

중학교 국어과 기초학력 진단평가에 관한 연구

윤여탁* · 민현식** · 김정우***

< 차 례 >

- I. 머리말
- II. 기초학력의 개념과 기준 설정
- III. 국어과 진단평가 설계의 기본 방향
- IV. 중학교 국어과 기초학력 진단평가 설계
- V. 맺음말

I. 머리말

1. 연구의 필요성

근대 이후의 주된 특징이라고 할 만한 공교육 제도는 ‘국가 혹은 준국가적 자치 조직의 통제와 관리와 지원에 의하여 국민 전체를 대상으로 하여 운영되는 교육’(서울대학교 교육연구소, 1998 : 138)을 지향한다.¹⁾ 이러한 제

* 서울대학교 교수(교신저자)

** 서울대학교 교수(공동저자)

*** 영남대학교 교수(주저자)

1) 공교육은 법적 해석, 학력 인정 여부, 정부 규제 여부, 이윤 추구 여부, 정부의 재정 지원 여부 등의 조건을 어떻게 보느냐에 따라 그 범위가 달라질 수 있다. 여기에서는 김영철(2000)에 따라 공교육을 교육관련 법령에 의해 교육과정, 교육수준 및 자격이 제도화되어 있고, 이와 같은 기준에 의해 정부의 통제를 받으며, 과정을 이수한 자의 학력이

도 하에서의 국가 교육은 국민들이 교육을 통해 일정한 수준에 도달할 수 있도록 최선을 다해야 하는 책무를 가지고 있다. 이러한 책무를 다하기 위해 국가는 교육이 보다 내실 있고 효과적으로 이루어질 수 있도록 다양한 정책을 개발하고 각 교육 실천 단위에서 필요로 하는 사안에 대해 충분한 지원을 해 나가야 한다.

목표 수준에 다수의 학습자들이 도달할 수 있게 하기 위하여 국가는 학습자의 상황을 정확히 진단하고, 그에 근거하여 학습자에게 맞는 교육을 효과적으로 설계하여 시행하는 일이 원활히 이루어질 수 있도록 지원해야 한다. 이는 교육의 성패를 결정짓는 매우 기초적인 부분으로, 국가에서 특히 많은 관심을 기울여야 할 부분이다. 학습자에 대한 진단이 정확히 이루어지지 않을 경우, 그에 맞는 교육이 효과적으로 이루어지기 어렵고, 그 결과 국가가 설정한 목표 수준에 도달하는 데에 다수의 학습자들이 어려움을 겪을 가능성이 크기 때문이다. 그러므로 학습자의 상태를 다양한 방법으로 정확하게 진단하고, 그 진단 결과에 따른 보정 교육이 충실히 이루어질 수 있도록 국가는 지원을 아끼지 않아야 한다.²⁾

우리나라에서 이러한 진단평가를 국가 단위에서 시행하고 그에 근거한 보정 교육의 프로그램을 제공한 것은 그리 오래되지 않았다. 대표적인 것으로 ‘초등학교 3학년’을 대상으로 하는 ‘국가수준 기초학력 진단평가’를 들 수 있다. 한국교육과정평가원(이하 평가원)은 지난 2002년부터 초등학교 3학년을 대상(표집)으로 하여 모든 학습의 기초가 되는 ‘읽기’, ‘쓰기’, ‘기초수학’ 세 분야에 대해 ‘국가수준 기초학력 진단평가’를 매년 9월 말~10월 초에 시행해 왔다. 평가 결과는 각 학습자들에게 개별 통보되는

인정되는 교육으로 본다. 그리고 공교육 담당 기관은 이윤보다 일반 이익을 추구하는 공공 기관이고, 공교육 운영을 위해 정부가 재원을 부담하는 것을 원칙으로 한다고 본다. 이렇게 볼 때, 논란의 여지가 있지만, 사립 중·고등학교는 상당 부분 공교육의 범주 안에 있는 것으로 볼 수 있고, 공교육의 반대 개념인 사교육은 주로 사설 학원을 뜻하는 개념으로 쓰이게 된다.

- 2) 국가가 이러한 보정 교육에 힘써야 한다는 것은 국민의 교육권 보장의 문제와 관련된다. 국민의 교육권이 단지 ‘취학’ 그 자체만을 의미한다고 보는 소극적 법 해석과 달리 교육권은 의미 있는 결과에 도달하는 수준을 보장받는 것이라는 적극적 해석이 헌법학과 교육학 쪽에서 다양하게 논의되고 있다(김수천, 2005 : 164).

데, 정해진 기준(기초학력)에 도달하였는지 여부와 각 내용 영역별로 도달과 미도달의 의미를 해석하여 현재 상황을 진단할 수 있는 정보를 제공하고 있다. 또한 평가 결과에 대해 영역별, 성별, 지역별 차이와 연도별 변화 추이를 분석하고, 그 결과와 상관관계가 있는 배경 변인들을 분석하며, 나아가 진단 결과의 활용 방안에 대한 연구까지 수행함으로써 각 학교에서 기초학력 증진을 위한 보정 교육을 실시할 수 있도록 지원해 왔다.³⁾

이에 비해 중등 국어과 교육의 기초학력⁴⁾에 대해서는 본격적인 국가 단위의 진단평가가 체계적이고 지속적으로 개발 운영되지 못했다. 2000년에 평가원에서 중등 국어과 교육의 기초학력 측정을 위한 평가를 연구하고 시행한 결과를 분석⁵⁾한 적이 있으나, 이후에 지속적으로 시행되지는 못하였다. 2000년에 시행된 이 기초학력 평가는 전국 학생들의 약 0.5% (중학교 88개교, 고등학교 92개교)를 표집하여 학습자들의 학력을 ‘듣기와 말하기, 읽기, 쓰기, 문학’의 네 영역에 걸쳐 평가하였다. 실제 문장 차원의 쓰기를 포함하는 수행평가형 지필 평가를 30% 정도 포함하는 평가였으며, 평가 결과를 해석하기 위한 기준점을 설정하여 우수학력 수준, 보통학력 수준, 기초학력 수준, 기초학력 미달 수준의 네 집단으로 수준을 설정하는 등 비교적 체계적이고 실제 학습자들의 능력을 타당하게 진단할 수 있는 기본 틀을 갖추었다. 그러나 기초학력을 진단하는 평가임에도 불구하고 전체적인 체제는 학업성취도를 평가하는 것과 유사한 부분이 있었고, 이후에 계속 시행되지 못하였다는 한계를 가진다.

중등교육에서 국가 단위의 진단평가가 제대로 이루어지지 않은 원인은 일차적으로 우리나라 사람들이 교육의 출발점이나 과정보다는 결과, 그것도 상대적으로 서열화한 결과에서 상위를 차지하는 데에만 관심을 기울여 왔다는 데에서 찾을 수 있을 것이다. 그렇지만 그러한 분위기나 문

3) 자세한 것은 남명호 외(2007) 및 한국교육과정평가원 교수학습개발센터(<http://classroom.re.kr/view.jsp?mcode=101410>)의 ‘기초학력·특별보충 프로그램’ 참고.

4) ‘기초학력’의 개념은 다양하게 쓰일 수 있다. 이에 대해서는 2장에서 자세히 살펴보기로 한다.

5) 자세한 것은 백순근·최미숙·유동엽·김광조(2000), 『중·고등학교 학생들의 국어 기초학력 평가 연구』 참고.

화 탓만 하면서 교육에서 매우 중요한 부분인 진단평가에 관한 연구나 그 실천을 소홀히 할 수는 없다. 부족한 부분을 하나씩 채워가고 새롭게 설계해 가면서 체계적인 진단평가 체제를 구축할 필요가 있다.

2. 국가 단위 진단평가 시행에 대한 우려와 바람직한 방향

국가 단위에서 진단평가를 실시하는 데 대해 비판적인 견해들이 많다. 국가 단위에서 대규모로 이루어지는 진단평가는 개별 학습자의 상황을 정확하고 세밀하게 측정하는 데에 아무래도 한계가 있을 수밖에 없으므로 그 진단의 실효성과 타당성에 의문을 가질 수 있다. 이와 함께 진단평가가 표집 방식이 아닌 전집 방식으로 시행될 경우, 그 결과를 바탕으로 각 학습자의 상황을 전국적으로 상대 서열화하는 데서 오는 폐해도 생각해 볼 수 있다. 이러한 점들을 충분히 고려하면서 국가 단위의 진단평가의 필요성을 점진해 보도록 한다.

대규모 국가 단위의 검사가 개별 학습자의 상태에 대해 풍부하고 충실한 정보를 제공하기 어려운 것은 분명히 사실이다. 그렇지만 현재 학교 현장에서는 계획적이고 내실 있는 진단평가를 아예 실시하지 않는 경우가 적지 않다는 점도 간과하기는 어렵다. 진단평가가 제대로 이루어지지 않는 이유는 진단평가 도구 개발 방법, 평가 결과를 수업에 반영하고 이용하는 방법 등에 대해 교사 양성 기관에서 그에 대한 충분한 교육이 이루어지지 않았고, 그에 관한 정보나 교사 재교육의 기회 등이 교사에게 충분히 제공되지 않고 있기 때문이다. 교사에게 그러한 능력이 갖추어져 있다고 하더라도 개별 학습자들을 위한 진단평가 도구를 개발하고 평가를 시행하기에는 교사 1인당 학생 수나 수업 외 업무량의 과다 등 여러 제약이 존재한다. 이러한 현실 속에서 국가 단위의 진단평가는, 비록 그 효과 면에서 충분하지는 않더라도 학습자의 기초학력에 대한 기본 정보를 제공하는 순기능을 담당할 수 있다. 각 교실에서 교수학습을 계획, 운영할 때 학습자들의 다양성을 염두에 둔 교육, 적어도 일정 기준에 미달하는 학생

들을 위한 보정 교육이 별도로 이루어질 가능성이 높아질 것으로 기대할 수 있는 것이다.⁶⁾ 아울러 국가에서 전문가 집단을 구성하여 개발한 진단 평가 도구라면 그 평가도구가 일정한 수준을 갖추고 있을 것으로 기대할 수 있다. 그러므로 각 교실에서 진단평가를 스스로 계획하는 교사들에게 의미 있는 예시 혹은 참고 자료가 되어줄 것이다.

한편 진단평가의 결과를 바탕으로 전국의 모든 학생을 한 줄의 서열에 자리매김하는 일이 생기지 않을까 하는 우려가 크다. 실제로 ‘우리 아이가 전국 몇 등이나 되나?’를 궁금해 하는 학부모의 알 권리를 충족시키고, ‘전국에서 몇 번째 이하인 학교에 우리 아이를 보내고 싶지 않다’는 학부모의 학교 선택권을 보장하기 위해 이러한 서열화의 정보를 당연히 국가가 제공해야 한다는 견해도 만만치 않게 제시된다.

그러나 우리 교육과정의 지향하는 교육은 교과별로 일정한 성취기준을 정하고, 그것에 도달하는 것을 목표로 하는 교육이다. 제7차 교육과정 해설서(교육인적자원부, 1999a : 171)는 “평가는 모든 학생들이 교육 목표를 성공적으로 달성하기 위한 교육의 과정(process)으로 실시하는 것”으로 “교육목표 성취도를 평가하여 학생의 목표 도달도를 확인”하기 위한 것이며, 이를 위해 “평가는 등급을 매기고 서열을 정하기 위한 활동이라는 지금까지의 인식에서, 궁극적으로 도달해야 할 목표에의 달성 여부를 확인하기 위한 활동이라는 개념으로 바뀌어야 한다.”고 분명히 밝히고 있다.⁷⁾ 그러므로 기초학력 진단평가 역시 상대적 서열 파악에 초점을 맞추는 규준참조평가가 아니라 성취기준 도달 여부에 초점을 맞추는 준거참조평가로서의 성격을 분명히 할 필요가 있다.

또한 학업성취도 평가의 경우이긴 하지만 교육과정해설서는 국가 단위의 평가의 시행에 대해 “학생의 학업성취도 평가는 시·도 또는 지역

6) 이때의 보정 교육은 흔히 말하는 ‘우열반’과는 그 목적과 취지가 다르다. 우열반은 결국 소수의 ‘우반’의 학습 능력 극대화를 위해 보통의 많은 학생들을 분리하는 것이다. 반면 여기서의 보정 교육은 소수의 기초학력 미도달 학생들이 정상적인 학업 과정에 참여할 수 있게 하기 위한 특별 프로그램이다.

7) 이 부분은 7차 교육과정뿐만 아니라 2007 개정 교육과정 해설서에서도 동일하게 반복되고 있다(교육과학기술부, 2008 : 183).

교육청별로 일정 수의 학교를 표집하여 몇 개의 학년과 몇 개의 교과를 대상으로 학생들의 학업성취도를 평가한다.”(교육과학기술부, 2008 : 179)고 하여 ‘전집’이 아닌 ‘표집’으로 시행할 것을 분명히 하고 있다. 교육과정에서 규정하고 있는 학업성취도 평가의 이러한 기본 방향을 굳이 진단평가에서만 전집 서열 평가로 바꿀 이유가 명확히 없는 한, 국가 단위에서 이루어지는 진단평가 역시 표집 형태로 시행하는 것이 바람직할 것이다. 문항과 평가내용 및 진단 정보 등은 모두 공개되므로 표집 대상이 아닌 학교나 학생이 아니라고 하더라도 평가 문항을 내려 받아 자체적으로 진단 평가를 시행하는 것은 얼마든지 가능하다. 설사 전집 평가를 한다고 하더라도 불필요한 논란을 불러일으키는 서열 평가가 아니라 학습자 개인이 해당 학년의 학습을 수행하는 데 필요한 기초 능력을 가지고 있는지 여부를 측정함으로써 학습자와 교사, 그리고 학부모에게 필요한 정보를 제공하는 평가로 그 성격을 명확히 하는 것이 필요하다.

이 연구는 이상에서 살핀 점들을 고려하면서 중학생을 대상으로 한 국가 수준의 교과학습 진단평가 도구 개발의 기초 작업을 수행하고자 한다. 기초학력의 개념과 도달 기준 설정 방법을 살펴보고, 국가 단위에서의 국어과 진단도구 개발의 바람직한 방향을 모색해 볼 것이다.

II. 기초학력의 개념과 기준 설정

진단평가는 그 목적에 따라 매우 다양하게 개발될 수 있다. 이 글에서 상정하고 있는 진단평가는 개별 학습자의 상태를 정교하게 분석하는 맞춤형 평가라기보다는 국가 단위에서 전국적으로 분포하는 다양한 학습자들의 상태를 일괄적으로 진단하는 평가이다. 평가 결과로 제공되는 정보는 개별 교육 실천의 단위에서 볼 때 다소 불충분할 수 있겠으나, 이 평가는 국가가 개발한 교육과정을 정상적으로 이수해 가는 데 필요한 최소한의 수준, 즉 기초학력을 갖추고 있는지, 각 단계의 출발점에서 학습자가 어떤

수준에 도달해 있는지의 여부를 가려내는 것을 주 목적으로 삼는다.

1. 기초학력의 개념

기초학력의 개념은 ‘기초’를 무엇으로 보는가에 따라 달라진다.⁸⁾

첫째, 모든 교과 내용을 익히기 위해 필요한 기초적·도구적 기능을 수행하는 학력을 뜻한다. 주로 언어와 수에 관한 능력을 뜻하는데, 이때 언어 능력은 듣기·말하기는 제외하고 읽기·쓰기만으로 한정하기도 한다. 모든 교과 학습에 기초가 되는 것이므로 특별히 교과의 구별이 필요하지는 않지만, 그러한 학력을 기르는 교과는 국어과와 수학과가 될 것이다.⁹⁾

둘째, 각 교과별로 기저가 되는 학력을 의미한다. 각 교과별로 그 교과의 토대가 되는 기초적 지식이나 기능을 설정하고 그에 근거하여 기초학력을 판단하는 것이다. 이 관점에서 보면 학교에서 배우는 모든 교과에는 각 교과별로 기초학력이 존재한다고 볼 수 있다. 국어과의 경우에도 물론 이러한 관점에서의 기초학력을 설정할 수 있겠으나, 그 경우 기초학력의 수준을 어느 정도로 잡을 것인가에 대해서는 국어교육관에 따라 매우 다양한 견해가 있을 수 있을 것이다.

셋째, 위계적 관점에서 학력을 파악할 때, 기초학력이란 특정 단계의 학력이 다음 단계 학습의 기초가 된다는 점을 부각시키는 것이라고 할 수 있다. 1학년의 학습을 통해 형성한 학력은 2학년 학습의 기초가 되고, 초등학교의 학습을 통해 형성한 학력은 중학교 학습의 기초가 된다고 보는 것이다. 이렇게 볼 때 중학교 1학년 학습자의 기초학력은 초등학교 6학년에서의 성취 정도와 밀접한 관련이 있다고 볼 수 있다.

넷째, 기초학력을 국민이 사회생활을 영위하는 데 꼭 필요한 최소한의

8) 기초학력의 개념에 대한 내용은 주로 남명호 외(2007 : 3-6)를 참고하였다.

9) 이렇게 ‘기초학력’을 단순한 수준의 읽기와 쓰기, 셈하기로 한정하는 입장에서는 기초학력과 다른 기본학력의 개념을 설정하여 사용하기도 한다. 기초학력이 비교적 저학년에서 습득하게 되는 것인 데 비해, 기본학력은 매 학년 또는 매 학교급에서 최소한으로 성취해야 할 기준이라는 의미로 쓰인다(현주, 2003 : 29).

능력으로 보고, 의무 교육 단계에서 형성되는 학력 전체를 기초학력으로 보는 관점이다. 학습자들이 의무 교육 기간에 배운 모든 교과와 학습 내용이 기초학력의 범주에 든다.

다섯째, 국민적 교양의 최소 성취기준을 설정하여 이를 기초학력을 보는 경우이다. 다소 상대적이긴 하지만, 예를 들어 고등학교 진학률이 99.5%에 달하는 상황에서는 고등학교 교육 종료 단계에서의 학력 수준을 현대 국민의 기초학력으로 보는 것이다.

이상의 기초학력에 대한 관점들 가운데 중학교 1, 2, 3학년 학생들을 대상으로 실시하는 진단평가로는 주로 세 번째 관점, 즉 특정 단계의 학력이 다음 단계 학습의 기초가 된다는 관점을 취하는 것이 적절하다. 이러한 관점을 취함으로써 매 학년별로 해당 학년의 학업을 이수하는 데 요구되는 기본적 능력의 수준과 범위를 정하고, 직전 학년의 학업성취도를 참고하면서 해당 학년의 학업을 이수할 수 있는 소양을 갖추고 있는지 진단하는 평가를 설계할 수 있다.

다만 이러한 관점을 취할 때에는 진단평가가 직전 학년의 학업성취도 평가와 동일한 성격의 평가로 인식될 가능성이 있다. 그러나 상급 학년에서 도달해야 하는 성취기준 및 그에 근거하여 설정하는 학습 목표를 달성하기 위해 필요한 기초학력이 이전 학년에서 모두 다루어진다고 볼 수는 없다. 중학교 2학년에서 배우는 학습 내용 가운데 1학년에서의 학습 내용과 관련성이 적거나 새로운 것이 있을 수 있기 때문이다. 그러므로 이전 학년에서 배운 학습 내용을 확인하는 것으로 진단평가를 한정할 수는 없으며, 해당 학년에서 배우게 될 내용과 관련되는 기초적 소양을 확인하는 평가로서의 성격을 분명히 할 필요가 있다.

한 가지 더 짚어야 할 점은 이것이 어디까지나 ‘기초적’ 소양에 대한 평가이어야 한다는 점이다. 진단평가가 시행되면 사교육의 이른바 ‘선행 학습’이 더욱 기승을 부릴 것이라는 우려의 목소리가 있는데, 진단평가가 말 그대로 해당 학년의 학습을 해 나가는 데 필요한 기초적 소양을 확인하는 데 충실하다면 ‘선행학습’은 이 평가에 큰 의미를 가지지 못할 것이다.¹⁰⁾ 물론 중학교 1, 2, 3학년의 교육과정을 이수하는 데 필요한 ‘기초적

소양'이 무엇인지에 대한 연구는 부족한 형편이다. 그렇지만 교사와 해당 분야의 전문가들의 협의를 거친다면 어느 정도의 기준선을 설정하는 것이 불가능한 것은 아니다. 이 협의와 연구 결과에 근거하여 평가 도구를 개발하고, 그것이 지속적으로 검증을 받고 다시 수정되는 과정을 거치면서 보다 더 과학적이고 타당하며 세밀한 기초학력의 내용과 기준이 마련될 수 있을 것이다.

정리해 보면 중학교 진단평가에서의 기초학력은 ‘해당 학년에서 배울 내용을 원활하게 학습하는 데 필요한 기초적 능력’으로 설정하는 것이 적절하며, 이는 이전 학년에 도달한 학력을 기반으로 하되, 거기에 해당 학년에서 새로 배우게 될 내용에 관계되는 부분이 결합되는 형태로 구성된다. 초점은 이전 학년 학습 내용에 대한 성취도가 아니라 해당 학년에서 배워할 내용과 관련되는 최소한의 학력의 보유 여부이다.

2. 기초학력의 도달 기준 설정

앞에서 해당 학년에서 배워야 할 내용, 또는 해당 학년에서 도달해야 하는 성취기준의 측면에서 기초학력의 개념을 설정하였다. 이와 함께 기초학력 진단평가에서는 기초학력의 도달 기준 설정이 중요한 과제이다. 즉 기초학력의 도달과 미도달의 경계점을 어떻게 결정할 것인가의 문제이다.

도달 기준을 설정하는 데에 여러 방법이 있으나, 1996년 루이스·미첼·그린에 의해 개발되어 현재까지 사용되고 있는 북마크 방법이 가장 보편적으로 사용되고 있는 도달 기준선 설정 방법으로 알려져 있다(이규민, 2003). 북마크 기법은 전문가들을 패널로 설정하고, 도달 기준선에 대한 패널들의 기대치를 반영하여 도달 기준선을 설정하는 방법이다. 곤란도를 기준으로 쉬운 문항에서부터 어려운 문항으로 순서화되어 있는 문항 책자

10) 유사한 우려 속에서 2002년부터 시행되었던 초등학교 3학년 기초학력 진단평가는 취지와 목적을 크게 벗어나지 않았으므로써 다행(!)하게도 큰 부담이 없는 시험으로 정착되어 가고 있다.

를 놓고, 각 패널들이 평가 대상 학생이 최소한의 능력에서 답을 맞힐 수 있는 수준을 판단하고, 그에 대해 토의하면서 여러 단계의 의사결정 과정을 거쳐 도달 기준선을 설정한다. 그 절차를 간단히 제시하면 아래와 같다.

- 1단계 : 곤란도별로 쉬운 문항부터 어려운 문항으로 구성된 ‘순서화된 문제집’을 마련한다.
- 2단계 : 패널 전체를 기본 단위의 소그룹으로 나눈 후, 각 소그룹별로 학생들이 문항을 맞히기 위해 갖추어야 할 지식과 기능이 무엇인지 토의한다.
- 3단계 : 토의 후, 각 패널로 하여금 기초 능력 수준 이상 학생 중 최저선, 다시 말해서 도달 기준선에 있는 학습자가 반드시 맞혀야 하는 문항을 곤란도 순으로 배열된 순서화된 문제집에서 고르게 한다. 예를 들어 50개의 문항 중 1~20번까지는 맞혀야 한다고 생각하면, 20번 문항을 선택하는 것이다. 그 문항의 척도점수가 그 패널이 기대하는 첫 번째 분할 점수이다.
- 4단계 : 소그룹 내에서 각 패널이 선택한 문항을 가지고 다시 토의한다. 이견을 조정하고, 합의된 분할 점수에 이르는 방향으로 두 번째로 ‘순서화된 문제집’에 자신의 선택을 표시한다. 이것이 그 패널의 두 번째 분할 점수이다.
- 5단계 : 두 개의 소그룹을 하나로 묶어 중그룹을 만든 뒤, 중그룹 회의를 거쳐 다시 순서화된 문제집에 각 패널이 의견을 표시한다.
- 6단계 : 마지막으로 전체 패널이 모여 토의를 하고, 각 패널은 전체의 의견을 청취하고 자신의 견해도 제시하면서 마지막으로 자신의 선택을 결정하게 된다.
- 7단계 : 최종적으로 표시된 문항의 척도 점수를 수합하고, 패널들이 지정한 문항들의 척도 점수의 중앙값 또는 평균을 계산하여 도달 기준선을 설정한다.
- 8단계 : 평가 결과 도달 기준선 위에 있는 사람이 수행할 수 있는 지식이나 기능을 열거하여 수행지표 목록을 작성한다. 수행지표 목록을 통해 학습자의 학력에 대한 구체적 정보를 확인할 수 있다.

이 북마크 방법의 장점은 관련 전문가들의 합의를 통해 기초학력 도

달 기준선을 설정할 수 있다는 것이다. 그리고 그러한 설정 과정이 막연히 감에 의해서 이루어지거나, 특정한 시각에 편중되어 결정되는 것이 아니라 학생들이 치른 시험 문항과 분석 결과를 활용하여 이루어진다는 점에서 객관성을 확보할 수 있다는 장점이 있다. 또한 교과전문가가 평가 이론에 관한 전문적 소양이 없더라도 교과에 대한 전문성만 갖추고 있다면 도달 기준선을 설정하는 데 큰 어려움 없이 참여할 수 있다는 점도 이 북마크 방법의 장점이라고 할 수 있다.

III. 국어과 진단평가 설계의 기본 방향

국가 단위에서 시행하는 진단평가는 국가가 고시한 교육과정에 근거한 교육이 원활하게 이루어질 수 있도록 지원하는 데 그 목적이 있다. 그러므로 근본적으로 국가 단위 교육과정에 근거한 평가를 설계하는 것으로 방향을 설정해야 한다.¹¹⁾ 교육과정에 제시된 학년별 내용들을 재구성하여 그러한 내용들을 충실히 학습해 가는 데 요구되는 기초학력을 설정하여 평가의 내용으로 삼을 필요가 있다. 그리고 그 결과를 바탕으로 국가 단위의 평가라는 특성 혹은 한계를 감안한 질 높은 평가도구를 개발하여야 한다. 이러한 진단평가 설계 과정을 하위 단계로 나누어 살펴보기로 한다.

11) 이처럼 현행 교육과정에 근거하여 그 내용을 바탕으로 평가를 설계하는 것에 대해 부정적인 평가도 존재한다(염은열, 2005). 부정적 평가의 주요 내용은 그것이 교육과정에 이미 제시되어 있는 것들을 반복하는 차원에 그친다는 점, 학습자의 특성, 교육 현장의 특수성, 우리 고유의 언어문화를 고려하는 등의 보다 기초적인 연구가 되기 어렵다는 점 등이다. 이러한 지적은 충분히 타당하며, 교육과정의 한계를 넘어서는 연구는 그 방향으로 지속적으로 이루어져야 한다고 본다. 그렇지만 국가 단위에서 개발한 교육과정의 시행 및 정착의 정도를 평가하면서 학습자들의 교육과정 이수율 돕는 방향의 연구 역시 지속적으로 이루어져야 한다.

1. 국어과 교육과정에 근거한 평가 내용 설계

이 연구는 중학생들의 국어과 교과학습 수행에 필요한 진단을 목적으로 출발하였고, 일차적으로 2008년을 기준으로 진행한 연구이므로 설계의 근간이 되는 교육과정은 제7차 교육과정이다. 물론 2007년에 교육과정이 개정 고시되었고, 그 교육과정을 전혀 고려하지 않은 것은 아니지만, 2007 개정 교육과정에 따른 교과서가 개발되어 본격적으로 시행되는 시기가 중학교 1학년 2010년, 2학년 2011년, 3학년 2012년이기 때문에 우선은 제7차 교육과정에 근거하여 평가를 설계하였다. 2007 개정 교육과정에 근거한 진단평가 연구 역시 이루어져야 하고, 향후 좀 더 많은 연구 인력이 참여해서 교육과정 개발 및 운영에 기반이 되어줄 수 있는 근본적 차원¹²⁾의 진단평가 도구 개발이 이루어져야 할 것이다.

7차 교육과정의 주요 특징은 국민 공통 기본 교육과정과 선택 중심 교육과정의 이원 체제로 되어 있다는 점과, 국어과의 경우 심화·보충형 수준별 교육과정을 운영하도록 설계되었다는 점이다. 두 가지 특징은 진단평가를 설계하는 데 기본적으로 고려해야 할 사항이다. 조금 더 구체적으로 살펴보기로 한다.

우선 국어과 진단평가 설계 시, 교육과정에 제시된 이전 학년의 학년별 내용을 어느 정도까지 반영할 것인가의 문제를 명확히 할 필요가 있다. 국민 공통 기본 교육과정에 해당하는 1학년에서 10학년의 경우, 초등학교, 중학교, 고등학교의 학교급별 목표가 총론 차원에서 다르게 설정되어 있기는 하지만, 국어과의 경우 학년별, 학교급별 목표를 따로 설정하지는 않고 있다. 1~10학년까지의 국어과 교육의 목표는 동일하게 설정되어 있으

12) 이 근본적 차원의 연구는 무엇보다도 지금 이 시대에 요구되는 ‘국어 능력’은 과연 무엇이고, 그것을 구성하는 요소들은 어떤 것이 있는지에 대한 다각도의 접근을 요구한다. 국가경쟁력을 높이는 차원에서의 국어 능력을 파악하고, 지식-기능의 균형을 통해 그것을 실현하고자 하는 구상(민현식, 2005)이나, ‘언어문화능력’을 상정하고, 그 구성 요소로 ‘비평적 이해 능력, 소통적 교섭 능력, 창조적 실천 능력’을 구안하는 연구(최인자, 2005) 등은 국어 능력 개념에 대한 새로운 이해를 가능하게 해 준다.

며, 초등학교와 중학교의 학년별 내용이 동일한 목표 하에 위계화의 원리를 바탕으로 구성되어 있다. 그렇기 때문에 매 학년의 기초학력 진단평가는 직전 학년의 학년별 내용을 충분히 고려할 필요가 있다. 중학교 1학년 기초학력 진단평가를 계획한다면, 초등학교 6학년에 설정되어 있는 학년별 내용이 반영되어야 한다.

그렇지만 앞에서 기초학력의 개념에 대해 살펴본 바와 같이 해당 학년의 학습을 위한 진단평가는 이전 학년의 학업성취도 평가와 동일한 것은 아니다. 직전 학년의 내용을 반영하되, 그것은 새로운 학년의 학습 내용을 이수하는 데 관련이 있는 내용들을 중심으로 선별되고 조직되는 것이 바람직하다. 여기서 새로운 학년의 내용과 관련이 있다는 것이 곧 선행학습을 통해 미리 그 내용을 학습한 정도의 상태를 의미하는 것은 아니며 그러한 내용을 원활하게 배워 가는 데 필요한 기초적 소양을 뜻한다.

한편 심화·보충형 수준별 교육과정을 운영하기 위해 교육과정의 학년별 내용마다 모두 [기본]과 [심화]로 수준별 학습 활동의 예를 마련한 것도 기초학력 평가를 설계하는 데 고려할 점이다. 해설서(교육인적자원부, 1999b)는 교육과정에 예로 제시된 [기본]만 충실히 이수하면 해당 교육 내용과 관련된 목표에 도달할 수 있게 구안했음을 밝히고 있다. 즉 [심화]가 아니라 [기본]이 모든 학습자들이 도달해야 하는 지점이다. 또한 하급 학년의 [심화]라 하더라도, 그것이 곧 상급 학년의 [기본]과 중복되거나 단순 반복되는 것이 아니어야 한다. 다시 말하면 하급 학년의 [심화]라 하더라도, 상급 학년의 [기본]과 같거나 그보다 높은 수준의 것일 수는 없다. 그러므로 해당 학년의 학습을 염두에 두고 기초학력을 측정한다고 할 때, 그것은 주로 해당 학년에서 [기본]에 도달할 수 있는 기초학력을 가지고 있는지를 측정하는 것이어야 한다. 또 이전 학년의 도달 수준을 확인하고자 할 때에도 주로 [기본]에 제시된 바를 충실히 성취하였는가에 중점을 두고 설계를 해야 한다.

종합하면, 기초학력 진단평가의 설계는 시행되고 있는 교육과정에 근거하여 설계할 필요가 있다. 이 연구에서는 주로 제7차 교육과정을 근거로 삼았으며, 그에 따라 수준별 교육과정을 염두에 두고 설정된 교육과정의 학습 활동의 예에서 주로 [기본] 수준에 초점을 맞추는 것으로 방향을 설정하였다.

2. 교육과정의 내용 영역의 재구성과 행동 영역 구성

제7차 국어과 교육과정은 ‘듣기, 말하기, 읽기, 쓰기, 국어지식, 문학’의 여섯 영역으로 그 내용이 이루어져 있고, 이는 다시 ‘본질, 원리, 태도, 실제’의 네 범주로 체계화되어 있다. 개정된 2007 교육과정의 경우에는 내용 영역은 여섯 영역에서 ‘국어지식 → 문법’으로의 명칭 변화가 있지만 큰 차이는 없었던 반면, 범주에서는 ‘실제, 지식, 기능, 맥락’으로 다소 큰 변화를 보였다. 학년별 내용은 제7차 교육과정은 주로 영역별로 중요하다고 판단되는 지식이나 기능을 구성하는 요소 중심으로 내용이 설계된 반면, 2007 개정 교육과정은 텍스트를 중심으로 하여 그 수용과 생산에 요구되는 지식과 기능이 결합된 성취기준의 형식으로 내용이 설계되어 있다.¹³⁾

교육과정을 근거로 하여 진단평가의 내용 영역을 구성한다고 할 때, 교육과정에 제시된 여섯 영역의 내용 영역을 벗어나는 설계를 하기가 쉽지 않아 보인다. 이미 5차 이후로 큰 변화 없이 이어져 온 여섯 영역 체계의 국어과 교육 내용 체계는 그 논리상의 문제에도 불구하고 어느 정도 안정된 현실적 힘을 가지고 있다는 점은 부인하기 어렵다. 중학교의 경우 7차 교육과정의 교과서는 명확하게 그 영역별 체제를 가지고 분권까지 된 상태이기 때문에 진단평가에서도 그 영역별로 학습자들의 능력을 진단하여 교실로 송환해 주는 것이 요구된다고 하겠다.

그러한 까닭에 기본적으로 기초학력 진단평가 역시 여섯 영역 체제를 평가의 내용 영역 설계에 반영하는 것이 필요하다. 다만 대규모로 이루어지는 평가라는 특성상 말하기나 쓰기 평가의 경우에는 평가하려는 내용을 간접적으로 측정할 수밖에 없다. 또한 듣기와 말하기의 긴밀한 연관성을 감안하여 말하기 영역은 듣기의 평가에 간접적으로나마 반영될 수 있도록 하여 두 영역을 하나로 묶는 것이 현실적인 방법이다. 그러므로 현재 중

13) 이러한 점에서 개정 국어과 교육과정은 ‘텍스트 수용·생산 중심 교육과정’, ‘성취기준형 교육과정’이라고 할 수 있다. 이재기(2007), 김정우(2007) 등 참고.

학생을 대상으로 한 기초학력 진단평가는 ‘듣기·말하기, 읽기, 쓰기, 국어지식, 문학’의 다섯 영역으로 하는 것이 적절하다.

평가의 큰 영역을 결정하였으면, 각 영역별로 평가해야 할 주요 내용을 선정하는 것이 그 다음 단계의 과제이다. 앞서 살펴본 바와 같이 각 학년 학습의 기초학력을 진단하는 평가는 해당 학년에서 배울 내용을 충분히 받아들일 수 있는 학력을 갖추고 있는지를 확인하는 데 초점을 맞추는 것이 바람직하다. 7학년 듣기의 경우 해당 학년에서 배워할 내용으로 제시된 것은 ‘(1) 듣기와 읽기의 공통점과 차이점을 안다, (2) 사실과 의견 구별하며 듣는다, (3) 내용의 짜임을 파악하며 듣는다, (4) 내용의 통일성을 평가하며 듣는다, (5) 메모하며 듣는 습관을 지닌다.’의 다섯 가지이다. 이러한 다섯 가지 내용을 해당 학년에서 충분히 학습하기 위해 필요한 사항이 무엇인지 판단하고 선정하여 평가의 내용으로 삼아야 한다. 이와 함께 이전 학년의 내용 중에서도 해당 학년의 학습에 밀접한 관련이 있는 내용들을 골라 평가 내용으로 삼아야 한다. 7학년 듣기를 예로 하여 간단히 표로 구성해 보면 다음과 같다.

<표 1> 7학년 듣기 영역 평가 내용과 관련된 이전 학년의 내용

내용 영역 7학년 듣기	교육과정에 제시된 이전 학년 듣기 영역 관련 내용	교육과정에 제시된 이전 학년 타 영역 관련 내용
7-(1) 듣기와 읽기의 공통점과 차이점을 안다.	6-(1) 듣기가 의미 형성 과정임을 안다.	
7-(2) 사실과 의견을 구별하며 듣는다.		읽기 영역 5-(3) 사실과 의견을 표현한 부분을 구별하며 글을 읽는다.
7-(3) 내용의 짜임을 파악하며 듣는다.	3-(3) 내용의 연결 관계를 파악하며 듣는다.	
7-(4) 내용의 통일성을 평가하며 듣는다.	4-(4) 주제를 파악하며 듣는다.	
7-(5) 메모하며 듣는 습관을 지닌다.	5-(2) 중요한 내용을 메모하며 듣는다.	

이전 학년의 관련 내용들을 모두 다 평가의 내용으로 삼을 필요는 물론 없을 것이며, 평가 문항을 제작할 때 반영해야 할 사항 정도로 처리해도 무리가 없을 것이다. 이전 학년 내용들 중 관련 내용이 어떤 것이 있는지 명확히 확인하고 반영은 하되, 그것을 하나하나 모두 반영하는 데 치중해서 정작 해당 학년에서 학습해야 할 내용에 관한 진단평가라는 평가의 성격에 혼란을 초래하지는 않게 하는 것이 필요하다.

내용 영역을 교육과정에 제시된 학년별 내용에 근거하여 설정한 후에는 이 내용 영역과 결합할 행동 영역의 범주를 설정하는 것이 필요하다. 2007 개정 국어과 교육과정을 보면 학년별 내용이 성취기준형으로 되어 있어서 이 행동 영역 구성에 어느 정도 시사점을 준다. 개정 교육과정의 성취기준은 ‘① 텍스트+② 내용 요소+③ 행동(이해, 분석, 해석, 평가, 조사 등)이 결합된 구조이다(이재기, 2007 : 92). 듣기 영역을 예로 7, 8, 9학년의 행동 술어를 모아 보면 아래와 같다.

<표 2> 듣기 영역 성취기준에 사용된 행동 술어(2007 중학교 교육과정)

7학년 듣기	8학년 듣기	9학년 듣기
7-(1) ... 고려하여 ... 메모한다.	8-(1) ... 질문한다.	9-(1) ... 비판적으로 이해한다.
7-(2) ... 파악한다.	8-(2) ... 반응한다.	9-(2) ... 비판적으로 평가한다.
7-(3) ... 분석한다.	8-(3) ... 이해한다.	9-(3) ... 이해한다.
7-(4) ... 파악한다.	8-(4) ... 평가한다.	9-(4) ... 관련짓는다.

이상의 술어는 여섯 개 영역에서 자주 반복되는 것도 있고, 영역에 따라 특수한 술어도 있다. 이들 모두를 분리하여 평가할 수는 없으므로 타당성 있게 범주화를 해야 한다. 이 범주화에 고려해야 할 사항 중 하나가 평가의 효율성이다. 몇 개 정도의 범주로 묶는 것이 가장 효과적이고 효율적인 평가를 가능하게 할 것인지도 생각해 보아야 한다는 뜻이다. 이 분야의 대표적인 예는 블룸(1956 / 1983)의 목표분류이론에 제시된 지식(knowledge), 이해(comprehension), 적용(application), 분석(analysis), 종합(synthesis), 평

가(evaluation)의 여섯 범주이다. 이 여섯 범주를 그대로 구성하고, 그 순서대로 위계적 관계를 드러내 보일 수 있는 평가를 설계하려면 최소한 문항 수가 2 내지 3배수의 문항이 필요하다. 기계적으로 각 행동 영역에 2배수만 잡아도 12문항이 되고, 듣기·말하기, 읽기, 쓰기, 국어지식, 문학의 다섯 영역에서 12문항을 내면 총 60문항이 된다. 여기에 다시 기존의 초등학교 3학년 기초학력 진단평가에서처럼 내용 영역 외에 <문제 상황>, <글의 유형> 등의 변수를 결합시킨다고 한다면, 다시 3배수가 늘어난다. ‘본질적’인 고려 사항은 아닐 수 있으나, ‘현실적’으로는 고려해야 하는 부분이다.

이 연구에서는 이러한 점을 감안하여 교육과정상에 나타나는 술어들을 바탕으로 진단평가의 행동 영역 범주를 크게 세 가지로 설정하였다. 1) 어떤 개념이나 텍스트의 내용에 대한 기본 이해에 초점을 맞추고 있는 경우, 2) 그 지식을 다른 텍스트나 상황에 적용하는 데 중점을 두는 경우, 그리고 3) 자신이나 상대방의 행위 혹은 상황에 대해 학습자의 평가를 수반하는 비판적 사고에 초점을 맞추는 것이다. 그리고 그것을 각각 ‘이해, 적용, 비판’이라는 명칭으로 범주화하는 것으로 평가의 틀을 설계하였다. <표 2>에서 보면 ‘이해’로 묶을 수 있는 것은 7-(2), 8-(3), 9-(3), ‘비판’으로 묶을 수 있는 것은 8-(4), 9-(1), 9-(2), 나머지는 ‘적용’으로 분류할 수 있을 것이다.¹⁴⁾ 이 분류의 틀이 절대적인 기준이 될 수는 없을 것이지만, 중학교 학습자의 국어과 기초학력을 평가하는 데 일정한 유용성이 있으리라고 판단된다. 명제적 지식이나 방법적 지식에 대한 기본적 이해, 그 이해를 바탕으로 다양한 텍스트에 적용할 수 있는 능력, 그리고 자신의 언어 행위를 성찰하고 타인의 언어와 상황을 평가하는 일에 합리성을 갖추는 것 등은 중학교 수준에서 반드시 갖추어야 할 국어 능력의 중요한 범주로 볼 수 있다.

14) 물론 이는 술어만으로 판단할 수 없는 경우가 많기 때문에 텍스트나 성취기준의 성격을 종합적으로 고려해야 한다.

3. 평가 문항의 유형

국어과의 기초학력을 진단하는 평가는 원칙적으로 수행평가가 많은 비중을 차지하는 것이 좋다. 말을 하고, 글을 읽고 쓰는 능력은 그것을 직접 수행한 결과 및 그 과정에 대한 평가를 통해 좀 더 정확하게 평가될 수 있을 것이기 때문이다.

그러나 문제는 채점에 있다. 진위형이나 선택형 등 빠른 속도로 정답자와 오답자를 가려낼 수 있는 방식이 아닌 형태의 문항은 응시자가 대규모인 평가에서는 매우 오랜 시간을 소요하고, 또 대규모의 채점자를 구성하여 채점을 벌이는 것이 쉽지 않기 때문에 제약이 많다. 결국 채점의 곤란함 때문에 대규모 평가에서는 선택형의 평가문항을 선택하게 된다. 장기적으로 볼 때 서답형 문항¹⁵⁾에 대한 인터넷 기반 채점 시스템이 안정적으로 구축된다면 이러한 제약을 극복할 수 있을 것이다.¹⁶⁾ 현재로서는 전국 단위 평가에서 서술형으로 대표되는 서답형 평가를 시행하기 위해서는 많은 비용이 요구된다. 사법시험, 중등교사임용시험 등 응시자가 이러한 비용을 부담하는 체제의 국가 시험이 없는 것은 아니지만, 중학교 학생들의 실력을 진단하는 평가에서 응시자에게 채점 비용의 부담을 지게 하기가 현실적으로 쉽지 않다. 이러한 현실적 제약을 해결할 수 있는 대안이 마련되지 않는 한 선택형 객관식 문항으로 대체하는 것이 불가피하다.

선택형 문항으로 유형을 결정한 다음에는 최소한 국어과 전 영역을 포괄하여 평가할 수 있는 최소 문항수와 중학교 학생들이 집중력을 가지

15) 여기서 서답형 문항은 응시자가 주어진 답지에 따라 정답을 선택하는 것이 아니고 직접 정답을 구성하여 답을 작성하는 문항 형태를 의미한다. 서답형 문항의 유형은 형태에 따라 단답형, 괄호형, 완성형, 서술형(논술형) 등으로 나누어 볼 수 있다(한국교육평가학회, 2004).

16) 이미 미국의 학업성취도 평가 채점에 웹 기반 채점 방식이 도입된 바 있으며, 그밖에 칠레 등에서도 국가 단위 대규모 평가에서의 서답형 문항에 대해 정보통신기술의 발달에 기반한 채점 시스템을 운영하고 있다. 이러한 채점 방식에 대한 상세한 정보는 조지민 외(2007)에 잘 소개되어 있다.

고 평가에 응할 수 있는 적정 문항 수 사이에서 적절한 문항 수를 결정해야 한다. 그리고 문항 수를 기준으로 하여 평가 영역의 비중을 진단평가의 특성에 맞게 합리적으로 구성해야 한다. 이에 대해서도 관점에 따라 다양한 안이 있을 수 있다. 이 연구에서는 사범대학 교수 3명, 연구기관 연구위원 1명, 중학교 교사 8명으로 구성된 전문가 패널에서 전체 문항 수 및 영역별 문항 수 구성안을 다양하게 설정한 후, ‘중학생’, ‘기초학력 평가’, ‘선택형’ 등의 요인을 고려하여 내용상의 타당성과 시행의 현실성을 갖출 수 있는 체제로 아래와 같은 체제를 구안하였다.

<표 3> 영역별, 학년별 문항 수

내용 영역 \ 학년별 문항수	1학년	2학년	3학년
	영역별 문항 수 (비율)	영역별 문항 수 (비율)	영역별 문항 수 (비율)
듣기·말하기	4 (13.3%)	4 (13.3%)	4 (13.3%)
읽 기	10 (33.3%)	10 (33.3%)	10 (33.3%)
쓰 기	4 (13.3%)	4 (13.3%)	4 (13.3%)
국어지식	6 (20.0%)	5 (16.7%)	4 (13.3%)
문 학	6 (20.0%)	7 (23.3%)	8 (26.7%)
합 계	30	30	30

모든 영역을 평가하기 위해 최소한 30문항은 필요하며, 중학생의 기초학력을 평가하는 데에는 읽기 능력이 가장 중요하다고 판단하여 읽기 영역의 비중을 33.3%로 가장 높게 설정하였다.¹⁷⁾ 듣기·말하기와 쓰기는 세 학년에서 동일하게 13.3%로 설정하였고, 국어지식과 문학 영역은 학년에 따라 약간의 변화를 두되, 두 영역을 합쳐서 매 학년 12문항(40%)으로 설정하였다. 이러한 문항 비율은 교육과정에 제시된 여섯 영역 체제에 근거하면서 중학교 기초학력 평가라는 평가목적에 감안하여 영역 간 비중의

17) 읽기의 ‘범교과적 도구성’과 그에 따른 기초학력 평가의 고려사항에 대해서는 천경록(1999), 염은열(2005) 등 참고.

균형을 적절히 맞춘 결과라고 할 수 있다.¹⁸⁾ 국어지식과 문학의 문항 수는 학년에 따라 한 문항씩 증감하도록 하였다. 7차 교육과정의 경우 7학년 국어지식 영역에 음절, 형태소, 낱말, 동음이의어와 다의어, 담화 등 언어를 체계적으로 이해하는 가장 기초가 되는 개념들이 다수 등장한다는 점을 고려하여 세 학년 중 1학년의 문항 수 비중을 높였다. 이에 비해 문학 영역에서는 함축적이고 예술적인 문학 작품들을 대상으로 한 기초학력 평가의 비중은 학년이 올라감에 따라 조금씩 높이는 것이 타당하다고 판단하여 학년이 올라갈수록 문항의 수를 한 문항씩 늘렸다.

IV. 중학교 국어과 기초학력 진단평가 설계

1. 성취기준 설정

구체적인 평가 설계 단계에서는 앞서 구안한 내용 영역(듣기·말하기, 읽기, 쓰기, 국어지식, 문학)과 행동 영역(이해, 적용, 비판)의 이원 체제를 기본 틀로 삼고, 7차 교육과정의 각 학년별 내용들을 종합적으로 재구성하는 작업이 요구된다.

교육과정의 내용에 근거하여 평가를 위한 성취기준을 설정한다고 할 때 다음 세 가지 방법이 있을 수 있다. ① 교육과정의 학년별 내용과 평가를 위한 성취기준을 일대일 대응시키는 방법,¹⁹⁾ ② 앞에서 설정한 문항

18) 듣기, 말하기, 읽기, 쓰기 등 언어사용 기능 영역의 비중을 60%로 잡고, 국어지식과 문학을 합쳐 40%로 영역 간 비중을 설정한 것이 전 학년에 동일하게 적용될 필요는 없다고 본다. 전문가 협의를 통해 중학교 기초학력 평가로는 이러한 이 비율이 가장 적절하다는 판단을 한 것이지만, 초등학교나 고등학교에도 이러한 비율이 동일하게 적용되는 것이 바람직하다고 볼 수는 없다. 중학교에서도 기초학력에 대한 진단평가라는 평가 목적이 달라질 경우라면 이 영역 간 비율 역시 다시 검토·설계될 필요가 있을 것이다.

19) 7차 국어과 교육과정의 중학교 학년별 내용은 여섯 영역에 걸쳐 1학년 35개, 2학년

수에 맞게 교육과정의 학년별 내용을 재구성하여 문항 수와 같은 30개의 성취기준으로 설정하는 방법, ③ 문항 수보다 적은 수의 성취기준을 설정하고, 각 성취기준별로 복수의 문항을 개발하는 방법 등이다. ①은 교육과정과의 일치도 면에서 장점이 있으나, 문항 수가 제한되어 있으므로 일부 내용은 평가 문항으로 개발되지 못한다는 단점이 있고, ③은 각 성취기준에 관련된 학습자의 상태를 복수의 문항으로 파악하여 타당도를 높일 수 있다는 장점이 있으나, 성취기준을 복합적으로 설정해야 하는 부담이 상대적으로 크고, 시행할 때마다 문항의 난도나 유형 등에서 변화가 클 수 있다는 한계가 있다. ②의 경우에는 ①과 ③의 단점을 최소화할 수 있다는 장점이 있는 반면, 하나의 성취기준에서 개발할 수 있는 문항 유형에 한계가 있어 해가 거듭할수록 문항의 유형이 단조로워지는 단점이 있다. 그렇지만 기초학력 진단평가의 경우 특별히 ‘기출문항’은 안 된다든가, 고차원적 사고력을 요하는 창의적 문항 개발의 필요성이 적기 때문에 ②의 단점이 크게 영향을 미치지 않는다고 볼 수 있다. 반면 국가 단위에서 시행하는 기초학력 진단평가가 매년 안정적으로 시행이 지속되고, 그 결과로부터 의미 있는 비교가 이루어지기 위해서는 동형검사 형태가 유지되는 것이 요구된다. ②는 그러한 요구를 비교적 안정적으로 충족시켜 줄 수 있는 방법이다. 이러한 점을 고려한 결과 이 연구에서는 ②의 방법을 취하여 교육과정에 있는 30여 개의 학년별 내용을 재구성하여 해당 학년의 학습을 해 나가기에 필요한 30개의 성취기준을 설정하였다. 문항 수와 동일하게 30개로 설정한 1학년 진단평가의 성취기준 설정 초안은 아래와 같다.²⁰⁾

36개, 3학년 37개이다.

20) 이것은 아직 하나의 ‘안’으로 예비평가 등의 통해 수정할 여지가 있는 상태이다.

<표 4> 중학교 1학년 국어과 진단평가 성취기준(안)

중학교 1학년 기초학력 진단평가 국어과 성취기준(안)				
번호		내용 영역	행동 영역	성취 기준
1	듣말	내용의 구성	이해	내용의 짜임을 파악하며 듣는다.
2	듣말	내용의 전개방식	이해	내용의 전개 방식을 파악하며 듣는다.
3	듣말	자료 활용	적용	주제에 알맞은 자료를 활용하여 말한다.
4	듣말	내용의 통일성	비판	말의 내용이 주제에서 벗어나지 않았는지 평가한다.
5	쓰	주제문 쓰기	이해	중심 내용과 의도를 반영하여 주제문을 쓴다.
6	쓰	내용의 구성	적용	통일성 있게 내용을 조직한다.
7	쓰	적절한 표현	적용	상황에 맞는 관용 표현을 사용한다.
8	쓰	근거의 타당성	비판	주장을 뒷받침하기에 알맞은 근거를 제시한다.
9	국지	자료 활용	이해	한글 자모의 순서를 고려하여 사전 찾는 순서를 안다.
10	국지	어휘의 의미	이해	동음이의어와 다의어의 개념을 안다.
11	국지	문장의 연결	이해	문장과 문장을 연결하는 어휘를 안다.
12	국지	어휘의 의미	적용	단어들의 의미 관계를 파악한다.
13	국지	어휘의 선택	적용	상대방을 고려하여 적절한 어휘를 선택한다.
14	국지	국어 태도	비판	국어 순화의 필요성을 안다.
15	읽	읽기 방법 이해	이해	글의 종류에 따른 적절한 읽기 방법을 안다
16	읽	중심내용 요약	이해	중심 내용을 요약한다.
17	읽	주제 파악	이해	주제를 파악하며 글을 읽는다.
18	읽	비유적 표현	이해	비유적 표현의 의미를 파악하며 글을 읽는다.
19	읽	정확한 이해	이해	문장 부호의 쓰임에 유의하며 글을 읽는다.
20	읽	배경지식활용	적용	배경 지식을 바탕으로 글의 내용을 이해한다.
21	읽	메모 활용	적용	필요하거나 중요한 내용을 메모하며 글을 읽는다.
22	읽	근거 판단	비판	주장에 대한 근거의 타당성을 판단하며 글을 읽는다.
23	읽	의도 파악	비판	글쓴이의 의도나 목적을 파악하며 글을 읽는다.
24	읽	내용의 통일성	비판	어휘나 문장이 주제와 연결되는지 평가하며 글을 읽는다.
25	문	작품의 특성	이해	표현의 특징을 파악하며 작품을 읽는다.
26	문	인물의 성격	이해	작품에 나타난 인물의 성격을 이해한다.
27	문	배경의 이해	이해	작품에 드러난 상황을 구체적으로 파악한다.
28	문	인물의 심리	이해	작품에 나타난 인물의 심리 상태를 파악한다.
29	문	내용의 구성	적용	사건의 전개 과정을 파악하고 생략된 내용을 추리한다.
30	문	인물의 성격	비판	작품 속에 드러난 내용을 바탕으로 인물의 성격을 평가한다.

듣기·말하기, 읽기, 쓰기, 국어지식, 문학의 각 내용 영역별로 교육과정
정에 제시된 해당 학년 및 이전 학년의 학년별 내용들을 재구성하여 하위
내용 요소들을 구체화하고, 그에 따라 구체적인 성취기준을 설정하였다.
분류의 타당성, 중요 내용 요소나 성취기준의 누락 여부 등에 대해 조금
더 보완이 필요한 상태이긴 하나 기본적으로 하나의 성취기준에 하나의
문항을 구성할 수 있도록 설계하였다.

이 성취기준들은 진단평가 후 학습자들에게 송환될 정보의 구체적인
하위 항목들이 된다. 각 성취기준에 근거하여 설계된 문항을 맞혔을 경우
해당 성취기준에서 제시하는 능력을 갖추고 있다고 판정하는 근거가 되
며, 틀렸을 경우 그러한 능력을 갖추고 있지 못하므로 그에 따른 보정 교
육이 요구된다는 정보를 제공하게 된다.²¹⁾

위 표에 포함된 성취기준의 분포를 살펴보면 아래와 같다.

<표 5> 중학교 1학년 성취기준의 내용영역/행동영역 분포(안)

	듣기/말하기			쓰기			읽기			국어지식			문학		
문항 수	이해	적용	비판	이해	적용	비판	이해	적용	비판	이해	적용	비판	이해	적용	비판
	2	1	1	1	2	1	5	2	3	3	2	1	4	1	1
	4			4			10			6			6		

각 내용영역별로 이해와 적용과 비판이 고루 분포하되 기본적으로 이
해의 비중이 높고, 비판의 비중이 낮도록 설계하였음을 볼 수 있다. 이는
기초학력 진단평가라는 평가 목적상 상대적으로 기본 개념이나 텍스트의
내용을 이해하는 능력에 중점을 두었기 때문이다. 이러한 방식으로 2학년
과 3학년도 구성을 하였으며, 기본 틀은 그대로 유지해 가되, 학년이 올라
갈수록 이해의 비중을 다소 줄이고, 적용과 비판의 비중은 다소 늘리는
쪽으로 설계하는 것이 바람직할 것이다. 각 학년별로 성취기준별 문항 구
성 초안의 행동영역별 분포는 다음과 같다.

21) 각 성취기준 미도달 학습자에게 제공할 구체적인 보정교육 프로그램은 이번 연구에서
는 개발하지 못하였다. 후속 연구가 필요한 부분이다.

<표 6> 성취기준의 행동영역별 분포(안)

	이해	적용	비판	계
	문항 수(비율)	문항 수(비율)	문항 수(비율)	문항 수(비율)
1학년	15(50.0%)	8(26.7%)	7(23.3%)	30(100%)
2학년	13(43.3%)	10(33.3%)	7(23.3%)	30(100%)
3학년	11(36.7%)	11(36.7%)	8(26.7%)	30(100%)

학년이 올라갈수록 이해의 비율은 줄이고, 적용과 비판의 비율을 높여 가는 구조로 구성되어 있다. 학년이 높아짐에 따라 기초학력 수준에서 요구되는 사고 작용이나 행동도 점차 단순한 쪽에서 복합적인 쪽으로 변화해야 함을 반영한 구도이다.

2. 기초학력 도달 기준점 설정

위 성취기준 표에 따라 문항을 제작하고, 기초학력 도달 기준점을 설정하기 위한 예비평가²²⁾를 시행한 다음, 그 자료를 바탕으로 앞서 언급한 북마크 기법을 이용하여 기초학력 도달 기준점을 설정하였다. 패널은 사범대학 교수 3명, 연구기관 연구위원 1명, 중학교 교사 8명으로 구성하였으며, 3인 1개조로 4개조를 편성하였다. 조별로 소그룹을 결성하여 각 그룹별로 토의를 거쳐 1차 분할점수를 설정하고, 다시 2개조가 하나의 중그룹을 이루어 2차 분할점수를 설정하였다. 마지막으로 전체가 모여 다시 토의를 하고, 각 영역별로 점검을 한 뒤, 최종적으로 분할점수를 설정하였

22) 예비평가는 2008년 3월, 중학교 2학년 약 8,000명, 중학교 3학년 약 7,700명을 대상으로 전국 16개 시도의 인구 비례에 따른 비율로 표집하여 시행하였다. 전국 중학생 수는 2008년 현재 총 2,063,159명으로, 표집 비율은 1% 정도에 해당한다. 앞에 구성한 문항 개발 계획에 따라 학년별로 30문항씩 출제하였으며, 문항의 난도는 기초학력을 측정하기 위한 문항이므로 대부분의 학생들은 ‘쉽다’고 느낄 수 있는 난도의 문항들로 주로 구성하였다. 현재 예비평가의 단계이고, 문항 및 분석 결과는 공개하지 않도록 되어 있으며, 본 논문의 목적은 개발의 방향을 설정하는 데 목적을 두고 있으므로 문항과 분석 결과는 제시하지 않기로 한다.

다. 각 학년별, 각 내용 영역별로 분할점수를 설정하였는데, 그 중 3학년 <문학> 영역의 분할점수 설정 과정을 보면 아래와 같다.

<표 7> 중학교 3학년 문학 영역의 기초학력 도달기준선 설정 결과

조	패널	소그룹	중그룹	전체
1	1	4	3	3
	2	4	3	3
	3	4	3	3
2	1	3	3	3
	2	3	3	3
	3	3	3	3
3	1	3	2	3
	2	3	2	3
	3	3	2	3
4	1	3	2	3
	2	3	2	3
	3	3	2	3
도달기준선 (중앙값)		3.25	2.5	3

중학교 3학년 문학 영역은 총 8문항으로 구성되어 있다. 각 문항에 대한 학습자들의 반응을 분석하여 정답률이 높은 문항부터 순서대로 재편집을 한 후, 각 패널별로 기초학력 수준의 학생이라면 몇 번 문항까지 맞힐 수 있다고 보는지를 판단하게 하는 것이다. 위 표에 나와 있는 소그룹의 결과 이전에 각 패널별로 개인 의견에 편차가 있다. 그것을 조별로 토의하여 1조는 4문항, 2, 3, 4조는 3문항을 맞히는 것이 기초학력의 수준이라고 결정하였다. 그런데 1조와 2조가 중그룹을 형성하여 토의를 한 결과, 1조의 패널들이 자신들의 의견을 수정하고 2조의 의견을 받아들여 기초학력은 3문항을 맞히는 수준이라고 결정한 반면, 3조와 4조는 논의 중에 자신들의 선택이 너무 기초학력을 높게 보고 있는 것이라고 판단하여 2문항으로 하향 조정하였다. 이는 논의과정에서 정답률과 변별도 등을 보면서

충분히 나뉠대로의 근거를 가지고 조정이 되는 경우이다. 최종적으로 전체 논의 과정에서 여러 근거를 동원한 논의들이 오고 간 끝에 다시 3문항으로 조정이 되었으며, 그 선을 중심으로 하여 기초학력 도달자와 미도달자를 변별할 수 있는 정보를 구성하였다.²³⁾

이상에서 북마크 방법으로 기초학력 도달기준선을 설정하는 작업의 과정을 간략하게 제시하였다. 일괄적으로 1%, 3% 등의 비율을 정하는 방법이나, 전적으로 출제자의 경험적 직관에 근거하여 문항별 예상정답률을 정하고 그에 따라 도달/미도달을 나누는 것에 비해 보다 더 객관적인 설정이라고 할 수 있다. 다양한 전문가의 안목을 종합할 수 있으며, 논의 과정에서 나온 토의의 내용들을 기록, 분석한 후 그것을 나중에 학습자들에게 평가 결과에 대한 정보를 제공할 때에 유의미한 정보로 반영하여 제공할 수 있다는 장점이 있다. 다만 현재 작업에는 모든 문항의 배점을 동일하게 1점으로 설정하였고, 패널의 수도 다소 부족하다는 한계가 있다. 앞으로 문항별 배점을 차등화하고, 단순히 문항 수로 결정하는 것이 아니라 척도화된 점수를 사용하는 등 좀 더 정교한 설계를 해 나감으로써 보다 더 정확한 기초학력 도달 기준선을 설정할 수 있을 것이다.

23) 이러한 과정을 영역별로 거친 후, 그것을 다시 종합하여 도달점을 결정하였다. 이번 예비평가 문항 분석 결과를 바탕으로 설정한 기준점은 중학교 2학년은 30개 문항 중 15개, 3학년은 30개 문항 중 13개이다. 즉 2학년의 경우 14개 이하를 맞았다면 중학교 2학년에 배울 내용을 이수하기 위한 기초학력을 갖추지 못하였다고 판단할 수 있고, 각 영역별로 설정된 기준점 및 문항 정보에 따라 부족한 부분을 학습자와 교사에게 알려 줄 수 있다. 기준점에 도달하지 못한 학생 수는 전체 예비평가 응시자 중 약 5%, 중학교 3학년의 경우는 약 7% 정도였다. 학급당 인원이 40명이라면, 대략 2명에서 3명 정도에 해당한다고 할 수 있겠다. 이는 ‘기초적 소양’을 평가하려고 한 평가의 의도에 어느 정도 부합하는 결과라고 할 수 있겠다.

V. 맺음말

정확한 진단 없이 막연히 모든 학습자들이 영점(零點)에 균등하게 분포하고 있을 것이라는 잘못된 가정에 근거하여 출발하는 교육은 다수 학습자의 필요나 요구를 충분히 반영하기 어렵다. 현실적으로 개별 교실에서 진단평가가 제대로 이루어지고 있지 않다면, 비록 학습자의 상태를 세밀하게 진단하기는 어렵다고 하더라도 최소 수준의 도달/미도달 여부를 판단하는 기본적인 정보를 국가에서 제공함으로써 모든 국민이 기초적인 학력을 갖추는 데 도움을 줄 수 있어야 할 것이다.

물론 국가에서 일괄적으로 시행하는 진단평가는 한계를 가질 수밖에 없다. 특히 선택형 문항으로 국어과 교육의 기초학력을 진단하는 체제로는 그 효과가 아무래도 제한적일 수밖에 없다. 그렇기 때문에 학습자들의 실제 언어 능력을 진단하는 평가 도구의 개발과 실시의 주체는 언제나 교사여야 한다(김석우 외, 2004 : 66)는 점을 분명히 해 두는 것이 필요하다. 앞에서 설계한 국가 수준에서의 국어과 기초학력을 진단하는 평가 도구 개발은 교사들의 진단평가를 위한 기초 자료 제공과 평가 도구 개발의 예를 제시하는 기능을 가진다는 점을 다시 확인해 둔다.

핵심적이라고 할 수 있는 평가 문항, 그리고 그 문항으로 예비평가를 실시한 결과에 대한 분석 자료를 공개할 수 없는 제약으로 인해 충분한 논의를 보이지 못하였다. 향후 국어과 기초학력 진단평가의 발전을 위한 시급한 과제를 든다면, 읽기, 쓰기, 문학 등을 위한 서답형 평가, 나아가 듣기, 말하기 영역을 위한 멀티미디어 기술을 이용한 평가가 시행되어야 한다는 점, 그리고 그에 대한 대규모 채점이 기술적, 제도적으로 충분히 높은 신뢰도 하에 이루어질 수 있는 체제가 개발되어야 한다는 점을 들 수 있겠다. 또한 각 교실과 학습자들의 특수 상황을 고려한 교사들의 진단 평가가 충분히 이루어질 수 있게 하기 위하여 다양한 교사 연수 프로그램이 개발되어야 할 것이다. 나아가 진단 후에 효과적으로 보정 교육이 실시될 수 있도록 진단 후 교육을 위한 커리큘럼이나 교재 등을 개발하는

일도 필요할 것이다. 그리고 국가 단위 진단평가를 매년 실시하게 된다면, 그 평가 도구 개발 단계에서 동등화 설계를 하여 매년 그 변화 추이를 파악할 수 있게 해야 한다. 그렇게 함으로써 국가수준 학업성취도 평가와 짝을 이루는, 우리 교육을 과학적으로 분석할 수 있게 하는 의미 있는 자료를 얻을 수 있게 될 것이다.

진단 없는 처방전이 때로 비극적 결과를 낳을 수 있다. 국가 단위에서, 또 각 교실에서, 진정한 학습자 중심 교육을 위한 정확한 학습자 진단에 많은 노력을 기울여야 할 것이다.*

* 본 논문은 2008. 6. 30. 투고되었으며, 2008. 7. 3. 심사가 시작되어 2008. 7. 31. 심사가 종료되었음.

■ 참고문헌

- 교육인적자원부(1999a), 『중학교 교육과정 해설(Ⅰ)－총론·특별활동』.
- 교육인적자원부(1999b), 『중학교 교육과정 해설(Ⅱ)－국어』.
- 교육인적자원부(2007), 『국어과 교육과정』(교육인적자원부 고시 제2007-79호[별책5])
- 교육과학기술부(2008), 『중학교 교육과정 해설(Ⅰ)－총론·특별활동』.
- 김대현·김석우·박소영(2000), “수준별 교육과정 평가 준거 개발을 위한 탐색적 연구”, 『교육과정연구』 18권 1호, 299-326.
- 김석우·김정섭·정성아(2004), “수준별 교육과정의 효율적 운영을 위한 학생평가 방안 연구”, 『교육과정연구』, 22권 1호, 45-74.
- 김수동·이화진·유준희·임재훈(1998), 『학습부진아 지도 프로그램 개발 연구(RRC-98-4)』, 한국교육과정평가원.
- 김수천(2005), “국민의 교육받을 권리와 기초학력 책임지도제”, 『교육과정연구』 23권 4호, 157-174.
- 김신영(2002), “현장교사의 평가전문성 연구”, 『교육평가연구』 15권 1호, 67-85.
- 김영철(2000), “공·사교육의 위상 재정립 및 발전 방향”, 한국교육미디어 21C교육연구소, 『공교육과 사교육의 보완적 발전방향 세미나』 자료집.
- 김정우(2007), “2007 개정 국어과 교육과정의 개정 방향과 특징”, 『한말연구』 20, 한말연구학회, 101-129.
- 남명호 외(2007), 『2006년 초등학교 3학년 국가수준 기초학력 진단평가 연구(연구보고 CRE 2007-3-1)』, 한국교육과정평가원.
- 노명완·정혜승·옥현진(2003), 『창조적 지식기반사회와 국어과 교육』, 박이정.
- 민현식(2005), “국어교육과 국가 경쟁력”, 『국어교육』 117, 한국어교육학회, 233-285.
- 백순근(2000), “교육평가의 개념에 대한 고찰 : ‘교육적 가치’를 중심으로”, 『교육평가연구』 13권 1호, 1-10.
- 백순근·최미숙·유동엽·김광조(2000), 『중고등학교 학생들의 국어 기초학력 평가 연구』, 한국교육과정평가원.
- 서울대학교 교육연구소(1998), 『교육학대백과사전』, 하우동설.
- 송현정(2000), “국어 읽기 부진의 진단에 대한 연구”, 『국어교육학연구』 11, 189-213.
- 염은열(2005), “읽기 기초학력 평가를 위한 선결 과제”, 『국어교육학연구』 22, 한국어교육학회, 213-241.
- 윤정일(1999), “교육개혁 추진 과정의 쟁점”, 『교육개혁 정책의 진단과 전망』, 한국교육학회 1999년 춘계 학술대회 발표 자료집, 41-60.
- 이규민(2003), “Bookmark 방법을 이용한 수행 기준 설정”, 『교육진흥』 16권 3호.

- 이규민(2005), “초등학교 3학년 국가수준 기초학력 진단평가 동등화 방안”, 『교육평가 연구』 제18권 1호, 한국교육평가학회, 125-151.
- 이도영(2001), “국어과 교육 평가 모형”, 『국어교육학연구』 13, 국어교육학회, 297-302.
- 이재기(2007), “2007년 개정 국어과 교육과정의 특징과 실행 방안”, 『청람어문교육연구』 36호, 청람어문교육학회, 81-108.
- 이종원(2002), “텍사스 주 교육평가제도의 변화와 사회와 평가”, 『한국지리환경교육학회지』 10권 3호, 73-85.
- 정구향(2005), “21세기 국어과 평가의 발전 방향”, 『새국어교육』 71, 한국국어교육학회, 335-370.
- 조지민 · 김경훈(2007), “서답형 문항의 인터넷 기반 채점시스템 설계 연구”, 『한국컴퓨터교육학회논문지』 10권 2호, 한국컴퓨터교육학회, 89-100.
- 천경록(1999), “읽기의 개념과 읽기 능력의 발달 단계”, 『청람어문학』 21호, 청람어문학회, 263-282.
- 최미숙(2006), “국어과 평가의 반성과 탐색”, 『국어교육』 121, 한국어교육학회, 79-105.
- 최인자(2005), “한국의 언어문화와 국어교육”, 『국어교육』 117, 한국어교육학회, 93-131.
- 한국교육과정평가원(2003), 『기초학력 보정교육 자료』, 한국교육과정평가원.
- 한국교육평가학회 편(2004), 『교육평가용어사전』, 학지사.
- 현주(2003), “의무교육 단계에서 기본 학업성취(학력) 보장을 위한 과제”, 『교육개발』 30권2호, 한국교육개발원, 27-31.
- 황정규(1998), 『학교학습과 교육평가』, 교육과학사.
- Bloom, B. S.(1956), 임의도 외 역(1983), 『教育目標分類學：教育 目標의 分類 및 評價의 實際』, 교육과학사.
- Klein, S. P., Hamilton, L. S., McCaffrey, D. F. & Stecher, B. M.(2000), “What do test scores in Texas tell us?”, *Education Policy Analysis*, 8(49), October 26.
<http://epaa.asu.edu/epaa/v8n49/>.
- Zieky, M. J.(2001), “So much has changed : How the setting of cutscores has evolved since the 1980s”, in G. J. Cizek ed., *Setting performance standards : concepts, methods, and perspectives*, Mahwah, NJ : Lawrence Erlbaum Associates.

<초록>

중학교 국어과 기초학력 진단평가에 관한 연구

윤여탁 · 민현식 · 김정우

이 논문은 중학교 국어과 기초학력 진단평가 개발을 위한 연구이다. 이를 위해 기초학력의 개념을 설정하고, 성취기준을 개발하며, 실제 문항 개발의 틀을 설정하여 제시하고자 하였다.

기초학력에는 여러 개념이 있는데, 이 논문에서는 해당 학년에서 배울 내용을 원활하게 학습하는 데 필요한 기초적 능력으로 그 의미를 규정하였다. 이전 학년에 도달한 학력을 기반으로 하되, 거기에 해당 학년에서 새로 배우게 될 내용에 관계되는 부분이 결합하는 형태로 구성된다. 초점은 이전 학년 학습 내용에 대한 성취도가 아니라 해당 학년에서 배워야 할 내용과 관련되는 최소한의 학력의 보유 여부이다.

기초학력 진단평가는 현재 시행되고 있는 국가 수준 교육과정에 근거하여 설계하는 것이 바람직하다. 제7차 국어과 교육과정에 근거하여 진단평가를 계획할 때에는 수준별 교육과정 체제의 [기본]에 초점을 맞추어 설계하는 것이 필요하고, 내용 영역은 ‘듣기·말하기’, ‘읽기’, ‘쓰기’, ‘국어지식’, ‘문학’의 다섯 영역, 행동 영역은 ‘이해’, ‘적용’, ‘비판’의 세 영역으로 하는 것이 적절하다.

성취기준은 문항 수와 동일하게 30개로 구성하였고, 기초학력 진단평가의 특성을 고려하여 내용 영역에서는 ‘읽기’, 행동 영역에서는 ‘이해’가 가장 많은 비중을 차지하도록 구성하였다.

이 평가는 준거참조평가로 기준점 도달 여부와 성취기준별 송환 정보를 제공한다. 기준점을 전문가들의 협의를 통한 북마크 기법으로 설정하는 것이 바람직하다.

【핵심어】 기초학력, 진단평가, 성취기준, 기준점, 북마크 기법

<Abstract>

A Study of the Diagnostic Assessment of Basic Competency
in Middle School Korean Language Learning

Yoon, Yeo-tak · Min, Hyun-sik · Kim, Jung-woo

This study aims to explore how to develop the diagnostic assessment of basic competency(DABC) in middle school Korean language learning, specially focused on the achievement standards for assessment and frame of developing test items.

DABC should be designed to be based on the current national curriculum. If DABC is developed based on the 7th National Curriculum Korean language, it is most desirable to focus on the 'basic' level of the contents, considering the characteristics of differentiated curriculum. And the frame of test items is able to be organized as the two dimensions of assessment, contents and behaviors, the former is constituted by 'listening and speaking', 'reading', 'writing', 'grammar', 'literature', and the latter is constituted by 'understanding', 'applying', 'critiquing'. Considering the characteristics of DABC, 'reading' and 'understanding' should be given more weight.

DABC is criterion-referenced evaluation and the result are reported to the students, teachers, schools with two levels such as Basic and Below-Basic. The cut-off score dividing the levels was determined by bookmark method.

【Key words】 basic competency, diagnostic assessment, achievement standard, cut-off score, bookmark method

【토론문】

“중학교 국어과 기초학력 진단평가에 관한 연구”에 대한 토론

박수자(부산교육대학교)

I. 연구의 요약

이 연구는, 2002년부터 초등학교 3학년을 대상으로 실시해 온 ‘국가 수준 기초학력평가’와 ‘보정 교육’을 중등학교로 확대하려는 연장선상에서 수행되었다고 했다. 우선, 지금까지 초등학교에 비해 중등학교에서 국가 단위의 진단평가가 제대로 이루어지지 않은 원인을, ‘진단평가가 개별 학습자에 대한 충분한 정보를 제공하기 어렵다는 점, 진단평가가 전집 방식으로 진행될 경우 발생하는 학생 서열화의 폐해’ 등으로 분석하였다.

그럼에도 불구하고 국가 수준의 교과학습 진단평가 도구 개발이 필요한 것은, 국가 단위에서 개발한 평가도구의 공신력을 바탕으로 수업 및 평가에 대한 교사의 업무 부담을 줄일 수 있다는 점 때문이라고 했다. 특히 전집이 아니고 표집으로 실시된다면, 서열화 논란을 잠재우고 수업에 필요한 경우 언제든 교사가 평가 문항과 결과(보정교육 : 교육적 처방)를 활용할 수 있다고 했다.

이런 맥락에서 이 연구가 제시하는 ‘기초학력’의 개념 규정과 기준 설정에 주목할 필요가 있다. 우선, 개념 규정 면에서 이 연구는 기초학력에 대한 여러 가지 개념 중에서 ‘위계적 관점에서 다음 단계 학습의 기초가 되는 능력’을 선택하되, 이 개념을 다소 변형하였다. 구체적으로 보면, ‘해당 학년의 기본적 능력의 수준과 범위’를 먼저 정하고, ‘직전 학년의 학업 성취도’를 참고하되 직전 학년의 학업성취도 평가와는 구별되는 선에서, 해당 학년에 대한 ‘기초학력 진단평가 문항’을 마련한다는 것이다. 결국

중학교 진단평가에서 가정하는 ‘기초학력’이란, ‘해당 학년에서 배울 내용을 원활하게 학습하는 데 필요한 기초적 능력’인데, 이는 이전 학년에 도달한 학력을 기반으로 하되, 거기에 해당 학년에서 새로 배우게 될 내용에 관계되는 부분이 결합되는 형태로 구성된다고 한다.

그리하여, 해당 학년에서 배워야 할 내용과 관련되는 최소한의 학력의 보유 여부로서 기초학력의 도달 기준을 선정하고자, 교육과정에서 추출한 <표 4>와 같은 초안을 토대로 문항을 제작한 후 전문가 협의와 합의를 거치는 북마크 기법을 사용하였다. 이때, 평가의 모든 근거를 현실적인 현장 시행을 염두에 두고 7차 교육과정의 [기본] 수준에 초점을 두었다고 한다. 구체적으로 교육과정 술어 연구를 통해 평가 범주를 ‘이해, 적용, 비판’으로 정하고, 내용영역을 ‘듣기·말하기, 읽기, 쓰기, 국어지식, 문학’으로 하되, 평가 채점과 비용 절감을 위해 선다형 평가 문항(30문항)으로 출제한다고 한다.

II. 연구에 대한 생각

이 연구를 읽으면서 드는 본질적인 의문점을 몇 가지 짚고 넘어가고자 한다.

첫째, 중학생에 대한 기초학력의 진단평가는, 표준지향적(상대평가)이어야 하는가, 아니면 준거 지향적(절대평가)이어야 하는가?

둘째, 현행과 같은 기초학력 진단평가는 국가 차원의 자료 수집이 목적인가? 아니면 학습자 개인에게 교육적 처방을 해 주기 위한 교육 실행의 일환인가?

이상의 의문은 매우 본질적이어서 이미 실행된 이 연구가 해결할 사안은 아니지만, 이 연구의 기반과 관련되는 것이므로 실제 연구자들의 고충과 의견을 다음과 같은 사항을 중심으로 들었으면 한다.

1. 평가도구 개발에서 가장 중요한 것은, 내용요소의 추출과 평가 유형의 선정이다. 이 연구는 평가자원의 출처를 교육과정의 내용항목에 두

고 있다. 교육과정은 학년별로 배울 내용을 구별하여 제시하기 때문에, 본질적으로 직전 학년과 다음 학년에 같은 내용항목이 반복 진술되지는 않는다. 따라서 직전 학년의 성취도를 다음 학년의 기초학력으로 연계하기가 쉽지 않았을 것으로 짐작된다. 또한, 교육과정의 [기본]이 모든 학생이 도달해야 할 수준이라고 해서, 그 수준이 곧 다음 학년의 ‘기초학력’이 될 수는 없다. 선행 학습을 찾기 위해 교육과정에서 관련 학년을 검토하다 보면, 자칫 누적적 평가가 될 수도 있다. 그렇다면, <표 4>보다는 좀 더 구체적인 성취기준의 설정이 있었을 것인데, 그 구체화 과정에 대한 설명을 해 주었으면 한다. 이때, 이 연구에서는 ‘기초학력’과 ‘기본학력’에 대한 개념을 어떻게 구별하였는지도 설명해 주었으면 한다.

2. <표 4>를 보면, 각 영역별로 ‘이해, 적용, 비판’이 골고루 분포하되 이해의 비중이 높고 비판의 비중이 낮도록 하였는데, 학년이 높아지면 적용과 비판의 비중을 높이는 것으로 설계하였다고 한다. 이는 논리적인 접근 같지만, 어쩌면 탈맥락적이고 작위적이다. 이러한 가정에서도 역시 기초학력과 기본학력의 경계 구분에 대한 오해의 여지가 있다. 초등학교 3학년 기초학력 진단평가 도구가 3%의 부진 학생을 가려내어 보정교육을 하는 것이라면, 그 연장선상에서 시행한다는 중등학교의 진단평가로서는 그 기초학력의 기준 설정이 엄청나게 높은 것이 아닌가 한다.

* 현재 초6, 중3, 고1 대상으로 실시되는 국가 수준 학업성취도평가에서 ‘선택형, 서답형, 수행평가’가 함께 사용되고, ‘우수학력, 보통학력, 기초학력, 기초학력 미달’로 그 등급을 구분한다.

3. 진단평가의 가장 중요한 목적은, 학생의 점수나 등급의 제시가 아니라 개별 학생에 대한 취약점을 지적하고 제시하는 것이다. 평가문항의 개발만큼이나 평가 문항별로 평가 결과 처리 방식에 대한 논의도 중요하다. 만일, 평가유형별로 학생의 취약점 서술 방법과 이후 보정학습에 대한 안내지침이 함께 공개적으로 제시되지 않는다면, 즉 교사든 학생이든 학생에 대한 정보가 충분히 파악되는 양식이 마련되지 않는다면, 이것은 진단검사가 아니라 배치고사(placement test)가 된다. 평가 문항이나 유형을 구

체적으로 공개할 수 없는 상황이지만, 평가 문항 개발에서 학생의 취약점 포착과 처방에 대해 어떻게 고려하였는지에 대한 설명을 간략히 해 주었으면 한다.

- * 배치고사는 학생의 현재 지식이나 능력 수준을 탐색하고, 특정 학습 내용에 대한 지식보다는 학생의 전반적인 능력을 탐색하고 추정하고자 할 때 실시된다.

이상에서 제기한 내용들은 사실 토론자의 소임을 다하기 위해 마련한 것이다. 정작 개발된 평가 문항을 직접 보지도 못하였기 때문이다. 국가 수준의 평가문항 개발이 얼마나 어려운 작업인지는 두말할 필요가 없다. 개발자의 노고에 갈채를 보낸다.