

# TIMSS 2011에 나타난 우리나라 초등학생들의 과학 성취 특성 분석

김지영 · 김수진<sup>†</sup>

(둔촌중학교) · (한국교육과정평가원)<sup>†</sup>

## An Analysis of Korean Elementary School Students' Science Achievement in TIMSS 2011

Kim, Jiyoung · Kim, Soojin<sup>†</sup>

(Doon Chon Middle School) · (Korea Institute for Curriculum and Evaluation)<sup>†</sup>

### ABSTRACT

This research purports to analyze released items and G4 students' science achievement from TIMSS 2011 according to their academic achievement levels and gender. By doing so, it aims to draw educational implications for Korea from analyses results. Korea showed a lower rate of students at the advanced international benchmark - the highest achievement level - compared to Singapore. The difference was the smallest in Life Science among three content domains and knowing among three cognitive domains. The results of analysis according to gender showed that male students' achievement was significantly higher in Physical Science and Earth Science, and their achievement was also higher in the cognitive domains of Knowing and Applying. From the analysis of the released items, it was revealed that the students' achievement was low in items related to classification of organisms, functions of heart, matters that combust or emit light, and the concept of rotation. Moreover, students drew some illogical conclusions based on their personal experience. Male students were found to show high achievements in items that were not included in curriculum, constructed-response items in the form of short-answer questions, and multiple-choice items in the Knowing domain. Female students were found to show high achievement in items that were included in curriculum, constructed-response items that require reasons and methods, and items that represent experimental situations. Male students showed high achievement in forces concept and movements concept of bodies in the universe, while female students showed high achievement in solubility concept.

**Key words** : science achievement, released items, international benchmark, content domain, cognitive domain

### I. 서론

국제 교육성취도 평가협회(International Association for the Evaluation of Educational Achievement; 이하 IEA)의 주관으로 시행되는 수학·과학 성취도 추이 변화 국제비교 연구(Trends in International Mathematic and Science Study; 이하 TIMSS)는 초등학교 4학년 및 중학교 2학년 학생들의 수학·과학 성취도 변화 추이를 국제적인 수준에서 비교하고, 교육 맥락 변

인들과 성취도 사이의 관계 파악을 목적으로 하고 있다. TIMSS의 평가들은 교육과정에 기반하고 있기 때문에, 각 국가의 교육과정의 타당성을 점검할 수 있으며, 과학 교육의 변화를 파악할 수 있는 중요한 기반이 된다.

TIMSS의 연구는 초등학교 4학년과 중학교 2학년 학생들을 대상으로 4년 주기로 시행되는데, 우리나라는 1995년부터 연구에 참여하고 있다. 중학교 2학년은 TIMSS 1995부터 TIMSS 2011까지 총 5

주기에 걸쳐 꾸준히 참여하였으나, 초등학교 4학년은 1995년에 참여한 후 계속적으로 참여하지 않다가 2011년에 TIMSS 연구에 다시 두 번째로 참여하였다. 1995년 이후 16년 만에 4학년 학생들이 다시 국제 성취도 평가에 참여하여 그 결과를 얻음으로써 우리나라 초등 과학교육을 국제적인 수준에서 점검할 수 있게 되었다.

TIMSS 2011의 과학 성취도 결과가 2012년 12월에 발표되었는데, 우리나라의 초등학교 4학년 과학 성취도는 587점으로 참여국 중 1위로 가장 높은 성취를 보였다. 싱가포르가 평균점수 583점으로 2위, 핀란드가 570점으로 3위를 차지한 것으로 나타났다. TIMSS 1995에서도 우리나라 4학년 학생들의 과학 성취도는 576점으로 1위를 차지하여 초등학교 4학년은 두 번의 참여에서 가장 높은 결과를 나타내었다. 더욱이 TIMSS 1995에 비해 TIMSS 2011에서는 11점 더 상승한 결과가 나타났다(김수진 등, 2012; Martin, Mullis & Foy, 2012).

TIMSS 2011 이전 주기의 과학 성취도 결과를 이용해 우리나라 교육의 현 주소를 파악하고, 개선방안에 대한 시사점을 도출하려는 연구들이 수행되어 왔다. 과학의 내용영역인 물리, 화학, 생물, 지구과학 중 한 영역의 성취를 심층적으로 분석하여 시사점을 도출한 연구나(곽영순과 정은영, 2007; 정은영, 2005; 전경문, 2003), 과학 성취도 결과에 따라 학생들을 성취수준별로 나누어 성취도를 분석한 연구가 수행되었다(홍미영 등, 2006). 또한 TIMSS의 과학 성취가 우수한 국가의 교육과정과 우리나라의 교육과정이나 교과서를 비교한 연구들이 진행되었다(김미영과 김경희, 2010; 김미영과 김경희, 2011; 김경희 등, 2009).

과학의 경우 국제 비교연구에서 우리나라는 지속적으로 성차가 나타나고 있는데, TIMSS의 8학년에서는 1995에서 28점, 2003에서 12점, 2007에서 8점으로 남학생의 점수가 높았다(김수진 등 2012; 신동희 등, 2006). TIMSS 4학년의 결과에서도 1995에서 13점, 2011년에서 8점으로 남학생의 성취가 높은 것으로 나타났다. TIMSS 2011에서 과학 성취도가 국제 2위와 3위에 해당하는 싱가포르와 핀란드에서는 성별 점수 차이가 각각 4점, 0점으로 그 차이는 유의하지 않아 우리나라가 성별 점수 차이가 큰 것으로 나타나(Martin *et al.*, 2012), 성별 차이에 대한 분석이 필요한 것으로 보인다.

가장 최근에 실시된 TIMSS 2011의 성취 결과를 활용한 연구는 아직 시작 단계에 있고, 우리나라 초등학교의 4학년 과학 성취도가 국제적인 수준에서 평가된 것은 1995년 이후 16년만의 시행이므로 TIMSS 2011의 결과를 이용한 분석은 현행 우리나라 초등학교 과학 교육에 대한 다양한 시사점을 도출해 낼 수 있을 것으로 판단된다. 이에 본 연구에서는 우리나라 초등학교 4학년 학생들의 과학 성취 특성을 알아보고자 학생들을 성취수준별, 성별로 구분하여 성취도에 나타난 특성을 분석하고자 한다. 또한 국가별로 공통된 교육과정에 근거하여 여러 국가의 전문가 집단에 의해 출제된 문항들을 통해 학생들의 성취 특성을 판단하는 것은 의의가 있으므로 TIMSS 2011의 공개 가능한 문항을 중심으로 학생들의 응답 특성을 분석하고자 한다. 공개 문항을 심층 분석함으로써 구체적 과학 내용 수준에서 나타나는 학생들의 응답 경향을 살필 수 있을 것으로 사료된다. 특히 우리나라의 성취가 미흡한 문항 및 성별 차이가 큰 문항들을 분석하여 학생들의 성취 특성을 알아보는 것은 우리나라 현행 교육을 진단하고, 좀 더 발전 가능한 방향으로 교육을 안내하는데 구체적인 자료가 될 것으로 판단된다.

## II. 연구 방법 및 내용

우리나라에서는 TIMSS 2011에서 4학년 학생들을 표집하기 위해 층화 군집표집을 사용하였다. 초등학교 4학년의 경우 총 150학교가 추출되었고, 평가 대상은 4,916명이었다(조지민 등, 2011a).

이 연구에서는 TIMSS 2011의 과학 성취수준별로 학생들의 성취도를 분석하기 위해 과학성취 점수가 625점 이상인 경우 수월 수준으로, 550점 이상 625점 미만인 경우 우수 수준, 475점 이상 550점 미만인 경우 보통 수준, 475점 미만인 경우 기초 수준으로 구분하였다. TIMSS 2011의 4학년 평가들에서는 과학의 내용영역을 물상과학, 생명과학, 지구과학의 세 가지 영역으로 구분하고 있고, 인지영역을 알기, 적용하기, 추론하기로 구분하여 제시하고 있어(Mullis *et al.*, 2009), 성취수준별로 내용영역별, 인지영역별 정답률을 분석하였고, 문항 유형에 따른 정답률도 분석하였다. TIMSS의 데이터 분석을 위해서는 데이터에 대한 정보를 제공하는 안내책을 참고하여 자료 분석을 실시하였다(Foy *et al.*,

2013).

성별 성취도 특성을 분석하기 위해서 학생들을 남학생과 여학생으로 구분하여 문항유형별, 내용영역별, 인지영역별로 정답률을 분석하였다. 이때 성별 정답률 차이를 확인하기 위하여 남녀 두 집단간  $t$ -검증을 실시하였다.

개별 문항 단위에서 나타난 학생들의 응답 특성을 분석하기 위해 공개문항별로 정답률을 국제 정답률과 비교 분석하였고, 공개문항별로 동일한 능력을 가지고 있음에도 불구하고 문항의 내용으로 인해 남학생과 여학생 간에 차이가 있는지 알아보기 위해 DIF(Differential Item Functioning) 분석을 실시하였다. DIF 분석은 각 문항이 동일 능력을 가진 두 집단에 있어서 어떤 특정 집단에게 유리하게 또는 불리하게 차별적으로 기능하고 있는지를 점검하는 방법이다. Mantel-Haenszel 방법(Holland & Thayer, 1998), 문항 난이도 모수 비교 방법(Camilli & Shepard, 1994), SIBTEST 방법(Shealy & Stout, 1993), 로지스틱 회귀분석 방법(Swaminathan & Rogers, 1990) 등의 다양한 방법들이 있으며, 각 자료에 적합한 방법을 사용하면 된다. 본 연구에서는 기존 연구(노언경 등, 2010)를 통해 균일적 차별기능문항과 비균일적 차별기능문항을 모두 추출하고, 비교적 많은 차별기능문항을 추출하고 있는 로지스틱 방법을 사용하여 분석을 실시하였다.

로지스틱 회귀분석 방법을 통해 집단( $G$ ), 검사 총점( $X$ ), 그리고 집단과 검사 총점의 상호작용( $GX$ )에 대한 회귀계수를 추정한다. 차별기능문항을 추출하기 위해 식(1)과 같은 3가지 모형을 설정하여 모형을 비교함으로써 유의도를 검증한다(Swaminathan & Rogers, 1990).

$$\begin{aligned} \text{완전모형: } \psi_i &= \delta + \tau_1 G_i + \tau_2 X_i + \tau_3 (G_i X_i) \\ \text{축소모형: } \psi_i &= \delta + \tau_1 G_i + \tau_2 X_i \\ \text{영모형: } \psi_i &= \delta + \tau_2 X_i \end{aligned} \quad (1)$$

축소모형과 완전모형을 비교하는 것은 집단과 검사 총점의 상호작용 회귀계수( $GX$ )의 차이를 나타내며, 이는 비균일적 차별기능문항을 추출하기 위함이다. 영모형과 축소모형을 비교하는 것은 집단( $G$ )의 회귀계수 차이를 나타내며, 균일적 차별기능문항을 추출하기 위한 것이다. 이는 식(2)와 같으며,  $G_1$ 과  $G_2$ 는  $\chi^2$  분포(자유도 1)를 따른다(Swaminathan & Rogers, 1990).

$$\begin{aligned} G_1 &= [-2L(\text{축소모형})] - [-2L(\text{완전모형})] \\ G_2 &= [-2L(\text{영모형})] - [-2L(\text{축소모형})] \end{aligned} \quad (2)$$

### III. 연구 결과 및 논의

#### 1. 성취수준별 분석에서 나타난 성취 특성

TIMSS 2011의 4학년 과학 성취도 결과에서 성취가 우수한 5개 국가의 성취수준별 학생 비율은 표 1과 같다. 우리나라는 과학 최상위권인 수월 수준 학생 비율은 29%로 33%로인 싱가포르에 이어 두번째로 높았다. TIMSS 2011의 수학 성취도 결과에서는 우리나라 수학의 수월 수준 학생 비율이 39%인 것과 비교했을 때 과학의 비율이 상대적으로 낮다는 것을 알 수 있다. 우수 수준 이상, 보통 수준 이상, 기초 수준 이상의 학생 비율은 각각 73%, 95%, 99%로 우리나라가 가장 높은 것으로 나타났다. 우리나라는 과학 성취가 가장 높았지만, 수월 수준의 비율이 싱가포르보다 낮았는데, 이전 주기인 TIMSS 2007의 중학교 2학년 과학 성취도 결과에서도 우리나라 수월 수준의 비율이 17%로 싱가포르의 32%, 대만 25%에 비해 낮은 것으로 나타났다(김경희 등, 2008c). 성취가 높은 수월 수준 학생들의 비율을 높이기 위한 노력이 필요한 것으로 보인다. 싱가포르는 중등학교부터 학생들의 수준을 구분하여 특화(special), 수월(express), 보통(normal) 과정을 운영하고, 능력에 따라 보다 심화된 교육을 받

표 1. TIMSS 2011 초4 과학 성취수준별 학생 비율

국가	국제 순위	수월 수준	우수 수준 이상	보통 수준 이상	기초 수준 이상
대한민국	1	29	73	95	99
싱가포르	2	33	68	89	97
핀란드	3	20	65	92	99
일본	4	14	58	90	99
대만	5	15	53	85	97

을 기회를 제공하고 있다(허경철 등, 2005). 초등학교부터 이러한 수준 구분에 대비하기 위한 교육을 하는 것이 싱가포르의 수월 수준의 비율이 높게 나타나는데 영향을 미친 것으로 판단된다.

TIMSS 2011 4학년 과학 성취도 결과를 문항유형, 내용영역, 인지영역에 따라 성취수준별로 나누어 정답률을 분석한 결과는 표 2와 같다.

문항 유형별로 분석한 결과, 네 수준의 학생들 모두 선다형이 구성형에 비해 정답률이 높게 나타나, 구성형보다 선다형에 더 쉽게 답한 것을 알 수 있다. 수월 수준에서 선다형과 구성형의 정답률 차이는 12.9%였으나, 기초 수준에서 선다형과 구성형의 정답률 차이는 18.8%로, 그 차이가 더 커서 성취수준이 우수한 학생보다 낮은 학생들이 선다형에 비해 구성형에 답하는 것을 더 어려워함을 알 수 있다. 구성형보다 선다형에서 학생들의 성취가 우수한 결과는 국가수준 학업성취도 결과와 TIMSS에 참여한 모든 국가에서 공통적으로 나타난 결과이다(김미경 등, 2011; 김미경 등, 2012; Martin et al., 2012). 문항 유형에 따른 과학 능력 추정의 효율성을 비교한 결과, 구성형이 선다형보다 난이도도 높고 변별도도 높아, 구성형이 학생들의 능력을 측정하는데 더 효과적인 것으로 보고되었다(박정, 2002). 선다형과 달리 구성형은 추측에 의해 답을 고를 수 없기 때문에 변별도와 난이도가 높게 나타난 것으로 판단된다.

내용영역별로 정답률을 비교한 결과, 생명과학의 정답률이 가장 낮았고, 물상과학의 정답률이 가장 높았다. 생명과학의 정답률이 낮은 것은 TIMSS의 성취가 우수한 싱가포르와 우리나라의 교육과정을 비교한 연구 결과에 비추어 볼 때 생명과학의 영역을 저학년에서 많이 다루고 있지 않기 때문인 것으로 판단된다(김미영과 김경희, 2010). 수준별로 내

용영역의 정답률을 비교한 결과, 수월 수준에서는 지구과학의 정답률이 가장 높았으나, 우수, 보통, 기초 수준에서는 물상과학의 정답률이 가장 높았다. 지구과학의 경우 수월 수준의 정답률이 가장 높았으나, 기초 수준의 정답률은 가장 낮아 수준별 차이가 가장 큰 영역으로 나타났다. 반면, 생명과학은 수준별 차이가 가장 작은 영역으로 나타났다. 공간 능력이 천체운동개념에 영향을 미친다고 보고되고 있는데(김상달 등, 2005; 이경훈과 임종욱, 2010), 성취수준이 낮은 학생들의 경우 공간 능력이 부족하여 지구과학 영역에서 성취가 낮게 나타난 것으로 판단된다.

인지영역별로 살펴본 결과, 전체적으로 알기와 추론하기의 정답률이 적용하기보다 높게 나타났다. 수준별로 살펴보면 수월 수준에서는 추론하기의 정답률이 가장 높았고, 기초 수준에서는 알기의 정답률이 가장 높아 수준이 높은 학생들은 추론하기에서 수준이 낮은 학생들은 알기에서 우수함을 알 수 있다. TIMSS 2011의 평가틀에 제시된 인지영역별 하위요소를 살펴보면 알기 영역의 경우는 회상/인식하기, 정의하기, 기술하기 등이 있고, 적용하기 영역의 경우 비교/대조/분류하기, 모델 사용하기, 관련짓기, 정보해석하기, 설명하기 등이 있으며, 추론하기 영역의 경우 문제분석하기, 통합/종합하기, 가설설정/예상하기, 설계하기, 결론도출하기, 일반화하기 등이 포함되어 있다(Mullis et al, 2009). 알기의 구성요소는 적용하기와 추론하기에 비해 상대적으로 좀 더 쉬운 요소로 구성되어 있어 인지수준이 낮은 학생들이 해결하는데 덜 어려움을 느꼈을 것으로 판단된다.

## 2. 성별 분석에서 나타난 성취 특성

TIMSS 2011의 4학년 학생들을 남, 여로 나누어

표 2. 초등학교 4학년 학생들의 성취수준별 정답률

수준	문항유형		내용영역			인지영역		
	선다형 M(SD)	구성형 M(SD)	생명과학 M(SD)	물상과학 M(SD)	지구과학 M(SD)	알기 M(SD)	적용하기 M(SD)	추론하기 M(SD)
수월	87.9(9.6)	75.0(12.9)	78.5(13.3)	84.7(12.2)	85.6(14.9)	83.1(12.1)	80.6(12.2)	83.8(18.5)
우수	73.9(12.3)	57.7(14.5)	61.8(15.5)	70.6(15.6)	69.3(20.0)	66.6(15.6)	66.3(14.1)	67.5(23.8)
보통	57.3(12.8)	39.7(14.3)	45.6(15.4)	53.1(16.3)	50.4(22.4)	50.1(15.6)	49.3(14.8)	49.3(24.9)
기초	39.9(14.6)	21.1(13.6)	31.2(16.6)	33.4(16.4)	30.8(21.8)	34.0(15.8)	31.0(16.8)	32.5(25.4)
계	72.9(17.3)	57.2(20.2)	61.8(20.0)	69.3(20.0)	68.3(24.2)	66.4(19.9)	65.3(19.0)	66.7(26.8)

표 3. 초등학교 4학년 학생들의 성별 정답률

성별	문항유형		내용영역			인지영역			전체
	선다형 M(SD)	구성형 M(SD)	생명과학 M(SD)	물상과학 M(SD)	지구과학 M(SD)	알기 M(SD)	적용하기 M(SD)	추론하기 M(SD)	
남	74.2(17.5)*	57.6(20.5)	62.3(20.2)	70.3(20.4)*	70.0(23.8)*	67.7(20.2)*	65.8(19.3)*	67.2(26.6)	66.6(16.7)*
여	71.5(17.0)	56.9(19.7)	61.2(19.7)	68.3(19.5)	66.7(24.6)	65.0(19.5)	64.6(18.7)	66.3(27.0)	64.8(15.9)

\* $p < .05$

문항 유형, 내용영역, 인지영역별 정답률을 분석한 결과는 표 3과 같다. TIMSS 2011 결과에 따르면 초등학교 4학년 과학의 성별 정답률은 남학생 66.6%, 여학생 64.8%로 그 차이는 유의한 것으로 나타났다.

문항유형별로 분석한 결과, 선다형 문항은 남학생 정답률 74.2%, 여학생 정답률 71.5%로 남학생이 2.7% 높은 것으로 나타났고, 이 차이는 유의수준 .05에서 유의한 것으로 나타났다. 구성형 문항은 남학생 정답률 57.6%, 여학생 정답률 56.9%로 성별 차이가 유의하지 않은 것으로 나타났다. 선다형에서 남학생의 성취가 우수한 결과는 2011 국가수준의 학업성취도 평가의 결과와 유사하였다(김미경 등, 2012). 구성형 문항의 경우 과학 지식을 활용하여 이유를 설명해야 하는 경우가 많이 있으므로 과학적 지식과 함께 문장을 서술하는 능력이 함께 평가된다고 할 수 있지만, 선다형 문항의 경우 과학 지식만 알고 있으면 답할 수 있다. 남학생들이 여학생들에 비해 과학 문항을 풀 때 과학 지식을 자주 활용한다고 보고되고 있는 반면, 언어와 관련된 수행능력에서는 여학생들이 더 우수하다고 보고되므로(신동희와 박병태, 2009; 조지민 등, 2011b) 과학 지식을 이용해서 답을 선택할 수 있는 선다형 문항에서 남학생들이 우수한 성취를 나타내었다고 할 수 있다.

내용영역의 성별 분석결과를 살펴보면, 생명과학에서 남학생의 정답률은 62.3%로 여학생에 비해 정답률이 1.1% 높았으나, 유의한 차이가 나타나지 않았다. ‘물상과학’ 영역에서 남학생의 정답률은 70.3%, 여학생 정답률은 68.3%로 남학생이 2.0% 높은 것으로 나타났으며, 그 차이는 유의수준 .05에서 유의한 것으로 나타났다. 지구과학에서 남학생 정답률은 70.0%, 남학생 정답률이 66.7%로 남학생이 3.3% 높은 것으로 나타났으며, 그 차이는 유의수준 .05에서 유의한 것으로 분석되었다. 이러한 결과에서 물상과학, 지구과학에서 남학생이 여학생에 비

해 성취가 우수함을 알 수 있다. 물상과학의 빛 관련 내용이나 지구과학의 태양계 관련 내용의 경우 시공간 지각력을 요하는 문항이 많다. 시공간 지각력을 요하는 문항에서 남학생의 성취가 우수한 것으로 보고되고 있으므로(신동희 등, 2006), 이러한 영향이 물상과학이나 지구과학에서 남학생의 성취가 높게 나타나는데 영향을 미쳤을 것으로 판단된다.

인지영역의 성별 분석 결과를 살펴보면, 알기는 남학생 67.7%, 여학생 65.0%로 남학생의 정답률이 2.7% 높았고, 그 차이는 유의한 것으로 나타났다. 적용하기는 남학생 65.8%, 여학생 64.6%로 역시 남학생의 정답률이 1.2% 높았으며, 그 차이는 유의한 것으로 나타났다. 추론하기의 경우 남학생 67.2%, 여학생 66.3%로 남학생과 여학생의 정답률 차이가 유의하지 않은 것으로 나타났다. 이러한 결과에서 알기와 적용하기에서 남학생의 성취가 우수함을 알 수 있다. 추론하기에 비해 알기와 적용하기에서 남학생의 성취가 우수한 결과는 이전 연구 결과의 경향성과 유사하였다(신동희 등, 2006; 정은영 등, 2006; 신동희와 박정, 2002). 과학 문제 풀이 과정에서 남학생들은 여학생들보다 더 빈번하게 과학 지식을 활용하는 것으로 나타나, 여학생들이 문제 풀이에 활용할 과학 지식이 상대적으로 부족한 것으로 보고되었다(신동희와 박병태, 2009). 알기나 적용하기의 경우 추론하기에 비해 과학 지식을 회상에서 푸는 문항이 더 많기 때문에, 이러한 결과가 나타난 것으로 판단된다.

### 3. 국제평균보다 성취가 낮은 공개문항 분석에서 나타난 성취 특성

공개문항 중에서 우리나라의 성취가 국제 평균보다 낮은 문항은 총 9문항으로 나타났다. 9문항 중에서 생명과학이 5개, 물상과학이 2개, 지구과학이 2개로 나타나, 생명과학이 성취가 낮은 문항이

가장 많았다. TIMSS 2011의 평가가 우리나라 초등 학교 4학년 학생들을 대상으로 2010년 말에 시행되었으므로 2010년의 초등학교 4학년 학생들에게 적용된 우리나라 과학과 교육과정은 3학년은 7차, 4학년은 2007 개정 교육과정에 해당된다. 문항의 내용이 교육과정에서 다루어지고 있는지 살펴본 결과 9문항 모두 4학년까지의 교육과정에서 다루어지지 않은 것들이었다(교육부, 1997; 교육인적자원부, 2007). TIMSS 2011의 과학 평가 문항 내용과 우리나라 교육과정에 제시된 과학 내용의 일치도는 33.7%로 국제 평균인 75.2%에 미치지 못하는 수준이었음에 비추어 볼 때, 학습하지 않은 많은 내용들 중에서도 8 문항에 대한 성취가 크게 더 낮다고 할 수 있다. 이 문항들의 내용을 분석하여 우리나라 학생들의 성취가 낮은 개념들의 내용영역별로 살펴본 결과는 다음과 같다. 본 연구에서 제시된 문항 내용은 국제 본부인 IEA에서 국제적으로 공통되게 공개된 문항을 사용한 것이다.

1) 생물과 관련된 분류 개념

S041003은 생물과 무생물을 분류하는 문항이고, S041163은 동물 중에서 파충류를 분류하는 문항인데, 이 두 문항에서 우리나라의 정답률이 국제 평균보다 낮았다. 즉, 우리나라 학생들은 생물과 관련된

분류 개념의 문항에서 성취가 낮음을 알 수 있다.


7차 과학 교육과정의 3학년과 2007 개정 교육과정의 4학년에서 생물과 무생물을 구분하는 내용을 구체적으로 다루고 있지 않았을 뿐 아니라, 파충류로 분류하는 내용도 전혀 다루고 있지 않아, 우리나라의 성취가 낮았다. 특히 파충류를 분류하는 문항인 S041163에서 우리나라 학생들은 오답인 1번에 대한 응답률이 66.5%로, 정답인 2번에 대한 응답률 25.1%보다 41.4% 높아, 우리나라 학생들의 과반수 이상이 개구리를 파충류로 잘못 알고 있음을 알 수 있다. 학생들이 오답으로 선택한 1번의 경우 동물 무리가 도마뱀, 개구리, 뱀으로 묶여 있는데, 개구리가 도마뱀이나 뱀의 피부 표면과 비슷한 피부 표면을 가지고 있어서 학생들은 개구리가 파충류라고 잘못 생각한 것으로 판단된다.

2) 자신의 경험에 근거한 잘못된 추론

S031325와 S041013A 문항은 학생들이 일상생활에서 경험한 사실들을 바탕으로 잘못된 추론을 통해 답을 얻은 문항으로 분류하였다.

S031325 문항에서는 체온이 높아진 이유를 알고 있는지 평가하고 있는데, 우리나라 학생들 중 체온이 높아진 이유를 아파서 열이 난 것과 관련지어 서술한 비율이 35.7%였으나, 외부 온도가 높아 체온

**S041003**  
그림은 연못을 나타낸 것입니다.



그림에서 생물 3가지, 무생물 3가지를 찾아 아래 빈칸에 쓰시오.

생물		무생물	
1.		1.	
2.		2.	
3.		3.	

	대한민국	국제
1.	46.5	47.1
2.		
3.		


**S041163**  
동물의 무리 중 파충류로만 묶인 것은 어느 것입니까?  
① 도마뱀, 개구리, 뱀    ② 거북, 도마뱀, 악어  
③ 문어, 달팽이, 거북    ④ 게, 지렁이, 뱀

①	②	③	④	무응답	대한민국	국제
66.5	25.1	2.6	5.0	0.8	25.1	39.6

**S031325**  
사람의 정상 체온은 약 37℃입니다. 어느 날 아침, 잠에서 깬 진호는 체온을 재어 보았습니다. 진호의 체온은 40℃였습니다. 진호의 체온이 정상 체온보다 높아진 이유를 한 가지 쓰시오.

대한민국	국제
35.7	38.9

**S041013A**  
그림은 호랑이 머리뼈와 쥐 머리뼈를 나타낸 것입니다.



호랑이 머리뼈                      쥐 머리뼈

호랑이는 송곳니라고 하는 매우 큰 이빨을 가지고 있습니다. 쥐는 앞니라고 하는 매우 큰 이빨을 가지고 있습니다. 호랑이와 쥐는 서로 다른 종류의 먹이를 먹습니다. 호랑이는 송곳니를 무엇을 하는 데 사용할까요?

대한민국	국제
23.5	31.3

이 높아졌다고 답한 비율이 37.6%로 정답률보다 높게 나타났다. 학생들은 외부 온도와 체온의 관계에 대해 오개념을 가지고 있음을 알 수 있다. 2007 개정 교육과정에서는 5학년 ‘우리의 몸’ 단원에서 건강을 각 기관의 기능과 관련지어 학습하도록 하여 건강과 관련된 내용을 다루고 있었지만, 시험을 치른 학생들은 4학년이므로 건강과 관련된 내용을 학습하지 않아 체온이 높아진 이유를 병에 걸린 것과 연관지어 생각하지 못한 것으로 판단된다. 외부 온도가 높은 더운 날에 학생들은 매우 더운 것을 경험하였기 때문에, 이때의 경험을 토대로 더운 날에는 체온이 올라간다는 잘못된 추론을 한 것으로 판단된다.

S041013A는 호랑이가 송곳니를 먹잇감을 찢거나 죽일 때 사용하는 것을 알고 있는지 평가하고 있는데, 우리나라 학생들의 정답률은 23.5%로 국제 평균보다 낮았다. 우리나라의 경우 호랑이가 먹이를 먹을 때 송곳니를 사용한다고 응답한 학생들이 많았는데, 이와 관련되어 잘못된 개념을 가지고 있음을 알 수 있다. 일반적으로 사람의 송곳니는 음식을 먹을 때 사용하므로 학생들은 이러한 자신의 경험과 관련지어 호랑이의 송곳니 역할을 잘못 추론했을 것으로 판단된다. 7차 교육과정의 4학년 ‘여러 가지 동물의 생김새’ 단원과 2007 개정 교육과정의 3학년 ‘동물의 세계’ 단원에서는 여러 가지 동물의 생김새와 특징을 배우고, 이를 사는 곳과 연관지어 학습한다. 이때 동물들의 이빨을 비교하는 내용은 거의 다루지 않는다. 또한 우리나라 학생들은 3학년은 7차 교육과정, 4학년은 2007 개정 교육과정의 적용하게 학습하였으므로 동물의 생김새와 특징에 대한 내용을 학습하지 않아 정답률이 낮게 나타났다.

### 3) 심장의 역할

S051057은 사람 몸의 각 부분이 하는 일을 보고, 이러한 일을 담당하고 있는 몸의 각 부분을 알고 있는지 평가하고 있는데, 우리나라의 정답률은 23.4%로 국제 평균 정답률인 50%에 비해 크게 낮았다. 학생들의 응답을 보면 심장만 빠르게 답한 경우 1.1%, 뇌만 빠르게 답한 경우가 64.5%로 우리나라 학생들은 심장의 기능을 잘 모르고 있는 것으로 나타났다. 2007 개정 과학과 교육과정에서는 5학년 ‘우리의 몸’ 단원에서 이 내용을 다루므로 우리나

#### S051057

표는 사람 몸의 여러 부분들이 하는 일을 세 가지 나타낸 것입니다. 각각의 일을 하는 몸의 부분은 무엇인지 옆의 빈 칸에 쓰시오. 첫 번째 칸은 예로 써 놓은 것입니다.

하는 일	몸의 부분		
몸을 지탱한다.	뼈대	대한민국	국제
피를 온몸으로 내보낸다.			
생각을 할 때 사용된다.			
		23.4	50.0

라 4학년 학생들은 사람 몸의 각 부분이 하는 일에 대한 내용을 학습하지 않아 심장이나 뇌와 같은 기관의 이름을 적지 못한 것으로 보인다. 본 문항에 대해 학생들이 쓴 답안을 분석한 결과, 피를 온몸으로 내보내는 곳을 심장이라고 쓰지 않고 혈관이라고 쓴 경우가 많았다. 학생들은 심장대신 혈관을 피를 온몸으로 내보내는 곳으로 잘못 알고 있는 것으로 판단된다.

### 4) 타거나 빛을 내는 물질 고르기

S031421은 타는 물질을 고르는 문항이고, S041120은 빛을 내는 물체를 고르는 문항인데, 이 두 문항에서 우리나라의 정답률이 국제 평균보다 낮았다.

S031421 문항을 살펴보면 문항에 제시된 물질 중에서 타는 물질로 나무와 휘발유를 선택한 학생 비율이 52.1%로 국제 평균 정답률보다 8.5% 낮았다. 물, 모래, 공기가 타지 않는 물질임에도 우리나라 학생들은 이 물질들이 탄다고 응답한 비율이 높

#### S031421

다음 물질 중 어떤 것은 타고, 어떤 것은 타지 않습니다. 타는 물질을 골라 옆의 네모 안에  하시오. (한 가지 이상 고를 수 있습니다.)

- 물
- 나무
- 모래
- 휘발유
- 공기

대한민국	국제
52.1	60.6

#### S041120

스스로 빛을 내는 물체 두 가지는 무엇입니까?

- ① 촛불과 달
- ② 달과 거울
- ③ 태양과 촛불
- ④ 거울과 태양

①	②	③	④	무응답	대한민국	국제
4.6	11.5	38.2	44.3	1.4	38.2	50.0

았다. 특히, 불을 끌 때 사용하는 물질인 물이나 모래가 탄다고 응답한 학생들의 비율이 14.5%나 되어 우리나라 학생들이 타는 물질과 타지 않는 물질을 잘 구분하지 못하고 있음을 알 수 있다. 연소와 관련된 내용은 7차의 6학년 ‘촛불 관찰’과 2007 개정의 6학년 ‘연소와 소화’ 단원에서 다루어지므로, 이와 관련된 내용을 학습하지 않아 성취가 낮게 나타났다.

S041120에서는 빛을 내는 물체를 고르도록 하고 있는데, 우리나라의 정답률은 38.2%로 국제 정답률보다 11.8%나 낮았다. 학생들의 응답을 보면 태양과 촛불을 고른 비율보다 거울과 태양을 고른 비율이 더 높아, 많은 학생들이 거울이 빛을 내는 물체라는 오개념을 가지고 있음을 알 수 있다. 2007 개정 과학과 교육과정의 초등학교 3학년 ‘빛의 직진’ 단원에서는 스스로 빛을 내는 물체에 대해 다루고 있지 않으며, 4학년 ‘태양계와 별’ 단원에서 태양이 지구의 에너지원임을 학습하고, 초등학교 6학년 ‘연소와 소화’에서 촛불을 관찰하게 된다. 따라서 초등학교 4학년 학생들은 이 문항에 대하여 어려움을 겪은 것으로 판단된다. 초등학생들이 스스로 빛을 내는 일차 광원과 빛을 반사하여 빛을 내는 이차 광원을 모두 빛을 내는 물질이라고 생각하는 오개념을 가지고 있다는 이전의 결과와 유사한 결과임을 알 수 있다(최병순, 1993).

5) 공전과 자전 개념 혼동

S031275는 자전으로 인해 나타나는 현상을 S041208은 자전의 주기를 알고 있는지 평가하고 있는데, 자전과 관련된 개념을 평가하는 이 두 문항에서 우리나라의 정답률이 낮았다. 자전으로 인해 낮과 밤이 나타나는 것을 알고 있는 학생의 비율이 20.3%로 국제 평균 정답률보다 낮았다. 오답에 대한 반응을 보면 자전이 아닌 공전으로 인해 낮과 밤이 나타난다고 응답한 학생들이 63.9%로 매우 높게 나타났다. 우리나라 학생들의 경우 자전의 주기가 24시간이라고 답한 비율이 51.5%로 나타났고, 1년이라고 답한 비율이 26.8%로 높았다. 2007 개정 교육과정의 5학년 ‘지구와 달’ 단원에서 지구상에 낮과 밤이 생기는 이유를 지구의 자전과 관련지어 이해하는 내용을 다루고, 지구 표면 위치에 따라 같은 순간에도 낮인 곳과 밤인 곳이 동시에 존재한다는 것을 다루므로, 우리나라 4학년 학생들은 자전과 관련된

S031275

지구에서 낮과 밤이 나타나는 이유를 옳게 설명한 것은 어느 것입니까?

- ① 태양이 지구 주위를 돈다.
- ② 지구가 태양 주위를 돈다.
- ③ 지구가 자신의 축을 중심으로 돈다.
- ④ 태양이 자신의 축을 중심으로 돈다.

①	②	③	④	무응답	대한민국	국제
14.0	63.9	20.3	1.2	0.2	20.3	34.6

S041208

지구는 자신의 축을 중심으로 얼마나 자주 회전합니까?

- ① 12시간에 한 번
- ② 24시간에 한 번
- ③ 한 달에 한 번
- ④ 일 년에 한 번

①	②	③	④	무응답	대한민국	국제
14.8	51.5	6.5	26.8	0.1	51.5	53.0

내용을 학습하지 않아 성취가 낮게 나타났음을 알 수 있다. 두 문항의 응답에서 알 수 있듯이 우리나라 학생들은 자전의 개념과 공전의 개념을 혼동하고 있음을 알 수 있다. 이전의 연구 결과에서 학생들은 낮과 밤이 생기는 이유에 대하여 지구가 돌기 때문이라고 애매하게 답한 학생들이 많았고(김효남, 1990), 태양이 뜨고 지는 이유에 대해 지구의 공전과 달의 공전 때문으로 오인하는 경우도 많았다(김순 등, 1999) 학생들은 공전이나 자전이 모두 지구가 도는 것이므로 돈다는 현상 자체가 유사하여 두 개념을 혼동하는 것이라고 판단된다. 자전과 공전의 정의를 각각 알고 있더라도 실제로 지구의 자전 개념을 바탕으로 하는 공간 변화의 개념은 수업을 통한 지식으로는 쉽게 형성되기 어렵다(김순 등, 1999). 다양한 모형이나 역할놀이 등을 통해 공전과 자전 개념이 혼동되지 않고 잘 이해될 수 있도록 도와주어야 할 것으로 판단된다.

4. 공개문항의 차별기능문항 분석에서 나타난 성별 응답 특성

공개문항에서 학생들의 성별 성취 차이가 나는 문항들을 살펴보기 위하여 DIF 분석을 실시한 결과는 표 4, 5와 같다. 표 4는 남학생에게 유리한 문항, 표 5는 여학생에게 유리한 문항을 제시하였다.

S041180의 경우, G1 값이  $\alpha=.05$  수준에서 자유도 1인 카이제곱 분포의 임계치인 3.84를 넘어 영가설을 기각하므로 비균일적 DIF가 된다. 또한 나머지 문항의 경우 모두 G1 값은  $\alpha=.05$  임계치를



표 4. DIF 분석 결과 남학생의 성취가 우세한 문항

문항	내용영역	문항유형	인지영역	G2 (영모형- 축소모형)	G1 (축소모형- 완전모형)	DIF 유형	성별회귀계수	유의확률	교육과정일치
S031044	지구과학	구성형	추론하기	7.824	1.833	균일	-.546	.005	No
S051156	지구과학	선다형	알기	13.953	.885	균일	-.641	.000	Yes
S041117	물상과학	선다형	알기	6.356	1.358	균일	-.730	.014	No
S041120	물상과학	선다형	알기	9.034	1.431	균일	-.512	.003	No
S041163	생명과학	선다형	적용하기	5.308	.387	균일	-.432	.022	No
S041208	지구과학	선다형	알기	4.157	1.207	균일	-.329	.042	No
S041180	생명과학	선다형	적용하기	4.525	3.861	비균일	-.557	.035	No
S031299	물상과학	구성형	알기	4.574	.591	균일	-.372	.033	No
S031281	생명과학	선다형	알기	4.597	.109	균일	-.493	.033	No
S031077	물상과학	선다형	적용하기	4.389	2.062	균일	-.464	.037	No
S031311	물상과학	선다형	알기	4.504	.090	균일	-.470	.035	No

표 5. DIF 분석 결과 여학생의 성취가 우세한 문항

문항	내용영역	문항유형	인지영역	G2 (영모형- 축소모형)	G1 (축소모형- 완전모형)	DIF 유형	성별회귀계수	유의확률	교육과정일치
S031068	물상과학	선다형	추론하기	4.348	.884	균일	.538	.039	Yes
S051119	물상과학	구성형	추론하기	8.803	1.413	균일	.578	.003	Yes
S031391Z	지구과학	구성형	알기	10.574	.716	균일	.605	.001	Yes
S041048	물상과학	구성형	추론하기	6.124	.994	균일	.459	.014	Yes
S031204	물상과학	구성형	적용하기	8.332	.248	균일	.593	.004	Yes

넘지 못하는 한편, G2 값이 임계치 3.84를 넘고 있으므로 균일적 DIF가 된다.

성별 차이에 대해서 살펴보면, 남학생은 0, 여학생은 1로 더미코딩 하였기 때문에 양수의 성별 회귀계수를 보이는 문항은 여학생에게 유리하고, 음수의 성별 회귀계수를 가진 문항은 남학생에게 유리하다. 남학생에게 유리한 문항은 전체 공개 가능한 74개의 문항 중에서 총 13개로 S041180을 제외하고는 모두 균일한 유형으로 나타났다. 여학생에게 유리한 문항은 5문항으로 모두 균일한 유형으로 나타났다. 남학생에게 유리한 문항과 여학생에게 유리한 문항의 내용을 분석한 결과, 두드러진 특징이 나타났다.

1) 교육과정에 문항 내용이 포함되어 있는지의 여부에 따른 특성

성별 차이가 나는 문항을 중심으로 문항 내용이 교육과정에 포함되어 있는지를 살펴보았다. 그 결

과, 남학생에게 유리한 문항의 경우에는 S051156을 제외하고는 모든 문항이 교육과정에 포함되어 있지 않은 것으로 나타났고, 여학생에게 유리한 문항의 경우에는 모든 문항이 교육과정에 포함되어 있는 것으로 나타났다. 교육과정에 포함되어 있지 않아 학생들이 정규 교육과정에서 직접적으로 학습하지 않은 내용에서 남학생의 성취가 더 우수한 것을 알 수 있다. 이러한 결과는 이전에 학습하지 않은 내용에서 남학생의 성취가 더 높게 나타났다는 결과와 유사함을 알 수 있다(신동희 등, 2006). 남학생들이 여학생들에 비해 과학 관련 서적이거나 학습 자료들을 더 많이 접했기 때문에, 이러한 결과가 나타났을 것이라고 판단된다. 국가에서 실시하는 학업 성취도 평가는 모두 배운 내용을 평가 하는데, 2011 국가수준 학업성취도 평가의 결과, 6학년의 경우 남학생이 201.77, 여학생이 202.76으로 여학생의 점수가 조금 더 높았다(김미경 등, 2012).

2) 인지영역과 문항유형에 따른 특성

인지영역과 문항유형에 따른 문항수별로 남학생이 우수한 문항수와 여학생이 우수한 문항수의 비율을 나타내면 결과는 표 6과 같다.

인지영역별로 살펴보면 알기에 해당하는 문항에서 남학생이 우수한 문항이 7개로 가장 많았고, 추론하기에 해당하는 문항에서 여학생이 우수한 문항이 3개로 가장 많았다. 공개 문항 중에서 알기의 선다형 문항이 총 20개인데, 이 중에서 여학생은 한 문항도 우세한 문항이 없었고, 30%에 해당하는 6문항에서 남학생의 성취가 우세한 것으로 나타나, 남학생들이 알기의 선다형에 해당하는 문항에서 높은 성취를 나타낸다고 할 수 있겠다. 남학생의 성취가 우수했던 알기의 선다형에 해당하는 문항은 다음에 제시된 S031311처럼 문항의 길이가 짧고 문항을 읽고 바로 과학 지식을 회상에서 풀 수 있는 문항이 대부분이었다. 과학 문항을 풀 때 남학생이 여학생에 비해 과학 지식을 더 많이 떠올린다는 이전의 연구 결과(신동희와 박병태, 2009)가 본 연구의 결과와 일맥상통한다고 할 수 있겠다.

남학생의 성취가 우수한 구성형 문항은 S031299와 S031044였는데, S031299는 햇빛이 여러 가지 색으로 이루어져 있음을 볼 수 있는 현상을 쓰는 것으로 무지개, 해가 질 때, 비눗방울 등으로 답하는 문항이고, S031044는 천체모형을 보고 태양, 지구, 달을 각각 바르게 찾아 표시하는 문항이다. 반면, 여학생이 우수한 구성형 문항은 모두 4문항이었는데,

표 6. 인지영역과 문항유형에 따른 남학생과 여학생의 우수 문항 비교

인지영역	문항유형	문항수	남학생	여학생
			우수 문항수	우수 문항수
알기	선다형	20	6(30.0%)	0( 0.0%)
	구성형	17	1(11.8%)	1( 0.0%)
적용하기	선다형	22	3(13.6%)	0( 0.0%)
	구성형	16	0( 0.0%)	1( 6.3%)
추론하기	선다형	6	0( 0.0%)	1(16.7%)
	구성형	5	1(20.0%)	2(40.0%)

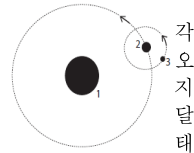
S031311

손으로 잡고 있던 물체를 놓았을 때 물체를 땅으로 떨어지게 하는 것은 무엇입니까?

- ① 자기력                      ② 중력
- ③ 공기 저항                 ④ 손이 미치는 힘

S031044(남학생의 성취가 우수한 구성형 문항)

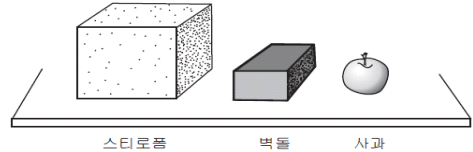
그림은 지구, 달, 태양을 나타낸 것입니다. 각 천체는 번호로 표시되어 있습니다. 화살표는 각 천체가 움직이는 방향을 나타냅니다.



각 천체를 나타내는 번호(1, 2, 3)를 쓰시오  
 지구를 나타내는 번호: \_\_\_\_\_  
 달을 나타내는 번호: \_\_\_\_\_  
 태양을 나타내는 번호: \_\_\_\_\_

S041048(여학생의 성취가 우수한 구성형 문항)

재훈이의 선생님은 그림과 같이 책상 위에 세 가지 물체를 놓았습니다. 선생님은 물체들을 부피 순으로 늘어놓았습니다.



재훈이는 부피가 큰 물체가 더 무거울 것이라고 생각했습니다. 재훈이의 생각에 동의합니까?(하나만 고르시오.)

- 예
  - 아니요
- 그 이유를 설명하십시오.

데, 이 문항들은 모두 방법이나 이유 등을 설명해야 하는 문항이 대부분이어서 같은 구성형에서도 남학생이 우수한 경우와 여학생이 우수한 경우에 차이가 나타났다. 이유나 방법들을 설명해야 하는 구성형 문항에서 여학생의 성취가 우수한 것은 언어와 관련된 수행능력에서 여학생들의 성취가 더 우수하기 때문에 이러한 결과가 나타난 것으로 판단된다.

3) 문항에서 실험 상황의 제시 여부에 따른 특성

여학생이 우세한 총 5문항 중에서 3문항이 모두 실험과 관련된 문항이었다.

S031068은 실험 설계를 보고 알아보고자 하는 물체를 파악하는 문항이고, S051119는 실험 결과를 해석하는 문항이며, S031204는 실험 결과를 예측하고 설명하는 문항이다. 특히 S031068이나 S051119는 과학적 지식과 크게 관련 없이 문제에 주어진 실험 상황을 잘 해석하기만 하면 풀 수 있는 문항인데, 이러한 실험의 해석과 관련된 문항에서 여학생의 성취가 높다고 할 수 있겠다. 남학생이 우세한 13문항 중에 단 한 문항이 실험 상황으로 제시되었는데 이 문항은 전기회로가 제시되어 있고 전기회로에 연결했을 때 불이 켜지는 물체는 무엇인가를 질

**S031068**

미란이는 소금과 물을 사용하는 실험을 설계하였습니다. 미란이의 실험 결과는 표와 같습니다.

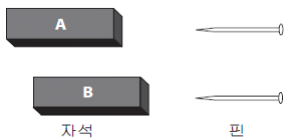
물에 녹은 소금의 양	물의 부피	물의 온도	혼합물을 저어 주었는가?
15 g	50 mL	25°C	예
30 g	100 mL	25°C	예
45 g	150 mL	25°C	예
60 g	200 mL	25°C	예

미란이가 이 실험에서 알아보고자 하는 것은 무엇입니까?

- ① 서로 다른 부피의 물에 소금이 얼마나 많이 녹을까?
- ② 서로 다른 온도에서 소금이 얼마나 많이 녹을까?
- ③ 젓기를 많이 하면 소금이 얼마나 빨리 녹을까?
- ④ 젓기를 적게 하면 소금이 얼마나 빨리 녹을까?

**S051119**

민주는 똑같은 금속 편 두 개와 자석 두 개(A와 B)를 가지고 있습니다. 민주는 책상 위에서 자석이 편을 끌어당길 때까지 자석 A를 밀었습니다. 그 다음, 민주는 책상 위에서 자석이 편을 끌어당길 때까지 자석 B를 밀었습니다.



민주는 자석 A가 15 cm 떨어져 있는 편을 끌어당기고, 자석 B는 10 cm 떨어져 있는 편을 끌어당긴다는 것을 알았습니다. 영수는 두 자석의 세기가 같다고 말합니다. 영수의 의견이 옳을까요? (하나만 고르시오.)

- 예
  - 아니요
- 그 이유를 설명하십시오.

문하고 있어 전류가 흐르는 물체를 알고 있는지 평가하고 있어 여학생이 우수한 문항과 차이가 있었다. 여학생이 우수한 실험 관련 문항의 경우 과학 지식을 적용해서 풀어야 하는 문항이기 보다는 문항에서 주어진 실험 상황을 해석하기만 하면 풀 수 있는 문항이다. 이전 연구에서도 과학 지식 자체를 묻는 문항보다 과학 하는 과정에 필요한 추론과 분석 문항에서 여학생들이 우월하다는 결과가 보고 되었는데(신동희 등, 2006; 신동희와 박정, 2002), 이러한 결과와 맥락을 같이 한다고 볼 수 있다.

**4) 평가하는 개념에 따른 응답 특성**

물체를 떨어뜨리게 하는 힘을 질문하는 S031311과 돛단배를 움직이게 하는 힘을 질문하는 S041117에서 남학생의 성취가 우수한 결과에서 남학생이 힘과 관련된 개념에서 우수한 성취를 나타낸다고

할 수 있다. 다음으로는 태양계 내 천체의 운동과 관련된 개념을 들 수 있다. S031044는 태양계에서 지구, 태양, 달의 운동을 알고 있는지 평가하고 있으며, S041208은 지구의 자전 운동의 주기를 알고 있는지 평가하고 있는데, 이 두 문항에서 남학생의 성취가 높은 결과에서 남학생이 태양계 천체의 운동 개념에서 높은 성취를 나타낸다고 할 수 있다. TIMSS 2003과 PISA 2006에서 남학생의 성취가 우수한 문항들을 분석한 결과에서도 태양과 달의 운동과 그 특징에 대한 내용에서 남학생이 우수한 성취를 나타낸 이전의 결과와 같은 결과임을 알 수 있다(김경희 등, 2008b). 중학교 학생들의 TIMSS의 성취 결과를 분석한 연구에서도 남학생들이 여학생에 비해 지구와 달의 운동과 관련된 천문학 문항의 성취가 유의하게 높게 나타났는데(신동희와 박정, 2002), 이러한 결과는 초등학교에서도 일관되게 나타났다. 공간 지각력의 성별 차이가 이러한 결과를 나타낸 것이라고 할 수 있겠다.

공개문항 중에서 정답보다 오답에 대한 응답 비율이 두드러지게 더 높은 문항은 S041163, S041120이었는데, 이 두 문항에서 남학생의 성취가 높았다. S041163 문항에서 개구리를 파충류로 답한 비율이 66.5%로 매우 높았고, S041120 문항에서 거울이 빛을 내는 물체라고 답한 학생의 비율이 44.3%로 높아 각각 정답에 대한 반응 비율인 25.1%, 38.1%보다 높았다. 오개념이 두드러지게 나타났던 파충류에 대한 개념과 빛을 내는 물체에 대한 개념에서 여학생보다 우수한 성취를 나타낸다고 볼 수 있다. 그 외에도 먹이사슬 관련 개념(S041180), 전류를 흐르게 하는 물체와 관련된 개념(S031077) 및 화석과 관련된 개념(S031044)에서 남학생의 성취가 높았다. 또한 햇빛이 여러 가지 색으로 되어 있다는 것을 알 수 있는 현상(S031299)과 새가 알을 품는 이유(S031281)에 대한 내용에서도 남학생의 성취가 높았다. 남학생은 여학생에 비해 동물에 대해 관심이 많아, 파충류나 먹이사슬과 관련된 개념 및 새가 알을 품는 이유의 성취가 높게 나타난 것으로 판단된다. 빛을 내는 물체와 햇빛이 여러 가지 색으로 이루어져 있음을 알 수 있는 무지개와 관련된 현상은 빛과 관련된 광학현상으로, 광학현상은 주변 세계를 이해하려는 자연스러운 시도로 인해 정규학습이 있기 전에 상당한 지식이 자연스럽게 이미 형성되므로 견고한 대안 개념이 형성되기도 한

다고 지적하였다(Galili & Hazan, 2000). 남학생들은 여학생들보다 주변 세계를 과학적으로 이해하려는 시도가 높은 것으로 보이며, 이로 인해 이러한 개념에서 남학생의 성취가 높은 것으로 판단된다.

여학생의 성취가 높게 나타난 두드러진 개념은 용해도 관련 개념이었다. S031068과 S031204는 각각 용해도 실험에서 알아보고자 하는 것, 입자의 크기와 용해속도에 대한 개념을 평가하고 있는데, 이 두 문항에서 여학생의 성취가 높은 결과에서 여학생이 용해도와 관련된 개념에서 우수한 성취를 나타낸다고 할 수 있다. 그 외에 자기력의 세기와 관련된 개념(S051119), 물체의 부피와 질량의 관계에 관련된 개념(S041048), 물을 낭비하지 않는 방법에 대한 내용(S031391Z)에서 여학생의 성취가 높았다. 이 개념들은 4학년까지의 교육과정에서 모두 다루어졌으며, 용해도와 자기력은 실험 상황으로 제시되었는데, 다른 상황으로 제시되었을 때에도 이 개념에서 여학생들의 성취가 높은가에 대한 연구가 필요하다고 판단된다. 물을 낭비하지 않는 방법을 쓰는 문항에서 여학생의 성취가 높았는데, 2007 공개문항인 S042153은 재활용이 중요한 이유를 쓰도록 하는 문항에서도 여학생의 성취가 높았다(김경희 등, 2008a). 일상생활에서 물을 아껴 쓰거나 재활용을 하는 것 등과 같이 생활에서 지켜야 하는 일과 관련된 것에서 여학생의 성취가 높음을 알 수 있다.

#### IV. 결론 및 제언

TIMSS 2011의 초등학교 4학년 성취도를 성취수준별로 분석한 결과, 우리나라는 과학 성취도가 국제적으로 가장 높은 수준이었지만, 최상위권인 수월 수준의 학생 비율이 싱가포르에 비해 낮은 것으로 나타났다. 이러한 결과는 우리나라의 과학 교육에서 성취가 우수한 학생들을 위한 수월성 교육이 지속적으로 강조되고 지원될 필요가 있음을 시사한다고 할 수 있다. 형평성을 고려한 교육도 중요하지만, 우수한 인재의 능력을 신장시키기 위한 노력도 국가경쟁력을 기르기 위한 차원에서 중요하다고 할 수 있겠다. 수월성 교육도 중요할 뿐 아니라, 과학에서 흥미가 있고 성취가 우수한 학생들이 과학 관련 진로에 대한 적절한 탐색이 이루어져 관련 분야에 진출할 수 있도록 과학과 관련된 진로교

육도 지속적으로 함께 이루어져야 하겠다.

TIMSS 2011의 내용영역별 성취를 비교한 결과, 생명 과학의 성취가 다른 영역에 비해 가장 낮게 나타났고, 성취수준별로 과학 성취도 차이를 비교한 결과, 생명과학이 성취수준별 차이가 가장 낮은 영역으로 나타났다. 우리나라는 TIMSS의 성취가 우수한 국가와 비교했을 때 저학년에서 생물 영역의 내용을 많이 다루고 있지 않은데, 본 연구의 결과에 비추어 볼 때 생명 과학 영역은 성취 수준별 차이가 크지 않으므로 인지 수준이 낮은 저학년에서 많이 다루어질 필요가 있겠다. 저학년에서는 성취 수준별 차이가 작은 내용영역 위주로, 고학년에서는 성취수준별 차이가 큰 내용영역 위주로 과학 교육과정을 구성할 필요가 있겠다.

공개문항을 분석하여 우리나라의 성취가 국제평균보다 낮은 문항을 살펴본 결과, 우리나라는 생물의 분류와 관련된 개념에서 성취가 낮았고, 자신의 경험에 근거하여 잘못된 추론을 하여 정답이 낮게 나타난 문항들이 있었다. 또한 심장의 역할에 대해 잘 알지 못하였고, 타거나 빛을 내는 물질을 잘 고르지 못하였으며, 공전과 자전 개념을 혼동하고 있는 것으로 나타났다. 국제 평균보다 낮은 문항들에서는 정답률보다 오답에 대한 응답비율이 더 높아, 우리나라 학생들이 가지고 있는 두드러진 오개념이 나타났는데, 개구리를 파충류로 알고 있는 것과 거울을 빛을 내는 물질로 알고 있는 것, 더운 날 사람의 체온이 높다는 오개념이었다. 오개념은 수업이 이루어진 후에도 잘 교정되지 않는 견고성이 있으므로, 본 연구의 결과에서 나타난 오개념들에 대한 교정이 잘 이루어질 수 있도록 수업이 구성되어야 하겠다. 또한 생물영역에서 생물과 무생물의 분류에 대한 내용이 우리나라 교육과정에서 구체적으로 다루어져야 하겠으며, 간단한 생물의 분류에 대한 내용이나 우리 몸의 각 부분의 간단한 역할 등은 3~4학년군에서 다루어질 필요가 있겠다. 지구과학 영역에서는 공전과 자전을 가르칠 때 두 개념을 혼동하지 않고 잘 구분하여 알 수 있도록 지도해야 하겠다.

TIMSS 2011의 초등학교 과학 성취도의 성별 차이를 분석한 결과, 남학생이 여학생에 비해 과학 성취가 유의하게 높은 것으로 나타났고, 내용영역별로는 물상과학과 지구과학에서 남학생의 성취가 유의하게 높았다. 공개문항을 대상으로 남학생과

여학생의 성취가 우수한 문항의 특성을 분석한 결과, 남학생과 여학생이 우수한 문항에서 두드러진 차이가 나타났다. 본 연구의 공개 문항 분석 결과에서 나타난 성별 차이의 특징은 성차를 해소하기 위한 수업 전략을 구성하는데 활용될 수 있을 것으로 판단된다. 남학생의 경우, 이유나 방법을 설명하는 구성형 문항이나 실험 상황을 해석하는 문항에 잘 답할 수 있도록 이유를 잘 서술하는 방법이나 실험 상황을 해석하는 과정에 대한 구체적인 지도가 필요하겠다. 여학생의 경우, 남학생보다 두드러지게 나타난 오개념에 대해 교정수업이 더 집중적으로 이루어져야 하겠다. 또한 힘에 대한 개념이나 천체의 운동에 대해 지도할 때 여학생들이 잘 이해할 수 있도록 적절한 수업 전략을 활용하여 지도하여야 하겠다. 천체의 운동에 대해 지도할 때 공간 지각력이 필요한 경우가 많으므로 이와 관련하여 여학생들이 잘 이해할 수 있도록 컴퓨터 시뮬레이션이나 모형실험 등을 통해 관련 개념을 잘 이해시키는 수업이 필요하겠다.

학교에서 배우지 않은 내용에 대해서는 남학생들의 성취가 우수한 결과에서 남학생들은 과학 관련된 내용에 대해 더 많이 접하고 있음을 시사하며, 이는 과학 관련 내용에 더 관심이 있다고 할 수 있겠다. 반면, 배운 내용과 관련된 문항에 대해 여학생의 성취가 더 우수했다는 것은 교실수업에서 여학생들이 더 수업에 대한 내용을 잘 받아들이고 있음을 시사한다고 할 수 있다. 여학생들에게는 과학에 대한 관심을 가질 수 있도록 과학관 견학이나 과학 관련 도서 등을 더욱 많이 접할 기회를 제공해야 하겠다. 남학생들에게는 과학 수업 태도가 바른지, 수업에서 배운 내용을 잘 받아들여 이해하고 있는지 등에 대해 스스로 점검하도록 할 필요가 있겠다.

## 참고문헌

박영순, 정은영(2007). TIMSS 2003 지구과학 영역 문항 및 성취도 분석. 한국지구과학학회지, 28(4), 405-414.  
 교육부(1997). 과학과 교육과정(교육부 고시 제 1997-15호) [별책 9].  
 교육인적자원부(2007). 과학과 교육과정(교육인적자원부 고시 제 2007-79호) [별책 9].  
 김경희, 김선희, 김지영, 박선용, 김수진, 김남희, 박효희, 정송(2008a). TIMSS 2007 공개문항 분석 자료집. 한

국교육과정평가원 연구보고 ORM 2008-32.  
 김경희, 김수진, 김남희, 박선용, 김지영, 박효희, 정송(2008b). 국제 학업성취도 평가(TIMSS/PISA)에서 나타난 우리나라 중·고등학생의 성취 변화의 특성. 한국교육과정평가원. RRE 2008-3-1.  
 김경희, 김수진, 김남희, 박선용, 김지영, 박효희, 정송(2008c). 수학 과학 성취도 추이변화국제비교연구-TIMSS 2007 결과보고서. 한국교육과정평가원. 연구보고 PRE 2008-3-3.  
 김경희, 김수진, 김미영, 김선희, 강민경, 박효희, 정송(2009). PISA와 TIMSS 상위국과 우리나라의 교육과정 및 성취 특성 비교 분석(RRE 2009-7-2).  
 김미경, 김동영, 이인호, 정은영, 강훈식(2012). 2011년 국가수준 학업성취도 평가 결과 분석: 과학, 한국교육과정평가원. 연구보고 RRE 2012-2-5.  
 김미경, 김현경, 이인호, 정진수, 신명경(2011). 2010년 국가수준 학업성취도 평가 결과 분석: 과학, 한국교육과정평가원. 연구보고 RRE 2011-3-5.  
 김미영, 김경희(2010). 우리나라, 싱가포르, 대만의 과학 교과서 비교 연구 -생명 영역을 중심으로-. 한국과학교육학회지, 30(4), 498-518.  
 김미영, 김경희(2011). TIMSS 평가들에 따른 우리나라 및 싱가포르 교과서의 생물 영역 내용 분석. 한국생물교육학회지, 39(2), 217-234.  
 김상달, 이용섭, 이상균(2005). 초등학교 학생들의 공간 능력과 천체운동개념 및 과학탐구능력과의 관계. 한국지구과학학회지, 26(6), 461-468.  
 김수진, 박지현, 김현경, 진의남, 이명진, 김지영, 안윤경, 서지희(2012). 수학 과학 성취도 추이변화 국제 비교 연구: TIMSS 2011 결과보고서. 한국교육과정평가원. 연구보고 RRE 2012-4-1.  
 김순, 강용희, 안병호, 윤일희, 정원우(1999). 중학교 과학교육에서 천체운동의 개념형성학습에 관한 연구. 과학교육연구지, 23, 15-30.  
 김효남(1990). 국민학교 아동의 과학개념에 대한 실태조사 및 교정을 위한 방법연구. 한국과학교육학회지, 10(2), 11-24.  
 노언경, 김진호, 김수진(2010). PISA 2003 문제해결력 영역에 대한 성별 차별기능문항 추출. 교육방법연구, 22(4), 165-194.  
 신동희 박정(2002). 평가 방법에 따른 과학 성취도에서의 성 차이. 한국과학교육학회지, 22(1), 40-53.  
 신동희, 권오남, 김희백(2002). TIMSS 2003 과학 공개 문항 내용 분석에서 나타난 성별 문항 응답 특성. 한국과학교육학회지, 26(6), 732-742.  
 신동희, 박병태(2009). 초등 과학 문제 풀이 과정에서의 성별 특성. 초등과학교육, 28(3), 229-244.  
 신동희, 박정(2002). 국제비교 연구에 나타난 우리나라 학생들의 지구 과학 성취도: 성 차이를 중심으로. 한

- 국지구과학회지, 23(3), 207-220.
- 이경훈, 임종옥(2010). 9학년 학생의 공간능력 차이에 따른 지구 자전 개념의 획득. 한국지구과학회지, 31(3), 267-275.
- 전경문(2003). 제 3차 수학 과학 성취도 국제비교 반복 연구(TIMSS-R) 중 화학영역 성취도 분석. 한국과학교육학회지, 23(3), 299-307.
- 정은영, 이미경, 홍미영(2006). TIMSS 2003 과학 성취도에서의 성 차이. 한국과학교육학회지, 26(4), 492-501.
- 조지민, 김수진, 이상하, 김미영, 옥현진, 임해미, 박연복, 이민희, 한희진, 손수경(2011a). 2011년 국제 학업성취도 평가연구(PISA/TIMSS): TIMSS 2011 본검사 시행 보고서. 한국교육과정평가원 연구보고 RRE 2011-4-1.
- 조지민, 김수진, 이상하, 김미영, 옥현진, 임해미, 박연복, 이민희, 한희진, 손수경(2011b). 2011년 국제 학업성취도 평가 연구(PISA/TIMSS): PISA 2009 결과에 기반한 읽기 영역 성취 특성 비교. 한국교육과정평가원 연구보고 RRE 2011-4-3.
- 최병순(1993). 빛. 한국교원대학교 과학교육연구소.
- 허경철, 차우규, 김선주, 권혁준(2005). 세계 각국의 최근 교육동향 교육과정 국제 비교. 한국교육과정평가원 연구보고 RRC 2005-14.
- 홍미영, 정은영, 이미경, 광영순(2006). TIMSS 2003 성취 수준에 따른 우리나라 중학생들의 과학 성취도 분석. 한국과학교육학회지, 26(6), 246-257.
- Camilli, G. & Shepard, L. A. (1994). *Methods for identifying biased test items*. Thousand Oaks, CA: Sage Publications.
- Foy, P., Arora, A. & Stanco, G. M. (2013). *TIMSS 2011 user guide for the international database science fourth grade released items*. MA: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College.
- Galili, I. & Hazan, A. (2000). Learners' knowledge in optics: Interpretation, structure and analysis. *International Journal of Science Education*, 22(1), 57-88.
- Holland, P. W. & Thayer, D. T. (1988). Differential item performance and the Mantel-Haenszel procedure. In H. Wainer & H. Braun (Eds.), *Test validity*. Hillsdale, NJ: Lawrence Erlbaum.
- Lynn, R. & Mikk, J. (2009). Sex differences in reading achievement. *Trames*, 13(63/58), 1, 3-13.
- Martin, M. O., Mullis, I. V. S. & Foy, P. (2012). *TIMSS 2011 International Results in Science*. MA: Boston College.
- Mullis, I. V. S., Martin, M. O., Ruddock, G. J., O'Sullivan, C. Y. & Preuschoff, C. (2009). *TIMSS 2011 Assessment Framework*. MA: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College.
- Shealy, R. T. & Stout, W. F. (1993). An item response model for test bias and differential test functioning. In P. Holland & H. Wainer (Eds.), *Differential item functioning* (pp. 197-239). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Swaminathan, H. & Rogers, H. J. (1990). Detecting differential item functioning using logistic regression procedures. *Journal of Educational Measurement*, 27, 361-370.