

## 인지진단모형을 적용한 TIMSS 8학년 수학 기하 영역의 성차 분석<sup>1)</sup>

이 현 숙\* · 고 호 경\*\*

수학에서의 성차는 교수·학습 환경에서 학습자에 대한 공평성을 추구하는 맥락에서 연구자들의 지속적인 관심을 받아 왔다. 수학의 여러 영역 중 특히 기하 영역은 전통적으로 남학생이 여학생에 비해 높은 성취를 보이는 영역으로 인식되어 왔으나, 최근에는 성차가 완화되거나 점차 사라지고 있다는 경험적 증거들이 종종 보고되고 있다. 본 연구에서는 2003년부터 2011년까지 3개 주기 동안 실시된 TIMSS 8학년 수학과 데이터를 활용하여 우리나라 중학생들이 기하 영역의 각 하위 인지요소에서 나타내는 성차를 인지진단모형을 활용하여 고찰하였다. 본 연구에서는 여러 가지 인지진단모형 중 교육 전문가에게 유용하고 해석 가능한 정보를 제공할 수 있는 Fusion 모형을 적용하였다. 연구 결과, 기하 영역의 세부 인지요소 중 ‘입체도형의 모양’에 있어서는 2003년과 2007년 각각 남학생이 여학생에 비해 높은 숙달 확률을 나타내었으나, 2011년에는 전체 인지요소에서 남녀 간에 차이가 없는 것으로 나타나, 성차가 완화되고 있다는 최근 연구들을 지지하는 하나의 경험적 증거를 제공하였다. 이밖에 인지진단모형을 적용하여 성차를 분석한 결과에 따라, 학생들의 인지요소 숙달 프로파일이 남녀 간에 어떠한 차이를 보이는지, 그리고 특정 문항을 푸는데 있어서 반드시 필요하다고 정의된 인지요소들이 성별에 따라 상대적으로 더 혹은 덜 중요하게 기능하는지 등에 대해 고찰하고 이에 근거하여 기하 영역의 교수·학습에 시사점을 제공하였다.

### 1. 서론

교수·학습 환경에서 학습자에 대한 공평성 문제는 오랜 기간 동안 수학교육 분야의 주요 관심 대상이었으며, 이에 대한 일환으로 성차를 줄이기 위한 수학교육 프로그램의 개발이나 교육환경의 변화 등 성공평성 추구는 지속적인 연구의 대상이 되어 왔다(예, 이대식·김수미, 003; 정경아, 2005; 이지혜, 2008; 최하나, 2007;

Cronin, 2005; Honigsfeld & Dunn, 2003). 성공평성을 추구하기 위해서는 성차가 어느 부분에서 나타나며 어떠한 요인에 기인하는지를 밝히는 것이 우선되어야 하므로 수학의 인지적·정의적 측면에서 성차의 존재에 대한 규명 및 성차 발생의 원인을 분석하기 위한 연구들이 그동안 꾸준히 진행되어 왔다(Bennet, 1996; Burton, 1986; Chipman, Brush & Wilson, 1985; Fennema, 1985; Fennema & Leder, 1990; Furnham, Clark & Bailey, 1999; Funham & Fong, 2000; Lynn,

\* 건국대학교, hyunysi@konkuk.ac.kr (제1 저자)

\*\* 아주대학교, kohoh@ajou.ac.kr (교신저자)

1) 이 논문은 2013학년도 건국대학교의 연구년교원 지원에 의하여 연구되었음

Irwing, & Cammocck, 2001). 성차 연구의 결과는 다양한 형태로 나타났는데, 예를 들어 Cronin (2005)과 Summers(2005)의 연구와 같이 성차를 의미있게 보고한 연구가 있는가 하면, 수학 교과에서 성차가 미미하거나 수학 불안과 같은 영역에서는 오히려 남학생이 더 높게 나타난다는 연구(예, Betz, 1997; Meece, 1981), 혹은 학년에 따라 성차의 양상이 다르게 나타난다는 연구 등 (Meece & Wigfield, 1988) 다양한 결론을 보이기도 한다.

성차의 발현 형태는 시대의 변화에 따라 조금씩 다른 형태로 나타나기도 하는데, 과거의 연구에서 나타난 성차가 근래에 들어 완화되거나 없어지기도 하였고(예, 박정 외, 2004; Ding, Song, & Richardson, 2007), 혹은 역전되는 경향도 관측되고 있다(이연옥, 1989; 이향란, 1991). 그러나 여전히 중·고등학생들의 수학 교수·학습과 관련하여 사람들이 초미의 관심을 보이는 부분은 수학의 세부적인 영역에서 성차의 존재 여부에 관한 것이다(예, 박효주, 2009; 서경숙, 2007; Armstrong, 1981; Bellisari, 1989; Finn, 1980; Royster, Hennis & Schoeps, 1999; Utsumi & Mendes, 2000). 이는 성차가 수학의 어느 영역에서 어느 정도의 강도로 나타나는지 파악하고 그에 대한 원인을 수학 내적 요인과 관련하여 추측함으로써, 성차를 배려한 구체적인 교수·학습 환경을 제공하기 위한 적극적인 시도를 추구하고 있기 때문이다(예, 고호경·고상숙, 2008; Davidson & Schofield, 2002; Honigsfeld & Dunn, 2003).

기존에 실시된 성차 연구에서는 주로 전체 총점 또는 영역별 총점에 기반하여 남녀 간의 차이가 있는지에 주목하고 있으나, 총점은 문항별 정답여부를 바탕으로 합산한 점수이기 때문에 문항 각각에 대한 개별적인 반응을 고려한 구체적인 정보를 반영하지 않는다는 점에서 한계를

가진다(Hartz & Roussos, 2008). 만일 개별 문항에 반응한 피험자의 응답 패턴을 그대로 반영하여 능력 정보를 산출하고, 각 문항이 측정하고 있는 지식과 기능들을 고려하여 피험자의 강점과 약점에 대한 구체적인 분석이 이루어진다면, 총점에 기반하여 단순히 성차의 존재 여부를 확인하는 것을 넘어서 성차에 대한 보다 정밀한 형태의 진단 정보를 제공할 수 있을 것이다. 인지진단모형(Cognitive Diagnostic Modeling)은 이와 같은 정보를 제공할 수 있는 측정학적 모형으로, 문항의 정답 확률을 피험자의 인지 숙달 패턴(cognitive mastery pattern)의 함수로 설명함으로써 피험자의 점수가 높거나 낮은 원인을 검사가 측정하고자 하는 잠재적 속성(latent trait)에 기반하여 설명할 수 있다는 장점을 가진다(Henson & Douglas, 2005; Henson, Templin, & Willse, 2009). 따라서 인지진단모형을 통해 성차를 분석하면 학생들의 인지요소 숙달 프로파일의 남녀 간에 어떠한 차이를 보이는지, 그리고 특정 문항을 푸는데 있어서 반드시 필요하다고 정의된 인지요소 각각이 성별에 따라 상대적으로 더 혹은 덜 중요하게 기능하는지 등에 관한 정보를 산출할 수 있어 성차에 대한 풍부하고 구체적인 고찰이 가능하다.

수학 교과에 검사가 측정하고자 하는 인지요소들이 다른 과목에 비해 비교적 명확하게 정의될 수 있다는 점에서 인지진단모형의 활용도가 다른 교과에 비해 높은 편이며, 최근 국내외에서 평가의 진단적(diagnostic)·형성적(formative) 기능이 강조되고 있는 흐름과 맥을 같이하여 국가수준 학업성취도평가나 TIMSS와 같은 대규모 성취도 데이터에 인지진단모형을 적용하는 사례가 증가하고 있다(예: 김희경, 김부미, 2011; 김희경, 한정아, 최숙기, 김부미, 2012; Im & Park, 2010; Lee, Park, & Taylan, 2011; Tatsuoka, Corter, & Tatsuoka, 2004). 예를 들어, Tatsuoka, Corter,

Tatsuoka(2004)는 TIMSS-R 8학년 수학과 데이터에 규칙장 모형(Rule Space Model; Tatsuoka, 1985)을 적용하여 내용, 과정, 문항 형식 등을 포괄한 23개 인지요소에 대해 20개국의 인지숙달 프로파일을 비교하는 연구를 수행하였고, Im과 Park(2010)은 TIMSS 2003 8학년 수학과 데이터에 규칙장 모형을 적용하여 한국과 미국 학생들이 5개의 내용 영역과 5개의 행동 영역 인지요소에서 수행의 차이를 보이는지 고찰하고, 인지요소 숙달 확률(mastery probability)을 예측하는 교수·학습 관련 변수들을 탐색하였다. 또한 Lee, Park, Taylan(2011)은 TIMSS 2007 4학년 수학과 데이터를 사용하여 미국 전체 표본과 매사추세츠주와 미네소타주 표본에 DINA 모형(Junker & Sijtsma, 2001)을 적용하여 인지요소 숙달 확률과 인지요소 숙달 패턴 등을 비교하는 연구를 수행하였다. TIMSS 데이터에 인지진단모형을 적용하여 성차를 분석하고자 하는 연구도 수행되고 있다. Chen, Ferron, Thompson, Gorin, Tatsuoka(2010)는 타이완 학생들의 TIMSS 1999년 수학과 데이터에 규칙장 모형을 적용하여 성차를 비교하는데 있어서 기존의 문항특성곡선과 유사한 속성특성곡선(Attribute Characteristic Curve)을 이용한 사례를 제시하였고, Johnson 외(2013)는 DINA 모형을 다집단 상황으로 확장한 다집단 DINA(Multi-group DINA) 모형을 제안하고 미국 학생들의 TIMSS 2007 8학년 수학과 공개문항 전체에 대해 성차를 비교한 바 있다.

이와 같이 인지진단모형을 적용하여 성차를 분석한 연구들은 총점을 통해 분석한 성차에 비하여 성차에 관한 다양한 스펙트럼을 제시한다는 측면에서 의의를 가지며, 수학과 내용 영역 전체를 포괄하는 성차 분석을 수행함에 따라 대영역별 성차에 대한 큰 그림을 제시하는 역할을 하였으나, 세부적인 속성에 대한 성차를 구체적으로 제시하고 있지는 않다. 따라서 정밀한 진단

적 정보 제공이 가능한 인지진단모형의 장점을 극대화하기 위하여 수학의 각 하위 영역별로 보다 정밀한 차원에서 성차에 대한 고찰을 할 필요가 있다. 특히 기하 영역은 전통적으로 남학생이 우세하다고 알려진 분야지만 최근에는 기하 영역의 성차가 좁혀지고 있다는 경험적 연구들이 발표됨에 따라(예: 고상숙·고호경, 2006), 기하 영역을 구성하는 세부적인 인지적 특성 중 어느 부분에서 성차가 여전히 존재하고 어느 부분에서 성차가 좁혀지고 있는지를 구체적으로 밝히는 것은 수학 교육에서 많은 시사점을 줄 것으로 판단된다.

따라서 본 연구에서는 TIMSS 8학년 수학과 데이터를 활용하여 우리나라 중학생들이 기하 영역의 각 하위 인지요소에서 어떠한 형태로 성별 차이를 나타내는지에 관해 인지진단모형의 관점에서 고찰하였다. 이를 위해, 본 연구에서는 2003년부터 2011년까지 3개 주기 동안 실시된 TIMSS 수학과 문항 중 기하 영역 공개 문항을 사용하였으며, 여러 가지 인지진단모형 가운데 교육 전문가에게 유용하고 해석 가능한 정보를 제공함에 따라 최근 활용도가 높은 Fusion 모형(Hartz, 2002)을 적용하여 분석하였다. 본 연구에서 구체적으로 고찰하고자 하는 분석 내용은 다음과 같다. 첫째, TIMSS 8학년 수학과 기하 영역의 각 인지요소별 숙달 확률 및 숙달 유형에 대한 성차는 연도별로 어떠한 경향으로 변화하고 있는가? 둘째, TIMSS의 각 문항 중 Fusion 모형을 적용하여 산출된 문항 모수의 성별 차이가 가장 큰 문항은 무엇인가? 본 연구는 성차 연구에 인지진단모형을 적용한 하나의 방법론적 사례를 제공하며, 더불어 기하 영역의 성차를 잠재적인 속성의 관점에서 대규모 데이터를 통해 확인한 하나의 경험적 증거가 될 것으로 기대된다.

## II. 이론적 배경

### 1. 기하영역 성차

수학성취도에서 성차가 존재하는가 하는 질문은 성차 연구에서 주된 논의의 대상이 되어왔다. 그동안 수학이 교육과정 및 평가와 선발에서 차지하는 비중이 높은 과목으로서 주목 받아온 만큼, 남학생이 여학생보다 우수한 수학 학업성취도를 나타내고 있다는 성차의 존재성을 다룬 초기 연구 결과들에 의거해 후속 연구들이 활기를 띠어 왔다(이연옥, 1989). 비교적 최근 연구들 중에서는 성차가 좁혀지고 있다는 증거들이 속속 제시되었는데, 예를 들어 윤희선(2000), 강인수(2000)에 따르면 남학생이 여학생보다 수학 성취도가 다소 높으나 유의미한 성별 차이가 나타나지 않다고 제시하였으며, 해외 연구에서 역시 수학 관련 학업성취도에서 남녀학생들의 차이가 줄어들고 있음을 보고하고 있는 예들을 찾아볼 수 있다(예, Friedman, 1989, Halpern, 2000, Eccles, Lord, Roeser, Barber, & Jozefowicz, 1997).

한편, 기하 영역에서의 성차는 꾸준히 보고되고 있는데, 예를 들어 Benbow와 Stanley(1980, 1982)는 기하 영역에서 성차가 가장 크게 나타나고 있다고 보고 하였으며, Battista(1990)의 연구에서는 기하 관련 문제해결 전략과 논리 추론적 능력의 경우 성차가 크게 나타나지 않았으나 시각적 공간 표상에서는 그 차이가 유의미하게 나타난다고 하였다. 한편 기하 영역에서 성차가 학년에 따라 다르게 나타난다는 보고도 있는데, 예를 들어 Leahey와 Guo(2001)는 추론과 기하영역에서 초등학교 때는 성차가 거의 나타나지 않다가 12학년 무렵에 이르러 남학생이 여학생보다 좀 더 높은 성취를 보인다고 하였다.

기하에서의 성차 관련 국내 연구를 살펴보면, 권오남과 박경미(1995)가 공간화 능력을 필요로

하는 기하 영역에서 남학생들의 성취도가 여학생들의 성취도보다 높은 것으로 나타났다고 하였으며, 인문계 고등학교 학생들을 대상으로 대수와 기하의 성취도에서의 성차를 조사한 김진만(2002)의 연구에서는 대수보다는 기하 영역에서 그 차이가 더욱 두드러지게 나타나고 있다고 보고하였다. 또한 류신열(1998)은 해석영역이나 추론에서는 확연히 나타나지 않던 성차가 기하영역에서는 남학생이 우수하였다는 연구 결과를 제시하였다.

성차의 원인에 관심을 기울인 연구들도 있는데, 성차가 남녀 고유의 생물학적 본질에서 온 것이라는 주장(예, Baron-Cohen 2003; Bellisari, 1989; Benbow, 1988; Pinker, 2002)이 있는 반면, 기하 영역의 성차에 대한 생물학적 원인을 보여주는 증거가 불충분함을 들어(Fennema, 1985) 교육의 환경적 요인에 연유한다는 연구들도 제시되고 있다. 예를 들어 Armstrong(1981)은 남학생들의 경우 어린 시절부터 블록 맞추기와 같은 놀이에 더 자주 접함으로써 여학생들보다는 공간화 능력의 발달에 유리했을 것이란 주장을 펼치며, 생물학적 차이보다는 사회화에 기인한다고 보고 있으며, 권오남(2000)의 연구에서도 후천적 변인들을 설명력 있는 요인으로 보고 있다.

이와 같이 기하 영역의 성차에 주목한 연구들은 기하 성취도의 성차 여부를 밝히는 것 외에도 기하 교육에 있어서 성차를 고려한 수학 교수·학습이나 환경 조성에 관심을 가짐으로써 성차를 좁히기 위한 방안을 고려한다는 점에서 의의를 가진다(홍경아, 1994; Battista, 1990).

### 2. Fusion 모형

Fusion 모형(Hartz, 2002)은 DiBello, Stout, Roussos(1995)가 개발한 Unified 모형의 모수들을 추정 가능한 형태로 재모수화한 것으로 RUM

(Reparameterized Unified Model) 모형이라고도 불린다. Fusion 모형은 다차원적 공간에서 정의된 인지요소에 기반한 피험자의 능력 모수와 문항 모수에 의해 피험자가 문항에 정답할 확률이 결정된다는 문항반응이론의 기본 가정에 기초하고 있으며, 문항과 인지요소 간의 관계를 정의한 Q행렬을 정답확률 산출시 반영함으로써 각 문항에 고유하게 정의된 인지요소별로 문항모수가 산출되도록 한다. 이 때, 어떤 문항에 정답을 하기 위해서는 해당 문항에서 요구하는 모든 인지요소를 숙달해야 한다는 결합적(conjunctive) 가정과 문항에서 요구하는 인지요소들 중 하나라도 숙달하지 못할 경우 정답을 할 수 없다는 비보상적(non-compensatory) 가정에 기초한다. Fusion 모형이 다른 인지진단모형에 비해 가지는 특징은 인지요소가 피험자의 수행을 변별하는 정도를 문항마다 다르게 추정하도록 허용한다는 점과, Q행렬 상에서 정의되지 않았지만 문항을 푸는데 중요한 역할을 하는 인지요소가 존재할 수 있다는 가정을 모형에서 명시적으로 반영한다는 점이다.

Fusion 모형에서 한 피험자 어떤 문항에 정답을 할 확률은  $\pi_i^*$ ,  $r_{ik}^*$ ,  $c_i$  등의 모수에 의해 결정된다.  $\pi_i^*$ 는 문항  $i$ 를 푸는데 필요하다고 가정된 인지요소를 모두 숙달한 학생이 자신이 숙달한 인지요소를 정확히 사용하여 문항  $i$ 의 정답을 맞힐 확률, 즉 Q행렬에 기반한 조건부 정답 확률을 의미하며, 1에 가까울수록 Q행렬에 명시된 인지요소들이 피험자의 수행을 잘 예측하고 있다고 할 수 있다.  $r_{ik}^*$ 는 인지요소  $k$ 의 숙달여부에 따른 정답 확률의 비, 즉 문항  $i$ 에 대해 고유하게 추정된 인지요소  $k$ 의 변별도를 의미한다.  $r_{ik}^*$  값이 1에 가까우면 인지요소  $k$ 의 숙달여부에 관계없이 정답 확률이 유사하다는 의미가 되므로 인지요소  $k$ 가 문항을 푸는데 크게 기여

하지 않는다고 해석할 수 있고, 반대로 0에 가까우면 인지요소  $k$ 의 숙달여부에 따라 정답 확률이 큰 차이를 보인다고 할 수 있다.  $c_i$ 는 문항이 Q행렬에 명시되지 않은 능력에 어느 정도 의존하고 있는지를 나타내며, 0에서 3 사이의 값으로 추정된다.  $c_i$ 가 0에 가까울수록 문항  $i$ 의 풀이에 요구된다고 연구자가 설정한 인지요소 외에 다른 인지요소가 존재할 가능성이 크다는 의미이므로, Q행렬의 타당성을 평가하기 위한 진단적 정보로서 중요한 기능을 한다(Hartz & Roussos, 2008). Fu(2005)는 이와 같이 Harz(2002)가 정의한 Fusion 모형을 다분 문항에 적용 가능하도록 확장하였다.

### III. 연구 방법

#### 1. 연구 자료 및 대상

본 연구에서는 기하 영역에서 우리나라 중학생들의 성차 변화 추이를 탐색하기 위하여 2003년, 2007년, 2011년의 3개 주기에 실시된 TIMSS 데이터를 이용하였다. 2003년부터 TIMSS는 각 주기별로 실시되는 총 14개의 문항 블록 중 6개 블록에 해당하는 문항을 일반에 공개하고 있는데, 본 연구에서는 공개 문항 중 문항이 측정하는 대표적 기능이 기하 영역으로 분류된 문항과, 대수나 측정 영역으로 분류되었지만 기하 영역의 인지요소가 반드시 필요한 문항들을 포함하여 분석 문항을 선정하였다. 연도간의 비교가 타당하기 위해서는 비교대상이 되는 연도 간에 문항의 동등성이 확보해야 한다. 본 연구에서는 연도별로 동일한 검사 명세화표를 바탕으로 개발된 TIMSS 문항을 사용하였고 인지진단모형은 기존의 총점 기반 분석 방법과 달리 잠재 변수인 인지요소 숙달 확률을 바탕으로 성차의 변화

<표 III-2> TIMSS 8학년 기하영역 인지요소 정의

인지요소	정 의
1. 도형의 기본 성질 (Basic Geometric Properties)	- 예각, 직각, 둔각의 성질, - 선분, 직선, 평행선, 각의 이등분선, 수선, 중점의 성질, - 맞꼭지각, 엇각, 동위각의 성질 등을 알고 문제 해결에 적용하는 것
2. 도형의 성질 (Geometric Properties)	- 평면 도형의 모양, - 이등변삼각형, 직각삼각형, 정삼각형, 평행사변형, 직사각형, 정사각형의 속성, - 이등변삼각형의 성질, 삼각형의 내각 및 외각의 성질, 피타고라스 정리 등을 알고 문제 해결에 적용하는 것
3. 입체 도형의 모양 (Three-dimensional Geometric Shapes)	- 입체도형의 모양, - 평면도형과 입체 도형과의 관계 등을 알고 문제 해결에 적용하는 것
4. 기하적 측정 (Geometric Measurement)	- 변의 길이, 도형의 둘레와 넓이, 부피 등을 구하는 것 - 컴퍼스, 각도기, 시계 등의 도구를 활용하여 각의 크기를 어렵거나 측정하는 것
5. 도형의 변환 (Geometric Transformations and Movement)	- 평면 도형의 기하학적 변환(회전변환, 점대칭변환, 선대칭변환, 평행이동) 등을 인지, - 평면 도형의 합동과 닮음을 인지하고 문제 해결에 적용하는 것
6. 대수를 활용한 기하 (Geometry Combined with Algebra)	- 변수나 미지수를 포함한 기하 문제, - 도형의 성질에 대한 일반화, - 좌표평면 등을 활용한 기하 문제를 해결하는 것

추이를 보는 것이므로 개별 문항의 특수성에 영향을 덜 받음에 따라 연도간 동등성은 확보되었다고 보았다. 한편, TIMSS는 각 학생들마다 일부 문항만 치르는 행렬 표집(matrix sampling)을 실시하고 있기 때문에 타당한 모수 추정을 위해서는 서로 다른 검사지를 치르는 학생들을 연결해주는 공통문항이 존재해야 한다. 따라서 공개된 문항 중 동일 척도 상에서 비교 가능한 연결된 하위집단(joint subset)에 실시된 문항만을 사용하여 분석하였다. 아울러, 문항수가 많을수록 더욱 신뢰로운 결과를 산출할 수 있기 때문에, 2007년과 2003년에 실시된 검사의 경우 해당 연도에는 공개되지 않았지만 후속 연도에 공개된 문항도 포함하여 분석하였다. 2003년의 경우 변별도가 매우 낮은 1개 문항은 삭제하여 총 50문항으로 분석하였다. <표 III-1>은 본 연구에서 사용한 TIMSS 문항의 수와 사례수를 정리한 것이다.

<표 III-1> 본 연구에 사용된 문항수 및 사례수

실시 연도	분석에 사용된 문항 블록	문항수	사례수		
			여	남	전체
2003년	01, 02, 03, 04, 05, 06, 07, 08, 09, 10, 11, 12, 13	50문항	2,533	2,776	5,309
2007년	01, 02, 03, 04, 05, 06, 07	28문항	992	1,124	2,116
2011년	05, 06, 07	14문항	743	722	1,465

## 2. 인지요소 추출 및 Q행렬 타당화

인지진단모형을 적용하기 위해서는 검사가 측정하고자 하는 구인의 정의를 바탕으로 핵심 인지요소들을 추출하고, 이를 바탕으로 각 문항이 어떤 인지요소를 측정하는지를 규정하는 Q행렬

이 제작되어야 한다. 이를 위해 본 연구에서는 측정평가 전공 교수 1인과 수학교육 전공 교수 1인, 중학교 2학년 수학을 가르친 경험이 10년 이상인 수학 교사 2인을 전문가 집단으로 구성하여 인지요소 추출 및 Q행렬 제작 절차를 진행하였다.

TIMSS 기하 영역을 구성하는 핵심 인지요소를 추출하는 과정은 TIMSS 수학과 평가들에 제시된 세부 목표를 기본으로 하되, 각 문항이 측정하고 있는 인지적 속성들을 모두 포괄하면서도 가능한 한 구체적인 진단적 정보를 제공할 수 있는 방식으로 추출하였다. 인지진단모형에서는 검사가 측정하고자 하는 지식과 기술, 능력을 통칭하여 인지요소(attribute)로 정의하고 있으며, 본 연구에서는 기하 영역의 내용적 측면을 구성하는 하위 요소에 초점을 두어 인지요소를 정의하였다. TIMSS 기하 영역은 도형, 기하학적 측정, 위치와 이동의 세 가지 하위 영역으로 구성되어 있고, 각 하위 영역별로 각각 5개, 2개, 2개씩 총 9개의 세부 평가목표가 정의되어 있다 (International Association for the Evaluation of Educational Achievement, 2011). 총 9개의 세부 평가목표 중 공개 문항에서 다루어지지 않은 세부 평가목표는 삭제하거나 유사한 세부 목표와 통합하는 방식으로 각각 4개, 5개, 6개의 인지요소 후보군을 정의하고, 피험자에게 제공되는 피드백의 구체성 및 유용성을 고려하여 6개 인지요소가 가장 적합하다고 판단하였다. 최종적으로 추출된 6개의 인지요소 및 정의는 <표 III-2>에 제시하였다.

이와 같이 추출된 6개의 인지요소를 바탕으로 각 문항이 어떠한 인지요소를 측정하고 있는지 연결하는 작업이 전문가 집단의 반복적인 협의 과정을 통해 진행되었다. 전문가 집단은 각 문항을 직접 풀면서 문항에서 요구되는 인지요소를 선택하는 작업을 독립적으로 진행한 다음, 불일

치되는 부분에 대해 논의하는 방식으로 의견을 수렴해나가는 절차를 세 차례 반복하여 Q행렬을 완성하였다. 완성된 Q행렬의 타당성은 Jaccard 유사성 계수(Jaccard, 1901)와 Fusion 모형을 적용해서 산출된 문항 모수의 적절성(Hartz & Roussos, 2008)을 고려하여 판단하였다.

Jaccard 유사성 계수는 인지요소의 중복성을 판단하기 위한 지수로, 각 인지요소를 두 개씩 조합하여 두 인지요소 중 하나가 선택된 문항 중에서 두 인지요소가 공통으로 선택된 문항의 비율로 판단한다. Park(2008)이 적용한 유사성 판단의 기준(.5이상)에 근거할 때, 본 연구에서 설정된 인지요소 모두 중복성 없이 고유하게 정의된 것으로 나타났다. Fusion 모형에서 산출된 문항 모수 중 조건부 정답확률인  $\pi_i^*$ 는 총 92문항 중 6문항을 제외하고는 .9이상으로 산출되어 본 연구에서 정의된 인지요소가 피험자의 수행을 대체로 타당하게 예측하고 있다고 판단하였다. 한편, 각 문항에 대해 인지요소별로 산출되는  $r_{ik}^*$ 가 높게 나타난 경우 해당 인지요소가 문항을 푸는데 있어서 기여도가 낮은 것으로 볼 수 있는데,  $r_{ik}^*$ 가 .9 이상인 문항에 대한 재검토 결과, 전체 집단에서  $r_{ik}^*$ 가 높게 나온 것이 아니라 특정 성별 집단에서 높게 나온 것이기 때문에 Q행렬의 오류보다는 해당 인지요소에 대한 성차의 증거로서 해석하는 것이 적합하다고 판단하였다.

### 3. 인지진단모형 모수 추정 및 성차 분석

본 연구에서는 다분문항으로 확장된 Fusion 모형(Fu, 2005)을 적용하여 피험자 모수 및 문항 모수들을 추정하였으며, 추정에는 Arpeggio 3.1(Dibello & Stout, 2010) 프로그램을 사용하였다. Fusion 모형에서는 Q행렬에 정의되지 않은 인지요소의 존재 가능성을 모형화하는 것이 가

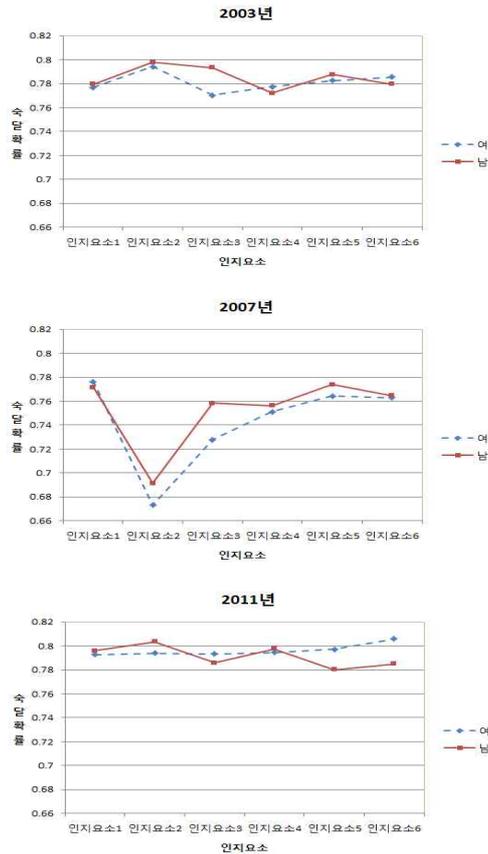
능하기 때문에, c모수를 자유모수로 설정하여 분석하였다. 인지요소 숙달 프로파일 산출을 위한 피험자 분류 기준은 숙달 확률이 .4이하인 경우 미숙달, .6이상인 경우 숙달로 판정하였으며, 숙달 확률이 .4와 .6 사이인 경우 숙달 여부가 불확실한 사례로 간주하였다. Fusion 모형 적용 결과에 따른 성차 분석은 인지요소별 숙달 확률 및 숙달 패턴의 차이, Fusion 모형에서 산출된 각 문항모수의 성별 차이 등을 주로 비교하였고, 인지요소 숙달 확률이 뚜렷이 나타난 인지요소에 대해서는 인지요소 특성곡선(Attribute Characteristic Curve)의 성별 차이를 추가로 탐색하였다.

#### IV. 연구 결과

##### 1. 인지요소 숙달 확률의 성별 차이

여학생과 남학생의 인지요소별 숙달 확률의 평균과 표준편차는 <표 IV-1>에 제시하였으며, 인지요소별 숙달 확률의 평균은 [그림 IV-1]에 시각적으로 제시하였다. 각 연도별 평균을 살펴 보면, 전반적으로 각 인지요소의 숙달 확률이 .7에서 .8 사이로 나타났으나, 2007년의 경우는 인지요소 2의 숙달 확률이 여학생과 남학생 모두 다소 낮게 나타났다. 남녀간의 인지요소 숙달 확률 차이의 전반적인 경향을 보면, 먼저 2003년에는 인지요소 3에서 남학생의 평균이 뚜렷이 높게 나타났고, 나머지 인지요소들에 있어서도 인지요소 4와 6을 제외하고는 남학생의 평균이 높았으나 큰 차이를 보이지 않았다. 2007년의 경우 전반적으로 남학생의 평균이 높았으며, 다른 인지요소에 비하여 인지요소 2와 3에서 상대적으로 평균 차이가 크게 나타났다. 반면, 2011년의 경우 인지요소 3, 5, 6은 여학생이 높았고, 나머지 인지요소에서는 남학생이 높았으나 남녀 간

의 차이가 크지는 않았다.



[그림 IV-1] 인지요소 숙달 확률의 성별 분포

인지요소 숙달 확률의 성차에 대한 통계적 유의미성을 검증하기 위해 MANOVA 분석을 실시한 결과는 <표 IV-2>에 제시하였으며, 유의수준 .05에서 통계적 유의성을 가지는 인지요소는 굵은 글씨로 표기하였다. MANOVA의 다변량 검정 결과는 3개년도 모두 Wilks Lambda가 유의수준 .001에서 유의미하여, 성별에 따라 인지요소 숙달 확률에 있어서 차이가 있는 것으로 나타났다. 성별에 따른 단변량 평균 차이 검증을 수행한 결과, 2003년과 2007년에는 인지요소 3에서 남학생의 숙달 확률이 통계적으로 유의미하게 높은 것으로 나타났으며, 2011년에는 모든 인지요소에

<표 IV-1> 인지요소 숙달 확률에 대한 기술통계치

연도	성별	N	평균 (표준편차)					
			인지요소 1	인지요소 2	인지요소 3	인지요소 4	인지요소 5	인지요소 6
2003	여	2,533	.7769 (.2681)	.7943 (.2786)	.7703 (.2314)	.7775 (.3089)	.7825 (.2512)	.7855 (.2223)
	남	2,776	.7795 (.2844)	.7978 (.2756)	.7933 (.2391)	.7721 (.3230)	.7875 (.2474)	.7795 (.2383)
	전체	5,309	.7783 (.2767)	.7961 (.2770)	.7823 (.2357)	.7747 (.3163)	.7851 (.2492)	.7823 (.2308)
2007	여	992	.7762 (.2813)	.6731 (.2927)	.7278 (.2340)	.7513 (.3055)	.7643 (.2199)	.7628 (.2377)
	남	1,124	.7716 (.2918)	.6914 (.3530)	.7582 (.2432)	.7562 (.3067)	.7739 (.2201)	.7645 (.2642)
	전체	2,116	.7738 (.2869)	.6828 (.3262)	.7439 (.2393)	.7539 (.3061)	.7694 (.2200)	.7637 (.2520)
2011	여	743	.7928 (.2592)	.7940 (.2197)	.7932 (.2106)	.7945 (.2826)	.7971 (.2462)	.8059 (.2741)
	남	722	.7960 (.2449)	.8036 (.2722)	.7858 (.2341)	.7976 (.3218)	.7802 (.2291)	.7850 (.3036)
	전체	1,465	.7943 (.2522)	.7987 (.2469)	.7895 (.2224)	.7960 (.3025)	.7887 (.2380)	.7956 (.2891)

서 통계적으로 유의미한 차이가 없는 것으로 나타났다.

## 2. 인지요소 숙달 유형의 성별 차이

<표 IV-3> 성별 인지요소 숙달의 주요 유형  
(상위 5개 유형)

연도	여		남	
	숙달 유형	비율	숙달 유형	비율
2003	000000	2.2%	000000	2.4%
	000001	1.1%	000001	1.0%
	001000	0.9%	001000	1.0%
	000010	0.8%	000010	0.8%
	111111	0.8%	100000	0.7%
2007	000000	1.9%	000000	2.7%
	000001	0.9%	000010	1.4%
	000010	0.9%	001000	1.0%
	001000	0.9%	001010	0.8%
	010000	0.9%	000001	0.7%
2011	000000	1.1%	000000	2.2%
	010000	0.7%	000010	1.0%
	111111	0.7%	100000	0.9%
	001000	0.6%	001000	0.8%
	000001	0.5%	010000	0.7%

인지요소 숙달의 주요 유형은 <표 IV-3>에 제시하였으며, 숙달을 1, 비숙달을 0으로 표시하였다. 3개년도 모두 남녀 공통으로 6개 인지요소를 모두 숙달하지 못한 학생들이 가장 높은 비율을 나타내었으며, 2007년과 2011년에는 주요 숙달 유형이 남녀 간에 다소 차이를 보였다. 2007년에는 남학생의 경우 5번 인지요소만 숙달한 학생의 비율이 1.4%로 두 번째로 높았으며, 여학생의 경우 1번 인지요소만 숙달한 학생의 비율이 여학생 0.9%로 두 번째로 높게 나타났다. 2011년에는 남학생의 경우 5번 인지요소만 숙달한 학생과 1번 인지요소만 숙달한 학생의 비율이 상대적으로 높았던 반면, 여학생의 경우는 이와 같은 숙달 유형이 상위 5개 유형 안에 들지 않았다.

<표 IV-2> 인지요소 숙달 확률의 성차에 대한 MANOVA 분석 결과

연도	분산원	종속변수	제곱합	df	평균제곱	F	유의확률
2003	성별	인지요소 1	.009	1	.009	.121	.728
		인지요소 2	.016	1	.016	.204	.652
		<b>인지요소 3</b>	.700	1	.700	12.622	.000
		인지요소 4	.039	1	.039	.388	.533
		인지요소 5	.033	1	.033	.534	.465
		인지요소 6	.048	1	.048	.901	.343
	오차	인지요소 1	406.515	5,307	.077		
		인지요소 2	407.307	5,307	.077		
		인지요소 3	294.278	5,307	.055		
		인지요소 4	530.948	5,307	.100		
		인지요소 5	329.589	5,307	.062		
		인지요소 6	282.658	5,307	.053		
	전체	인지요소 1	406.524	5,308			
		인지요소 2	407.322	5,308			
		인지요소 3	294.978	5,308			
		인지요소 4	530.987	5,308			
		인지요소 5	329.622	5,308			
		인지요소 6	282.706	5,308			
2007	성별	인지요소 1	.011	1	.011	.145	.713
		인지요소 2	.177	1	.177	1.664	.197
		<b>인지요소 3</b>	.488	1	.488	8.543	.004
		인지요소 4	.012	1	.012	.133	.716
		인지요소 5	.048	1	.048	1.001	.317
		인지요소 6	.001	1	.001	.023	.880
	오차	인지요소 1	174.061	2,114	.082		
		인지요소 2	224.830	2,114	.106		
		인지요소 3	120.664	2,114	.057		
		인지요소 4	198.106	2,114	.094		
		인지요소 5	102.293	2,114	.048		
		인지요소 6	134.337	2,114	.064		
	전체	인지요소 1	174.072	2,115			
		인지요소 2	225.007	2,115			
		인지요소 3	121.152	2,115			
		인지요소 4	198.118	2,115			
		인지요소 5	102.342	2,115			
		인지요소 6	134.339	2,115			
2011	성별	인지요소 1	.004	1	.004	.058	.809
		인지요소 2	.033	1	.033	.548	.459
		인지요소 3	.020	1	.020	.401	.527
		인지요소 4	.004	1	.004	.040	.842
		인지요소 5	.105	1	.105	1.849	.174
		인지요소 6	.161	1	.161	1.923	.166
	오차	인지요소 1	93.094	1,463	.064		
		인지요소 2	89.228	1,463	.061		
		인지요소 3	72.397	1,463	.049		
		인지요소 4	133.949	1,463	.092		
		인지요소 5	82.796	1,463	.057		
		인지요소 6	122.228	1,463	.084		
	전체	인지요소 1	93.098	1,464			
		인지요소 2	89.261	1,464			
		인지요소 3	72.417	1,464			
		인지요소 4	133.953	1,464			
		인지요소 5	82.901	1,464			
		인지요소 6	122.388	1,464			

### 3. Fusion 모형 문항 모수의 성별 차이

#### 가. 조건부 정답확률( $\pi_i^*$ )의 성별 차이

$\pi_i^*$ 는 한 문항을 푸는 데 필요한 인지요소를 모두 숙달한 학생이 그 문항에 정답을 할 확률 이므로, 어떤 문항의  $\pi_i^*$  모수가 특정 성별에서 낮게 나타난다는 것은 해당 성별의 학생들이 그 문항을 푸는 데에는 이미 정의된 인지요소 외에 다른 요소가 더 필요하다거나, 그 문항이 가지고 있는 고유한 특징으로 인해 실수로 틀릴 확률이 더 높다고 해석할 수 있다. 본 연구의 분석 결과, 조건부 정답확률을 나타내는  $\pi_i^*$  모수가 .9 이하로 나타난 6개의 문항 중 5개 문항은 남녀 간에 동일하게 낮게 나타났으나, 1개 문항(M032097)은 여학생에게 있어서만 낮게 나타났다.

[그림 IV-2]에 제시된 M032097번(2007년도 블록 07-7번) 문항은 인지요소 4와 6을 측정하는 문항으로 여학생의  $\pi_i^*$  모수가 낮게 산출되었다.

이 문항은 원의 넓이를 구하여 어렵값을 산출하는 측정 영역과 비례식을 푸는 요소로 구성되어진 문항이다. 이 문항을 해결하는 과정에서 여학생이 남학생에 비하여 특히 계산 실수를 더 하였거나 도형의 닮음으로 오인하여 실수를 범했을 가능성이 더 높다고 보기는 어렵고, '1m<sup>2</sup>당 평균 2마리'라는 기본 단위(unit)의 통계적 용어로 인하여 문항을 해석하는 데 어려움을 주었을 가능성을 유추해 볼 수 있다. 그러나 본 문항 하나로 이와 같은 해석을 일반화하기에는 한계가 있으며, 이러한 유형의 문항에 대한 성차 연구를 추후 더 진행하는 것이 바람직해 보인다.

반지름이 10m인 원 모양의 연못이 있다. 이 연못에는 1m<sup>2</sup>당 평균 2마리의 개구리가 산다고 한다. 이 연못에 사는 개구리는 대략 몇 마리인가? (단, π는 약 3.14이다.)

① 120  
 ② 300  
 ③ 600  
 ④ 2400

[그림 IV-2] M032097번 (2007년도 블록 07-7번) 문항

<표 IV-4> 인지요소별 변별도가 가장 높은 문항

인지요소	성별	문항 ID	블록내 번호	$r_{ik}^*$ 모수
인지요소 1	여	M032691	2007년도 블록 03-15	.1493
	남	M032414	2007년도 블록 07-9	.1051
인지요소 2	여	M042301B	2007년도 블록 02-7B	.0210
	남	M042301B	2007년도 블록 02-7B	.0297
인지요소 3	여	M032344	2007년도 블록 05-4 (2003년도 실시)	.4398
	남	M032344	2007년도 블록 05-4 (2003년도 실시)	.3914
인지요소 4	여	M022234A	2007년도 블록 01-11A	2점: .0616 1점: .0123
	남	M022234A	2007년도 블록 01-11A	2점: .0816 1점: .0129
인지요소 5	여	M032331	2011년도 블록 07-6	.2767
	남	M032331	2011년도 블록 07-6	.5412
인지요소 6	여	M042300B	2011년도 블록 06-10B	.1287
	남	M042300A	2011년도 블록 06-10A	.0651

나. 인지요소 수준 변별도( $r_{ik}^*$ )의 성별 차이

다음으로  $r_{ik}^*$  모수는 인지요소 수준의 변별도를 의미하며,  $r_{ik}^*$ 의 값이 0에 가까울수록 문항을 푸는데 그 인지요소가 중요한 역할을 하고 높고 1에 가까울수록 기여도가 낮다고 해석할 수 있다. 각 인지요소를 측정하는데 있어 가장 변별도가 높은 문항은 남녀별로 구분하여 <표 IV-4>에 제시하였다. 6개의 인지요소 중 인지요소 3과 5는 변별도가 다른 인지요소들에 비해 상대적으로 낮게 나타나, 분석에 사용된 TIMSS 문항들이 이 인지요소의 숙달 여부를 민감하게 반영하는 정도가 다른 인지요소에 비해 상대적으로 낮음을 알 수 있다. 이를 수학교수 학습 측면으로 생각해 보면, 입체도형의 모양과 도형의 변환에 대한 내용은 인지적 내용을 숙달하는 것에 의해 문제해결 능력이 좌우될 가능성이 다른 요소들보다 작고, 직관과 같은 또 다른 능력에 의해 좌우될 가능성이 있다는 해석이 가능하다. 따라서 향후 입체도형의 모양과 도형의 변환 내용에 대한 교수·학습은 내용 요소의 숙달 외에 또 다른 능력을 개발해 나가야 한다는 시사점을 던져 주고 있다.

인지요소 1과 6을 제외한 4개의 인지요소는 해당 인지요소를 가장 잘 변별하는 문항이 남학생과 여학생

간에 차이가 없었던 반면, 인지요소 1과 6에서는 남학생과 여학생 간에 가장 잘 변별하는 문항에 차이가 있었다. 인지요소 1을 측정하는데 있어서 여학생은 M032691번 문항이, 남학생은 M032414번 문항이 가장 잘 변별하는 것으로 나타났으며, 인지요소 6을 측정하는데 있어서 여학생은 M042300B번 문항이, 남학생은 M042300A번 문항이 가장 잘 변별하는 것으로 나타났다.

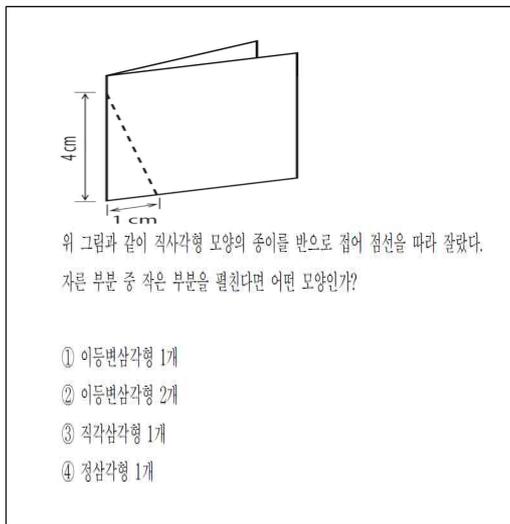
$r_{ik}^*$  모수의 성차가 크게 나타난 문항은 <표 IV-5>에 제시하였으며, 성차의 판단 기준은 남학생과 여학생 간의  $r_{ik}^*$  값이 .3 이상의 차이를 보이는 경우 큰 차이를 보인다고 판단하였다. 성차가 크게 나타난 6개 문항을 검토한 결과, 각 문항별로 남학생과 여학생의 접근 방식에 있어 다소 차이가 있는 것으로 유추할 수 있었다.

예를 들어, [그림 IV-3]에 제시된 M032679번 (2011년 블록 07-8번) 문항의 경우, 인지요소 2는 여학생보다 남학생들이 문항을 푸는데 더 중요하게 기능하였다. 이 문항을 풀기 위해서는 경험적으로 접어서 자르면 앞장과 같은 모양이 대칭을 이루게 된다는 모양을 추측해서 이등변삼각형이 된다고 직관적으로 판단할 수도 있으나, 분석적 관찰에 의해서 ‘두 삼각형이 합동’을 이루거나, ‘꼭지각의 수직이등분선’, ‘밑각이 서로 같

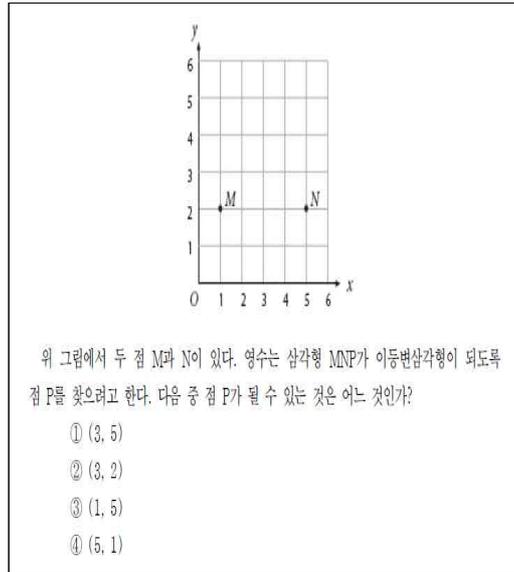
<표 IV-5>  $r_{ik}^*$  모수의 성차가 크게 나타난 문항

연도	문항 ID	블록내 번호	측정하는 인지요소	$r_{ik}^*$ 모수		
				여	남	차이 (여-남)
2003	M012026	03-2	2	.4184	.7279	-.3095
			5	.8764	.6001	.2762
	M022142	03-9	1	.7275	.2780	.4495
			2	.7355	.6965	.0390
	M032689	10-11	4	.6556	.6137	.0418
5			.3772	.7952	-.4180	
2007	M042148	02-11	6	.8653	.5298	.3355
	M032294	07-10	2	.9228	.4979	.4249
			6	.7579	.6519	.1060
2011	M032679	07-8	2	.5707	.2388	.3319

은 삼각형'이 된다는 속성을 떠올리게 되어 이등변삼각형이 됨을 찾을 수도 있다. 따라서  $r_{ik}^*$  모수의 차이에 비추어 볼 때, 남학생은 직관적 관찰에 의해 문제를 푸는 경우보다 인지요소 2(도형에 관한 수학적 정리)를 활용하여 이등변삼각형의 속성에 대한 보다 확실한 지식에 의해 문제를 푸는 경향이 있는 반면, 여학생은 경험상 혹은 직관적인 감각으로 문제를 푸는 경향이 상대적으로 강했다고 해석할 수 있다. 또한 [그림 IV-4]에 제시된 M032294번(2007년도 블록 07-10번) 문항의 경우, 2번 인지요소는 여학생에게 있어 거의 기능하지 않고 남학생에게만 중요한 인지요소로 기능하였으며, 6번 인지요소는 여학생과 남학생 모두 문제를 푸는데 큰 역할을 하지 않았다고 볼 수 있다. 따라서 M032679번 문항과 마찬가지로, 여학생은 이등변삼각형의 속성을 몰라도 직관적으로 문제를 해결하는 경향이 강하다는 해석이 가능하며, 남학생의 경우 인지요소 2(도형에 관한 수학적 정리)를 활용하여 이등변삼각형의 속성에 대한 지식과 적용이 더 중요하게 작용했거나 이를 이용해서 해결하려는 의지가 더 강했다고 추측할 수 있다.



[그림 IV-3] M032679번(2011년도 블록 07-8번) 문항



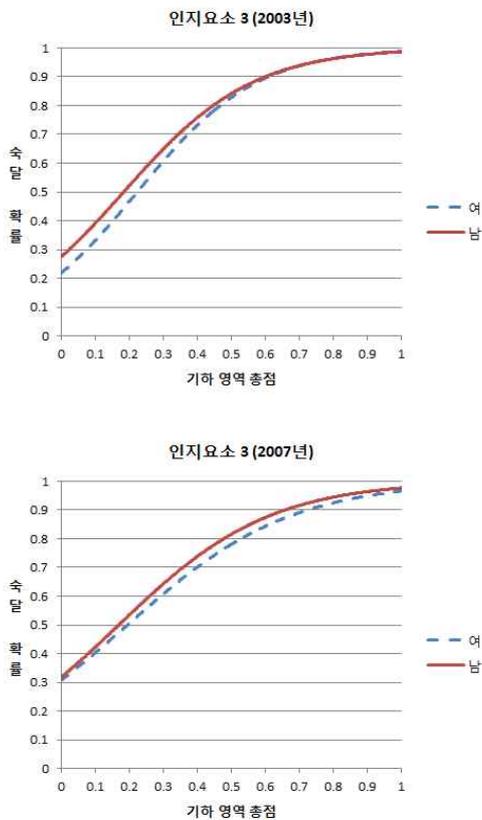
[그림 IV-4] M032294번(2007년도 블록 07-10번) 문항

#### 다. 인지요소 특성곡선의 성별 차이

다음으로 인지요소 숙달 확률의 성차가 통계적으로 유의미하게 나타난 2003년도 인지요소 3과 2007년도 인지요소 3에 대해서 어떤 능력대에 분포한 학생들의 성차가 크게 나타나는지 확인하기 위해 인지요소 특성곡선(Attribute Characteristic Curve)을 비교하였다. 인지요소 특성곡선은 피험자의 능력대에 따라 인지요소의 숙달 확률이 어떠한 패턴으로 변화하는지를 나타낸 곡선으로, 본 연구에서는 피험자의 능력을 대표하는 지표로서 TIMSS 기하 영역에 대한 총점 중 첫 번째 유의측정값(plausible value)을 사용하여 피험자의 능력별 인지요소 숙달 확률을 2모수 로지스틱 함수를 사용하여 추정하였다. 기하 총점은 0에서 1 사이의 값으로 변환하여 사용하였다.

[그림 IV-5]에 제시된 바와 같이, 2003년도에 측정된 인지요소 3의 경우 피험자의 능력이 .7이상인 상위권 학생들은 남녀 간에 숙달 확률에 있어서 차이가 없는 반면 성적이 하위권으로 갈

수록 인지요소 숙달 확률의 차이가 점차 벌어지고 있음을 알 수 있다. 또한 2007년에 측정한 인지요소 3의 경우는 상위권과 하위권 학생들보다는 중위권 학생들에게 있어서 성차가 상대적으로 크게 나타났으나 그 크기는 2003년도의 하위권에서 추정된 숙달 확률의 차이에 비해 다소 작았고, 지면 관계상 그림에서는 생략하였으나 2011년에는 모든 능력대에서 인지요소 3의 성차가 나타나지 않았다.



[그림 IV-5] 인지요소 특성곡선(ACC)의 성별 차이

## V. 논의 및 결론

수학에서의 성차는 교수·학습 환경에서 학습

자에 대한 공평성을 추구하는 맥락에서 수학교육 분야에서 지속적인 관심을 받아 왔다. 수학은 전통적으로 남학생이 여학생에 비해 높은 성취를 보이는 영역으로 인식되어 왔으나, 최근에는 수학에서의 성차가 여전히 존재한다는 연구들과 더불어 성차가 완화되거나 점차 사라지고 있다는 경험적 증거들도 종종 보고되고 있다(예, 강인수, 2000; 박정 외, 2004; 윤희선, 2000; Eccles, Lord, Roeser, Barber, & Jozefowicz, 1997; Friedman, 1989, Halpern, 2000). 이와 같은 일관되지 않은 결과는 수학의 여러 하위 영역에서 성차가 서로 다른 형태로 발현되었기 때문일 수도 있기 때문에, 최근 수학교육 분야에서의 성차 연구 동향은 수학 학업성취도 전체에 성차가 있는가를 조사하기 보다는 좀 더 구체적인 영역과 관련하여 어떤 차이가 있는가를 분석하는 추세라고 할 수 있다. 이와 같은 연구 동향을 반영하기 위하여 본 연구에서는 오랜 기간 성차가 존재한다고 알려져 있는 기하 영역에서의 성차를 인지진단모형이라는 비교적 새로운 접근법을 통하여 재조명했다는 점에서 의의가 있다 하겠다.

본 연구는 2003년부터 2011년까지 3개 주기 동안 실시된 TIMSS 8학년 수학과 데이터를 활용하여 우리나라 중학생들이 기하 영역의 각 하위 인지요소에서 어떠한 형태로 성별 차이를 나타내는지에 관해 인지진단모형의 관점에서 고찰하였다. 분석 결과를 살펴보면, 우선 2003년에는 인지요소 3(입체도형의 모양)에서 남학생의 평균이 높게 나타났으며 그 밖의 인지요소에 있어서도 인지요소 4(기하적 측정)와 6(대수를 활용한 기하)을 제외하고는 남학생의 평균이 다소 높은 것으로 나타났다. 인지요소 4에 해당되는 문항은 대부분 도형의 넓이와 부피 등을 구하는 문제로서 공식을 활용하여 푸는 문제들이 주를 이루고 있으며 인지요소 6은 대수 개념을 포함한 기하 문제로서, 본 연구의 결과는 전통적으로

공식을 활용한 문제해결이나 대수에서 여학생의 성취도가 높거나(예, 김진만, 2002) 성차가 나타나지 않는다는 연구 경향(예, Ma, 1993)을 반영하고 있다.

2007년의 경우에도 여전히 남학생의 평균이 전반적으로 높은 것으로 나타났으나 2011년의 경우 인지요소 3, 5, 6은 여학생이 높았고, 나머지 인지요소에서는 남학생이 높았으나 남녀 간의 차이가 크지는 않은 것으로 나타났다. 그간 수학교육 분야에서는 남녀 학생의 기하 관련 문제해결 전략과 논리 추론적 능력에서는 별 차이가 없으나 시각적 공간 지각력에서는 그 차이가 유의미하게 나타나기 때문에 공간 지각력 학습에서 성차를 고려한 내용을 반영할 것이 제안되어 왔다(Battista, 1990). 본 연구의 결과에서 고무적인 것은 전통적으로 공간 감각이나 도형의 변환 측면에서는 남학생이 우세하다는 기존의 연구 결과(예, Benbow & Stanley, 1980, 1983; 이향란 1991)와는 다르게, 최근 들어 공간 지각력을 다룬 인지요소 3(입체도형의 모양)과 인지요소 5(도형의 변환)에서 여학생의 성취도가 향상되었다는 점이다.

그동안 수학교육계는 학생들의 공간 감각을 키우기 위한 방안으로 다양한 학습 프로그램을 시도해 오고 있다. 이는 단지 지식적 관점에서만 기하를 배우는 것이 아니라, 도형에서 패턴 찾기, 투시도 그리기, 추측하기, 종이접기를 이용해 펼친 그림 만들기, 기하판(Geoboard) 이용하기, 공간 방향 생각하기, 전개도 만들기 등을 이용하여 교수학습 자료를 구성하고 이를 통해 기하학적 감각을 키우는데 노력을 기울이고 있다. 더불어 성차와 관련해서는 성차에 대한 공평성을 염두에 두고 잘 설계된 수업 프로그램을 권장해 오고 있다. 예를 들어, 테셀레이션 활동이 수학의 심미성과 여학생들의 긍정적인 수학적 태도와 흥미 형성에 도움을 주었다는 연구(예; 이경

미, 2003; 오혜원, 2000)에 기반하여, 기하적 변환을 테셀레이션의 미적 요소를 강조해서 가르치는 활동(고상숙·고호경, 2006) 등 보다 적극적인 방법으로 성차를 줄일 수 있는 방안을 추구하기 위한 노력을 꾸준히 기울이고 있다. 본 연구의 결과는 이러한 교육계와 사회의 노력에 따라 변화되어지고 있는 결과라고도 볼 수 있을 것이다.

한편 방법론적인 측면에서 볼 때, 인지진단모형을 적용하여 성차를 분석하면 기존의 성차 연구에서 얻을 수 없었던 세부적인 정보를 구체적으로 고찰할 수 있다는 장점을 가진다. 구체적으로, 인지진단모형을 적용하면 인지요소별로 남학생과 여학생 간의 숙달 확률을 파악할 수 있다는 점 외에, 동일한 문항을 푸는데 있어서 남녀 간에 주로 사용하는 인지요소 또는 접근법이 어떻게 다른지를 확인하거나 특정 성별의 학생들이 문항을 푸는데 있어서 추가로 필요한 인지요소의 존재 여부를 탐색하는 등의 정보를 알 수 있다. 또한 인지요소별로 가장 변별도가 높은 문항을 확인하여 그 원인을 유추해냄에 따라 교수학습의 시사점을 발견하거나, 진단적 기능을 강조하는 검사의 문항을 개발할 때 변별력을 최대화할 수 있는 문항을 선별할 수 있는 기본 정보를 제공한다는 점에 있어서도 유용성을 가진다. 따라서 이와 같은 분석 결과를 통하여 향후 기하 학습에서의 시사점 등을 도출해 내는 것은 더욱 의의가 있는 작업이라 할 수 있다.

예를 들어, 문항의  $\pi_i^*$  모수의 상대적인 크기를 바탕으로 어떠한 문항을 푸는 데에 이미 정의된 인지요소 외에 다른 요소가 더 필요하다거나, 그 문항이 가지고 있는 고유한 특징으로 인해 학생들이 실수로 틀릴 확률을 해석해 낼 수 있다. 본 연구에서 각 인지요소별로 학생들의 수행을 가장 잘 변별하는 문항에 있어서, 인지요소 1의 경우 여학생은 M032691번 문항이, 남학생은

M032414번 문항이 가장 잘 변별하는 것으로 나타났다. M032691번 문항에 포함되어 있는 인지적 영역은 1요소로서 각에 대한 문항이다. 문항을 자세히 살펴보면, 예각과 둔각의 모양에 대한 간단한 내용만 들어 있어서 복잡한 내용에 속하지는 않으나, 정형화된 문항이 아닌 열린 문항(open ended) 문항이라 할 수 있다. 반면 M032414번 문항은 직각과 각의 이등분선이라는 주어진 조건을 활용하여 푸는 정형화된 문항이기는 하나 3단계의 논리를 적용해야 하는 문항이다. 따라서 여학생은 인지요소 1의 숙달 외에 정형화되지 않은 문항이 출제됨에 따라 이에 대한 변별도가 높아졌다고 볼 수 있다. 반면에 남학생의 변별력을 높인 문항은 그 문항이 가지고 있는 인지요소 1의 난이도에 의한 것이라 추측할 수 있다. 다시 말하면, 여학생은 열린 문항에 대한 교수학습에서, 남학생은 복잡성이 있는 문항에 대해 변별이 생긴 것이다. 따라서 단순한 인지요소를 포함한 문항임에도 불구하고 비정형화된 문항이나 열린 문항에 대한 여학생의 문제해결력이 낮은 것에 대한 교수·학습 방법을 고려할 필요가 있을 것이다.

또한 앞에서 논의한 M032679번 문항의 예시와 같이 여학생은 이등변삼각형의 속성을 몰라도 직관적으로 문제를 해결하는 경향이 강하며, 남학생의 경우 인지요소 2(도형에 관한 수학적 정리)를 활용하여 문제를 푼 경향이 강하다는 해석이 가능하다. 이는 여학생은 도형의 구체적인 성질 보다는 경험이나 시각적 직관에 의해 문제를 푸는 경향이 남학생보다 강하며 남학생은 구체적인 도형의 성질에 의거하여 문제를 풀려는 경향이 강하다고 해석할 수 있다. 이와 같이 성차에서의 다양한 차이점을 제시하는 것은 어떤 접근 방법에 대한 장단점이나 효율성을 논하기 보다는 남녀 학생의 특성을 반영하여 교수학습에서 보다 다양한 접근을 시도해야함을 제안해

볼 수 있다.

성차 외에도 전체적인 경향에서 입체도형의 모양과 도형의 변환 내용에 대한 기하 학습은 인지적 내용에 대한 숙달보다는 직관과 같은 능력에 의해 문제해결 능력이 달라질 수 있다는 해석에 따라, 향후 이에 대한 어떠한 교수학습을 적용해야 할지에 대한 추후 연구 과제를 던져주고 있다.

학교에서 제공되는 교육과정은 여학생 뿐 아니라, 소외계층 또는 장애학생과 같은 다양한 주체들의 욕구를 충족시킬 책임이 있다는 주장(Sullivan, 1994)을 우리가 여전히 받아들인다면, 수학에서 성 불평등의 존재 및 그와 관련된 지속적이고 다양한 추수 연구들을 통해 성차를 줄이기 위한 노력을 해야 할 것이다(Else-Quest et al. 2010). 특히 본 연구에서는 인지진단모형을 적용하여 TIMSS 8학년 기하 영역에서 나타나는 성차의 변화 연도별 추이를 탐색하였으며, 그 결과 성차가 점차 완화되는 경향을 발견하였다. 그러나 이러한 경향이 우리나라 중학생들에게 보편적으로 관측되는 현상인지, 아니면 TIMSS 검사 문항 또는 표집된 피험자들의 특수성으로 인한 것인지를 분명히 하기 위해서는 다른 대규모 데이터를 바탕으로 한 추가적인 연구가 추후 수행될 필요가 있다.

## 참고 문헌

- 강인수(2000). **남녀별 수학 성취도와 비판적 사고력의 상관관계에 관한 연구 : 중학교 3학년 남녀 학생을 중심으로**. 경남대학교 대학원 석사학위논문.
- 고상숙, 고희경(2006). 컴퓨터를 이용한 기하 변환학습에서 남녀성차에 따른 연구. **한국학교수학회논문집**, 9(4), 539-556.
- 고호경·고상숙(2008). 중학교 함수의 수학적 과

- 정에서의 성차 연구. **수학교육**, 47(3), 273-297.
- 권오남·박경미(1995). 수학생취도에 있어서의 성별차이에 대한 고찰, **한국여성학**, 11, 202-232.
- 권오남(2000). 양성평등을 위한 수학교육의 과제와 전망. **수학교육논문집**, 10, 417-440.
- 김진만(2002). 인문계고등학교의 대수학 및 기하학 영역에서의 남녀별 성취도에 관한 연구. 서남대학교 대학원 석사학위논문.
- 김희경, 김부미 (2011). 인지진단모형을 활용한 수학 학업성취 결과 분석 - 2011년 국가수준 학업성취도 평가 자료를 중심으로. **학교수학**, 15(2), 289-314.
- 김희경, 한정아, 최숙기, 김부미 (2012). 인지진단모형을 적용한 학업성취 프로파일 분석 및 결과 보고 방안. 연구보고 RRE 2012-7. 한국교육과정평가원.
- 류신열(1998). 학년별에 따르는 남녀에 대한 수학적 능력에서의 비교 연구. 국민대학교 교육대학원 석사학위논문.
- 박정 외(2004). 남·여학생의 성취도 차이 해소 방안 : TIMSS와 PISA 2003 결과를 반영하여. 한국교육과정평가원 연구보고 RRE 2004-10.
- 박효주(2009). 성별과 학업성취도에 따른 중학생의 수학불안요인의 차이에 관한 연구: 중학교 2학년을 대상으로. 강원대학교 대학원 석사학위논문.
- 서경숙(2007). 중학생의 성별과 수학생취도에 따른 수학불안 요인의 차이에 관한 연구. 전북대학교 대학원 석사학위논문.
- 오혜원(2000). 중학교 CA활동에서 테셀레이션 도입을 통한 기하학습의 효과 연구. 이화여자대학교 대학원 석사학위논문.
- 윤희선(2000). 수학 문제상황이 남녀 학생의 문제해결에 미치는 영향. 이화여자대학교 대학원 석사학위논문.
- 이경미(2003). 중학교 기하수업에서 테셀레이션 활용을 통한 수학적 흥미 유발에 관한 연구. 신라대학교 대학원 석사학위논문.
- 이대식·김수미(2003). 수학학습에서의 성차에 대한 초등학교 학생 및 교사의 인식조사, **초등교육연구**, 16(1). pp.297-315.
- 이연옥(1989). 수학 성취에서 남녀차이에 영향을 주는 요인에 관한 연구. 서울대학교 대학원 석사학위논문.
- 이지혜(2008) 중학교 여학생의 수학적 성향개발을 위한 재량활동 자료개발연구. 성신여자대학교 대학원 석사학위논문.
- 이향란(1991). 남녀간의 수학능력차에 관한 연구. 서울대학교 대학원 석사학위논문.
- 정경아(2005). 중등학교의 수학에서의 성별격차 및 해소방안 연구. **젠더리뷰**, 9, 92-96.
- 최하나(2007). 학교 유형별 및 성별에 따른 수학적성에 관한 비교 연구. 경남대학교 대학원 석사학위논문.
- 홍경아(1994). 공간 능력 신장을 위한 교과서 지도 방안과 그 효과에 관한 연구: 중학교 기하 중심으로. 이화여자대학교 대학원 석사학위논문.
- Armstrong, J. M.(1981). Achievement and participation of women in mathematics: results of two national survey. *Journal for Research in Mathematics Education*, 15(5), 356-372.
- Baron-Cohen, S. (2003). *He essential difference: The truth about the male and female brain*. N. Y.: basic Books.
- Battista, M. T. (1990). Spatial visualization and gender differences in high school geometry. *Journal of Research in Mathematics Education*, 21(11), 47-60.
- Bellisari, A. (1989). Male superiority in mathematical aptitude: An artifact. *Human Organization*, 48,

- 273-279.
- Benbow, C. P. (1988). Sex differences in mathematical reasoning ability in intellectually talented preadolescents: Their nature, effects, and possible causes. *Behavioral & Brain Science, 11*, 169-232.
- Benbow, C. P., & Stanley, J. C. (1980). Sex differences in mathematical ability: Fact or artifact. *Science, 210*(4475), 1262-1264.
- Benbow, C. P., & Stanley, J. C. (1982). Consequences in high school and college of sex differences in mathematical reasoning ability: A longitudinal perspective. *American Educational Research Journal, 19*(4), 598-622.
- Bennet, M. (1996). Men's and women's self-estimates of intelligence. *The Journal of social Psychology, 136*, 411-412.
- Betz, N. E. (1997). What stops women and minorities from choosing and completing majors in science and engineering. In D. Johnson (Ed.) *Minorities and girls in school: Effects on achievement and performance*. Thousand Oaks CA; Sage. Series on Leaders in Psychology.
- Burton, L. (1990). *Gender and mathematics: An international perspective*. London, England: Cassell.
- Chen, Y. H., Ferron, J. M., Thompson, M. S., Gorin, J. S., & Tatsuoka, K. K. (2010). Group comparisons of mathematics performance from a cognitive diagnostic perspective. *Educational Research and Evaluation, 16*(4), 325-343.
- Chipman, S. F., Brush, L. R., & Wilson, D. M. (1985). *Women and Mathematics: Balancing the Equation*, Hillsdale, NJ: Lawrence Erlbaum Associates.
- Cronin, H. (2005). The vital statistics: evolution, not sexism, puts us at a disadvantage in the sciences. *The Guardian, 3*, pp.21.
- Davidson, L. & Schofield, W. (2002). Female Voice in Virtual reality: Drawing Young Girls into an Online World, In K. Renninger & W. Shumar (Eds.) *Building Virtual Communities: Learning and Change in Cyberspace* (pp.34-59). NY: Cambridge University Press.
- DiBello, L., & Stout, W. (2010). *Arpeggio version 3.1 [Computer Program]*. Chicago: Applied Informative Assessment Research Enterprises.
- DiBello, L. V., Roussos, L. A., & Stout, W. (2007). Review of Cognitively Diagnostic Assessment and a Summary of Psychometric Models. *Handbook of Statistics, 26*, 1-52.
- DiBello, L. V., Stout, W., & Roussos, L. A. (1995). Unified cognitive/psychometric diagnostic assessment likelihood-based classification techniques. In: Nichols, P.D., Chipman, S.F., Brennan, R.L. (Eds.), *Cognitively Diagnostic Assessment*. Erlbaum, Mahwah, NJ, pp. 361 - 389.
- Ding, C. S., Song, K., & Richardson, L. I. (2007). Do mathematical gender differences continue? A longitudinal study of gender difference and excellence in mathematics performance in the U. S.. Educational Studies: *Journal of the American Educational Studies Association, 40*(3), 279-295.
- Eccles, J. S., Lord, S. E., Roeser, R. W., Barber, B. L., & Jozefowicz, D. M. (1997). The association of school transitions in early adolescence with developmental trajectories through high school. In J. Schulenberg & J. Maggs & K. Hurrelmann (Eds.), *Health risks and developmental transitions during adolescence*, (pp. 283-320). New York: Cambridge

- University Press.
- Else-Quest, N., Hyde, J., & Linn, M. (2010). Cross-National Patterns of Gender Differences in Mathematics: A Meta-Analysis. *American Psychological Association, 236*(1), 103 - 127.
- Fennema, E. (1985). Explaining Sex-related Differences in Mathematics: Theoretical Models. *Educational Studies in Mathematics, 16*, pp.303-320.
- Fennema E. & Leder, G. (1990). *Mathematics and gender*. New York: Teachers College Press.
- Finn, J. (1980). Sex differences in educational outcomes: A cross-national study. *Sex Roles, 6*, 9-25.
- Friedman, L. (1989). Mathematics and the gender gap: A meta-analysis of recent studies on sex differences in mathematical tasks. *Review of Educational Research, 59*(2), 185-213.
- Fu, J. (2005). *A polytomous extension of the Fusion Model and its Bayesian parameter estimation*. Unpublished doctoral dissertation. University of Wisconsin- Madison.
- Furnham, A., Clark, K., & Bailey, K. (1999). Sex differences in estimates of multiple intelligences. *European Journal of Personality, 13*, pp.24
- Furnham, A., & Fong, G. (2000). Self-estimated and psychometrically measured intelligence; A cross-cultural and sex differences study of British and Singaporean students. *North American Journal of Psychology, 2*, 191-200.
- Halpern, D. F. (2000). *Sex differences in cognitive abilities* (3rd ed.). Mahwah, NJ: Erlbaum.
- Hartz, S. (2002). *Bayesian framework for the unified model for assessing cognitive abilities: Blending theory with practicality*. Unpublished doctoral dissertation. Champaign, IL: University of Illinois.
- Hartz, S. & Roussos, L. (2008). *The Fusion Model for Skills Diagnosis: Blending Theory with Practicality*. ETS RR-08-71. Educational Testing Service.
- Henson, R. & Douglas, J. (2005). Test construction for cognitive diagnosis. *Applied Psychological Measurement, 29*(4), 262-277.
- Henson, R. & Templin, J., & Willse, J. (2009). Defining a family of cognitive diagnosis models using log-linear models with latent variables. *Psychometrika, 74*(2), 191-210.
- Honigsfeld, A., & Dunn, R. (2003). High school male and female learning-style similarities and differences in diverse nations. *Journal of Educational Research, 96*(4), 1-12.
- Im, S. & Park, H. J. (2010). A comparison of US and Korean students' mathematics skills using a cognitive diagnostic testing method: linkage to instruction. *Educational Research and Evaluation, 16*(3), 287-301.
- International Association for the Evaluation of Educational Achievement (2011). *TIMSS 2011 Mathematics Framework*. IEA. <http://timssandpirls.bc.edu>.에서 2014년 1월 2일 자료 얻음.
- Jaccard, P. (1901). Distribution de la flore alpine dans le bassin des Dranses et dans quelques régions voisines. *Bulletin de la Société Vaudoise des Sciences Naturelles, 37*, 241-272.
- Johnson, M., Lee, Y. S., Sachdeva, R. J., Zhang, J., Waldman, M., & Park, J. Y. (2013). *Examination of gender differences using the multiple groups DINA model*. Paper presented at the 2013 Annual Meeting of the National Council on Measurement in Education in San Francisco, CA.
- Junker, B. & Sijtsma, K. (2001). Cognitive

- assessment models with few assumptions and connections with nonparametric item response theory. *Applied Psychological Measurement*, 25, 258-272.
- Leahey, E. & Guo, G. (2001). Gender Differences in Mathematical Trajectories. *Social Forces*, 80(2), pp. 713-732.
- Lee, Y. S., Park, Y. S., and Taylan, D. (2011). A cognitive diagnostic modeling of attribute mastery in Massachusetts, Minnesota, and the U.S. national sample using the TIMSS 2007. *International Journal of Testing*, 11, 144-177.
- Lynn, R. Irwing, P., & Cammock, T. (2001). Sex differences in general knowledge, *Intelligence*, 30(1), pp.27-39.
- Ma, X. (1993). *An International Study of Gender Differences in Mathematics Achievement*. Master's thesis, University of British Columbia. Published as Ma (1995b) in Refereed Articles.
- Meece, J. (1981). *Individual differences in the affective reactions of middle and high school students to mathematics: A social cognitive perspective*. Unpublished doctoral dissertation, University of Michigan.
- Meece, L. & Wigfield. A. (1988). Math Anxiety in Elementary and Secondary School Students. *Journal of Educational Psychology*, 80(2) 210-216.
- Park, C. (2008). *A multilevel IRT model for group-level diagnostic assessment with application to TIMSS*. Unpublished doctoral dissertation. University of Wisconsin-Madison.
- Pinker, S. (2002). *The black state: The modern denial of human nature*. N.Y.: Viking.
- Royster, D. C., Herris, M. k., & Schoeps, N. (1999). Dispositions of college mathematics students. *International Journal of Mathematical Education in science & Technology*, 30, 317-333.
- Rupp, A. A., Templin, J. L., & Henson, R. A. (2010). *Diagnostic Measurement: Theory, Methods, and Applications*. NY: The Guilford Press.
- Sullivan, E. (1994). Achieving equity in mathematics education. *Thrust for Educational Leadership*, 23(7), pp.35-39.
- Summers, L. (2005, January 14). Remarks at NBER conference on diversifying the science and engineering workforce. Retrieved April 5, 2005 from <http://www.president.harvard.edu/speeches/2005/nber.html>.
- Tatsuoka, K.K. (1985). A probabilistic model for diagnosing misconceptions in the pattern classification approach. *Journal of Educational Statistics*, 12, 55 - 73.
- Tatsuoka, K. K., Corter, J. E., & Tatsuoka, C. (2004). Patterns of diagnosed mathematical content and process skills in TIMSS-R across a sample of 20 countries. *American Educational Research Journal*, 41(4), 901-926.
- Utsumi, M.C., & Mendes, C. R. (2000). Researching the attitudes towards mathematics in basic education. *Educational Psychology*, 20(2), 237-243.
- Zhou, J., Gierl, M. J., & Cui, Y. (2009). *Attribute Reliability in Cognitive Diagnostic Assessment*. Paper presented at the Annual meeting of the National Council on Measurement in Education, San Diego, CA.

# Gender Differences in Geometry of the TIMSS 8th Grade Mathematics Based on a Cognitive Diagnostic Modeling Approach

Yi, Hyun Sook (Konkuk University)

Ko, Ho Kyoung (Ajou University)

Gender differences have been given major attention in mathematics education in the context of pursuing gender equity in instructional and learning environment. It had been traditional belief that male students would outperform female students in mathematics, especially in the areas as geometry. This belief has been given doubts by cumulated empirical evidences that gender differences are gradually diminishing or even reversing its direction as time goes on. In this study, gender differences in geometry were explored using TIMSS 8th grade mathematics data administered in TIMSS 2003, 2007, and 2011, based on a cognitive diagnostic modeling(CDM) approach. Among various CDM models, the Fusion model was employed. The Fusion model has advantages over other CDM models in that it provides more detailed information about gender

differences at the attribute level as well as item level and more mathematically tractable. The findings of this study show that Attribute 3(Three-dimensional Geometric Shapes) revealed statistically significant gender differences favoring male students in TIMSS 2003 and 2007, but did not show significant differences in TIMSS 2011, which provides an additional empirical evidence supporting the recent observation that gender gap is narrowing. In addition to the general trends in gender differences in geometry, this study also provided affluent information such as gender differences in attribute mastery profiles and gender differences in relative contributions of each attribute in solving a particular item. Based on the findings of the CDM approach exploring gender differences, instructional implications in geometry education are discussed.

\* Key Words : cognitive diagnostic modeling(인지진단모형), gender difference(성차), TIMSS mathematics (TIMSS 수학), Geometry(기하)

논문접수 : 2014. 5. 11

논문수정 : 2014. 5. 28

심사완료 : 2014. 5. 31