大韓數學敎育學會誌 數學敎育學硏究 第 12 卷 第 2 號
The Journal of Educational Research in Mathematics
Vol. XII, No.2, 291-311, May 2002

학업성취도 국제 비교 평가(PISA)에 나타난 우리나라 학생의 수학적 소양 수준 분석

박경미* 최승현** 노국향***

I. 서론

최근 들어 국경을 초월한 무한 경쟁 시대에 진입하면서 국가 내 학생의 성취 수준에 대한 정보만으로는 자국 교육 체제의 장점과 단점을 종합적으로 파악하기 어렵다는 인식이 높아지고 있다. 그 결과 세계 각국은 자국 학생의 성취 수준을 국제적인 맥락에서 파악하고자 하는 노력을 기울이고 있다. 특히 지식 기반 사회의 도래와 더불어 물리적 자원보다는 부가 가치 생산 능력과 같은 고등 사고 능력이 국가의 경쟁력을 결정하는 핵심적인 요소로 부각되면서, 교육 성과를 점검하는 평가가 분절화된 지식보다는 창의력, 문제해결 능력 등을 위주로 해야 한다는 요구가 높아지고 있다. 이러한 필요성에 부응하기 위한 학업 성취도 국제 비교 평가(Program for International Students Assessment, 이하 PISA로 약칭)는 경제협력개발기구인 OECD가 주관하는 연구로, 각 국가의 인적 자원의 수준을 나타내는 지표의 산출을 목적으로 한다.

본 고는 PISA 수학 평가의 특징을 알아보고 수학 평가틀을 소개한 후, PISA 결과에 나타난 우리나라 학생의 수학적 소양 수준과 그 분포를 알아볼 것이다. 더불어 수학 문항에 대한 우리나라 학생의 정답율과 반응 빈도를 분석함으로써 주목할 만한 특징을 추출하고자 한다. 또한 설문 조사 결과에 나타난 배경 변인과 수학적 소양의 관계를 종합하여, PISA 결과가 우리나라 수학 교육에 주는 시사점을 도출할 것이다.

II. PISA 수학 평가 개관

PISA는 대부분의 OECD 국가에서 의무 교육이 종료되는 시점인 만 15세 학생을 대상으로 하는 국제 비교 연구로, 1998년부터 2006년까지 3년을 주기로 세 차례에 걸쳐 총 9년 동안 진행된다. PISA의 1주기에서는 읽기가 주영역이고 수학과 과학은 보조영역이 되며, 2주기와 3주기에서는 수학과 과학이 각각 주영역이고 나머지 두 영역은 보조영역이 된다. PISA 1주기에는 OECD 가입국 28개국과 비가입국 4개국을 포함하여 총 32개국이 참여하였으며, 이 중 네덜란드는 표집의 문제로 인해 결과 분석에서 제외되었다. 우리나라는 2000년 7월 전국 146개 중·고등학교 학생 5,131명을 대상으로 PISA 1주기의 검사를 실시하였다.

PISA의 평가 도구는 읽기 137문항, 수학 32문항, 과학 35문항으로 구성되며, 검사를 효율

* 홍익대
** 한국교육과정평가원
*** 한국교육과정평가원

적으로 실시하기 위하여 전체 문항을 9종의 검사지에 분리하여 배치하였다. 학생은 9종의 검사지 중 1종을 무작위로 배정 받았으며, 검사 시간은 읽기, 수학, 과학 문항이 포함된 검사지 1종에 120분, 설문 조사에 35분, 휴식 시간 25분으로 총 3시간이 소요되었다. PISA의 서술형 문항의 채점은 단수 채점과 복수 채점의 두 과정을 거쳐 이루어졌으며, 네 명의 채점자가 독립적으로 채점하도록 한 복수 채점은 평가의 신뢰성을 높이기 위한 방안의 하나로, 전체 답안지의 10%가 복수 채점의 대상이 되었다(한국교육과정평가원, 2001).

1. PISA 수학 평가의 특징

우리나라는 지금까지 TIMSS[1], TIMSS-R[2], IAEP[3] 등 여러 차례에 걸쳐 국제 비교 연구에 참여해 왔는데, PISA는 기존의 연구와 몇 가지 측면에서 차별화된다(OECD, 2000).

첫째, 이전의 국제 비교 연구는 각 국가가 학교 교육을 통해 제공하는 지식을 토대로 하는 교육과정 지향적인 평가라고 한다면, PISA는 교육과정과 긴밀하게 연계되지 않는 내용을 대상으로 하는 평가이다. PISA의 목적은 학생이 미래의 건전한 민주 시민으로 생활을 영위하는 데 필요한 읽기 소양, 수학적 소양, 과학적 소양을 측정하고자 하는 것이다. 여기서의 '소양(literacy)'이란 각 교과의 단편적인 지식을 수용하여 단순하게 재생산하는 능력이 아니라, 핵심 개념에 대한 풍부한 이해를 바탕으로 다양한 상황에 적절하게 활용할 수 있는 능력을 말한다. 이와 같이 PISA는 여러 분야에 대한 소양의 측정을 목적으로 하기 때문에 교과 내용과 긴밀하게 연계되는 교과서적인 문항보다는, 일상 생활에서 쉽게 접할 수 있는 다양한 소재와 배경을 동원하는 실생활 맥락의 문항이 주류를 이룬다.

둘째, 기존의 국제 비교 연구는 객관성과 신뢰성을 우선시 하기 때문에 선택형 문항을 위주로 하였지만, PISA는 학습자의 구성적인 반응을 요구하는 서술형 문항을 적극적으로 도입하였다. 이는 평가 결과의 객관성을 관건으로 하는 국제 비교 연구라는 점에 비추어 볼 때 이례적이라고 할 수 있다.

셋째, PISA는 학년 단위가 아닌 연령(15세) 단위의 표집이자 학교 내의 학급을 선택하지 않고 개인을 선택하는 방식이라는 점에서 기존 연구와 차이를 갖는다. PISA의 표집은 두 단계로 이루어지는데, 우선 학교에 대한 표집으로 12개의 중학교와 134개의 고등학교를 선정하고, 이어 표집된 학교 내의 학생을 무작위 추출 방법에 의해 선정한다. 이러한 절차를 거쳐 선정된 5131명의 학생 중 고등학교 1학년 학생이 98%, 중학교 3학년과 고등학교 2학년 학생이 각각 1%씩 포함되어 있다. 실제 PISA는 우리나라의 고등학생이 참여한 국제 비교 연구라는 점에서도 주목할 만하다. 이전의 국제 비교 연구는 대개 초등학교나 중학교 학생을 중심으로 이루어졌으며, TIMSS의 평가 대상에는 초등학교, 중학교, 고등학교가 모두 포함되었지만 우리나라는 초등학교와 중학교만 참여하였다. 따라서 PISA는 우리나라 고등학교 학생의 수학적 수준을 국제적인 수준에서 점검한다는 점에서 의미를 지닌다.

넷째, 기존의 국제비교 연구가 주목할 만한 이론적 배경을 제시하지 않은 점에 비추어 볼

---

1) 제 3차 수학 · 과학 성취도 국제 비교 연구 (Third International Mathematics and Science Study)
2) 제 3차 수학 · 과학 성취도 국제 비교 반복 연구 (TIMSS-Repeat)
3) 국제 교육 발전 평가 (International Assessment of Educational Progress)

때, PISA는 비교적 이론과 실제가 잘 조화된 평가라고 할 수 있다. 수학 평가틀에서 중심적인 역할을 하는 '영역통합적 사고'는 수학 문제 상황의 발생이라는 점에 주목하여 설정된 것으로, 프로이덴탈(Freudenthal)의 교수학적 현상학과 관련지을 수 있다. 교수학적 현상학에 따르면 수학은 실재를 구조화하고 정리하는 하나의 인지적 도구로(Freudenthal, 1983; 정영옥, 1997), PISA의 수학 문항에 포함된 맥락은 수학을 통해 조직될 수 있는 현상을 제공한다. 이러한 점에 비추어 볼 때 PISA의 수학 평가틀은 심도 있는 이론적 배경을 토대로 하고 있으며, 이론과 실제의 접목이라는 측면에서 PISA 수학 평가의 의의를 찾아볼 수 있다.

2. PISA 수학 평가틀

PISA 수학 평가틀은 '내용', '과정', '상황'의 삼원적 요소로 구성된다.

(1) 내용 요소

내용 요소는 '수학 교육과정적 요소(mathematical curricular strand)'와 '영역통합적 사고(overarching idea)'의 두 가지 관점에서 설명될 수 있다. 수학 교육과정적 요소와 영역통합적 사고는 모두 수학 내용을 분류하는 틀로, 전자는 수학의 내용 체계를 엄격하게 따르는 반면 후자는 수학으로 해결할 수 있는 문제 상황의 생성이라는 측면에 주목한다. 기존의 분류와 유사한 수학 교육과정적 요소는 '수,' '대수', '측정', '기하', '어림', '함수', '확률', '통계'와 같이 학교 교육과정에서 전형적으로 다루어지는 내용 영역을 중심으로 구성된다.

PISA는 기존의 평가 연구와 달리 실생활을 배경으로 문제 상황을 구성하는 데 우선권을 두며, 문제 상황은 특정 내용 영역에 국한하여 발생하는 것이 아니기 때문에 영역간의 통합을 도모할 필요가 있다. 이러한 특징에 따라 PISA의 평가틀에서 설정한 영역통합적 사고는 '가능성', '변화와 증가', '공간과 도형', '계량적 추론', '불확실성', '종속과 관계'의 여섯 가지 하위 영역으로 구성된다. 그런데 PISA 1주기에서 수학은 보조영역이기 때문에 문항 수의 제약으로 인해 위에 열거한 6개의 영역통합적 사고 중 비교적 다양한 내용을 포괄할 수 있는 '변화와 증가'와 '공간과 도형'의 두 가지만 취급한다(박경미, 1999).

우리나라의 제7차 수학과 교육과정은 내용 영역으로 '수와 연산', '도형', '측정', '문자와 식', '규칙성과 함수', '확률과 통계'의 여섯 가지를 설정하고 있다. PISA 평가틀의 수학 교육과정적 요소는 우리나라 제7차 교육과정의 영

<표 1> 수학 교육과정의 내용 영역과 PISA 평가틀의 내용 요소의 관계

제7차 교육과정의 내용 영역	수학 교육과정적 요소	영역통합적 사고
수와 연산	수	변화와 증가
문자와 식	대수	
규칙성과 함수	함수	
확률과 통계	확률	
	통계	
도형	기하	공간과 도형
	어림	
측정	측정	

역 구분과 거의 유사하다. 한편 영역통합적 사고의 '변화와 증가'에 대응되는 교육과정의 내용 영역은 '수와 연산', '문자와 식', '규칙성과 함수'이며, '공간과 도형'과 부합되는 교육과정의 내용 영역은 '도형'과 '측정'이라고 할 수 있다. 또 '확률과 통계'는 특정한 하나의 영역 통합적 사고와 직접적으로 연결짓기는 어렵지만 확률이나 통계와 관련된 상황은 '변화와 증가'에 해당하는 문항을 전개하는 소재를 제공할 가능성이 높다.

(2) 과정 요소

과정 요소를 이루는 수학적 능력은 '수학적 사고 기능', '수학적 논증 기능', '모델링 기능', '문제 제기와 해결 기능', '표상 기능', '기호화와 형식화 기능', '의사소통 기능', '보조 교구와 도구의 사용 기능'의 8가지로 구성된다. 이 요소 중 첫 번째의 수학적 사고 기능은 다른 것보다 포괄하는 범위가 넓으며, 마지막의 보조 교구와 도구의 사용 기능은 기존의 관점에서는 대개 수학적 능력의 범위에 포함시키지 않았기 때문에 다소 이질적이라고 할 수 있다.

한편 특정한 수학 문항을 해결하기 위하여 요구되는 수학적 능력의 개수는 단수가 아니라 복수라고 할 수 있다. 다시 말해 각각의 문항은 하나의 수학적 능력과 일대일로 대응되기보다는 수학적 능력 중 여러 요소와 동시적으로 대응되는 경우가 대부분이다. 또한 문항을 해결하는 데 필요한 수학적 능력은 다음 준거에 의하여 세 수준으로 구분될 수 있다(박경미, 1999).

· 1수준 : 재생, 정의, 계산
· 2수준 : 문제해결을 위한 연결과 통합
· 3수준 : 수학적 사고, 일반화와 통찰

위와 같이 수준을 조작적으로 정의할 수 있지만, 수준을 구분하는 또 다른 중요한 기준은 프로이덴탈의 교수학적 현상학을 이끄는 핵심 개념인 '수학화(mathematization)'의 정도이다. 즉 1수준은 수학화 과정이 결여되어 있거나 미약한 경우를, 2수준은 표준적인 수학화 과정이 일어나는 경우를, 3수준은 복잡한 수학화 과정을 수반하는 경우를 말한다. 그러나 이 세 수준이 엄격한 위계 관계를 이루는 것은 아니며, 따라서 난이도와 직접적인 관련성을 맺지는 않는다. 예를 들어 1수준의 문항이라도 난이도가 높을 수 있으며, 3수준이 문항이라도 난이도가 낮은 경우가 발생할 수 있다.[4)]

(3) 상황 요소

상황 요소는 문항이 배경으로 하고 있는 소재에 따라 '개인적 상황', '교육적 상황', '공적인 상황', '직업적 상황', '수학적 상황'[5)]의 다섯 가지로 구분된다. 이 때 앞의 네 가지는 학생의 주변에서 흔히 일어날 수 있는 일상적인 상황으로부터 교육을 염두에 둔 상황, 공동체와 사회, 직업 세계에서 나타날 수 있는 상황으로 점차 확대된다. 수학적 상황에 해당하는 문항은 구체적인 실생활 장면이 결합되지 않은 순수 수학적인 경향을 띤다. 따라서 개인적인 상황에서 수학적인 상황으로 갈수록 학생이 일상적으로 경험하는 것과의 거리는 점점 더 멀어진다고 할 수 있다.

4) 1, 2, 3 수준이라는 용어를 사용할 경우 세 개의 수준이 서로 엄격한 위계를 이룬다는 선입견을 주기 때문에, 그 대신 '재생군(reproduction cluster)', '연결군(connection cluster)', '반성군(reflection cluser)'이라는 용어를 각각 사용하기도 한다.

5) 수학적 상황에서 '수학적'이란 'scientific'을 번역한 것으로, 이를 '과학적'이라고 번역할 경우 물리, 화학, 생물, 지구과학과 같은 협의의 과학을 연상하는 경향이 있기 때문에 이러한 혼란을 방지하기 위하여 '수학적'으로 의역하였다.

3. PISA 수학 문항 유형

PISA의 수학 문항은 유형에 따라 선택형, 제한적 구성반응형, 열린 구성반응형으로 구분된다. 선택형(multiple choice)에는 일반적인 문항 유형 분류에서와 마찬가지로 참과 거짓을 판단하는 진위형과 여러 답지 중 하나를 선택하는 선다형이 포함된다. 제한적 구성반응형(closed constructed response)에는 단답형과 더불어 서술하는 것을 요구하는 문항이라도 반응의 방향이 고정되어 있는 경우는 이 유형으로 분류된다. 열린 구성반응형(open constructed response)에는 서술형 중 답안이 하나로 수렴되지 않는, 즉 다양한 반응이 가능한 문항이 포함된다.

<표 2> PISA의 수학 문항 유형 분류

문항 유형	하위 유형
선택형	진위형
	선다형
제한적 구성반응형	단답형
	서술형 (제한된 반응)
열린 구성반응형	서술형 (다양한 반응)

4. PISA 수학 검사 도구

수학 검사 도구는 과제(task)와 이에 수반되는 문항(item)으로 구성된다. 하나의 과제는 적게는 1개의 문항, 많게는 4개의 문항들로 이루어진다. 과제는 대개 일련의 조건이나 상황으로 시작되는데, 수학의 경우 교과의 특성상 과제에 제시된 지문이 길지 않다.

2000년의 본검사에 앞서 1999년 실시된 예비검사의 수학 검사 도구는 총 34개의 과제와 77개의 문항으로 구성되었으나 이 중 본검사의 수학 검사 도구에 포함된 것은 16개의 과제와 32개의 문항이다. 본검사 과제와 문항의 취사선택은 통계적 측면과 내용적 측면에서 이루어진다. 통계적 측면의 정량적 기준에는 난이도, 변별도, 답지 반응 분포, 모형 적합도 등이 있

<표 3> 수학 평가틀과 문항 유형에 따른 PISA 본검사 문항의 분류

분류			하위 분류	문항수	문항의 비율
삼원적 요소	내용	영역통합적 사고	변화와 증가	18	56.3%
			공간과 도형	14	43.7%
		수학 교육과정적 요소	수	1	3.1%
			대수	5	15.6%
			측정	7	21.9%
			기하	8	25.0%
			함수	5	15.6%
			통계	6	18.8%
	과정	수학적 능력의 수준	1수준	10	31.3%
			2수준	20	62.5%
			3수준	2	6.3%
	상황	상황	개인적	12	37.5%
			교육적	6	18.8%
			직업적	3	9.4%
			공적	4	12.5%
			수학적	7	21.9%
문항의 유형			선택형	11	34.4%
			제한적 구성반응형	18	56.2%
			열린 구성반응형	3	9.4%
총계				32	100%

으며, 정량적 기준은 예비검사 결과를 국제 연구 본부에서 점검할 때 적용된다. 정량적인 기준을 보완하는 정성적인 기준은 주로 내용적인 측면에 대한 것으로, 문항의 적절성(교육과정과의 연계성, 내용 친숙도, 흥미도), 편파성(문화적, 성별), 기술적인 문제(번역, 질문 의도의 명료성, 채점 기준), PISA 목적에의 부합 정도 등이 있다. 정성적인 기준을 충족시키는지의 여부는 참여국 각국의 의견 개진을 통해 이루어진다.

본검사 문항을 수학 평가틀과 문항 유형에 따라 분류하면 다음과 같다.

5. PISA 수학 문항과 채점 기준의 실제

PISA 1주기에서 사용된 문항 중 일부는 2주기에서 반복적으로 사용될 예정이기 때문에 대부분의 문항은 비공개를 원칙으로 하나, <부록>에 제시된 '사과나무' 과제는 공개가 허용된 문항이다. 이 과제는 수학 평가틀에 비추어 볼 때, '영역통합적 사고' 중 '변화와 증가'에 해당하며, '수학 교육과정적 요소'에서는 '대수'의 내용을 취급한다. '상황'은 '교육적 상황'에 해당하며, 첫 번째와 세 번째 문항의 유형은 제한적 구성반응형이며 두 번째 문항은 열린 구성반응형이다.

수학 채점 기준은 두 자리 코딩 방식을 취하는데, 이 방식은 학생들의 반응을 점수화하기 편리할 뿐 아니라 피험자의 다양한 반응 양식을 효과적으로 추적할 수 있다는 이점 때문에 보편화된 방법이다. 두 자리 코딩 방식에서 십의 자리는 정답과 오답이나 부분정답을 나타내는데, 10점대의 코드에 해당하는 반응보다는 20점대의 코드에 해당하는 반응이 수학적으로 더 정답에 가까우며, 더 높은 점수를 받을 가치를 지닌다. 이에 반해 일의 자리는 동일한 점수에 해당하는 반응을 유형화하여 코드를 부여한 것으로, 십의 자리 코드와 달리 일의 자리 코드 사이의 관계는 위계적이지 않다. 일의 자리까지 코드는 학생이 문제에 접근하는 방법이나 전략, 정답과 오답의 유형, 학생이 가지고 있는 오개념에 대한 의미있는 정보를 제공할 수 있다.

III. PISA 수학 평가 결과

1. 수학적 소양 수준

우리나라 학생의 수학적 소양의 수준은 전체 평균을 500점, 표준편차를 100으로 하였을 때 평균 547점으로, 일본에 이어 2위를 기록하였다 (OECD, 2001). 수학적 소양 점수는 미래 사회를 이끌어갈 미래의 주역의 수학적 역량을 예측하는 지표라고 할 수 있기 때문에, 우리나라가 2위를 기록했다는 사실은 일단 상당히 고무적인 현상으로 받아들일 수 있다.

<표 4>에는 결과 분석에 포함된 31개국 중 OECD 가입국 27개국에 대한 수학적 소양 점수의 다변량 분석 결과가 제시되어 있는데, 명암이 ▒로 표시된 국가는 평균이 OECD 평균보다 유의미하게 높은 경우로 일본, 한국을 포함한 14개국이 이에 해당한다. 또한 명암이 ▓로 표시된 아일랜드, 노르웨이, 체코, 미국은 평균이 OECD 평균과 유의미하게 차이가 나타나지 않는 경우이며, 명암이 □인 국가는 독일과 헝가리를 포함하는 9개국으로, 평균이 OECD 평균보다 유의미하게 낮은 경우이다.

행과 열이 만나는 위치에 표시된 ▲는 행에 해당하는 국가의 수학적 소양 점수의 평균이 열에 해당되는 국가의 평균보다 유의미하게 높은 경우이며, ▼는 그 반대의 의미이다. 또 행

<표 4> 수학적 소양 점수 평균의 다변량 비교

국가명	평균	표준오차	일본	한국	뉴질랜드	핀란드	호주	캐나다	스위스	영국	벨기에	프랑스	오스트리아	덴마크	아이슬란드	스웨덴	아일랜드	노르웨이	체코	미국	독일	헝가리	스페인	폴란드	이탈리아	포르투갈	그리스	룩셈부르크	멕시코
			557	547	537	536	533	533	529	529	520	517	515	514	514	510	503	499	498	493	490	488	476	470	457	454	447	446	387
			(5.5)	(2.8)	(3.1)	(2.1)	(3.5)	(1.4)	(4.4)	(2.5)	(3.9)	(2.7)	(2.5)	(2.4)	(2.)3	(2.5)	(2.7)	(2.8)	(2.8)	(7.6)	(2.5)	(4.0)	(3.1)	(5.5)	(2.9)	(4.1)	(5.6)	(2.0)	(3.4)
일본	557	(5.5)		○	○	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲
한국	547	(2.8)	○		○	○	○	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲
뉴질랜드	537	(3.1)	○	○		○	○	○	○	○	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲
핀란드	536	(2.1)	▼	○	○		○	○	○	○	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲
호주	533	(3.5)	▼	○	○	○		○	○	○	○	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲
캐나다	533	(1.4)	▼	▼	○	○	○		○	○	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲
스위스	529	(4.4)	▼	▼	○	○	○	○		○	○	○	○	○	○	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲
영국	529	(2.5)	▼	▼	○	○	○	○	○		○	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲
벨기에	520	(3.9)	▼	▼	▼	▼	○	▼	○	○		○	○	○	○	○	▲	▲	▲	○	▲	▲	▲	▲	▲	▲	▲	▲	▲
프랑스	517	(2.7)	▼	▼	▼	▼	▼	▼	○	▼	○		○	○	○	○	▲	▲	▲	○	▲	▲	▲	▲	▲	▲	▲	▲	▲
오스트리아	515	(2.5)	▼	▼	▼	▼	▼	▼	○	▼	○	○		○	○	○	▲	▲	▲	○	▲	▲	▲	▲	▲	▲	▲	▲	▲
덴마크	514	(2.4)	▼	▼	▼	▼	▼	▼	○	▼	○	○	○		○	○	▲	▲	▲	○	▲	▲	▲	▲	▲	▲	▲	▲	▲
아이슬란드	514	(2.3)	▼	▼	▼	▼	▼	▼	○	▼	○	○	○	○		○	▲	▲	▲	○	▲	▲	▲	▲	▲	▲	▲	▲	▲
스웨덴	510	(2.5)	▼	▼	▼	▼	▼	▼	▼	▼	○	○	○	○	○		○	○	▲	○	▲	▲	▲	▲	▲	▲	▲	▲	▲
아일랜드	503	(2.7)	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	○		○	○	○	▲	○	▲	▲	▲	▲	▲	▲	▲
노르웨이	499	(2.8)	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	○	○		○	○	○	○	▲	▲	▲	▲	▲	▲	▲
체코	498	(2.8)	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	○	○		○	○	○	▲	▲	▲	▲	▲	▲	▲
미국	493	(7.6)	▼	▼	▼	▼	▼	▼	▼	▼	○	○	○	○	○	○	○	○	○		○	○	○	▲	▲	▲	▲	▲	▲
독일	490	(2.5)	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	○	○	○		○	▲	▲	▲	▲	▲	▲	▲
헝가리	488	(4.0)	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	○	○	○	○	○		○	○	▲	▲	▲	▲	▲
스페인	476	(3.1)	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	○	▼	○		○	▲	▲	▲	▲	▲
폴란드	470	(5.5)	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	○	▼	○	○		○	○	○	▲	▲
이탈리아	457	(2.9)	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	○		○	○	▲	▲
포르투갈	454	(4.1)	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	○	○		○	○	▲
그리스	447	(5.6)	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	○	○	○		○	▲
룩셈부르크	446	(2.0)	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	○	○	○		▲
멕시코	387	(3.4)	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	

과 열이 만나는 위치에 표시된 ○는 행과 열에 해당하는 국가들 사이의 수학적 소양 점수의 평균 차이가 통계적으로 유의미하지 않음을 나타낸다. 이러한 해석에 따를 때, 최상위를 기록한 일본, 한국, 뉴질랜드 학생의 수학적 소양 점수의 평균은 각각 557점, 547점, 537점으로 평균에 있어서는 10점씩 차이가 있지만, 학생이 보유하고 있는 수학적 소양 점수 평균의 차이는 통계적으로 유의미하지 않다. 또 2행 1열부터 5열을 제외한 나머지 열에는 모두 ▲가 표시되어 있으므로, 우리나라 학생은 5개국을 제외한 나머지 22개국 학생보다 통계적으로 유의미하게 높은 수학적 소양 점수의 평균을 나타냈다.

우리나라를 비롯한 동아시아 국가의 학생은 일련의 국제 비교 연구에서 높은 성취 수준을 유지해 왔다. 그런데 이전의 국제 비교 연구에서 검사 도구로 삼았던 문항은 대개 단순 계산을 위주로 하는 것들이고(물론 문제해결이나 추론으로 분류되는 문항도 있었지만, 엄밀한 기준을 적용할 때 계산의 범주를 크게 벗어나지 않는다), 우리나라 학생은 학교 수업이나 사교육 등의 다양한 경로를 통해 이러한 문항을 해결하는 훈련을 받기 때문에 상위의 수준을 보인다고 해석되어 왔다.

이에 반해 PISA의 문항은 실생활 소재를 바탕으로 한 비정형적인 문제해결이 주류를 이루고, 한국과 일본의 학생은 이러한 문제에 접할 기회가 많지 않았다는 점에서 PISA의 수학적 소양 수준에 대한 기대감은 그리 높지 않았다. 즉 이전의 국제 비교 연구와 다른 경향을 지닌 검사 도구를 이용하는 PISA에서는 동아시아 국가의 우위가 유지되지 않을 것이라는 예측이 많이 제기되었다. 그러나 드러난 바와 같이 일본이 1위, 한국이 2위를 기록함으로써, 지금까지 기록한 국제 비교 연구의 높은 성취 수준이 낮은 인지 단계의 사고를 요구하는 문항의 특성 때문만은 아닌 것으로 드러났다. 즉 TIMSS와 같이 교육과정과 교과서에 긴밀하게 연계된 수학 내용에 대한 평가이건, 아니면 PISA와 같이 일반적인 수학적 소양을 측정하는 평가이기 때문에 사교육을 통한 훈련의 효과가 직접적으로 나타나지 않는 평가이건 상관없이 한국과 일본 등 동아시아 국가의 학생은 높은 수준을 보유하고 있다고 할 수 있다.

그런데 PISA의 결과는 이전의 국제 비교 연구의 성취도와 결부시켜 해석할 필요가 있다. 1995년 중학교 2학년을 대상으로 실시된 TIMSS에서 우리나라와 일본의 성취도 평균은 581점으로 동일하였다. TIMSS와 4년 시차를 두고 실시된 TIMSS-R에서 일본은 5위를 기록하였는데, 일본의 성취도 평균인 579점은 2위인 우리나라 학생의 평균 보다 8점이 낮은 점수이다. 이와 같은 TIMSS와 TIMSS-R의 결과에 비추어 볼 때, PISA에서 우리나라 학생의 수학적 소양 점수가 일본보다 낮은 것에 대해서 경각심을 가질 필요가 있다. 특히 1999년 2월에 실시된 TIMSS-R의 대상이 중학교 2학년이고, 2000년 7월에 실시된 PISA의 대상이 만 15세(대부분 고등학교 1학년)이므로, TIMSS-R과 PISA의 표집 대상은 다르지만 모집단은 동일하다. 이러한 점을 고려할 때 TIMSS-R에서의 우위가 1년여의 시간이 흐른 후 PISA에서의 열세로 돌아선 점에 대해서 주목할 필요가 있다.

2. 수학적 소양 점수 분포

수학적 소양 점수의 평균과 더불어 관심을 기울여야할 측면은 수학적 소양 점수가 어떤 분포를 보이는가 하는 점이다. <표 5>에는 PISA 연구에 참가한 국가 중 분석 대상이 된 OECD 가입국 27개국의 수학적 소양 점수 백

<표 5> 수학적 소양 점수의 백분위 분포

국가명	평균	표준편차	백분위					
			5%	10%	25%	75%	90%	95%
일본	557	87	402	440	504	617	662	688
한국	**547**	**84**	**400**	**438**	**493**	**606**	**650**	**676**
뉴질랜드	537	99	364	405	472	607	659	689
핀란드	536	80	400	433	484	592	637	664
호주	533	90	380	418	474	594	647	679
캐나다	533	85	390	423	477	592	640	668
스위스	529	100	353	398	466	601	653	682
영국	529	92	374	412	470	592	646	676
벨기에	520	106	322	367	453	597	646	672
프랑스	517	89	364	399	457	581	629	656
오스트리아	515	92	355	392	455	581	631	661
덴마크	514	87	366	401	458	575	621	649
아이슬란드	514	85	372	407	459	572	622	649
스웨덴	510	93	347	386	450	574	626	656
아일랜드	503	84	357	394	449	561	606	630
노르웨이	499	92	340	379	439	565	613	643
체코	498	96	335	372	433	564	623	655
미국	493	98	327	361	427	562	620	652
독일	490	103	311	349	423	563	619	649
헝가리	488	98	327	360	419	558	615	648
스페인	476	91	323	358	416	540	592	621
폴란드	470	103	296	335	402	542	599	632
이탈리아	457	90	301	338	398	520	570	600
포르투갈	454	91	297	332	392	520	570	596
그리스	447	108	260	303	375	524	586	617
룩셈부르크	446	93	281	328	390	509	559	588
멕시코	387	83	254	281	329	445	496	527
OECD 평균	500	100	326	367	435	571	625	655

분위 분포 결과가 제시되어 있다 (OECD, 2001).

위의 표는 각 백분위에 도달한 학생의 수학적 소양 점수에 대한 정보를 제공한다. 예를 들어 우리나라 학생 전체를 100%로 놓았을 때 소양 점수가 하위 5%에 해당하는 학생의 수학적 소양 점수는 400점이고, 10%, 25%, 75%, 90%, 95%에 해당하는 수학적 소양 점수는 각각 438점, 493점, 606점, 650점, 676점이다.

백분위에서 25%와 75%에 해당하는 수학적 소양 점수의 차이는 중간층을 이루는 50% 학생이 수학적 소양 수준에서 큰 차이를 보이는지 그렇지 않은지에 대한 정보를 제공한다. PISA 연구를 통해 밝혀진 사실 중의 하나는 수학적 소양 점수의 분포와 수학적 소양 점수의 평균 사이에 모종의 관련성이 있다는 점이

다. 즉 수학적 소양 점수의 분포에서 25%부터 75%에 해당하는 중간 50% 학생의 점수 차이가 크지 않은 경우, 전체적인 수학적 소양 점수의 평균은 대체적으로 높다. 예를 들어 25%와 75% 사이의 수학적 소양 점수의 차이가 적은 캐나다, 덴마크, 핀란드, 아이슬란드, 일본, 한국은 모두 OECD 평균보다 높다. 수학적 소양 점수의 편차가 크지 않은 동질적인 수준을 보인 국가 중에는 한국, 일본과 같은 동아시아 지역 국가와 사회 제 분야에 대한 전반적인 평등이 실현되어 있는 북유럽의 덴마크, 핀란드, 아이슬란드가 포함되어 있다. 이에 반해 중간 50% 학생의 수학적 소양 점수 차이가 크게 나타난 국가의 수학적 소양 점수는 낮은 경향이 있다. 가장 불평등한 분포를 보인 5개국인 독일, 그리스, 헝가리, 폴란드, 벨기에 중 벨기에를 제외하고는 모두 수학적 소양 점수의 평균이 OECD 평균보다 낮다.

전체적으로 우리나라 학생의 수학적 소양 점수의 평균은 2위를 기록하였지만, 상위 5% 학생의 점수는 뉴질랜드, 일본, 스위스, 호주, 영국에 이어 6위이다. 즉 전반적인 수학적 소양 점수에 비하여 상위 학생의 수학적 소양 수준은 상대적으로 낮은 편이다. 이러한 경향성은 TIMSS-R에서도 찾아볼 수 있는데, 우리나라는 전체적인 수학 성취도에서 2위를 기록하였지만 상위 5% 학생의 성취도 평균은 3위이므로, 상위 학생의 성취도는 상대적으로 약간 저조하다고 할 수 있다(Mullis et al, 2000a). 상위 수준 학생의 상대적인 부진이라는 TIMSS-R의 경향은 PISA에서 보다 극명하게 나타났으므로 학교급과 학년이 높아질수록 상위권 학생의 수학 성취도가 저조해지는 현상을 심각하게 받아들일 필요가 있다. 한 국가의 과학 기술의 수준은 수학 발달의 수준을 넘어서지 않으며, 국가 발전의 원동력을 제공하는 수학의 발달은 최상위권 학생에 의해 이루어진다는 점을 고려할 때, 상위 5% 학생의 수학적 소양 점수가 상대적으로 낮은 것은 일종의 적신호라고 할 수 있다. 국가 경쟁력을 높일 수 있는 집단은 최상위 집단이며, 이들의 수준에 부합되는 맞춤 교육을 제공하기 위해서 교육에서 수월성의 추구는 더 이상 미룰 수 없는 과제라고 하겠다.

3. 문항 정답율 비교

PISA 1주기 검사에 포함된 수학 문항의 수는 제한적이기 때문에, 수학 평가틀에 의거한 분석을 통해 경향성을 찾는 것이 용이하지 않다. 뿐만 아니라 어떠한 경향성을 포착하였다 하더라도 적은 수의 문항에서 추출한 특성이기 때문에 일반화하기는 어렵다. 이러한 점을 감안하여 본 논문에서는 문항 단위로 관심의 범위를 국지화시켜, 국내 정답률이 국제 정답률보다 낮은 문항이나 월등하게 높은 문항을 조사하였다. 다음 <표 6>에는 우리나라 학생의 정답률이 국제 평균보다 낮거나 유사한 문항을 제시하였다.

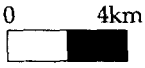
<표 6> 국내 정답률이 국제 정답률보다 낮거나 유사한 문항

문항명	정답률		수학적 능력 수준	영역통합적 사고	교육과정적 요소	상황	문항의 유형
	국내	국제					
보폭(2)	11.7	18.9	2수준	변화와 증가	대수	개인적	제한적 구성반응형
대륙의 넓이(2)	15.2	19.7	2수준	공간과 도형	측정	개인적	제한적 구성반응형
방의 가구 배치도	74.0	73.9	1수준	공간과 도형	기하	교육적	선택형

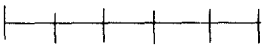
수학 평가 결과와 관련하여 흔히 제기될 수 있는 가설은 우리나라 학생들은 사지선다형의 문항에 익숙하기 때문에 선택형에서는 높은 정답률을 나타내지만, 생소한 열린 구성반응형 문항에 대해서는 정답률이 낮을 것이라는 점이다. 또한 수학적 능력에 따른 세 수준의 구분이 엄격한 위계를 이루고 3수준 문항의 난이도가 1수준 문항보다 반드시 높은 것은 아니지만, 단순 계산에 가까운 1수준 문항에서는 우리나라 학생의 정답률이 높지만 고등 사고 기능과 통합적 사고를 요구하는 3수준 문항에서는 그렇지 못할 것이라는 예측이 제기될 수 있다. 그러나 국내 정답률이 국제 정답률보다 낮거나 유사한 문항을 수학 평가틀에 의해 분류할 때 1, 2수준의 문항과 선택형, 제한적 구성반응형 문항이 포함되어 있어 위와 같은 가설과 예측이 성립하지 않음을 알 수 있다. 따라서 국내 정답률이 국제 정답률보다 낮거나 유사한 문항의 특징을 내용적인 측면에서 분석할 필요가 있다.

첫째, 보폭 과제의 두 번째 문항은 주어진 식과 보폭을 이용하여 분속과 시속을 구하는 문항으로, 답을 구하기 위해서는 복잡도가 높은 계산이 요구된다. 그런데 우리나라 학생은 검사시 계산기를 사용하지 않았고, 거의 모든 참여국에서 검사시 계산기를 이용하였기 때문에 정답률의 차이는 여기에서 기인한 것으로 판단된다. 우리나라에서는 계산기의 활용이 보편화되어 있지 않기 때문에, 익숙하지 않은 계산기를 검사에서 이용하는 것이 혼란을 초래할 수 있다는 국내 연구진의 판단에 따라 PISA 검사를 시행할 때 계산기의 사용을 허용하지 않았다. 대부분의 문항에서는 계산기의 사용 여부가 심각한 영향력을 미치지 않았으나, 보폭(2) 문항의 경우는 수반되는 계산의 복잡성 때문에 계산기의 활용 여부가 정답률에 영향력을 미쳤을 가능성이 높다.

둘째, 대륙의 넓이 과제의 두 번째 문항은 지도에 제시된 축척에 대한 정보와 자를 이용하여 남극 대륙의 넓이를 근사적으로 구하는 문항으로, 우리나라 학생들의 정답률이 낮은 이유는 두 가지 측면에서 설명할 수 있다. 우선 지도에 표시된 축척 정보가 학생들에게 익숙하지 않다는 점이다. 우리나라 학생들은 초등학교 4학년부터 지도를 읽는 방법을 학습하는데, 이 때 축척을 나타내는 방법은 비례법(예를 들어 1:200,000으로 축척을 표시)과 분수법(예를 들어 1/200,000로 축척을 표시)과 축척자의 세 가지이다. 문제에 제시된 방법은 축척자의 방법인데, 우리나라 학생들이 지도에서 주로 접하는 축척 정보는 왼쪽에 표시된 방식으로, 문제에 제시된 오른쪽의 방식에는 익숙하지 않다.



우리나라 학생에게 익숙한 축척 표현



문제에 제시된 축척 표현

뿐만 아니라 문항에 표시된 축척 정보에서 한 칸은 0.8cm 정도인데, 학생들은 다른 지도에서 경험해 온 바와 같이 한 칸이 1.0cm이라고 가정하고 답을 구한 경우가 많았다. 또한 축척 정보의 생소함 이외에 보폭(2) 문항과 마찬가지로 계산기의 사용 여부도 영향을 미쳤을 것으로 예상할 수 있다. 정확하게 축척 정보를 파악한 경우라도 정답에 이르기 위해서는 복잡한 계산이 수반되기 때문이다.

셋째, 방의 가구 배치도는 제시된 가구들의 측면도를 보고 평면도를 찾는 문항으로, 우리

나라와 국제 정답률이 거의 같다. 이 문항에 대한 우리나라 학생의 정답률이 국제 정답률보다 낮지는 않지만 다른 문항에 비해 상대적으로 저조한 정답률을 보인 문항이라고 할 수 있다. 우리나라도 생활 양식이 점차 서구화되면서 입식 생활과 관련된 식탁, 소파와 같은 가구가 보편화되고 있다. 그러나 PISA 검사를 위한 표집은 전국적으로 이루어졌고, 농어촌에서는 좌식 생활을 유지하는 경우가 적지 않기 때문에 식탁과 소파가 생소하게 받아들여지는 경우가 없지 않았을 것이다. 즉 방의 가구 배치도에서 우리나라 학생의 정답률이 상대적으로 낮은 것은 문제 상황에의 친숙도라는 일종의 문화적 편파성의 입장에서 부분적으로 설명할 수 있을 것이다.

한편 국내 정답률이 국제 정답률보다 10점 이상 높은 것은 10문항이다. 우리나라 학생이 우위를 보인 문항은 대부분 교과서에서 반복적으로 취급되거나 다른 검사 문항과 유사성이 높은 경우이다. 상식적인 논의이지만 피험자의 문항에 대한 익숙도가 문항의 정답률에 큰 영향을 미쳤다고 할 수 있다.

그 외에 수학 평가틀에서 설정한 '상황' 요소와 관련하여 한 가지 특징적인 현상을 찾아볼 수 있다. 국내 정답율이 국제 정답율에 비하여 월등하게 높은 문항 중 개인적 상황의 문항은 없고 수학적 상황과 교육적 상황의 비율이 높다. 개인적인 상황의 문항은 생활과 밀착된 다양한 상황이 수반한다. 이에 반해 수학적 상황의 문항이란 구체적인 배경을 동원하지 않고 기호와 수식으로 이루어진 탈맥락적인 문항을 말하며, 교육적 상황의 문항이란 문항이 어떠한 배경 하에 제시되기는 하지만 교육의 장면을 염두에 둔 정형화된 문항을 말한다. 우리나라 수학 수업은 대개 형식화, 추상화된 내용 설명과 문제 제시로 일관하는 경향이 있기 때문에 학생들은 구체적인 상황이 결부된 수학 문항을 다루는 경험을 충분히 하지 못하며, 이러한 교수-학습 관행은 정답률이 높은 문항 중 개인적인 상황의 문항은 전무하고, 그 대신 수학적인 상황과 교육적인 상황의 문항이 다수 포함된 것과 인과 관계를 이룬다고 할 수 있다.

4. 문항별 반응 빈도 분석

본검사에 포함된 문항 중 대부분의 제한적 구성반응형과 열린 구성반응형의 문항에 대해서 두 자리 코딩 방식을 적용하기 때문에 문항별 반응 빈도를 추적하는 것이 가능하다. 그 중 <부록>에 제시된 사과나무 과제와 교통사고 과제에 대하여 국내 학생의 구체적인 반응 빈도와 특징적인 현상을 조사하였다 (국내 분석이기 때문에 외국 자료는 포함되어 있지 않음).

<표 7> 사과나무(1) 문항의 코드별 반응 분포

채점 코드		코드 설명	반응 빈도 (%)	
오답	01	$n=5$인 경우의 두 값이 모두 틀림	110 (6.6%)	264
	02	기타	154 (9.2%)	(15.8%)
부분정답	11	$n=5$인 경우의 값 중 하나가 틀림	114 (6.8%)	142
	21	$n=2, 3, 4$인 경우의 값 중 하나가 틀림	28 (1.7%)	(8.5%)
정답	21	모든 칸의 답이 맞음	1204 (72.0%)	
오답	99	공란	62 (3.7%)	
계			1672 (100%)	

사과나무 과제의 첫 번째 문항은 주어진 표의 빈 칸에 값을 채우는 문항으로 한 칸도 예외 없이 모두 정확하게 답한 경우에만 정답으로 처리된다. 이와 같이 엄격한 기준에도 불구하고 우리나라 학생의 반응 중 72%가 정답으로 분류되었다. 최근 들어 수학 교과서에서 강조하고 있는 것은 주어진 몇 가지 사례로부터 규칙성을 찾아 일반화하는 것으로, 초등학교에서부터 연습의 기회를 빈번하게 갖기 때문에 높은 정답율을 보인 것으로 판단된다. 부분정답 중 코드 11이 코드 12보다 높은 빈도를 나타냈는데, 이를 통해 $n=5$인 경우의 값을 구하는 것이 $n=2$, 3, 4에 비하여 훨씬 어렵다는 것을 알 수 있다.

사과나무 과제의 두 번째 문항은 사과나무와 소나무의 그루 수가 같아지는 n을 구하는 문제로, n2=8n이라는 대수식으로 접근한 비율이 예상보다 낮아 14.6% 밖에 되지 않았다. 이에 반해 규칙성을 확장하기 위하여 표를 이용한 빈도의 비율이 38.5%로 가장 높았다. 그림 그리기 등의 방법을 이용하는 빈도는 아주 낮아 그 비율이 1.5% 밖에 되지 않았는데, 학교에서 이루어지는 평가에서 그림 그리기가 하나의 방법으로 인정되기보다는 보조적인 역할을 하는 것으로 인식되기 때문인 것으로 해석할 수 있다.

사과나무 과제의 세 번째 문항은 주어진 정보에 기초하여 사과나무와 소나무 중 어느 것의 그루 수가 더 빠르게 증가하는지 구하고 그

<표 8> 사과나무(2) 문항의 코드별 반응 분포

채점 코드		코드 설명	반응 빈도 (%)
오답	00	기타 (예를 들어 '$n=0$ 포함'이나 '$n^2=8n$' 이라는 식만 쓴 경우)	189 (11.3%)
정답	11	대수적 방법을 이용하여 $n=8$을 구함 ($n^2=8n$, $n^2-8n=0$, $n(n-8)=0$, $n=0$과 $n=8$, 따라서 $n=8$)	244 (14.6%)
	12	규칙성을 확장하기 위하여 표를 이용하여 $n=8$을 구함	644 (38.5%)
	13	그림 그리기 등 기타 방법을 이용하여 $n=8$을 구함	25 (1.5%)
	14	대수적인 방법을 이용하여 $n=0$, 8을 구함	45 (2.7%)
오답	99	공란	525 (31.4%)
계			1672 (100%)

<표 9> 사과나무(3) 문항의 코드별 반응 분포

채점 코드		코드 설명	반응 빈도 (%)	
오답	01	사과나무라고 답하였으나 답을 뒷받침하는 설명이 불충분하거나 틀렸거나 없음	363 (21.7%)	835 (49.9%)
	02	기타 (소나무, 모르겠다 등)	472 (28.2%)	
부분정답	11	구체적인 예를 들거나 표를 확장함으로써 사과나무라고 답함	47 (2.8%)	170 (10.2%)
	12	사과나무라고 답하였으며, n^2 과 $8n$ 사이의 관계를 이해한다는 증거를 보여주기는 하였으나 코드 21과 같이 명확하게 표현하지 못함.	123 (7.4%)	
정답	21	타당한 설명과 함께 사과나무라고 답함	337 (20.2%)	
오답	99	공란	330 (19.7%)	
계			1672 (100%)	

<표 10> 교통사고 문항의 코드별 반응 분포

채점 코드		코드 설명	반응 빈도(%)	
오답	01	'아니오'라고 답했으나 타당성있는 설명이 수반되지 않음	164 (14.7%)	590 (52.7%)
	02	그래프의 외형에 초점을 맞춰 '예'라고 답함	54 (4.8%)	
	03	설명이 없거나 코드 02와 다르게 설명함	174 (15.6%)	
	04	기타 오답	198 (17.7%)	
부분정답	11	'아니오'라고 답했으나 상세한 설명이 부족함	264 (23.6%)	266 (23.8%)
	12	'아니오'라고 답했으나 비나 퍼센트의 계산 과정이 잘못됨	2 (0.2%)	
정답	21	그래프의 일부만 보여주고 있다는 점에 초점을 맞추어 '아니오'라고 답함	74 (6.6%)	167 (15.0%)
	22	증가의 비나 퍼센트를 사용하여 '아니오'라고 답함	32 (2.9%)	
	23	판단을 위해서는 추세 변동 자료가 필요하다는 설명과 함께 '아니오'라고 답함	61 (5.5%)	
오답	99	공란	95 (8.5%)	
계			1118 (100%)	

이유를 설명하는 문항이다. 채점 기준을 보면 사과나무라고 옳게 답한 경우라도 설명의 유무나 수준에 따라 정답, 부분정답, 오답으로 분류된다. 우리나라 학생의 반응 중 정답인 코드 21에 해당하는 경우가 20.2%이고, 부분정답에 해당하는 경우가 10.2%이므로 전체적으로 30% 정도의 학생이 옳은 답과 함께 완전하거나 부분적인 설명을 시도하였다. 그런데, 사과나무라는 옳은 답을 하였지만 설명이 불충분하거나 틀리거나 없는 경우는 코드 01에 해당하는데, 이에 대한 반응 빈도 비율이 21.7%나 된다. 코드 01로 분류되는 답을 한 학생 중에는 추측에 의하여 사과나무라고 적은 경우도 있지만, 학생 나름의 수학적인 판단에 의하여 사과나무임을 알았지만 이를 타당한 논리에 의해 명료하게 설명하지 못한 경우도 다수 포함되어 있을 것이다. 일부의 학생은 "소나무는 생명력이 길어서 무성하게만 자랄 뿐 사과나무의 그루수가 더 빨리 증가할 것 같다"와 같이 수학적인 타당화가 아닌 일상적인 의미의 설명을 시도한 경우도 있다. 이를 통해 우리나라 학생에게 상대적으로 결핍되어 있는 것은 자신의 사고를 수학적 방식으로 설득력있게 정당화하는 능력임을 알 수 있다.

교통사고 문항[6]은 연간 교통사고의 변화를 나타내는 세로 막대그래프를 제시하고 교통사고가 급격하게 증가하였다는 해석이 적절한지 그렇지 않은지 답하고 그 이유를 설명하는 문제이다. 교통사고의 건수는 508에서 516 정도로 증가하였는데 그래프의 y축의 눈금 중 0~500 정도까지 생략되어 있다. 따라서 이 문항에 답하기 위해서는 그래프 상으로 크게 증가한 것처럼 보이지만 실제 자료는 그렇지 않다는 것을 인식하고 자신의 판단의 정당화할 수 있어야 한다.

이 문항은 그래프의 눈금이 어떻게 설정되어 있는지를 확인하거나 전체 수치에 비하여 변화의 정도가 미미하다는 것을 인식하면 그리 어렵지 않게 답할 수 있는 문항임에도 불구하고 완전 정답에 해당하는 반응이 15% 밖에 되지 않는다. 여기서 주목할 만한 현상은 '아니오'라는 사실을 알고 있지만 설명이 부족한 경우가

6) 이 문항은 고등학생을 대상으로 한 TIMSS 검사 도구를 PISA에서 활용한 경우로, TIMSS에서의 국제 정답율 평균은 100점을 기준으로 남학생은 24점, 여학생은 15점이었다(Mullis, et al, 2000b).

<표 11> 배경 변인이 수학적 소양에 비치는 영향 지표

국가명	사회 경제적 지위		물질적 재산		수학에 대한 흥미도	
	사회 경제적 지표	사회 경제적 지위의 영향 지표	물질적 재산 지표	물질적 재산의 영향 지표	수학에 대한 흥미도 지표	수학에 대한 흥미도의 영향 지표
일본	50.3	12.2	-0.13	9.0		
한국	**43.0**	**22.4**	**-0.28**	**26.2**	**-0.27**	**26.7**
뉴질랜드	52.4	31.4	0.22	22.2	0.09	13.5
핀란드	50.0	19.6	0.22	15.5	-0.07	25.0
호주	52.3	31.4	0.42	17.0	0.04	15.0
캐나다	52.9	21.5	0.42	13.8		
스위스	49.0	38.1	0.04	12.6	-0.03	6.9
영국	51.2	35.4	0.44	18.5		
벨기에	49.1	37.6	-0.10	16.2	-0.11	12.5
프랑스	48.4	27.7	-0.16	25.5		
오스트리아	49.8	31.7	0.25	13.9	-0.23	7.8
덴마크	49.8	25.1	0.49	16.4	0.47	17.1
아이슬란드	52.4	17.0	0.54	-7.7	0.11	22.9
스웨덴	50.3	31.5	0.64	14.1	-0.21	16.0
아일랜드	48.0	26.1	0.05	21.9	-0.01	7.8
노르웨이	53,7	26.1	0.57	10.4	-0.28	22.6
체코	48.2	43.0	-0.87	19.8	-0.07	13.5
미국	52.3	38.2	0.61	42.0	0.08	12.0
독일	49.1	41.0	0.18	26.5	-0.07	8.9
헝가리	49.8	40.9	-0.86	27.4	-0.04	14.0
스페인	44.9	28.3	-0.13	19.0		
폴란드	45.9	35.8	-0.98	16.2		
이탈리아	46.9	21.6	0.10	15.5	0.00	9.3
포르투갈	44.2	35.9	-0.12	30.8	0.26	15.5
그리스	48.3	30.1	-0.43	23.4		
룩셈부르크	44.4	34.7	0.31	28.2	-0.18	0.5
멕시코	42.7	31.2	-1.44	22.7	0.39	9.5
OECD 평균	48.9	30.2	0.00	19.2	0.00	10.7

* '수학에 대한 흥미도'에 대한 설문 조사는 일부 국가에서만 실시되었기 때문에, 모든 국가에 대한 자료가 제시되어 있지 않다.

23.6%, 설명을 수반하지 않은 경우가 14.7%에 달해, 우리나라 학생에게 결여되어 있는 것은 자신의 의견을 체계적으로 설명하는 능력임을 알 수 있다. 설명을 시도한 경우라도 "자동차가 많아짐으로써 교통 사고의 위험이 커진다", "운전자들의 이기적이고 개인적인 성향이 강해졌기 때문이다"와 같이 질문의 의도를 벗어난 지엽적인 답변을 한 경우도 다수 발견할 수 있다.

5. 배경 변인과 수학적 소양 수준

학생을 둘러싼 다양한 배경 변인이 수학적 소양 점수에 영향을 미친다는 것은 주지의 사실이다. PISA의 설문 조사에 포함된 변인 중 수학적 소양에 미치는 영향력이 높은 세 가지 변인에 대한 지표는 다음과 같다(OECD, 2001).

첫째, 사회 경제적 지표는 아버지와 어머니가 어떤 종류의 직업에 종사하며 직장 내에서 어떤 일을 하는지에 대한 설문을 통해 산출되며, 값이 클수록 사회 경제적 지위가 높음을 의미한다. 사회 경제적 지위의 영향 지표는 사회 경제적 지표가 수학적 소양 점수에 미치는 영향력을 수치화한 것이다. 우리나라의 경우 수학적 소양 점수에 대한 부모의 사회 경제적 지위의 영향 지표는 22.4로 OECD 평균에 비하여 낮다. 사회 경제적 지위의 영향력이 가장 미약한 국가는 일본이며, 그 뒤를 이은 것이 아이슬란드나 핀란드와 같은 북유럽의 국가들이다. 주지하고 있는 바와 같이 북유럽의 국가에는 사회보장 제도가 철저하게 갖추어져 있기 때문에 교육을 비롯한 대부분의 것이 균등화, 동질화 되어 있어 실질적인 차이가 크게 존재하지 않는다. 이러한 점에서 사회 경제적 지위의 영향 지표가 낮은 것을 이해할 수 있다.

한편 우리나라나 일본과 같은 동아시아의 국가에서는 부모의 지위 고하를 불문하고 높은 교육열을 지니고 있는 것으로 잘 알려져 있다. 부모가 자녀의 교육에 최상의 가치를 두는 이유는 쉽게 생각할 수 있는 바와 같이 자녀가 장차 사회적으로 의미 있는 기여를 하거나 출세를 하는 경로는 대개 대학 진학으로 귀결되기 때문이다. 사회 경제적 지위가 낮은 직업을 가지고 있는 부모나 사회적으로 명망을 얻는 높은 지위에 있는 부모가 공통적으로 염원하는 것은 자녀의 미래를 보장해주는 좋은 대학에 진학하는 것이기 때문에, 우리나라와 일본에서는 사회 경제적 지위에 따른 수학적 소양 점수의 차이가 크게 나타나지 않는 경향이 있다.

둘째, 물질적 재산 지표는 가정에서 보유하고 있는 물질적인 재산의 정도를 나타내는 값으로, OECD 전체의 평균을 0으로 하므로 양수의 값을 가지면 OECD 평균보다 물질적 재산이 많음을 나타내고 음수의 값이면 그 반대의 경우가 된다. 여기서 물질적 재산은 사회 경제적 지위와 밀접하게 연관되어 있기는 하지만 양자가 반드시 일치하는 것은 아니다. 사회 경제적 지위가 높다고 해서 반드시 경제적인 보상이 뒤따르는 것이 아니며, 또 동일한 사회 경제적 지위에 있더라도 가치관에 따라 물질적 부의 보유 정도가 다를 수 있다.

우리나라의 경우 물질적 재산이 수학적 소양 점수에 미치는 영향력은 OECD 평균보다 큰 것으로 나타났다. 앞에서 지적한대로 우리나라와 일본은 사회적 지위와 무관하게 높은 교육열을 가지고 있지만, 그 교육열이 학생에게 구체적인 영향을 미치기 위해서는 경제력이 뒷받침되어야 한다. 이러한 점을 고려할 때, 물질적인 재산은 사회 경제적 지위보다 수학적 소양 점수에 보다 실제적인 영향력을 행사할 수 있다.

그런데 우리나라의 영향 지표인 26.2는 OECD 평균 19.2보다 7점이 높은 값으로, 물질적 재산이 미치는 영향력의 순위가 27개국 중 여섯 번째이다. 만일 PISA가 교육과정과 교과서에 긴밀하게 연관되어 있다면 사교육의 유무와 정도가 보다 직접적인 영향력을 미칠 것이며, 따라서 물질적인 재산과의 관련성도 더욱 높아질 것이다. 한편 물질적 재산이 수학적 소양 점수에 가장 강력한 영향력을 미치는 국가는 미국인데, 자본주의가 극도로 발달되어 있는 국가로 재산의 정도에 따라 교육과 관련된

수혜의 기회가 차별화된다는 점에서 이해할 수 있다.

셋째, 수학에 대한 흥미도 역시 물질적 재산과 마찬가지로 OECD 전체의 평균을 0으로 설정하기 때문에 양수의 값을 가지면 수학적 흥미도가 높다는 것을, 음수이면 그 반대임을 나타낸다. PISA의 설문 조사 결과 우리나라는 노르웨이에 이어 수학에 대한 흥미도가 두 번째로 낮은 것으로 나타났다. 사실 높은 수학적 성취도와 대비되는 수학에 대한 낮은 태도(흥미, 자신감)는 기존의 국제 비교 연구에서 지속적으로 나타난 경향으로(Beaton et al, 1996; Mullis et al, 2000a), 매번 이에 대한 많은 비판과 자성이 뒤따랐다. 앞에서 지적한 바와 같이 우리나라는 사회에 진출하는 다중적인 경로가 열려 있는 서양과 달리 일정 수준의 직업을 취득하기 위해서는 대학 졸업이라는 기본 조건이 갖추어져야 한다. 따라서 학교교육의 구심점이 대학 진학에 놓여 있으며, 수학에서의 높은 성취도는 대학 진학을 위한 필요충분조건에 해당하기 때문에 수학에 대한 부정적인 태도에도 불구하고 수학 학습에 매진한다고 볼 수 있다.

그런데, 우리나라 학생들의 부정적인 태도를 약간 다른 시각에서 바라볼 수도 있다. 즉 동양에서는 겸손을 중요한 미덕 중의 하나로 간주하기 때문에 자신의 능력을 평가할 때 한 단계 절하하거나, 혹은 중용의 미를 강조한 영향으로 극단적인 응답보다는 미온적인 응답을 선호하는 경향이 있다. 이러한 성향은 학생이 설문 조사에 응할 때 반영되었을 가능성이 높다. 또한 동양의 오랜 가르침에서 스승은 학생이 자만하지 않도록 적당히 채근하고 독려하는 것을 이상적으로 여기기 때문에, 동양의 교사는 서양의 교사에 비하여 칭찬에 인색하고, 이로 인해 학생의 수학에 대한 자신감이 감소되었을 가능성이 높다(Leung, 2001; Park, 2001). 이러한 사실은 설문 조사에서 왜 동양의 학생이 서양의 학생에 비하여 저조한 반응을 나타냈는지에 대한 부분적인 설명이 될 수 있다. 물론 우리나라 학생이 수학에 대한 부정적인 태도를 소유하고 있다는 것에 대해서는 이론의 여지가 없지만, 태도 조사 결과를 수용할 때 동양적인 정서를 감안하여 그 결과를 약간 조정할 필요가 있다.

넷째, 우리나라의 경우 수학에 대한 흥미도가 수학적 소양 점수에 미치는 영향 지표가 가장 높은 것으로 나타났다. 즉 수학에 대한 흥미도는 수학적 소양 점수에 영향을 미치는 결정적인 요인이므로, 수학적 흥미를 유발시키기 위한 다양한 장치를 동원해야 한다는 시사점을 얻을 수 있다. 여기서 수학적 흥미를 북돋우기 위한 방법에 대해 재고할 필요가 있다. 보통 학생의 흥미를 유발시킨다고 하면 수학 교수·학습에 재미있는 소재를 이용한다는 생각을 떠올리는 것이 대부분이다. 그런데 추상적인 수학을 친근한 상황이나 재미있는 도입으로 포장하더라도, 학습자가 언젠가는 그 포장지를 벗기고 형식화된 수학을 직접적으로 대면해야 한다는 점을 생각하면 과도한 포장이 바람직하지만은 않다.

또한 다양한 상황과 맥락을 통해 수학을 포장할 때 학습자의 순간적인 흥미는 촉발시킬 수 있지만 수학에 대한 영속적인 흥미는 결국 추상적인 수학을 어렵게 이해해 나가는 가운데 습득되는 것이라는 점에서 상황과 맥락의 동원은 한계를 갖는다고 할 수 있다. 따라서 피상적으로만 흥미로워 보이는 일차적인 자료를 수학 내용의 본질이 담길 수 있는 이차적인 자료로 가공하는 절차가 필요하다. 다시 말해 수학에 대한 흥미를 증진시키기 위한 방안의 모색은 재미있는 소재를 발굴하는 정도의 단순한 작업이 아니며, 재미있는 소재가 수학 내용적

인 측면과 유기적으로 결합되어 시너지 효과를 낼 수 있도록 장치하는 복합적인 과정이라고 할 수 있다.

VI. 논의 및 시사점

국제 비교 연구의 의의는 각 국가 학생의 능력을 교과별로 한 줄로 세워 순위를 정하는데 있는 것이 아니라, 국제 비교 결과에 비추어 각 국의 교육 제도, 교육 내용, 방법의 전반을 점검하는데 있다. 현재 우리에게 필요한 것은 세계화 시대의 도래와 더불어 우리나라 학생의 수준을 국제적인 잣대에 의해 판정하는 것이기 때문에, PISA 결과로부터 유용한 시사점을 추출할 수 있을 것이다.

우선 PISA는 우리에게 상당히 고무적인 결과를 제공한다. PISA는 교육과정과 교과서의 내용을 충실히 반영하고 선택형을 위주로 하는 기존의 국제 비교 연구와 차별화되기 때문에, 우리나라를 비롯한 동아시아의 국가가 지금까지 보여온 최상위의 기록이 유지되지 못할 것이라는 부정적인 예측이 설득력있게 제기되었다. 다시 말해 우리나라를 비롯한 동아시아의 국가에서는 대학 입시가 교육의 구심점을 이루고 이를 위한 사교육이 팽배해 있으며, 공교육이나 사교육에서 강조하는 것은 단편적이고 분절화된 수학적 기능과 단순 계산 위주이며, 따라서 고등 사고 기능이 요구되지 않는 문항에만 능하다는 비판을 받아왔다. 그러나 PISA 수학 결과에서 일본과 한국이 나란히 1, 2위를 기록함으로써 이러한 비판이 유효하지 않다는 것을 입증하였다.

이와 같이 일단은 긍정적인 해석을 내려볼 수 있지만, PISA 결과에는 많은 부정적인 측면이 공존한다. 첫째, TIMSS나 TIMSS-R에서 일본과 비슷하거나 우위를 보였던 우리나라 학생의 수학적 소양 점수가 1위를 기록한 일본보다 10점이 낮다는 점이다. 물론 그 차이가 통계적으로 유의미하지는 않지만 다음 세대의 주역이 될 15세 학생의 수학적 소양 점수가 일본에 뒤졌다는 점에 대해서는 문제 의식을 가질 필요가 있다.

둘째는 우리나라 상위 5% 학생의 수학적 소양 점수가 6위에 머물렀다는 점이다. 다시 말해 우리나라가 2위의 수준을 보유할 수 있었던 것은 중위권에 수학적 소양 점수가 밀집해 있기 때문이다. 물론 이는 '모든 사람을 위한 수학(Mathematics for all)'이라는 민주주의의 정신에 부합되는 결과로 보일 수도 있지만, 국가 경쟁력은 수학의 발달 수준에 종속되며, 이는 다시 상위 학생의 수학적 수준에 의해 결정된다는 점에 비추어 볼 때 경각심을 가질 필요가 있다. 상위권 학생을 위한 수월성의 추구는 일면 교육 수혜의 평등이라는 대원칙에 위배되는 것처럼 오인되기도 하지만, 우리가 추구해야 할 바는 무조건적인 평등이 아니라 능력에 따른 평등이라는 점을 분명히 해야 한다.

셋째, 이전의 국제 비교 연구에서 드러났듯이 우리나라는 높은 수학적 소양 수준에도 불구하고 수학에 대한 흥미도는 참여국 중 두 번째로 낮은 것으로 밝혀졌다. 우리나라 학생이 설문 조사에 답할 때에 동양 고유의 겸허함이나 중용의 미덕이 작용하여 부정적인 태도가 실제보다 더 과장되어 나타났다는 점을 고려하더라도, 저조한 수학 흥미도는 분명 바람직한 현상이 아니다.

넷째, 우리나라 학생은 풍부한 맥락을 배경으로 하는 문장제에서 두드러진 열세를 보이지는 않았지만, 국내 정답률이 국제 정답률에 비하여 월등하게 높은 문항은 상황을 동반하지 않는 순수 수학적 문항이거나 정형화된 소재를

배경으로 하는 교과서적인 문항이다. 학습자가 새로운 개념과 원리를 자신의 기존 인지 구조에 의미 있게 연계시키기 위해서는 실제적이고 구체적인 상황과 더불어 학습하는 것이 효과적이라는 점을 고려할 때, 실생활 맥락의 수학을 강력하게 지향할 필요가 있다.

다섯째, 제한적 구성반응형과 열린 구성반응형 문항에 대한 반응 분석 결과, 우리나라 학생에게 결핍되어 있는 것은 자신의 사고를 타당하게 설명하고 답을 정당화하는 의사소통 능력이다. 몇몇 문항의 경우 다수의 학생이 옳은 답을 적었지만 설명이 생략되어 있어 부분정답이나 오답으로 분류된 비율이 높다. 물론 이 중에는 추측에 의해 우연히 정답을 적은 경우도 있지만 상당수의 학생들은 수학적으로 사고하여 답을 도출했지만 이를 적절한 방법으로 표현하는 능력을 구비하고 있지 못하였기 때문으로 판단된다. 또한 설명을 시도한 경우라도 질문의 의도를 제대로 포착하지 못하고 질문과 무관한 일상적인 이유를 열거한 경우도 다수 발견할 수 있었다. 따라서 자신의 수학적 판단을 적절한 근거에 의해 논리적으로 정당화하는 문항을 풍부하게 경험하도록 하는 것이 필요하다.

마지막으로 평가에서 PISA의 채점 기준과 방식을 적극적으로 참고할 필요가 있다. PISA는 서술형 문항에 대하여 나타날 수 있는 가능한 반응을 반영하여 상세한 채점 기준을 작성하고, 채점의 정확도와 신뢰도 높이기 위하여 복수 채점을 하였다. 최근 몇 년 동안 우리나라에서 광범위하게 시도된 수행평가 실시의 최대 걸림돌은 채점의 객관성 문제였다. PISA에서는 채점자의 자의적인 판단을 최소화할 수 있도록 정교한 채점 기준을 마련하여 피험자의 다양한 반응을 비교적 공정하게 채점할 수 있는 장치를 마련하였다. 특히 두 자리 코드에 의한 채점 기준에서 십의 자리와 일의 자리가 각각 유의미한 정보를 제공할 수 있도록 체계화한 것은 우리의 평가 상황에 유용한 시사점을 제공한다. 물론 이러한 채점은 시간과 비용이 많이 드는 평가이기 때문에 단시간에 수백 명의 답안지를 채점해야 하는 우리의 학교 현실에 적용하기에는 어려움이 많다. 그러나 평가 결과가 상급 학교 진학을 위한 전형 자료의 일부로 활용되는 점을 고려할 때 평가자의 주관적 판단에 의한 오차를 최소화해야 하며, 따라서 평가의 객관성을 높일 수 있는 PISA의 채점 방식을 참고할 필요가 있다.

참 고 문 헌

박경미 (1999). OECD/PISA 학생성취도 지표 개발 연구의 소개. 학교수학 제1권 1호. pp. 367-382.

정영옥 (1997). Freudenthal의 수학화 학습-지도론 연구. 서울대학교 대학원 박사학위논문.

한국교육과정평가원 (2001). PISA 2000 수학 평가 결과 분석 연구. 한국교육과정평가원 연구보고 RRE 2001-9-3.

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

Freudenthal, H. (1983). *Didactical phenomenology of mathematical structures.* D. Reidel Publishing Company.

Leung, F.S. (2001). In search of an east asian identity in mathematics education. *Educational Studies in Mathematics, 47.* 35-51.

Mullis, I. V. S. *et al (2000a). TIMSS 1999 International mathematics report.* Boston, MA.: The International Study Center. Chestnut Hill, MA: Boston College.

Mullis, I. V. S. *et al (2000b). Gender differences in achievement, IEA's third international mathematics and science Study. TIMSS International Study Center.* Chestnut Hill, MA: Boston College.

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.

Park, K. (2001). The findings and implications from the TIMSS and TIMSS-R Korean data. 수학교육학연구 11(1), 123-146.

A Mathematical Literacy Profile of Korean Students in PISA

Kyung-Mee Park (Hongik University)
Seung-Hyun Choe (KICE)
Kooghyang Ro (KICE)

The PISA(Program for International Student Assessment), an international comparative study supervised by OECD, is aimed at producing reliable and internationally comparable indicators of students' literacy in reading, mathematics, and science. In mathematical literacy, Korean students ranked the 2nd out of the 32 participating countries in PISA. This result is very encouraging in the sense that the scores in the mathematical literacy is the forecasting indicator for the mathematical level of future citizens who are supposed to lead their countries in every fields.

In contrast to the high performance in mathematics, Korean students possess extremely low level of interest in mathematics. On the other hand, Korea's highest top 5% students are not on a par with the best students in New Zealand, Japan, Swiss, Australia, and England. This finding urges Korean education to put serious efforts to pursue academic excellence.

Overall, PISA results suggest that Korean mathematics education has been effective and efficient. Korean education system should set goals to meet the world standard and the challenges of the lifelong learning society.