

## 학업 성취도 국제 비교 연구에 나타난 우리나라 학생들의 수학 성취도 심층 분석<sup>1)</sup>

박 경 미\*

본 연구는 최근에 이루어진 대표적인 수학 성취도 국제 비교 연구인 PISA와 TIMSS-R의 성취도를 표준화하여 비교함으로써 한국과 일본, 그리고 미국과 호주를 비롯한 영어권 국가들이 각각 유사한 성취 경향을 보인다는 사실을 확인하였다. 이를 보다 체계화하기 위해 PISA의 결과를 문항반응이론에 의해 분석하여 국가별, 문항별 난이도를 조사하였다. 문항반응이론에 근거하여 각 국가의 학생들에게 상대적으로 더 어렵고 쉬운 문항을 판별한 결과, 30개 중 16개의 문항에서 한국과 일본, 그리고 미국과 호주가 각각 동일한 성취 경향을 보이는 것으로 나타났다. 이러한 성취 경향의 유사성을 심층적으로 분석하기 위하여 한국과 일본, 그리고 미국과 호주가 각각 강세를 보인 문항들을 추출하여 그 공통점을 탐색하였다. 미국과 호주의 학생들은 열린 구성반응형 문항을 해결하는데 상대적으로 능숙하고, 한국과 일본의 학생들은 함수와 통계 영역의 문항에서 성취 수준이 높지 않으며, 형식화된 수학적 지식을 요구하는 문항일수록 한국과 일본 학생들에게 유리하게 작용하는 것으로 나타났다.

### 1. 서 론

우리나라는 최근에 이루어진 학업 성취도 국제 비교 연구인 PISA(Program for International Student Assessment)의 수학 영역에서 2위, TIMSS(Third International Mathematics and Science Study)와 그 반복 연구인 TIMSS-R(Repeat)의 수학 영역에서 역시 2위의 성취도를 기록하면서 국내외 수학교육 연구자의 관심의 초점이 되어 왔다(OECD, 2001; Beaton, et al, 1996; Mullis, et al, 2000; 노국향 외, 2001, 박정 외 2001). PISA는 OECD가 주관하는 국제 비교 연구로 학생들이 건전한 민주시민으로 살아가는 데 필요한 '수학적 소양'의 측정을 목적으로 하며, 계산을

위주로 측정하는 이전의 검사와 달리 고등 사고 기능을 요구하는 문항이 다수 포함되어 있어 이전의 연구와 차별화 된다(OECD, 2000). 또한 TIMSS와 TIMSS-R은 현재까지 이루어진 국제 비교 연구 중 참여국과 표집 대상의 규모가 가장 크고 연구 방법면에서도 가장 체계성을 띤 신뢰할 만한 연구로, 4년 시차를 두고 유사한 검사를 반복하여 학생들의 성취도 변화 추이를 추적할 수 있기 때문에 풍부한 정보의 원천을 제공한다.

우리나라의 TIMSS-R 검사 대상은 1999년 2월 중학교 2학년 재학생으로, 이 학생들은 PISA가 실시된 2000년 7월에 고등학교 1학년으로 진급하여 PISA의 검사 대상인 만 15세가 되었

\* 홍익대학교, kpark@hongik.ac.kr

1) 이 논문은 2002년도 한국학술진흥재단의 지원에 의하여 연구되었음 (KRF-2002-003-B00258).

기 때문에 두 연구의 모집단은 동일하다. 이러한 점을 고려할 때 학생 성취도의 변화 과정을 탐색하기 위하여 PISA와 TIMSS-R의 결과를 연결지어 종합적으로 분석할 필요가 있다. 본 연구에서는 PISA와 TIMSS-R의 결과를 하나의 척도로 표준화하고 두 검사에 모두 참가한 국가들의 성취 수준을 비교하였다.

국제 비교 연구의 일차적인 목적은 우리나라 학생의 수학적 수준을 국제적인 잣대에 의해 판정하는 것이지만, 그 진정한 의의는 국제 비교 연구 결과에 비추어 수학 교육과정과 교과서의 내용과 수업 방법 등 수학교육의 전반을 점검하는 데 있다. 본 연구는 PISA에 대한 한국, 일본, 미국, 호주의 수학 성취도 결과를 문항반응이론에 의거하여 심층적으로 분석하고 우리나라 학생들의 강점과 취약점을 파악하면서 수학교육의 방향에 대한 시사점을 도출하고자 한다.

## II. PISA와 TIMSS-R의 수학 성취 수준 비교

PISA에는 32개국이 참가하였고, TIMSS-R의 참가국은 38개국이며, 이 두 연구에 모두 참여한 국가는 13개국이다. PISA와 TIMSS-R의 결과를 비교하기 위하여 공통으로 참가한 13개국가 내에서의 수학 성취 순위를 조사하였다. 전체 참가국에서의 순위와 두 검사에 모두 참여한 13개국 내에서의 순위 사이에는 약간의 차

이가 있다. 예를 들어 우리나라는 TIMSS-R에서 2위를 기록하였으나, 1위였던 싱가포르가 PISA에 참여하지 않았기 때문에 비교 대상에서 제외되었고, 따라서 13개국 중 우리나라의 수학 성취도 순위는 1위이다. <표 II-1>에 제시한 PISA와 TIMSS-R의 성취도 순위를 통해 두 검사의 결과를 대략적으로 비교할 수 있다. 예를 들어 뉴질랜드는 TIMSS-R에 비해 PISA의 순위가 훨씬 높은 반면 헝가리나 러시아는 PISA보다 TIMSS-R의 성취 수준이 상대적으로 높다. 그러나 이는 두 검사의 순위를 중심으로 한 단순한 비교 결과이므로, 보다 정확한 비교를 위해서는 추가의 분석이 필요하다.

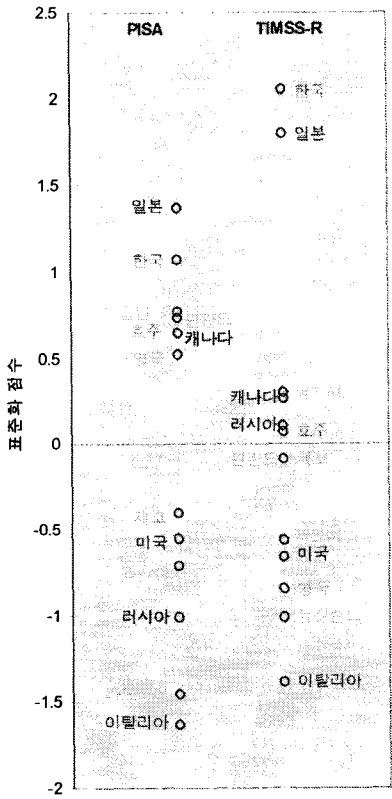
### 1. PISA와 TIMSS-R 수학 성취 수준의 표준화 점수 비교

PISA와 TIMSS-R은 평균과 표준편차가 다르고 서로 다른 척도를 사용하여 점수화하기 때문에 두 연구 결과를 비교하기 위해서는 동일한 방식으로 표준화할 필요가 있다. 본 연구에서는 우선 PISA와 TIMSS-R에 대한 비교 대상 13개국의 성취도 평균과 표준편차를 각각 구했다. 그리고 각 국가의 수학 성취도 평균과 13개국의 수학 성취도 평균의 차이를 표준편차로 나누어 각 국가의 표준화 점수를 산출하였다. [그림 II-1]을 보면 PISA에서의 일본의 표준화 점수는 1.37인데, 이 수치는 일본의 PISA 수학 성취도 평균이 13개국의 수학 성취도 평균보다 표준편차의 1.37배 높음을 의미한다. 마찬가지로 한국의 TIMSS-R 표준화 점수는 2.05인데,

<표 II-1> PISA와 TIMSS-R의 수학 성취도 순위

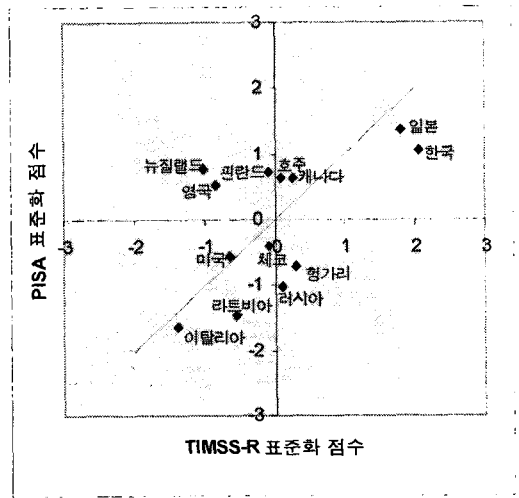
국가 검사	일본	한국	뉴질랜드	핀란드	캐나다	호주	영국	체코	미국	헝가리	러시아	라트비아	이탈리아
PISA	1	2	3	4	5	6	7	8	9	10	11	12	13
TIMSS-R	2	1	12	8	4	6	11	7	10	3	5	9	13

이는 한국의 TIMSS-R 수학 성취도 평균이 13개국의 수학 성취도 평균보다 표준편차의 2.05배 높다는 의미이다. [그림 II-1]의 표준화 점수를 비교해 보면 몇 가지 주목할만한 경향성을 발견할 수 있다. 우선 TIMSS-R에서는 한국 및 일본과 나머지 국가들 사이에 큰 차이가 있지만, PISA에서는 그 차이가 현격하게 감소하였다. 또한 PISA에서 일본과 한국에 뒤이은 국가 중 핀란드를 제외하고는 뉴질랜드, 캐나다, 호주, 영국과 같이 모두 영어권 국가들이다. 한편 헝가리, 체코, 러시아와 같은 동구권의 국가들은 PISA의 성취 수준이 TIMSS-R에 비해 낮은 편이며, 이탈리아는 두 검사 모두에서 13개국 중 가장 낮은 성취 수준을 보였다.



[그림 II-1] PISA와 TIMSS-R의 표준화 점수 분포

이런 경향성은 [그림 II-2]를 통해 더욱 명료하게 드러난다. TIMSS-R과 PISA의 성취도를 각각  $x$  축과  $y$  축으로 하고, 각 국가의 TIMSS-R과 PISA의 표준화 점수의 위치를 표시한 후 기울기가  $45^\circ$ 인 직선을 그린다. 그러면 일부의 국가들은 직선 위에, 일부의 국가들은 직선 아래에 표시된다. 기울기가  $45^\circ$ 인 직선 아래에 위치하는 국가들은 TIMSS-R의 성취 수준이 PISA의 성취 수준보다 높은 국가로 아시아의 한국과 일본, 그리고 동유럽의 체코, 헝가리, 라트비아, 러시아가 여기에 해당된다. 서유럽의 국가로는 유일하게 이탈리아가 직선 아래에 위치하고 있다. 이와 반대로 직선 위에 표시된 국가들은 PISA의 성취 수준이 상대적으로 높은 경우로 호주, 뉴질랜드, 캐나다, 영국, 미국과 같은 영어권 국가와 핀란드가 이에 해당된다.



[그림 II-2] PISA와 TIMSS-R의 표준화 점수 산점도

## 2. PISA와 TIMSS-R의 상이한 수학 성취 경향의 원인

PISA와 TIMSS-R의 결과에서 나타난 상이한

성취 수준의 원인은 무엇일까? 우선적으로 생각할 수 있는 측면은 PISA와 TIMSS-R의 표집 대상의 차이이다. TIMSS-R의 검사 대상은 참가국의 8학년(우리나라의 경우 중학교 2학년) 학생들이고, PISA는 만 15세 학생들이다.

TIMSS-R은 학년을 기준으로 한 표집(grade-based sampling) 방법을 따르기 때문에 8학년 표집 대상 중에는 상이한 연령의 학생들이 포함되어 있다. 그에 반해 나이를 기준으로 표집하는(age-based sampling) PISA의 만 15세 표집 대상에는 상이한 학년의 학생이 혼재되어 있다.

그런데 이러한 차이가 두 검사에서의 성취 수준의 차이를 유발시킬 수는 있지만 결정적인 영향력을 행사했다고 보기는 어려울 것이다.

일례로 동유럽 국가에서는 다른 참가국에 비해 학생들의 수학적 능력에 따른 속진 제도가 활성화되어 있다. 그렇다면 PISA의 표집 대상인 만 15세 학생 중에는 속진을 하여 고등학교 2, 3학년인 학생이 일부 포함될 수 있고, 이는 PISA에서 동유럽 학생들의 성취 수준을 높이는 요인으로 작용할 가능성이 높다. 그럼에도 불구하고 헝가리와 체코와 러시아에서 PISA의 성취 수준이 TIMSS-R에 비해 상대적으로 낮았다는 점을 고려할 때, 표집이 학년 중심으로 이루어졌는지 나이에 근거했는지가 성취 결과에 그리 심각한 영향을 미치지 않은 것으로 보인다. 물론 이에 대해 추가로 분석할 필요가 있을지도 모르지만 본 연구에서는 표집 방법의 차이를 일차적인 탐구 주제로 삼지는 않았다.

한편 PISA와 TIMSS-R의 참가국마다 취학 연령과 학제가 다르기 때문에 TIMSS-R의 8학년과 PISA의 만 15세의 위계가 다르기는 하지만, 대부분의 국가에서는 PISA의 검사 대상이 TIMSS-R의 검사 대상보다 1~2년 정도 상위 학년이다. 이로부터 PISA의 성취 수준이

TIMSS-R보다 상대적으로 높은 국가는 학교급과 학년이 높아짐에 따라 학생들의 수학 성취 수준이 상승하는 경향이 있는 반면, PISA의 성취 수준이 TIMSS-R에 비해 상대적으로 낮은 국가는 학교급과 학년의 상승에 따라 학생들의 성취 수준이 하락하는 경향이 있다고 해석할 수도 있다. 그렇지만 불과 1~2년 차이로 학생들의 성취 수준의 변화 추이를 논하는 것은 그리 적절하지 않을 것이다. 다시 말해, 표집 방법이나 연령이 검사 결과의 차이를 유발한 결정적인 요인이라고 보기는 어렵다.

그렇다면 이제 PISA와 TIMSS-R의 성취 수준이 상이하게 나타난 원인으로 두 검사의 내용적인 측면에 주목할 필요가 있다. TIMSS-R은 교육과정에 기반을 둔 검사이므로 문항의 제작에 앞서 참여국들의 교육과정을 조사하여 공통적인 내용 요소를 추출하는 작업을 진행한다. 이에 기초하여 각 국가에서 다루는 공통분모의 내용에서 문항을 출제하기 때문에 TIMSS-R은 PISA에 비해 소위 교과서에서 제시한 전형적인 문항을 다수 포함한다. 이에 반해 PISA는 ‘수학적 소양(mathematical literacy)’의 측정을 목적으로 하므로, 수학이 활용되는 자연스러운 상황을 위주로 한다.

기존의 연구가 지적하고 있는 바와 같이 우리나라의 수학 교과서는 수학과 실생활의 연결 고리를 충분히 제시하지 않고 수학의 계통성과 엄밀성을 충실히 따르면서 형식적인 연역 체계로 수학을 제시하는 경향이 있다(박경미 2000). 물론 제7차 교육과정과 그에 따른 교과서는 실생활 맥락의 수학 내용 도입을 강조하고 있지만, 여전히 서구 국가에 비하면 수학의 내용적인 측면이 강조되어 있다. 일본 역시 이러한 경향성을 지니기 때문에 우리나라와 일본이 교육과정과 긴밀하게 연계된 검사인 TIMSS-R에

서 상대적으로 더 높은 성취 수준을 보인 것이라고 추측할 수 있다. 비슷한 맥락에서, 실생활을 중시하는 서구의 교과서로 수학을 학습한 학생들은 실제 상황을 중심으로 문항을 제작한 PISA에서 높은 성취 수준을 보일 가능성이 높다. 이처럼 교육과정과 교과서가 수학적인 측면과 실생활과의 관련성 중 어느 쪽에 무게중심을 두느냐에 따라 PISA와 TIMSS-R에서 상대적인 성취 수준이 달라질 수 있다.

### III. PISA 수학 성취 수준의 심층 분석

PISA와 TIMSS-R의 수학 성취도 결과 비교에서 알 수 있는 바와 같이 아시아의 한국과 일본, 그리고 영어권의 국가들은 비슷한 성취 경향을 보인다. 이러한 동질성을 확인하기 위해 한국과 일본, 그리고 영어권의 국가 중에서 미국과 호주를 선택하고 PISA의 문항에 대한 각국의 정답률과 난이도를 심층적으로 비교·분석하였다.<sup>2)</sup>

#### 1. PISA 수학 검사 문항의 상대적 난이도

PISA의 검사지에 포함된 수학 문항은 모두 31문항이나 일본의 번역 오류로 제외된 한 문항을 예외로 하면 모두 30문항이 된다. 이 30문항에 대한 한국, 일본, 미국, 호주의 정답률은 <표Ⅲ-1>과 같다. PISA의 종합적인 수학 성취도 순위는 ‘일본 > 한국 > 호주 > 미국’이므로, 각 문항의 정답률에서도 이 순위가 어느 정도 유지된다. 총 30개의 문항 중에서 14개의

<표Ⅲ-1> PISA 수학 문항의 정답률

문항 코드	문항 번호	일본	한국	호주	미국
M033Q01	1	81.5	74.0	77.4	72.5
M034Q01T	2	56.5	51.3	43.9	29.0
M037Q01T	3	81.6	70.9	66.8	46.4
M037Q02T	4	85.6	80.0	63.4	59.8
M124Q01	5	46.1	41.2	30.8	25.5
M124Q03T	6	37.2	11.7	19.9	17.6
M136Q01T	7	81.5	73.4	61.7	53.4
M136Q02T	8	50.8	60.6	25.4	23.9
M136Q03T	9	21.1	30.4	19.3	14.8
M144Q01T	10	84.9	78.6	72.9	52.7
M144Q02T	11	41.7	35.5	29.6	12.4
M144Q03	12	85.8	78.8	86.1	74.3
M144Q04T	13	49.9	49.8	43.0	34.8
M145Q01T	14	72.6	63.7	64.6	52.4
M148Q02T	15	23.3	15.2	26.7	21.1
M150Q01	16	76.6	77.5	64.3	50.8
M150Q02T	17	77.5	86.7	73.1	61.2
M150Q03T	18	45.0	48.3	63.7	57.2
M155Q02T	19	63.5	68.0	73.5	64.2
M155Q03T	20	22.5	22.9	19.2	18.1
M155Q04T	21	62.6	60.2	59.8	54.3
M159Q01	22	82.2	75.5	75.4	62.3
M159Q02	23	90.2	90.9	90.7	83.2
M159Q03	24	87.9	86.9	88.9	81.6
M159Q05	25	53.9	32.8	36.0	22.6
M161Q01	26	72.4	63.1	60.3	45.5
M179Q01T	27	24.9	28.0	37.5	30.0
M192Q01T	28	59.0	52.0	46.0	28.0
M266Q01T	29	42.6	35.9	24.4	12.8
M273Q01T	30	68.3	57.3	56.4	47.6

2) 미국과 호주라는 두 국가의 분석을 통해 영어권, 혹은 서구의 국가들의 경향으로 확대해석하고 객관화하는 것은 적절하지 않을 수 있으나, 심층적인 문항 분석을 위해 자료 분석의 범위를 두 국가로 제한하였다.

문항(2, 3, 4, 5, 7, 10, 11, 13, 21, 22, 26, 28, 29, 30)은 전체적인 성취 수준의 순위를 따라 일본의 정답률이 가장 높고 이어 한국, 호주, 미국의 순서로 정답률이 낮아진다. 그러나 정답률의 순위가 다른 경우도 적지 않다. 예를 들어 문항 8, 9, 17의 경우는 한국의 정답률이 일본보다 높으며, 문항 18의 경우는 한국과 일본보다 호주와 미국의 정답률이 훨씬 높은 것으로 나타났다. 정답률의 순위가 전체적인 성취 수준의 순위를 그대로 따르는 문항이 있고, 그 순위가 뒤바뀌는 문항이 있다는 것은 국가에 따라 학생들에게 상대적으로 어렵고 쉬운 문항이 있음을 의미한다.

문항반응이론을 이용하여 문항의 난이도를 계산하면 특정 문항이 특정 국가의 학생들에게 상대적으로 어렵게 혹은 쉽게 작용했는지 판정할 수 있다. 우리나라와 일본은 전반적인 성취 수준이 높기 때문에 각 문항에 대한 정답률도 대체적으로 높지만, 높은 성취 수준을 감안할 때에도 상대적으로 정답률이 더 높은 문항이 있다. 또 한국과 일본의 정답률이 호주와 미국보다 높기는 하지만, 높은 성취 수준에 비취볼 때 상대적으로 정답률이 낮은 문항이 있다. 문항반응이론에 의거하여 난이도를 계산함으로써 이러한 문항을 선별할 수 있다.<sup>3)</sup>

문항반응이론에서는 각 문항을 분석하기 위해 문항특성곡선(item characteristic curve)을 구한다. 문항특성곡선의 그래프에서  $x$  축은  $-a$ 부터  $a$ 까지로 학생의 능력을 나타내는데 능력값이 클수록 우수한 학생이 되며 평균적인 학생

의 능력값은 0이 된다. 문항특성곡선의 그래프에서  $y$  축은 학생이 해당 문항을 맞출 확률을 나타내며, 일반적인 문항에서의 문항특성곡선은 누운 S자 곡선이 된다. 문항반응이론에 따른 난이도는 해당 문항을 맞출 확률이 이론상 1/2이 되는 점에 해당하는 능력값을 나타낸다. 따라서 문항의 난이도 값이 클수록 어려운 문항이고, 난이도 값이 작을수록 쉬운 문항이다(김재춘 외, 2003). 예를 들어 문항 1은 네 국가의 정답률이 72.5~81.5에 분포하는 비교적 쉬운 문항으로 <표Ⅲ-2>에 의할 때 난이도 평균은 -1.294이다. 즉 이 문항을 맞출 확률이 이론상 1/2이 되는 점의 능력값은 -1.294이므로 기준이 되는 능력값이 비교적 낮다. 이에 반해 문항 2의 정답률은 29.0~56.5에 분포하므로 비교적 어려운 문항이며, 난이도 평균은 0.511이다. 문항 2를 맞출 확률이 1/2이 되는 점의 능력값이 0.511이므로 문항 1에 비해 기준점의 능력값이 훨씬 높음을 알 수 있다.

문항반응이론에 의한 <표Ⅲ-2>의 난이도는 그 국가의 전반적인 성취 수준에 비추어 볼 때 각 문항의 상대적인 난이도를 보여준다. 예를 들어 문항 4의 정답률은 일본이 85.6%, 한국이 80.0%, 호주가 63.4%, 미국이 59.8%이다. 즉 일본과 한국은 전반적인 성취 수준이 높은 것이 상으로 호주와 미국보다 이 문항에서 높은 정답률을 보였다. 따라서 일본의 난이도는 -1.700, 한국의 난이도는 -1.514로 네 국가의 난이도 평균인 -1.233보다 낮으며, 미국은 -1.173, 호주는 -0.543으로 난이도 평균보다 높은 값을 보인다.

3) 최근 교육평가 분야에서 활발하게 연구되고 있는 문항반응이론(IRT, Item Response Theory)은 고전검사이론의 한계를 극복하기 위해 제안된 이론으로, 문항의 특성과 피험자 특성 사이의 함수관계를 이용하여 피험자가 문항에 정답을 할 확률을 수학적으로 나타내는 측정 모형의 하나이다(백순근, 채선희, 1998). 일반적으로 볼 때 동일한 문항이라도 능력이 높은 집단에게는 난이도가 낮은 문항이 되고 능력이 낮은 집단에게는 난이도가 높은 문항이 되기 때문에 문항의 난이도는 검사를 치르는 집단에 따라 달라진다. 그런데 본 연구에서 계산한 난이도는 문항반응이론에 기초하여 피험자의 능력 수준을 고려한 상대적인 수치이므로 피험자의 능력 수준에 비취볼 때 어려운 문항인지 쉬운 문항인지에 대한 정보를 제공한다.

즉 문항 4는 일본과 한국 학생들에게 상대적으로 쉽고, 호주와 일본의 학생들에게는 상대적으로 어려운 문항이고 할 수 있다.

이를 보다 정교하게 분석하기 위해 각 문항에 대한 난이도 평균에서 각 국가의 난이도를 뺀 값을 계산하여 [그림Ⅲ-1]에 제시하였다. 수평축을 기준으로 아래에 위치하는 국가는 (난이도 평균)-(각 국가의 난이도) 값이 마이너스이므로 해당 문항이 상대적으로 쉬움을 의미하고, 반대로 축 위에 위치하는 국가는 해당문항이 상대적으로 어려움을 의미한다.

여기서 주목할만한 현상은 한국과 일본, 미국과 호주가 동일한 경향을 보인 경우가 문항 1, 2, 4, 7, 8, 10, 12, 15, 16, 18, 19, 23, 24, 27, 28, 29로 문항의 개수가 14개나 된다는 사실이다. 그런데 이러한 결과가 무작위로 나타나기란 쉽지 않다. 한국, 일본, 미국, 호주의 4개 국가가 두 국가씩 짝을 짓는 경우는 (한국-일본, 미국-호주), (한국-미국, 일본-호주), (한국-호주, 일본-미국)의 모두 세 가지이다. 각 경우가 나타날 확률이 1/3라고 하면 30개의 문항 중 16개의 문항에서 동일하게 짝 지워질 확률은

<표Ⅲ-2> PISA 수학 문항의 문항반응이론에 의한 난이도

문항 코드	문항 번호	일본	한국	호주	미국	평균
M033Q01	1	-1.197	-0.925	-1.456	-1.598	-1.294
M034Q01T	2	0.315	0.281	0.595	0.852	0.511
M037Q01T	3	-1.399	-0.839	-0.867	-0.446	-0.888
M037Q02T	4	-1.700	-1.514	-0.543	-1.173	-1.233
M124Q01	5	0.602	0.804	1.013	0.802	0.805
M124Q03T	6	0.953	2.334	1.475	1.482	1.561
M136Q01T	7	-1.208	-0.909	-0.450	-0.669	-0.809
M136Q02T	8	0.536	-0.088	1.558	1.086	0.773
M136Q03T	9	2.015	1.201	1.668	1.355	1.560
M144Q01T	10	-1.707	-1.332	-1.137	-0.854	-1.258
M144Q02T	11	0.832	1.021	1.024	1.914	1.198
M144Q03	12	-1.865	-1.389	-2.267	-2.046	-1.892
M144Q04T	13	0.417	0.272	0.332	0.270	0.323
M145Q01T	14	-0.301	-0.244	-0.365	-0.271	-0.295
M148Q02T	15	1.915	2.256	1.441	1.297	1.727
M150Q01	16	-0.694	-0.936	-0.449	-0.321	-0.600
M150Q02T	17	-0.384	-1.016	-0.878	-0.896	-0.794
M150Q03T	18	0.980	0.617	-0.349	-0.620	0.157
M155Q02T	19	0.114	-0.423	-0.857	-1.015	-0.545
M155Q03T	20	1.829	1.616	1.938	1.528	1.728
M155Q04T	21	0.121	-0.172	-0.106	-0.424	-0.145
M159Q01	22	-0.919	-0.874	-1.017	-0.774	-0.896
M159Q02	23	-1.786	-2.120	-2.247	-2.205	-2.090
M159Q03	24	-1.479	-1.641	-2.098	-2.068	-1.822
M159Q05	25	0.617	1.455	1.189	1.456	1.179
M161Q01	26	-0.189	-0.100	-0.011	-0.002	-0.076
M179Q01T	27	2.167	1.520	1.086	1.020	1.448
M192Q01T	28	0.257	0.284	0.509	1.016	0.517
M266Q01T	29	1.237	1.279	1.869	2.160	1.636
M273Q01T	30	-0.081	0.179	0.220	-0.015	0.076

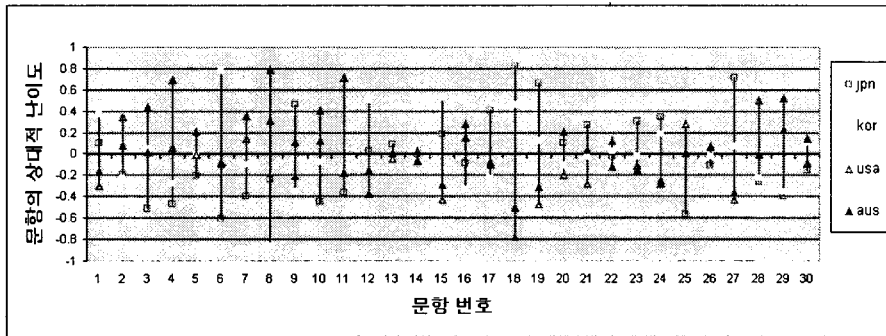
${}_{30}C_{16} \left(\frac{1}{3}\right)^{16} \left(\frac{2}{3}\right)^{14} = 0.019$  이므로 2% 보다도 작다. 이처럼 30개의 문항 중 16개의 문항에서 한국과 일본, 미국과 호주의 동일한 짝이 지워질 가능성은 극히 희박하기 때문에, 한국과 일본, 미국과 호주의 성취 결과 사이에는 모종의 동질성이 있다고 볼 수 있다.

## 2. 한국과 일본 학생들에게 상대적으로 쉬운 문항

한국과 일본의 성취 수준이 미국이나 호주의

성취 수준보다 상대적으로 높은 문항은 [그림 III-1]에서 한국을 나타내는 ■와 일본을 나타내는 □가 모두 수평축 아래에 위치한 8개 문항으로, 이 문항들에 대한 정보를 정리하면 다음과 같다. PISA 1주기 검사에서 수학 문항의 개수가 많지 않으므로 한국과 일본의 정답률이 상대적으로 높은 위의 문항들의 공통적인 특징이 명료하게 드러나지는 않으나 다음과 같은 특징을 추출할 수 있다.

첫째, 문항의 유형의 측면에서 볼 때 열린 구성반응형 문항이 포함되어 있지 않다. PISA



[그림III-1] 문항별 상대적 난이도 (난이도 평균 - 각 국가의 난이도)

### <표III-3> 한국과 일본 학생들에게 상대적으로 쉬운 문항

문항 번호	문항 제목	문항 유형	영역통합적 사고	수학적 능력의 수준	수학 내용 요소	형식화된 수학적 지식의 요구 정도
2	비공개 <sup>4)</sup>	제한적 구성반응형	공간과 도형	2	기하	아니오
4	농장(2)	제한적 구성반응형	공간과 도형	2	측정	예
7	사과나무(1)	제한적 구성반응형	변화와 증가	2	대수	약간
8	사과나무(2)	제한적 구성반응형	변화와 증가	2	대수	예
10	비공개	제한적 구성반응형	공간과 도형	1	기하	아니오
16	비공개	제한적 구성반응형	변화와 증가	1	수	약간
28	비공개	선택형	변화와 증가	2	측정	예
29	비공개	선택형	공간과 도형	2	측정	예

4) PISA는 3년을 주기로 반복되는 검사이며, 상당 수의 문항은 2000년 PISA와 2003년 PISA에 공통적으로 포함되는 연계 문항이므로 비공개를 원칙으로 한다. 따라서 공개 문항인 '사과나무'와 '농장' 이외에는 문항의 제목을 제시하지 않았다.



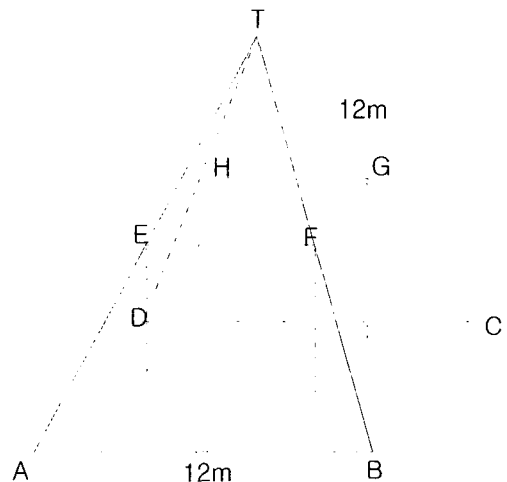
의 문항은 선택형, 제한적 구성반응형, 열린 구성반응형의 세 가지로 구분된다. 선택형은 진위형과 선다형을 포함하고, 제한적 구성반응형(closed constructed response)은 모범 답안이 하나 혹은 몇 가지로 정해져 있는 서술형 문항과 단답형을 포함한다. 열린 구성반응형(open constructed response)은 서술형 문항이지만 제한적 구성반응형과 달리 피험자의 다양한 반응이 예상되는 문항을 말한다. PISA에서 한국과 일본이 상대적 우위를 보인 문항 중 열린 구성반응형에 해당되는 문항이 없음은 한국과 일본의 학생들이 일반적으로 다양한 반응을 요구하는 서술형 문항에 취약하다는 통념에 부합되는 결과이다.

둘째, 수학 내용 요소의 측면에서 볼 때 '함수'와 '통계'에 대한 문항이 포함되어 있지 않다. TIMSS의 교육과정 국제 비교(Schmidt et al, 1997) 등의 연구에서 일관되게 제시하는 결과에 따르면 한국과 일본의 수학 교육과정과 교과서는 '통계' 영역을 비교적 소홀하게 취급하는 경향이 있다. 이는 PISA에서 한국과 일본 학생들에게 쉽게 체감된 문항 중 '통계' 영역에 해당되는 문항이 없다는 점과 인과 관계를 이룬다. 마찬가지로 다양한 변화 현상을 탐구하는 '함수'도 서구의 교과서에서 보다 중점적으로 다양하게 다루어지는 경향이 있다. 이에 반해 우리나라와 일본은 함수 개념의 도입에서부터 상당한 수학적 측면을 강조하면서 엄밀하게 다루기 때문에, 자연스러운 상황 속에서의 변화 현상을 묻는 함수 문항에서 취약함을 보였다고 할 수 있다.

셋째, 한국과 일본 학생들이 상대적 우위를

보인 문항 중에는 형식화된 수학적 지식을 요구하는 경우가 많다. 물론 어떤 문항이 형식화된 수학적 지식을 요구하는지의 여부는 판단하기 어려운 미묘한 문제이지만, 수학 교육과정이나 교과서에서 공식화된 교육 내용으로 다루는지를 기준으로 대략적인 구분이 가능하다.

예를 들어 문항 4의 '농장(2)'는 다음 그림과 같은 정사각뿔 모양의 농장에서 선분 AB의 길이가 주어지고, 점 E와 점 F가 각각 선분 AT와 선분 BT의 중점일 때, 선분 EF의 길이를 구하는 문항이다. 한국과 일본의 학생들은 선분 EF의 길이가 선분 AB의 길이의 1/2이라는 것을 일종의 공식화된 수학 내용으로 중학교 때 학습하기 때문에 짧은 시간에 정확하게 답할 가능성이 높다. 그렇다고 보면 공식화된 지식을 요구하는 문항에서 한국과 일본 학생들이 강세를 보인 것은 당연한 귀결이라고 할 수 있다.



5) PISA는 실생활에서 접하게 되는 다양한 상황을 배경으로 문항을 제작하기 때문에 분절화된 수학 내용 요소들을 통합하여 다루는 경향이 강하며, 그 결과 '영역통합적 사고'라는 평가틀을 설정하였다. '변화와 증가'라는 영역통합적 사고는 '수', '대수', '함수', '통계'의 수학 내용 요소를 포함하며, '공간과 도형'이라는 영역통합적 사고는 '기하'와 '측정'의 수학 내용 요소를 포괄한다.

### 3. 미국과 호주 학생들에게 상대적으로 쉬운 문항

미국과 호주의 학생들이 한국과 일본의 학생들보다 쉽다고 인식한 문항은 [그림Ⅲ-1]에서 미국을 나타내는  $\triangle$ 와 호주를 나타내는  $\blacktriangle$ 가 모두 수평축 아래에 위치한 8개의 문항으로, 이 문항들의 특징을 정리하면 다음과 같다.

첫째, 문항의 유형으로 볼 때 한국과 일본 학생들에게 상대적으로 쉬운 문항과 달리 미국과 호주의 학생들이 높은 성취 수준을 보인 문항에는 열린 구성반응형이 2문항 포함되어 있다. PISA의 문항 중 열린 구성반응형이 모두 3 문항밖에 포함되어 있다는 사실을 고려할 때, 미국과 호주의 학생들이 열린 구성반응형 문항에서 강세를 보인다고 할 수 있다.

둘째, 내용적인 측면에서 볼 때 영역통합적 사고에서는 ‘변화와 증가’, ‘공간과 도형’의 두 가지가 모두 포함되나, 수학 내용 요소 중 ‘대수’에 대한 문항은 포함되어 있지 않다. TIMSS의 교육과정 국제 비교 연구(Schmidt et al, 1997)

에 따를 때 한국과 일본의 교육과정은 대수 영역에 높은 비중을 두는 반면, 미국과 호주의 교육과정은 공식화, 알고리즘화 된 수학적 지식을 이용하는 경향이 강한 대수 내용을 강조하여 중점적으로 다루지 않는다. 이런 점에 비추어 미국이나 호주의 학생들에게 유리하게 작용한 문항 중 대수 관련 문항이 없다는 것을 설명할 수 있다.

셋째, 미국과 호주의 학생들에게 쉬운 것으로 나타난 문항은 대부분 형식화된 수학적 지식을 요구하지 않는다. 예를 들어 문항 15 ‘대륙의 넓이’는 한국의 학생들에게 특별히 어려운 문항으로 비정형적인 모양을 가진 남극 대륙을 제시하고 그 넓이를 근사적으로 구하는 것을 요구한다. 이 문항에 대한 성취 수준이 가장 높은 국가는 정답률이 26.7%인 호주이고 그 뒤를 이어 일본과 미국의 정답률은 각각 23.3%와 21.1%이다. 한국의 정답률은 15.2%로 PISA 검사의 전체 수학 문항 중 두 번째로 정답률이 낮다. 이 문항은 <부록 1>에 제시한 바와 같이 교육과정이나 교과서에서 전형적으

<표Ⅲ-4> 미국과 호주 학생들에게 상대적으로 쉬운 문항

문항 번호	문항 제목	문항 유형	영역통합적 사고	수학적 능력의 수준	수학 내용 요소	형식화된 수학적 지식의 요구 정도
1	비공개	선택형	공간과 도형	2	기하	아니오
12	비공개	선택형	공간과 도형	2	기하	아니오
15	대륙의 넓이	제한적 구성반응형	공간과 도형	2	측정	약간
18	비공개	열린 구성반응형	변화와 증가	2	통계	약간
19	비공개	제한적 구성반응형	공간과 도형	2	통계	약간
23	경주용 차	선택형	변화와 증가	1	함수	약간
24	경주용 차	선택형	변화와 증가	2	함수	약간
27	비공개	열린 구성반응형	변화와 증가	2	함수	약간

6) 호주는 남반구에 위치하는 국가이기 때문에 사회나 지리 교과에서 남극 대륙을 접할 기회가 많은 편이다. 남극 대륙의 지도에 대한 친숙함이 긍정적인 작용을 한 문화적 편파성의 소지가 있는 문항이다.

로 다루는 내용이 아닐 뿐 아니라 축척 정보를 제시한 방식이 우리나라 학생들에게 익숙하지 않다는 점도 부정적인 요인으로 작용했을 것으로 판단된다. <부록 2>에 제시한 문항 23과 24의 '경주용 차'는 그래프를 해석하는 능력을 요구하는 문항으로, 마찬가지로 교육과정과 교과서에서 본격적으로 다루는 내용은 아니다. 이처럼 미국과 호주의 학생들이 강세를 보인 문항은 대체적으로 형식화된 지식을 요구하지 않거나 미약한 정도로만 요구하는 문항이다.

#### IV. 논 의

본 연구에서 실시한 TIMSS-R과 PISA의 결과 분석을 통해 한국과 일본, 미국과 호주의 수학 교육 사이에 모종의 동질성이 있음을 추론할 수 있다. 수학을 가르치고 배우는 수학 수업은 그 구성원들의 가치관과 사고가 담긴 일종의 문화라고 할 때, 우리나라와 지리적으로 가깝고 교육 제도에서도 많은 것을 공유하는 일본이 수학교육의 제 현상에 있어 우리와 높은 유사성을 보이는 것은 일면 자연스럽다. PISA에서 한국과 일본의 학생들이 강세와 약세를 보이는 문항들이 상당 부분 일치하는 현상은 수학교육에 있어 한국과 일본의 인접성으로 부분적인 설명이 가능할 것이다. 물론 세밀한 비교를 한다면 한국과 일본 사이에 차이가 존재하지만, 미국이나 호주에 비하면 한국과 일본의 동질성이 더 강하다고 볼 수 있다.

국제 비교 연구의 의의는 그 결과로부터 우리 교육의 강점과 약점을 국제적인 수준에서 객관적으로 파악하여 강점을 지속시키고 취약점을 보완하는 계기로 삼는 데 있다. PISA의 성취 수준이 TIMSS-R의 성취 수준보다 상대적으로 낮은 것과 관련하여 수학과 실생활과의

연관성을 높이고 영역통합적인 접근을 적극적으로 시도하는 PISA의 경향을 우리나라의 수학교육에 적극적으로 반영할 필요가 있다는 시사점을 얻을 수 있다. 물론 여기서 고려해야 할 것은 수학교육에 완벽한 사조와 경향이란 존재할 수 없기 때문에 어떤 것을 따르더라도 장점과 더불어 취약점을 내포하기 마련이라는 점이다. 예컨대 상황과의 관련성을 강조하는 서구의 교과서는 학생들로 하여금 수학의 유용성을 인식하게 하는 데에는 도움이 되지만, 경우에 따라서는 수학과 실제 상황과의 연결 고리가 확실하지 않아 수학적 의미를 명료하게 전달할 수 없는 경우가 있다. 이에 반하여 수학적 내용 설명을 강조하고 실생활 맥락이 상대적으로 덜 반영되어 있는 한국과 일본의 교과서는 수학 내용을 경제적으로 전달할 수는 있지만, 중하위권 학생이 수학에 대한 호감을 갖도록 하여 수학 학습의 장(場)으로 유도하는 것이 쉽지 않다(박경미, 임재훈, 2002). 따라서 각 경향이 갖는 특징을 종합적으로 고려하여 최적의 방향으로 수학 교육과정과 교과서를 구성할 필요가 있을 것이다.

수학 내용 요소의 구성 비율과 관련해서도 두 가지 관점이 성립할 수 있다. 한국과 일본은 대수 영역을 상대적으로 강조하고 통계 영역이 상대적으로 빈약하며, 함수의 경우 실제적인 변화 현상을 다양하게 다루기보다는 수학적 엄밀성을 추구하는 방향으로 내용을 전개하는 경향이 있다. 그렇다면 국제적인 표준에 맞추어 대수 영역을 약화하고 통계 영역을 강화할 것인가? 이에 대하여 단정적으로 결론을 내리기는 쉽지 않다. 내용 요소의 구성 비율은 한 국가가 고유하게 결정해야 할 문제이지 전 세계적인 표준을 정하고 이를 일률적으로 따르도록 강요할 성질의 것이 아니기 때문이다. 그렇지만 고유한 전통을 보존한다는 수구적인 입

장에 서서 국제적인 조류를 받아들이지 않는 것도 문제는 있으므로 진취적인 입장에서 세계적인 경향이 어떻게 흘러가고 있는지 반영할 필요도 있어 보인다. 결과적으로 수학 교육과정의 내용 구성에 있어 세계적인 표준을 따르는 것과 국가별 고유성을 고수하는 것 사이의 절묘한 조화가 필요하다.

1999년에 실시된 TIMSS-R에 뒤이은 TIMSS-2003, 그리고 2000년에 실시된 PISA 1주기 검사에 이어 2003년의 2주기 PISA의 결과가 앞으로 발표될 것이다. 본 연구에서 TIMSS-R과 PISA 1주기 수학 성취 수준을 연계하여 분석함으로써 우리나라 학생들의 수학 수준을 보다 입체적으로 파악하고자 시도했다.

이처럼 TIMSS-R과 TIMSS-2003, PISA 1주기와 2주기 성취 수준을 종단적으로 비교하고, TIMSS와 PISA의 성취 수준을 연계하여 분석하는 시도들이 계속되어 국제 비교 연구로부터 많은 시사점을 얻고 우리나라의 수학교육의 수준을 향상시키는 계기가 되어야 할 것이다.

## 참고문헌

김재춘, 부재율, 소경희, 채선희(2003). *교육과정과 교육평가*. 서울: 교육과학사.

노국향, 최승현, 박경미(2001). *PISA 2000 수학 평가 결과 분석 연구*. 한국교육과정평가원 연구보고 RRE 2001-9-3.

박경미(2000). 중학교 중학교 수학 교육과정과 교과서 내용의 양과 난이도에 대한 -- 考察. *수학교육학연구*, 10(1), 35-55.

박경미, 임재훈(2002). 한국, 일본과 미국, 영국의 수학 교과서 비교, *학교수학*, 4(2), 317-331.

박정, 홍미영, 나귀수(2001). *TIMSS-R 국제기준에 따른 우리나라 중학생의 수학·과학 성취도 분석*. 한국교육과정평가원 연구보고 RRE 2001-10.

백순근, 채선희(1998). *컴퓨터를 이용한 개별 적응검사*. 서울: 원미사.

Beaton, A. E. et al(1996). *Mathematics achievement in the middle school years*. Boston: Center for the Study of Testing, Evaluation and Educational Policy, Boston College.

Schmidt W. H. et al(1997). *Many visions, many aims (volume 1)*. Dordrecht: Kluwer Academic Publishers.

Organisation for Economic Co-operation and Development(2000). *Measuring student knowledge and skills*. Paris: OECD Publications.

Organisation for Economic Co-operation and Development(2001). *Knowledge and skills for life first results from PISA 2000*. Paris. OECD Publications.

Mullis, I. V. S., Martin, O. M., Gonzalez, E. J., Gregory, K., D., Garden, R. A., O'Connor, K. M., et al(2000). *TIMSS 1999 international mathematics report*. International Study Centre, Boston College, Chestnut Hill.

# An In-depth Analysis of the Result of the International Comparative Study of Mathematics

Park, Kyung Mee (Hongik University)

The recent international comparative studies such as PISA(Program for International Student Assessment) and TIMSS-R(Third International Mathematics and Science Study-Repeat) provide results of relative mathematics achievement of participating countries. The purpose of this paper is to compare the mathematics results of PISA and TIMSS-R. To make PISA and TIMSS-R results comparable, they were standardized. The close investigation of these standardized results reveals that the two Asian countries(Korea and Japan) and several English speaking countries have the commo-

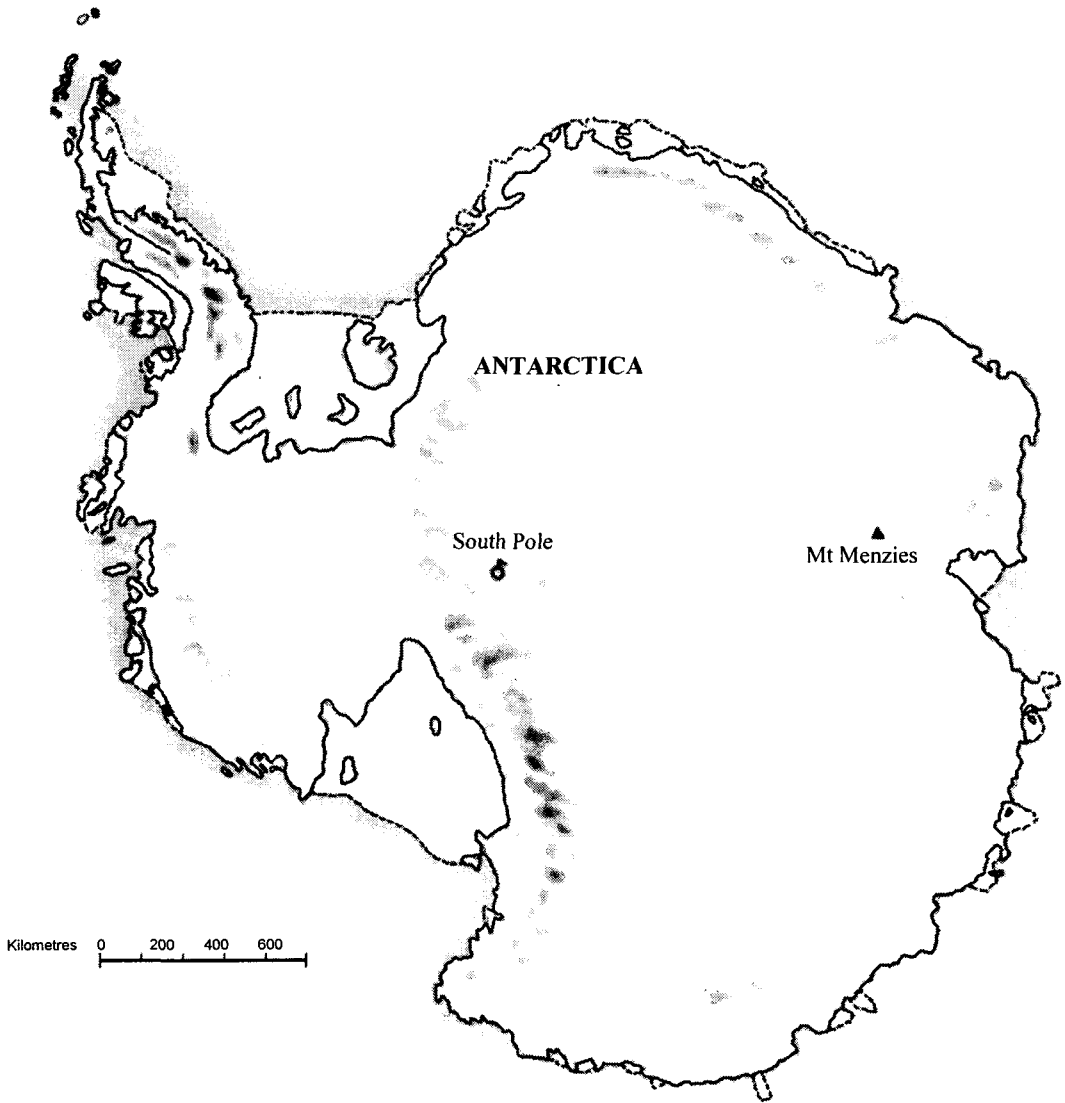
ality in mathematics achievement. Thus this study looks for patterns and similarities within a group of Asian countries(Korea and Japan) and Western countries(the U.S and Australia) by in-depth analysis of PISA mathematics achievement based on item response theory. As a result, it was noted that Western countries tend to perform well on open constructed items and are likely to perform better when an item involves less formal mathematics. On the other hand, Asian countries perform well when an item involves numeric or algebraic computation related to curriculum-based content, but they are relative poor at an item calls for verbal explanations or interpretations of graphs.

\* **Key words:** international comparative study(국제비교연구), TIMSS-R(Third International Mathematics and Science Study-Repeat), PISA(Program for International Student Assessment), item response theory(문항반응이론)

논문접수 : 2004. 8. 23

심사완료 : 2004. 10. 20

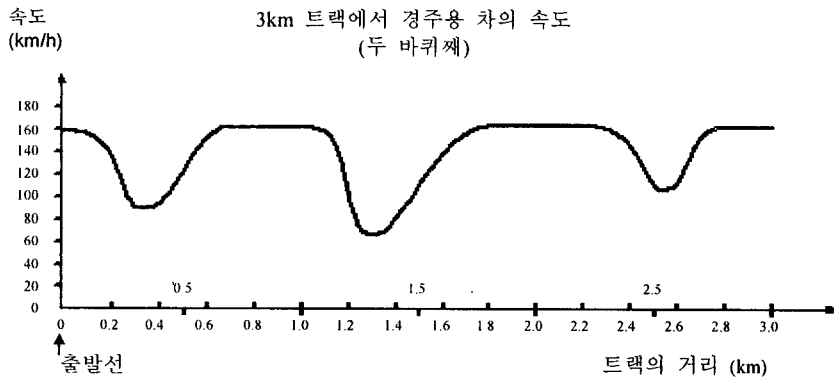
<부록 1> 문항 15 '대륙의 넓이'



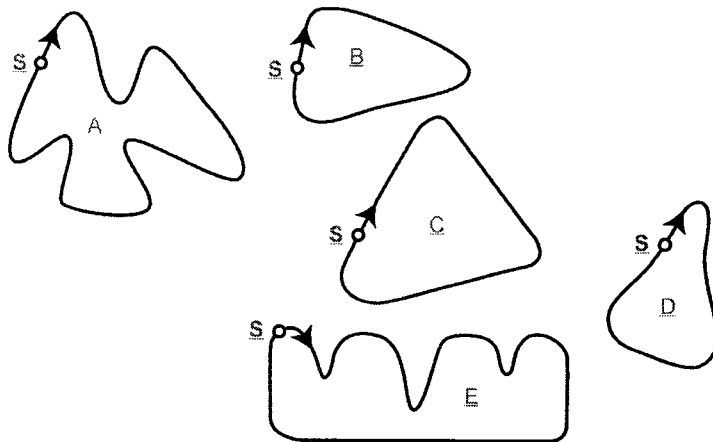
지도의 축척을 이용하여 남극 대륙의 넓이를 어림하여라.  
어떻게 어림하였는지 그 과정을 설명하여라.

<부록 2> 문항 23, 24 ‘경주용 차’

다음 그래프는 경주용 차가 3km의 평탄한 트랙에서 두 바퀴째를 돌 때의 속도 변화를 나타낸 것이다.



다음의 다섯 가지 트랙 중 앞에 제시한 속도 그래프를 만족시키는 트랙은 어느 것인가?



**S:** 출발점