



인지진단이론에 근거한 TIMSS 2011의 과학 결과 분석을 통한 인지 속성의 국제비교

김지영¹, 김수진², 동효관^{2*}

¹둔촌중학교, ²한국교육과정평가원

International Comparison of Cognitive Attributes using Analysis on Science Results at TIMSS 2011 Based on the Cognitive Diagnostic Theory

Jiyoung Kim¹, Soojin Kim², Hyokwan Dong^{2*}

¹Doonchon Middle School, ²Korea Institute for Curriculum and Evaluation

ARTICLE INFO

Article history:

Received 24 January 2015

Received in revised form

27 February 2015

7 April 2015

Accepted 13 April 2015

Keywords:

scientific attributes,

Q matrix,

cognitive diagnosis theory,

cognitive attribute

ABSTRACT

This research purports to find out the characteristics of Korean students cognitive attributes and compare it with that of high-achieving countries who took TIMSS 2011 based on the Cognitive Diagnostic Theory. Based on TIMSS 2011 Science framework, nine cognitive attributes were extracted and the researcher analyzed that 216 of the TIMSS 2011 science items require these attributes. This analysis was conducted to come up with a Q-matrix. After producing the Q-matrix, multi-level IRT was used to figure out each countries' characteristics for each of the cognitive attribute. According to the study results, four attributes, such as 'Use Models,' 'Interpret Information,' 'Draw Conclusions,' and 'Evaluate and justify' were easier attributes for Korean middle school students. However, the other five attributes such as 'Recall/Recognize', 'Explain', 'Classify', 'Integrate', 'Hypothesize and Design' were considered as harder attributes compared to other countries. Korean students also considered 'Interpret Information' as the easiest attributes, and 'Explain' as the hardest attributes of all. For Korean students, those attributes considered to be easy were the easiest and hard attributes as the hardest compared to other countries, showing very extreme cases. Therefore, to give students more meaningful learning experience, it is better to use all the attributes altogether rather than use specific attributes while constructing Science curriculum or textbooks.

1. 서론

과학 교수 학습에서 평가는 궁극적으로 도달해야 할 교육 목표를 달성하기 위해 이용되는 도구로 학습자가 목표에 도달하기까지의 사고과정 중심의 평가가 이루어져야 한다(Ministry of Education, Science and Technology, 2011). 평가 결과 분석에 활용되고 있는 인지진단이론은 평가도구에 의해 측정되는 여러 가지 지식 및 기능에 대하여 학생들의 숙달 정도를 파악하고 다양한 측면에서 학생들의 학습 정도를 평가하여 교수 학습 과정에 실질적인 안내와 도움을 주고자 개발된 이론이다(DiBello *et al.*, 1995; Hartz, 2002; Tatsuoka, 1995). 인지진단이론을 적용한 평가 결과는 학생이나 학부모 및 교사에게 학생들의 학습 진행 상황에 대한 진단적인 정보를 제공할 수 있을 뿐 아니라 국가의 교육정책이나 교육과정 운영에 필요한 유용한 정보를 제공해준다.

인지 진단 이론을 활용한 평가 결과의 분석은 크게 두 가지로 나눌 수 있다. 하나는 인지진단모형을 고려하여 인지 속성을 평가하는 문항을 개발한 후 인지진단 모형을 통해 결과를 분석하는 방법이고(Fu & Li, 2007), 또 다른 하나는 이미 개발된 평가문항을 토대로 인지 속성을 추출한 후 각 문항에서 요구하는 인지 속성이 무엇인지에 해당

하는 행렬을 작성한 후 인지진단모형을 적용하는 방법이다(Hartz, 2002). 여기서 인지적 속성이란 학생들이 문항을 정확히 푸는데 필요한 능력(ability) 지식(knowledge), 기능(skills)이나 인지적 과정(cognitive process)등과 같이 다양하다(Tatsuoka, 1995). 두 가지 방법 중에서 이미 개발된 평가 문항을 토대로 결과를 분석하는 방법은 현장에서 활용가능성이 더 높은 방법이다. 실제로 국내에서 연구된 인지진단 관련 연구들을 살펴보면 모두 이미 개발되어 현장에서 시행된 평가의 결과를 활용한 연구들이 대부분이다. 학교 현장에서 시행된 중간고사나 기말고사의 결과를 인지 진단 모형으로 분석한 연구나, 국가수준의 학업성취도 결과나 대학수학능력 시험결과를 분석한 연구들이 주로 진행되었으며 이때 사용된 인지진단모형은 주로 Fusion 모형과 DINA 모형이었다(Kim *et al.*, 2008; Kim & Kim, 2009; Kim & Seong, 2011; Seong *et al.*, 2011). 이 연구들에서는 평가 과정에서 성취도 결과를 총점으로만 제시하는 현행의 결과 제시 방법에서 벗어나 인지적 숙달 여부에 대한 구체적인 정보를 제공해 줄 수 있음을 시사하고 있다.

인지진단이론에 근거한 모형을 활용하면 학생 개개인의 인지적 특성에 대한 정보를 파악할 수 있을 뿐 아니라 학습자 집단이나 국가 간의 특성을 비교하여 파악할 수 있다. 국가 간의 특성을 파악하기 위해서는 국제적으로 시행된 평가의 결과가 활용되어야 하는데 국제

* 교신저자 : 동효관 (hkdong@kice.re.kr)

** 본 논문은 2013년도 한국교육과정평가원의 연구 보고서인 'TIMSS 2011 결과에 따른 수학 과학 교육 현황 국제비교'의 일부 내용을 수정 보완한 것임.
http://dx.doi.org/10.14697/jkase.2015.35.2.0267

적인 수준에서 과학 성취를 평가하고 있는 평가로는 TIMSS와 PISA가 있다. 국제 교육성취도 평가 협회(International Association for the Evaluation of Educational Achievement)의 주관으로 시행되고 있는 수학·과학 국제 비교 평가인 TIMSS는 4년을 주기로 중학교 2학년과 초등학교 4학년을 대상으로 실시되고 있다. TIMSS는 각 국가의 공통된 교육과정을 기반으로 한 평가이므로 학생들의 수학·과학 성취를 국제적인 수준에서 비교할 수 있을 뿐 아니라 각 국가의 수학·과학 교육과정의 타당성을 점검할 수 있다는 데 그 의의가 있다. 가장 최근에 실시되어 결과가 보고된 TIMSS 2011의 과학 분야에서 우리나라 중학교 2학년 학생들은 싱가포르와 대만에 이어 3위에 해당하는 높은 성적을 나타내었고(Martin *et al.*, 2012), 그 이전 주기에서도 계속해서 우수한 성취를 나타내었다. TIMSS의 과학 성취도 평가 결과는 총점 외에 내용영역인 물리, 화학, 생물, 지구과학 및 인지요소인 알기, 적용하기, 추론하기에 대한 평가결과를 함께 제시하고 있다. 하지만 TIMSS의 평가들에서는 알기, 적용하기, 추론하기의 인지영역의 하위 요소를 총 19개로 더 세분화하여 제시하고 있는데 이러한 인지 요소와 관련된 국가별 성취결과가 구체적으로 제공되고 있지 않다. 이에 TIMSS의 평가를 이용해 국가간 인지 요소별 특성을 파악하려는 연구들이 진행되었는데 모두 수학성취도와 관련된 연구였다(Kim *et al.*, 2008; Park, 2008). 또한 TIMSS의 과학 성취도 결과를 활용한 대부분의 연구는 과학의 내용 영역별 성취 특성을 파악하거나 과학 성취가 우수한 국가의 교육과정이나 교과서를 비교하는 연구들이어서(Jeon, 2003; Jeong, 2005; Kim & Kim, 2010; Kim & Kim, 2011; Kwak & Jeong, 2007), 국제적인 수준에서 과학성취와 관련된 인지적 특성을 구체적으로 비교한 연구들이 거의 없었다.

우리나라 학생들의 과학 관련 인지적 특성을 국제적인 수준에서 분석하고 파악하는 것은 기존의 연구에서 얻을 수 없었던 새로운 정보를 제공해 줄 수 있을 것이다. 과학의 여러 인지 요소 중에서 우리나라 학생들이 어떠한 부분에서 어려움을 느끼고 있는지를 구체적으로 파악하여 교육과정이나 교과서의 구성에 반영하는 것은 의미 있는 일이라고 할 수 있다. 이에 본 연구에서는 TIMSS 2011의 과학 성취 특성을 인지진단이론에 근거하여 분석하여 우리나라 학생들의 인지적인 특성을 국제적인 수준에서 비교해보고자 한다. 이를 위해 TIMSS 2011의 평가들에 제시된 인지영역 하위 요소들을 재구조화하여 학생들이 행해야 하는 인지적 속성(attributes)을 추출하고, 이러한 인지적 속성들을 토대로 하여 국가 간 인지 특성을 파악해 보고자 한다. 우리나라 학생들이 과학에서 우수한 성취를 나타낸다고 알려져 있지만 우리나라 학생들이 TIMSS 2011 평가들의 인지 영역 중에 어떠한 부분에서 강점과 약점이 있는지를 파악하는 것은 우리나라의 과학 교수학습 과정이나 과학 교육과정의 개정이나 운영에 필요한 시사점을 도출하는 것을 가능하게 할 것이다.

II. 연구 방법 및 내용

인지진단이론을 이용해 학생의 평가 결과를 분석하기 위해 중요한 것은 타당한 인지속성을 추출하여 Q행렬을 작성하는 것이다. Q행렬은 검사 문항과 인지적 속성 간의 관계를 제시한 $k \times n$ 행렬(k =인지적 속성의 수, n =문항의 수)이다(Tatsuoka, 1983; Tatsuoka, 1990). 본 연구에서는 먼저 타당한 인지 속성을 추출하기 위해 과학 교육전문가

Table 1. Cognitive attributes extracted using cognitive elements of TIMSS 2011 science framework

추출된 인지적 속성	기존 TIMSS 2011 인지요소
회상/인식하기	회상/인식하기, 정의하기
설명하기	예를 이용하여 설명하기, 과학적 도구의 사용법 설명하기, 설명하기
분류하기	비교/대조/분류하기
모형사용하기	모델 사용하기
통합하기	관련짓기, 대답찾기, 통합/종합하기
자료 분석하기	정보 해석하기, 문제 분석하기
가설설정 및 실험설계하기	가설 설정/예상하기, 설계하기
결론도출하기	결론 도출하기, 일반화하기
평가 및 정당화하기	평가하기, 정당화하기

1인과 경력 10년 이상의 과학 교사 2인이 TIMSS 2011의 과학 평가들에 제시된 세 가지 인지영역인 알기, 적용하기, 추론하기의 총 19개의 하위요소들 살펴보고 유사한 인지 요소를 통합하고 재구조화하여 9개의 인지 속성을 추출하였다(Table 1).

1. Q 행렬 타당화

TIMSS 2011의 중학교 2학년 과학 검사 문항 총 216개에 대하여 각 문항이 9개의 인지 속성 중 어떠한 속성들을 요구하는지 판단하여 첫 번째 Q행렬을 작성하였다. Q행렬의 작성은 과학교육전문가 1명과 현장교사 2명이 함께 참여하였다. 문항별 인지속성을 판단하기 위하여 먼저 TIMSS의 평가들을 충분히 숙지한 후 문항별로 논의를 거쳐 문항에 해당하는 속성을 부여하였다. Q행렬 작성을 위한 문항 분석의 예를 살펴보면, TIMSS 2011의 공개문항인 S042304는 자료분석하기와 결론도출하기의 속성이 부여되었다(Figure 1). 과학개념이나 원리에 비추어 그래프 형태의 정보를 해석하는 것이므로 자료 분석하기의 속성이, 자료의 경향을 바탕으로 결론을 타당하게 추론하는 내용이 포함되어 있으므로 결론도출하기의 속성이 부여되었다. S042017은 모형으로 나타낸 세포 소기관을 알고 있는지 평가하고 있으므로 모형사용하기의 속성이, 생물체의 세포 소기관의 기능을 알고 있는지 평가하고 있으므로 회상/인식하기 속성이 부여되었다.

TIMSS의 과학 문항 일부에 대하여 작성한 Q 행렬은 Table 2와 같으며, 문항을 해결하는 데 필요한 인지적 속성은 1, 필요하지 않은 인지적 속성은 0으로 표기하였다.

작성한 Q행렬의 타당성을 점검하기 위해 두 가지 방법이 활용되었다. 첫 번째, 자카드 유사성 계수(Jaccard similarity coefficient)를 산출하고 두 번째, Fusion 모형에 의한 문항 모수치를 추정하였다. 자카드 계수는 Q행렬에 서로 겹치거나 유사한 인지적 속성이 있는지를 살펴 보기 위해서 산출한 것으로 그 값이 .50보다 클 경우 인지 속성 간에 유사성이 있다고 판단하여 Q-행렬을 수정하게 하는 정보를 제공한다(Hennig, 2004). Q-행렬에 기초하여 Fusion 모형에서 나온 문항난이도 모수인 π_i^* 는 문항 i 를 풀기 위해 요구되는 모든 속성을 숙달한 학생이 문항 i 를 풀 때 해당 속성을 정확히 적용할 확률이다. 속성 k 와 관련된 문항변별도 모수인 r^*_{ik} 는 속성 k 를 숙달했을 때 문항 i 에 해당 속성을 적용할 확률에 대한 속성 k 를 숙달하지 못했을 때 문항 i 에 속성 k 를 적용할 확률의 비율이다(Hartz, 2002).

자카드 계수를 산출하여 그 값이 .50보다 작는지 확인하여 인지

S042304

철수가 운동하기 전에 자신의 맥박수를 재었더니 1분에 70회였다. 철수는 1분 간 운동을 하고 다시 맥박수를 재었고, 그 후 몇 분 동안 1분마다 맥박수를 재었다. 철수는 맥박수를 측정한 결과를 그래프로 나타내었다.

이 결과로부터 얻을 수 있는 결론은?

- ① 맥박수는 1분당 50회가 증가했다.
- ② 맥박수는 빨라지는 것보다 느려질 때 더 적은 시간이 걸렸다.
- ③ 4분 후에 맥박수는 1분당 80회였다.
- ④ 맥박수가 6분 이내에 정상으로 돌아왔다.

S042017

식물 세포를 나타낸 위의 그림에서 X로 표시된 부분의 기능은?

- ① 수분을 저장한다.
- ② 영양분을 만든다.
- ③ 에너지를 흡수한다.
- ④ 세포의 활동을 조절한다.

Figure 1. Examples of TIMSS 2011 items(S042304 and S042017)

Table 2. Example of Q-matrix

문항 ID	회상/인식하기	설명하기	분류하기	모형사용하기	통합하기	자료 분석하기	가설설정 및 실험설계하기	결론도출하기	평가 및 정당화하기
S032611	1	0	0	0	0	0	0	0	0
S032614	0	1	0	0	0	0	0	0	0
S032451	0	1	1	0	0	0	0	0	0
S032156	0	0	0	0	0	0	0	1	0
S032056	0	1	0	0	1	0	0	0	0
S032087	1	0	0	0	0	0	0	0	0

속성간의 유사성이 없는지를 판단하였고 Fusion 모형의 문항모수치가 $\pi_i^* < .4$ 인 경우와 $r_{ik}^* > .9$ 인 경우에 해당하는 문항들에 대해 총 5차례의 검토와 수정을 반복하여 최종적인 Q행렬을 얻었다. 분석 과정에서는 Arpeggio Version 3.1을 사용하였다. TIMSS의 과학은 총 14개의 문항군(Block)으로 이루어져 있으며 하나의 검사지에 2개의 문항군이 포함되어 있어 한 학생은 2개의 문항군만 풀게 되므로 12개의 문항군에 대해서는 모두 결측 처리가 되어 분석이 실행되지 않아 문항군별로 참가국의 모든 학생들의 응답을 하나의 자료로 구성하여 분석을 실시하였다. 타당성 분석을 통해 최종적으로 결정된 Q행렬의 인지적 속성별 문항수는 회상/인식하기 82개, 설명하기 99개, 분류하기 35개, 모형사용하기 17개, 통합하기 31개, 자료분석하기 71개, 가설/실험설계하기 20개, 결론도출하기 18개, 평가 및 정당화하기 2개로 나타났다.

2. 인지 속성에 따른 국가별 특성 분석

인지 속성에 따른 국가별 특성을 비교하기 위해 대상 국가로 TIMSS 2011 과학 성취 상위 15개국을 선정하였다. 상위 15개국은 성취순위별로 싱가포르, 대만, 대한민국, 일본, 핀란드, 슬로베니아, 러시아연방, 홍콩, 영국, 미국, 헝가리, 호주, 이스라엘, 리투아니아, 뉴질랜드에 해당한다. 우리나라의 경우 중학교 2학년을 대상으로 한 TIMSS 2011은 2010년 12월에 시행되었으므로 평가에 참여한 학생들은 초등학교와 중학교에서 모두 7차 교육과정의 적용을 받은 학생들이다.

상위 15개국 간의 인지 속성에 따른 특성을 비교하기 위해서 인지적 속성 타당화를 거쳐 최종적인 Q행렬을 작성한 후, 인지 속성에 따른

국가별 성취 특성을 파악하기 위해서 다층 IRT 분석을 실시하였다. Q-행렬과 다층 IRT 등을 사용한 인지진단적 분석은 학생들이나 특정 능력을 보이는 집단에 좀 더 많은 피드백을 주고자 세계적으로 다양한 연구에서 많이 활용되고 있는 추세이다. Q-행렬을 활용한 인지진단모형은 수학 교과에 많이 적용되었지만 과학 교육 분야에서는 그렇지 않다. 다층 IRT를 과학 교과에도 적용하여 과학 교과의 인지적 특성을 분석하여 피드백을 제공하는 것은 과학 교육에 새로운 시사점을 제공 해줄 수 있다. 대부분의 인지진단모형 적용 연구가 학생 수준에서 이루어지는 반면에 본 연구에서는 각 국가를 분석 집단으로 하여 연구를 실시하였다. 따라서, 학생을 기준으로 분석하는 다른 인지진단모형(예, Fusion 모형 또는 DINA 모형)을 사용하였을 때보다 집단별 분석을 하기 위해 설계된 다층 IRT 모형을 사용하는 것이 추정의 오차가 작고 설명을 할 때에도 더 명확하게 할 수 있는 장점이 있다(Park, 2013). 다음은 다층 IRT에 있어서의 각 수준에 대한 설명이다.

[1 수준]

$$p(X_{ijk} = 1 | \theta_{jk}; b_{ik}) = \frac{\exp(\theta_{jk} - b_{ik})}{1 + \exp(\theta_{jk} - b_{ik})}$$

$X_{ijk} = 1$ 은 문항 i 에 대하여 k 국가에 속한 학생 j 가 바르게 응답하는 것을 나타낸다(부분점수 문항의 경우, 0, 1 = 0 그리고 2 = 1로 리코드함). θ_{jk} 는 k 국가에 속한 학생 j 의 능력 수준을 의미한다. b_{ik} 는 k 국가에 속한 학생들에게 실시된 문항 i 의 난이도 모수라고 할 수 있다.

Table 3. International comparison of cognitive attributes that are difficult to Korean students according to analysis result of multi-level IRT

구분	회상/인식하기	설명하기	분류하기	통합하기	가설설정 및 실험설계하기
싱가포르	-0.24	-0.43	0.12	-0.25	-0.45
대만	-0.04	0.14	0.19	-0.67	-0.01
대한민국	0.17	0.84	0.50	0.66	0.63
일본	0.21	-0.06	0.14	-0.28	0.21
핀란드	0.18	-0.13	-0.19	0.25	0.16
슬로베니아	-0.09	0.16	-0.10	-0.01	-0.17
러시아연방	-0.36	-0.26	-0.36	-0.07	-0.16
홍콩	0.00	-0.06	0.05	-0.33	-0.01
영국	-0.17	-0.35	-0.16	-0.07	-0.05
미국	-0.09	-0.21	-0.24	0.10	-0.29
헝가리	0.23	0.03	-0.45	-0.13	0.04
호주	0.26	-0.17	0.29	0.22	0.07
이스라엘	-0.18	0.15	-0.04	-0.23	-0.22
리투아니아	0.02	0.25	0.18	0.19	0.30
뉴질랜드	0.09	0.09	0.07	0.62	-0.06

[2수준] $\theta_{jk} = \mu_k + E_{jk}$

μ_k 는 k 국가의 평균 능력 수준이고 E_{jk} 는 평균0, 분산 1의 정규 분포를 가정한다.

[3수준] $\mu_k = \gamma_0 + U_k$
 $b_{ik} = \delta_{i1} + \sum_l w_{kl} q_{il}, \quad k = 1, 2, \dots, K,$

K는 총 국가 수이며, δ_{il} 는 국가 1(기준 집단)에 대한 문항 i의 난이도를 나타낸다. q_{il} 는 문항 i와 관련된 인지 속성 1 ($l = 1, \dots, L$)을 나타내는 변수이다. w_{kl} 는 k 국가에 대한 문항들의 난이도에 대한 인지 속성 l의 효과를 나타내는 연속변수이다; $w_{11}, \dots, w_{1L} = 0$. 각 국가 간 문항 난이도들은 국가들 간 θ 에 대한 비교 가능한 해석을 하기 위해 기준 국가의 난이도에 상대적으로 정의된다. 모형은 전체 국가들의 평균(γ_0)과 기준 국가의 문항 난이도(δ_{i1})를 고정 효과로 포함하고 있다. w_{kl} 는 각 속성에 관련된 임의 효과라고 할 수 있다. 표준화되었을 때, 이러한 임의 효과는 평균 0, 추정된 분산 τ_1^2 인 정규분포를 따른다. U_k 는 평균 0, 분산 τ_0^2 인 정규분포를 따른다.

다층 IRT 분석 결과는 인지 속성 간에 국가 간 비교가 가능한 수치로 제시되는데 그 수치의 범위는 -1에서 1사이이다. 특정 인지 속성에서 양수가 나온 경우 해당 인지 속성에 대해 그 국가의 학생들이 어려워하는 경우이고 음수가 나온 경우는 해당 인지 속성을 잘 숙달하여 쉬운 경우에 해당하므로 다층 IRT에서 계산된 수치를 통하여 인지적 속성에 대해 각 국가가 어느 정도 수준인지를 비교할 수 있다.

III. 연구 결과 및 논의

1. 우리나라 학생들에게 어려운 인지 속성의 국가별 비교

인지 속성에 따른 국가별 성취 특성을 파악하기 위해서 다층 IRT 분석을 실시한 결과 우리나라 학생들이 어려워하는 인지속성은 회상/인식하기, 설명하기, 분류하기, 통합하기, 가설설정 및 실험설계하기인 것으로 나타났다. 다층 IRT 분석결과 우리나라 학생들이 어려워하는

Table 4. Percentage of theme that students of each nation learned among science theme in TIMSS 2011

구분	전체(%)	생물(%)	화학(%)	물리(%)	지구과학(%)
싱가포르	65	63	80	83	31
대만	68	92	98	59	5
대한민국	54	38	42	79	64
일본	57	35	86	76	41
핀란드	59	35	91	60	67
슬로베니아	63	61	56	57	81
러시아연방†	-	-	-	-	-
홍콩	56	54	61	76	32
영국	87	86	91	89	83
미국	84	88	79	77	92
헝가리	83	75	98	86	75
호주	58	47	66	63	61
이스라엘	74	70	94	80	53
리투아니아	72	69	65	65	91
뉴질랜드	48	40	62	58	38

† 비교할만한 자료가 없음

인지속성에 대한 국가별 비교 결과는 Table 3과 같다.

회상/인식하기는 TIMSS 2011의 인지요소 중 회상/인식하기와 정의하기가 통합된 속성으로 과학적 사실, 관계, 과정, 개념에 대한 정확한 진술을 알고 특정 생물체나 물질 및 과정의 특징이나 성질을 아는 것과 과학적 용어의 정의를 아는 것과 관련이 있다(Cho *et al.*, 2011; Mullis *et al.*, 2009). 회상/인식하기에 대한 분석 결과를 살펴보면 과학 성취도 상위 15개국 중 7위에 해당하는 러시아 연방의 학생들이 가장 쉬워하는 것으로 나타났고 호주의 학생들이 가장 어려워하는 것으로 나타났다. 헝가리, 일본, 핀란드에 이어 우리나라 학생들도 회상/인식하기를 상대적으로 어려워하는 것으로 나타났다.

회상/인식하기에서 나타난 결과와 연관지어 살펴보기 위해 TIMSS 2011의 문항에 제시된 주제를 학습한 학생들의 비율을 분석하였다. TIMSS의 과학 평가들에서는 과학의 내용 영역을 생물, 화학, 물리, 지구과학의 네 개로 구분하고 각 영역에 여러 개의 주제를 포함하여 제시하고 있다(Mullis *et al.*, 2009). TIMSS 2011의 문항은 각 영역에 해당하는 주제에서 골고루 출제되고 있으며, 각 문항별로 어떤 영역에 해당하는 어떠한 주제에서 출제되어 있는지 그 정보가 제시되어 있다. TIMSS 2011에서는 학생들의 성취도를 배경 변인과 연결하여 분석할 수 있도록 설문을 통해 배경 변인을 조사하며, 아울러 TIMSS 2011에 참여하는 학교의 교사들을 대상으로 TIMSS 평가들의 내용 영역 요소들을 학교에서 수업을 통해 가르쳤는지를 질문하여 평가들에 제시된 과학 주제를 학습한 비율을 조사하였다. Table 4는 TIMSS 2011에 참여한 과학 성취 상위 15개국의 학생들이 평가들에 제시된 과학 주제를 학습한 비율을 나타낸 것이다.

학생들이 TIMSS 2011의 주제를 학습한 비율과 회상/인식하기를 어려워하거나 쉬워하는 정도를 비교하여 살펴보면 회상/인식하기를 쉬워하는 국가인 싱가포르, 영국, 미국, 이스라엘의 경우 TIMSS의 주제를 학습한 비율이 각각 65%, 87%, 84% 74%로 높게 나타났다. 반면 회상/인식하기를 어려워하는 국가인 호주, 일본, 핀란드, 우리나라는 TIMSS의 주제를 학습한 비율이 각각 58%, 57%, 59%, 54%로 낮은 것으로 나타났다. TIMSS 2011의 평가들에 제시된 주제를 학습한 비율이 낮은 국가가 회상/인식하기를 어려워함을 알 수 있다. 회상/인식하기는 과학적 사실, 관계, 과정, 개념에 대한 정확한 진술을 알고 이를

기억해 내야하는 속성이므로 TIMSS의 주제를 학습한 비율이 낮은 학생들일수록 관련 개념이나 사실을 알지 못해 회상/인식하기를 어려워하는 것으로 나타났음을 알 수 있다.

설명하기는 TIMSS 2011의 인지요소 중 예를 이용하여 설명하기, 과학적 도구의 사용법 설명하기, 개념에 근거하여 설명하기가 통합된 속성이다. 설명하기는 적절한 예를 이용하여 사실이나 개념을 진술하거나 과학적 실험 장치나 도구 등의 사용법을 설명하는 것과 생물체나 물질의 성질이나 구조 및 기능에 대한 지식을 증명하거나 과학 개념이나 원리, 법칙, 이론 등에 근거하여 자연 현상을 설명하는 것 등과 관련이 있다(Cho *et al.*, 2011; Mullis *et al.*, 2009). 설명하기에 대한 분석 결과 과학 성취도 상위 15개국 중 1위에 해당하는 싱가포르의 학생들이 가장 쉬워하는 것으로 나타났고 다음으로는 영국, 러시아 연방, 미국 등이 쉬워하는 것으로 나타났다. 반면 우리나라 학생들이 다른 비교국가에 비해 설명하기를 가장 어려워하는 것으로 나타났으며 다른 인지 속성에 비해 더 어려워하는 것으로 나타났다. 우리나라의 학생들이 설명하기의 속성을 어려워하는 데에는 TIMSS의 주제 중 학생들이 학습한 비율이 낮은 것이 영향을 미쳤다고 볼 수 있지만 호주, 일본, 핀란드의 경우 TIMSS의 주제를 학습한 비율이 낮았지만 설명하기는 어려워하는 속성이 아닌 것으로 나타났다.

싱가포르와 우리나라의 중학교 과학 교과서를 비교한 이전 연구에서 우리나라 교과서는 과학 개념들을 단편적으로 제시한 경우가 많고 대부분의 내용이 탐구나 활동 중심으로 구성되어 있으며 활동을 주요 개념들과 관련짓는 설명이 부족하거나 제시되지 않은 경우가 많다고 지적하였다. 반면 싱가포르의 교과서는 교과서에 제시된 설명을 읽는 것만으로도 주요 개념을 이해할 수 있도록 상세하고 구체적으로 기술되어 있다고 하였다(Kim *et al.*, 2009; Kim & Kim, 2010). 또한 미국과 우리나라 초등 과학 교과서를 비교한 연구에서도 우리나라는 미국에 비해 과학교과서에서 다루는 개념이나 지식의 양이 적을 뿐 아니라 개념 제시가 분절적이고 직접적인 설명이 많이 제시되지 않았다고 하였다(Kim & Park, 2009; Suh, 2007). 우리나라에 비해 교과서의 개념이 연계되어 제시되어 있고 구체적이고 상세하게 설명되어 있는 미국과 싱가포르 학생들은 설명하기 인지 속성을 쉽게 여기는 것으로 나타났다. 과학 교과서에 제시된 설명방식의 차이가 이러한 결과에 영향 미칠 수 있다고 판단되며 우리나라의 과학교과서도 개념을 연계적으로 이해하고 학생들이 스스로 읽고 이해할 수 있도록 서술하려는 노력이 계속되어야 하겠다. 또한 탐구 활동이 탐구로만 끝나는 것이 아니라 활동을 통해 얻은 원리나 개념에 대해 구체적으로 이해할 수 있도록 설명해줄 필요가 있으며 과학 수업에서 학생들에게 개념이나 원리 등을 이용하여 자연 현상을 설명하거나 예를 이용하여 개념을 진술하는 활동들을 자주 할 수 있도록 해주어야 하겠다.

분류하기는 TIMSS 2011의 인지요소 중 비교/대조/분류하기에 해당하는 속성이다. 분류하기는 여러 가지 생물체나 물질, 과정들 간의 유사점과 차이점을 찾아내고 주어진 특징이나 성질에 근거하여 물체, 물질, 생물체, 과정을 구별하고 분류하거나 순서를 정하는 것과 관련이 있다(Cho *et al.*, 2011; Mullis *et al.*, 2009). 분류하기에 대한 분석 결과 과학 성취도 상위 15개국 중 11위에 해당하는 헝가리의 학생들이 가장 쉬워하는 것으로 나타났고 다음으로는 러시아연방이 쉬워하는 것으로 나타났다. 반면 우리나라의 학생들이 비교국가에 비해 분류하기를 가장 어려워하는 것으로 나타났다. 분류와 관련된 속성은 TIMSS

2011의 내용 영역 평가들과 관련지어 보면 생물체의 특징에 근거하여 개체들을 분류하는 내용이나 인간과 다른 생물체의 특성을 비교하여 유사점과 차이점을 찾아내는 것이 포함된다. 또한 물질의 물리적 화학적 성질에 따라 물질을 금속이나 비금속으로 구분하는 내용이나 순물질인 원소, 화합물 및 혼합물인 균일혼합물과 불균일 혼합물 등을 서로 구별하는 것, 물리변화와 화학변화의 구분, 산과 염기의 구분, 재생가능한 자원과 재생불가능한 자원의 구분 등이 주로 포함된다(Mullis *et al.*, 2009). 우리나라의 7차 과학 교육과정에서는 산과 염기의 구분, 생물체의 분류 등을 8학년 이전까지 다루고 있지만 TIMSS의 평가들에 제시된 내용의 수준까지 다루고 있지 않으며 금속과 비금속을 물리적, 화학적 성질에 따라 구분하는 내용, 원소와 화합물의 구분, 화학변화와 물리변화의 구분, 재생 가능한 자원과 재생 불가능한 자원에 대해서는 8학년까지 거의 다루고 있지 않다(Ministry of Education, 1997). 분류와 관련된 과학 내용 지식을 많이 다루지 않은 것이 분류하기의 인지속성을 어렵게 인식하는데 영향을 미쳤을 것으로 판단된다. 2007 개정 과학 교육과정에서는 8학년까지의 교육과정에서 금속과 비금속을 물리적 성질이나 화학적 성질에 따라 구분하는 내용, 원소와 화합물의 구분 등이 포함되었지만 생물의 특성에 따른 분류나, 물리변화와 화학변화의 구분, 재생 가능한 자원과 재생 불가능한 자원에 대한 내용이 구체적으로 다루어지지 않고 있다(Ministry of Education, 2007). 특히 생물의 특성에 따른 분류 내용이 중학교 이후에는 공통교육과정에서 거의 다루어지지 않고 있어 교육과정을 구성할 때 이와 관련된 내용에 대한 고려가 필요하겠다.

통합하기는 TIMSS 2011의 인지요소 중 관련짓기, 해당찾기, 통합/종합하기가 통합된 속성으로, 과학 지식을 관찰이나 추론을 통해서 알 수 있는 생물체, 물체나 물질의 용도, 성질, 행동 특성과 관련짓는 것과 해답을 찾기 위해 과학적 관계, 방정식, 공식 등을 이용하는 것, 그리고 과학의 여러 영역에 걸친 통합적 개념이나 주제에 대한 이해를 설명한다. 과학 문제해결을 위하여 수학적 개념이나 절차를 통합하는 것과 관련이 있다(Cho *et al.*, 2011; Mullis *et al.*, 2009). 통합하기에 대한 분석 결과를 살펴보면 과학 성취도 상위 15개국 중 2위에 해당하는 대만의 학생들이 가장 쉬워하는 것으로 나타났고 다음으로는 일본의 학생들이 쉬워하는 것으로 나타났다. 반면 우리나라 학생들이 통합하기를 가장 어려워하는 것으로 나타났고 다음으로는 뉴질랜드 학생들이 어려워하는 것으로 나타났다. 우리나라의 7차와 2007 개정 과학 교육과정에서는 법교과 학습을 강조하였고, 2009개정 과학과 교육과정에서는 과학을 기술, 공학, 예술, 수학 등 다른 교과와 관련지어 통합적이고 창의적으로 사고할 수 있는 능력을 신장시키도록 하는 목표를 제시하여 통합교육을 강조하고 있다(Ministry of Education, 2007; Ministry of Education, Science and Technology, 2011). 하지만 7차와 2007개정 과학 교육과정에서는 3학년에서 10학년의 내용 체계를 4개의 영역으로 구분하여 제시하고 있고, 2009 개정에서는 내용체계를 ‘에너지와 물질’, ‘생명과 지구’로 묶어 제시하였으나 실제 내용은 에너지, 물질, 생명, 지구라는 기존의 분과적 영역으로 뚜렷이 구분할 수 있다(Bang *et al.*, 2013). 반면, 싱가포르 과학 교육과정과 교과서의 경우 주제 중심으로 내용체계가 구성되어 내용이 통합적으로 구성되어 있다(Yu, 2011). 과학과 그 외의 영역간의 통합도 중요하지만 우리나라의 과학 교육과정도 과학 영역간의 통합을 위한 구성이 필요하다고 판단되며, 과학 영역내의 통합이 이루어질 수 있는 방안에 대한

연구가 지속되어야 하겠고 이와 관련된 실제 지도방안에 대한 교사 교육이 필요하겠다. 더 나아가 과학 영역과 다른 영역에 걸친 통합적 개념이나 주제에 대해 이해하는 내용을 다루는 활동이나 여러 요인과 관련된 개념들을 이용해 문제를 해결하는 활동들도 병행해야 할 것으로 보인다.

가설설정 및 실험 설계하기는 TIMSS 2011의 인지요소 중 가설설정/예상하기와 설계하기가 통합된 속성으로, 관찰을 통해 얻은 정보나 지식을 이용해 검증 가능한 가설을 설정하는 것과 증거에 근거한 조건의 변화 예상과 관련이 있다. 또한 이 속성은 가설의 검증을 위한 탐구 설계와 탐구 절차 및 측정에 대한 의사 결정과 관련이 있다(Cho *et al.*, 2011; Mullis *et al.*, 2009). 가설설정 및 실험설계하기에 대한 분석 결과를 살펴보면 과학 성취도 상위 15개국 중 과학 성취도가 1위인 싱가포르의 학생들이 쉬워하는 것으로 나타났고 다음으로 미국의 학생들이 쉬워하는 것으로 나타났다. 반면 우리나라의 학생들이 가설설정 및 실험설계하기를 가장 어려워하는 것으로 나타났고 다음으로 리투아니아, 일본의 학생들도 가설설정 및 실험설계하기를 어려워하는 것으로 나타났다. 우리나라 학생들은 가설설정 및 실험 설계하기와 관련된 문항을 잘 풀지 못하고 있음을 알 수 있다. 우리나라의 과학 교과서에 제시된 탐구 과정 요소를 분석한 연구에서 가설 설정하기나 실험설계와 관련된 변인통제를 거의 다루고 있지 않다고 보고되고 있다(Kim *et al.*, 2008; Kim & Yeau, 2004; Yeau *et al.*, 2003). 또한 가설의 제안에서부터 가정을 검증하기 위해 변인을 통제하여 실험을 설계하고 수행하여 결론을 얻는 과정의 실험은 하나도 없다고 보고되며(Park & Cho, 2003; You & Cho, 2003), 우리나라 중학교 2학년 학생들의 과학 탐구 능력을 조사한 결과에서도 가설 설정 능력이 제일 낮은 것으로 나타났다(Kim & Yeau, 2004). 또한 미국, 영국, 캐나다, 싱가포르, 호주, 뉴질랜드 등의 과학교육과정에 나타난 탐구 기준을 우리나라와 비교한 결과에서도 외국의 교육과정에서 가설설정과 변인 통제 등의 요소를 우리나라에 비해 강조하고 있다고 지적하였다(Lee, 2006). 싱가포르와 우리나라의 중학교 과학 교과서에 제시된 탐구 활동의 특징을 비교한 결과에서도 싱가포르는 변인설정, 실험계획, 변인 통제와 같은 활동을 우리나라에 비해 훨씬 많이 하고 있는 것으로 보고되었다(Kim *et al.*, 2007). 싱가포르의 학생들이 가설설정 및 실험 설계하기 인지속성을 가장 쉽게 여기고 있는 것으로 나타났으므로 싱가포르의 교과서를 참고하여 가설설정이나 변인 통제를 비롯한 실험 설계하기 활동들이 어떻게 구성되어 있는지 분석하여 적절히 반영할 필요가 있겠다.

2. 우리나라 학생들에게 쉬운 인지 속성의 국가별 비교

인지 속성에 따른 국가별 성취 특성을 파악하기 위해서 다층 IRT 분석을 실시한 결과 우리나라 학생들이 쉬워하는 인지속성은 모형사용하기, 자료분석하기, 결론도출하기, 평가 및 정당화하기인 것으로 나타났다. 다층 IRT 분석결과 우리나라 학생들이 쉬워하는 인지속성에 대한 국가별 비교 결과는 Table 5와 같다.

모형 사용하기는 TIMSS 2011의 인지요소 중 모델사용하기에 해당하는 속성이다. 모형 사용하기는 과학 개념, 구조, 관계, 과정, 생물학적 계와 순환(먹이그물), 물리적 계와 순환(전기회로, 물의순환, 태양계, 원자구조)에 대하여 이해한 것을 다이어그램이나 모형을 사용하여

Table 5. International comparison of cognitive attributes that are easy to Korean students according to analysis result of multi-level IRT

국가	모형 사용하기	자료 분석하기	결론 도출하기	평가 및 정당화하기
싱가포르	0.09	0.13	0.01	0.06
대만	0.12	0.11	0.09	0.06
대한민국	-0.27	-0.85	-0.32	-0.12
일본	-0.02	0.32	0.07	0.03
핀란드	0.03	0.16	-0.05	0.01
슬로베니아	0.07	0.17	0.03	0.03
러시아연방	-0.05	-0.18	-0.02	-0.05
홍콩	-0.01	0.12	0.08	-0.01
영국	0.03	0.00	0.05	-0.05
미국	-0.04	-0.18	-0.15	-0.03
헝가리	-0.02	0.12	-0.04	0.00
호주	-0.07	0.00	0.09	0.02
이스라엘	0.01	0.01	-0.05	-0.02
리투아니아	0.00	0.02	0.00	-0.04
뉴질랜드	0.13	0.04	0.21	0.10

나타내는 것과 관련이 있다(Cho *et al.*, 2011; Mullis *et al.*, 2009). 모형사용하기에 대한 분석 결과 국가별로 큰 차이가 나타나지 않았지만 비교국가에 비해 우리나라의 학생들이 이 속성을 가장 쉬워하는 것으로 나타났다. 반면, 뉴질랜드의 학생들이 모형사용하기를 가장 못하는 것으로 나타났고 다음으로는 과학 성취도가 2위인 대만이 모형사용하기를 어려워하는 것으로 나타났다. 모형 사용하기는 연관적 사고나 비유적 사고와 관련된 속성이므로 우리나라 학생들이 연관적 사고나 비유적 사고를 잘하고 있음을 알 수 있다.

자료 분석하기는 TIMSS 2011의 인지요소 중 정보 해석하기와 문제 분석하기가 통합된 속성으로, 과학 개념이나 원리에 비추어 글이나 그래프 등의 정보를 해석하는 것과 관련 문제, 개념, 문제해결 단계를 결정하기 위하여 문제를 분석하는 것과 관련이 있다(Cho *et al.*, 2011; Mullis *et al.*, 2009). 자료 분석하기에 대한 분석 결과 비교국가에 비해 우리나라의 학생들이 이 속성을 가장 쉬워하는 것으로 나타났다. 총 9개의 인지 속성 중에서 자료 분석하기가 우리나라 학생들에게 가장 쉬운 속성으로 나타나, 우리나라 학생들은 자료 분석하기와 관련된 문항을 쉽게 해결하고 있음을 알 수 있다. 우리나라를 이어 러시아연방과 미국의 학생들에게 자료 분석하기가 쉬운 속성으로 나타났다. 반면, 과학 성취도 4위 국가인 일본의 학생들이 자료 분석하기를 가장 어려워하는 것으로 나타났다. 우리나라는 중학교 과학 교과서의 탐구가 자료해석의 형태가 많았고 싱가포르의 교과서에 비해 자료 처리하는 방법을 학습하거나 결론을 지지하기 위해 자료를 이용하는 법을 학습하는 빈도가 더 높다고 보고되었는데(Kim *et al.*, 2007; Kim & Yeau, 2004), 우리나라가 싱가포르보다 자료해석을 훨씬 쉬워하는 것으로 나타나 교과서의 탐구가 영향을 미쳤을 것으로 판단된다.

결론 도출하기는 TIMSS 2011의 인지요소 중 결론 도출하기와 일반화하기가 통합된 것으로, 자료의 파악과 경향의 기술, 증거를 바탕으로 한 타당한 추론과 문제나 가설에 부합하는 적절한 결론의 도출을 통해 원인과 결과에 대한 이해를 보여주는 속성이다. 또한 실험 상황 또는 주어진 조건 외에서도 적용 가능한 일반적인 결론을 도출하는 것을 포함하는 속성이다(Cho *et al.*, 2011; Mullis *et al.*, 2009). 결론도출하기에 대한 분석 결과를 살펴보면 중학교 과학 성취도 상위 15개국 중 우리나라의 학생들이 가장 쉬워하는 것으로 나타났다. 반면, 뉴질랜드

드의 학생들이 결론도출하기를 가장 어려워하는 것으로 나타났다.

한국과 미국의 과학 교과서를 비교 분석한 이전의 연구 결과 우리나라는 미국에 비해 과학 지식보다는 과학 탐구를 더 많이 포함하고 있다고 하였다(Kim & Park, 2009). 또한 우리나라의 초등학교와 중학교에서 실시되는 실험의 유형을 분석한 결과 초등학교와 중학교 모두 실험의 과정을 학생들이 스스로 계획하는 실험은 거의 이루어지지 않고 주어진 과정대로 실험을 실시하여 알고 있는 지식을 통해 예상되는 실험결과를 확인하거나 결과를 도출해내는 실험이 가장 많다고 보고되었다(Yang *et al.*, 2007). 우리나라 중학교 1학년과 2학년 교과서의 탐구 영역을 분석한 결과 관찰 자료나 이미 알고 있는 몇 가지의 구체적인 지식으로부터 포괄적인 결론을 이끌어내는 귀납적 일반화의 과정, 또는 과학적 법칙이나 이론으로부터 특정한 사실이나 법칙을 도출하는 연역적 활동에 해당하는 추리의 탐구 과정 요소가 가장 빈도가 높거나 관찰 다음으로 빈도가 높게 분석되었다(Kim & Yeau, 2004; You & Cho, 2003). 또한 초등학교 5학년 과학 교과서를 분석한 결과에서도 관찰에 이어 추리가 가장 높은 빈도로 분석되었다고 보고되었다(Yeou *et al.*, 2003). 우리나라 학생들은 주어진 과정대로 실험을 실시하여 이로부터 얻은 결과에서 결론을 도출해내는 실험을 주로 실시하므로 가설 설정 및 실험설계하기는 어려워하고 결론도출은 쉬운 것이라고 볼 수 있고 추리의 탐구 과정 요소의 활용 빈도가 높으므로 결론도출을 잘하는 것이라고 판단된다.

평가 및 정당화하기는 TIMSS 2011의 인지요소 평가하기와 정당화하기를 통합한 속성이다. 이 속성은 문제해결 전략을 개발하고 의사 결정을 위하여 장점과 단점을 검토하며 과학과 기술의 영향력을 평가하기 위해 각 요인을 고려하고 대안적 설명이나 문제해결책을 평가하는 것과 관련이 있고 증거와 과학적 이해를 사용하여 문제의 해결과 설명을 정당화하는 것과 관련이 있다(Cho *et al.*, 2011; Mullis *et al.*, 2009). 평가 및 정당화하기에 대한 분석 결과 우리나라의 학생들이 가장 쉬워하는 것으로 나타났고 뉴질랜드의 학생들이 어려워하는 것으로 나타났다. 평가 및 정당화하기는 다른 인지 속성에 비해 국가별 차이가 크지 않았다.

IV. 결론 및 제언

인지진단이론에 근거하여 국가별로 인지적 속성의 숙달 정도를 비교한 결과 우리나라 중학교 학생들이 쉬워하는 인지속성은 모형사용하기, 자료분석하기, 결론도출하기, 평가 및 정당화하기였고 어려워하는 인지속성은 회상/인식하기, 설명하기, 분류하기, 통합하기, 가설설정 및 실험설계하기였다.

TIMSS 2011에 참여한 상위 15개국 중에서 우리나라는 두 번째로 TIMSS의 평가틀에 제시된 내용 영역별 주제를 학습한 비율이 낮았으며 TIMSS의 주제를 학습한 비율이 낮은 국가들이 대체로 회상/인식하기를 어려워하는 것으로 나타났다. 우리나라는 TIMSS에 제시된 영역별 주제를 많이 다루고 있지 않은 국가에 해당되며 특히 생물영역과 화학 영역의 주제가 많이 다루어지지 않았다. 7차 과학교육과정에서 2007개정이나 2009개정 과학 교육과정으로 바뀌면서 TIMSS에 제시된 주제 중에서 다루어지지 않던 내용들이 우리나라의 과학 교육과정 내용 영역에 일부 포함되기는 했지만, 아직도 생물이나 화학 영역 등에 포함되지 않는 것들이 있다. 따라서 우리나라의 교육과정에 제시된

과학 내용 중 TIMSS에 제시된 주제에 비해 어떤 부분이 빠져있고 포함되어 있는지를 파악하고 빠져있는 부분에 대한 포함 여부를 논의할 필요가 있다.

우리나라 학생들은 가설설정 및 실험설계하기와 통합하기 및 분류하기 속성을 어려워하였는데 이것은 우리나라 학교의 과학수업에서 가설을 세우거나 실험을 설계하는 실험 수업이 많이 이루어지지 않는다는 점과 과학의 여러 영역에 걸친 통합적 개념이나 주제에 대한 내용, 생물의 분류나 물질의 분류와 같은 내용이 많이 다루어지지 않는 것 등이 영향을 미친 것이라고 볼 수 있다. 가설설정 및 실험설계와 같은 인지 요소를 활용하기 위해서는 확인실험이나 과정이 주어지는 실험이 아닌 열린 탐구 형태의 실험이 이루어져야 한다. 이러한 실험이 이루어질 수 있도록 학교에서 이루어지는 탐구 수업의 일부를 열린 탐구로 구성할 필요가 있으며 열린 탐구 활동이 시간적 제약을 받지 않도록 블록타임제 등의 활용이 용이하게 되도록 교육과정 편성이나 운영에 있어서 정책적 지원이 이루어질 필요가 있다. 또한 통합하기의 속성이 활용되기 위해서는 과학 내의 통합이나 과학과 타교과의 통합이 이루어질 수 있는 주제들을 선정하여 수업을 구성할 필요가 있다. 2009개정 교육과정에서 STEAM이 새롭게 대두되면서 많은 자료들이 개발되어 보급되었는데, 현장에서 활용이 가능하도록 수업컨설팅 지원단이나 교사 연수 등을 통해 이러한 교육에 대한 지속적인 안내와 교육이 필요하겠다. 통합을 위한 노력은 과학 교육과정 구성에서도 지속적으로 필요하며 긴 시간적 안목을 가지고 계속적으로 연구가 진행될 필요가 있다.

우리나라 학생들의 경우 가설설정 및 실험설계하기와 통합하기 및 분류하기 속성을 비교 국가 중에서 가장 어려워하는 것으로 나타난 반면 모형사용하기, 자료 분석하기, 결론도출하기, 평가 및 정당화하기는 비교국가에 비해 가장 쉽게 여기는 것으로 나타나 속성간의 차이가 크게 나타났다. 이러한 결과는 인지 속성에 해당되는 능력을 제고할 수 있도록 인지 속성을 수업에 활용하는 정도의 차이에 기인한다고 볼 수 있다. 특정한 인지 속성의 활용빈도가 매우 높고 일부 다른 인지 속성의 활용빈도가 극히 적은 경향에서 벗어나 다양한 인지 속성들이 활용될 수 있도록 하여야겠다. 즉, 과학 교육과정의 편성이나 교과서의 제작 과정에서 여러 가지 인지 속성이 적절히 활용될 수 있도록 여러 인지 속성을 활용해 해결할 수 있는 다양한 탐구나 문제 등을 포함할 필요가 있다. 과학과의 인지 속성은 과학 교육과정에 제시되고 있는 기초탐구 능력(관찰, 분류, 측정, 예상, 추리)과 통합탐구능력(문제 인식, 가설 설정, 변인 통제, 자료 변환, 자료 해석, 결론 도출, 일반화)의 여러 가지 요소와 일치한다. 일반적으로 과학교육의 목표는 과학 지식과 탐구 기능의 습득, 과학적 태도와 문제 해결력의 신장과 이를 통한 과학적 소양의 함양으로 받아들여지고 있으므로 과학 교육과정에 내용의 이해와 간단한 활동 수행 외에도 과학 탐구 기능 습득과 문제 해결에 요구되는 능력을 성취 기준으로 제시할 필요가 있다. 2009 개정 과학과 교육과정에서는 내용의 이해나 간단한 활동 수행을 성취 기준으로 제시하고 있지만 단원별로 다루어야 하는 탐구 활동이 제시되어 있지만 학습 내용과 탐구 기능이나 문제 해결을 연관지어 성취기준이 제시되지 않고 있다. 과학 탐구 기능과 문제 해결력과 관련된 과학 능력 요소들을 교육과정의 성취 기준으로 제시하는 것은 미국의 교육과정(NGSS; Next Generation Science Standard)에도 반영되고 있다. 우리나라의 과학 교육과정에서도 학습 내용을 탐구기능이나 문제

해결력과 연관지어 성취 기준을 제시함으로써 여러 가지 탐구 기능이나 문제 해결력과 관련된 과학 능력 요소들이 골고루 활용될 수 있도록 할 필요가 있다.

인지 속성별로 우수한 국가를 살펴본 결과 국가별로 우수한 속성이 다른 것으로 나타났다. 회상/인식하기는 러시아가 설명하기와 가설설정 및 실험 설계하기는 싱가포르가 분류하기는 헝가리가 통합하기는 대만이 가장 우수한 국가로 나타났다. 특히 TIMSS 2011의 과학 성취가 가장 우수한 싱가포르의 경우 우리나라 학생들이 어려워하는 인지 속성인 회상/인식하기, 설명하기, 가설설정 및 실험설계하기, 통합하기에서 모두 우수한 것으로 나타났다. 우리나라도 자료 분석하기, 결론 도출하기, 평가 및 정당화하기에서 가장 우수한 것으로 나타났다. 여러 나라의 TIMSS 2011 결과를 비교할 때 우리나라의 우수한 점을 유지하면서 다른 국가의 장점을 벤치마킹할 필요가 있다. 인지 속성별로 강점을 나타내는 국가에서 이러한 인지 속성들이 어떻게 다루어지고 있는지에 대해 교과서나 교육과정을 근거로 살펴볼 필요가 있으며, 분석 결과 중에서 우리나라의 교육에 도입이 필요한 내용을 선별하여 반영할 필요가 있다. 교과서 집필이나 교육과정의 개정에 있어서 국내 상황이나 실정을 반영해야 하므로 무조건 도입하기보다는 우리나라의 실정에 맞는 방향으로 현행의 장점을 살리고 단점을 보완하여 개선하려는 노력이 필요하다.

우리나라 학생들이 어려워하는 인지 속성에 대해 활용기회의 부족뿐 아니라 활용기회가 주어졌을 때 어떠한 어려움을 느끼고 있는지 구체적으로 분석해볼 필요가 있다. 학생들이 어려워하는 인지 속성의 활용이 이루어지는 교수 학습 과정을 질적으로 분석하여 어떠한 단계나 과정에서 학생들이 일반적으로 어려움을 느끼는지에 대한 분석이 필요한 것으로 보인다. 이러한 분석을 통해 학생들의 인지 속성을 잘 숙달 할 수 있는 수업 지도 방안에 대한 연구가 이루어질 필요가 있다.

국문요약

본 연구의 목적은 우리나라 학생들의 인지적 속성의 특징을 인지진단 이론에 근거하여 국제적인 수준에서 비교 분석하고 우리나라 교육에 주는 시사점을 찾는 것이다. TIMSS 2011의 평가틀에 근거하여 9개의 인지 속성을 추출하였고 TIMSS 2011의 과학 문항 216개에 대해 각 문항이 어떠한 인지 속성들을 요구하는지 판단하여 Q행렬을 작성하였다. 총 5차례의 검토와 수정을 반복하여 타당도를 점검하고 최종적인 Q행렬을 작성한 후, 다층 IRT 분석을 실시하여 인지 속성에 따른 국가별 특성을 파악하였다. 분석을 통해 나타난 우리나라의 인지적 속성을 TIMSS 2011의 과학성취도 상위 15개국과 비교한 결과 우리나라 학생들에게 모형사용하기, 자료분석하기, 결론도출하기, 평가 및 정당화하기는 쉬운 속성에 해당하였고 회상/인식하기, 설명하기, 분류하기, 통합하기, 가설설정 및 실험설계하기는 어려운 속성에 해당하였다. 우리나라 학생들이 가장 쉬워하는 인지 속성은 자료해석하였고 가장 어려워하는 인지 속성은 설명하기였다. 우리나라의 경우 쉬운 인지 속성의 대부분은 비교국 중에서 가장 쉽게 여기고 어려운 인지 속성은 가장 어렵게 여기는 것으로 나타나 극단적인 특성을 보였다. 따라서 특정한 인지 속성이 많이 활용되도록 하기 보다는 여러 가지 인지 속성이 골고루 활용되도록 과학 교육과정과 과학 교과서를

구성할 필요가 있겠다.

주제어 : 인지진단이론, TIMSS 2011, 과학 인지 속성

References

- Bang, D., Park, E., Yoon, H., Kim, J., Lee, Y., Park, J., Song, J., Dong, H., Shim, B., Lim, H., & Lee, H. (2013). The design of curricular framework for integrated science education based on big idea. *Journal of Korean Association for Science Education*, 33(5), 1041-1054.
- Cho, J., Kim S., Lee S., Kim, M., Ok, H., Rim, H., Park, Y., Lee, M., Han, H., & Son, S. (2011). The Trends in International Mathematics and Science Study (TIMSS 2011) : A Technical Report of the Main Survey in Korea (KICE Research Report RRE 2011-4-1). Seoul: KICE.
- DiBello, L., Stout, W., & Roussee, 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*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Fu, J., & Li, Y. (2007). Cognitively diagnostic psychometric models: An integrative review. In annual meeting of the National Council on Measurement in Education, Chicago, IL.
- Hartz, S. (2002). A Bayesian framework for the unified model for assessing cognitive abilities: Blending theory with practicality. Unpublished doctoral dissertation, University of Illinois at Urbana-Champaign.
- Hennig, C. (2004). Cluster-wise assessment of cluster stability. Unpublished research report, Department of Statistical Science, University of College London.
- Jeon, K. (2003). An Analysis of Chemistry Achievement in the Third International Mathematics and Science Study-Repeat(TIMSS-R). *Journal of Korean Association for Science Education*, 23(3), 299-307.
- Jeong, E. (2005). An Analysis of Korean Middle School Students' Achievement of Life Science in TIMSS 2003. *Journal of Korean Biology Science Society*, 33(3), 277-290.
- Kim, H., & Park, D. (2009). Elementary Science Textbook Analysis of Korea and the United States. *Journal of Science Education*, 33(2), 258-270.
- Kim, H., & Yeau, S. (2004). An Analysis on Scientific Inquiry Process in the Middle School 8th Grade Science Textbook published by 7th National Curriculum and Students' Current Abilities of Scientific Inquiry Skill. *Journal of Korean Biology Science Society*, 32(4), 390-397.
- Kim, J., & Kim S. (2009). Analysis of Test Result at Secondary Science Using Cognitive Diagnosis theory. *Journal of Korean Association for Science Education*, 29(8), 812-823.
- Kim, K. M., Park, Y. S., Choe, S. U. (2008). Analysis of Scientific Inquiry Activities in the Astronomy Section of School Science Textbooks. *Journal of Korean Earth Science Society*, 29(2), 204-217.
- Kim, K., Kim, S., Kim M., Kim S., Kang, M., Park, H., & Jung, S. (2009). Comparative analysis of curriculum and achievement characteristics between Korea and high performing countries in PISA & TIMSS (KICE Research Report RRE 2009-7-2). Seoul: KICE
- Kim, M., & Kim, K. (2010). A Comparative Study of Science Textbooks in Korea, Singapore, and Taiwan - Focus on the field of Biology -. *Journal of Korean Association for Science Education*, 30(4), 498-518.
- Kim, M., Kim, K. (2011). A Content Analysis of Biology Domain of Korean and Singaporean Textbooks Based on the TIMSS Framework. *Journal of Korean Biology Science Society*, 39(2), 217-234.
- Kim, S., & Seong, M. (2011). Diagnosis of knowledge states using large scale assessments - An application of DINA model -. *The Journal of Curriculum and Evaluation*, 14(1), 177-200.
- Kim, S., Kim S., Seong, M. (2008). Using Cognitive Diagnosis Theory to Analyze the Test Results of Mathematics. *School Mathematics*, 10(2), 259-277.
- Kim, T., Lee, J., Shin, K., Park, J., Kim, D., & Lee, S. (2007). A Comparative Study on Physics Inquiry Activities of Science Textbooks for Secondary

- School in Korea and Singapore. *Journal of Korean Association for Science Education*, 27(7), 547-558.
- Kwak, Y., & Jeong, E. Y. (2007). An analysis of Earth Science Items and Achievement in TIMSS 2003. *Journal of Korean Earth Science Society*, 28(4), 405-414.
- Lee, B. (2006). Analysis of Inquiry Standards in Foreign National Science Curricula. *Journal of Korean Association for Science Education*, 25(7), 873-883.
- Martin, M. O., Mullis, I. V. S., & Foy, P. (2012). *TIMSS 2011 International Results in Science*. MA: Boston College.
- Ministry of Education (1997). *Science Curriculum 1997-15(9)*. Seoul: Daehan Textbook Publishing.
- Ministry of Education and Human Resources Development (2007). *Science Curriculum 2007-79(9)*. Seoul: Ministry of Education and Human Resources Development.
- Ministry of education, science and technology (2011). *Science Curriculum 2011-361(9)*. Seoul: Ministry of education, science and technology.
- Mullis, I. V., Martin, M. O., Ruddock, G. J., O'Sullivan, C. Y., & Preuschoff, C. (2009). *TIMSS 2011 Assessment Framework*. MA: Boston College.
- Park, C. (2008). *A Multilevel IRT Model for Group-Level Diagnostic Assessment with Application to TIMSS*. Doctoral dissertation, University of Wisconsin-Madison.
- Park, C. (2013). *A Comparison of the Methods of Estimating Intraclass Correlation Coefficients for Hierarchically Structured Dichotomous-Item Data*. *Journal of Educational Evaluation*, 26(2), 459-476.
- Park, H., & Cho, H. (2003). Analyses of Scientific Inquiry in Science 8. *Journal of Korean Association for Science Education*, 23(3), 239-245.
- Song, M., Lee, Y., & Park, Y. (2011). *An analysis and score reporting based on cognitive diagnostic models using the National Assessment of Educational Achievement*. (KICE Research Report RRE-2011-8). Seoul: KICE.
- Suh, Y. (2007). A Comparative Study on Elementary Science Textbooks in Korea and the U.S. : Focusing on 3rd Grade Scientific Concepts and Inquiry Process in 'Matter' Units. *Journal of Elementary Science Education*, 26(5), 509-524.
- 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*. 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*. Hillsdale, NJ: Lawrence.
- Yang, I., Kim, S., & Cho, H. (2007). Analysis of the Types of Laboratory Instruction in Elementary and Secondary Schools Science. *Journal of Korean Association for Science Education*, 27(3), 235-241.
- Yeau, S., Kim, H., & Kim, M. (2003). An Analysis on Scientific Inquiry Process in the Elementary 5th Grade Science Textbook Published by 7th National Curriculum and Students' Current Abilities of Scientific Inquiry Skill. *Journal of Korean Biology Science Society*, 31(3), 214-223.
- You, M., & Cho, H. (2003). Analysis of Scientific Inquiry in Science VII. *Journal of Korean Association for Science Education*, 23(5), 494-504.
- Yu, S. A. (2011). An analysis of Singaporean primary-school science textbooks as a preliminary study for the development of grade-cluster textbooks in the 2009 revised Korean National Curriculum system. *The Journal of Curriculum Studies*, 29(3), 147-171.