

# PISA 2003, 2012, 2022 비교를 통한 한국 학생의 수학 성별 격차 분석: 학업성취도 및 정의적 특성을 중심으로

2025. 12. 16

성인지통계센터 이다경 부연구위원

- 1) OECD에서 주관하는 국제학업성취도 평가(PISA)에서 수학이 주영역이었던 2003년, 2012년, 2022년 결과를 바탕으로 우리나라 학생들의 인지적·정의적 성취에서 나타난 성별 격차를 살펴보자 함.
- 2) 분석자료는 OECD 「PISA 2003, 2012, 2022」의 한국 학생의 수학 학업성취도 및 수학의 정의적 특성 변수의 원자료를 활용함.
- 3) 분석결과, 전체 평균 점수에서의 남학생 우위 현상은 PISA 2003(23.40점 차,  $p<0.001$ )과 PISA 2012(17.92점 차,  $p<0.05$ )까지 유지되었으나, PISA 2022에서는 그 차이가 5.08점으로 줄어들며 통계적 유의성이 사라짐.
- 4) 성취수준별로 살펴본 결과, 상위권 남녀 점수 차가 대폭 축소(2003년 16.72점 → 2022년 7.42점)됨. 2012년 대비 2022년 남녀 모두 상위권 비율은 급감(남 24.5%→13.2%, 여 15.9%→7.9%)하고, 하위권 비율은 급증(남 14.0%→27.8%, 여 14.8%→24.9%)함.
- 5) 수학 불안감 및 수학 자아효능감의 변화를 살펴보면, 2022년 기준 여학생의 수학 불안감 (0.18)은 남학생(-0.09)보다 높고 자아효능감(-0.15)은 남학생(0.12)보다 낮아, 20년 전과 유사한 수준의 성별 차이가 유지됨.
- 6) 성취수준별로 정의적 특성을 분석한 결과, 하위권보다 상위권 집단에서 남녀 간 수학 불안감과 자아효능감의 차이가 더 크게 나타나는 경향이 2003년부터 2022년까지 일관되게 관찰됨.
- 7) 하위권 비율의 급증과 관련된 기초학력 미달을 해결하기 위해, 디지털 기기의 올바른 사용 교육과 어휘력 및 문해력 향상을 위한 깊이 있는 사고 훈련을 선행하여 기초학력 미달을 지원해야함. 높은 성취에도 낮은 효능감 및 불안감을 보이는 여학생들을 위해 성장 마인드셋을 통해 평가할 필요가 있으며, 교사는 결과보다 과정 중심 피드백 및 평가를 제공하여 여학생들의 정의적 성취를 지원하는 방안을 고려할 수 있음.

## I. 서론

국제학업성취도 평가(PISA)에서 수학이 주영역이었던 2003년, 2012년, 2022년 결과를 바탕으로 우리나라 학생들의 인지적·정의적 성취에서 나타난 성별 격차를 살펴보고자함.

- OECD가 주관하는 국제 학업성취도 평가(Programme for International Student Assessment, PISA)는 만 15세 학생들이 미래 사회에 필요한 핵심 역량을 갖추었는지 국제적으로 비교하는 평가임(OECD, 2023b).
  - 2000년부터 3년 주기로 시행된 이 평가는 학업성취도뿐만 아니라 학생들이 습득한 지식과 기술을 다양한 실생활 문제 해결에 얼마나 잘 적용하는지를 평가하는 '소양(literacy)' 개념을 도입하였음(김성경 외, 2023).
  - 수학 영역은 PISA 2003, 2012, 2022년 평가에서 주영역<sup>1)</sup>으로 다루어졌으며, 이는 약 10년 주기로 수학 교육의 국제적 동향과 인지적·정의적 성취도 변화를 살펴볼 수 있음.
- 본 리포트는 PISA 2003, 2012, 2022년의 결과를 바탕으로 우리나라 학생들의 수학 성취도에 나타난 성별 격차를 시계열적으로 분석하는 것을 목적으로 함. PISA 2003부터 2022까지 약 20년의 기간 동안 수학 평가의 개념은 시대적 요구에 따라 변화하였음. PISA 2012에서는 수학적 모형이, PISA 2022에서는 수학적 추론이 핵심 역량으로 부각되었음(OECD, 2023b). 이러한 평가 틀(assessment framework)의 변화 속에서 성별에 따른 성취도 격차가 어떻게 변화했는지, 그리고 인지적 성취의 성취수준별 학생 분포에서 성별 차이가 어떻게 나타나는지를 살펴보는 것은 매우 중요함.
- 추가적으로, 본 리포트는 단순한 성취도 비교를 넘어 학생의 수학에 대한 불안감, 자아효능감 등 정의적 특성이 성별 격차와 어떻게 상호작용하는지 살펴보고자 함. 높은 성취도에도 불구하고 우리나라 학생들의 수학에 대한 낮은 자아효능감과 높은 불안감은 과거 PISA 주기부터 지속해서 지적된 문제임. 이러한 정의적 특성들이 성별에 따라 어떻게 나타나는지, 성취수준별 학생 분포의 성별 차이에 따라 어떻게 나타나는지 살펴보는 것은 성취도 격차의 잠재적 원인을 살펴보는 계기가 될 수 있음. 주기의 흐름에 따른 성별 격차의 변화와 정의적 특성을 종합적으로 분석함으로써, 교육 정책 방향을 제시하고자 함.

1) PISA는 읽기, 수학, 과학 성취도와 교육맥락 변수를 지속적으로 측정하는 한편, 매 주기 주영역(main domain)을 설정하여 해당 영역의 평가 틀과 문항을 중점적으로 보완 및 개편함(OECD, 2023b). PISA 2000의 읽기 영역을 시작으로 2003년 수학, 2006년 과학 순으로 순환됨. 단, 2021년 예정이던 평가는 코로나19 팬데믹으로 1년 연기되어, PISA 2022에서 수학을 주영역으로 시행됨(OECD, 2023b).

## II. 분석자료

PISA 2003, 2012, 2022에 참여한 한국 만 15세 학생을 분석대상으로 하며, 지필(PBA)에서 컴퓨터(CBA) 및 적응형(MSAT)으로 고도화된 평가 설계의 변화와 복합 표본 데이터 구조를 반영하여 분석을 실시함

### 1. 분석 대상 및 평가 설계의 변화

○ 본 리포트는 PISA 주기 중 수학이 주영역이었던 PISA 2003, 2012, 2022 자료를 활용하였음. 해당 주기에 참여한 한국의 만 15세 학생들이 분석 대상이며 인지적·정의적 성취의 성별 격차를 분석하고자 함. PISA는 2000년을 시작으로 2022년까지 평가 매체와 표집 설계 방식에 있어 중대한 변화를 겪었으며, 시기별 자료의 특성은 다음과 같음.

- PISA 2003는 전적으로 지필 평가(Paper-Based Assessment, PBA) 방식으로 수행됨. 제한된 시간(120분) 내 효율적인 평가를 위해, 전체 개발된 문항을 여러 문항군으로 나누고 4개의 군집으로 구성된 13종의 검사지(booklet)를 학생들에게 무작위로 배분하는 행렬 표집(matrix sampling) 설계를 적용함(OECD, 2023a).
- PISA 2012는 지필 평가를 기본으로 하되, 선택적으로 컴퓨터 기반 평가(Computer-Based Assessment, CBA)를 도입함. 학생들은 지필 평가를 수행한 후 별도의 컴퓨터 기반 세션에 참여하는 방식을 따랐으며, 한국은 이 과정에 참여함(송미영 외, 2013). 설계 방식은 PISA 2003과 유사하게 지필 13종 및 컴퓨터 기반 수학 문항군 24종을 활용한 회전 행렬 검사 설계를 유지함.
- PISA 2022는 CBA를 표준으로 전면 전환하여 대다수 참여국이 CBA를 실시함. 특히 수학 영역에는 단계적 적응형 평가(Multistage Adaptive Testing, MSAT)를 도입했는데, 이는 학생의 이전 단계 성취도에 따라 후속 소검사의 난이도가 동적으로 결정되는 방식임(테스틀릿 기반 총 144개 경로). 이러한 적응형 설계는 개별 학생의 능력에 최적화된 문항을 제공하여 측정 오차를 최소화하고, 문항 수를 줄여 학생의 평가 부담을 완화하는 데 목적이 있음(OECD, 2023a).

### 2. 데이터 구조

○ PISA는 국가 전체를 대표할 수 있도록 설계된 복합 표본설계(Complex Survey Design) 방식을 따름(OECD, 2023a). 이에 따라 본 리포트에서는 분석과정에서 표본 설계의 특성을 반영한 가중치와 분산 추정 기법을 필수적으로 적용함.

- PISA 2003, 2012, 2022는 국제 표준에 따라 학교를 먼저 추출한 뒤, 그 학교 안에서 학생을 추출하는 방식을 통해 표본추출을 시행함.
- 1단계(학교): 전국의 학교를 지역, 설립 유형(공·사립), 도시화 정도 등에 따라 여러 층(strata)으로 나눔. 그 후, 확률 비례 추출(PPS) 방식으로 학생 수가 많은 큰 학교일수록 표본으로 뽑힐 확률을 높게 설정함.
  - 2단계(학생): 추출된 학교 내에서 만 15세 학생 명단을 작성, 무작위로 대상 학생을 선정함.

- 위와 같은 표본 추출은 어떤 학생은 뽑힐 확률이 높고 어떤 학생은 낮아지는 불균형이 발생할 수 있음. 이에 가중치를 활용하여 대표성을 확보함. 본 리포트는 산출된 최종 학생 가중치를 적용하여, 표본의 데이터가 전체 모집단의 특성을 편향 없이 대표하도록 보정하였음.
- PISA는 표본 추출 오차뿐만 아니라, 후술할 유의측정값(plausible value, PV) 사용에 따른 측정 오차까지 고려해야 함.
  - 따라서 균형 반복 복제방법(Balanced Repeated Replication, BRR)을 사용하여 오차를 추정함.
  - 이는 데이터를 여러 번 반복적으로 섞어서 분석해 보며 결과의 변동성을 계산하는 기법으로, 복잡한 표본 설계하에서도 통계적 유의성을 정확하게 검증할 수 있게 해줌(OECD, 2023a).

### 3. 성취도 변수의 특성

- PISA에서 사용하는 학업성취도 점수는 단일 점수가 아닌 유의측정값으로 산출된 변수임 (OECD, 2023a). 이에 따른 자료의 통계적 특성과 분석 시 고려사항은 다음과 같음.
- PISA는 학생이 모든 문항을 풀지 않는 행렬 표집 방식을 따르므로, 학생 개인의 능력을 하나의 점수로 단정할 수 없음. 따라서 학생의 문항 반응과 배경 변인을 결합하여 학생이 가질 수 있는 능력값의 확률 분포를 추정하는데, 이 분포에서 무작위로 추출한 값들을 유의측정값 (PV)라고 함(OECD, 2023a). 이는 측정 오차에 의한 분산의 과대 추정을 막고 모집단 특성을 정확히 추정하기 위해 필수적임.
  - PV는 고정된 값이 아니라 측정 오차를 포함한 가능한 점수들(plausible values)임. 따라서 연구자는 제공된 모든 PISA 2003과 2012는 5개, PISA 2022는 10개에 대해 각각 분석을 수행한 뒤 결과를 종합해야 함. 2022년 MSAT는 학생마다 거치는 경로가 달라 측정의 불확실성이 다를 수 있어, 측정 오차를 보다 정밀하게 포착하기 위해 기존 5개였던 PV를 10개로 확대 제공함(OECD, 2023a).
- 또한 본 리포트의 정의적 성취도 변수인 수학 불안감 및 수학 자아효능감은 학생 설문 문항에 대한 자기보고식 응답을 바탕으로 산출되었으며, 국가 간 비교를 위해 주기별 PISA 참여국 전체 학생의 평균이 0, 표준편차가 1이 되도록 표준화하여 OECD에서 제공하는 변수임 (OECD, 2024). 따라서 변수값이 0에 가까울수록 OECD 평균 수준에 근접한 것으로 해석할 수 있음.

### 4. 성취수준 구분 기준

- 집단 내 분포의 변화를 심층적으로 파악하기 위해 본 리포트는 수학 학업성취도를 PISA 성취 수준별로 세분화하여 분석함.
- PISA는 점수 구간에 따라 1~6수준으로 성취도를 구분하며, OECD는 2수준을 사회생활에 필요한 최소한의 능력을 갖춘 기초 성취수준(baseline level)으로 정의함(OECD, 2023b).

이에 따라 본 분석에서는 성취수준을 다음과 같이 세 집단으로 재범주화하여 살펴봄.<sup>2)</sup>

〈표 2-1〉 PISA 수학 학업성취도 성취수준 구분 기준

구분		최저점수
상위	6수준	669
	5수준	607
중위	4수준	545
	3수준	482
	2수준	420
하위	1수준	358

출처: OECD(2023b, p. 92)를 바탕으로 재구성.

2) PISA 2022는 1수준을 세분화하여 총 8수준으로 구분하나(PISA, 2023b), 본 분석에서는 이전 주기(2003, 2012) 와의 기준 통일을 위해 1수준을 세분화하지 않고 분석함.

### III. 분석결과

#### 1. 주기별 성별 수학 성취수준 분포

PISA 2012 대비 PISA 2022 상위권 비율은 급감하였고, 하위권 비율 급증함

〈표 3-1〉 PISA 2003, 2012, 2023 성별에 따른 수학 성취수준별 분포 추이

구분		빈도	비율
2003	여학생	전체	533,504
		전체	216,230
		상위	21,160
		중위	154,690
		하위	40,380
	남학생	전체	317,274
		상위	52,213
		중위	220,933
		하위	44,128
2012	여학생	전체	603,632
		전체	281,164
		상위	44,587
		중위	195,028
		하위	41,549
	남학생	전체	322,468
		상위	78,856
		중위	198,598
		하위	45,015
2022	여학생	전체	428,012
		전체	203,883
		상위	16,141
		중위	137,066
		하위	50,677
	남학생	전체	224,129
		상위	29,610
		중위	132,284
		하위	62,235

주: 가중치 적용

자료: OECD, 「PISA 2003, 2012, 2022」, 원자료 분석.

○ 우리나라 학생들의 주기별 성별 분포를 살펴본 결과는 다음과 같음.

- 분석 대상 모집단은 2003년 약 53만 명에서 2012년 약 60만 명으로 증가했으나, 2022년에는 약 42만 8천 명으로 감소함.
- 2003년에는 남학생 비율(59.5%)이 여학생(40.5%)보다 월등히 높았으나, 2022년에는 남학생(52.4%)과 여학생(47.6%)의 격차가 줄어듦.

- 다음으로 주기별 성별 및 성취수준별 분포를 살펴본 결과는 다음과 같음. 각 주기별 성별 중위 학생의 비율은 유사하게 나타났으나, 상위권과 하위권 분포에서 변동이 있었음.
- 2003년에서 2012년 사이, 남녀 학생 모두 상위권 비율이 대폭 상승함(여학생: 9.8% → 15.9%, 남학생: 16.5% → 24.5%).
  - 동일 기간 여학생의 하위권 비율은 18.7%에서 14.8%로 감소함.
  - 반면, 2022년에는 2012년 대비 상위권 비율이 급감함(여학생: 15.9% → 7.9%, 남학생: 24.5% → 13.2%).
  - 이와 대조적으로 하위권 비율은 급증하였으며, 여학생: 14.8%에서 24.9%로, 남학생은 14.0%에서 27.8%로 나타남.
  - 코로나19 팬데믹의 영향 및 어휘력·문해력 저하로 인한 기초학력 저하 현상을 고려해 볼 수 있음.

## 2. 수학 학업성취도 성별 차이

PISA 2003과 2012년에 통계적으로 유의했던 남학생의 성취도 우위가 PISA 2022에서는 사라져 전체 평균 점수에서의 성별 격차가 해소됨

- PISA 2003, 2012, 2022의 수학 학업성취도 성별 차이를 살펴봄. 우리나라 학생들의 수학 성취도 성별 격차는 2012년까지 유의미하게 유지되었으나, PISA 2022에서 통계적 유의성이 사라지는 경향을 보임.
- PISA 2003 여학생의 수학 학업성취도 평균점수는 528.31, 남학생은 561.71로 나타남. 두 집간 간 차이는 23.40점으로 통계적으로 유의한 차이( $p<0.001$ )인 것으로 나타남.
  - PISA 2012 여학생의 수학 학업성취도 평균점수는 544.19, 남학생은 562.11로 나타남. 두 집단 간 점수 차이가 17.92점으로 다소 완화되었으나, 여전히 남학생이 여학생보다 통계적으로 유의미하게 높은 수준( $p<0.05$ )을 유지함.
  - PISA 2022 여학생의 수학 학업성취도 평균점수는 524.64, 남학생은 529.72로 나타남. 두 집간 간 차이는 5.08점으로 통계적으로 유의한 차이가 없는 것으로 나타남.

〈표 3-2〉 PISA 2003, 2012, 2023 수학 학업성취도 성별 평균 및 차이

구분	여학생		남학생		차이(남-여)	
	평균점수	표준오차	평균점수	표준오차	평균점수	표준오차
2003	528.31	5.35	551.71	4.36	23.40***	6.90
2012	544.19	5.13	562.11	5.82	17.92*	7.76
2022	524.64	3.66	529.72	5.59	5.08	6.68

주: 가중치 적용.

자료: OECD, 「PISA 2003, 2012, 2022」, 원자료 분석.

### 3. 수학 학업성취도 성취수준별 성별 차이

상위권 남녀 점수 차이가 대폭 줄어들며, 모든 성취수준 구간에서 성별에 따른 유의미한 차이가 소멸함

- 전체 평균점수에서 나타난 성별 격차의 변화가 성취수준별로는 어떻게 나타나는지 구체적으로 살펴보기 위해, 성취수준별 남녀 학생 간 평균점수와 그 차이를 분석함.
- 분석 결과, 상위 성취수준 집단에서는 2003년과 2012년에 남학생이 여학생보다 통계적으로 유의하게 높은 성취도를 보였음. 2022년 상위 성취수준에서 남학생과 여학생 격차가 7.42점으로 줄어들며 통계적으로 유의하지 않은 것으로 나타남.
  - PISA 2003 상위권 남학생은 645.23점 여학생은 628.51점으로 차이점수는 16.72점이었으며, 이는 통계적으로 매우 유의한 차이( $p<0.001$ )를 보임.
  - PISA 2012 상위수준 남학생(658.11점)이 여학생(645.19점)보다 12.92점 높게 나타나 남학생 우위 경향이 지속됨( $p<0.01$ ).
  - PISA 2022에서는 남학생은 622.54점, 여학생 615.12점으로 점수 차이가 7.42점으로 줄어들었으며, 통계적으로 유의한 차이가 없는 것으로 나타나 상위권의 성별 격차가 해소되는 경향을 보임.
- 중위수준과 하위수준에서는 2003년 중위권을 제외하고는 지난 20년간 남녀 간 유의미한 성취도 차이가 발견되지 않음.
  - PISA 2003에서는 남학생이 5.28점 높게 나타나 유의미한 차이( $p<0.05$ )를 보였으나, PISA 2012(4.82점 차)와 PISA 2022(2.52점 차)에서는 모두 통계적으로 유의한 차이가 없는 것으로 나타남.
  - PISA 2003(-2.52점), 2012(-3.53점), 2022(-2.43점) 모든 주기에서 남녀 간 점수 차이는 오차 범위 내에 있어 통계적으로 유의하지 않음.
  - 2022년에는 모든 성취수준 구간에서 성별에 따른 점수 차이가 통계적으로 유의하지 않아, 수학 성취도에서의 성별 격차 완화 현상이 특정 수준에 국한되지 않고 전반적으로 나타났음을 확인할 수 있었음.

〈표 3-3〉 PISA 2003, 2012, 2023 성취수준별 수학 학업성취도 성별 평균 및 차이

구분	여학생		남학생		차이(남-여)	
	평균점수	표준오차	평균점수	표준오차	평균점수	표준오차
2003	상위	628.51	3.42	645.23	3.81	16.72***
	중위	530.14	2.15	535.42	2.28	5.28*
	하위	425.33	3.66	422.81	4.05	-2.52
2012	상위	645.19	2.89	658.11	3.12	12.92*
	중위	545.03	1.98	549.85	2.05	4.82
	하위	438.74	3.45	435.21	3.88	-3.53
2022	상위	615.12	3.74	622.54	3.92	7.42
	중위	526.31	2.56	528.83	2.71	2.52
	하위	410.55	4.12	408.12	4.45	-2.43

주: 가중치 적용.

자료: OECD, 「PISA 2003, 2012, 2022」, 원자료 분석.

#### 4. 수학의 정의적 특성(불안감, 자아효능감) 성별 차이

여학생이 남학생보다 수학 불안감이 높고 자아효능감이 낮은 정의적 특성의 성별 불균형은 20년간 지속됨

- PISA에서는 수학 학업성취도와 같은 인지적 결과뿐만 아니라, 학생들이 수학 학습 과정에서 느끼는 심리적·정서적 태도인 정의적 특성을 중요하게 다룸. 본 리포트에서는 PISA 2003, 2012, 2022에서 나타난 정의적 지표인 수학 불안감과 수학 자아효능감의 성별 차이를 분석 함.
- 수학 불안감 분석 결과, 여학생의 수학 불안감은 지난 20년간 남학생보다 일관되게 높은 수준을 유지함.
  - PISA 2003에서 여학생의 불안감은 0.35, 남학생은 -0.05로 차이값 0.40은 통계적으로 유의함 ( $p<0.001$ ).
  - PISA 2012에서 여학생은 0.42, 남학생 0.08 간 0.34 차이값을 보였으며, 차이는 통계적으로 유의함 ( $p<0.001$ ).
  - PISA 2022에서는 전체적인 불안감이 감소하는 추세를 보였으나, 여학생은 0.18로 OECD 평균보다 높은 불안을 보였음.
  - 반면 남학생은 -0.09로 OECD 평균보다 낮아, -0.27의 유의미한 성별 격차가 지속됨.
  - 이는 인지적 성취의 향상이 심리적 안정감으로 직결되지 못하고 있음을 시사함.
- 수학 자아효능감 분석 결과, 남학생은 모든 주기에서 OECD 평균 이상의 자신감을 보인 반면, 여학생은 평균 이하에 머물렀음.
  - PISA 2003에서 남학생(0.15)과 여학생(-0.31)의 자아효능감 격차는 0.46로 매우 컸으며, 통계적으로 유의함( $p<0.001$ ).
  - PISA 2012에서 여학생은 -0.25, 남학생은 0.15로 나타났으며, 남녀 차이는 0.40로 통계적으로 유의함 ( $p<0.001$ ).
  - PISA 2022에서 성취도의 성별 차이가 사라졌음에도 불구하고, 자아효능감에서는 남학생(0.12)이 여학생 (-0.15)보다 0.27포인트 높게 나타나 통계적으로 매우 유의한 차이( $p<0.001$ )를 유지함.

〈표 3-4〉 PISA 2003, 2012, 2023 수학 정의적 특성 성별 평균 및 차이

구분	여학생		남학생		차이(남-여)		
	평균값	표준오차	평균값	표준오차	평균값	표준오차	
수학 불안감	2003	0.35	0.02	-0.05	0.02	-0.40***	0.03
	2012	0.42	0.02	0.08	0.02	-0.34***	0.03
	2022	0.18	0.03	-0.09	0.03	-0.27***	0.04
수학	2003	-0.31	0.02	0.15	0.02	0.46***	0.03

자아효능감	2012	-0.25	0.02	0.15	0.02	0.40***	0.03
	2022	-0.15	0.03	0.12	0.03	0.27***	0.04

주: 1) 가중치 적용.

2) 수학 불안감은 점수가 높을수록 불안이 높음을, 자아효능감은 점수가 높을수록 자신감이 높음을 의미함.

3) 수학 불안감, 수학 자아효능감 변수는 평균이 0, 표준편차가 1이 되도록 표준화된 값임(OECD, 2024).

자료: OECD, 「PISA 2003, 2012, 2022」, 원자료 분석.

## 5. 성취수준별 수학의 정의적 특성(불안감, 자아효능감) 성별 차이

상위권 집단에서 오히려 남녀 간 정의적 특성(불안감, 효능감)의 격차가 가장 크게 벌어짐

- 정의적 특성의 성별 격차가 학생의 인지적 성취 수준에 따라 다르게 나타나는지, 그리고 지난 20년간 어떻게 변화했는지 파악하기 위해 PISA 2003, 2012, 2022의 성취수준별 수학 불안감 및 자아효능감을 살펴보았음.
- 수학 성취도가 가장 높은 상위권 집단에서 불안감, 효능감의 격차가 가장 크게 나타남.
  - 수학 불안감에서 2003년은 차이값이 -0.47, 2012년은 -0.43, 2022년은 -0.33로 나타남. 상위권 여학생은 남학생보다 훨씬 높은 불안감을 보임.
  - 2022년 상위권 남학생(-0.55)이 OECD 평균보다 상당히 높은 값을 보였지만, 여학생(-0.22)은 상대적으로 높은 불안을 느림. 이는 우수한 여학생들이 실력에 비해 높은 불안을 느끼고 있음을 보여줌.
  - 수학 자아효능감에서 2003년은 차이값이 0.53점, 2012년은 0.47점, 2022년은 0.37점으로 나타남.
  - 상위권에서 남녀 간 자아효능감 격차가 가장 컸음. 2022년 상위권 여학생의 효능감(0.48)은 높은 성취에도 불구하고 남학생(0.85)의 절반 수준으로 나타남.
- 중위권 학생들은 20년 동안 불안감과 자아효능감 모두에서 일관된 성별 격차를 보이고 있으나, 최근 들어 그 차이가 다소 완화되는 추세임.
  - 불안감은 2003년(-0.36)과 2012년(-0.30)에 비해 2022년(-0.23)에는 중위권 격차가 감소하는 추세를 보임. 2022년 중위권 여학생의 불안감(0.18)은 상위권 여학생(-0.22)보다 높아, 성취수준이 낮아질수록 불안이 증가함.
  - 자아효능감은 중위권 여학생이 2003년 -0.35, 2012년 -0.32로 나타나 OECD 평균보다 매우 낮은 효능감을 보였으며, 2022년은 -0.25로 소폭 상승함. 남학생(0.05)과의 격차(0.30)는 유의미하게( $p<0.001$ ) 유지됨.
- 하위권 학생들은 성별과 관계없이 전반적으로 높은 불안감과 낮은 효능감을 보이며, 상위권에 비해 상대적으로 성별 격차가 작게 나타남.
  - 하위권의 불안감 성별 격차는 2003년 -0.27, 2012년 -0.23, 2022년 -0.17로 나타남. 2022년 격차값은 상위권 격차(-0.33)의 절반 수준임. 이는 하위권에서는 성별 요인보다 낮은 학업 성취 자체가 불안의 주원인이기 때문으로 생각할 수 있음.
  - 자아효능감 성별 격차는 2003년 0.34, 2012년 0.23, 2022년 0.20로 나타남. 2022년 여학생은 -0.85, 남학생은 -0.65로 낮은 효능감을 보였으며, 격차(0.20) 또한 상위권(0.37)에 비해 작게 나타남.

- 2003년 대비 2022년의 결과를 비교할 때, 격차 값은 점진적으로 감소하는 추세를 보이고 있음.
  - 시계열에 따른 성취수준별 불안감 및 자아효능감의 격차는 모두 통계적으로 유의( $p<0.001$ )한 것으로 나타남.
  - 인지적 성취 격차가 2022년에 해소된 것과 다르게, 정의적 성취에서의 격차 완화 속도는 매우 더디며 구조적인 문제로 고착화되어 있음을 시사함.

〈표 3-5〉 PISA 2003, 2012, 2023 성취수준별 수학 정의적 특성 성별 평균 및 차이

구분	여학생		남학생		차이(남-여)	
	평균값	표준오차	평균값	표준오차	평균값	표준오차
2003	상위	불안감	0.05	0.03	-0.42	0.03
		효능감	0.12	0.03	0.65	0.03
	중위	불안감	0.38	0.02	0.02	-0.36***
		효능감	-0.35	0.02	0.05	0.40***
	하위	불안감	0.72	0.04	0.45	0.04
		효능감	-0.92	0.04	-0.58	0.04
2012	상위	불안감	0.08	0.03	-0.35	0.03
		효능감	0.25	0.03	0.72	0.03
	중위	불안감	0.45	0.02	0.15	0.02
		효능감	-0.32	0.02	0.08	0.02
	하위	불안감	0.85	0.04	0.62	0.04
		효능감	-0.88	0.04	-0.65	0.05
2022	상위	불안감	-0.22	0.04	-0.55	0.04
		효능감	0.48	0.04	0.85	0.04
	중위	불안감	0.18	0.03	-0.05	0.03
		효능감	-0.25	0.03	0.05	0.03
	하위	불안감	0.52	0.05	0.35	0.05
		효능감	-0.85	0.05	-0.65	0.05

주: 1) 가중치 적용.

2) 수학 불안감은 점수가 높을수록 불안이 높음을, 자아효능감은 점수가 높을수록 자신감이 높음을 의미함.

3) 수학 불안감, 수학 자아효능감 변수는 평균이 0, 표준편차가 1이 되도록 표준화된 값임(OECD, 2024).  
자료: OECD, 「PISA 2003, 2012, 2022」, 원자료 분석.

## IV. 요약 및 시사점

- 본 리포트는 PISA 2003, 2012, 2022년 자료를 바탕으로 우리나라 학생들의 수학 학업성취도 및 정의적 특성에 나타난 성별 격차를 시계열적으로 분석하였음. 우리나라 학생들의 인지적·정의적 성취 변화는 다음과 같음.
- 첫째, 수학 학업성취도에서 전통적으로 존재했던 남학생 우위 현상은 PISA 2022에 이르러 통계적 유의성이 사라짐으로써 성별 격차는 해소되었음. 이는 여학생의 성취 향상 및 상위권 남학생의 성취 하락과 더불어 남녀 모두에서 나타난 하위권 급증이 맞물린 결과임. PISA 2022에서 인지적 성취 측면에서의 성별 격차가 해소된 것으로 보이나, 전반적인 학력 저하의 경향이 나타남.
- 둘째, 인지적 성취의 균형에도 불구하고 수학에 대한 불안감과 자아효능감 등 정의적 특성에서의 성별 불균형은 구조적으로 고착화되어 있음. 여학생은 남학생보다 높은 수학 불안감과 낮은 자아효능감을 보고하였으며, 이러한 경향은 성취도 격차가 사라진 2022년에서도 유의미한 개선을 보이지 않았음. 이는 인지적 성취의 향상이 학생들의 내면적 태도나 자신감의 변화로 반영되지 않음을 시사함.
- 셋째, 상위권 여학생 집단에서 성취와 정서 간의 괴리가 가장 극명하게 나타났음. 객관적 성취도가 상위 수준임에도 불구하고, 동일 수준의 남학생 대비 자아효능감은 현저히 낮고 불안감은 높은 현상이 지속되고 있음이 확인되었음. 이는 우수 여학생들의 잠재력 발휘를 저해하는 심리적 장벽이 여전히 높음을 보여줌.
- 이상의 결과를 토대로, 다음과 같은 지원방안을 고려해 볼 수 있음.
  - PISA 2022에서 확인된 하위권 비율의 급증은 PISA 2022 직전의 코로나19 팬데믹의 영향과 디지털 과의존 현상이 복합적으로 작용한 결과로 생각할 수 있음. 팬데믹 기간 급증한 디지털 기기 노출은 학생들이 짧고 자극적인 콘텐츠에 익숙해지는 계기가 되었으며, 결과적으로 긴 호흡의 글을 이해하는 어휘력과 문해력의 저하를 초래함. 이러한 기초학력 미달 문제는 PISA 결과뿐만 아니라 교육 현장 전반의 공통된 문제임<sup>3)</sup>. 따라서 팬데믹 이후 고착화된 디지털 기기의 올바른 사용 교육과 어휘력 및 문해력 향상을 위한 깊이 있는 사고 훈련을 선행하여, 학습의 기초가 되는 문해력을 회복시키는 지원방안이 시급함.
  - 높은 성취에도 불구하고 낮은 효능감을 보이는 여학생들을 위해, 자신의 지능과 능력이 고정되어 있지 않고 노력을 통해 성장할 수 있다는 믿음인 성장 마인드셋(Growth Mindset)을 통한 피드백 및 평가체계가 필요함. 이를 통해 학생들이 “나는 수학을 잘하는 사람”이라는 결과 중심의 정체성 대신 “나는 어려운 수학 문제를 끈기 있게 파고드는 사람”의 과정 중심적 정체성을 강화함으로써, 학업 탄력적인 태도를 내면화하도록 지원이 필요함.
  - 이러한 과정 중심 평가 체계는 교사의 역할이 중요함. 교사는 정답 여부나 점수 등 결과에 초점을 맞춘 피드백을 지양하고, 학생이 문제를 해결하기 위해 시도한 전략과 논리적 사고 과정, 그리고 끈기 있는 태도

3) 서울특별시교육청 보도자료(2025.01.14). 2024 서울 학생 문해력·수리력 진단검사 시행 결과.

를 구체적으로 칭찬하고 평가해야 함. 이러한 과정 중심 피드백과 평가 방식의 변화는 학생들이 자신의 노력이 성장의 밑거름이 됨을 체감하게 하여, 학습에 대한 효능감을 높이는 데 기여할 것임.

○ 본 리포트는 다음과 같은 한계점이 있어 해석의 유의가 필요함.

- 제언에서 언급된 문해력 저하와 기초학력 미달 간의 연관성은 교육현장과 PISA 2022의 전반적인 경향성에 근거한 제언임. 본 리포트의 분석 과정에서 학생들의 디지털 기기 사용 시간이나 디지털 리터러시 수준을 변인으로 투입하여 수학 성취도와의 직접적인 상관관계를 실증적으로 검증한 것은 아니므로 해석의 확장에 유의가 필요함.
- 수학 불안감 및 자아효능감과 같은 정의적 특성 변수는 학생의 자기보고식 응답에 기초함. 이는 학생의 주관적 판단이나 응답 성향(예: 사회적 바람직성 등)의 영향을 받을 수 있음. 특히 성취도는 높으나 효능감이 낮은 현상은 이러한 사회·문화적 특성이 반영되었을 가능성이 있으므로 신중한 해석이 필요함.
- 본 리포트는 성별과 성취수준을 주요 변수로 설정하여 분석하였으나, 인지적·정의적 성취에 영향을 미칠 수 있는 부모의 사회·경제적 지위(ESCS), 부모의 양육 태도 등 PISA에서 제공하는 다양한 배경 변수들을 통제하지 못하였음. 향후 연구에서는 이러한 변수들을 포함한 분석을 통해 성별 격차의 원인을 정밀하게 규명할 필요가 있음.

## 참고문헌

- 송미영, 임해미, 최혁준, 박혜영, 손수경(2013). OECD 국제 학업성취도 평가 연구: PISA 2012 결과 보고서. 한국교육과정평가원 연구보고 RRE 2013-6-1.
- 김성경, 김명화, 김인숙, 이신영, 백혜선(2023). OECD 국제 학업성취도 평가 연구: PISA 2022 결과 보고서. 한국교육과정평가원 연구보고 RRE 2023-10.
- OECD. (2023a). *PISA 2022 assessment and analytical framework*. OECD Publishing. <https://doi.org/10.1787/dfe0bf9c-en>
- OECD. (2023b). *PISA 2022 results (Volume I): The state of learning and equity in education*. OECD Publishing. <https://doi.org/10.1787/53f23881-en>
- OECD. (2024). *PISA 2022 technical report*. OECD Publishing. <https://doi.org/10.1787/01820d6d-en>