

수준별 영어수업에 대한 평가방안 연구

이경인*, 김현주†

단국대학교

Lee, Kyoung In & Kim, Hyun-Ju. (2011). A study of the evaluation system for level-differentiated English classes. *Modern English Education*, 12(3), 305-321.

The purpose of this study is to investigate student and teacher perception of level-differentiated tests as opposed to traditional English tests and to examine the validity of the level-differentiated tests used in this study. Toward that purpose, this study had three groups of students—'approaching level', 'on level', and 'beyond level'. The students who participated in this study took two different kinds of level-differentiated tests. One (Type B) included all three levels' test items and allowed students to choose any test item suitable for them. The other (Type C) included only level-differentiated test items according to students' levels and students were not able to choose the items. The results of this study indicated that Type C is more valid than Type B and the traditional test in the approaching level group, but in other levels there was no significant difference between Types B and C. In terms of student and teacher perception, they strongly perceived the necessity of level-differentiated tests and most of them preferred Type B since they could choose the items suitable for their levels. This study implies an urgent need for an implementation of level-differentiated assessment and institutional framework in our education system.

[Level-differentiated assessment, traditional English tests, validity, perception/수준별 평가, 전통식 영어 평가, 타당도, 인식]

I. 서론

수준별 수업은 우리 교육에서 끊임없이 제기되고 있는 획일적인 교육 방식에 대한 문제를 해결하고자 나타난 맞춤형 교육방식이라고 볼 수 있다(Oakes, Gamoran & Page, 1992). 즉 과목별 학습 능력이 뛰어난 학생에게는 심화된 교수·학습 활동을 통해 학습에 대한 도전의식과 능력을 더욱 더 발전시킬 수

* 제1저자

† 교신저자

있는 기회를 제공하고, 학습 능력이 다소 떨어지는 학생에게는 보충적 교수·학습 활동을 통해 학습 부진의 누적을 예방하고 학생에게 자신감을 부여하여 수준별 교육의 효율성을 제고한 것이다. 특히, 영어수업은 ‘학습활동책(English Activity Book)’을 통해 수준별 수업에 대한 자료 및 활동 절차가 자세히 소개되어 있어 수준별 수업이 가능하도록 2007년 개정 교육과정은 제시하고 있다(교육인적자원부, 2007). 하지만 수준별 수업이 진행된 후 시행되는 평가는 학교 별로 다소 차이는 있지만, 대부분의 학교에서는 공통 평가를 통해 영어평가가 이루어지고 있다. 이는 현행 교과과정의 수준별 수업을 원활히 진행하기에는 평가상의 문제가 있음을 보여주는 것이다.

수준별 교육과정에서의 수업내용과 평가방식 간의 불일치로 인한 문제들을 해결하지 못하고 수준별 평가방법이 도입되지 않는다면 앞으로도 수준별 수업의 발전과 그에 따른 성과는 기대하기 어려울 것이다. 수준별 이동수업에서 학생들은 각자의 수준에 맞는 서로 다른 내용을 학습했음에도 불구하고 평가는 학생의 수준과는 관계없는 동일 수준의 공통 평가지로 시행되는 현실적 한계는 이미 지적된 바 있다(조난심, 김재춘, 허정철, 박순경, 1998). 교육정책 당국자, 학교 행정가, 학부모 등의 교육 평가에 대한 인식부족, 교사의 학생 평가에 대한 전문적 지식 부족으로 인해 올바른 수준별 교육과 평가가 이루어지지 않았으며(김신영, 백순근, 채선희, 1998; 배호순, 2001), 지금까지의 교육과정이 전반적으로 원론의 내용에 주목해 왔기 때문에 각 교과 영역에서의 학급별, 학년별 목표수준을 체계화하지 못해 학생 평가의 소홀함 또한 지적되고 있다(김명숙, 2003).

우리나라 수준별 수업에 따른 수준별 평가의 실태를 조사한 결과에 따르면(진경애, 송미영, 김화경, 2009), 지필평가와 수행평가 모두 동일하게 평가하는 학교가 62.8%, 지필평가는 동일하게 시행하고 수행평가만 수준별 평가를 시행하는 학교가 28.0%, 지필평가만 수준별 평가를 하는 학교가 6.0%로 나타났다. 그리고 지필평가와 수행평가 모두 수준별 평가를 시행하는 학교는 3.2%로 수준별 평가는 매우 드물게 시행되고 있다는 것을 알 수 있다. 또한 수준별 수업에 따른 평가상의 어려움에 대해서 교사들은 현실적으로 교사가 가지고 있는 업무 부담으로 인한 타당한 평가 시행이 어렵고, 수준별 수업에 대한 차별화된 평가시스템 구축과 객관적인 평가기준의 마련이 시급하다는 점을 지적하였다. 이는 아직까지도 수준별 수업에 따른 수준별 평가의 시행이 매우 미흡하며, 현장에서 시행 가능한 평가방법과 명확한 평가기준 및 시스템 구축의 필요성을 시사한다. 즉, 학생들에게 서로 다른 내용을 가르치면 평가의 내용 또한 각 과정에 맞는 적합한 평가를 실시하여, 가르치는 내용과 평가하는 내용과의 괴리문제를 해결해야 한다. 이러한 점을 고려해 볼 때, 수준별 교육과정의 성공적인 운영을 위해서는 학교 현장에서 적용 가능한 타당한 평가 방안에 대한 연구가 요구된다. 하지만 지금까지 수준별 교육과정의 학생평가에 관한 선행 연구들은 대부분 수준별 수업에 따른 평가의 당위성에 대해

논하거나 특정 평가만을 제시하고 있어 수준별 다양한 평가방안을 제시하지는 못하고 있다.

따라서 본 연구에서는 영어 보조 교과서를 활용한 평가에 주안점을 두고 실용가능한 수준별 평가방안을 제안하고자 한다. 또한 학생 및 교사를 대상으로 기존의 동일 평가와 수준별 평가에 대한 인식을 살펴보고, 평가결과를 통해 수준별 평가에 대한 신뢰성 및 타당성을 살펴보고자 한다.

II. 이론적 배경

1. 수준별 평가의 의미와 필요성

우리나라의 수준별 수업은 학생의 성적에 의해 나누어진 수준별 반에서 심화, 기본, 보충 내용을 각각 지도하는 것을 말하며(김주훈 등, 1998), 수준별 평가는 수준별 수업에서 이루어진 학습에 대한 평가를 의미한다(진경애 등, 2009). 따라서 수준별 교육과정에 따른 수준별 평가가 이루어지는 것이 타당하나, 현재 우리나라에서 시행되는 수준별 평가는 학생들이 학습한 내용과 일치하지 못한다는 점에서 문제가 야기되고 있다. 내용 타당도 측면에서 교수·학습 내용과 평가의 내용이 일치되어야 하는 것은 매우 중요한 일이며(Brown, 2002), 이는 또한 학습효과를 극대화시킬 수 있는 평가의 긍정적 역류현상을 일으키기도 한다. 수준별 수업의 목표는 학습자의 개인차와 상황을 고려하여 수업 목표를 설정하고 이에 따른 교수·학습이 이루어지는 것이므로 이에 대한 타당한 평가가 이루어지는 것이 마땅하다.

제7차 영어과 교육과정의 목표가 학습자 중심의 교육과정의 운영에 있으며(교육인적자원부, 2007), 이를 위한 수준별 수업에 대해 수준별 평가가 이루어져야 하는 것은 당연한 일일 것이다. 하지만 1997년 수준별 이동수업이 적극 권장되기 시작한 이래 지금까지 수준별 수업에서의 평가 문제는 여전히 해결되지 않은 상태로 남아 있다. 수준별 이동 수업을 실시함에도 불구하고 가르치는 내용과 평가하는 내용이 괴리되는 이유는 첫째, 모든 동급 학년 학생들에게는 동일한 내용을 가르치도록 현행 교육과정에서 규정하고 있고(국민공통기본 교육과정), 둘째, 이러한 교육과정에서 동일한 평가를 수행하여 평가결과를 학년별 또는 동일 교과이수자 전체를 대상으로 상대적 등급·석차를 산출하도록 규정하고 있기 때문이다(김재춘, 2009). 이렇듯 대부분의 학교에서 수준별 수업을 실시함에도 불구하고 이에 맞는 수준별 평가가 제대로 이뤄지고 있지 않기 때문에 수준별 교육의 목표와 결과는 다를 수밖에 없다.

평가는 또 다른 학습의 시작으로써 교수·학습과는 때려야 뗄 수 없는

유기적 관계에 놓여있다. 수준별 수업의 목표를 달성하기 위한 수준별 평가의 문제를 해결함으로써 공정하고 타당한 평가가 이루어져야 할 것이다.

2. 수준별 수업에 대한 평가 방안

국내 현행 평가 사례를 토대로 정리하면 수준별 수업에 대한 평가는 크게 두 가지로 설명된다.

1) 단일 형 공통 평가

단일 형 공통 평가는 현재 현장에서 가장 선호하는 방식의 평가로서 수업은 수준별로 하지만 평가는 수준과 상관없이 공통된 단일 문항으로 평가하는 방식이다. 따라서 학생들은 자신의 수준과 맞지 않는 문항들을 풀어야 하는 결과가 나타나 학생들에게 학습 동기를 부여하는데 도움을 주지 못하고 있다(김홍원, 조인진, 나장함, 차동춘, 김재춘, 2004). 즉 수준별 수업이 시행되기는 하지만 보충 반에 속한 학생들의 사기를 저하시킬 수 있다는 문제점 등을 고려하여 보충 반과 기본 반에 속한 학생들의 수준에서 문제를 출제하는 경우가 발생하거나 변별력을 고려하여 심화반에 속한 학생들의 수준에서 문제를 출제하는 경우가 빈번하다. 이는 학습자의 학습 동기를 저하시킬 수 있으며 수준별 수업의 실효성 또한 저하시키는 일이라고 분석된다(김미숙, 2008).

2) 수준별 평가

수준별 수업의 효과성을 높이기 위해 제안된 여러 가지 평가방법이 있지만(김재춘, 1999; 김홍원 등, 2004; 박선화, 2005), 이는 크게 부분적 수준별 평가방법, 전면적 수준별 평가방법, 혼합식 평가방법으로 나누어 설명할 수 있다. 부분적 수준별 평가방법은 대부분 공통문항으로 출제하되 일부 문항만을 수준별 문항으로 출제하여 학습자가 자신에게 맞는 수준의 문항을 골라 평가 받을 수 있는 평가방식이다. 전면적 수준별 평가방법은 학습자의 수준을 반영하여 수준별로 학습자 자신이 속한 과정의 문제들만 풀 수 있도록 하여 학생의 수준과 수준별 평가결과를 함께 학교생활기록부에 기록하는 방식이다. 마지막으로 혼합식 평가방법은 공통평가와 전면적 수준별 평가를 모두 실시한 후 학교생활기록부에 모두 기록하는 방식이다.

3. 수준별 평가지의 종류

학생들의 수준을 고려하여 진행되는 수준별 수업에 대한 평가지는 크게 부분적 수준별 평가지와 전면적 수준별 평가지로 구분된다.

1) 부분적 수준별 평가지

현재 수준별 이동 수업을 시행하고 있는 현장에서 학생들에게 공정한 평가를 적용하려는 적극적인 노력의 한 형태로 실천되고 있는 수준별 평가방식이라 할 수 있다(진경애, 송미영, 반재천, 2010). 수행평가뿐 만 아니라 중간고사나 기말고사와 같은 정기고사에서도 수준별 평가를 적용하려는 노력으로 평가지의 대부분 문항은 공통으로 출제하지만 일부 문항을 학생들의 수준(보충, 기본, 심화)을 고려하여 각 과정별 수준에 맞는 문항을 선택형 세트문항으로 출제하여 학생들이 자신이 풀 수 있는 수준의 문항을 자유롭게 선택하여 풀 수 있게 한다. 배점은 각 과정의 수준별로 문항에 따라 차이가 있고, 선택 문항에 따라 총점이 달라지게 된다. 만약, 모든 선택형 세트문항에서 심화 수준의 문항을 선택하고 만점을 받은 경우는 100점이 되지만, 세트문항에서 기초나 기본 수준의 문항을 선택한 경우는 100점 만점이 아닌 각 과정 별 수준에 맞는 만점이 주어지는 방식이다. 100점 만점이 되는 경우는 심화수준의 문항을 선택하여 모두 맞춘 경우에만 성립될 수 있는 반면 기본 반이나 보충 반에 속한 학생들도 자신이 풀 수만 있다면 자신이 속한 과정보다 높은 수준의 문제를 풀어 맞았을 경우에는 보다 높은 점수를 받을 수도 있는 평가 방식이다. 주관식과 객관식문항 모두 출제 가능하며 문항의 배점도 다양하고 만점의 기준도 각기 다르다. 복수선택도 허용되지만 선택 시 보다 높은 과정의 배점만을 인정하는 방식이다. 예를 들어, 그림 1은 D중학교의 2011년 1학기 영어 기말고사의 세트문항이다.

* 다음은 수준별 문항입니다. (A형, B형, C형 문제 중 자신 있게 풀 수 있는 문항을 선택하여 풀되, 2문항 이상 풀었을 경우 배점이 높은 문제의 점수가 인정됩니다.)

방 안 그림

A> 위 그림의 내용에 맞게 알맞은 표현을 고르시오. (2점)

There is a mouse (in / on) the cup.

B> 위 그림의 내용에 맞게 빈 칸에 들어갈 표현을 쓰시오. (4점)

(1) There is a mouse _____ the door.

(2) There is a mouse _____ the chair.

C> 위 그림의 내용에 맞게 주어진 단어를 넣어 3문장을 쓰시오. (6점)

(1) _____

(2) _____

(3) _____

그림 1. D중학교 1학년 (2011-1학기 기말고사).

이와 같은 부분적 수준별 평가지에 대해, 최용익(2009)은 현실 적용 가능성이 높은 “수준별 세트문항”을 이용한 평가방법이 수준별 평가로서 타당하고 효과적인 평가방안인지 살펴보았다. 그 결과, 수준별 세트문항을 포함한 평가가 단일 평가지보다 학생들의 성취도가 높게 나타난 것을 발견하였다. 즉, 상 수준의 학생들은 세트문항 내에서 심화 문제를 선택하여 고득점을 얻고, 하 수준의 학생들은 세트문항 내에서 보충 문제를 선택하여 성적이 오른 것으로 나타났다. 따라서 “수준별 세트문항”을 이용한 평가가 수준별 평가방법으로 타당하고 효과가 있음을 보여주고 있다. 하지만 이 연구는 서술형 평가에만 국한되어 있어 현재 현장에서 사용하고 있는 평가방식으로 적용하기에는 어려움이 있다. 또한, 수준별 평가 문항에 대한 연구(유세희, 2010)에서는 문제 출제 과정에서 문항 분석 과정이 반드시 이루어져야 함을 제시하고 있다. 예를 들어, 수준별 평가 문항을 제작할 당시 출제 교사들의 예상과는 달리 정답률이 높으리라고 생각했던 낮은 배점의 문제보다 학생들이 어려워할 것이라고 예상했던 배점이 높은 문제의 정답률이 높게 나타나는 경우가 종종 있다. 따라서 출제자들은 왜 이러한 결과가 발생했는지에 대한 평가 문항 분석 과정을 반드시 거쳐야 함을 제시하고 있다(유세희, 2010).

2) 전면적 수준별 평가지

전면적 수준별 평가지는 수준별로 나누어진 각기 다른 수준의 독립적 평가문항으로 구성된 평가지를 말한다. 이는 수준별 수업의 효과성을 극대화시킬 수 있는 평가지로 보여지나, 우리나라의 현 내신 체제에는 적용시키기에 다소 무리가 있어 보인다. 왜냐하면 이를 적용시키기 위해서는 학생생활기록부에 성적을 기록하는 방식까지 전면적 혁신이 요구되기 때문이다(진경애 등, 2010).

전면적 수준별 평가지의 만점은 수준별로 보통 100점, 90점, 80점 정도로 차별화가 된다. 이러한 경우에는 상 수준 평가지에서의 90점과 중 수준 평가지의 90점이 동일한 능력을 보여주는 것이라야 한다. 따라서 이러한 평가지는 교사들 간의 충분한 논의와 합의를 거치고 문항에 대한 철저한 검토 후에 시행되어야 할 것이다. 또는 통계적 점수 조정을 거쳐 전면적 수준별 평가지의 점수를 산출할 수도 있다. 하지만 이러한 방식 또한 전체 학생을 상대평가하는 우리나라의 내신 체제에서는 적용하기가 매우 어려운 것이 현실이다(진경애 등, 2010).

III. 연구방법

본 연구는 제7차 교육과정에 따른 수준별 수업을 위한 학습 자료를 보완하기 위해 개발된 수준별 선택 세트형 평가 문항을 제작하고 이를 현장에서 적용한 연구이다. 2010년 10월, 경기도에 소재한 일반계 고등학교 한 곳을 선정하여 1학년 영어교과를 담당하는 교사들과 1학년 학생들을 대상으로 연구를 수행하였다.

1. 연구대상

본 연구는 고등학생 총 99명과 영어과 교사 10명을 대상으로 수행되었다. 본 연구의 1학년 영어과 수준별 이동 수업 분반 형태는 표 1과 같으며 본 연구에 참여한 학생들은 심화반(1반: 32명), 기본반(8반: 33명), 보충반(5반: 34명)으로 구성된 총 99명이다.

표 1

1학년 영어과 수준별 이동 수업 분반 형태

반	1반	2반	3반	4반	5반	6반	7반	8반	9반
인원	32명	31명	33명	33명	34명	33명	35명	33명	45명
수준별 수업반	심화반	심화반	심화반	보충반	보충반	보충반	보충반	기본반	기본반

2. 연구 도구

1) 평가지

본 연구는 전통적 동일형 평가방식과 수준별 문항을 활용한 평가방식을 비교하기 위하여 난이도를 달리한 수준별 문항을 각 수준별로 15문항, 총 45문항을 개발하여 본 실험에 앞서 24명의 타 고등학교 학생들을 대상으로 사전연구를 수행하였다. 그 결과, 문항의 난이도 및 변별도에 문제가 있는 문항들을 수정하여 수준별 평가지 두 세트—유형 2(B형 평가지)와 유형3(C형 평가지)—의 평가도구를 개발하였다. 유형 2는 학생들이 수준에 맞는 문항을 선택하여 풀도록 한 부분적 수준별 평가지이며, 유형 3은 학생들의 수준에 맞는 문항만을 풀도록 한 전면적 수준별 평가지이다(표 2 참조). 유형 1은 전통적 평가방식의 공통평가지로서 똑같은 평가 문항을 가지고 수준에 상관없이 문제를 풀도록 하는 것이다.

표 2

유형 2와 유형3의 평가지 구성 예

평가지 유형	내용	구성 예시
		*11~25번까지는 주어진 세 문항 중, 수준에 맞는 문제를 골라 지시에 맞는 답을 하시오. (세 개의 문항을 모두 풀 수 있으나 가장 높은 점수만 인정함)
유형 2 (B형 평가지)	11-1. 주어진 단어에 대한 의미로 알맞은 것을 찾아 연결하십시오. (3점) 11-2. 다음 밑줄 친 단어와 같은 의미로 바르게 연결되지 않은 것을 고르시오. (4점) 11-3. 우리말에 맞는 알맞은 표현으로 문장을 완성하십시오. (5점)	
		*11~ 25번까지는 자신에게 속한 과정의 문제만 푸시오. (다른 과정 문제를 풀었을 경우 점수 인정 안 됨)
유형 3 (C형 평가지)	11.<보충> 주어진 단어에 대한 의미로 알맞은 것을 찾아 연결하십시오. 11.<기본> 다음 밑줄 친 단어와 같은 의미로 바르게 연결되지 않은 것을 고르시오. 11.<심화> 우리말에 맞는 알맞은 표현으로 문장을 완성하십시오.	

2) 설문지

본 설문은 유세희(2010)의 수준별 영어 독해 수행평가에 대한 연구 및 서영원(2007)의 교과서 평가모형 개발을 위해 조사된 수준별 수업에 대한 인식 설문내용들을 통해 도출된 결론을 바탕으로 학생용과 교사용으로 각각 제작되었다. 학생용 설문지의 1~5번 문항은 현재 정기고사에서 실시되는 공통평가지에 대한 내용을, 6~10번 문항은 수준별 수업에 따른 수준별 평가지에 대한 내용으로 구성하였다. 또한, 11번과 12번 문항은 개발된 유형 2와 유형 3의 수준별 평가지 중 학생들이 선호하는 수준별 수업에 따른 수준별 평가지는 무엇이며 그렇게 생각하는 이유는 무엇인지 살펴보았다. 교사용 설문지의 1~4번 문항은 현재 정기고사에서 시행되고 있는 평가방식에 대한 만족도를 살펴보고, 5~8번 문항은 수준별 수업에 따른 수준별 평가에 대한 교사의 인식을 살펴보았으며 9번 문항은 정기고사에서 실시 가능한 수준별 평가유형 중 교사들이 생각하는 가장 바람직한 평가유형을 살펴보고 12번 문항을 통해 이에 대한 이유를 살펴보고자 하였다. 요약하자면, 각각의 설문은 크게 세 가지—현행 평가지에 대한 만족도, 수준별 평가지에 대한 인식, 앞으로 시행되기를 바라는 평가지 및 그에 대한 이유—들로 구성되었다. 각 설문지의 영역별 신뢰도를 살펴보기 위하여 Cronbach α 계수를 산출한 결과, 영역별 α 계수가 낮은 부분도 있었으나, 전체적으로는 학생용 설문지는 .827, 교사용 설문지는 .928로 높게 나타났다.

3. 연구 절차 및 자료 분석

먼저 본 연구자들에 의해 제작된 수준별 문항에 대한 타당도와 신뢰도를 검증하기 위해 본 연구에 참여하지 않은 24명의 학생들을 대상으로 평가를 시행하였다. 문항의 난이도, 변별도 및 신뢰도를 산출하고 문제가 있는 문항들은 수정한 후 본 연구에 사용되었다. SPSS 18.0 통계 프로그램을 활용하여 평가결과에 대한 자료 분석은 각 집단별 기술통계, 일원배치 분산분석, 사후검정을 하였으며 설문지 결과에 대한 자료 분석은 빈도분석을 통해 제시하였다. 기술 통계에 앞서 평가 점수에 대한 Cronbach α 계수를 산출하여 신뢰도를 분석하였으며 평가 및 설문결과를 토대로 수준별 수업에 대한 평가방안을 제안하고자 하였다.

IV. 연구 결과 분석

1. 신뢰도와 변별도 분석

본 연구에서는 사전조사를 거쳐 신뢰도가 검증된 문항들을 이용하여 평가지 유형 B와 유형 C를 제작하였다. 제작된 평가문항에 대한 신뢰도 검정을 한 결과는 표 3과 같다. 사전조사는 평가지의 Cronbach α 계수보다 각 문항의 '항목이 삭제된 경우의 Cronbach α 계수'가 큰 경우 전체 신뢰도에 좋지 않은 영향을 주는 것으로 간주하여 삭제한 후 수정된 문항으로 다시 신뢰도계수와 각 문항의 항목이 삭제된 경우의 신뢰도 계수를 비교한 결과이다.

표 3
수준별 문항의 신뢰도 통계량

문항	Cronbach α	항목 수
보충 문항	.785	15
기본 문항	.964	15
심화 문항	.947	15

항목이 삭제된 경우 Cronbach α	변별도	항목이 삭제된 경우 Cronbach α	변별도	항목이 삭제된 경우 Cronbach α	변별도
보충 1 .774	.338	기본 1 .963	.519	심화 1 .941	.820
보충 2 .781	.341	기본 2 .956	.833	심화 2 .939	.786
보충 3 .767	.587	기본 3 .963	.519	심화 3 .940	.845
보충 4 .779	.276	기본 4 .959	.762	심화 4 .937	.839
보충 5 .774	.338	기본 5 .954	.977	심화 5 .935	.898
보충 6 .783	.206	기본 6 .956	.899	심화 6 .943	.749
보충 7 .758	.606	기본 7 .960	.722	심화 7 .942	.934
보충 8 .767	.419	기본 8 .962	.590	심화 8 .946	.634

보충 9	.733	.705	기본 9	.957	.835	심화 9	.940	.774
보충 10	.774	.338	기본 10	.957	.865	심화 10	.936	.886
보충 11	.781	.259	기본 11	.957	.793	심화 11	.939	.845
보충 12	.783	.141	기본 12	.958	.884	심화 12	.933	.852
보충 13	.773	.488	기본 13	.959	.835	심화 13	.940	.831
보충 14	.779	.276	기본 14	.957	.757	심화 14	.934	.888
보충 15	.774	.338	기본 15	.957	.803	심화 15	.938	.840

보충반, 기본반, 심화반의 문항 전체 신뢰도를 나타내는 Cronbach α 계수는 0.785, 0.964, 0.947으로 높은 신뢰도를 보이고 있으며 항목이 삭제된 경우의 Cronbach α 계수를 통해서도 각각의 문항들의 신뢰도 또한 전체 신뢰도를 높여주는 문항들임을 알 수 있었다. 문항간 변별도는 심화 문항이 보충, 기본 문항보다 훨씬 높은 변별도를 보였다.

2. 평가지 유형에 따른 점수 비교

기존의 공통 평가지(A형)와 수준별 평가를 위해 제작된 평가지(B형과 C형)를 활용하여 평가를 시행한 결과는 표 4와 같다.

표 4

평가지별 점수 비교

	평가지 종류	평균	표준편차
보충반	기존 공통 평가지(A형)	57.88	12.977
	유형 2(B형 평가지)	71.89	7.849
	유형 3(C형 평가지)	85.95	14.965
기본반	기존 공통 평가지(A형)	79.57	14.938
	유형 2(B형 평가지)	87.94	12.749
	유형 3(C형 평가지)	91.20	8.034
심화반	기존 공통 평가지(A형)	88.36	9.972
	유형 2(B형 평가지)	93.55	12.035
	유형 3(C형 평가지)	94.73	7.938

평가지에 따른 점수 차이를 살펴보기 위해 세 집단(보충반, 기본반, 심화반) 모두 1차 평가와 2차 평가를 일주일 간격으로 치렀다. 1차 평가는 수준별로 배점을 달리하여 수준별 선택 세트문항을 제공해 학생들이 자유롭게 자신의 수준에 맞는 문항을 선택하여 풀게 하는 방식의 평가지(B형)를 2차 평가는 자신이 속한 수준의 문제만을 풀게 하는 수준별 평가방식의 평가지(C형)를 치르게 하였다. 기존의 공통 평가지(A형)는 대상 학교 학생들이 미리 치른 중간고사의 성적을 기준으로 기록한 것이다.

일원배치 분산분석을 통해 집단 내 점수 결과를 비교한 결과, 평가지에 따라 각각의 평균차이가 있음을 알 수 있었다(표 5 참조). 또한, 이러한 결과 차이가 어떠한 평가지 간의 차이에서 기인하였는지를 검토하기 위해 추가적인 통계분석인 사후분석(Tukey HSD)을 하였다(표 6~8 참조). 그 결과, 보충반·기본반·심화반 모두 기존의 평가지(공통 평가지)에 비해 본 연구에서 제작된 수준별 평가지에서 더 높은 영어점수를 보여주었다. 또한 보충반에서는 수준별 평가지 간의 차이 또한 발견할 수 있었다. 이는 일반적으로 학습 능력이 떨어지는 보충반 학생들에게는 자신의 수준에 맞는 문제를 자유롭게 선택하여 풀게 하는 평가방식이 있다 하더라도 높은 배점의 중·상 수준의 문제를 푸는데 어려움이 있어 C유형의 평균이 B유형의 평균보다 높게 나타난 것으로 해석되었다. 따라서 보충반 학생들에게는 기존의 평가지(A형)에 비해 자신이 속한 과정의 문제만을 풀게 하는 C형 평가지가 더 쉬울 수 있으며 기본반과 심화반 학생들에게는 기존의 평가지(A형) 보다는 수준별 평가지 B형 또는 C형이 더 쉬운 평가지임을 알 수 있었다.

표 5

수준별 평가지간 분산분석

		제곱합	자유도	평균 제곱	F	유의확률
보충반	집단-간	12474.430	2	6237.215	39.051	.000
	집단-내	15333.219	96	159.721		
	합계	27807.649	98			
기본반	집단-간	2819.674	2	1409.837	27.952	.000
	집단-내	4993.411	99	50.438		
	합계	7813.085	101			
심화반	집단-간	731.231	2	365.616	17.795	.000
	집단-내	1910.767	93	20.546		
	합계	2641.998	95			

표 6

보충반 사후분석 결과

평가지 종류(I)	평가지 종류(J)	평균차(I-J)	표준오차	유의확률	95%신뢰구간	
					하한값	상한값
1	2	-14.03*	3.112	.000	-21.440	-6.626
	3	-27.49*	3.475	.000	-34.900	-20.087
2	1	14.03*	2.198	.000	6.626	21.440
	3	-13.46*	3.154	.000	-20.867	-6.053
3	1	27.49*	3.198	.000	20.087	34.900
	2	13.46*	3.122	.000	6.053	20.867

표 7
기본반 사후분석 결과

평가지 종류(I)	평가지 종류(J)	평균차(I-J)	표준오차	유의확률	95%신뢰구간	
					하한값	상한값
1	2	-8.36*	1.693	.000	-12.397	-4.338
	3	-11.63*	2.347	.000	-15.664	-7.605
2	1	8.36*	1.691	.000	4.338	12.397
	3	-3.26	1.002	.136	-7.297	.761
3	1	11.63*	1.998	.000	7.605	15.664
	2	3.26	1.012	.136	-.761	7.297

표 8
심화반 사후분석 결과

평가지 종류(I)	평가지 종류(J)	평균차(I-J)	표준오차	유의확률	95%신뢰구간	
					하한값	상한값
1	2	-5.17*	1.133	.000	-7.870	-2.472
	3	-6.35*	1.237	.000	-9.055	-3.657
2	1	5.17*	1.145	.000	2.472	7.870
	3	-1.18	1.172	.550	-3.883	7.514
3	1	6.35*	1.987	.000	3.657	9.055
	2	1.18	1.172	.550	-1.514	3.883

3. 수준별 수업을 위한 평가방식에 대한 인식

기존의 공통 평가지와 수준별 평가지에 대한 학생들과 교사들의 인식을 알아보기 위해 제작된 설문 문항에 대한 빈도분석 결과는 표 9와 같다. 현재 사용되고 있는 공통 평가지에 대한 인식을 살펴보면, 보충반과 기본반의 학생들은 수준별 수업과 상관없이 시행되는 공통 평가지에 대해 매우 부정적인 인식을 가지고 있는 반면, 심화반 학생들은 공통평가지의 난이도, 동기부여도, 학업 성취도 등의 측면에서 만족스러워하고 있는 것으로 나타났다. 교사들은 과반수이상인 현재 시행되고 있는 공통 평가방식에 대해 부정적인 인식을 가지고 있는 것으로 나타났다.

수준별 평가지에 대한 인식을 살펴보면, 대다수의 학생들이 자신의 수준에 맞는 문항으로 구성된 평가지에 대해 매우 긍정적인 인식을 가지고 있는 것으로 나타났다. 교사들 또한 수준별 평가지가 학생들에게 자신감, 동기부여, 성취감 등을 향상시켜주는데 긍정적인 역할을 할 것이라고 대다수 인식하고

있는 것으로 나타났다. 이러한 결과를 종합해보면, 기존의 공통평가지 문항들은 다소 상위권 학생들의 수준에 적절한 문항들로 구성되어 보충반이나 기본반 학생들에게는 어려움이 있었으나 심화반 학생들에게는 수준별 평가지의 상 수준의 문항이나 기존 공통평가지의 문항에서 큰 차이를 느끼지 못하여 기존의 공통평가지와 수준별 평가지 모두에서 긍정적인 인식을 나타낸 것으로 보인다. 반면, 중·하위 수준의 학생들에게는 공통 평가지로는 영어 학습에 대한 자신감이나 동기부여를 하는데 어려움이 있을 수 있겠다. 따라서 수준별 수업에 따른 수준별 평가의 도입이 강하게 요구되고 있다는 것을 확인할 수 있었다.

표 9
공통 평가지와 수준별 평가지에 대한 인식

평가 지	문항	부정적		보통		긍정적	
		학생	교사	학생	교사	학생	교사
기 본 공 통 평 가 지	만족도	보충	33(97.1%)		1(2.9%)		
		기본	33(100%)	8(80%)		2(20%)	0(0%)
		심화	4(12.5%)		1(3.1%)		27(84.4%)
	동기 부여	보충	33(97.1%)		1(2.9%)		
		기본	28(84.9%)	8(80%)	5(15.1%)	2(20%)	0(0%)
		심화	4(12.5%)		1(3.1%)		27(84.4%)
	난이도	보충	33(97.1%)		1(2.9%)		
		기본	27(81.8%)	8(80%)	6(18.2%)	2(20%)	0(0%)
		심화	5(15.6%)		4(12.5%)		23(71.9%)
자신감	보충	33(97.1%)		1(2.9%)			
	기본	29(84.8%)	8(80%)	5(15.2%)	2(20%)	0(0%)	
	심화	4(13.5%)		8(25.0%)		20(62.6%)	
영어 성취도	보충	32(94.1%)		2(5.9%)			
	기본	15(45.5%)	8(80%)	17(51.5%)	2(20%)	1(3.0%)	
	심화	2(6.3%)		10(31.3%)		20(62.5%)	

수 준 별 평 가 지	요구 도	보충		2(5.9%)		32(94.1%)	
		기본	1(10%)	1(3.0%)		32(97.0%)	9(90%)
		심화	2(6.3%)	9(28.2%)		23(65.6%)	
수 준 별 평 가 지	동기 부여	보충		2(5.9%)		32(94.1%)	
		기본		1(3.0%)	3(30%)	32(97.0%)	7(70%)
		심화	2(6.3%)	8(25.0%)		22(68.7%)	
수 준 별 평 가 지	자신 감	보충		1(2.9)		33(97.1%)	
		기본	1(10%)	1(3.0)	2(20%)	32(97.0%)	7(70%)
		심화	2(6.3%)	9(28.1)		21(65.5%)	
수 준 별 평 가 지	영어 성취 도	보충		2(5.9)		32(94.1%)	
		기본		3(9.1)	3(30%)	30(90.9%)	7(70%)
		심화	2(6.3%)	8(25.0)		22(68.7%)	

수준별 평가방식으로 제안된 두 가지 유형 중 어떠한 것을 더 선호하는 지에 대한 결과를 살펴보면, 보충반과 기본반의 학생들은 자신의 수준에 속한 문제만을 풀 수 있도록 하는 전면적 평가방식보다는 어떤 수준의 문항이든지 자신이 선택하여 풀 수 있도록 하는 부분적 수준별 선택 평가방식을 더 선호하는 것으로 나타났다. 반면에 심화반의 학생들은 전면적 수준별 평가방식을 더 선호하는 것으로 나타났다(표 10 참조). 이에 대한 이유를 분석한 결과, 상위권의 학생들은 ‘어차피 배점 높은 문항만 골라서 풀 것이기 때문에’, ‘선택의 고민을 하지 않아도 되니까’ 등의 이유로 전면적 수준별 평가방식을 선호하는 것으로 나타났다.

표 10

선호하는 수준별 평가방식

반	학생	
	부분적	전면적
보충	29(85.3%)	5(14.7%)
기본	23(69.7%)	10(30.3%)
심화	10(31.3%)	22(68.7%)

선호하는 평가방식에 대해서 교사들의 인식을 살펴보면, 대부분(10명 중 7명) 부분적 수준별 선택 방식을 더 선호하는 것으로 나타났다. 그 이유는 하위권

학생들에게 자신이 속해있는 수준의 문제보다 높은 수준의 문제를 풀었을 때의 성취감과 자신감을 향상시킬 수 있는 좋은 매체로 사용될 수 있으며, 상위권 학생들에게도 다양한 문제를 접하게 할 수 있다는 점에서 선택형 수준별 평가방식을 선호하는 것으로 나타났다. 학생과 교사의 인식을 종합해보면, 수준별 수업에 대한 적절한 수준별 평가의 요구가 교사와 학생들 모두에게 매우 크게 나타나고 있다는 것을 알 수 있다.

V. 결론 및 논의

수준별 수업은 학생들의 능력에 맞게 맞춤형 수업을 한다는 점에서 학습자 중심의 교육이라는 장점을 갖고 있지만 수준별 수업에 따른 평가의 부재에 대한 문제점은 끊임없이 제기되어 왔다(진경애 등, 2010). 이에 본 연구에서 수준별 수업의 실효성을 더할 수 있는 평가 방안을 제안하고자 하였다. 본 연구 결과에 따르면, 수준별 집단에 따라 평가문항이 차별화된 수준별 평가방식에 대해 교사·학생 모두 긍정적인 인식을 갖고 있는 것으로 나타났다. 하지만 상위권의 학생들의 경우는 기존의 공통평가지에 의한 동일한 평가방식에 대해서도 긍정적인 인식을 가지고 있는 것으로 나타났다. 이는 기존의 공통평가지가 상위권 학생들의 수준에 맞게 구성된 평가방식이라고 미루어볼 수 있다. 이에 본 연구는 수준별 수업의 활성화를 위한 노력으로 수준별 평가방안을 살펴 본 후, 가장 현실 적용이 용이하고 타당성이 높은 평가방안을 제안하고자 하였다.

이에 본 연구자들에 의해 제작된 두 가지 유형의 평가방식과 기존의 동일 평가방식을 비교한 결과, 일부 문항들을 수준별로 배점을 달리한 수준별 선택 평가방식과 학생의 수준에 맞는 문제만을 풀 수 있게 한 전면적 수준별 평가방식이 기존의 동일 평가방식보다 학생의 능력을 다소 높게 평가하는 것으로 나타났다. 특히 보충반 학생들의 점수는 수준별 평가방식에서 월등히 높은 점수를 보여주었다. 이는 기존의 공통문항만으로 구성된 동일 평가방식은 중·하위권 학생들의 영어 성취도를 타당하게 평가하는데 문제가 있었음을 보여주는 것이다. 또한 보충반 학생들에게는 선택형 수준별 평가방식보다는 전면적 수준별 평가방식에서 더 좋은 점수를 보여주었다. 반면에 기본반과 심화반의 경우, 두 가지 유형의 수준별 평가방식에서 의미있는 점수 차를 보여주지 않았지만 기존의 동일 평가방식보다는 수준별 평가방식에어 훨씬 높은 점수를 나타냈다. 이는 기본반과 심화반 학생들에게도 기존의 동일 평가방식보다는 수준별 평가방식이 보다 타당하게 학생의 성취도를 평가할 수 있다는 것을 시사한다.

수준별 평가방식에 대한 인식을 조사한 결과에 따르면, 대부분의 학생들과 교사들은 기존의 동일 평가방식에 대해 부정적 인식을 가지고 있으며 선택형

수준별 평가에 대해 가장 높은 선호도를 보였다. 여기서 특이한 점은 보충반과 기본반 학생들은 부분적 수준별 선택 평가방식을, 심화반 학생들은 전면적 수준별 평가방식을 선호하는 것으로 나타난 점이다. 이는 학생들의 수준에 따라 타당한 평가방식에 대한 추후 연구가 요구되며, 수준별 평가방식을 선택할 수 있는 체계적인 평가 시스템의 마련이 시급하다는 점을 시사한다.

본 연구의 제한점은 통계분석만을 통해 결과를 제안한 점으로 일반화된 결론을 이끌기에는 다소 무리가 있다. 이러한 이유로 후속 연구에서는 질적 자료를 수집하여 면밀히 검토한다면, 야기될 수 있는 정량적 연구의 문제를 어느 정도 해결할 수 있으리라 생각된다. 본 연구 결과를 토대로, 앞으로 수준별 교육과정을 보다 효율적으로 운영하기 위한 수준별 평가방식에 대한 연구가 지속적으로 이루어질 필요가 있겠다. 또한 수준별 평가를 시행하기에 앞서 수준별 평가에 대한 점수 산출 방식에 대한 연구 또한 이루어져 학생들에게 보다 공정한 방식의 평가가 이루어지기를 기대한다.

참고문헌

- 교육인적자원부. (2007). *초·중등학교 교육과정*. 서울: 교육인적자원부.
- 김명숙. (2003). 국가수준 학업성취도 평가연구의 주요 쟁점과 발전 방향. *교육평가연구*, 16(1), 241-264.
- 김미숙. (2008). *수월성 제고를 위한 수준별 학습 확대 방안*. 연구보고 RR 2008-33. 서울: 한국교육개발원
- 김신영, 백순근, 채선희. (1998). 국가 수준의 '성취기준 및 평가기준' 개발에 대한 고찰. *교육평가연구*, 11(1), 47-73.
- 김주훈, 김재춘, 박경미, 최진황, 진경애, 정구향, 강운선, 유준희. (1998). *제 7차 교육 과정 개정에 따른 영어과 수준별 교육 과정 적용 방안과 교수·학습 자료 개발 연구*. 연구개발 RDM 98-6-2. 서울: 한국교육과정평가원.
- 김재춘. (2009). *수준별 교육과정의 이해: 제 6차 및 제 7차 교육과정에서 수준별 수업 운영 방안*. 서울: 교육과학사.
- 김홍원, 조인진, 나장합, 차동춘, 김재춘. (2004). *수준별 이동수업 내실화 방안 연구*. 연구보고 RR 2004-17. 서울: 한국교육개발원.
- 박선화 (편). (2005). *수학·영어과 교육과정 개정 시안 및 수준별 수업 활성화 방안 공청회*. 연구보고 ORM 2005-41. 서울: 한국교육과정평가원.
- 배호순. (2001). *교육과정 평가론*. 서울: 교육과학사.
- 서영원. (2007). *수준별 수업을 위한 교과서 평가 모형에 관한 연구*. 석사학위논문, 영남대학교.

- 유세희. (2010). *영어과 수준별 평가 분석과 수준별 평가에 대한 인식조사*. 석사학위논문, 서울시립대학교
- 조난심, 김재춘, 허경철, 박순경. (1997). *제7차 교육과정 개정에 따른 수준별 교육과정 편성 및 운영방안에 관한 연구*. 연구보고, CR 97-34. 서울: 한국교육개발원.
- 진경애, 송미영, 김화경. (2009). *수준별 수업에 따른 학생 평가 방안 연구—수학, 영어 교과를 중심으로*. 연구보고 PRE 2009-5. 서울: 한국교육과정 평가원.
- 진경애, 송미영, 반재천. (2010). *영어·수학 수준별 수업에 따른 학생평가 방안*. *교육과정평가연구*, 13(2), 335-358.
- 최용익. (2009). *수준별 선택 세트형 문항을 활용한 평가와 전통적 평가의 성취도 비교: 고등학교 집합과 명제를 중심으로*. 석사학위논문, 충북대학교.
- Brown, H. D. (2004). *Language assessment: Principles and classroom practice*. New York: Pearson Longman.
- Oakes, J., Gamoran, A., & Page, R. (1992). Curriculum differentiation: Opportunities, outcomes, and meanings. In P. W. Jackson. (Ed.), *Handbook of research on curriculum* (pp. 570-608). New York: Macmillan.

이경인

단국대학교 교육대학원

448-701 경기도 용인시 수지구 죽전동 126번지

Tel: 010-9227-6053

Email: chil616@nate.com

김현주

단국대학교 영어영문학과

448-701 경기도 용인시 수지구 죽전동 126번지

Tel: (031) 8005-3069

Email: hyunjukim@dankook.ac.kr

Received 15 July 2011

Revised 24 August 2011

Accepted 4 September 2011