

TIMSS 2003 과학 성취도에서의 성 차이

정은영* · 이미경 · 홍미영

한국교육과정평가원

Gender Differences in TIMSS 2003 Science Achievement

Jeong, Eun-Young* · Lee, Mee-Kyeong · Hong, Mi-Young

Korea Institute of Curriculum and Evaluation

Abstract: Gender differences in TIMSS 2003 science achievement by item type, benchmark, and content area were examined by producing a Gender Differences Index (GDI) in this study. International trends identified that male students performed better than female students in TIMSS 2003 science achievement in all types of items. The overall achievement of Korean male students was better than Korean female students, especially in multiple-choice type items. Male students outperformed females in three benchmarks, including advanced, high, and intermediate international benchmark, but they did not outperform females in the low international benchmark when gender differences of the international average as well as the Korean average were taken into account. The results of the analysis of the international average and the Korean average by content area showed that gender differences were the greatest in earth science and smallest in chemistry. In life science, female students excelled when considering the international average while male students excelled when considering the average of Korean students' performance. In addition, the number of items in which male students outperformed females was larger in both factual knowledge and the conceptual understanding domain. Implications for reducing gender differences in science achievement in Korea based on the results were provided.

Key words: TIMSS 2003, gender differences, science achievement

I. 서 론

2004년 12월에 전세계적으로 동시에 발표된 2003년 수학·과학 국제 비교 평가 연구(Trends of International Mathematics and Science Study: TIMSS 2003) 결과와 2003년 학업 성취도 국제 비교 연구(Programme for International Student Assessment: PISA 2003) 결과는 과학 성취도에 있어서 우리나라 남·여학생의 차이가 여전히 크게 존재함을 보여주고 있다. TIMSS 2003에서 우리나라 남·여학생의 과학 성취도 차이는 12점으로 국제 평균인 6점에 비하여 높았으며(Mullis *et al.*, 2004), PISA 2003에서는 그 차이가 18점으로 역시 OECD 국가별 평균인 6점보다 높았다(이미경과 광영순 등, 2004).

만 13세, 즉 중학교 2학년을 대상으로 한 TIMSS에서는 추이 분석을 하고 있으므로, 그 결과를 통하여 우리나라 남·여학생의 성취도 차이가 어떻게 변화해 왔는지를 알 수 있다. 과학 성취도에 있어서 남·여학

생의 성취도 차이는 1995년에는 28점, 1999년에는 21점, 2003년에는 12점으로 그 차이가 많이 감소하고 있으며, 이와 같이 차이가 감소하는 이유는 여학생들의 성취도가 남학생에 비하여 높아지고 있기 때문인 것으로 나타났다. 즉 여학생의 과학 성취도는 빠른 속도로 향상되고 있는 반면, 남학생의 성취도는 그 상승 폭이 작았다(박정 등, 2004).

만 15세 학생(우리나라의 경우 대부분 고등학교 1학년)을 대상으로 하는 PISA의 경우는 2000년의 결과와 2003의 결과를 비교할 수 있는데, 2000년에는 남·여학생의 성취도 점수 차이가 19점, 2003년의 경우에는 18점으로 그 격차가 거의 그대로 유지되고 있는 것으로 나타났다(이미경과 광영순 등, 2004).

이상의 국제 비교 연구 결과들은 과학 성취도에서 우리나라 남·여학생의 성차가 매우 크며, 성차가 나타나는 원인을 찾아내어 성차를 줄이려는 체계적인 노력이 필요함을 시사한다. 최근에 우리나라 남·여학생의 과학 성취도 격차에 대한 관심이 높아지면서 이

*교신저자: 정은영(jey@kice.re.kr)

**2005.10.04(집수) 2006.04.05(1심통과) 2006.05.26(2심통과) 2005.06.07(최종통과)

와 관련한 여러 가지 평가 및 개발 연구가 이루어지고 있다. 국가수준에서 이루어지고 있는 연구로는 ‘국가 수준의 학업성취도 평가 연구’가 있는데, 이 연구에서는 초등학교 6학년, 중학교 3학년, 고등학교 1학년을 대상으로 매년 과학 성취도 평가를 수행하고, 남·여학생의 성취도 차이를 분석하고 있다. 이 외에도 평가 방법에 따른 과학 성취도에서의 성 차이를 분석한 연구(신동희, 김동영, 2003), 교사의 성인지적 교육활동 능력 제고를 위한 프로그램 개발 연구(정해숙, 정경아, 2003), 여학생 친화적인 과학 교육 프로젝트의 내용 비교 및 분석 연구(진희정, 2002) 등 다양한 연구들이 이루어지고 있다. 그러나 구체적으로 과학의 어떤 내용에서 성취도 격차가 나타나는지, 여학생들이 부족한 부분이 과학에서 요구하는 능력 중 어떤 것들인지에 대한 심층적인 연구는 부족한 형편이다. 남·여학생의 성취도 격차를 해소하기 위해서는 이와 같은 연구들을 통하여 그 원인을 정확하게 파악하는 것이 먼저 이루어져야 할 것으로 판단된다.

이 연구는 2003년에 만 13세 학생을 대상으로 검사를 실시한 TIMSS 2003 결과에서 남·여학생의 성취도 차이를 심층적으로 분석하여, 이로부터 학교 과학 교육에서 성차를 줄이는 데 도움이 되는 시사점을 찾아냄으로써, 성차를 줄이기 위한 과학 프로그램이나 교수·방법의 구안에 도움이 되는 기초 자료를 산출하는 데 그 목적이 있다.

본 연구의 구체적인 내용은 다음과 같다.

- 1) 문항 유형에 따른 남·여학생의 성취도 차이를 분석한다.
- 2) 성취수준에 따른 남·여학생의 성취도 차이를 분석한다.
- 3) 내용 영역에 따른 남·여학생의 성취도 차이를 분석한다.
- 4) 성 차이가 크게 나타나는 문항들의 특성을 분석한다.

II. 연구 방법

TIMSS 2003의 본검사 시행 과정, 자료 분석 과정, 평가틀 및 평가 도구 등 일반적인 연구 방법은 박정등(2003)의 보고서에 상세하게 기술되어 있으므로, 이 연구에서는 분석 대상인 평가 문항과 이 연구에서 독자적으로 사용한 분석 방법에 관해서만 간략하게 소개하였다.

1. 분석 대상 문항

TIMSS 2003의 과학 평가도구는 5개 내용 영역(생

물, 화학, 물리, 지구과학, 환경)의 총 189문항(선택형 109문항, 자유반응형 80문항)으로 구성되어 있다. 각 내용 영역별 분포를 살펴보면 생물 영역이 54문항(29%)으로 가장 많았고, 그 다음으로 물리 영역이 46문항(24%)으로 많았으며, 화학 영역과 지구과학 영역은 각각 31문항씩(16%)이었고, 환경 영역이 27문항(14%)으로 가장 적었다. 참고로, TIMSS 2003의 공개 문항은 TIMSS & PIRLS International Study Center 홈페이지(<http://timss.bc.edu/timss2003i/released.html>)에서 다운받을 수 있다.

TIMSS 2003은 만 9세와 만 13세 학생을 대상으로 하였으며, 전 세계 46개국이 참가하였다. 우리나라는 만 13세 학생 대상의 연구에만 참여하였는데, 중학교 3학년 학생들을 대상으로 유층군집 표집 방법에 의해 표집을 하여 2003년 4월에 본검사를 실시하였으며, 최종적으로 자료가 정상 처리된 학생은 전국 149개 중학교의 5,478명(남학생 2,883명, 여학생 2,595명)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). 또한 GDI 수치를 통해 TIMSS-R의 과학 성취도에서의 성차 결과(이미경과 홍미영 등, 2004)와 비교할 수 있다.

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.

$$D = -\beta X \cdot Y \times 100 = -\frac{(P_0 - P_1)}{(P_0 + P_1)(2 - P_0 - P_1)} \times 100$$

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 본부에서 발행한 문항별 정답률 자료로부터 TIMSS 2003의 과학 총 189 문항 각각에 대해서 우리나라와 국제 평균의 남·여학생의 정답률 정보를 얻었다. 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 및 논의

1. 문항 유형별 남·여학생의 성취도 차이

TIMSS 2003의 과학 성취도에서 문항 유형에 상관 없이 우리나라와 국제 결과 모두 남학생이 여학생보다 성취도가 높은 것으로 나타났다(Table 1). 또한 두 가지 문항 유형에서 모두 우리나라의 GDI 값이 국제 GDI 값보다 큰 것으로 나타나 문항 유형에 따른 성취도 차이가 국제 평균보다 큼을 알 수 있다.

문항 유형에 따라 비교하여 보면, 우리나라와 국제 모두 자유응답형보다 선다형에서 남학생들이 여학생들보다 상대적으로 성취도가 높았다. 우리나라 결과에서는 선다형의 GDI 값이 4.08이며, 자유응답형에서 3.16으로 선다형 문항에서 여학생들이 취약함을 나타내었으며, 이는 TIMSS-R의 결과(이미경과 홍미영 등, 2004)에서도 확인된 바 있다. 반면 국제 결과를 보면 선다형의 GDI 값은 1.91이나 자유응답형에서는 0.35로 나타나, 우리나라와 달리 자유응답형에서는 성취가 미미하다고 볼 수 있다.

문항 형태에 따라 남학생과 여학생의 성취도가 달라진다는 것에 대하여 상반된 연구 결과가 보고되고 있다. 학업 성취도에 있어서 선다형 문항은 남학생들에게 유리한 반면, 서술형 문항은 여학생들에게 유리하다는 연구 결과가 다수 있으며(신동희, 김동영, 2003; Bolger & Kellaghan, 1990; DeMars, 1998; Din et al., 2004), 이 연구들에서는 성차를 줄이기 위해서는 선다형 대신 수행 평가 등 다양한 방법으로 평가를 실시할 것을 제안하고 있다. 한편 Wester & Henriksson(2000)은 문항 형태의 변화가 성취에 영향을 주는지를 알아보기 위하여 1995년에 실시된 TIMSS 수

학 문항 중 10개의 선다형 문항과 이 문항들을 짧은 응답형으로 바꾼 문항들에 대하여 동일 집단으로 간주할 수 있는 남녀 학생 집단의 성취도를 비교하였다. 그 결과, 선다형 문항과 짧은 응답형 문항에서 모두 여학생들이 약간 높은 성취도를 보였으며, 문항 형태가 성취에 영향을 주지 않는 것으로 나타났다.

여학생들이 선다형보다 서술형 문항에서 상대적으로 높은 성취도를 보이는 원인에 대해서는 남학생보다 언어 기능 특히 작문 기능이 우수하기 때문이라는 주장(Breland et al., 1994)이 있으나, TIMSS 2003의 서술형 문항들은 주로 짧은 단어나 단문 형태의 답을 요구하므로 성취의 원인을 작문 기능으로 보기에선 무리가 있으므로, 문항 내용이나 학습 경험 등에 대하여 보다 면밀히 살펴볼 필요가 있다.

2. 성취수준별 남·여학생의 성취도 차이

TIMSS의 결과는 문항반응이론(IRT) 방법을 사용하여 평균 500, 표준편차 100의 척도로 나타나고, 각 내용 영역의 국제 평균은 474점으로 맞추어, 이를 기준으로 국가별로 표준 점수를 부여하였다. 그리고 TIMSS에서는 학생들의 수행 능력에 대하여 의미 있게 서술하기 위하여, 전체 성취 점수에서 네 개의 점수를 선정하여 ‘성취수준(international benchmarks)’으로 활용하고 있다. 성취수준은 수월 수준, 우수 수준, 보통 수준, 기초 수준으로 구분되고, 각 성취수준에 해당하는 학생이 무엇을 알고 무엇을 할 수 있는지를 제시하고 있다. TIMSS 2003에서는 성취수준 설정을 위해서 75점의 동급간의 고정된 네 개의 점수인 625, 550, 475, 400을 사용하였다. 각 점수는 수월 수준, 우수 수준, 보통 수준, 기초 수준의 기준점이다. 2000년 보고서까지는 상위 10%, 상위 25%, 상위 50%, 하위 25% 수준으로 구분하였는데, TIMSS 2003부터 이전의 구분 방식과 점수를 고려하면서 이와 같이 고정된 점수를 사용하여 성취수준을 구분하였고, 이후의 추이변화 결과에서 이 점수를 기준으로 학생 비율을 표기하는 것으로 결정하였다. TIMSS 2003 국제 결과 보고서에는 내용 지식의 폭과 깊이, 문제 상황, 과학 탐구 기능의 수준, 표와 그래프의 복잡성, 응답 내

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	61.32	64.79	49.70	51.44	4.08	1.91	2.17
Constructed-response	46.76	49.51	33.28	33.59	3.16	0.35	2.81

용의 수준 등에 근거하여 성취수준을 설정하고, 각 수준에 해당되는 문항을 명시하였는데, 수월 수준 95문항, 우수 수준 61문항, 보통 수준 23문항, 기초 수준 10문항이다(Martin *et al.*, 2004).

각 성취수준에 해당되는 문항들에 대한 우리나라 남·여학생들의 정답률 및 GDI 값을 국제 평균과 비교한 결과는 Table 2와 같다.

우리나라와 국제 평균의 경우 모두 기초 수준을 제외하고 다른 3개의 성취수준에서 남학생의 성취도가 여학생의 성취도에 비하여 높게 나타났다. 이에 따라서 GDI 값이 기초 수준의 경우 음수의 값이 나왔고, 다른 성취수준에서는 양수의 값이 나왔다. 우리나라의 경우 보통 수준에서 GDI 값이 7.52로 가장 크게 나타났는데, 이는 TIMSS-R에서 상위 50%에 해당되는 문항들에 대한 GDI 값이 가장 크게 나타난 것(이미경과 홍미영 등, 2004)과 유사한 결과이다. 한편 기초 수준에서는 GDI 값이 -1.19로 가장 작았다. 이러한 경향은 국제 평균에서도 나타났는데 우리나라의 경우보다 GDI 값이 전반적으로 작게 나타났다. 이와 같이 우리나라의 GDI 값이 모든 성취수준에서 국제 평균 GDI 값보다 높게 나타나서, 우리나라 남·여학생의 성취도 차이가 상대적으로 더 크게 나타남을 알 수 있다.

우리나라의 GDI와 국제 평균의 GDI의 차이를 비교해 보면 보통 수준에서 3.91로 가장 크게 나타났고 기초 수준에서는 -0.79로 가장 작게 나타났다. TIMSS-R에서도 상위 50%에 해당되는 문항들에 대해서 우리나라와 국제 평균의 GDI 값의 차이가 가장 크게 나타났다(이미경과 홍미영 등, 2004). 이러한 결과를 볼 때 우리나라 여학생들이 남학생들에 비하여 보통 수준에 해당되는 문항들에서 상대적으로 더 큰 어려움을 겪고 있음을 알 수 있다.

TIMSS 연구에서 보통 수준은 기본적인 과학 지식을 인식하고 지식을 간단하게 적용하고 의사소통할 수 있는 것으로 그 특징을 서술하고 있다(Martin *et al.*,

2004). 이러한 능력이 다른 나라 여학생들에게도 비교적 부족한 것으로 나타났는데, 우리나라 여학생의 경우 상대적으로 더 부족하다고 할 수 있다. 한편 보통 수준에 해당되는 문항을 살펴보면 총 23문항 중에서 지구과학 영역과 물리 영역의 문항이 각각 일곱 문항씩으로 많이 포함되어 있고, 생물 영역과 환경 영역의 문항이 각각 네 문항씩, 화학 영역의 문항이 한 문항이 포함되어 있다. 지구과학 영역의 경우 일곱 문항 중에서 다섯 문항에 대한 GDI 값이 10 이상으로 나타난 반면에, 다른 내용 영역의 경우 GDI 값이 10 이상이 되는 문항이 없거나 한 문항 또는 세 문항이 있었다. 이러한 결과를 통하여 우리나라 여학생들이 보통 수준의 문항들 중에서 특히 지구과학 영역을 어려워하고 있음을 알 수 있다.

한편, 여학생들이 기초 수준에서는 상대적으로 성취도가 더 높게 나타났다. 기초 수준에서는 과학의 일부 기본적인 사실을 인식하는 능력을 요구하고 있다. 기초 수준에 해당되는 문항의 수가 10개뿐이어서 전반적인 경향을 살펴보기에는 무리가 있기는 하나, 생물 영역에 해당되는 네 문항 모두 여학생의 정답률이 높게 나타나서 여학생들이 기초 수준의 문항들 중에서 특히 생물 영역에서 강세를 나타내었다고 할 수 있다.

3. 내용 영역별 남·여학생의 성취도 차이

TIMSS 2003 결과에서 내용 영역별로 남·여학생의 성취도를 비교한 결과를 Table 3에 제시하였다.

내용 영역별로 GDI 값을 비교해 보면 우리나라의 경우 지구과학 영역에서 GDI 값이 8.30으로 가장 크게 나타났고, 그 다음으로 환경 영역과 물리 영역의 GDI 값이 크게 나타났고, 화학 영역의 GDI 값이 가장 작았다. 국제 평균의 경우에도 지구과학 영역의 GDI 값이 가장 크고 화학 영역의 GDI 값이 가장 작았다. 그런데 우리나라의 GDI 값이 모든 내용 영역에서 국제 평균 GDI 값보다 높게 나타나서, 우리나라 여학생

Table 2
Gender differences in science achievement by benchmark

TIMSS international benchmark	% correct				GDI		
	Korean		International		Korean	International	Korean -International
	Female	Male	Female	Male			
Advanced International Benchmark	42.58	45.52	31.75	32.42	3.32	0.80	2.52
High International Benchmark	64.37	67.49	48.01	49.30	3.60	1.25	2.35
Intermediate International Benchmark	69.43	74.91	60.18	63.52	7.52	3.61	3.91
Low International Benchmark	87.21	87.06	74.18	73.58	-1.19	-0.40	-0.79

Table 3
Gender differences in science achievement by content area

Content Area	% correct				GDI		
	Korean		International		Korean	International	Korean -International
	Female	Male	Female	Male			
Life Science	55.85	57.57	43.61	42.53	2.10	-1.26	3.36
Chemistry	46.43	47.89	39.69	40.08	1.99	0.34	1.65
Physics	62.77	65.94	43.70	46.12	3.71	2.73	0.98
Earth Science	53.36	60.81	44.82	48.19	8.30	3.56	4.25
Environmental Science	49.87	53.27	38.01	39.60	3.82	2.02	1.80

의 성취도가 남학생의 성취도에 비하여 상대적으로 더 낮음을 알 수 있다. 한편 국제 평균에서는 생물 영역의 경우 다른 내용 영역과는 달리 GDI 값이 -1.26으로 여학생의 성취도가 높게 나타난 반면에, 우리나라에서는 모든 내용 영역에서 남학생의 성취도가 높게 나타났다. 이는 여학생들이 생물 영역의 학습에 흥미가 있고 강세를 보인다는 기존의 연구 결과(Bell, 2001; Lee & Burkam, 1996; Taber, 1991)와 상반된 것이라고 할 수 있다. 이러한 결과의 원인을 진단하고 여학생의 성취도 향상 방안을 모색하기 위하여 계속적인 연구가 수행될 필요가 있다(정은영, 2005).

우리나라의 GDI와 국제 평균의 GDI의 차이를 비교해 보면 지구과학 영역에서 4.25로 가장 크게 나타났고 물리 영역에서 0.98로 가장 작게 나타났다. 지구과학 영역에 대해서는 국제 평균과 우리나라 모두 GDI 값이 크게 나타났는데, 우리나라와 국제 평균의 GDI 차이가 다른 내용 영역에 비하여 더 크게 나타나서 우리나라 여학생들이 지구과학 영역에서 특히 더 큰 어려움을 겪고 있음을 알 수 있다.

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, 문항 내용, 우리나라 학생들의 GDI, 국제 평균 GDI, 내용 영역, 인지 영역, 문항 유형, 성취 수준, 학습 경험 등이 제시되어 있다. 총 189개의 TIMSS 문항 중 약 21%에 해당하는 39개의 문항이 이에 속하였다.

이들 문항의 분포를 내용 영역별로 살펴보면 생물 영역의 문항이 8개, 물리 영역의 문항이 10개, 지구과학 영역의 문항이 13개, 화학 영역의 문항이 5개, 환경 영역의 문항이 3개인 것으로 나타났다. TIMSS 2003 문항 중 생물 영역의 문항이 29%, 물리 영역의 문항이 24%, 지구과학과 화학 영역의 문항이 각각

16%, 환경 영역의 문항이 14%인 것을 고려할 때, 이와 같은 결과는 우리나라 남·여학생의 성취도 차이가 큰 문항 중 남학생들이 높은 성취도를 보이는 문항은 지구과학 영역에서 가장 많고, 물리 영역에서도 상대적으로 많음을 의미한다. 남학생들이 여학생들보다 높은 성취도를 나타내는 문항들 중 공개 문항을 중심으로 그 내용을 영역별로 살펴보면 다음과 같다.

지구과학 영역의 경우, 일식 동안의 달의 위치, 공이 떨어지는 방향, 항성의 예로서 태양, 목성과 달의 모습, 화석 연료와 관련된 문항에서 남학생들의 성취도가 상대적으로 높았다. 지구과학 영역에서는 GDI가 10 이상인 문항들이 13개인 것으로 나타났는데, 그 중에서 소주제인 ‘태양계와 우주 속의 지구’에 속하는 문항들이 6개였다. 이는 우리나라 여학생들이 천문학 영역에서 남학생들과 큰 차이를 보이며(신동희, 박정, 2002), 남학생이 여학생보다 지구와 달의 운동에 관하여 더 발달된 개념을 갖고 있다는 연구 결과(우종욱 등, 1995)와 일관된다고 할 수 있다.

물리 영역의 경우, 천둥 소리를 듣기 전에 번개를 먼저 보게 되는 이유, 거울에 반사된 양초의 상의 위치, 나무판에 박혀있는 못을 뽑아낼 때 따뜻해지는 이유, 바깥의 밝은 곳에 있는 사람에게 어두운 실내에 있는 사람이 잘 안 보이는 이유, 헬륨 풍선이 위로 올라가는 이유 등에 관한 문항에서 남학생들의 성취도가 여학생들보다 상대적으로 높은 것으로 나타났다. 물리 영역에서 남학생들의 성취도가 높은 10개의 문항을 인지 영역별로 분류해 본 결과, 사실적 지식을 묻는 문항이 1개, 개념 이해를 묻는 문항이 6개, 추론과 분석 능력을 평가하는 문항이 3개로 나타났다. 이와 같은 결과는 물리 영역의 경우 남학생들이 개념 이해를 묻는 문항에서 상대적으로 강세를 보이고 있음을 시사한다.

생물 영역의 경우, 세균을 파괴하는 세포, 양쪽 귀로 듣는 것이 한 쪽 귀로 듣는 것보다 유리한 점, 광합성 결과 산소가 발생하는 것을 확인하는 실험, 생물

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

item	GDI		Content Area*	Cognitive Domain**	Item Type***	Benchmark****	Learning Experience*****	
	Korean	International						
1401	Percentage of fresh water on earth	22.92	8.93	Earth S	FK	MC	A	J
1108	Water level in heated glass container	21.65	7.58	P	RA	CR	A	B
1404	Iron nail with wire coiled around it	20.35	8.17	P	FK	MC	A	B
1302	Position of the moon during solar eclipse	18.87	11.65	Earth S	CU	CR	I	N
0507	Energy for Earth's water cycle	18.43	8.36	Earth S	FK	MC	I	E
1306	Direction dropped ball will fall	17.58	4.81	Earth S	CU	MC	I	J
0909	Lightning seen before thunder heard	17.47	8.64	P	CU	CR	H	B
0203	Candle position reflected on grid	17.07	5.88	P	RA	MC	I	B
0911	Sun is an example of a star	16.19	5.96	Earth S	FK	MC	I	N
0311	Appearance of Jupiter and Moon	15.88	3.61	Earth S	CU	CR	L	N
1208	Gaps between metal rail spans	15.73	5.81	P	RA	MC	A	B
1112B	Two changes in atmospheric conditions	14.95	1.66	Earth S	CU	CR	A	N
0116	Angle of reflected light ray	14.83	5.79	P	CU	MC	H	B
0205	Increased carbon dioxide in atmosphere	14.45	7.48	ES	FK	MC	H	N
1112A	Two changes in atmospheric conditions	14.14	1.07	Earth S	CU	CR	A	J
1101	Cells that destroy bacteria	13.87	3.59	LS	FK	MC	I	J
0115	Advantage of having two ears	13.44	2.78	LS	CU	CR	A	N
0903	Photosynthesis experiment	13.37	2.73	LS	RA	CR	A	J
1410B	Life in Oceans two conditions at the bottom of ocean	13.14	6.98	LS	CU	CR	A	N
0304	Atoms removed from chair	13.09	4.00	C	CU	MC	H	N
0601	Earth year	12.54	8.48	Earth S	FK	MC	H	E
0112	Fossil fuels	12.46	5.00	Earth S	FK	MC	H	N
0213	Nail pulled out of a wooden board	12.40	8.31	P	CU	CR	I	J
0306	Nonrenewable natural resource	12.05	4.51	ES	CU	MC	A	N
1109	Gravity acting on parachute jumper	11.87	4.11	P	CU	MC	A	J
0312	Neutral atom gains electron	11.80	0.10	C	FK	MC	A	N
0606	Light from sun and moon	11.59	5.00	Earth S	CU	CR	H	N
1211	Location of jungle on temperature/elevation diagram	11.35	6.14	Earth S	RA	MC	H	N
0210	Seeing person in a dark room	11.24	2.44	P	CU	MC	I	J
0512	Heart rate increasing with exercise	11.24	3.04	LS	RA	CR	A	J
0806	Energy released during a chemical reaction	11.04	4.98	C	CU	CR	A	N
1203	Difference in shell colors of snails	10.87	-6.91	LS	CU	CR	H	N
0206	Substance Not a fossil fuel	10.77	1.48	Earth S	FK	MC	I	N
0910	Why helium balloon moves upward	10.64	5.13	P	CU	MC	H	J
0801	Classification of animals into two categories	10.61	6.62	LS	CU	MC	A	E
0501	Gas needed for rust to form	10.45	4.47	C	FK	MC	A	N
0113	Main cause of acid rain	10.41	2.96	ES	FK	MC	A	N
0404	Candles burning in 3 jars	10.22	5.22	C	RA	CR	H	E
0402	How glasses/contact lenses work	10.21	2.01	LS	RA	CR	H	N

* P: Physics, C: Chemistry, LS: Life Science, Earth S: Earth Science, ES: Environmental Science

** FK: Factual Knowledge, CU: Conceptual Understanding, RA : Reasoning and Analysis

*** MC: Multiple-Choice Item, CR: Constructed-Response Item

**** A: Advanced International Benchmark, H: High International Benchmark, I: Intermediate International Benchmark, L: Low International Benchmark

***** B: Elementary & Junior School, E: Elementary School, J: Junior School, N: None

을 살기 어렵게 하는 해양 바닥의 조건, 운동을 하면 심장 박동이 빨라지는 이유, 달팽이 껍질의 색깔 차이가 달팽이의 생존에 유리한 점, 동물 분류의 기준, 시력이 나쁜 사람이 안경이나 콘택트렌즈를 끼면 훨씬 선명하게 보이는 이유 등에 대한 문항에서 남학생들이 여학생들보다 더 높은 성취도를 나타내었다. 생물 영역에서 GDI가 10 이상인 총 8개의 문항들 중에서 1101 문항을 제외하고 나머지 7개의 문항들은 모두 자유응답형 문항이었다. 이는 여학생들이 남학생들보다 서술형 문항에서 더 높은 성취도를 얻는 경향이 있다는 기존의 연구 결과(신동희, 김동영, 2003; Bolger & Kellaghan, 1990; DeMars, 1998; Din *et al.*, 2004)와는 상반된다. 그런데 이러한 문항들의 경우, 우리나라 과학 교육과정에서 포함되어 있지 않거나 수업 시간에 직접적으로 다루지 않는 내용에 관한 것이어서, 우리나라 여학생들은 학습하지 않은 내용에 관한 문항에서 남학생들보다 어려움을 더 겪는다고 할 수 있다. 그리고 채점 기준표를 살펴보면 과정에 대한 이해를 요구하고 있는데, 여학생들은 과정을 이해하기 보다는 정답을 제시하는 데 중점을 두는 경향이 있어서(Haggerty, 1991), 이러한 문항들에서 어려움을 겪은 것으로 생각된다.

화학 영역의 경우에는 원자 개념과 원자 구조, 화학 반응에서의 에너지 출입, 부식의 원인이 되는 기체 등 TIMSS 2003의 대상 학생들이 과학 수업에서 배우지 않은 내용에서 남학생들의 성취도가 높게 나타났다. 또한 환경 영역의 문항 중에서는 재생 불가능한 천연 자원, 산성비의 주된 원인에 대하여 묻는 문항에서 남학생들의 성취도가 높았다.

우리나라 학생들의 GDI가 큰 문항 중 남학생들이 잘한 문항을 인지 영역별로 살펴보면 사실적 지식을 묻는 문항이 12개, 개념 이해를 묻는 문항이 19개, 추론과 분석 능력을 요구하는 문항이 8개인 것으로 나타났다. TIMSS 2003 과학 문항의 인지 영역별 비율에서 사실적 지식을 묻는 문항이 30%, 개념 이해를 묻는 문항이 39%, 추론과 분석을 평가하는 문항이 31%인 것을 고려할 때, 이와 같은 결과는 남학생들이 강세를 보이는 부분은 개념 이해를 요구하는 문항과 사실적 지식을 묻는 문항들임을 시사한다. 한편 추론과 분석 능력을 요구하는 문항의 비율이 31%인 것에 비하여 남학생들이 잘한 문항은 8개로 나타나서 이 영역에서는 상대적으로 남·여학생의 성취도 차이가 크지 않았다.

Table 4의 문항을 문항 유형별로 살펴보면 총 39 문항 중에서 23문항이 선택형, 16문항이 자유응답형이었다. 이러한 결과는 문항 유형별로 문항의 GDI 평

균을 비교한 결과(Table 1 참조)에서와 마찬가지로, 남학생들이 선택형 문항에서 더 강세를 나타냄을 의미한다.

Table 4의 문항을 성취 수준별로 살펴보면, 수월 수준에 해당하는 문항이 17개, 우수 수준에 해당하는 문항이 12개, 보통 수준에 해당하는 문항이 9개, 기초 수준에 해당하는 문항이 1개였다. TIMSS 2003 문항 중 각 성취수준별 문항 수(수월 수준 95문항, 우수 수준 61문항, 보통 수준 23문항, 기초 수준 10문항)를 고려해 볼 때, 우리나라 남학생들은 보통 수준의 문항들에서 상대적으로 더 높은 성취도를 나타내었음을 알 수 있다. 이와 같은 결과는 성취 수준별로 문항의 GDI 평균을 구하여 비교한 결과(Table 2 참조)와 일치한다. 한편 TIMSS-R 결과를 대상으로 한 연구(이미경과 홍미영 등, 2004)에서는 남·여학생의 GDI 차이가 큰 문항들이 상위 25%와 하위 25%에 해당하는 경우가 상대적으로 많았다. TIMSS 2003에서 성취수준별 문항의 비율과 TIMSS-R에서 그 비율(상위 10%에 해당하는 문항 36%, 상위 25%에 해당하는 문항 28%, 상위 50%에 해당하는 문항 22%, 하위 25%에 해당하는 문항 14%)에 차이가 있고 성취수준을 구분하는 방법이 달라져서 비교하기에 무리가 있는데, TIMSS-R과 TIMSS 2003 결과에서 공통적인 것은 수월 수준(상위 10%)에서는 남학생과 여학생의 성취도 차이가 상대적으로 적게 나타난 점이다. 즉, 우리나라 여학생들은 복합적이고 추상적인 과학 개념과 기본적인 과학 탐구의 이해, 기본적인 과학 원리를 적용한 정량적인 문제의 해결, 과학 지식을 글로 설명하기 등에서는 남학생들과의 차이가 상대적으로 적다고 할 수 있다.

Table 4의 문항 분포를 학습 경험을 기준으로 살펴 보면, 초등학교와 중학교 2학년 때까지 배운 문항이 6개, 초등학교에서만 배운 문항이 4개, 중학교 1, 2학년 과정에서만 배운 문항이 10개, 교육과정에서 다루어지지 않은 문항이 19개인 것으로 나타났다. 이와 같이 교육과정에서 다루어지지 않은 문항의 비율이 비교적 높게 나타난 것은 학교 밖에서 습득한 지식에 대해서는 남학생의 성취도가 더 높게 나타난다는 Hamilton(1998)의 연구 결과와 유사하다. 한편, TIMSS 2003의 생물 영역 문항들 중에서 학습하지 않은 내용의 문항들에 대하여 우리나라 남학생과 여학생의 정답률 차이가 .05 수준에서 유의미하게 나타났다(정은영, 2005). 이러한 결과를 볼 때 우리나라 여학생들이 수업 시간에 다루지 않는 내용에 대하여 상대적으로 낮은 성취도를 나타내는 경향이 있음을 알 수 있다. 이러한 경향은 남학생들의 과학에 대한 관심이 여학

Table 5

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

item	GDI		Content Area*	Cognitive Domain**	Item Type***	Benchmark****	Learning Experience*****
	Korean	International					
0103 Time/temperature table	-18.94	-0.38	Earth S	RA	MC	I	E
0708 Sound of plucked guitar string	-15.01	-2.85	P	FK	MC	A	J
0612 Different materials sorted into two groups	-11.30	-2.27	C	CU	MC	A	B
0706 Diagram of a person doing work	-10.73	-1.88	P	CU	MC	L	J
0701 Which organ system	-10.63	-2.67	LS	FK	MC	L	J

* P: Physics, C: Chemistry, LS: Life Science, Earth S: Earth Science, ES: Environmental Science

** FK: Factual Knowledge, CU: Conceptual Understanding, RA: Reasoning and Analysis

*** MC: Multiple-Choice Item, CR: Constructed-Response Item

**** A: Advanced International Benchmark, H: High International Benchmark, I: Intermediate International Benchmark, L: Low International Benchmark

***** B: Elementary & Junior School E: Elementary School J: Junior School N: None

생들보다 높은 데서 기인한 것이라고 볼 수도 있을 것이다.

2) 여학생들의 성취도가 높은 문항

Table 5는 우리나라 학생들의 GDI가 10 이상인 문항 중 여학생들의 성취도가 높은 문항에 대한 정보를 나타낸 것이다. TIMSS 2003 문항 중 여기에 해당하는 문항은 총 189문항 중 5문항에 불과한 것으로 나타났으며, 물리 영역에서 두 문항, 화학, 생물, 지구 과학 영역에서 각각 한 문항씩 포함되었다. 이 문항들 중에서 수월 수준에 속하는 문항이 두 개, 보통 수준에 속하는 문항이 한 개, 기초 수준에 속하는 문항이 두 개로, 모든 성취수준의 문항이 비교적 골고루 포함되었다. 한편 여학생들의 성취도가 높았던 문항의 유형은 모두 선택형이었다. 문항의 인지 영역을 살펴보면 추론과 분석에 속하는 문항이 한 개, 사실을 묻는 문항과 개념 이해를 묻는 문항이 각각 두 개씩인 것으로 나타났다. GDI가 큰 문항 중 여학생들이 높은 성취도를 보이는 문항이 다섯 문항에 불과하여 여학생들이 강세를 보이는 문항들로부터 어떤 경향성을 찾기는 어렵다.

여학생이 강세를 보이는 문항을 구체적으로 살펴보면, 시간에 따른 기온 변화를 나타낸 표를 해석하는 문항, 기타줄을 튕겼을 때 나는 소리에 관한 문항, 금속과 비금속을 전기 전도도에 따라 분류하는 문항, 그림을 보고 일을 한 것에 해당하는 것을 고르는 문항, 기관의 종류를 제시하고 어떤 기관계에 해당되는지를 선택하도록 한 문항 등이었다.

IV. 결론 및 제언

이 연구에서는 TIMSS 2003 과학의 총 189 문항

의 GDI(Gender Difference Index) 값을 산출하여 문항 유형, 성취수준, 내용 영역에 따른 남학생과 여학생의 성취도 차이를 살펴보았다. 그 결과, TIMSS 2003의 과학 성취도에서 우리나라와 국제 결과 모두 남학생의 성취도가 더 높게 나타난 경향이 있었는데, 우리나라의 GDI 값이 국제 평균 GDI 값보다 높게 나타났으므로 우리나라 남·여학생의 성취도 차이가 상대적으로 더 크다고 할 수 있다.

문항 유형별 성차를 살펴본 결과, 국제 결과를 보면 자유응답형 문항에서는 비교적 성차가 미미하였는데, 우리나라 여학생들은 특히 선다형 문항에서 남학생보다 성취도가 낮았다.

성취수준별 성차를 살펴본 결과, 남·여학생의 성취도 차이가 모든 성취수준에서 동일한 양상으로 나타나지는 않았다. 우리나라 여학생들은 기초 수준에서는 상대적으로 높은 성취도를 나타낸 반면에, 보통 수준에서 상대적으로 낮은 성취도를 나타내었다. 이는 우리나라 여학생들이 과학의 기본적인 사실을 인식하는 능력에서는 더 우수하나, 과학 지식을 적용하고 의사소통하는 능력에서는 부족함을 의미한다. 이러한 결과는 여학생들에게 과학을 지도하는 과정에서 다양한 맥락과 예를 제시하면서 과학 지식을 적용하여 해결하도록 안내를 하고 자신의 생각을 논리적으로 표현하도록 독려할 필요가 있음을 시사한다.

내용 영역별 성차를 살펴본 결과, 우리나라와 국제 평균 모두 지구과학 영역에서 차이가 크게 나타났고 화학 영역에서 그 차이가 가장 작았다. 우리나라와 국제 평균의 GDI 값 차이가 다른 내용 영역에 비하여 지구과학 영역에서 더 크게 나타나서 우리나라 여학생들이 지구과학 영역에서 더 큰 어려움을 겪고 있다고 할 수 있다. 한편 국제 평균에서는 생물 영역의 경

우 여학생의 성취도가 다른 내용 영역과는 달리 더 높게 나타난 반면에 우리나라에서는 모든 내용 영역에서 남학생의 성취도가 높게 나타났다. 우리나라 여학생들의 경우 생물 영역에서도 상대적으로 낮은 성취도를 나타내는 원인에 대해서 분석하기 위한 연구가 수행될 필요가 있다.

GDI가 10 이상인 문항 중 남학생들의 성취도가 높은 문항들을 살펴본 결과, 지구과학 영역에 속하는 경우가 가장 많았으며, 물리 영역의 문항도 상대적으로 많았다. 학교 교육과정에서 전혀 배우지 않은 문항들의 비율 또한 비교적 높았다. 인지 영역별 분포를 볼 때 사실적 지식과 개념 이해를 묻는 문항에서 남학생들이 강세를 보이는 경향이 있었으며, 추론과 분석 능력을 평가하는 문항에서는 상대적으로 그 차이가 작았다. 이러한 경향은 여학생들의 과학 성취도를 향상하기 위한 교수·학습 방법 개발에 시사하는 바가 크다. 과학 수업에서 배운 내용을 잘 이해하고 있는지를 점검하고, 스스로 이해한 내용을 설명하게 하는 기회를 제공하는 등 학습 과정에서 학습 결손이 이루어지지 않도록 주의하는 것은 여학생들의 성취도를 향상시키기 위한 방안이 될 수 있을 것이다.

한편, 내용 영역, 문항 형태에 따른 성차 연구에서 평가 대상이나 평가 주제 등에 따라 일관되지 않은 결과가 보고되고 있다는 고려해 볼 때, 성차를 줄이기 위한 방안으로 여학생 친화적인 평가 방법이나 소재 선택 등에서 한발 더 나아가는 것이 필요하다. TIMSS 자료를 이용하여 성취도가 국제 최상위 수준에 해당하는 싱가포르, 한국, 홍콩, 일본 아시아 4개국의 수학에서의 성차를 연구한 결과(Soh & Quek, 2001)에 의하면, 다른 나라와 달리 싱가포르의 오히려 여학생들의 성취도가 높게 나타났는데, 이에 대한 원인으로는 싱가포르의 다문화적, 다인종적 환경으로 인하여 아시아 국가임에도 불구하고 여학생과 남학생을 차별하지 않는(non-sexist) 교육환경을 꼽고 있다. 따라서 성차를 줄이기 위한 방안 중 하나로 남학생과 여학생이 평등하게 대우받는(gender-equitable) 교육 환경이 확립되어야 할 것이다.

국문 요약

이 연구에서는 TIMSS 2003 과학 문항의 GDI (Gender Difference Index) 값을 산출하여 문항 유형, 성취수준, 내용 영역에 따른 남학생과 여학생의 성취도 차이를 살펴보았다.

TIMSS 2003의 과학 성취도에서 문항 유형에 상관 없이 우리나라와 국제 결과 모두 남학생이 여학생보

다 성취도가 높았으며, 우리나라 여학생들은 특히 선다형 문항에서 남학생보다 성취도가 낮았다. 성취수준에 따른 남·여학생의 성취도를 비교한 결과, 우리나라와 국제 평균 모두 기초 수준을 제외하고 다른 3개의 성취수준에서 남학생의 성취도가 여학생의 성취도에 비하여 높게 나타났고, 보통 수준에서 그 차이가 가장 크고 기초 수준에서 가장 작았다. 내용 영역별로 남·여학생의 성취도를 비교한 결과, 우리나라와 국제 평균 모두 지구과학 영역에서 차이가 크게 나타났고 화학 영역에서 그 차이가 가장 작았다. 생물 영역의 경우 국제 평균에서는 여학생의 성취도가 높게 나타난 반면에, 우리나라에서는 남학생의 성취도가 높게 나타났다. 남학생들의 성취도가 높은 문항들은 지구과학 영역에 속하는 경우가 상대적으로 많았으며, 학교 교육과정에서 다루어지지 않은 문항들의 비율 또한 비교적 높았다. 인지 영역별 분포를 볼 때는 남학생들이 높은 성취도를 보이는 문항은 사실적 지식과 개념 이해를 묻는 문항인 경우가 많았다. 연구 결과에 근거하여 과학 교육에서 성차를 줄이기 위한 방안들을 논의하였다.

참고 문헌

- 박정, 홍미영, 이봉주, 전현정 (2003). 수학과 과학 성취도 추이변화 국제비교 연구(TIMSS- 2003) - 본검사 시행보고서. 한국교육과정평가원.
- 박정, 정은영, 김경희, 한경혜 (2004). 수학과 과학 성취도 추이변화 국제비교 연구 -TIMSS 2003 결과 보고서. 한국교육과정평가원.
- 신동희, 김동영 (2003). 평가 방법에 따른 과학 성취도에서의 성 차이. 한국과학교육학회지, 23(3), 265-275.
- 신동희, 박정 (2002). 국제비교연구에 나타난 우리나라 학생들의 지구과학 성취도: 성 차이를 중심으로. 한국지구과학교육학회지, 23(3), 207-220.
- 우종욱, 이항로, 민준규 (1995). 계통도를 이용한 중, 고등학생들의 지구와 달의 운동에 관한 개념 유형 연구. 한국과학교육학회지, 15(4), 379-393.
- 이미경, 박영순, 민경석, 채선희, 최성연 (2004). PISA 2003 결과 분석 연구 -수학적 소양, 읽기 소양, 과학적 소양 수준 및 배경변인 분석-. 한국교육과정평가원 연구보고 RRE 2004-2-1.
- 이미경, 홍미영, 정은영 (2004). TIMSS-R 과학 성취도에서의 성 차이. 한국과학교육학회지, 24(6), 1235-1244.
- 정은영 (2005). TIMSS 2003에서 우리나라 중학생들의 생물 영역 성취도 분석. 한국생물교육학회지, 33(3), 277-290.
- 정해숙, 정경아 (2003). 교사의 성인지적 교육활동 능력 제고를 위한 프로그램 개발. 한국어성개발원.

진희정 (2002). 여학생 친화적인 과학 교육 프로젝트의 내용 비교 및 분석. 이화여자대학교 대학원 석사 학위 논문.

Bell, J. F. (2001). Investigating gender differences in the science performance of 16-year-old pupils in the UK. *International Journal of Science Education*, 23(5), 469-486.

Bolger, N., & Kellaghan, T. (1990). Method of measurement and gender differences in scholastic achievement. *Journal of Educational Measurement*, 28, 23-35.

Breland, H. M., Danos, D. O., Kahn, H. D., Kobota, M. Y., & Bonner, M. W. (1994). Performance versus objective testing and gender: An explanatory study of an advanced placement history examination. *Journal of Educational Measurement*, 31, 37-50.

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.

Din, Y. Y., Ming, M. C., & Esther, S. C. H. (2004). Hong Kong student achievement in OECD-PISA study: Gender differences in science content, literacy skills, and test item formats. *International Journal of Science and Mathematics Education*, 2(1), 91-106.

Hamilton, L. S. (1998). Gender differences on high school science achievement tests: Do format and content matter? *Educational Evaluation and Policy Analysis*, 20(3), 179-195.

Haggerty, S. M. (1991). Gender and school science: Achievement and participation in Canada. *Alberta Journal of Educational Research*, 37(3), 195-208.

Lee, V. E., & Burkam, D. T. (1996). Gender differences in middle grade science achievement: Subject domain, ability level, and course emphasis. *Science Education*, 80(6), 613-650.

Martin, M. O., Mullis, I. V. S., Gonzalez, E. J., & Chrostowski, S. J. (2004). TIMSS 2003 international science report: Findings from IEA's Trends in International Mathematics and Science Study at the fourth and eighth grades. Boston: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College.

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.

Soh, K. C., & Quek, K. S. (2001). Gender difference in TIMSS mathematics achievement of four Asian nations: A secondary analysis. *Studies in Educational Evaluation*, 27, 331-340.

Taber, K. S. (1991). Gender differences in science preferences on starting secondary school. *Research in Science and Technological Education*, 9(2), 245-252.

Wester, A., & Henriksson, W. (2000). The interaction between item format and gender difference in mathematics performance based on TIMSS data. *Studies in Educational Evaluation*, 26, 79-90.