

## 수학 평가 결과의 분석을 위한 인지 진단 이론의 활용

김 선 희\* · 김 수 진\*\* · 송 미 영\*\*\*

본 연구는 인지 진단 이론을 활용하여 수학 평가 결과를 분석하고 교수·학습에 활용하는 방안을 모색하고자 하였다. 2003~2006년에 실시된 국가수준 학업성취도 평가의 중학교 3학년 수학 검사에서 30개의 선다형 문항을 선정하여 검사지를 재구성하고 검사를 실시하였고 인지 진단 이론의 한 모형인 Fusion Model을 적용하여 평가 결과를 분석하였다. 검사 문항을 통해 학생들이 숙달한 수학적 속성을 판별하고, 학생 전체와 성취수준별로 숙달한 속성과 그 속성의 개수를 분석하였다. 그리고 학생 개인의 수학적 강점과 약점을 분석하여 교사들에게 학생 개인의 수학적 능력에 대한 정보를 구체적으로 알려줄 수 있었다. 이 결과는 학생들의 수학 학습에 대한 진단과 처방, 추후 학습 지도에 유용한 정보로 활용될 수 있을 것이다.

### 1. 서 론

교수·학습 상황에서 평가는 학생의 학습과 성취에 대한 교사의 의사결정을 돕기 위하여 정보를 수집하고 해석하여 활용하는 활동이다 (McMillan, 2004). 수학 교수·학습에서 평가는 교육과정에서 의도한 교육 목표를 달성하기 위해 이용되는 도구로서, 학생을 선발하고 서열화하기보다 수학 교수·학습 과정의 중요한 부분으로서 시행되고, 평가 결과가 차후 연계되는 학습 지도를 위한 자료로 활용되어야 한다. 따라서 수학 학습의 평가는 학생의 인지적 영역과 정의적 영역에 대한 유용한 정보를 제공하여 학생 개인의 수학 학습과 전인적인 성장을 돕고 교사의 교수 활동과 수업 방법을 개선하는 데 활용되어야 한다(교육인적자원부, 2007).

평가가 이러한 방향에서 활용되기 위해서 교사는 수학 학습 평가 결과를 정확하고 타당하게 분석 및 해석하고, 수업 및 학생에 대한 교육적 의사결정에 활용하며, 학생·학부모·교육 관련자와 정확히 의사소통할 수 있어야 한다. 검사 문항을 통한 평가에서도 평가 결과가 점수로만 제공되는 것에 그치지 않고 더 많은 정보를 내포해야 할 것이다. 예를 들어, 두 학생이 10개의 검사 문항에서 똑같이 8문항을 맞았다고 한다면, 이 두 학생은 맞힌 문항 수가 같기 때문에 같은 성적을 받게 되는 것에서 끝나는 것이 아니라, 그 학생이 어떤 문제를 틀렸고, 틀린 이유가 개념을 몰라서인지, 문제해결의 방법을 몰라서인지, 아니면 문제해결 과정에서의 오류인지를 교사는 판단하고 학생에게 알려야 한다. 이를 위해서는 평가 도구로 사용된 문항이 각각의 측면들을 살필

\* 신라대학교(mathsun@silla.ac.kr)

\*\* 한국교육과정평가원(sjkim@kice.re.kr)

\*\*\* 한국교육과정평가원(mysong@kice.re.kr)

수 있는 요소를 포함하고 있어야 하며, 교사는 각 문항에 대한 학생의 반응에 따라 개념을 바로잡아 주고 피드백을 제공해야 한다. 즉, 학생에게 '10문제 중 8개를 맞았으니 네 점수는 80점이다'라는 통보가 아니라, 이 문제를 틀린 원인이 무엇이며 어떻게 교정하거나 보충해야 하는지를 알려줄 수 있는 정보 제공으로서의 평가가 이루어져야 한다. Clarke(1997)이 주장한대로, 평가는 교수·학습 행동의 변화를 촉진시키고 예고해야 하는 것이다.

하지만 지금까지의 전통적인 교육 평가는 잘 만들어진 검사 문항을 구성하여 평가를 실시하고 학습자의 능력을 척도화된 점수로 만들기 위하여 정교하고 정확한 측정을 하는 데 치중해 있었다. 이러한 평가는 학생의 석차를 매기고 미래의 성과를 예측하기 위해서 능력을 비교할 수 있었으나 학생들이 수행을 더 잘하기 위해서 학습해야 하는 것들에 대한 유용한 진단적 정보를 제공하지 못하였다. 이에 따라 학생과 학부모들은 학습 결과를 점수에 의지하게 되고 점수를 향상시키기 위한 노력을 할 뿐 수학을 잘 하기 위해 어떤 개선점이 필요한지에 대해서는 알지 못한다. 최근 학생들의 복잡한 인지 활동을 포함하는 수행 과정 혹은 결과를 평가하는 수행평가<sup>1)</sup>가 전통적인 평가를 보완하는 관점에서 강조되고는 있으나 점수 위주의 결과가 학생들에게 보고되는 현실에는 변함이 없다.

이에 본 연구는 교사들이 수학 평가 결과를 분석하고 그 결과를 활용할 수 있는 방안을 인지 진단 이론(Cognitive Diagnosis Theory)에 터하여 소개하고자 한다. 인지과학은 측정

· 평가 분야에서 기본적인 지식과 인지 기능(cognitive skills)의 이해와 더불어 측정 내용에 대한 인지적 정보를 줄 수 있는 이론적 토대를 만들어 왔으며, 피험자에 대한 학습 정보 보완을 위하여 인지 진단 이론을 개발하였다. 인지 진단 이론은 검사에 의해 측정되는 특정 속성을 학습자가 어느 정도 이해했는지 알려줄 수 있는 이론이다. 여기서 특정 속성은 문항의 정확한 답을 하기 위해 요구되는 학습자의 능력(ability)이나 인지 과정(cognitive process), 지식(knowledge), 기능(skills) 등을 의미한다. 인지 진단 이론의 모형을 사용하면, 개개인의 학습자는 속성이 포함된 검사 문항을 통해서 각 속성을 숙달하였는지의 여부에 대한 결과를 보고받게 된다. 이것은 교사가 학생의 지식이나 기능 수준에 대한 이해를 하는데 도움이 되며 추후 수학 학습에 대한 지도에 있어서 유용한 자료가 될 뿐 아니라, 학생들은 자신의 수학 학습을 진단하고 자기 주도적 학습을 추진해 나가는 데 도움을 받게 될 것이다.

본 연구는 인지 진단 이론의 한 모형인 Fusion Model을 이용하여 수학 평가 결과를 분석하는 새로운 방법을 제안하고 중학교 3학년 학생들이 어떤 수학적 속성을 숙달하고 있는지 파악할 것이다. 그리고 학생 개개인의 수학적 속성 숙달 여부에 대한 프로파일을 작성하여 이것이 수학 교수·학습에 어떻게 활용될 수 있을지 현직 교사의 의견을 바탕으로 제안할 것이다.

1) Osterlind(2006)는 수행평가의 특성을 세 가지로 제시하면서 수행평가의 정의를 대신하였다. 첫째, 수험자로 하여금 무엇인가를 만들거나 어떤 방식으로 수행할 것을 요구하고 둘째, 복잡한 사고 기능을 측정하고, 셋째, 전문적 판단을 필요로 한다. 따라서 수학 문제 풀이를 하는 수행과정이나 수행의 결과물을 평가하는 수행검사도 수행평가의 개념으로 볼 수 있다(남명호, 2007).

## II. 인지 진단 이론

인지 진단 이론은 학생의 지식 및 기능에 대한 이해 상태를 상세하게 파악하여 학생의 학습을 돕고자 개발된 측정 이론이다. 인지 진단 이론의 목적은 개별 학생의 지식 및 기능에 대한 숙달 상태를 진단하여, 검사에 의해 측정되는 유용한 정보를 학생이나 교사, 학부모 등에게 제공하는 것이다. 이 이론은 학생 개개인이 어떤 속성을 습득했는지 알려주어 교사와 학생에게 적절한 피드백을 제공한다는 장점이 있다. 즉, 학생이 습득한 속성을 정확하게 진단하고 학생의 학습 발전 및 진전 상황을 정확하게 추정해준다(Embretson, 1990; DiBello, Stout, & Rousses, 1995; Tatsuoka, 1995).

검사를 통해 다양한 능력이나 속성을 평가할 수 있다면, 학생마다 능력이나 속성의 상태에 대한 정보를 파악하는 데 도움을 줄 것이며, 학부모와 교사에게 검사는 교육적 의사소통에 유용한 도구가 될 수 있다. 또한, 특정 속성에 대한 진단 정보는 학생으로 하여금 학습이 더 필요한 영역을 알고 비중을 두어 학습할 수 있게 만드는 계기를 제공할 수 있을 것이다.

수학·과학 성취도 추이 변화 국제비교 연구인 TIMSS(Trends in International Mathematics and Science Study)의 경우에도 인지 진단 모형을 사용하여 각국 학생들의 속성을 분석하는 연구가 진행되었으며(Tatsuoka, Corter, & Tatsuoka, 2004), Dogan & Tatsuoka(2008)는 TIMSS-R 결과를 가지고 터키와 미국 학생들의 수학적 속성 숙달 정도를 비교한 결과를 발표하기도 하였다. 그리고 김수진·송미영·김선희(2008)는 중학교 3학년 학생들의 수학 능력을 진단하기 위하여 Fusion Model을 적용하여 학생들의 속성 숙달 정도를 분석한 바 있다. 학생들에게 전반적인 능력 수준을 나타내는 총합

점수를 제공하는 것보다 자신의 지적인 강점과 약점을 알 수 있도록 특정 속성을 숙달하였는지의 여부에 대한 정보를 주는 것은 더 효과적인 학습을 계획할 수 있게 할 것이다.

본 연구에서는 Fusion Model을 이용하여 수학 평가 결과를 분석할 것이다. Fusion Model을 활용할 때는 몇 가지 절차가 수반되는데, 첫째는 학생들의 수학 능력을 잘 측정할 수 있는 문항을 개발하는 것이고, 둘째는 각 문항이 측정하는 속성을 정하는 것이다. 학생 개개인이 숙달한 속성은 문제를 푸는 데 관여된 여러 인지 행동을 포함해야 하므로 구체적이고 상세해야 한다. 이것은 문항을 개발하고 분석하는 수학교과 전문가의 몫이 될 것이며, 학생들을 평가하는 교사들이 해야 할 일이다.

셋째, 일단 속성이 정해지면 Q-행렬이 작성되는데, Q-행렬은 문항과 문항이 재고자 하는 속성 사이의 관계를 나타내는  $K \times n$  행렬이다(Tatsuoka, 1983, 1990, 1995). Q-행렬의 원소는 속성을 습득하는 것이 문항을 푸는 데 요구되면 1, 그렇지 않으면 0의 값을 가지며, 여기서  $K$ 는 측정하고자 하는 문항의 수를,  $n$ 은 속성의 수를 가리킨다. <표 II-1>은 Q-행렬의 예를 나타낸다.

<표 II-1> Q-행렬의 예

	속성1	속성2
문항1	1	0
문항2	0	1
문항3	1	1

문항 1을 풀기 위해서는 속성1이 필요하고, 문항 2를 풀기 위해서는 속성2가 필요하다. 문항 3의 경우에는 두 속성이 모두 필요하다. 학생들이 문항에 성공적으로 답하기 위해서는 그와 관련된 속성을 모두 숙달하고 있어야 한다. <표 II-1>의 경우에서 가능한 문항 반응 패턴의 예로 <표 II-2>을 예상해 볼 수 있을 것이

다. 문항 반응 패턴은 문항을 맞히면 1, 틀리면 0으로 나타낸다.

<표 II-2> 숙달한 속성에 따라 예상되는 문항 반응 패턴

숙달한 속성		문항 반응 패턴		
속성1	속성2	문항1	문항2	문항3
0	0	0	0	0
0	1	0	1	0
1	0	1	0	0
1	1	1	1	1

넷째는 Q-행렬을 바탕으로 <표 II-2>에 제시된 것과 같은 문항 반응 패턴에서 문항 모수, 각각의 속성에 대한 숙달 정도를 나타내는 학생 모수를 추정하는 것이다. 문항 모수는 두 가지로 측정되는데, 하나는 문항의 풀이에 요구되는 모든 속성을 숙달한 학생이 그 문항을 풀 때 그 속성을 정확히 적용할 확률( $\pi_i^*$ )이다. 다른 하나는 어떤 속성을 숙달하지 못하면서 그 속성이 추출된 문항에 정답을 맞힐 확률에 대한, 속성을 숙달했으면서 속성이 추출되는 문항에 정답을 맞힐 확률의 비율( $r_{ik}^*$ )로, 속성과 관련된 문항 변별도에 해당하는 것이다. 그리고 학생 모수는 학생이 각 속성을 숙달하였는지를 나타내는 수학적 속성 추정치이며, 이 추정치에 따라 속성의 숙달 여부를 결정하게 된다.

### III. 연구 방법

#### 1. 검사 도구

본 연구의 검사 도구는 문항 하나하나가 수

학적 속성을 갖고 있으면서 교육과정에 충실하게 개발된 것을 활용하고자 하였다. 이에 2003~2006년에 실시된 국가수준 학업성취도 평가<sup>2)</sup>의 문항을 활용하였다. 성취도 평가의 중학교 3학년 수학 검사에서 30개의 선다형 문항을 검사 문항으로 선정하였다. 성취도 평가는 선다형 문항과 구성형 문항(constructed response item)이 혼합된 검사이나, 각 문항 유형에서 요구되는 인지 행동이 차별화될 수 있기 때문에 (노국향, 김신영, 2000; Birenbaum & Feldman, 1998; Thissen, Wainer, & Wang, 1994) 본 연구에서는 선다형 문항만을 선택하였다. 성취도 평가의 문항은 여러 차례 수학 교사와 수학 교과 전문가들이 검토하여 개발된 것으로 예비검사를 통해 확정된 것이다. 성취도 평가의 중학교 3학년 수학 검사 문항의 내용은 교육과정에 근거한 것이며, 중학교 수학 교육과정의 전체 내용영역, 즉 수와 연산, 도형, 측정, 확률과 통계, 문자와 식, 규칙성과 함수 영역을 측정한다.

본 연구에서는 내용 영역의 범위를 수와 연산, 문자와 식, 측정, 규칙성과 함수 영역에 한정하였다. 이는 중학교 3학년 1학기까지의 교육과정으로, 본 연구가 7월에 실시되었기 때문에 검사 시기와 학습 진도를 고려하여 3학년 2학기에 학습하게 되는 도형, 확률과 통계 영역을 제외한 것이다. 단, 측정 영역은 중학교 1, 2학년 내용만 대상으로 검사문항을 선정하였다.

10개의 문항으로 구성된 문항군(A, B, C)을 3종의 검사지(가, 나, 다)에 중복되게 설계하였으며, 각 검사지에는 2개의 문항군 즉, 20문항이 포함되어 있다. 각 문항군이 다루는 내용영

2) 국가수준 학업성취도 평가는 국가 교육과정에서 규정하고 있는 교육목표에 비추어 학생의 도달도, 교육과정의 문제점과 정착 정도, 연도별 성취도 변화추이 등을 파악하기 위하여 초등학교 6학년, 중학교 3학년, 고등학교 1학년을 대상으로 매년 10월 시행되는 전국 규모의 평가이다. 이하 성취도 평가라 함.

역과 통계적 특성은 거의 동일하도록 설계하였다. 이를테면 A1, B1, C1 문항은 모두 일차방정식에 대한 문항이며, 정답률은 각각 85%, 72%, 71%로서 큰 차이가 나지 않는다. 학생들은 하나의 검사지를 할당받아 20개의 문항을 풀게 된다. 문항 설계는 [그림 III-1]과 같으며, 문항은 <부록>에 가형만 제시하였다.

가형	문항군 A (A1-A10)	문항군 B (B1-B10)	
나형		문항군 B (B1-B10)	문항군 C (C1-C10)
다형	문항군 A (A1-A10)		문항군 C (C1-C10)

[그림 III-1] 검사지 설계

## 2. 문항의 속성 추출

검사 도구에 포함된 각각의 문항은 그 문제를 풀기 위해 필요한 속성을 갖고 있다. 수학 문제를 해결하는 데 필요하면서 수학적으로 의미 있는 행동을 본 연구에서 분석할 속성으로 추출하였다. 본 연구에서는 TIMSS 2007의 인지 영역 평가틀(Mullis et al., 2005)에서 하위 요소를 선별하여 속성을 추출하였다. 연구 설계에서 문항의 내용과 통계적 특성이 유사하게 되도록 검사지를 구성하였기 때문에, 총 30개의 문항이 가진 속성의 개수를 정하는 데에는 한계가 있다. TIMSS의 평가틀은 인지 영역 내에 세부적인 하위 요소가 있어, 너무 포괄적이지 않으면서 특수한 문항에 한정되지 않도록 속성 추출을 가능하게 하는 데 도움이 된다. 이 평가틀은 문항의 설계시 행동 영역을 나타내는 것이며, 평가 결과의 분석에서 각 문항이 어떤 속성을 갖고 있는지의 근거로도 활용될 수 있다. TIMSS의 인지 영역은 알기, 적용하기, 추론하기의 3개 영역으로 구성되어 있는데 각각의

영역 내에는 <표 III-1>과 같이 하위요소가 있다. 본 연구에서는 3종의 검사지에 속한 문항에서 공통적으로 평가하고 있다고 판단할 수 있는 속성을 이 하위 요소에서 도출하였다.

<표 III-1> TIMSS 2007 인지 영역 평가틀

인지 영역	하위 요소
알기	회상하기 인식하기 계산하기 재발견하기 측정하기 분류/정렬하기
적용하기	선택하기 표현하기 모델링하기 실행하기 정형적인 문제해결 하기
추론하기	분석하기 일반화하기 종합/통합하기 정당화하기 비정형적인 문제해결하기

<표 III-1>의 인지 영역 하위 요소 중에서 본 연구의 검사 도구에 포함된 문항에 적합한 속성으로 계산하기, 모델링하기, 분석하기, 인식하기, 재발견하기를 추출하였다. 각각의 속성은 TIMSS 평가틀에서 다음과 같이 설명된다.

- 계산하기: 범자연수, 분수, 소수, 정수를 가지고 +, -, ×, ÷ 및 혼합계산과 관련된 알고리즘 절차를 수행한다. 근삿값을 취하여 계산 결과를 어렵한다. 정형적인 대수적 절차를 수행한다.
- 모델링하기: 정형적인 문제를 풀기 위해 방정식이나 다이어그램과 같은 적합한 모델을 만든다.
- 분석하기: 수학적 상황에서 변수나 대상들 사이의 관계를 결정하고 묘사하거나 사용한다. 주어진 정보로부터 타당한 추론을 한다.

- 인식하기: 수학적 대상, 모양, 수와 식을 인식한다. 수학적으로 동치인 수학적 실체를 인식한다(예를 들면, 동치 분수, 소수, 백분율; 여러 방향에서 본 간단한 도형).
- 재발견하기: 그래프, 표, 또는 다른 자료로부터 정보를 읽어낸다; 간단한 눈금을 읽는다.

성취도 평가에서는 행동 영역의 평가틀을 계산, 이해, 문제해결, 추론, 의사소통의 다섯 가지로 구분하고 있는데(조지민 외, 2006), 위의 속성은 성취도 평가의 행동영역의 내용을 보다 상세하게 한 것으로, 계산하기는 ‘계산’, 모델링하기는 ‘이해’나 ‘문제해결’, 인식하기/재발견하기는 ‘이해’, 분석하기는 ‘추론’의 범주에 속할 수 있다. 다른 중요한 수학적 속성이 문항을 통해 평가될 수 있지만, 여기서는 선정된 검사 도구에 따라 5개의 속성만을 측정해본다. 그리고 본 연구에서 사용한 문항은 중학교 교육과정에 기초한 것이므로, 위의 속성 또한 교육과정과 일관되어야 한다. 따라서 ‘계산하기’는 중학교 교육과정의 수와 연산, 문자와 식 영역에서 다루는 계산인 경우로 한정하였다.

### 3. Q-행렬 작성

Fusion Model을 활용하기 위해서는 각각의 문항과 속성을 연결하는 Q-행렬이 필요하다. Q-행렬을 작성하기 위해 먼저 연구자가 문항에 속성을 부여하였다. 이때에는 위의 5가지 속성 외에 ‘선택하기’<sup>3)</sup>의 속성이 추가되어 있었다. 이때 Fusion Model의 분석을 하여 6개 속성에 대한 문항 모수의 수렴이 적절치 않음이 나타났다. 수학 교사들과의 협의회를 통해 ‘선택하기’ 속성을 제외한 5개 속성의 Q-행렬

을 확정하였고, 문항과 속성 연결을 수정하였다.

문항에 연결된 속성의 예는 다음과 같다. [그림 III-2]는 문항번호 A4로 분석하기와 인식하기의 속성이 부여되었다. 절댓값의 개념을 ‘인식’하고, 주어진 <보기>에서 타당한 추론을 하여 옳은 것을 선택해야 하므로 ‘분석하기’의 속성을 측정한다고 합의된 것이다.

4. <보기>에서 옳은 것을 모두 고른 것은?

<보 기>

㉠ 절댓값이 3보다 작은 정수는 5개이다.

㉡ 절댓값이 같고 차가 7인 두 수 중 큰 수는 3이다.

㉢  $a$ 의 절댓값이 5이고  $b$ 의 절댓값이 3일 때,  $a+b$ 의 값 중 가장 작은 것은  $-8$ 이다.

① ㉠                      ② ㉠, ㉡                      ③ ㉠, ㉢  
 ④ ㉡, ㉢                      ⑤ ㉠, ㉡, ㉢

[그림 III-2] 문항 A4

이와 같은 방법으로 모든 문항에 속성을 추출한 것은 검사를 시행하는 중학교 교사 2명의 검토를 받아 완성되었다. 교사들의 검토 의견은 “속성의 수가 5개인 것은 충분하며, 너무 많은 속성이 있다 해도 복잡할 것”이라 하였다.

30개의 문항과 계산하기, 모델링하기, 분석하기, 인식하기, 재발견하기 등 5개의 수학적 속성들을 연결하기 위해 작성된 Q-행렬은 <표 III-2>와 같다.

3) 선택하기는 이미 알려진 절차나 방법이 있는 문제를 해결하는 데 필요한 연산, 방법, 전략을 선택하는 것을 말한다.

<표 III-2> 5개의 속성으로 구성된 Q-행렬

문항	계산하기	모델링하기	분석하기	인식하기	재발견하기	속성수
A1	1	0	1	0	0	2
A2	0	1	0	0	0	1
A3	1	1	0	0	0	2
A4	0	0	1	1	0	2
A5	0	0	1	1	1	3
A6	0	1	0	0	0	1
A7	0	0	0	1	1	2
A8	1	0	1	0	0	2
A9	0	0	1	0	1	2
A10	0	0	0	1	1	2
B1	1	0	0	0	0	1
B2	0	1	0	0	0	1
B3	1	0	0	0	0	1
B4	1	0	0	1	0	2
B5	0	1	1	0	1	3
B6	0	1	0	0	0	1
B7	0	0	1	1	1	3
B8	1	0	0	0	0	1
B9	1	0	1	0	1	3
B10	0	0	1	0	1	1
C1	1	0	0	0	0	1
C2	1	1	0	0	0	2
C3	1	1	0	0	0	2
C4	1	0	1	0	0	2
C5	0	0	1	1	0	2
C6	0	1	0	0	0	1
C7	0	0	0	1	1	2
C8	1	0	1	0	0	2
C9	0	0	1	0	1	2
C10	0	0	1	1	1	3
합계	13	10	14	9	11	57

#### 4. 검사 시행

본 연구를 위해 서울의 Y, M, P중학교와 온양에 있는 H중학교가 표집 되었고, 각 학교의 3학년 학생들이 검사에 임했다. 서울의 학교에서는 28개 학급, 온양의 학교는 8개 학급 학생들이 참여하였다. 학급마다 한 종의 검사지가 할당되어, 가형은 501명, 나형은 456명, 다형은 303명이 검사에 임했다. 검사는 2007년 7월에 실시되었고, 검사 시행에는 40분의 시간이 주어졌다. 문항당 풀이 시간을 평균 2분 정도로

보고, 중학교 현장의 교육과정 운영시 시간 단위 내에서 검사를 시행할 수 있도록 한 것이다.

#### 5. 분석 방법

Arpeggio 컴퓨터 프로그램(Hartz, Roussos, & Stout, 2002)의 MCMC (Markov Chain Monte Carlo) 추정법을 사용하여 Fusion Model의 문항 모수 및 학생 모수를 추정하였다. 그리고 검사에 참여한 학생 전체가 어떤 속성을 잘 숙달했는지 살펴보고, 학생들의 성취수준별로 숙달한 속성과 속성의 수를 조사하였다. 그리고 학생 개인별로 평가 결과를 제공할 수 있는 프로파일을 작성하였다.

MCMC 추정의 경우 모수 추정치의 자기상관계수(autocorrelaiton)를 기초로 수렴 여부를 결정하는데, 본 연구에서 추정된 문항 모수와 학생 모수의 자기상관계수가 대부분 0.2보다 적게 나타나 수렴하였다.  $\pi_i^*$  문항 모수 추정치는 0.92~0.99의 높은 값을 나타나 Q-행렬에 명시된 수학적 속성이 강하게 요구되는 문항들이라 할 수 있었다.  $r_{ik}^*$  문항 모수 추정치는 대체로 0.26~0.85로 각 문항과 관련된 각각의 속성을 잘 변별한다고 할 수 있었다. 문항 모수 추정을 통해 본 연구에 사용된 문항은 5가지 속성을 요구하면서 각 속성을 잘 변별한다고 할 수 있다.

#### 6. 교사 면담

검사의 시행과 분석이 종료된 후 본 연구의 결과를 가지고 Y, M 중학교 교사 2명과 면담을 실시하였다. 연구자가 준비한 설문을 바탕으로 반구조화된 면담 형식으로 진행되었고, 면담 시간은 30분 정도 소요되었다.

## IV. 연구 결과

본 연구에 참여한 학생들의 검사 점수는 <표 IV-1>과 같다. 검사지 가, 나, 다는 동형검사로 평균과 표준편차가 유사했다. 전체 평균은 13.9이고 표준편차는 5.4이다.

<표 IV-1> 검사지별 평균 및 표준편차

검사지	학생 수	평균	표준편차
가	501	14.8	5.2
나	456	13.4	5.7
다	303	13.1	4.9
합계	1260	13.9	5.4

다음에서는 학생들의 수학적 속성 숙달 여부를 전체와 성취수준에 따라 분석하고, 학생 개인별 프로파일을 작성하며, 이 결과를 어떻게 활용할 수 있을지 생각해 본다.

### 1. 학생들의 속성 숙달 정도

#### 가. 전체

본 연구에 참여한 1,260명이 숙달한 속성을 검사지 종별로 나타내면 <표 IV-2>와 같다.

<표 IV-2> 수학적 속성 숙달 학생 수

속성	( ) : 검사지종 내 백분율			
	가	나	다	합계
계산하기	362 (72.3)	282 (61.8)	193 (63.7)	837 (66.4)
모델링하기	294 (58.7)	256 (56.1)	144 (47.5)	694 (55.1)
분석하기	348 (69.5)	266 (58.3)	176 (58.1)	790 (62.7)
인식하기	338 (67.5)	245 (53.7)	182 (60.1)	765 (60.7)
재발견하기	329 (65.7)	269 (59.0)	150 (49.5)	748 (59.4)

중학교 3학년 학생들이 각 수학적 속성을 숙

달할 기대 확률은 47%에서 72%로 나타났다. 가, 나, 다 검사지에는 공통된 문항군이 있지만, 모든 문항이 동일한 것은 아니므로 분포가 조금씩 다르게 나타났다. 가형에서 모든 속성의 숙달 비율이 가장 높았다. 전체적으로는 계산하기, 분석하기, 인식하기, 재발견하기, 모델링하기의 순으로 학생들이 숙달한 속성이 드러났다. 이는 5개의 수학적 속성 중에 '계산하기'가 가장 쉬우며 '모델링하기'가 가장 어려운 것임을 보여준다. 중학교 교육과정의 수준에서 중학교 3학년 학생들은 66.4%가 '계산하기'를, 55.1%가 '모델링하기'를, 62.7%가 '분석하기'를, 60.7%가 '인식하기'를, 59.4%가 '재발견하기'를 숙달하고 있는 것으로 나타났다.

#### 나. 성취수준

학생들의 성취수준별로 숙달한 수학적 속성이 어떻게 다른지 살펴보았다. 상위 27%, 하위 27%를 기준으로 학생들의 성취수준을 상, 중, 하로 구분하였다. 19점 이상인 학생이 상 수준으로 353명 28.0% 있었고, 9점 이하인 325명의 학생이 25.8%로 하 수준에 해당하였다. 성취수준별로 학생들이 숙달한 속성을 조사한 것은 <표 IV-3>과 같다.

<표 IV-3> 성취수준별 속성 숙달 분포

속성	( ) : 성취수준 내 백분율			
	상	중	하	전체
계산하기	353 (100.0)	474 (81.4)	10 (3.1)	837 (66.4)
모델링하기	353 (100.0)	326 (56.0)	15 (4.6)	694 (55.1)
분석하기	353 (100.0)	430 (73.9)	7 (2.2)	790 (62.7)
인식하기	353 (100.0)	398 (68.4)	14 (4.3)	765 (60.7)
재발견하기	353 (100.0)	391 (67.2)	4 (1.2)	748 (59.4)
합계	353 (28.0)	582 (46.2)	325 (25.8)	1260 (100.0)



19점 이상의 상 수준 학생들은 5개의 속성을 모두 숙달하고 있는 것으로 나타났다. 하지만 하 수준의 학생들은 계산하기를 3.1%, 모델링하기를 4.6%, 분석하기를 2.2%, 인식하기를 4.3%, 재발견하기를 1.2%만 숙달하고 있었다. 중 수준의 학생들이 숙달한 속성은 개인별로 다양한데, 계산하기를 81.4%로 가장 많이 숙달하였고, 모델링하기가 56.0%로 가장 적게 숙달된 속성이었다.

학생들의 성취수준별로 숙달한 속성의 개수를 알아보면 <표 IV-4>와 같다.

<표 IV-4> 성취수준별 숙달한 속성의 개수  
( ): 성취수준 내 백분율

숙달한 속성의 수	상	중	하	전체
0	0 (0.0)	28 (4.8)	283 (87.1)	311 (24.7)
1	0 (0.0)	51 (8.8)	34 (10.5)	85 (6.7)
2	0 (0.0)	76 (13.1)	8 (2.5)	84 (6.7)
3	0 (0.0)	92 (15.8)	0 (0.0)	92 (7.3)
4	0 (0.0)	135 (23.2)	0 (0.0)	135 (10.7)
5	353 (100.0)	200 (34.4)	0 (0.0)	553 (43.9)
합계	353 (100.0)	582 (100.0)	325 (100.0)	1260 (100.0)

하 수준의 학생들은 숙달한 속성이 전혀 없는 경우가 많았고, 1개의 속성만 숙달하고 있는 경우가 10.5%, 2개의 속성을 숙달한 경우가 2.5%이었다. 중 수준의 학생들은 속성을 3~5개 숙달하고 있는 경우가 많았으며, 2개의 속성을 숙달한 경우가 13.1%, 1개의 속성을 숙달한 경우가 8.8% 있었다. 상, 중, 하에 따라 숙달한 속성의 개수에 차이는 있는 편이지만, 숙달한 속성의 개수가 학생들의 성취수준을 말해

주지는 않았다. 예를 들어 중 수준에서도 숙달한 속성의 개수가 0개인 학생들이 28명이나 있었고, 숙달한 속성의 개수가 5개인 학생이 200명 있었다. 따라서 수학 학습 평가 결과에서 학생들이 숙달한 속성의 개수는 점수 이외의 다른 정보를 제공한다고 볼 수 있다.

위의 평가 결과는 중학교 3학년 학생들이 전반적으로 숙달한 수학적 속성이 무엇이고, 성취수준별로 어떤 속성을 많이 숙달했는지 등 집단에 대한 평가 결과 해석을 가능하게 한다. 이제는 학생 개인별 평가 결과를 분석해 본다.

## 2. 학생 개인별 프로파일

학생 개개인의 수학적 능력에 대하여 각 속성별 학생 모수치를 추정된 결과를 예시하면 <표 IV-5>와 같다. 원점수 총점이 15점인 학생들 중에서 문항 반응 패턴이 다른 5명의 수학적 속성 추정치, 즉 개별 학생이 각 속성을 숙달할 확률에 대한 추정치를 구한 것이다. 학생들의 총점은 15점으로 동일하지만 수학적 속성 추정치가 서로 다른 것을 관찰할 수 있는데, 이는 문항 반응 패턴이 다르기 때문이다.

<표 IV-5> 총점이 15점인 학생들의 수학적 속성 추정치

이름	문항 반응 패턴	계산하기	모델링하기	분석하기	인식하기	재발견하기
박승*	1101100011111111110	.93	.09	.77	.75	.87
김진*	11110111111101010101	.87	.96	.38	.83	.71
박태*	11111101011111101100	.99	.64	.84	.87	.08
안정*	11110101011111101101	1.00	.83	.73	.79	.18
유정*	11011111011111011100	.98	.69	.62	.96	.32

위와 같은 수학적 속성 추정치를 기초로 학생 개인별 수학적 속성의 숙달 상태에 대한 프로파일을 작성하였다. 각 속성 추정치가 .50 이

상이면 주어진 속성을 숙달한 것으로 보고 1로, .50 미만이면 주어진 속성을 숙달하지 않은 것으로 판정하고 0으로 표현하였다.

Fusion Model을 이용하여 나온 학생 모수치에 따라 학생 개개인이 습득한 수학적 속성을 프로파일로 작성하였다. 이 프로파일은 각각의 기능에 대해 학습자가 숙달을 했는지 안 했는지를 나타내는 것으로, 이를 통해 평가 결과에 대한 상세한 피드백이 가능하다. 프로파일의 예는 <표 IV-6>과 같다. 각 속성에서 1은 숙달하였음을, 0은 숙달하지 않았음을 보여주는 것이다.

<표 IV-6> 수학적 속성의 프로파일

이름	총점	계산하기	모델링하기	분석하기	인식하기	재발견하기
고광*	19	1	1	1	1	1
고수*	20	1	1	1	1	1
권아*	8	0	0	0	0	0
김동*	0	0	0	0	0	0
박서*	18	1	1	1	1	1
박승*	15	1	0	1	1	1
신주*	7	0	0	0	0	0
정필*	5	0	0	0	0	0
주승*	17	1	1	1	1	0
유주*	12	1	0	0	0	0
장명*	12	0	0	1	0	0

기존의 평가 결과가 각 학생의 점수만을 보고한 반면, <표 IV-6>에서는 문항을 풀 때 사용한 수학적 속성의 숙달 여부가 나타나 학생들의 강점과 약점을 구체적으로 파악할 수 있다.

상, 하 수준의 학생들이 숙달한 수학적 속성은 수가 많거나 적은 차이가 있었지만, 중 수준의 학생들은 각자 숙달한 속성이 서로 다른 편이었다. 전체평균이 13.9이었으므로 중간에 해당하는 14점을 받은 학생의 프로파일을 M중

학교에서 살펴보기로 한다(<표 IV-7>).

<표 IV-7> M중학교에서 14점을 받은 학생들의 프로파일 분포

계산하기	모델링하기	분석하기	인식하기	재발견하기	학생수
0	0	0	1	1	1
0	1	0	0	1	1
1	1	0	0	0	2
0	1	1	1	0	3
1	0	1	0	1	3
1	0	1	1	0	1
1	0	1	1	1	1
1	1	0	1	1	1
1	1	1	0	1	1
1	1	1	1	0	4
1	1	1	1	1	1

M중학교에서 14점을 받은 학생은 모두 19명 있었다. <표 IV-7>을 보면 같은 점수라도 숙달한 속성의 프로파일이 다름을 알 수 있다. 총점이 14점으로 동일하지만 문항 반응 패턴이 학생마다 다르기 때문에 프로파일 또한 다른 것이다. 교사는 프로파일이 동일한 유형에 따라 학생들에게 피드백을 줄 수 있을 것이다. 예를 들어 프로파일이 00011<sup>4)</sup>인 학생은 정수, 유리수, 무리수의 사칙계산과 혼합 계산, 다항식의 계산 등을 능숙하게 하는 연습과, 실생활 문제에서 수학적 식을 끌어내는 문제를 많이 풀어보는 경험이 필요하며, 주어진 정보를 바탕으로 수학적 지식을 활용한 추론을 요하는 능력을 키워야 함을 교사가 제안해줄 수 있을 것이다. 이렇게 학생들에게 각각의 프로파일을 제공한다면 점수만 보고하는 것보다 많은 교육적 정보를 담게 되며, 학생들 각자 어떤 능력이 숙달되어 있고 부족한지에 대한 상세한 정보를 얻을 수 있을 것이다.

4) <표 IV-7>의 속성 순서에 따라 숙달 여부를 나타낸 것으로, 00011은 인식하기와 재발견하기의 속성을 숙달한 것으로 볼 수 있다.

### 3. 평가 결과의 활용

본 연구에서 분석된 자료가 활용될 수 있는 방법을 모색하기 위해 연구에 참여한 학생들을 지도하는 수학 교사 2명과 면담을 실시하였다. 본 연구에서 분석된 학생 개인별 프로파일이 교사가 학생에 대해 판단하는 것과 관련이 있는지, 본 연구의 결과가 어떻게 쓰일 수 있을지, 그리고 이 결과가 상용화되기 위해 어떤 것이 필요한지 알아보려 하였으며, 교사들에게 <표 IV-6>과 같은 학생 개인 프로파일을 제공하였다. 프로파일을 통한 평가 결과에 대해 교사들의 의견은 다음과 같다.

프로파일의 전반적인 내용에 대해서 교사들은 학생들의 학교 수학 성적이 우수할수록 그들이 숙달하고 있는 수학적 속성의 개수도 많은 경향이 보인다고 말하였다.

성적이 상 수준인 학생들은 수학적 기능을 많이 갖고 있는 것으로 나왔네요. 수준별 수업에서 중반과 상반을 옮겨 다니는 학생들은 뭔가 부족한 게 있는데요. 제가 부족하다고 느낀 만큼 0의 개수가 있네요.

교사들은 학생들을 관심 있게 관찰하면 수업 태도나 성실성 등에 대해 판단할 수도 있지만, 실제로 교사가 학생 개개인의 수학적 능력에 대해 자세하게 파악하고 있지 못하고 있으며 그것은 어려운 일이라고 말하고 있었다. 면담에 임한 2명의 교사 모두 학생 개개인이 어떤 능력을 갖고 있는지 잘 알지 못하였고, 관심을 끄는 학생 몇 명의 수학적 능력에 대해서 대체적인 경향을 말할 수 있을 뿐이었다.

교사들은 프로파일에 근거하여 학생의 강·약점을 되뇌는 모습을 보였다. 교사들은 프로파일을 보면서 학생 개개인의 모습을 생각해 보니 각 학생이 그러한 속성을 갖고 있다고 판단하기 쉬웠다고 한다. 이 프로파일이 평가

결과로 제공된다면, 학생 개개인의 특성을 파악하는 데 도움이 되고 학생 면담과 학부모 상담에 유용한 자료로 활용될 수 있으리라 기대하였다.

(학생의 능력이) 뭔가 부족하다는 것은 느꼈지만 무엇인지는 잘 모르겠어요. 하지만 이걸 보니 그게 무엇인지 알 것 같네요. 솔직히 교사가 학생 개개인의 세부 능력을 파악하는 것은 어려운 일이에요.

수학을 잘 하고 싶어 하는 학생이 많은데, 프로파일을 가지고 이런 부분이 부족하니 이렇게 공부하라고 조언해줄 수 있을 것 같아요. 학교에서는 답지 반응 분포를 보고 전반적인 오개념을 파악하고는 있는데, 이걸 학급이나 학년 전체에 해당되는 것이죠. 학생별로는 피드백을 줄만한 것이 점수밖에 없었는데, 이런 자료를 보니 반갑네요.

학생의 입장에서 프로파일 결과를 통해 자신의 학습 수준이 어느 정도이며 이를 개선하기 위해서 어떤 내용을 더 보완하는 학습을 해야 하고 어떤 방법을 활용할 수 있는지에 대한 정보를 제공받을 수 있어, 스스로 학습을 계획하고 실천해 갈 수 있을 것이다.

프로파일이 나오게 된 과정을 연구자가 교사들에게 설명하고 이런 결과가 상용화되기 위해 무엇이 필요한지에 대한 의견을 질문하였을 때, 교사들은 자신들이 이러한 평가 결과를 도출하기는 어려운 일이라 하였다. 프로그램이 주어진다고 하더라도 교사가 문항과 속성을 연결하는 것이 어려운 일이 될 것이라 하였다. 즉, 각 문항에 대해 무엇을 평가하고 있는지에 대한 판단을 하는 것이 어렵다는 것이다.

학교 평가에서도 이원분류표를 작성하지만, 각 문항이 지식, 이해, 적용 어떤 것에 해당하는지

(문항을) 볼 때마다 달라져요. 지식, 이해, 적용이 적절한 분류인지도 의심스럽고요.

수학 학습 목표에 비추어 각 문항이 어떤 것을 평가하고 있는지에 대한 분석이 문항을 출제할 때부터 고려되고, 교사가 문항과 속성을 연결해보는 경험과 자신감이 갖추어질 때 본 연구 결과의 현장 적용성도 높아질 것이라 여겨진다.

## V. 결론 및 제언

본 연구는 학교 교육에서 수학 성취도 평가 결과가 의미 있게 활용되고 학생과 학부모에게 유용한 교육적 정보를 제공하도록 Fusion Model을 이용하여 결과를 분석하였다. 수학적 속성은 문항을 통해 측정될 수 있었으며, 성취도 평가에 사용된 문항으로 우리나라 중학교 3학년 학생들이 어떤 속성을 많이 숙달하고 있는지, 성취수준별로 숙달한 속성과 속성의 수는 어떠한지 알 수 있었다.

본 연구 결과의 가장 중요한 내용은 학생 개인에게 무엇이 강점이고 약점인지에 대해 알려줄 수 있는 프로파일을 제공할 수 있었다는 점이다. 이 프로파일을 통해 학생은 자신의 능력을 개선시키는 데 필요한 부분을 구체적으로 인식할 수 있게 되어 보충·심화 학습을 추진해 나가는 데 도움을 받게 될 것이다. 뿐만 아니라 교사들은 학생들의 지식이나 기능 습득 정도에 대해 상세히 이해할 수 있게 되어 개별 학생들의 요구에 역점을 둘 수 있을 것이다. 평가 결과가 학생의 학습에 도움이 되기 위해서는 점수 이외에 더 많은 정보를 줄 수 있어야 하고, 교사는 학생에 대한 상세한 정보를 바탕으로 교육적 의사결정을 할 수 있어야 한

다. 따라서 검사를 실시하였을 때 그 결과가 전체 점수 이외의 다른 정보, 즉 학생들이 숙달한 속성에 대한 정보를 포함하는 것은 보다 의미 있는 평가 결과로 교육 현장에 피드백될 수 있을 것이다.

본 연구의 결과를 바탕으로 몇 가지 제언을 하고자 한다. 먼저, 이 결과가 학교 현장에서 활용되기 위해서는 후속 작업으로 Fusion Model을 프로그램화한 소프트웨어가 보급되어야 할 것이다. 학교에서 학생들의 성적을 분석하는 프로그램이나 NEIS 정보망에서 학생의 문항 단위 응답 결과가 입력되면 이러한 프로파일이 제공되는 환경을 상상해 볼 수 있을 것이다. 이를 위해서는 교육과학기술부와 교육청 단위의 행정적인 노력도 뒷받침되어야 할 것이고, 소프트웨어의 접근 용이성에 대한 연구도 뒤따라야 할 것이다.

다음으로, 문항을 해결하는 데 필요한 속성을 교사들이 연결시킬 수 있도록 전문성을 개발시켜 주어야 할 것이다. 교사는 학습목표에 따라 어떤 능력, 과정, 지식, 기능을 평가해야 하는지 결정하고, 그에 따라 문항을 개발할 수 있는 능력이 필요하다. 교사가 평가하기로 결정한 속성에서 문항의 내용과 인지 행동이 벗어나지 않고 일관성을 유지하도록 하는 것은 교육과정과 교수·학습, 평가가 일관되게 시행되는 데에도 기본이 될 것이다. 교사는 이러한 점을 염두에 두고 수업을 운영하고 평가 문항을 개발함으로써 평가 결과에 따른 수업 개선의 피드백 또한 받을 수 있을 것이다.

또한 면담에 참여한 교사의 의견처럼, 문항을 개발할 때 학생들의 어떤 속성을 평가하고 싶은지 분명하게 판단하고 그에 따라 각 문항의 특성을 파악하여 분석할 때 활용할 수 있는 능력을 교사들에게 키워주는 것이 필요하다. 문항을 개발한 후 이원분류표에서 각 문항과

인지 행동을 일대일로 대응시키는 것은 곤란한 일일 수 있다. 각 문항이 하나의 인지 행동만 다루고 있지 않기 때문이다. 본 연구에서는 문항마다 몇 개의 속성을 함께 연결시킬 수 있었으므로 이에 대한 어려움이 다소 완화될 수 있을 것이다. 하지만 문항의 특성을 결정할 수 있는 능력은 교사에게 반드시 필요한 일이다. 교사는 출제 의도에 따라 문항을 개발하고 학생들의 응답에 따른 문항의 분석과 해석 등을 할 수 있는 능력을 갖추어야 할 것이다.

교사에게 학생 평가의 전문성이 요구되고는 있지만, 현재 교사교육에서는 학생들의 평가 결과를 어떻게 분석하고 결과를 해석해야 할지에 대해 다루지 못하면서, 응당 해야 할 일로 강조만 되고 있는 듯하다. 교사들은 평가의 목적에 따라 평가를 실시하더라도 그 평가 결과를 어떻게 해석해야 할지, 수업 개선과 학생 지도에 어떻게 활용해야 할지 잘 모르고 있다. 2004년 실시된 설문조사 결과에서 우리나라 수학교사들의 22.8%는 평가 결과의 분석 및 활용 능력이 자신에게 필요하다고 여겼고, 21.8%는 평가 결과의 활용 능력이 부족하다고 하였다(이인제 외, 2004). 김선희(2006)는 수학교사 예비교육에서 실질적인 학생평가 활동이 다루어지지 못하고 있으며, 교사들과의 면담을 통해 평가 결과가 수업의 수준이나 진도의 빠르기를 조정하고, 학생들에게 평가에 대한 설명을 다시 하고, 향상된 학생들에게 칭찬을 해주고 수준이 낮은 학생들을 따로 관리하는 방법으로는 활용되지만, 평가결과에 대한 설명을 학생들이나 학부모에게 올바르게 의사소통하지 못하고 있음을 지적하면서, 이러한 현실을 보여주었다. 본 연구의 프로파일은 제공될 때 교사들이 학생들에게 그 결과를 의사소통하는 부분에 있어서도 전문성이 요구된다고 하겠다. 교사는 프로파일에 나온 속성의 숙달 여부에

따라 학생의 강점과 약점을 파악하여 후속적인 학습 지도를 하고, 그 내용을 학생과 학부모에게 잘 설명할 수 있어야 할 것이다.

본 연구의 결과는 교사들에게 가장 유용하게 활용될 수 있을 것이다. 프로파일 결과를 통해 교사는 학생들에 대하여 더 많은 정보를 얻고 그 정보를 바탕으로 수업과 평가를 연계함으로써 기존의 수업을 개선된 방향으로 발전시킬 수 있을 것이다. 학생들에게 많이 부족한 능력이 무엇인지에 초점을 두어 수업에서 강조를 하고, 학생들의 프로파일별로 그룹 또는 개별 지도를 시도할 수 있으며, 학부모와의 면담 자료로도 쓸 수 있을 것이다.

마지막으로, 본 연구는 성취도 평가의 선다형 문항만을 대상으로 검사 도구를 선정하였으나, 구성형 문항을 통해서도 학생들의 속성을 파악할 수 있을 것이다. 그리고 교사들의 질적 평가와 Fusion Model을 이용한 평가 결과 분석 사이의 관계에 대한 탐색도 후속 연구에서 이루어질 수 있으리라 본다.

## 참고문헌

- 교육인적자원부(2007). 수학과 교육과정. 교육인적자원부 고시 제 2007 - 79호 [별책 8].
- 김선희(2006). 학생평가 전문성을 갖춘 수학교사 양성을 위한 수학학습평가 강좌의 교육 내용과 방법에 대한 제안. *학교수학*, 8(3), 301-326.
- 김수진 · 송미영 · 김선희(2008). Fusion Model에 의한 수학 능력 진단을 위한 Q-행렬의 정교화. *교육평가연구*, 21(2), 115-139.
- 남명호(2007). 다시 생각해 보는 수행평가: 회고와 전망. *교육과정 운영에서의 수행평가의 적용 실태와 개선 방안*(pp.3-19). 한국교

- 육과정평가원 연구자료 ORM 2007-8.
- 노국향, 김신영(2000). 문항의 형태에 따른 피험자의 인지적·정의적 반응의 차이에 관한 연구. *교육평가연구*, 13(1), 181-194.
- 이인제, 이범홍, 박정, 진재관, 김옥남, 서수현, 김신영(2004). 교사의 학생평가 전문성 신장 모형과 기준. 한국교육과정평가원 연구보고 RRE 2004-5-2.
- 조지민·김선희·권점례·고정화·박정·김수진(2006). 2005년 국가수준 학업성취도 평가 연구 -수학. 한국교육과정평가원 연구보고 RRE 2006-1-3.
- Birenbaum, M. & Feldman, R. A. (1998). Relationships between learning patterns and attitudes toward two assessment formats. *Educational Research* 40(1), 90-97.
- Clarke, D. J. (1997). *Constructive Assessment in Mathematics: Practical Steps for Classroom Teachers*. Key Curriculum Press.
- DiBello, L., Stout, W., & Rousses, L. (1995). Unified cognitive/psychometric diagnostic assessment likelihood-based classification techniques. In P. D. Nichols, S. F. Chipman, and R. L. Brennan (Eds.), *Cognitively Diagnostic Assessment* (pp.361-389). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Dogan, E. & Tatsuoka, K. (2008). An international comparison using a diagnostic testing model: Turkish students' profile of mathematical skills on TIMSS-R. *Educational Studies in Mathematics*. 68(3), 263-272.
- Embretson, S. (1990). Diagnostic testing by measuring learning processes: Psychometric considerations for dynamic testing. In N. Frederiksen, R. L. Glasser, A. M. Lesgold, and M. G. Shafto (Eds.), *Diagnostic monitoring of skills and knowledge acquisition* (pp.453-486). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Hartz, S., Roussos, L., and Stout, W. (2002) *Skills Diagnosis: Theory and Practice. User Manual for Arpeggio software*. ETS.
- McMillan, J. H. (2004). *Classroom assessment principle and practice for effective instruction*. Boston : Allyn & Bacon.
- Mullis, I. V. S., Martin, M. O., Ruddock, G. J., O'Sullivan, C. Y., Arora, A. & Erberber, E. (2005). *TIMSS 2007 Assessment Framework*. TIMSS & PIRLS International Study Center, Lynch School of Education. Boston College.
- Osterlind, S. J. (2006). *Modern measurement: Theory, principles, and applications of mental appraisal*. Pearson Education, Inc.
- Tatsuoka, K. K. (1983). Rule space: An approach for dealing with misconceptions based on item response theory. *Journal of Educational Measurement* 20(4), 345-354.
- Tatsuoka, K. K. (1990). Toward integration of item response theory and cognitive error diagnoses. In N. Frederiksen, R. L. Glasser, A. M. Lesgold, and M. G. Shafto (Eds.), *Diagnostic monitoring of skills and knowledge acquisition* (pp.453-486). Hillsdale, NJ: Lawrence Erlbaum Associates.

- Tatsuoka, K. K. (1995). Architecture of knowledge structure and cognitive diagnosis: A statistical pattern recognition and classification approach. In P. D. Nichols, S. F. Chipman, and R. L. Brennan (Eds.), *Cognitively Diagnostic Assessment*(pp. 327-361). Hillsdale, NJ: Lawrence Erlbaum Associates.
- 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. *Educational Research Journal*. 41(4), 901-926.
- Thissen, D. Wainer, H. & Wang, X. (1994). Are Tests Comprising Both Multiple-Choice and Free-Response Items Necessarily Less Unidimensional than Multiple-Choice Tests? An Anaysis of Two Tests. *Journal of Educational Measurement*. 31(2), 113-123.

# Using Cognitive Diagnosis Theory to Analyze the Test Results of Mathematics

Kim, Sun Hee (Silla University)

Kim, Soo Jin (KICE)

Song, Mi Young (KICE)

Conventional assessments only provide a single summary score that indicates the overall performance level or achievement level of a student in a single learning area. For assessments to be more effective, test should provide useful diagnostic information in addition to single overall scores. Cognitive diagnosis modeling provides useful information by estimating individual knowledge states by assessing whether an examinee has mastered specific attributes measured by the test(Embretson, 1990; DiBello, Stout, & Rousses, 1995; Tatsuoka, 1995). Attributes are skills or cognitive processes that are required to perform correctly on a particular item.

By the results of this study, students, parents, and teachers would be able to see where a student stands with respect to mastering the attributes. Such information could be used to guide the learner and teacher toward areas requiring more study. By being able to assess where they stand in regard to the attributes that compose an item, students can plan a more effective learning path to be desired proficiency levels. It would be very helpful to the examinee if score reports can provide the scale scores as well as the skill profiles. While the scale scores are believed to provide students' math ability by reporting only one score point, the skill profiles can offer a skill level of strong, weak or mixed for each student for each skill.

\* key words : mathematical attributes(수학적 속성), cognitive diagnosis theory(인지 진단 이론), Fusion Model, Q-matrix

논문접수 : 2008. 4. 30

심사완료 : 2008. 6. 11



<부록> 가형 검사지

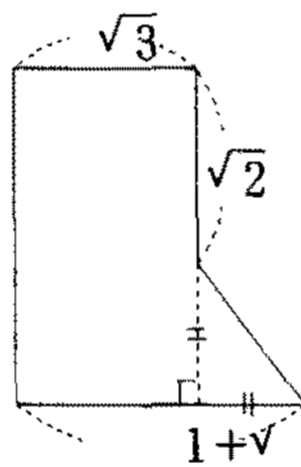
1. 다음 방정식 중에서 해가  $x=3$  인 것은?

- ①  $3x-10=1$
- ②  $2x-6=3$
- ③  $2x=4x-10$
- ④  $4(x-2)=4$
- ⑤  $x+2=3x-5$

2. 은경이는 20000원으로 500원짜리 아이스크림과 700원짜리 아이스크림을 섞어서 30개를 사려고 한다. 700원짜리 아이스크림은 최대 몇 개까지 살 수 있는가?

- ① 22개 ② 23개 ③ 24개 ④ 25개 ⑤ 26개

3. 오른쪽 그림은 직사각형과 직각이등변삼각형을 이어 붙여 만든 도형이다. 이 도형의 둘레의 길이는?



- ①  $5\sqrt{2}+\sqrt{3}$
- ②  $2+3\sqrt{5}$
- ③  $2+\sqrt{2}+2\sqrt{3}$
- ④  $2+3\sqrt{2}+2\sqrt{3}$
- ⑤  $3+3\sqrt{2}+2\sqrt{3}$

4. <보기>에서 옳은 것을 모두 고른 것은?

< 보 기 >

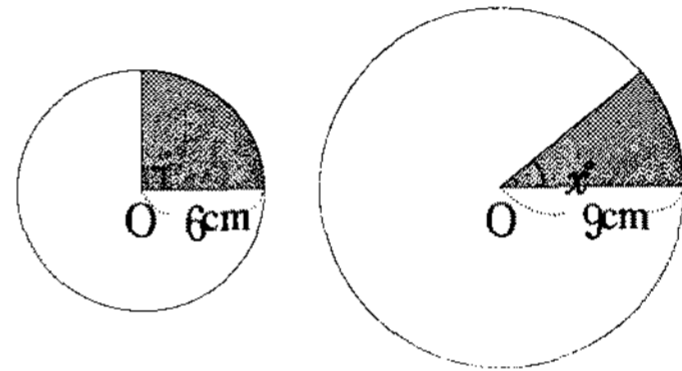
㉠ 절댓값이 3보다 작은 정수는 5개이다.  
 ㉡ 절댓값이 같고 차가 7인 두 수 중 큰 수는 3이다.  
 ㉢  $a$ 의 절댓값이 5이고  $b$ 의 절댓값이 3일 때,  $a+b$ 의 값 중 가장 작은 것은  $-8$ 이다.

- ① ㉠                      ② ㉠, ㉡                      ③ ㉠, ㉢
- ④ ㉡, ㉢                      ⑤ ㉠, ㉡, ㉢

5. 일차함수  $y=x+3$ 의 그래프에 대한 설명으로 옳지 않은 것은?

- ①  $x$ 절편은 3이다.
- ② 기울기는 1이다.
- ③  $y$ 축 위의 점  $(0, 3)$ 을 지난다.
- ④ 점  $(1, 4)$ 는 그래프 위의 점이다.
- ⑤ 일차함수  $y=x$ 의 그래프와 평행하다.

6. 반지름의 길이가 각각  $6\text{cm}$ ,  $9\text{cm}$ 인 두 원  $O$ 와  $O'$ 에서 색칠한 부채꼴의 넓이가 같을 때,  $x$ 의 값은?



- ① 20    ② 25    ③ 30    ④ 35    ⑤ 40

7. 기울기가  $\frac{3}{2}$ 인 일차함수의 그래프에서  $x$ 의 값이 1에서 5까지 증가할 때, 다음 중  $y$ 의 값의 변화를 바르게 설명한 것은?

- ① 3만큼 증가한다.
- ② 4만큼 증가한다.
- ③ 4만큼 감소한다.
- ④ 6만큼 증가한다.
- ⑤ 6만큼 감소한다.

8.  $y-x=2$ 일 때, 다음 식을  $x$ 에 관한 식으로 나타낸 것은?

$3(y-2x)+4x$

- ①  $x+6$     ②  $x-6$     ③  $5x+6$
- ④  $y+4$     ⑤  $y-4$

9.  $x, y$  에 관한 연립방정식  $\begin{cases} y=x+a \\ y=-bx-1 \end{cases}$  의 해를 구하기 위하여 각 방정식의 해를 그래프로 나타내었더니 다음 그림과 같았다.  $a$ 의 값을 구하면?

- ① 1      ② 2      ③ 4  
④ 6      ⑤ 8

10. 오른쪽 그림은 이차함수  $y=x^2-4x+k$ 의 그래프이다. 점 A, B사이의 거리가 6일 때,  $k$ 의 값은?

- ① -6      ② -5      ③ -4  
④ -3      ⑤ -2

11. 다음 일차방정식의 해는?

$$\frac{x+3}{2} = \frac{x+4}{3}$$

- ①  $x=-3$     ②  $x=-2$     ③  $x=-1$   
④  $x=0$       ⑤  $x=1$

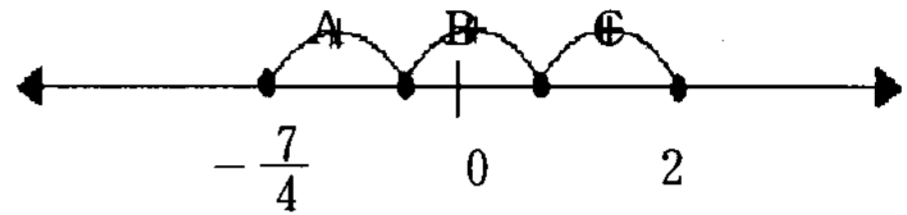
12. 어느 중학교 학생들이 수련회를 가서 짝짓기 놀이를 하였다. 5명씩, 6명씩, 8명씩 짝을 지었을 때 항상 3명이 남았다고 한다. 수련회에 참가한 학생의 수는? (단, 수련회 시설은 학생을 150명까지 수용할 수 있다.)

- ① 112명    ② 117명    ③ 120명  
④ 123명    ⑤ 132명

13.  $a=\sqrt{8}$ ,  $b=\sqrt{24}$  일 때,  $\frac{ab}{\sqrt{3}} - \sqrt{2}(a+b)$ 의 값을 구하면?

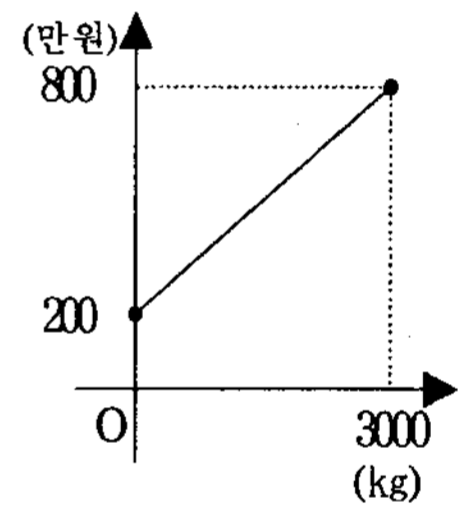
- ①  $4-4\sqrt{3}$     ②  $8-4\sqrt{3}$     ③  $4+4\sqrt{3}$   
④  $4+2\sqrt{6}$     ⑤  $\sqrt{2}+4\sqrt{3}$

14. 수직선 위에 네 점 A, B, C, D가 있고, 점 A와 점 D가 나타내는 수는 각각  $-\frac{7}{4}$  과 2이다.  $\overline{AB}=\overline{BC}=\overline{CD}$ 일 때, 점 B가 나타내는 수는?



- ①  $-\frac{1}{4}$     ②  $-\frac{1}{3}$     ③  $-\frac{1}{2}$   
④  $-\frac{2}{3}$     ⑤  $-\frac{3}{4}$

15. 오른쪽 그림은 어떤 공장에서 생산되는 물건의 생산량과 비용 사이의 관계를 그래프로 나타낸 것이다. 이 그래프에서 보



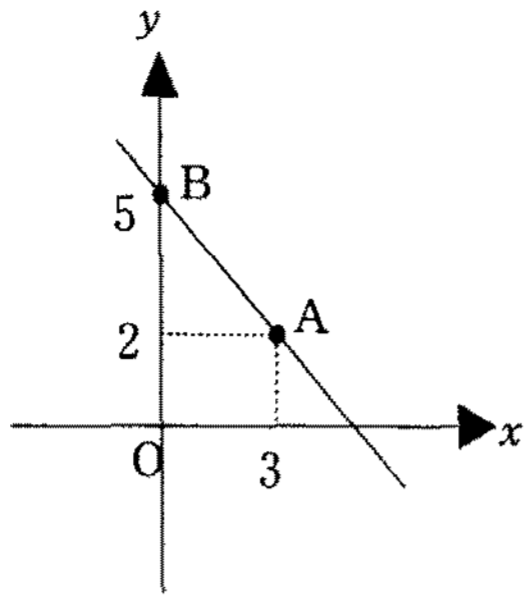
여지는 것과 같은 관계가 지속된다고 할 때, 5000kg의 물건을 생산하는 데 필요한 비용을 구하면?

- ① 900만원    ② 1000만원    ③ 1100만원  
④ 1200만원    ⑤ 1500만원

16. 원기둥을 회전축을 포함하는 평면으로 잘랐더니 그 단면이 한 변의 길이가 2cm인 정사각형이었다. 이 원기둥의 겉넓이는?

- ①  $5\pi\text{cm}^2$     ②  $6\pi\text{cm}^2$     ③  $10\pi\text{cm}^2$   
④  $12\pi\text{cm}^2$     ⑤  $16\pi\text{cm}^2$

17. 다음 그림과 같이 두 점 A, B를 지나 는 직선을 그래프로 하는 일차함수의 식은?



- ①  $y = x + 5$
- ②  $y = -x + 5$
- ③  $y = \frac{2}{3}x + 5$
- ④  $y = -\frac{2}{3}x + 5$
- ⑤  $y = -\frac{3}{2}x + 5$

18. 다음 식을 간단히 하면?

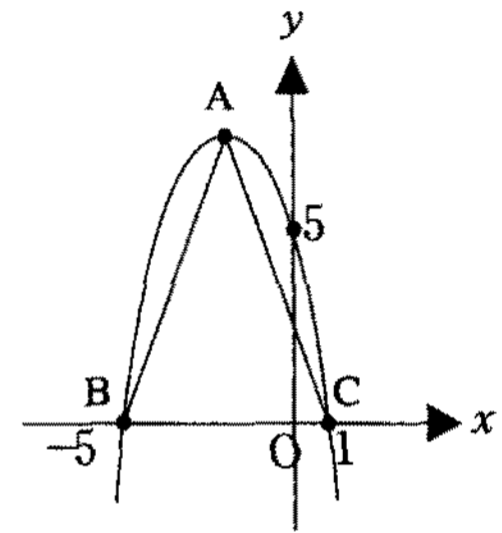
$$3a(a-2b) + (4a^3 - 6a^2b) \div (-2a)$$

- ①  $5a^2 - 9ab$
- ②  $a^2 - 3ab$
- ③  $5a^2 - 3ab$
- ④  $a^2 - 9ab$
- ⑤  $5a^2 + 3ab$

19.  $x, y$ 에 관한 연립방정식  $\begin{cases} x-2y=4 \\ 3x-y=a \end{cases}$ 의 해가  $2x-y=a$ 를 만족할 때, 주어진 연립방정식의 해를 구하면?

- ①  $x=0, y=2$
- ②  $x=0, y=-2$
- ③  $x=-2, y=0$
- ④  $x=-\frac{2}{5}, y=\frac{11}{5}$
- ⑤  $x=-\frac{2}{5}, y=-\frac{11}{5}$

20. 오른쪽 그림은 꼭 지점이 A이고,  $x$ 축과의 교점이 B(-5, 0), C(1, 0)인 이차함수의 그래프이다.  $\triangle ABC$ 의 넓이를 구하면?



- ① 18
- ② 21
- ③ 24
- ④ 27
- ⑤ 30