TIMSS-R 과학 성취도에서의 성 차이

이미경 · 홍미영 · 정은영
(한국교육과정평가원)

Gender Differences in TIMSS-R Science Achievement

Lee, Mee-Kyeong · Hong, Mi-Young · Jeong, Eun-Young
(Korea Institute of Curriculum and Evaluation)

ABSTRACT

The purpose of the study was to produce resources to help develop instructional methods and programs for school science to reduce gender differences in science achievement by analyzing TIMSS-R results according to item type, benchmark, and content category.

Korean male students performed higher than Korean female students and gender differences of Korean students were higher than international means in all item types including multiple-choice, short answer, and extended response type. GDI(Gender Difference Index) of Korean students was lower than international mean in top 10% benchmark but higher than international means in other benchmarks. Korean male students also exhibited higher performance than Korean female students in all content categories except scientific inquiry and the nature of science category.

Regarding items that GDIs were higher than 10, the number of items belonged to life science and earth science category was relatively larger and a high percentage of them was not included in school science curriculum. In addition, the items were equally distributed in each performance category.

In sum, the study showed various gender differences according to item type, benchmark, and content category. The results could be used to find appropriate instructional methods to reduce gender differences in science achievement.

Key words: TIMSS-R, gender differences, science achievement

I. 서 론

미국의 국제교육성취도평가협회(International Association for the Evaluation of Educational Achievement: IEA)가 주관하여 실시하고 있는 수학 · 과학 국제 비교 평가 연구(Trends of International Mathematics and Science Study: TIMSS)나 OECD가 주관하여 실시하고 있는 학업 성취도 국제 비교 연구(Programme for International Student Assessment: PISA) 결과는 과학 성취도에 있어서 우리나라 남 · 여학생의 차이가 큼을 보여주고 있다(Mullis *et al.*, 2000; OECD, 2001).

1995년에 검사가 실시된 TIMSS에서 우리나라 중학교 2학년 학생들의 성차는 전체 39개 참여국 중 여섯 번째로 큰 것으로 나타났으며, 남학생 평균은 576점, 여학생 평균은 551점이었다. 또한, 1999년에 검사가 실시된 TIMSS-Repeat(이하 TIMSS-R)에서도 이와 같은 성차가 그대로

*2004.8.4(접수) 2004.9.10(1심통과) 2004.11.11(최종통과)
**이미경(mklee@kice.re.kr)

유지되어 우리나라 남·여학생의 성차는 전체 38개국 중 9위를 차지하였다(Mullis *et al.*, 2000).

과학적 소양의 측정에 목적이 있는 PISA에서도 이와 같은 성차는 그대로 나타났는데 2000년에 검사가 실시된 PISA 2000의 결과에 의하면 과학 성취도에서 우리나라 남·여학생들의 성차는 32개 참여국 중 가장 큰 것으로 나타났다(OECD, 2001).

이상의 국제 비교 연구 결과들은 과학 성취도에서 우리나라 남·여학생의 성차가 우려할 만한 수준이며, 성차가 나타나는 원인을 찾아내어 그 차이를 줄이려는 체계적인 노력이 필요함을 시사한다.

과학 기술 분야에 여성 참여의 중요성이 인식되면서 정부에서는 '여성 과학 기술인 채용 목표제' 도입, '여성 과학 기술인 육성 및 지원에 관한 법률' 제정 등 여성 과학 기술인을 육성하기 위한 정책들을 의욕적으로 추진하고 있다(안재헌, 2004). 그런데 우리나라 여학생의 과학 성취도가 남학생보다 낮게 나타난다는 것은 미래의 여성 과학 기술인 양성에 적신호가 될 수 있다. 미래의 우수한 여성 과학 기술인을 육성하기 위해서는 현재의 여학생들의 과학 성취도를 높이려는 노력이 필요하다. 이를 위하여 우리나라 남·여학생의 과학 성차의 실태를 분석하여 이러한 분석 결과를 기초로 여학생의 과학 성취도 향상 방안을 모색할 필요가 있다.

국내에서 이루어진 남·여학생의 성차와 관련된 연구들을 살펴보면, 초·중·고 학생들의 과학 성차 문제를 심층적으로 조명하는 연구가 부족함을 알 수 있는데, 특히 TIMSS 결과를 활용한 연구는 많지 않다. TIMSS 결과와 관련된 연구들은 주로 화학(홍미영과 전경문, 2002; 전경문, 2003), 물리(유준희, 2001), 지구과학(명전옥과 홍미영, 2002), 과학 탐구 및 과학의 본성(홍미영, 2002) 등 내용 영역별 성취도를 중심으로 하였으며 성차에 관한 내용은 포함되지 않았다. TIMSS 결과를 활용한 성차 관련 연구로는 지구과학 영역에서 내용 소영역, 탐구 기능, 문항 유형에 따른 남학생과 여학생의 정답률을 비교한 연구(신동희와 박정, 2002)가 있을 뿐이다.

본 연구는 1999년에 검사가 실시된 TIMSS-R 결과에서 남·여학생의 성취도 차이를 문항 유형별, 성취수준별, 내용 영역별로 분석하고, 이로부터 학교 과학 교육에서 성차를 줄이는 데 도움이 되는 시사점을 찾아냄으로써, 성차를 줄이기 위한 과학 프로그램이나 교수·방법을 구안하는 데 도움이 되는 기초 자료를 산출하는 데 그 목적이 있다.

본 연구의 구체적인 내용은 다음과 같다.

1) 문항 유형에 따른 남·여학생의 성취도 차이 분석
2) 성취수준에 따른 남·여학생의 성취도 차이 분석
3) 내용 영역에 따른 남·여학생의 성취도 차이 분석
4) 성 차이가 크게 나타나는 문항들의 특성 분석

Ⅱ. 연구 방법

TIMSS-R의 본검사 시행 과정, 자료 분석 과정, 평가틀 및 평가 도구 등 일반적인 연구 방법은 다른 연구(김성숙 등, 1999; 홍미영 등, 2001)에 이미 상세하게 기술되어 있으므로, 본 연구에서는 분석 대상인 평가 도구와 이 연구에서 사용한 분석 방법에 관해서만 간략하게 소개하였다.

1. 분석 대상 문항

TIMSS-R의 과학 평가도구는 6개 내용 영역(화학, 물리, 생물, 지구과학, 환경과 자원, 과학의 탐구와 본성)의 총 147문항으로 구성되어 있다.

TIMSS-R은 만 13세 학생을 대상으로 하였으며, 전 세계 38개국의 6,347학교의 191,730명의 표본 중에서 최종적으로 6,076학교의 180,700명이 참가하였다. 우리나라에서는 1999년 2월 본검사를 실시하였으며, 최종적으로 자료가 정상 처리된 학생은 전국 150개 중학교의 2학년 6,114명(남학생 3,069명, 여학생 3,045명)이었다.

2. 분석 방법

TIMSS 연구 본부에서는 성취도에서 성 차이를 조사하기 위하여 GDI(Gender Difference Index)를 산출하고 있다. GDI는 문항 수준에서 성 차이를 알아보기 위한 통계적 수치로, 각 문항에 대하여 정답을 한 남학생의 비율과 여학생의 비율의 차이에 근거하고 있으므로, 각 문항에서 나타나는 성차를 비교할 수 있다(Mullis *et al.*, 2000).

GDI를 산출하는 방법은 아래와 같다. 남학생이 여학생보다 더 잘한 문항의 경우 GDI 값이 양수로 나타나게 되고, 여학생이 남학생보다 더 잘한 문항의 경우 GDI 값이 음수로 나타나게 된다.

$$D = -\beta X.\ Y^{*}100 = -\frac{(P_0 - P_1)}{(P_0 + P_1)(2 - P_0 - P_1)}*100$$

D = (오답한 학생들 중에서 여학생의 비율)-(정답을 한 학생들 중에서 여학생의 비율)

β = 회귀 계수

X = 성(gender)에 대한 dummy variable로 여학생에 대해서는 X=0, 남학생에 대해서는 X=1로 정함

Y = 결과(outcome)에 대한 dummy variable로 오답에 대해서는 Y=0, 정답에 대해서는 Y=1로 정함.

P_0 = 정답을 한 여학생들의 상대적 빈도

P_1 = 정답을 한 남학생들의 상대적 빈도

TIMSS-R 과학 총 147문항에 대한 국내와 국제 남녀 학생의 정답률 자료(박정 등, 2001; IEA, 2000)를 근거로 하여, 남학생과 여학생의 정답률을 위의 식에 대입하여 각 문항의 GDI를 산출하였다. 그리고 문항 유형, 성취수준, 내용 영역에 따른 남학생과 여학생의 성취도 차이를 살펴보았다.

Ⅲ. 분석 결과

1. 문항 유형별 남 · 여학생의 성취도 차이

TIMSS-R 문항 유형은 선다형(104문항)[1], 짧은 응답형(28문항), 설명형(16문항)[2]으로 되어 있다. 문항 유형에 상관없이 우리나라와 국제 결과 모두 남학생이 여학생보다 성취도가 높은 것으로 나타났다(Table 1). 또한, 문항 유형에 상관없이 우리나라의 GDI 값이 국제 GDI 값보다 큰 것으로 나타나 문항 유형에 따른 성차도 우리나라가 국제 평균보다 큼을 알 수 있다. 우리나라 결과에서는 선다형(5.13), 짧은 응답형(4.21), 설명형(2.27) 순으로 GDI 값 차이가 크게 나타난 반면, 국제 결과에서는 짧은 응답형에서 GDI 값이 가장 크며(3.41), 다음으로 선다형(3.20), 설명형(1.65) 순으로 나타났다. 우리나라와 국제 GDI 값의 차이를 보면, 선다형에서 1.93으로 가장 큰 차이를 나타내고 있어 우리나라 여학생들의 성취도가 남학생에 비하여 특히 선다형 문항에서 낮다고 볼 수 있다. 이는 과학 성취도에서 선다형에 비하여 주관식 문항이나 수행평가를 활용하는 것이 여학생들에게 유리하다는 신동희와 김동영(2003)의 연구 결과와, 선다형에서는 남학생의 성취도가 높으나 서술형 문항에서는 성차가 나타나지 않거나 여학생의 성취도가 높게 나타난다는 DeMars(1998)의 연구 결과와 유사하다.

2. 성취수준별 남 · 여학생의 성취도 차이

TIMSS 연구에서는 학생들의 성취수준을 네 단계(상위 10%, 상위 25%, 상위 50%, 하위 25%)로 구분한다. 그리고 특정한 성취수준에 해당되는 학생들이 정답을 하는 경향이 크고 그 바로 아래의 성취수준에 해당되는 학생들이 정답을 잘 못하는 문항들을 선정하여 각 성취수준에 해당되는 문항들끼리 분류를 한다. 또한 전문가들이 문항 내용을 검토하여 각 성취수준에 해당되는 학생들의 특성을 서술하여 제시한다(Martin *et al.*, 2000). 각 성취수준에 해당되는 문항들에 대한 우리나라 남 · 여학생들의 정답률 및 GDI 값을 국제 평균과 비교한 결과는 Table 2와 같다.

우리나라와 국제의 경우 모두 성취수준에 관계없이 남학생들의 성취도가 여학생의 성취도에 비하여 높게 나타났다. 그런데 우리나라의 경우 상위 50%에 해당되는 문항들에 대한 GDI 값이 가장 크고 상위 10%에 해당되는 문

Table 1. Gender differences in science achievement by item type

Item type	% correct				GDI		
	Korean		International		Korean	International	Korean -International
	Female	Male	Female	Male			
Multiple-choice	63.6	67.8	53.8	56.7	5.13	3.20	1.93
Short answer	51.3	54.6	38.2	41.1	4.21	3.41	0.80
Extended response	58.2	59.7	46.4	47.9	2.27	1.65	0.62

1) TIMSS-R의 선다형 문항수는 104개였으나, 본 연구에서는 우리나라 학생들의 정답률 결과가 나오지 않은 문항 A11을 분석에서 제외하였다.
2) 신동희와 박정(2002)의 연구에서는 짧은 응답형과 설명형을 묶어서 서술형 문항으로 분석하였다. TIMSS-R의 서술형 문항수는 42개이나, 본 연구에서는 한 문항에 여러 개의 문항이 포함된 경우 각각의 문항별로 분석하였으므로 짧은 응답형과 설명형 문항수의 합이 44개가 되었다.

항들에 대한 GDI 값이 가장 낮게 나타난 반면에, 국제의 경우에서는 그 반대의 현상, 즉 상위 10%에 해당되는 문항들에 대한 GDI 값이 가장 크고 상위 50%에 해당되는 문항들에 대한 GDI 값이 가장 낮게 나타났다.

상위 10%에 해당되는 문항들의 경우 우리나라의 GDI 값은 3.55인데 국제 GDI 값은 4.01로 나타났다. 이는 우리나라 남학생들이 상위 10%에 해당되는 문항들에 대해 여학생보다 성취도가 높지만, 그 차이가 국제의 경우보다는 크지 않음을 의미한다. TIMSS 연구에서 상위 10% 수준은 다소 복잡하고 추상적인 과학 개념을 이해하는 것으로 그 특징을 서술하고 있다(Martin *et al.*, 2000). 이러한 측면에서 본다면 우리나라 여학생들은 추상적인 과학 개념에 대한 이해 측면에서 남학생들보다 성취도가 낮게 나타나기는 하나, 다른 성취수준에 비해서는 성차가 적게 나타났으며, 국제 GDI 값보다 낮게 나타났음을 알 수 있다.

그 이외의 성취수준에 해당되는 문항들에 대해서는 우리나라의 GDI 값이 국제의 경우보다 모두 높게 나타나서, 우리나라 남·여학생의 성 차이가 상대적으로 크게 나타남을 알 수 있다. 특히, 상위 50%에 해당되는 문항들에 대해서 우리나라 GDI 값은 5.77인데 국제 GDI 값은 1.17로 나타나서, 그 차이가 다른 수준에 비하여 크게 나타났다. 상위 50%에 해당되는 문항들에서는 기본 과학 지식을 인식하고 실제 상황에 지식을 적용하거나 의사 소통할 수 있는 능력을 요구하고 있다. 다른 나라들에서는 성별에 따라 이러한 능력의 측면에서 그 차이가 상대적으로 적은 반면에, 우리나라 여학생들은 이러한 능력이 남학생들에 비하여 상대적으로 부족함을 알 수 있다. 한편 상위 50%에 해당되는 문항은 총 32개인데, 그 중에서 물리 영역에 해당되는 문항이 13개로 가장 많은 비율을 차지하고 있고, 이 문항들에 대한 남·여학생의 정답률에 차이가 크게 나타난 것을 볼 때, 여학생들이 물리 영역에서 어려움을 겪고 있음을 알 수 있다.

3. 내용 영역별 남·여학생의 성취도 차이

총 147문항으로 된 TIMSS-R 문항의 내용 영역별 분포를 보면 생물 영역이 40문항(27%), 지구과학 영역이 22문항(15%), 물리 영역이 39문항(29%), 화학 영역이 23문항(14%), 환경과 자원 영역이 14문항(9%), 과학 탐구와 과학의 본성 영역이 12문항(8%)으로 구성되어 있다. 즉, 물리와 생물 영역의 문항이 제일 높은 비율을 차지하고 있다. 내용 영역별 남·여학생의 성취도 차이는 Table 3과 같다.

내용 영역별 GDI 값을 볼 때 우리나라 학생들의 경우 '과학 탐구와 과학의 본성' 영역을 제외하고 모든 내용 영역에서 남학생들의 성취도가 높음을 알 수 있으며, 국제적인 경향을 종합했을 때는 모든 내용 영역에서 남학생들의 성취도가 높음을 알 수 있다. '과학 탐구와 과학의 본성' 영역을 제외하고는 모든 내용 영역에서 우리나라 GDI 값이 국제 GDI 값보다 컸는데, 이는 이 영역을 제외한 모든 내용 영역에서 우리나라 학생들의 성차가 국제적인 성차보다 큼을 의미한다.

우리나라 학생의 GDI가 가장 큰 영역은 지구과학 영역이었으며, 환경과 자원 영역이 그 뒤를 이었다. 물리 영역, 생물 영역, 화학 영역의 GDI 또한 상대적으로 컸으며, 우리나라 여학생들이 더 높은 성취도를 보인 과학 탐구와 과학의 본성 영역의 GDI 값은 상대적으로 매우 작았다. 이는 우리나라 학생들이 가장 큰 성차를 보이는 영역은 지구과학 영역과 환경과 자원 영역이며, 가장 작은 성차를 보이는 영역은 과학 탐구와 과학의 본성 영역임을 의미한다.

이와 같은 결과는 국제적인 경향과는 차이가 있는데 특

Table 2. Gender differences in science achievement by benchmark

TIMSS international benchmark	% correct				GDI		
	Korean		International		Korean	International	Korean -International
	Female	Male	Female	Male			
top 10% benchmark	46.7	49.6	33.5	36.9	3.55	4.01	2.95
upper quarter benchmark	61.3	65.9	50.1	52.9	4.99	2.88	2.11
median benchmark	77.1	81.1	73.0	73.7	5.77	1.17	4.60
lower quarter benchmark	72.4	75.9	62.4	65.1	5.25	3.00	2.26

Table 3. Gender differences in science achievement by content category

Content Category	% correct				GDI		
	Korean		International		Korean	International	Korean -International
	Female	Male	Female	Male			
Earth Science	62.69	67.99	52.55	56.47	6.36	4.48	1.88
Life Science	63.02	66.47	54.24	54.95	4.58	0.70	3.88
Physics	60.42	64.58	49.99	54.40	5.46	5.18	0.28
Chemistry	58.13	60.84	45.05	48.1	3.38	3.31	0.07
Environmental and Resource Issues	58.11	63.31	48.25	52.05	6.02	4.10	1.92
Scientific Inquiry and the Nature of Science	57.52	57.42	42.23	42.24	-0.26	0.26	-0.52

히 생물 영역에서의 차이가 심하다. 생물 영역에서의 우리나라와 국제 GDI 값의 차이는 3.88로서, 우리나라 학생들은 생물 영역에서 남학생들의 성취도가 높은데 비하여 국제적인 경향은 남 · 여학생의 성취도 차이가 거의 없는 것으로 나타났다. 생물 영역은 일반적으로 여학생들이 선호하는 내용 영역으로 알려져 있으며, 성취도도 여학생이 높을 것으로 기대되는 영역이다. 그러나 TIMSS 연구는 이러한 기대와는 상반된 결과를 보여준다. 따라서 구체적인 생물 영역 문항의 내용을 검토하고, 이를 토대로 여학생들의 성취도를 높일 수 있는 방법을 구안하려는 노력이 필요하다고 생각된다.

내용 영역별 남녀 성취도 차이를 살펴본 결과에 의하면, 남 · 여학생의 성차는 특정 내용 영역에 국한된 것이 아니고 대부분의 내용 영역에 해당되는 문제인 만큼 과학 내용 전반에 걸쳐서 과학에서 성차가 나타나는 이유에 대한 심층적인 연구가 요구된다.

4. 성취도 차이가 큰 문항들의 특성

1) 남학생들의 성취도가 높은 문항

Table 4는 우리나라 학생들의 GDI가 10 이상인 문항 중 남학생들의 성취도가 높은 문항에 대한 정보를 나타낸 것으로, 우리나라 학생들의 GDI, 국제 평균 GDI, 내용 영역, 성취 수준, 학습 경험, 수행 영역 등이 제시되어 있다. 총 147개의 TIMSS-R 문항 중 약 21%에 해당하는 31개의 문항이 이에 속하였다.

이들 문항의 분포를 내용 영역별로 살펴보면 생물 영역의 문항이 11개, 물리 영역의 문항이 7개, 지구과학 영역의 문항이 7개, 화학 영역의 문항이 3개, 환경과 자원 영역의 문항이 3개인 것으로 나타났으며, 과학 탐구와 과학의 본성에 해당하는 문항은 없었다. TIMSS-R 문항 중 생물 영역의 문항이 27%, 물리 영역의 문항이 29%, 지구과학 영역의 문항이 15%, 화학 영역의 문항이 14%, 환경과 자원 영역의 문항이 9%인 것을 고려할 때, 이와 같은 결과는 우리나라 남 · 여학생의 성취도 차이가 큰 문항 중 남학생들이 높은 성취도를 보이는 문항은 생물과 지구과학 영역에 상대적으로 많음을 의미한다.

남학생들이 여학생들보다 높은 성취도를 나타내는 문항들을 영역별로 구체적으로 살펴보면 다음과 같다. 생물 영역의 경우 늑대가 자기 구역을 알리는 방법에 관한 문제, 지구상에서 가장 오랜 기간 동안 살아온 동물 종류를 찾는 문제, 단백질이 건강에 좋은 이유, 운동을 하면 심장 박동이 빨라지는 이유, 곤충의 공통적인 특징, 인체에서 감각을 판단하는 곳, 초식동물의 생김새의 특징, 식물의 종자가 만들어지는 곳, 겨울잠을 자는 동물들에게 일어나는 변화, 동물의 분류 기준, 곤충 개체군의 생물학적 통제 등에 대한 문항들에서 남학생들이 여학생들보다 성취도가 높은 것으로 나타났다.

지구과학 영역의 경우, 지구의 내부 구조, 지구에서 보는 목성과 달의 크기, 지구의 1년에 해당되는 시간, 물의 순환에 필요한 에너지, 바닷물이 비가 되어 땅에 내리는 과정, 등고선 지도에서 강물이 흐르는 방향 찾기, 빛이 태양과 달에서 지구로 올 때 걸리는 시간 등에 대한 문항들이었다.

Table 4. Male higher-performing items with the Largest Gender Difference Index(GDI)

Item	GDI		Content Category*	Benchmark**	Learning Experience***	Performance Category****
	Korean	International				
P01 Determination of speed from graph	22.3	6.5	PH	U	B	UT
J05 Gravity acting on rocket	21.4	10.0	PH	T	J	SI
L05 Wolves marking territory	21.4	8.6	LS	M	J	SI
B01 Layers of Earth	19.9	7.9	ES	L	J	SI
L07 Effectiveness of insecticides	19.7	7.7	ER	U	N	CI
O14 Appearance of Jupiter and Moon	18.4	4.0	ES	M	E	TA
Q11 Earth year	15.7	10.7	ES	U	E	CI
I17 Animal on Earth longest time	15.6	7.8	LS	T	J	SI
I16 Energy for Earth's water cycle	15.4	9.8	ES	M	J	SI
J07 Reason for protein in diet	15.3	-0.3	LS	L	J	SI
Y01 Heart rate increasing with exercise	14.8	7.4	LS	T	B	TA
O10 Hole in ozone layer	14.8	5.1	ER	U	N	TA
J02 Feature shared by all insects	14.6	3.7	LS	U	J	CI
F03 Interpretation of senses	14.5	5.0	LS	L	B	SI
X01 Conversion of electrical/light energy	14.2	8.6	PH	T	N	TA
Z02 Diagram of rain from sea	13.9	6.6	ES	U	B	TA
D03 Contour map showing river	13.6	8.9	ES	T	N	TA
L03 Physical characteristic of prey	13.5	5.6	LS	T	N	CI
K19 Angle of reflected light ray	13.1	6.2	PH	L	E	SI
C11 Increased carbon dioxide in atmosphere	12.4	9.9	ER	U	N	CI
I14 Transmission of sound on the moon	12.1	5.6	PH	M	N	CI
D06 Seed development from plant part	11.6	1.6	LS	L	B	SI
P04 Hibernating animals	11.4	2.6	LS	T	N	CI
E10 Characteristics of animals groups	11.4	2.9	LS	U	J	UT
Q16 Light from sun and moon	10.9	6.1	ES	U	E	SI
Q14 Sugar dissolving in water	10.9	2.8	CH	T	B	SI
J04 Evaporation rate by surface area	10.7	1.9	PH	L	B	CI
Q12 Water level in U-tube	10.6	13.7	PH	T	J	CI
K12 Biological control of insect population	10.3	1.9	LS	U	N	TA
G10 Atoms removed from chair	10.2	6.1	CH	U	J	CI
I10 Gas needed for rust to form	10.1	7.2	CH	T	N	SI

*PH: Physics CH: Chemistry LS: Life Science ES: Earth Science ER: Environmental and Resource Issues
**T: Top 10% International Benchmark U: Upper Quarter International Benchmark
M: Median International Benchmark L: Lower International Benchmark
***B: Elementary & Junior School E: Elementary School J: Junior School N: None
****SI: Understanding Simple Information CI: Understanding Complex Information
TA: Theorizing, Analyzing and Solving Problems
UT: Using Tools, Routine Procedures and Science Processes

물리 영역의 경우에는 그래프를 보고 속도를 찾아내는 문제, 로켓에 작용하는 중력, 에너지 전환, 빛의 반사각, 달에서 소리의 전달, 표면의 넓이에 따른 증발률, U자관에서 물의 높이에 대한 문항들인 것으로 나타났다.

또한, 화학 영역의 경우에는 설탕의 용해, 의자의 모든 원자를 제거하면 남는 것, 부식의 원인이 되는 기체에 대한 문항, 그리고 환경과 자원 영역의 경우에는 살충제의 효과, 오존층, 지구 온난화에 대한 문항에서 남학생들의 성취도가 높게 나타났다.

Table 4의 문항을 성취 수준별로 살펴보면, 상위 10%에 해당하는 문항이 10개, 상위 25%에 해당하는 문항이 11개, 상위 50%에 해당하는 문항이 4개, 하위 25%에 해당하는 문항이 6개였다. TIMSS-R 문항 중 상위 10%에 해당하는 문항이 36%, 상위 25%에 해당하는 문항이 28%, 상위 50%에 해당하는 문항이 22%, 하위 25%에 해당하는 문항이 14%인 것을 고려할 때, 상위 25%와 하위 25%에 해당하는 문항의 경우에 남학생들의 성취도가 높음을 알 수 있다. 이와 같은 결과는 성취 수준별로 문항의 GDI 평균을 구하여 비교한 결과(Table 2)와는 차이가 있다. 성취 수준별로 문항 전체의 GDI 평균을 비교한 결과에서는 상위 50%에서 GDI 평균값이 가장 크게 나타났다. 이는 각 성취수준에 해당되는 모든 문항의 GDI 값을 평균했을 때는 상위 50%에 해당하는 문항들에 대한 남 · 여학생의 성취도 차이가 가장 큼을 의미한다. 그러나 GDI가 10 이상이면서 남학생들의 성취도가 높은 문항의 개수를 비교했을 때는 상위 25%와 하위 25%에 해당하는 문항이 많았다. 즉, 남 · 여학생의 GDI 차이가 큰 문항들은 중간 수준이 아니라 상대적으로 높은 인지 능력을 요구하는 문항들과 낮은 인지 능력을 요구하는 문항들인 경우가 많았다. 이는 고차적인 인지 능력을 요구하거나 반대로 낮은 인지 능력을 요구하는 내용을 가르칠 때 여학생이 그 내용을 이해하는지 여부에 좀 더 관심을 가질 필요가 있음을 시사한다.

Table 4의 문항 분포를 학습 경험을 기준으로 살펴보면, 초등학교와 중학교 때 배운 문항이 7개, 초등학교에서만 배운 문항이 4개, 중학교에서만 배운 문항이 10개, 전혀 배우지 않은 문항이 10개인 것으로 나타났다[3]. 전혀 배우지 않은 문항의 비율이 비교적 높은 것은 남학생들의 과학에 대한 관심이 여학생들보다 높은 데서 기인한 것이라고 볼 수도 있을 것이다.

우리나라 학생들의 GDI가 큰 문항 중 남학생들이 잘한 문항을 수행 영역별로 살펴보면 단순 정보를 묻는 문항이 12개, 복합 정보를 묻는 문항이 10개, 이론화하고 분석하는 능력을 요구하는 문항이 7개, 기기 및 과학의 과정 이용에 관한 문항이 2개인 것으로 나타났다. 즉, 단순 정보를 묻는 문항과 복합 정보를 묻는 문항에서 남학생들이 상대적으로 높은 성취를 나타내었다. 그러나 단순 정보를 묻는 문항이 39%, 복합 정보를 묻는 문항이 31%, 이론화하고 분석하는 능력을 묻는 문항이 19%, 기기 및 과학의 과정 이용에 관한 문항이 7%, 자연 세계에 대한 조사와 관련된 문항이 4%인 TIMSS-R의 전체 문항 분포를 고려할 때 특별히 남학생들이 강세를 보이는 수행 영역이 있다고 보기는 어렵다. 다만 남학생들이 높은 성취를 보이는 문항들이 전 수행 영역에 비교적 골고루 분포되어 있는 것으로 볼 수 있다.

2) 여학생들의 성취도가 높은 문항

Table 5는 우리나라 학생들의 GDI가 10 이상인 문항 중 여학생들의 성취도가 높은 문항에 대한 정보를 나타낸 것이다. TIMSS-R 문항 중 여기에 해당하는 문항은 총 147문항 중 3문항에 불과한 것으로 나타났으며, 생물, 물리, 과학 탐구와 과학의 본성 영역에서 각각 1 문항씩 포

Table 5. Female higher-performing items with the Largest Gender Difference Index(GDI)

Item	GDI		Content Category	Benchmark	Learning Experience	Performance Category
	Korean	International				
Q18 Mass of freezing water	-10.7	3.2	PH	T	J	SI
P07 Replication of measurements	-11.1	0.3	SN*	T	N	UT
O17 Transmission of cold in classroom	-11.6	-9.9	LS	U	N	TA

* SN: Scientific Inquiry and the Nature of Science

3) TIMSS-R을 위한 검사는 1999년도에 시행되었으므로 문항 분포의 학습 경험 여부는 제6차 과학과 교육과정에 근거하여 구분하였다.

함되었다. 이 세 문항 중 한 문항은 상위 25% 수준에 속하는 문항이며, 두 문항은 상위 10% 수준에 속하는 문항으로 세 문항 모두 고차원적인 능력을 요구하는 문항에 속한다. 세 문항 중 물리와 과학 탐구와 과학의 본성 영역에 속하는 두 문항은 국내 정답률은 여학생이 높았으나 국제 정답률은 남학생이 높았다.

문항의 분포를 수행 영역별로 살펴보면, 이론화하고 분석하는 능력을 요구하는 문항, 단순 정보를 묻는 문항, 기기 사용에 관한 문항이 각각 1개씩인 것으로 나타났다. GDI가 큰 문항 중 여학생들이 높은 성취도를 보이는 문항이 세 문항에 불과하여 여학생들이 강세를 보이는 문항들로부터 어떤 경향성을 찾기는 어렵다.

Ⅳ. 결론 및 제언

이 연구에서는 TIMSS-R의 총 147개 과학 문항의 GDI(Gender Difference Index) 값을 산출하여 문항 유형, 성취수준, 내용 영역에 따른 남학생과 여학생의 성취도 차이를 살펴보았다.

문항 유형에 상관없이 우리나라와 국제 결과 모두 남학생이 여학생보다 성취도가 높은 것으로 나타났다. 또한 모든 문항 유형에서 우리나라의 성차가 국제 평균보다 크게 나타났는데, 특히 우리나라에서는 선다형 문항에서 성차가 크게 나타났다.

성취수준에 따른 우리나라와 국제의 GDI 값을 비교해 보면, 상위 10%에 해당되는 문항들에 대해서는 우리나라의 GDI 값이 국제의 경우보다 낮게 나타났고, 그 이외의 성취수준에 대해서는 우리나라의 GDI 값이 국제의 경우보다 높게 나타났는데, 특히 상위 50%에 해당되는 문항들에 대해서 남학생과 여학생의 성취도 차이가 국제의 경우에 비하여 상대적으로 큰 것으로 나타났다. 이러한 결과를 볼 때 우리나라 여학생들은 다소 복잡하고 추상적인 개념의 이해 측면에서는 남학생들과 큰 차이가 나지 않지만, 기본 과학 지식을 인식하고 실제 상황에 지식을 적용하거나 의사 소통할 수 있는 능력이 상대적으로 부족함을 알 수 있다.

내용 영역별로 남·여학생의 성취도를 비교했을 때는 과학 탐구와 과학의 본성 영역을 제외하고는 모든 내용 영역에서 우리나라 남학생들의 성취도가 여학생들보다 높았으며, GDI 값 역시 국제 평균보다 높았다. 이와 같은 결과는 국제적인 경향과는 차이가 있는데 특히 생물 영역에서의 차이가 심하다. 생물 영역의 경우에 우리나라는 남학생들의 성취도가 높은데 비하여 국제의 경우에는 남·여학생의 성취도 차이가 거의 없는 것으로 나타났다.

GDI가 10 이상인 문항 중 남학생들의 성취도가 높은 문항들은 생물과 지구과학 영역에 속하는 경우가 상대적으로 많았으며, 학교 교육과정에서 전혀 배우지 않은 문항들의 비율 또한 비교적 높았다. 수행 영역별 분포를 볼 때 남학생들이 높은 성취도를 보이는 문항은 전 수행 영역에 걸쳐서 비교적 골고루 분포되어 있었다.

GDI가 10 이상인 문항 중 여학생들의 성취도가 높은 문항들은 그 수가 매우 적어 어떤 경향성을 찾아내는 데 한계가 있었다.

연구의 결과를 고려하여 남·여학생의 성차를 줄이기 위한 제언을 하면 다음과 같다.

첫째, 문항 유형별 남·여학생의 성취도를 비교해 본 결과, 짧은 응답형과 설명형 문항에서 그 차이가 상대적으로 작게 나타났는데, 이는 과학적 능력을 다양한 평가 방법을 활용하여 다각적으로 측정하는 것이 타당하고 필요하다는 것을 의미하며, 이러한 측정 방법의 다각화는 과학 성취도에서의 성차에 대한 고정 관념을 줄이는 데도 기여할 것이다.

둘째, 남학생과 여학생의 성취도 차이가 네 단계의 성취수준에서 동일한 양상으로 나타나지 않았다. 따라서 각 성취수준에서 요구되는 능력을 고찰하여 전체 학생들의 성취도 향상을 위한 노력을 기울임과 동시에, 특히 남·여학생의 차이가 크게 나타난 성취수준에 대해서는 여학생에게 부족한 능력을 보완하기 위한 방안을 모색할 필요가 있다.

셋째, 남학생들의 성취도가 높은 문항들에 대하여 학습 경험을 기준으로 살펴본 결과 전혀 배우지 않은 문항의 비율이 비교적 높게 나타났는데, 이는 남학생들의 과학에 대한 관심이 여학생들보다 높은 데서 기인한 것이라고 볼 수도 있다. 따라서 여학생들이 과학에 대하여 흥미와 관심을 갖도록 하기 위하여 학교 과학 수업에서뿐만 아니라 과학 캠프나 과학 클럽 등 여학생을 위한 수업 외의 다양한 활동을 개발할 필요가 있다.

넷째, 좀 더 구체적인 내용 수준에서 남·여학생의 성취도를 분석하여 여학생들이 취약점을 보이는 부분을 찾아낼 필요가 있다. 이 연구에서는 물리, 화학, 생물, 지구과학, 환경과 자원, 과학 탐구와 과학의 본성으로 영역을 나누어 남·여학생의 성취도를 분석하였으나 이처럼 폭넓

은 수준에서 성 차이를 살펴보는 것은 구체적인 교수-학습 방법의 개발에 도움이 되기 어렵다. 따라서, 여학생이 어려움을 겪는 문항들을 분석함으로써 성차가 나타나는 내용을 좀 더 구체적인 수준에서 찾아낼 필요가 있다.

다섯째, TIMSS 자료 외에도 여러 가지 다양한 성취도 평가 자료를 활용하여 남 · 여학생의 과학 성취도 성차의 원인을 진단하고, 해결 방안을 모색하려는 노력이 필요하다.

국 문 요 약

학업성취도 국제 비교 연구 결과에서 우리나라 남 · 여학생의 과학 성취도 차이가 큰 것으로 보고되고 있다. 이 연구에서는 TIMSS-R 과학 성취도에서의 성 차이를 문항 수준에서 살펴보기 위하여 각 문항의 GDI(Gender Difference Index) 값을 산출하였다. 그리고 문항 유형, 성취수준, 내용 영역에 따른 GDI 값을 비교하였다.

문항 유형에 상관없이 우리나라와 국제 결과 모두 남학생이 여학생보다 성취도가 높은 것으로 나타났으며, 우리나라의 성차가 국제 평균보다 크게 나타났다. 특히 우리나라 여학생들은 선다형 문항에서 상대적으로 낮은 성취도를 나타내었다. 성취수준에 따른 우리나라와 국제의 GDI 값을 비교해 보면, 상위 10%에 해당되는 문항들에 대해서는 우리나라의 GDI 값이 국제의 경우보다 낮게 나타났으나 그 이외의 성취수준에 해당되는 문항들에 대해서는 우리나라의 GDI 값이 더 높게 나타났고 특히 상위 50%에 해당되는 문항들의 경우 우리나라와 국제의 차이가 크게 나타났다. 내용 영역별로 남 · 여학생의 성취도를 비교했을 때는 과학 탐구와 과학의 본성 영역을 제외하고는 모든 내용 영역에서 우리나라 남학생들의 성취도가 여학생들보다 높았다.

GDI가 10 이상인 문항 중 남학생들의 성취도가 높은 문항들은 생물과 지구과학 영역에 속하는 경우가 상대적으로 많았으며, 학교 교육과정에서 전혀 배우지 않은 문항들의 비율 또한 비교적 높았다. 그리고 남학생들이 높은 성취도를 보이는 문항은 전 수행 영역에 걸쳐서 비교적 골고루 분포되어 있었다.

이상에서와 같이 문항 유형, 성취수준, 내용 영역에 따라 과학에서의 남 · 여학생의 성취도 차이가 다르게 나타났다. 이러한 결과는 여학생들의 과학 성취도를 향상시키기 위한 방안을 모색하는 과정에서 보다 구체적인 방향을 제시할 것으로 기대된다.

참 고 문 헌

김성숙, 유준희, 서동엽, 이춘식, 임찬빈(1999). 제3차 수학 · 과학 성취도 국제 비교 반복 연구 국내 평가 시행보고서. 한국교육과정평가원 연구보고서. 연구보고 RRE-99-7-2.

명전옥, 홍미영(2002). 비교 반복 연구의 지구과학 성취도 분석. 한국과학교육학회지, 22(3), 649-659.

박정, 홍미영, 나귀수, 김성숙(2001). 제3차 수학 · 과학 성취도 국제 비교 반복 연구(TIMSS-R) 공개문항 분석 자료집. 연구자료 ORM 2001-9. 서울: 한국교육과정평가원.

신동희, 김동영(2003). 평가 방법에 따른 과학 성취도에서의 성 차이. 한국과학교육학회지, 23(3), 265-275.

신동희, 박정(2002). 국제비교연구에 나타난 우리 나라 학생들의 지구과학 성취도: 성 차이를 중심으로. 한국지구과학교육학회지, 23(3), 207-220.

안재현(2004). 이학 분야 남녀 학력 차이와 여성 인적 자원 활용. 남 · 여학생의 학력 차이, 무엇이 문제인가?, 15-25. 이화여자대학교와 한국교육과정평가원. 연구자료 ORM 2004-5.

유준희(2001). 제3차 수학 · 과학 학업성취도 국제 비교 반복 연구(TIMSS-R) 물리 영역 성취도 분석을 통한 교육과정 개선 연구 과제의 추출. 한국과학교육학회지, 21(4), 757-772.

전경문(2003). 제3차 수학 · 과학 성취도 국제 비교 반복 연구(TIMSS-R) 중 화학 영역 성취도 분석. 한국과학교육학회지, 23(2), 299-307.

최경희(2001). 과학 교과에서의 양성평등 교육을 위한 교수학습 전략 및 자료 개발 방안 연구. 한국과학교육학회지, 21(1), 213-231.

홍미영, 박정, 김성숙(2001). 제3차 수학 · 과학 성취도 국제 비교 반복 연구(TIMSS-R) 과학 성취도 분석. 한국과학교육학회지, 21(2), 328-341.

홍미영(2002). 우리나라 중학생들의 과학적 탐구 및 과학의 본성 영역에서의 국제 성취도 분석. 한국과학교육학회지, 22(2), 336-344.

홍미영, 전경문(2002). 제3차 수학 · 과학 성취도 국제 비교 반복 연구(TIMSS-R) 중 제6차 교육과정상의 화학 문항에 대한 우리나라 중학생의 응답 분석. 대한화학회지, 46(2), 169-176.

DeMars, C. E. (1998). Gender difference in mathematics and science on a high school proficiency exam: The role of response format. *Applied Measurement in Education*, 11(3), 279-299.

IEA (2000). *TIMSS 1999 Percent of responses by item category for the science items*. Chestnut Hill, MA: Boston College.

Martin, M. O., Mulis, I. V. S, Gonzales, E. J., Gregory, K. D., Smith, T. A., Chrostowski, S. J., Garden, R. A., & O'Connor, K. M. (2000). *TIMSS 1999 International Science Report: Findings from IEA's Repeat of the Third International Mathematics and Science Study at the Eighth Grade*. International Study Center Boston College Chestnut Hill, MA, USA.

Mullis, I. V., Martin, M. O., Fierros, E. G., Goldberg, A. L., & Stemler, S. E. (2000). *Gender difference in achievement; IEA's third international mathematics and science study(TIMSS)*. MA: International Study Center.

OECD. (2001). *Knowledge and Skills for Life: First Results from PISA2000*. Paris. OECD Publication.