.31	.900	.47
12	.860	.43	.910	.48	.878	.51	.822	.35	.912	.51	.903	.48	.788	.26	.870	.37	.900	.51
13	.861	.34	.909	.56	.878	.51	.819	.48	.913	.45	.902	.59	.791	.19	.871	.31	.900	.50
14	.859	.48	.911	.33	.877	.53	.819	.46	.913	.48	.904	.45	.784	.34	.869	.42	.900	.52
15	.860	.43	.910	.50	.877	.51	.818	.50	.911	.60	.904	.44	.782	.43	.868	.46	.901	.40
16	.860	.40	.910	.51	.878	.51	.819	.46	.913	.41	.904	.41	.790	.24	.871	.28	.900	.51
17	.859	.51	.910	.42	.878	.52	.821	.42	.915	.24	.904	.43	.786	.36	.869	.41	.902	.33
18	.860	.40	.910	.49	.879	.41	.819	.45	.913	.37	.903	.54	.785	.33	.872	.22	.901	.44
19	.860	.43	.909	.52	.877	.57	.824	.30	.911	.63	.903	.57	.782	.37	.871	.29	.899	.55
20	.860	.46	.911	.38	.883	.14	.819	.48	.911	.58	.906	.27	.793	.13	.868	.48	.901	.40
21	.860	.45	.909	.58	.880	.32	.817	.49	.913	.37	.904	.40	.787	.25	.872	.20	.900	.51
22	.860	.41	.910	.44	.878	.50	.818	.46	.912	.59	.905	.39	.782	.36	.868	.47	.899	.56
23	.859	.50	.910	.47	.879	.43	.824	.57	.912	.58	.904	.41	.785	.31	.871	.28	.900	.45
24	.858	.51	.910	.50	.878	.47	.821	.40	.912	.48	.904	.41	.783	.37	.868	.44	.904	.23
25	.860	.40	.910	.48	.882	.23	.818	.51	.912	.51	.905	.35	.780	.32	.871	.29	.902	.33
26	.861	.37	.908	.63	.879	.39	.828	.38	.912	.54	.905	.37	.774	.41	.870	.38	.903	.28
27	.858	.49	.910	.43	.878	.49	.806	.67	.914	.34	.906	.27	.781	.33	.869	.43	.903	.31
28	.858	.51	.909	.55	.880	.35	.810	.65	.911	.60	.904	.46	.788	.36	.873	.13	.905	.14
29	.860	.38	.910	.45	.878	.48	.811	.69	.912	.45	.904	.47	.788	.40	.868	.44	.901	.41
30	.860	.42	.910	.46	.879	.41	.814	.64	.912	.46	.902	.56			.869	.41	.901	.43
31	.863	.44	.908	.59	.878	.43			.911	.58	.902	.59			.864	.55	.899	.47
32	.859	.47	.908	.61	.878	.50			.911	.58	.902	.63			.865	.61	.899	.58
33	.856	.55	.909	.54	.875	.66			.909	.68	.900	.71			.861	.62	.898	.59
34	.860	.44	.906	.72	.877	.51			.909	.72	.901	.64			.866	.49	.897	.63
35	.861	.41	.907	.69	.876	.63			.911	.65	.902	.64					.899	.60
36	.859	.46	.908	.67	.877	.57			.908	.78	.903	.61					.899	.53
37	.856	.58	.909	.63	.876	.59			.912	.61	.900	.69						

문항	2008						2009						2010					
	초		중		고		초		중		고		초		중		고	
	신뢰	변별	신뢰	변별	신뢰	변별	신뢰	변별	신뢰	변별	신뢰	변별	신뢰	변별	신뢰	변별	신뢰	변별
38	.858	.69	.909	.59	.876	.60			.910	.67	.901	.73						
39	.853	.66	.911	.68	.875	.60			.910	.72	.900	.71						
40	.861	.64	.907	.69	.877	.61			.910	.66	.905	.68						

* 두줄 선 안은 서답형 문항임.

** 음영 표시된 문항은 검사지 신뢰도를 낮추는 문항임.

*** 굵은 선으로 표시된 문항은 요인적재량 0.2미만인 문항임.

【부록 2】 검사지 문항별 요인적재량

순위	2008			2009			2010		
	초	중	고	초	중	고	초	중	고
1	.543	.635	.547	.661	.721	.680	.586	.694	.699
2	.530	.617	.520	.570	.670	.659	.535	.691	.662
3	.528	.607	.513	.570	.655	.652	.529	.652	.626
4	.490	.587	.500	.508	.630	.647	.508	.607	.620
5	.478	.557	.483	.508	.630	.596	.476	.560	.619
6	.471	.546	.482	.476	.617	.580	.476	.547	.605
7	.465	.545	.481	.468	.590	.571	.472	.544	.591
8	.455	.537	.473	.462	.582	.562	.467	.542	.590
9	.448	.530	.472	.450	.579	.541	.466	.521	.580
10	.444	.524	.469	.448	.578	.533	.451	.509	.574
11	.443	.523	.467	.435	.571	.532	.447	.506	.570
12	.427	.506	.465	.433	.571	.532	.446	.501	.567
13	.404	.468	.463	.422	.568	.525	.435	.484	.553
14	.395	.466	.455	.410	.530	.520	.427	.484	.544
15	.394	.460	.454	.408	.513	.519	.421	.481	.513
16	.382	.459	.450	.402	.511	.464	.420	.452	.493
17	.380	.459	.438	.399	.507	.448	.415	.447	.490
18	.380	.456	.433	.396	.505	.442	.409	.430	.474
19	.376	.455	.424	.394	.489	.428	.395	.429	.468
20	.371	.453	.414	.386	.486	.428	.374	.420	.454
21	.365	.448	.402	.375	.470	.419	.357	.408	.451
22	.361	.443	.378	.347	.470	.418	.336	.387	.446
23	.359	.435	.377	.317	.462	.418	.331	.372	.445
24	.352	.429	.373	.316	.446	.398	.318	.372	.429
25	.347	.429	.365	.285	.409	.393	.303	.365	.411
26	.344	.428	.364	.259	.407	.389	.281	.344	.402
27	.340	.424	.363	.257	.404	.384	.244	.337	.382
28	.330	.411	.334	.207	.402	.383	.242	.331	.372
29	.327	.407	.330	.203	.388	.376	.160	.326	.356
30	.321	.402	.319	.188	.375	.372		.318	.353
31	.315	.390	.309		.366	.364		.297	.351
32	.313	.372	.289		.360	.359		.243	.340
33	.301	.365	.278		.359	.349		.240	.323
34	.298	.326	.269		.329	.339		.149	.301
35	.298	.268	.266		.314	.314			.272
36	.291	.266	.216		.283	.312			.152
37	.274	.255	.148		.267	.306			
38	.256	.225	.144		.251	.301			

	2008			2009			2010		
순위	초	중	고	초	중	고	초	중	고
39	.213	.209	.101		.193	.195			
40	.198	.113	.088		.168	.187			

* 요인적재량의 순위별로 제시함.

** 음영 표시된 부분은 요인적재량이 0.2미만인 문항임.

【부록 3】 분석 대상 문항

* 지면 관계상 문항의 일부 내용을 생략함.
자세한 내용은 한국교육과정평가원 홈페이지를 참조 바람.

08-초-8. <자료>는 영화네 가족이 어젯밤 겪은 일을 그림으로 나타낸 것입니다.
다. 영화네 가족의 행동 변화를 잘 드러나게 쓴 것은 어느 것입니까?



- ① 영화네 가족은 대화를 하고 있었다. 그런데 정전이 되었다.
- ② 영화네 가족은 텔레비전 보기를 좋아한다. 그래서 대화하는 시간이 부족하다.
- ③ 영화네 가족은 텔레비전 보기를 좋아한다. 그래서 늦게까지 텔레비전을 보았다.
- ④ 영화네 가족이 텔레비전을 보고 있었다. 그런데 정전이 되어서 텔레비전이 꺼졌다.
- ⑤ 영화네 가족은 텔레비전을 보고 있었다. 그런데 정전이 되어서 대화를 하게 되었다.

※ 이번에는 방송 인터뷰를 들려 드립니다. 잘 듣고 물음에 답하십시오.

기자(여자) : 지금까지 지도력의 필요성에 대해 이야기를 해 보았습니다. 그럼, 어떤 사람을 지도자라고 할 수 있을까요? 이에 대해 세 분의 말씀을 들어 보겠습니다.
성인남자 : 인생에서 무엇이 가장 중요한지를 아는 사람이지요. 또, 그것을 이루기 위해 자기가 할 일이 무엇인지를 발견하는 사람일 겁니다.
여학생 : 흔히 지도자라고 하면 높은 자리에 앉아 다른 사람들에게 명령이나 내리는 사람을 떠올리지만, 사람들 위에 있다고 다 지도자는 아니죠. 진정한 지도자는 사람들을 바람직한 방향으로 이끌 수 있는 윤리 의식도 지녀야 한다고 생각합니다.

남학생 : 지도자는 뚜렷한 목표를 가지고 다른 사람들과 함께 그 목표를 향해 나아 가는 사람이 아닐까요?

기자(여자) : 지금까지 세 분의 말씀을 들어 보았습니다. 종합해 보면, 지도자는 ‘_____’라고 할 수 있겠군요. 우리나라에 앞으로 이러한 지도자가 많이 나타났으면 좋겠습니다.

08-중-3. (물음) 인터뷰 내용을 종합하는 기자의 말 중 생략된 내용으로 가장 적절한 것은?

- ① 자기 인생에 대한 책임감이 강한 사람이다.
- ② 대인 관계가 좋아서 인기가 많은 사람이다.
- ③ 자신을 바로 세우고 세상을 옳게 이끌어가는 사람이다.
- ④ 사회적 지위가 높아서 다른 사람들을 지휘하는 사람이다.
- ⑤ 인생에서 가장 가치 있는 것이 무엇인지 아는 사람이다.

※ 이번에는 강연을 들려 드립니다. 잘 듣고 물음에 답하십시오.

요즘 새롭게 등장한 용어 중 하나가 ‘생태맹’입니다. 생태맹이란 자연 생태계에 대한 지식 부족을 뜻하는 것이 아닙니다. 자연의 중요성과 신비함, 아름다움, 오묘함을 느끼지 못하는 감성이 부족한 상태를 뜻합니다. 이러한 감성이 부족하기 때문에 생태맹들은 자연과 더불어 살아갈 능력 또한 없습니다. 많은 현대인들이 이러한 생태맹에 해당하는데, 현대인의 삶과 우리 조상들의 삶을 비교해 보면 이 점을 잘 알 수 있습니다. 우리 조상들은 과일을 수확하더라도 나무 꼭대기에 두어 개쪼은 남겨 놓았습니다. 겨울에 먹이를 구하기 어려운 새의 양식으로 남겨 놓은 것입니다. 이뿐 아니라 뜨거운 물을 수채에 버릴 때에도 그곳에 살고 있는 벌레들을 생각해서 식혀서 버리곤 했습니다.

현대인들의 삶에는 이러한 배려가 부족합니다. 우리는 생태맹을 극복하기 위해 노력할 필요가 있습니다. 저는 생태맹을 벗어날 수 있는 가장 효과적인 방법으로 ‘숲을 가까이 하자’고 제안합니다. 숲은 자연 그 자체로서 생태적 감성을 기를 수 있는 좋은 체험장이기 때문입니다. 숲에서 자연과 교감하다 보면, 감성이 결여된 자연과학 지식은 자연과 함께 살아가는 데 필요한 지혜로 바뀔 수 있습니다. 숲에서의 자연 체험은 자연과 더불어 살아가는 지혜를 얻게 해 주기 때문입니다.

08-고-4. (물음) 강연의 목적을 고려할 때, 내용의 신뢰성을 높이기 위한 방안 중 가장 중요한 것은?

- ① 생태맹이란 용어의 기원에 대한 추가 설명
- ② 숲 생태계에 대한 구체적인 정보와 지식의 제시
- ③ 생태맹 극복을 위해 제시한 방법의 타당성 구체화

- ④ 현대 사회가 현대인을 생태맹으로 만드는 구체적 근거 제시
- ⑤ 우리 조상들의 삶의 방식 속에 들어 있는 지혜에 대한 추가 설명

08-고-9. '동맥경화 예방의 중요성과 예방법'에 관한 초고를 <자료>와 같이 작성하였다. 이를 수정하여 건강 관련 전문학술지에 게재하려고 한다. 검토 의견으로 적절하지 않은 것은?

<자 료>

- (가) (생략) 하지만 이제는 심장·뇌혈관 질환으로 인한 사망 원인이 1위 자리를 ㉠ 차지하고 있다. 정부는 동맥경화를 예방하고 관리하기 위해 전년도 예산 29억 원보다 134% 늘어난 68억 원을 책정하였다.
- (나) 그러나 국가가 동맥경화 환자를 관리하는 데는 한계가 있다. 환자 스스로 위험을 관리하지 않는다면 헛일일 뿐이다. 보통 혈관 안지름이 70% 이상 좁아지면 스트레스 상황에서 혈관이 막혀 장기가 정상적으로 작동하지 않게 된다. 이런 상황에서 특단의 대책을 세워도 혈관이 원상태로 회복되지 않는다. ㉢ 미미미리 노후 대책을 세워 놓지 않으면 안 되며, 식습관 및 운동습관 같은 생활습관을 바꿔야 할 이유도 여기에 있다.
- (다) 만일 혈압이나 콜레스테롤 수치가 정상에서 벗어났다면 식습관을 서둘러 바꾸어야 한다. 되도록 육류 섭취를 줄이고 녹황색 채소 위주의 식사를 하면 동맥경화 가능성을 40% 이상 줄일 수 있다. 녹황색 채소의 카로티노이드나 비타민 C, 셀레늄 등이 항산화 작용을 하며 '실포라페인' 같은 물질이 항염증 작용을 하기 때문이다.
- (라) '소리 없는 살인자'라 불리는 동맥경화이지만 어떤 식으로든 전조는 있게 마련이다. 다만 대부분의 사람들이 혈관의 신음소리에 주의를 기울이지 않을 뿐이다.

- ① (가)의 ㉠을 글의 성격에 맞게 '차지하고 있다'로 바꿔야겠어.
- ② (나)의 ㉢는 글의 통일성을 해치는 내용이므로 삭제해야겠어.
- ③ (다)의 전문적 용어들은 독자층을 고려하여 쉽게 풀어 써야겠어.
- ④ 글 전체의 흐름을 고려할 때, (다)에서 '운동을 통한 예방법'도 소개하는 것이 좋겠어.
- ⑤ (라)는 결론 문단이므로 예방의 중요성을 강조하는 내용을 간략히 추가해야겠어.

08-고-20. ㉠에 대한 설명 중 적절하지 않은 것은?

㉠ 아프리카 중부에 위치한 콩고는 콜탄이 많이 생산되는 나라이다.

- ① ㉠의 주성분은 2개이다.
- ② ㉠의 전체 주어는 '콩고는'이다.

- ③ ㉠은 관형절과 부사절이 있는 안은 문장이다.
 ④ ㉠의 첫 번째 안긴 문장의 주어는 문장 전체의 주어와 일치한다.
 ⑤ ㉠의 두 번째 안긴 문장은 ‘주어 + 부사어 + 서술어’의 구조로 되어 있다.

빈 도시락마저 들지 않은 손이 홀가분해 좋긴 하였지만, 해방촌 고개를 추어 오르기에는 뱃속이 너무 허전했다. (중략) ㉠저만큼 골목 막다른 곳에, 누런 시멘트 부대 종이를 흰 실로 열기설기 문살에 얹어댄 철호네 집 방문이 보였다. 철호는 때에 절어서 마치 가죽끈처럼 된 형견이 달린 문걸쇠를 잡아당겼다. 손가락이라도 드나들 만치 엉성한 문이면서 찌각찌각 집혀서 잘 열리지 않을었다. 아래가 잔뜩 집힌 채 비틀어진 문 틈으로 그의 어머니의 소리가 새어 나왔다.

⑥ “가자! 가자!”

미치면 목소리마저 변하는 모양이었다. 그것은 이미 그의 어머니의 조용하고 부드러운 그 목소리가 아니고, 쟁쟁하고 간사한 게 어떤 판 사람의 목소리였다. (중략)

두 달 전까지만 해도 ㉡철호는 저녁때 일터에서 돌아오면, 어머니야 알아들긴 말건 그래요, 어머니 지금 돌아왔습니다. 하고 인사를 하곤 하였었다. 그러나 요즘은 그것마저 안 하게 되었다. 그저 한참 물끄러미 굽어보고 섰다가 그대로 뒷방으로 올라와 버리는 것이었다.

킁킁한 구석에 앉아 있던 철호의 아내가 슬그머니 일어섰다. 담요 바지 무릎을 한쪽은 꺼명, 또 한쪽은 회색으로 기웠다. 만식이 되어서 꼭 바가지를 얹어 놓은 것 같은 배를 안은 아내는 몽유병자처럼 철호의 앞을 지나 나갔다. 부엌으로 나가는 것이었다. 분명 병어리는 아인데 아내는 말이 없었다.

“아버지.”

철호는 누가 꼭대기를 쿵 쥐어박기나 한 것처럼 흠칫했다.

바로 옆에 다섯 살 난 딸애가 눈을 동그랗게 뜨고 철호를 쳐다보고 있었다. 철호는 어린것에게로 얼굴을 돌렸다. 웃어 보이려는 철호의 얼굴이 도리어 흉하게 이지러졌다.

“나아, 삼춘이 나이롱 치마 사준됐다.” / “응.”

“그리구 구두두 사준됐다.” / “응.”

“그러면 나 엄마하고 화신* 구경 간다.” / “.....”

㉢철호는 그저 어린것의 노랑게 뜬 얼굴을 바라보고 있을 뿐이었다. 철호의 흰 셔츠 허리통을 잘라서 위에 끈을 꿰어 스커트로 입은 딸애는 짹짹 양말 목달이예다 어디서 주운 것인지 가는 고무줄을 끼웠다.

“가자! 가자!”

아랫방에서 또 어머니의 그 저주 같은 소리가 들려 왔다. 벌써 칠 년을 두고 들어와도 전연 모를 그 어떤 판사람의 목소리.

철호는 또 눈을 꼭 감았다. 머릿속의 낫줄이 팽팽히 헤워졌다. ㉣두 주먹으로 무엇이건 콧매려 부수고 싶은 충동에 철호는 어금니를 바서저라 맞췄었다.

— 이범선, <오발탄>

〈자 료〉

전쟁의 폐허 속에 미국의 경제 원조에 의지하여 겨우 유지되던 1950년대 사회현실은 당대인들에게 궁핍의 보편화를 강요하였다. 이러한 상황 속에 가장 고통을 당한 계층은 전쟁을 통하여 뿌리 뽑힌 계층으로 전락한 월남 난민, 고아, 미망인, 도시 빈민 등이었다.

- ① 철호 가족이 궁핍한 원인은 사회적 요인이 크군.
- ② 철호 어머니의 고통은 고향을 잃어버렸기 때문이겠군.
- ③ 철호가 사는 해방촌은 도시 개발에 따라 형성된 곳이겠군.
- ④ 철호 가족 모두 어머니처럼 남북통일에 대한 소망이 간절했겠군.
- ⑤ 중학교에 다녔다는 사실에 의하면 철호 가족은 전쟁 전에 유복했겠군.

09-초-서답 4. ㉠에 들어갈 알맞은 낱말을 쓰고, ㉡에 들어갈 낱말은 <보기>에서 골라 쓰시오. [2점]

〈보 기〉

도시, 하우스, 시골, 다이어리, 거실, 구두

구분	뜻	예
고유어	우리말에 본디부터 있던 말이나 그것에 기초하여 새롭게 만들어진 말	㉡
㉠	다른 나라 말에서 빌려 와서 우리말처럼 쓰이는 말	구두

㉠ : _____ ㉡ : _____

“육철은?”

선생님의 구구단 질문에 기남이는 답을 못하고 고개를 숙였다.

“수남이 넌, 동생 구구단도 안 가르치고 뭐했어!”

6학년인 형 수남이는 동생 때문에 선생님께 혼난 적이 한두 번이 아니었다. 4학년인 기남이와 다른 교실에서 배웠으면 했지만 소용없는 일이었다.

집으로 오는 길에 수남이는 개울에서 먹을 감고 있는 아이들을 보았다. 수남이도 먹을 감으려고 옷을 벗는데, 장터에서 오십 원에 새로 산 동생의 검정 고무신이 눈에 들어왔다.

‘기남이 이 자식, 혼 좀 나 봐라.’

수남이는 동생을 골탕 먹일 생각에 고무신 한 짝을 옥수수 밭에다 슬쩍 밀어 넣었다.

“어, 내 신 한 짝…….”

먹을 감고 나온 기남이는 신발을 찾지 못하여 신발 한 짝만을 책보에 넣고, 형이 소를 먹고 있는 산으로 갔다.

수남이는 시치미를 떼고 놀란 척 물었다.
 “나는 발이 와 그라노?”
 기남이는 신발을 잃어버렸다는 말을 차마 하지 못했다.
 “내가 신 찾아 주면 나는 뭐 해 줄래?”
 형이 신발을 숨겼다는 것을 눈치 챈 기남이는 화가 나서 형의 얼굴에 흠 한 줌을 뺨 뿌렸다.
 “너 이 자식, 거기 안 있을래?”
 수남이는 동생을 쫓아갔지만 잡지 못했다. 분한 마음에 되돌아와 보니 소가 보이질 않았다.
 얼마 후 기남이도 돌아왔다.
 “기남아, 나는 저쪽으로 가 봐라.”
 “알았다.”
 “기남아, 내 신 신고 가라.”
 수남이는 동생이 걱정되어 신발을 벗어 주었다. 기남이는 형의 고무신을 신고 소를 찾기 시작하였다. 수남이의 발은 돌부리에 치여 상처가 났지만 아픈 줄도 몰랐다.
 얼마나 헤맸을까, 기남이의 귀에 워낭 소리가 들려왔다. 소를 찾은 기남이는 워낭을 힘껏 흔들었다.
 “다행이다. 기남아.”
 “형, 발은 괜찮나? 안 아프나? 이제 형이 이 신 신어라.”
 기남이는 신발을 벗어 형에게 주었다.
 “기남아, 집에 가면 나랑 같이 구구단 외우자.”

09-초-19. 이야기에 나타난 당시의 삶의 모습을 잘못 이해한 것은 어느 것입니까?

- ① 형제가 한 교실에서 공부한 것으로 보아 학년 구분이 없었다.
- ② 먹을 감는다는 것으로 보아 아이들은 개울에서 물놀이를 했다.
- ③ 형제가 산에서 소를 먹이는 것으로 보아 아이들도 일을 하였다.
- ④ 고무신이 오십 원인 것으로 보아 오십 원의 가치는 지금과는 달랐다.
- ⑤ 장터에서 고무신을 산 것으로 보아 아이들은 고무신을 신고 다녔다.

(생략)

이러한 수면 부족은 인간의 건강에 심각한 영향을 끼친다. 우선 수면 부족은 비만을 유발한다. 영국의 한 대학에서 성인 1,000명을 대상으로 수면 시간을 10시간에서 5시간으로 줄이는 실험을 한 결과, 피실험자의 체중이 평균 4퍼센트 늘었다. 그들의 혈액을 분석해 보니 수면 시간을 반으로 줄였을 때, 식욕을 자극하는 성분은 증가한 반면 억제하는 성분은 그만큼 줄었다는 사실이 밝혀졌다.

- (가) 또한 수면 부족은 노화 현상과 유사한 반응을 일으킨다. 잠을 제대로 자지 못하면 정
 신적·신경적 측면에서 젊은 사람도 노인과 비슷해진다. 36시간 동안 잠 못 잔 20
 대 성인은 잠이 부족하지 않은 60대의 사람들과 유사한 행동을 하게 되는 것이다.

수면 부족 때문에 빚어지는 문제는 그뿐이 아니다. 옛 소련의 체르노빌과 미국 스리마일 섬의 원자력발전소 사고는 수면 부족에 의한 인재였다. 졸음으로 인한 기계 오작동이 원인이었던 것이다. 한국도로공사가 발표한 2007년 교통사고 원인들 중에서도 졸음운전이 23.3 퍼센트를 차지하고 있다. 졸음운전은 먼저 취소 수준을 훨씬 넘는 혈중 알코올 농도 0.17 퍼센트의 음주 운전과 상태가 비슷하다는 것이 전문가들의 분석이다. (이하 생략)

09-중-17. 위 글의 통일성을 위해 (가) 부분을 보완하고자 할 때, 그 보완 방향으로 가장 적절한 것은?

- ① 노화 현상의 원인을 분석하는 내용을 추가한다.
- ② 식욕이 수면 부족과 관련된 예를 덧붙여 내용을 보완한다.
- ③ 수면 부족과 노화 현상이 일으키는 반응을 항목별로 비교하며 재진술한다.
- ④ 수면 부족이 유발하는 다양한 신체적 부작용을 밝히는 내용으로 교체한다.
- ⑤ 동일 연령대의 사람들에게 수면 부족이 유발하는 현상을 중심으로 서술한다.

09-중-8. 다음 중 어법에 맞게 표현한 것은?

- ① 나는 형보다 키가 적다.
- ② 저 산 넘어 무지개가 있을까?
- ③ 밥을 먹던지 말던지 네 마음대로 해라.
- ④ 아버지께서 할아버지를 데리고 오셨다.
- ⑤ 민영이는 수박을 별로 좋아하지 않는다.

(앞부분 생략)

㉠로봇 청소기를 소비자에게 실제로 판매하는 회사는 최근까지 몇몇뿐이다.

(뒷부분 생략)

09-고-20. ㉠에 대한 다음 설명 중 적절하지 않은 것은?

- ① ㉠은 관형절을 안은 문장이다.
- ② ㉠의 안긴 문장은 또 하나의 문장을 안고 있다.
- ③ ㉠의 안은 문장과 안긴 문장의 주어는 일치한다.
- ④ ㉠에서 안은 문장의 주성분만 따지면 ‘주어+서술어’의 구조이다.
- ⑤ ㉠에서 안긴 문장의 주성분만 따지면 ‘주어+목적어+서술어’의 구조이다.

— 채만식, 〈레디메이드 인생〉

09-고-27. 위 글을 바탕으로 1930년대 사회상을 추론한 것 중 적절하지 않은 것은?

- ① 사회적으로 공장 노동자를 낮게 보고 있었군.
- ② 열차나 전보와 같은 서양 문물이 보급되어 있었군.
- ③ 체면이나 위신보다는 실질을 숭상하는 기풍이 일반화된 듯하군.
- ④ 지식인이나 노동자 모두 현실이 어려웠던 것은 별 차이가 없었겠군.
- ⑤ 학교에 다니지 못하고 일찍부터 공장에 나가 일하는 어린이들이 많았겠군.

(생략)

제주도 고인돌은 대부분 단독으로 자리한다. ㉠비록 고인돌이 모여 있다 할지라도 수십에서 수백 미터의 일정한 간격을 유지하고 있다. 전남 지방의 고인돌이 약 330제곱미터 안에 모두 모여 있는 모습과는 ㉡매우 대조적이다.

제주도 고인돌은 만드는 방식도 ㉢여간 독특하다. 바닥에 네 개의 기둥 돌을 놓고 그 위에 덮개돌을 얹어 만들었다. 그리고 덮개돌 윗면에는 여러 개의 홈을 파 놓았다. 이 홈은 다산과 풍요를 상징한다.

현재 제주도에 남아 있는 고인돌의 수는 120여 기*이다. 제주도의 고인돌이 우리나라에서 가장 늦게까지 만들어졌음을 고려하면 남아 있는 수가 ㉣그리 많지 않은 편이다. 그 까닭은 제주도 고인돌의 재료가 현무암이라서 다른 지역의 고인돌에 비해 깨지기 쉽고, 고인돌이 있던 지역에 촌락이 생겨 ㉤많이 훼손되었기 때문이다. 세계적 문화유산으로 가치를 인정받는 제주도 고인돌을 보존하려는 노력이 필요하다. *기: 무덤, 비석, 탑 따위를 세는 단위

10-초-20. 밑줄 친 ㉠~㉤ 중에서 잘못 쓰인 낱말은 어느 것입니까?

- ① ㉠ ② ㉡ ③ ㉢ ④ ㉣ ⑤ ㉤

10-중-28. 밑줄 친 단어 중, <자료>의 ㉦의 예로 적절하지 않은 것은?

— <자 료> —

한국어의 단어 중에는 ㉦공정문에서는 잘 쓰이지 않고 부정문에서 주로 쓰이는 특정한 단어들 있다. 따라서 문장을 쓸 때에는 문장 성분 간의 호응을 잘 살펴보아야 한다.

- ① 의문이 아직 가시지 않았다.
- ② 나는 그가 그다지 달갑지 않다.
- ③ 나는 하고 싶은 말이 별로 없었다.

- ④ 그의 행동을 도저히 참을 수 없다.
- ⑤ 이런 일은 비단 어제오늘의 일이 아니다.

(가) 맹사성(孟思誠), 〈강호사시가(江湖四時歌)〉
(나) 작자 미상의 사설 시조

- 10-고-24. (가), (나)에 대한 설명으로 적절하지 않은 것은?
- ① (가)는 계절의 상징성을, (나)는 반복적 일상을 통해 주제를 구현하고 있다.
 - ② (가)는 선경후정의 방식으로, (나)는 시간의 흐름에 따라 시상을 전개하고 있다.
 - ③ (가)는 과거형으로 서술하여 회상의 느낌을 주고, (나)는 현재형으로 서술하여 생동감을 주고 있다.
 - ④ (가)는 시적 화자의 삶의 태도를 직접 드러내는 데 비해, (나)는 인물의 삶의 태도를 간접적으로 드러내고 있다.
 - ⑤ (나)는 (가)와 달리 청각 효과를 매개로 정서를 드러내고 있다.

10-고-28. <자료>의 ㉠~㉣에 대한 설명으로 적절하지 않은 것은?

— < 자 료 > —

(가) 판매원 : 어머, ㉠**사장님!** 이 셔츠가 굉장히 잘 어울리시네요. ㉡**얼굴이 하얀 편이셔서** 집은 색깔이 참 좋으시고 이목구비가 확 살아 보이는 것 같으세요. 여기 잠시 기다리시고 계시면 새 상품으로 가져다 드릴게요. ㉢**진짜** 손해 보는 장사지만, 고객님한테만 특별히 싸게 해 드리는 거예요.

손님 : 아, 그래요? 그럼 이거 하나 살게요.

(나) 군대 지휘관 : ㉣**제군들에게** 필요한 것은 정신력입니다. 연병장을 열 바퀴 돌기 전에, 먼저 전투복 상의를 탈의합니다. 전원 ㉤**복장 해제!** 실시!

- ① ㉠은 손님의 실제 직함과 관계없이, 손님의 기분을 좋게 하기 위해서 쓴 말이다.
- ② ㉡은 ‘-(으)시-’를 적절하게 사용하여 손님을 높이는 뜻을 분명히 드러낸다.
- ③ ㉢은 사실 여부와 관계없이, 손님의 구매 의욕을 자극하기 위해서 한 말이다.
- ④ ㉣은 엄격한 분위기를 형성하기 위해서 문장을 ‘-에요’로 끝내지 않고 ‘-니다’로 끝내고 있다.
- ⑤ ㉤은 군복을 벗으라는 의미를 전달하기 위해 사용되고 있다.