



순위 선다형 문항을 이용한 천문 시스템 학습 발달과정 개발 및 타당화 연구

맹승호¹, 이기영^{1*}, 박영신², 이정아³, 오현석¹

¹강원대학교, ²조선대학교, ³서울대학교

Development and Validation of a Learning Progression for Astronomical Systems Using Ordered Multiple-Choice Items

Seungho Maeng¹, Kiyong Lee^{1*}, Young-Shin Park², Jeong-A Lee³, Hyunseok Oh¹

¹Kangwon National University, ²Chosun University, ³Seoul National University

ARTICLE INFO

Article history:

Received 26 August 2014

Received in revised form

21 December 2014

Accepted 26 December 2014

Keywords:

learning progressions,
astronomical systems,
astronomical thinking,
Rasch model, validity evidence

ABSTRACT

This study sought to investigate learning progressions for astronomical systems which synthesized the motion and structure of Earth, Earth-Moon system, solar system, and the universe. For this purpose we developed ordered multiple-choice items, applied them to elementary and middle school students, and provided validity evidence based on the consequence of assessment for interpretation of learning progressions. The study was conducted according to construct modeling approach. The results showed that the OMCs were appropriate for investigating learning progressions on astronomical systems, i.e., based on item fit analysis, students' responses to items were consistent with the measurement of Rasch model. Wright map analysis also represented that the assessment items were very effective in examining students' hypothetical pathways of development of understanding astronomical systems. At the lower anchor of the learning progression, while students perceived the change of location and direction of celestial bodies with only two-dimensional earth-based view, they failed to connect the locations of celestial bodies with Earth-Moon system model, and they could not recognize simple patterns of planets in the solar system and milky way. At the intermediate levels, students interpreted celestial motion using the model of Earth rotation and revolution, Earth-Moon system, and solar system with space-based view, and they could also relate the elements of astronomical structures with the models. At the upper anchor, students showed the perspective change between space-based view and earth-based view, and applied it to celestial motion of astronomical systems, and they understood the correlation among sub-elements of astronomical systems and applied it to the system model.

1. 서론

이 연구는 초중등 지구과학 교과의 핵심 주제 중의 하나인 천문 시스템의 이해에 대한 평가 문항을 개발하고, 이를 초등학교생들과 중학교생들에게 적용한 검사 결과를 통해 천문 시스템의 학습 발달과정을 조사하여 그 타당성 검증 과정을 제시하였다. 이 연구에서 천문 시스템(astronomical systems)은 '지구, 지구-달 계, 태양계, 그리고 은하와 우주를 포함하여 천체의 구조와 운동을 종합'하여 지칭한 것이다. 또한, 학습 발달과정(learning progressions)은 "학생들이 한 주제에 대해 오랜 기간 동안 탐구하고 학습함에 따라 그 주제에 대한 이해 및 관련된 사고 능력이 순차적으로 점점 더 정교해지는 경로들을 기술한 것"(National Research Council [NRC], 2007, p. 219)으로서, '학습을 위한 평가'(assessment for learning, Black & Wiliam, 1998)의 관점에서 평가의 결과를 바탕으로 과학 교사의 교수활동 및 과학 교육과정을 일관성 있게 구성하는데 중요한 역할을 할 수 있어서 최근 과학교육 연구뿐만 아니라 국가 과학교육 정책을 입안하는 과정에서도 중요하게 논의되고 있다(e.g., Next Generation Science Standards [NGSS], Achieve Inc., 2013; Framework for K-12 Science Education [NRC

Framework], NRC, 2012).

천체의 운동은 지구과학 교과의 핵심 주제 중의 하나로서, 학생들이 움직이는 지구에 있는 관측자로서 여러 천문 현상을 설명하고, 직접 관찰할 수 없는 지구의 운동을 개념적으로 이해하기 위한 필수적인 내용이다(Plummer & Krajcik, 2010). 천체의 운동은 NRC Framework와 NGSS에서 지구우주과학(Earth and Space Science)의 3대 핵심 주제(disciplinary core idea) 중의 하나인 우주에서 지구의 위치(Earth's place in the Universe)의 주요 내용으로 선정되었으며, NGSS에 기반한 새로운 교육기준(standards) 개발을 위한 지구우주과학의 5가지 표준 주제(standard topics)의 첫 번째인 우주 시스템(space system)의 중심적 내용에 해당한다. 우리나라의 2009개정 과학과 교육과정(교육과학기술부, 2011)에서도 천체의 운동은 핵심적인 주제로서 5-6학년 군에서 '지구와 달의 운동' 단원과 '태양계와 별 단원', 그리고 7-9학년 군에서 '태양계' 단원과 '외권과 우주개발' 단원이 포함되었다.

미국의 NGSS에서는 천체의 운동과 관련된 기초공통 개념(crosscutting concepts)으로서 패턴, 규모(scale), 비율(proportion) 및 정량화, 안정성과 변화를 이해하는 것이 중요함을 밝힌 바 있다(Achieve Inc., 2013). 패턴은 3차원 공간의 우주에서 형성되는 천체의

* 교신저자 : 이기영 (leeky@kangwon.ac.kr)

** 이 논문은 2013년 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (NRF-2013S1A5A2A03045044).

http://dx.doi.org/10.14697/jkase.2014.34.8.0703

운동에 대한 메커니즘과 원리를 인과적으로 설명할 수 있는 모델을 파악할 때 필요하다. 그리고 천체의 운동은 시간적, 공간적 규모가 매우 큰 조건에서 발생하는 현상이므로 운동의 규모와 비율을 인식하고 이를 정량화하는 것이 매우 중요하다. 또한, 천체의 운동이 나타내는 현상은 전체적인 규모를 한 눈에 관찰하여 이해하기 쉽지 않으며, 하나의 천체만이 아니라 여러 천체들의 운동이 시스템적으로 관련되어 형성되는 것이므로, 천체의 운동을 이해하려면 천체들 간의 구조에 대한 이해가 필요하다. 이에 따라 이 연구에서는 학습 발달과정의 중심 주제로서 천체의 운동과 천체의 구조를 종합적으로 고려한 천문 시스템을 선정하였다.

2007년 미국의 과학 학습에 대한 종합 연구 보고서 *Taking Science to School* (NRC, 2007)에서 공식적으로 제기되었던 학습 발달과정 연구는 이후 Consortium for Policy Research in Education이 발표한 보고서 *Learning Progressions in Science* (Corcoran, Mosher, & Rogat, 2009)에서 그 의미와 과학교육적 의의가 구체화되었고, 현재까지 미국을 중심으로 다양한 연구가 진행되고 있으며, 독일, 중국, 한국에서도 일부 연구자들에 의해 연구의 영역이 넓혀지고 있다. 학습 발달과정의 선행 연구들은 물리학(e.g., Alonzo & Steedle, 2009), 화학(e.g., Stevens, Delgado, & Krajcik, 2010), 생물학(e.g., Duncan *et al.*, 2009; Furtak, 2012), 지구과학(e.g., Plummer & Krajcik, 2010)의 핵심적 주제들을 다루고 있을 뿐 아니라, 과학적 모델링(Schwarz *et al.*, 2009)이나 과학 논증활동(Berland & McNeill, 2012), 증거기반 추론활동(Songer & Gotwals, 2012) 등 과학의 실행(science practices) 영역에 대한 발달과정을 조사하기도 하였다. 이러한 연구의 경향에 대하여 Duschl, Maeng, and Sezen (2011)은 학습 발달과정 연구는 과학의 기초적인 지식(foundational knowledge)을 과학적 맥락에 활용하는 능력을 다루어야 하며, 이것을 과학의 탐구실행과 통합적으로 구성하여 진행되어야 함을 주장한 바 있다. 과학의 내용지식과 과학적 탐구실행의 통합적인 학습 발달과정에 대한 문제의식은 이후 Black *et al.* (2011), Shea and Duncan (2013), Gotwals and Songer (2013) 등에 의해 지속적으로 논의되어 왔다. 이것은 과학적 탐구실행에 참여하여 과학의 문화를 체득하는 것을 과학적 과정으로서 중요하게 평가하는 현대 과학교육의 철학적 지향(Duschl & Grandy, 2008)과도 부합하는 것으로 볼 수 있다.

이 연구에서 조사하는 천문 시스템의 학습 발달과정은 과학 내용지식으로서 천체의 운동과 구조를 다루는 것과 함께, 학생들이 천체의 운동과 구조를 효과적으로 학습하기 위한 과학적 탐구실행으로서 천문학적 사고(astronomical thinking)를 포함하였다. 이 연구에서 천문학적 사고는 공간적 사고 (spatial thinking)와 시스템 사고 (systems thinking)로 구성된다. 공간적 사고는 학생들이 천문 현상에 대한 이해를 증진시키고 천문 현상과 관련된 과학적 활동에 참여를 확장시킬 수 있는 근본적 토대가 된다(Plummer, 2014).

공간적 사고에 대한 중요한 보고서 *Learning to Think Spatially* (NRC, 2006)에서는 “공간에 대한 의미를 이해하는 공간적 개념, 사물의 공간적인 형태나 구조를 시각화하는 방식, 그리고 공간적 표현 방식을 활용하여 사물의 구조와 성질, 기능 등을 이해하고 설명하는 추론 과정까지 포함하는 종합적인 사고 능력”으로서 공간적 사고를 정의하고 있다(p. 3). 공간적 사고는 또한, 사물의 공간적 관계를 시각화하기, 공간적 크기/방향의 변화를 상상하기, 및 사물을 상상하여 회전시키기

등과 같은 내적인 인지적 과정과 이것을 다양한 표상으로 외적으로 표현하는 능력을 포함한다(NRC, 2006). 학생들이 공간적 사고 활동에 참여하는 것은 공간 능력(spatial ability)뿐만 아니라, 학생들이 공간 능력을 사용할 때 동반하는 과학 지식 및 관련된 경험에 따라 달라질 수 있다(Plummer, 2014). 선행 연구들에서 규명된 공간적 사고에 대한 세부적인 사항들을 바탕으로 이 연구에서는 학생들이 천문 시스템을 학습할 때 활용되는 공간적 사고로서 ‘천체의 위치 변화와 지리적 방향을 판단하는 능력, 2차원의 평면 자료를 3차원의 공간적 자료로 또는 그 반대로 서로 전환하는 능력, 지구에서 보는 관점에서 파악한 천문 현상을 우주에서 보는 관점에서 재구성하고 이것을 천체의 상대적 운동으로 파악하는 능력 등을 포함하는 것’으로 정의하였다. 공간적 사고에 대한 이러한 정의는 천문 시스템의 이해 중에서 주로 천체의 운동을 이해하고 설명하는 과정에서 주로 적용된다.

시스템 사고는 “천체 시스템을 구성하는 어느 한 부분의 작동이나 변화 또는 특정한 기능이 어떻게 전체 시스템에 영향을 주어 전체 시스템이 작동하게 되는지 이해하는 능력으로서, 시스템 작동 과정에서 서로 다른 각각의 요소들이 상호작용하는 것에 대한 판단과 의사결정, 시스템 분석, 시스템 평가 및 추상적인 추론 과정을 포함한다.”(NRC, 2010, p. 3). 시스템 사고와 관련하여 Ben-Zvi Assaraf and Orion (2005)은 하나의 시스템을 이해하는데 필요한 시스템 사고의 특성을 “한 시스템의 구성 요소와 그 시스템에서 발생하는 현상들의 과정을 파악하는 능력(A), 시스템의 구성 요소들 간의 간단한 관계를 파악하는 능력(B), 시스템 내의 역동적 관계를 파악하는 능력(C), 시스템의 구성요소와 과정 및 그 상호작용을 연관성의 측면에서 조직하는 능력(D), 시스템 내에서 물질과 에너지의 순환을 파악하는 능력(E), 자료의 패턴과 상호관련성을 바탕으로 표면적으로 드러나지 않는 시스템의 내적 특성을 인지하는 능력(F), 시스템의 메커니즘을 이해하고 일반화하는 능력(G), 미래지향적/과거역산적 추론을 통한 시간적 사고 능력(H)”(p. 523) 등의 8가지 능력으로 구분하여 정리하였다. 이들은 후속 연구에서 시스템 사고에 포함되는 각각의 사고 특성들 중에서 사고특성 A는 시스템의 구성요소를 ‘분석하기’에 해당하고 사고특성 B~E는 시스템 구성요소를 ‘종합하기’에 해당하며, 사고특성 F~H는 시스템을 ‘실행하기’에 해당하는 것으로 구분하였다(Ben-Zvi Assaraf & Orion, 2010). 또한, 시스템 사고는 시스템 구성요소들에 대한 분석과 종합, 그리고 시스템적 실행이 위계적 관계를 형성하며 순차적으로 발달한다(Ben-Zvi Assaraf & Orion, 2005; Orion & Basis, 2008)고 주장하였다. 천체의 운동 및 구조와 관련된 현상들은 작은 규모에서 형성된 부분적인 결과를 종합하여 전체적인 큰 규모에 해당하는 설명 모델로서 파악되는 경우가 많으며, 여러 천체들의 역학적 관계가 시스템적으로 이루어져 형성된다. 따라서 천체의 운동과 그 구조를 이해하기 위해서는 천문 시스템의 시스템적 특징을 모형화하는 시스템 모델링 및 그와 관련된 시스템적 사고가 중요하게 요구된다. Ben-Zvi Assaraf & Orion이 분류한 시스템 사고의 특성들은 천문 시스템에서 천체들 간의 구조를 이해하는 과정에서 부분적으로 재구성되어 적용될 수 있다. 예를 들어 천문 시스템의 하위 시스템으로서 태양계를 이해하려면, 태양계를 구성하는 하위 요소로서 행성들을 구별하고 행성을 관측하여 행성의 공전을 파악하고(사고특성 A), 태양계를 구성하는 행성, 위성, 소행성들 간의 관계를 파악하고, 각 천체들 간의 역학적 상호작용을 이해해야(사고특성 B, C, D) 하며, 지구에서 행성들을 관

측할 때 각 행성의 상대적인 위치의 변화에서 행성 운동의 패턴을 찾고 각 행성의 패턴 결과를 종합하여 태양 중심의 공전 궤도 모델을 수립하고 설명할(사고특성 F, G) 수 있어야 한다. 이와 같은 맥락에서 이 연구는 학생들이 천문 시스템을 학습할 때 활용되는 시스템 사고를 ‘천문 시스템 내에서 천문현상의 패턴을 이해하고, 시스템 구성요소들 간의 상호관계를 이해하며, 각 구성요소들의 부분적인 정보를 전체적으로 종합하여 시스템의 규모와 비율 및 변화를 이해하고 설명할 수 있는 모델을 구성하는 과정’으로 정의하였다. 시스템 사고에 대한 이와 같은 정의는 천문 시스템의 이해 중에서 주로 천문 시스템의 구조를 이해하고 설명하는 과정에서 주로 적용된다.

천문학적 사고의 수행을 통한 천문 시스템의 학습 발달과정은 과학자나 과학교육자, 또는 과학 정책 입안자들의 전문성에 근거하여 학습 주제의 위계나 학습 내용의 순서를 선형적으로 예단하거나 결정한 것이 아니다. 학습 발달과정은 “경험적 연구에 근거하고, 학습 평가를 통해 검증할 수 있는 가설적 발달경로를 의미한다.”(Corcoran *et al.*, 2009). 따라서 학습 발달과정을 조사하고 기술하는 과정에서 학생들의 학습 수준을 정확하게 측정하기 위한 평가 도구를 고안하고, 그것의 타당성을 검증하는 것이 매우 중요하게 제기된다(Wilson, 2009; Maeng, Seong, & Jang, 2013). 아울러 학습 발달과정 연구에서 타당성 검증(validation)은 단지 검사도구인 평가 문항의 내용 타당성(content validity) 검증에 국한된 것이 아니라, “제한된 가설적인 학습 발달 경로의 적절성을 검증하기 위한 증거를 얼마나 모을 수 있는가, 그리고 제한된 학습의 발달 경로에 대한 가설이 실제로 학습자의 발달 경로에 대한 결과를 산출하는데 효과적이냐”(Corcoran *et al.*, 2009, p. 41)를 검증할 수 있어야 한다. 즉, 학습 발달과정의 타당성 검증은 학습평가의 결과를 통해 학습자가 학습 발달과정 중 어느 위치에 도달하였는지 알려주는 증거들을 제공하고, 학습평가에 제시된 학습 과제의 곤란도(difficulty)에 따라 학습자의 학습 수행 능력이 순차적으로 배열되는 증거들을 제공할 수 있음을 밝혀야 한다.

이상의 논의를 바탕으로 이 연구에서는 천문 시스템의 학습 발달과정을 조사하기 위하여 천문학적 사고의 실행과 천체의 운동 및 구조에 대한 이해를 통합적으로 반영한 천문 시스템 평가 문항을 개발하였다. 개발된 평가 문항을 초등학교와 중학교 학생들에게 적용하여 그 평가 결과를 바탕으로 천문 시스템에 대한 가설적 학습 발달과정을 도출하였다. 또한, 평가 결과를 바탕으로 천문 시스템 학습 발달과정 연구의 타당화 과정에 대하여 논의하고자 한다. 이하 본문에서는 평가 문항 개발 및 적용 과정과 평가 결과 해석에 대한 논의를 제시한다.

II. 연구 방법

천문 시스템의 학습 발달과정을 조사하기 위한 평가 문항 개발 및 타당성 검증 과정은 일련의 평가 시스템에 맞추어 진행되었다. 평가 시스템은 학습평가의 내용과 과제 및 구인을 선정하고, 이에 대한 평가 도구 또는 평가 문항을 개발하여 학생들의 학습 상황을 관찰하고, 그 결과를 해석하여 학생들의 학습 양상을 규명하는 일관적인 체계를 말한다. 이 연구에서는 천문 시스템에 대한 학습 발달과정을 조사하기 위한 평가 시스템으로서 Berkeley Evaluation and Assessment Research (BEAR) Center에서 제안한 구인 모델링 방식(construct modeling approach, Wilson, 2005)을 적용하였다. 구인 모델링 방식은

‘구인특화, 평가문항 개발, 평가결과 기술, 측정 모델 적용’으로 이루어진 4단계의 구성단위로 이루어진다. 각 단계별 자세한 내용을 아래에 제시하였다.

1. 구인 특화 단계

구인 특화(specifying construct) 단계에서는 평가하고 측정할 요소, 즉 구인(construct)을 선정하고 구인에 대한 선행 연구 결과를 바탕으로 구인구성도(construct map, Wilson, 2005)를 작성한다. 구인구성도는 평가 구인에 대한 이해 및 실행의 정도가 낮은 수준에서 높은 수준에 이르기까지 질적으로 구분되는 연속적인 도식을 말한다(Maeng *et al.*, 2013). 구인구성도는 평가 문항을 개발할 때 기본적인 밑받침이 되며, 평가 문항에 대한 학생들의 문항 반응 결과를 구인구성도와 비교하여 응답자를 그 구인의 이해 및 실행 수준에 따라 배치하는 준거가 된다. 그러므로 구인구성도는 그 자체로 학습 발달과정의 초안 역할을 할 수 있다.

이 연구의 구인구성도는 크게 세 부분으로 되어 있다. 구인구성도에서 가장 낮은 수준의 천문학적 오개념 사례 및 이와 관련된 공간적 사고와 시스템 사고의 실행 수준을 하위 정착점(lower anchor, Maeng *et al.*, 2013; NRC, 2007)으로 보았다. 하위 정착점은 학생들이 천문 시스템에 대한 학습 발달의 출발점으로서 학습 과정 이전에 이미 가지고 있던 일상적 개념 이해 수준 및 천문학적 사고 수준을 말한다. 반면에 구인구성도에서 가장 높은 수준의 천문학적 개념 이해 사례와 천문학적 사고 수준은 상위 정착점(upper anchor, Maeng *et al.*, 2013; NRC, 2007)에 해당한다. 상위 정착점은 천문 시스템의 학습 발달과정의 마지막 단계에서 학생들이 학습하게 될 거라고 예상되는 학습의 기대치 또는 천문학적 사고의 실행 수준을 말한다. 하위 정착점과 상위 정착점 사이에는 다양한 수준의 중간 단계(intermediate steps)의 과정이 있다. 한 주제에 대한 학습 발달과정에서 학생들의 학습 발달 경로의 중간 단계를 최대한 현실에 근접하게 조사하고 규명하는 것이 학습 발달과정 연구에서 가장 중요하다. 반면에 학습 발달의 중간 단계들은 하나의 정해진 경로가 있는 것이 아니며, 그 자체로도 매우 복잡하고 다양하게 제시될 수 있어서 “복잡한 중간 단계”(messy middle, Gotwals & Songer, 2013)라고 불리기도 한다.

구인구성도를 작성하기 위하여, 지구의 자전과 공전 운동, 밤과 낮의 변화, 계절의 변화, 달의 운동, 달의 위상 변화, 지구와 달의 운동, 지구-달 계의 구조, 태양계 행성의 운동, 태양계의 생성과정, 태양계의 구조, 은하계, 우리 은하의 구조, 우리 은하의 운동, 우주의 구조, 우주의 팽창, 우주의 진화 등을 주요 검색어로 하여 이 주제들에 대한 학생들의 개념 이해를 조사한 선행연구들을 검토하였다. 문헌 검토의 기본 방향은 이 주제들에 대해 학생들이 소유한 오개념의 사례와 개념 이해의 특성들을 조사하는 것이었다. 천문 시스템과 관련된 하위 개념들에 대하여 학생들이 가지고 있는 오개념 사례들은 모두 과학적 개념과 비교하면 올바르게 않은 사고들이지만, 각 오개념 사례를 세분하여 보면 과학적 개념에 근접한 오개념이 있고, 과학적 개념과는 차이가 많지만 일상적 개념 이해에 더 가까운 오개념들이 있었다. 그래서 연구에 참여한 공동연구자들이 함께 학생들의 오개념 사례들을 다시 검토하여 오개념들 간의 수준을 구분하여 구인구성도를 작성하였다.

이 연구에서 다루는 천문 시스템의 학습 발달과정은 천문 시스템을

구성하는 천체의 운동과 구조에 대한 개념적 이해뿐만 아니라, 천문 시스템을 이해하는데 필요한 천문학적 사고로서 공간적 사고와 시스템 사고를 포함한다. 그러므로 천문 시스템에 대한 학생들의 오개념 사례들의 수준을 구분할 때 각 오개념 사례들에 내포된 공간적 사고의 수준과 시스템 사고의 수준을 함께 파악하여 구인구성도를 작성하는 것이 필요하다고 판단되었다. 그래서 가장 낮은 천문학적 사고에 해당하는 수준 0은 천체의 운동과 구조에 대한 개념 이해 과정에서 공간적 사고와 시스템 사고를 적용하지 못하는 경우에 해당한다. 천문학적 사고에 대한 하위 정착점으로서 수준 1은 관측된 천체의 위치와 변화를 2차원 평면에서 사고하는 수준이며, 천체의 구조에 대한 단순한 패턴을 파악할 수 있는 시스템 사고가 적용되는 경우에 해당한다. 수준 2는 관측 현상을 지구에서 바라보는 관점에서 보이는 그대로 기술하는 사례로서, 운동하는 천체의 공간상의 위치나 방향의 변화만을 사고하거나, 관측 현상에서 천체의 구조를 나타내는 요소들을 파악하지만 그 요소들 간의 규모와 비율 관계를 이해하지 못하는 수준이다. 수준 3은 교과서에 제시된 과학 지식 정도의 천문학적 사고를 나타내는 수준이며, 우주에서 바라보는 관점에서 천체의 운동을 설명하는 모델에 대한 지식에 근거하여 관측 현상을 서술할 수 있다. 또한, 관측 현상에서 천체의 구조를 나타내는 요소들 간의 규모와 비율을 고려한 상호연관성을 파악하여 설명할 수 있는 수준이다. 마지막 수준 4는 상위 정착점으로서 공간적 사고와 시스템 사고가 충분히 발현되어 천체의 운동과 구조를 설명하는 사례로서, 2차원적 평면 자료를 3차원적 공간 자료로 자유롭게 전환할 수 있으며, 지구에서 보는 관점으로 표현된 관측 자료를 우주에서 보는 관점으로 표현하는 천문학적 모델을 만들어 설명할 수 있는 경우이다. 또한, 시스템 사고의 측면에서는 천체의 구조를 구성하는 하위 요소들을 구분하고, 그 요소들 간의 상호 관련성을 이해할 뿐만 아니라, 각 요소들을 종합하여 천체의 구조의 규모와 비율, 변화 등과 같은 전체적인 시스템의 특성을 설명할 수 있는 경우를 말한다.

2. 문항 개발 단계

문항 개발(item design) 단계에서는 앞서 작성한 구인구성도에 맞추어 구체적인 평가 문항을 개발하고 점검한다. 이 연구에서는 천문 시스템의 학습 발달과정을 조사하기 위하여 순위 선다형 문항(ordered multiple-choice items)을 개발하였다. 순위 선다형 평가 문항은 선다형 문항에 포함된 선택지들이 학생들의 개념 이해의 수준 및 과학적 탐구 실행의 위계에 따라 순위를 정하여 구성되는 형태를 띤다(Briggs et al., 2006; Briggs & Alonzo, 2012). 그러므로 순위 선다형 평가 문항은 하나의 선택지를 선택하는 문항이지만 학생들의 정답과 오답을 구분하는 것이 아니라, 어느 선택지를 선택하더라도 학생들의 선택 결과가 그 학생의 개념 이해 및 탐구 실행의 수준에 대한 정보를 제공해 줄 수 있다.

이 연구에서 순위 선다형 평가 문항 개발은 ‘평가 맥락 및 평가 장면 선정 - 선택지 구성 - 평가 문항 검토’의 순서로 진행되었다. 평가의 맥락으로서 지구, 지구-달 계, 태양계, 은하와 우주로 천문 시스템의 범위를 구분하고, 각각의 맥락에 대하여 지구에서 관측하여 볼 수 있는 관측 자료를 제시하는 형태로 평가의 장면을 구성하였다. 문항의 형태는 관측 자료를 보고 그 자료를 해석한 것을 나타내는 선택지들

중에서 학생들의 생각과 가장 일치하는 것을 선택하게 하였다. 이것은 학생들이 각 문항에서 정답을 고르는 것이 아니라, 자신의 생각과 가장 가까운 선택지를 고르게 하여 학생들의 수준을 표현하게 해 준다. 순위 선다형 문항의 선택지들은 천문 시스템의 개념 이해와 천문학적 사고의 실행 수준에 따라 배열된 구인구성도의 수준별 진술을 바탕으로 작성되었다.

평가 문항의 검토는 이 연구에 참여한 공동 연구진의 전체 회의를 통해 각 문항들이 천문 시스템을 구성하는 천체의 구조와 운동의 맥락에 적합하게 구성되었는지, 천문학적 내용 지식의 측면에서 오류는 없는지, 천체의 구조를 나타내는 현상에 시스템 사고가 적절하게 반영되었는지, 천체의 운동을 나타내는 현상에 공간적 사고가 적합하게 반영되었는지, 선택지들의 문장 구성과 배열이 구인구성도에 부합하는지 등을 검토하였다. 검토 과정에서 오류나 부적절한 부분이 발견되면 공동 연구진의 협의를 통해 수정하거나, 새로운 문항으로 대체하여 다시 검토하였다. 문항 검토 결과 최종적으로 12개의 평가 문항이 확정되었다.

3. 평가 결과 기술 단계

이 연구를 진행하던 시기에 초등학교는 2009 개정 과학과 교육과정 이 적용되고, 중학교 1학년은 2009 개정 과학과 교육과정, 2학년과 3학년은 2007 개정 과학과 교육과정이 적용되었다. 시행된 과학 교육 과정에서 천문 시스템 관련 내용은 초등학교 5학년에 ‘지구와 달의 운동’ 단원과 ‘태양계와 별’ 단원과 초등학교 6학년에 ‘계절의 변화’ 단원이, 그리고 중학교 2학년에 ‘태양계’ 단원과 ‘별과 우주’ 단원이 편성되었다. 이에 따라 이 연구에서 개발된 순위 선다형 평가 문항들을 서울, 인천, 경기, 강원, 광주 지역의 초등학교 5학년과 6학년, 중학교 2학년 학생 1,289 명을 임의표집하여 적용하고 응답 결과를 수집하였다. 검사지에 응답한 학생들의 인원수 현황은 아래 Table 1 과 같다.

평가 결과(outcome spaces) 기술 단계는 개발된 평가 문항을 학생들에게 적용하여 평가 구인에 대한 검사 결과를 작성하는 과정을 말한다. 이 단계에서는 평가 문항에 대한 학생들의 응답 결과에 적절한 점수를 할당하여 학생 개인별로, 그리고 문항별로 점수를 부여하게 된다. 학습 발달과정 연구에서 평가 문항으로 열린 응답 검사지(open-ended written test)나 임상적 면담(clinical interview)을 사용할 경우 학생들의 응답 결과가 매우 다양하게 제시될 수 있으므로 그 결과를 검토하여 유사한 응답끼리 분류하고 그것을 세분하여 범주화하는 과정이 필요하다. 이와 같이 응답 결과를 분류하고, 분류된 범주에 따라 점수를 부여하여 학생들의 검사 결과를 구체적으로 기술하는 것이 평가 결과 기술 단계에서 중요한 과제이다. 이 연구에 적용한 순위 선다형 평가

Table 1. The number of students participated in this study

학교, 학년, 성별	서울	경기	인천	강원	광주	합계	
초등학교 (5학년)	남	24	64		162	54	304
	여	22	49		156	43	270
초등학교 (6학년)	남	120				48	168
	여	127				51	178
중학교 (2학년)	남	111		61	147		319
	여			50			50
총 인원						1,239	

문항은 구인구성도에서 정리된 개념 이해와 탐구 실행의 수준에 맞추어 문항의 선택지를 작성하였으므로 선택지마다 부분점수를 순차적으로 부여하여 비교적 수월하게 문항별 배점을 구성할 수 있다. 예를 들면 모두 4개의 선택지 중에서 상위 정착점에 해당하는 선택지는 4점, 그보다 한 단계 낮은 수준에 해당하는 선택지는 3점, 그 아래는 2점, 끝으로 하위 정착점에 해당하는 선택지는 1점을 부여하는 방식이다.

4. 측정 모델 적용 단계

구인 모델링 방식의 마지막 단계인 측정 모델(measurement model) 적용 단계는 평가 결과로 얻은 학생들의 점수를 평가 구인에 대한 성취 수준과 연결하기 위하여 통계적 해석으로 응답자의 점수나 성취 정도를 측정하는 것이다. 평가 문항을 구인구성도에 근거하여 개발하였으므로 학생들의 응답 결과가 구인구성도에 부합할 수 있다면, 구인구성도가 적절하다고 할 수 있을 것이다. 따라서 응답자의 점수나 성취 정도를 구인구성도의 어느 위치에 배치할 수 있는지 결정하게 해 주는 적합한 통계적 측정 모델을 선정하여 적용하는 것이 중요하다. 학습 발달과정의 선행 연구들에서 학습평가 결과와 학습자의 발달 경로 간의 일치 여부와 타당성을 검증하기 위한 측정 모델의 적용은 문항반응 이론(item response theory)에 근거한 통계적인 심리측정(psychometric) 방법(Briggs & Alonzo, 2012; Wilson, 2012)이 주로 사용되었다. 문항 반응 이론에 의하면, 검사 문항은 측정하고자 하는 구인에 대하여 고유한 특성을 가지고 있으며, 피험자 역시 그 구인과 관련된 잠재적인 특성(latent trait)을 지니고 있다. 그러므로 각 문항이 나타내는 고유한 문항 특성과 그 문항에 대한 피험자의 반응을 수학적으로 계산하여 피험자가 지니고 있는 잠재적 특성을 측정할 수 있다는 것이다(Ji & Chae, 2000; Seong, 2002). 문항 반응 이론에서 고려하는 문항의 고유한 특성으로 문항의 곤란도(difficulty), 문항 변별도, 문항 추측도 등이 있는데, 덴마크의 수학자 Georg Rasch는 그 중 문항의 곤란도만을 문항의 모수(parameter)로 반영하여 응답자의 문항 응답 확률을 계산하는 Rasch 모델을 제안하였다. Rasch 모델은 문항의 곤란도와 피험자의 능력의 관계를 계산할 수 있게 해 주므로 그동안 학습 발달과정 연구의 측정 모델로 가장 널리 사용되어 왔다(e.g., Jin & Anderson, 2012; Neumann *et al.*, 2013; Gotwals & Songer, 2013). 이 연구에서도 순위 선다형 평가문항을 적용하여 천문 시스템에 대한 학습 발달과정을 조사하는데 Rasch 모델을 적용하였다. 특히, 순위 선다형 평가 문항의 각 선택지마다 부분 점수를 부여하였으므로 Rasch 모델의 세부적인 측정 방식 중에서 부분점수 모형(partial credit model, Masters, 1982)을 사용하였다. Rasch 모델 및 부분점수 모형에 대한 자세한 수학적 설명은 이 논문의 범위를 벗어나므로 생략한다. 연구 결과에서 문항 반응 분석 결과 및 그것을 해석하는 과정을 자세히 기술할 것이다.

학습 발달과정 연구에서 구인 모델링 방식의 4가지 단계는 하나의 사이클 형태로 이루어진다. 즉, 측정 모델 단계에서 학생들의 문항 응답 결과를 구인구성도와 연결하여 도출된 결과는 다시 구인구성도를 수정하는데 적용되며, 그 수정된 구인구성도는 잠정적인 학습 발달 과정의 초기 형태로 제시된다. 이렇게 정리된 학습 발달과정의 초기 형태는 새로운 구인구성도로서 2차 문항 개발을 위한 근거가 되어

다시 구인 모델링 방식의 2차 사이클이 진행된다. 이 연구는 3년 동안 구인 모델링 방식의 사이클을 반복적으로 진행하도록 설계되었으며, 이 논문에서는 그 중 첫 번째 사이클의 연구를 통해 얻은 결과를 바탕으로 천문 시스템의 학습 발달과정을 개발하고자 하였다.

5. 검사 문항 및 검사 결과의 타당성

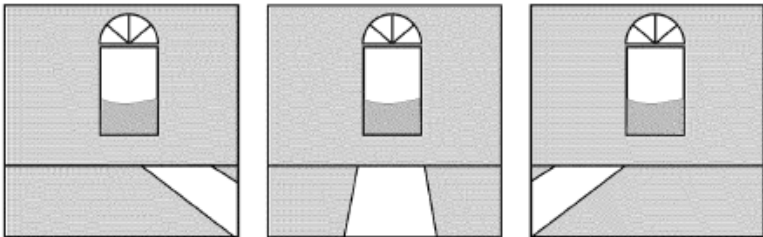
타당도 또는 타당성에 대한 증거와 관련하여 그동안 국내 과학교육계에서는 타당도를 검사도구나 평가 문항의 속성으로 보고 검사도구의 타당도를 주로 다루었다. 그러나 Standards for Educational and Psychological Testing (AERA, APA, & NCME, 1999)에서는 타당도를 “제한된 목적에 맞게 검사를 적용하여 얻은 검사 점수를 해석할 때 그 해석이 증거나 관련 이론에 의해 뒷받침 되는 정도” (p. 9)로 정의하였다. 이 정의에 따르면 타당성 검증 과정은 검사 도구 또는 검사 자체의 타당도를 측정하기보다는 “검사 결과를 해석하는데 필요한 합당한 과학적 기초를 제공해 주는 근거나 증거를 축적하는”(Seong, 2002, p. 37) 과정을 포함해야 한다. 따라서 검사 결과를 포함하지 않은 채 단순히 검사도구나 평가 문항의 내적 오류 여부를 검토하는 ‘내용 타당도’ 또는 ‘안면 타당도’만을 수행하는 것은 타당성 검증으로 충분하지 않음을 알 수 있다.

이 연구에서 타당성 검증 과정은 구인 모델링 방식의 4단계에 맞추어 진행되었다. 즉, 구인 특화 단계와 문항 개발 단계에서 검사 문항의 내용에 기반한 타당성의 근거와 내적 구조에 기반한 타당성의 근거를 확보할 수 있다. 또한, 평가 결과 기술 단계 및 측정 모델 적용 단계에서 학습 발달과정 개발을 위한 기초 자료로서 검사 결과에 기반한 타당성의 근거를 수집하여 검사 결과 해석의 타당성을 검증하였다. 검사 문항이 측정을 위하여 규정된 내용을 잘 표현하는 정도를 의미하는 내용에 기반한 타당성의 근거 (validity evidence based on test content, AERA, APA, & NCME, 1999)를 알아보기 위하여, 구인 특화 단계에서 먼저 평가 구인으로서 천문 시스템을 구성하는 천체들의 운동과 천문 시스템의 구조, 그리고 이를 이해하는데 필요한 공간적 사고와 시스템적 사고의 내용을 상세화하였다. 내용에 기반한 타당성의 근거는 일반적으로 그 내용에 대한 전문가들의 집단적 판단에 근거하여 이루어지므로(Seong, 2002), 전문가의 협의과정과 그 결과물로 작성된 구인구성도에 포함된 하위 정착점과 상위 정착점 및 중간 단계들의 내용적 타당성을 바탕으로 판단할 수 있다. 이에 이 연구에 참여한 공동연구원 5명(중등 지구과학교육 전공 교수 2명, 중등 교사 경력 10년 이상인 중등 지구과학교육 전공 박사급 연구원 2명, 초등 교사 경력 10년 이상인 지구과학교육 전공 박사급 연구원 1명)이 지속적인 세미나와 토의를 거쳐 평가 구인의 내용을 정리하고 합의하였다. 문항 개발 단계에서는 개발된 순위 선다형 평가 문항들을 검토하는 과정에서 각 문항들의 질문과 선택지들이 구인구성도에 제시된 평가 구인의 내용과 그것의 수준 구분을 얼마나 잘 반영하고 있는지를 파악하게 된다. 이를 통해 개발된 문항과 그 문항에서 평가 및 검사하려는 요소들의 관계가 평가의 구인과 합치되는 정도를 파악하는 내적 구조에 기반한 타당성의 근거 (validity evidence based on internal structure, AERA, APA, & NCME, 1999)를 확보할 수 있게 하였다. 평가 결과 기술 단계와 측정 모델 적용 단계에서는 검사를 실시하여 얻은 응답 자료를 분석한 결과를 이용하여, 검사 결과가 측정하고자 했던 구인들

Table 2. Overview of the ordered multiple-choice items on the contents and practices of astronomical systems

문항 번호	4E	5E	6E	7EM	8EM	9EM	10EM	11S	12S	13S	14GU	15GU	
평가구인	영역	지구	지구	지구	지구-달계	지구-달계	지구-달계	지구-달계	태양계	태양계	태양계	은하와 우주	은하와 우주
	개념	자전	공전	낮의 길이	일주 운동	일주 운동	달의 공전	달의 위상	행성의 공전	행성의 공전	행성의 특징	우주의 팽창	우리 은하
	실행	공간적 사고	공간적 사고	시스템 사고	공간적 사고	시스템 사고	공간적 사고	시스템 사고	공간적 사고	공간적 사고	시스템 사고	공간적 사고	시스템 사고

4. 다음 그림 (가)는 겨울철 어느 날, 우리나라 건물의 창문으로 비친 햇살의 모습을 아침, 점심, 저녁에 관찰한 것이고, 그림 (나)는 같은 날 인공위성에서 지구를 찍은 영상을 시간 순서대로 나타낸 것이다.



아침 점심 저녁

그림 (가)




그림 (나)

그림 (가)와 (나)를 보고 태양과 지구의 움직임에 대하여 서술한 것 중 자신의 생각과 가장 일치하는 것을 고르시오.

- 태양은 왼쪽에서 오른쪽으로 움직이고 지구는 왼쪽에서 오른쪽으로 움직인다.
- 지구는 반시계 방향으로 자전하고, 태양은 동쪽에서 서쪽으로 움직인다.
- 태양은 서쪽에서 동쪽으로 움직이고 지구는 서쪽에서 동쪽으로 움직인다.
- 지구가 반시계 방향으로 자전하므로 태양이 그것과 반대로 움직이는 것처럼 보인다.

Figure 1. An example of the ordered multiple-choice items in this study: Earth's rotation

의 관계를 얼마나 잘 표현하는지, 그리고 검사의 결과가 검사의 목적에 얼마나 부합하는지를 검증하는 검사결과에 기반한 타당성의 근거 (validity evidence based on the consequence of testing, AERA, APA, & NCME, 1999; Seong, 2002; Seong & Si, 2006)를 확보할 수 있다. 즉, 평가 결과 기술 단계에서 각 문항별로 점수를 부여하고 이를 종합한 결과를 측정 모델 단계에서 Rasch 모델을 적용하여 학생들의 응답 결과가 구인구성도에서 잠정적으로 구분했던 하위 정착점 - 중간 단계 - 상위 정착점으로 이어지는 학습 발달의 경로를 잘 보여줄 수 있는지를 파악하여 검사 결과의 해석과 관련된 타당성의 증거를 제공하게 된다.

III. 연구 결과

구인 모델링 방식 평가 시스템의 1차 순환 과정을 거치는 동안 순위 선다형 평가 문항을 개발하였고 이를 초등학교와 중학교에 적용하였다. 또한, 그 결과를 Rasch 모델로 측정하여 해석한 내용을 바탕으로 도출된 검사 결과에 기반한 타당성의 근거와 천문 시스템에 대한 1차적인 가설적 학습 발달과정을 연구 결과로 제시하였다.

1. 순위 선다형 평가 문항 개발

이 연구에 적용된 천문 시스템 평가 문항은 천문 시스템의 세부

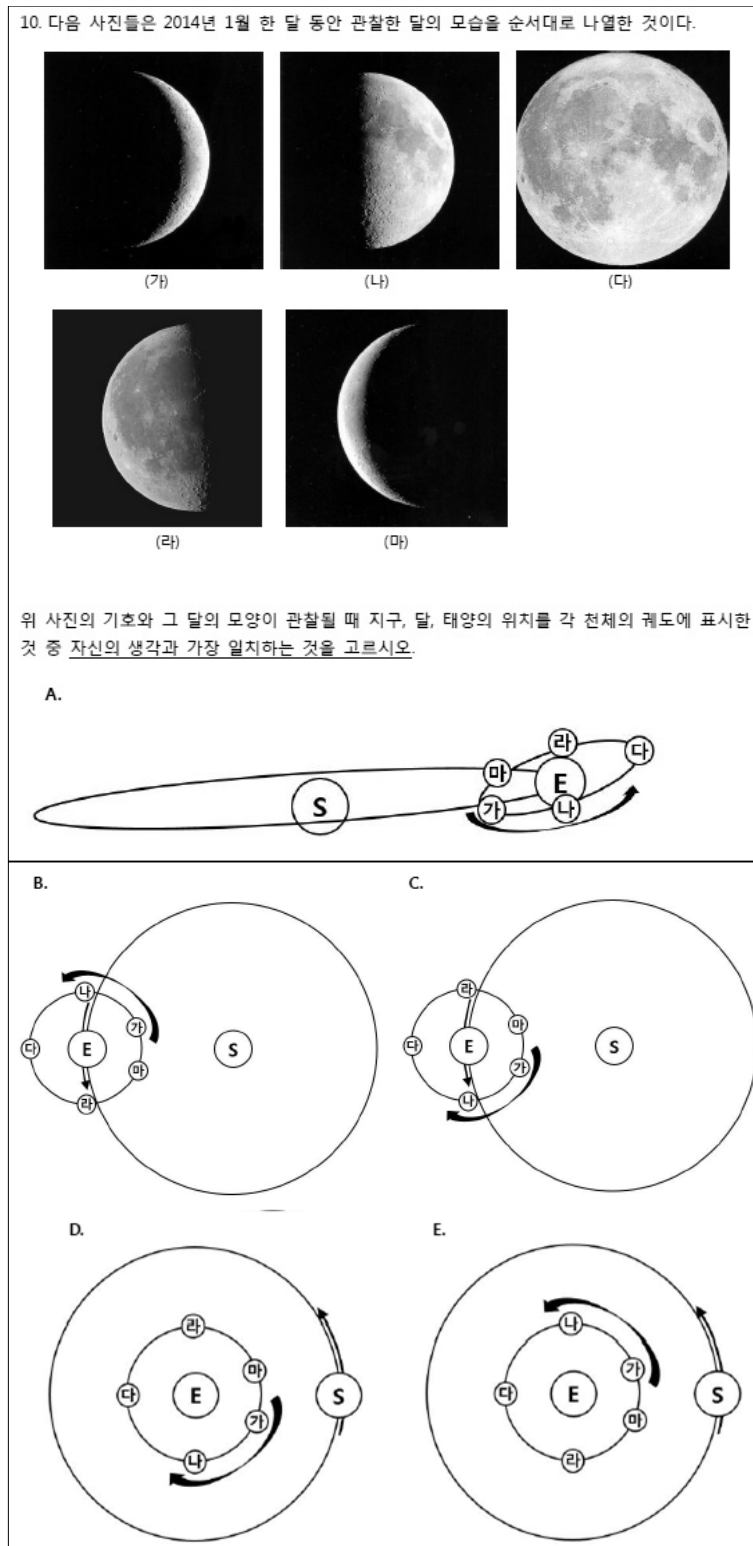


Figure 2. An example of the ordered multiple-choice items in this study: Earth-Moon system

영역으로서 지구의 운동에 대해 3 문항, 지구-달 계의 운동과 구조에 대해 4 문항, 태양계의 운동과 구조에 대해 3 문항, 은하와 우주의 운동과 구조에 대하여 2 문항이 개발되었다. 개발된 전체 문항은 15 문항이었으나, 그 중 1번 ~ 3번까지 3 문항은 공간적 방향에 대한 일반적인 공간 능력을 평가하는 문항으로서 천문 시스템에 대한 이해를 위한 독립변수의 기능을 하는 문항이어서 이 연구의 분석에는 제외되었다. Table 2에 본 문항에 해당하는 12 문항의 문항별 평가 구인을 천문 시스템의 영역별로 정리하였다. Table 2에서 E(Earth)는 지구의

운동, EM(Earth-Moon)은 지구-달 계, S(Solar system)는 태양계, GU(Galaxy and the Universe)는 은하와 우주에 대한 문항임을 표시한 것이다. 아래에 각 세부 영역별로 대표적인 문항 하나씩을 연구 결과로 제시하였다.

가. 지구의 운동에 대한 문항 예시

지구의 운동 영역의 순위 선다형 평가 문항의 예시(4E)를 Figure

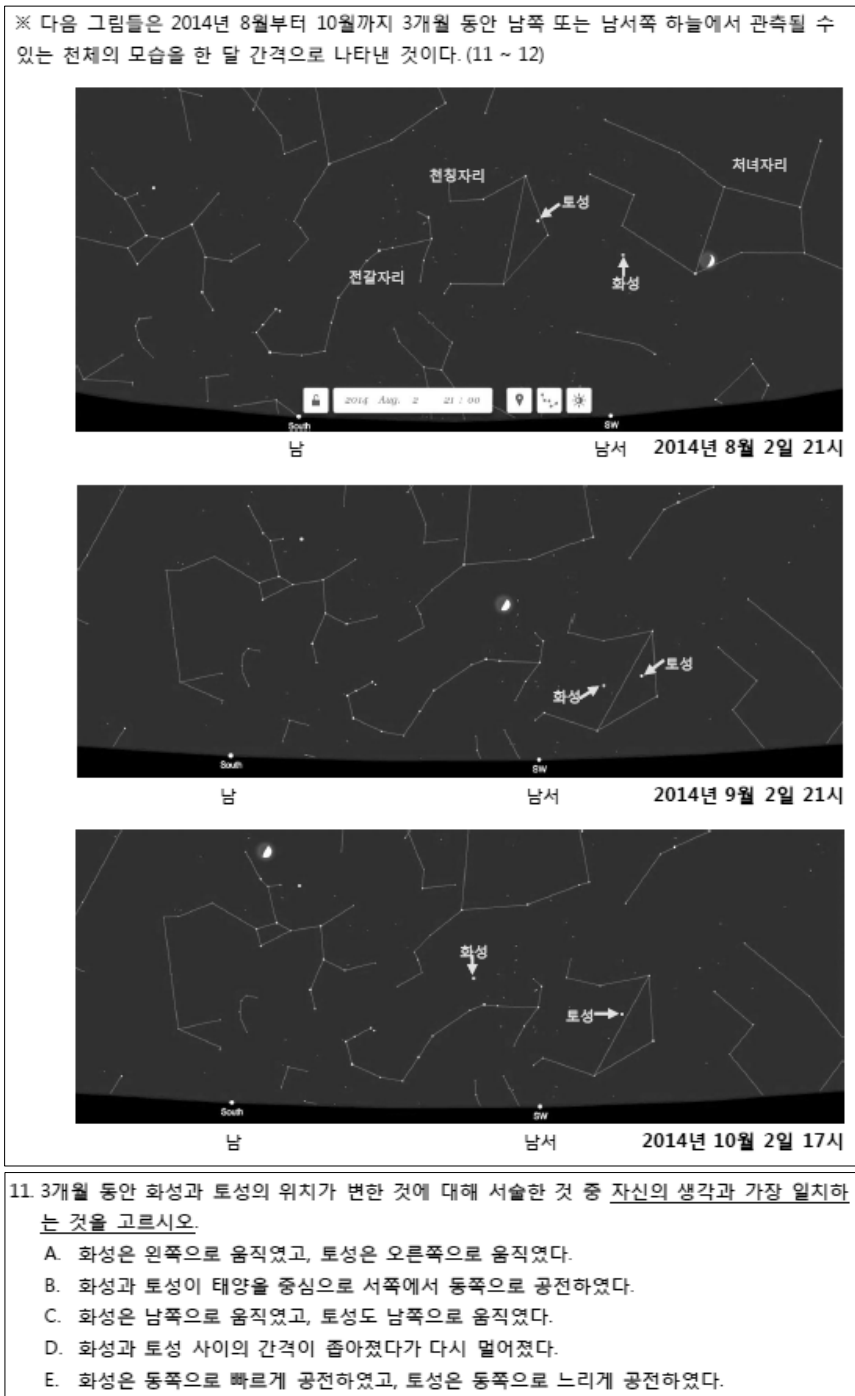


Figure 3. An example of the ordered multiple-choice items in this study: Solar system

1 에 제시하였다. 이 문항은 하루 동안 아침, 점심, 저녁에 창문에 비친 햇빛의 방향으로 태양의 일주운동을 관찰하고, 이것을 같은 날 우주의 인공위성에서 지구를 내려다 본 영상 자료와 연관시켜 이해하는 맥락을 제시하였다. 이 맥락에서 지구의 자전과 일주운동의 개념에 대한 이해를 천문 시스템의 내용적 평가 구인으로 선정하였고, 천체 운동의 방향을 판단하고 지구에서 보는 관점과 우주에서 보는 관점 간의 공간적 관점의 변환을 포함하는 공간적 사고를 천문학적 탐구실행의 평가 구인으로 선정하였다. 이 문항에서 선택지 A는 하위 정착점(수준 1)에 해당하며, 태양과 지구의 운동을 평면적으로만 이해하는 수준을 나타낸다. 선택지 C는 수준 2에 해당하며, 지구에서 보는 관점에서 태양과 지구의 위치 변화와 운동 방향을 보이는 그대로 기술할 뿐이다. 선택지 B는 우주에서 바라보는 관점에서 태양과 지구의 운동을 서술하는 수

준 3에 해당한다. 수준 3의 이해는 천문학적 지식에 근거한 모델을 도입하지만, 그 설명 모델을 지구에서 보는 관점에서 관측된 태양과 지구의 운동과 직접 연결시키지 못한다. 선택지 D는 상위 정착점(수준 4)에 해당하는 이해 사례로서, 우주에서 보는 관점에서 도입된 천문학적 설명 모델을 지구에서 보는 관점에서 태양과 지구의 운동을 관측한 자료를 연관시켜 설명할 수 있다.

나. 지구-달 계의 운동과 구조에 대한 문항 예시

지구 - 달 계의 운동과 구조에 대한 순위 선다형 평가 문항 중 예시 문항(10EM)을 Figure 2에 제시하였다. 이 문항의 맥락은 한 달 동안 달의 위상을 나타낸 사진을 보고 그 때 지구, 달, 태양의 위치를 각

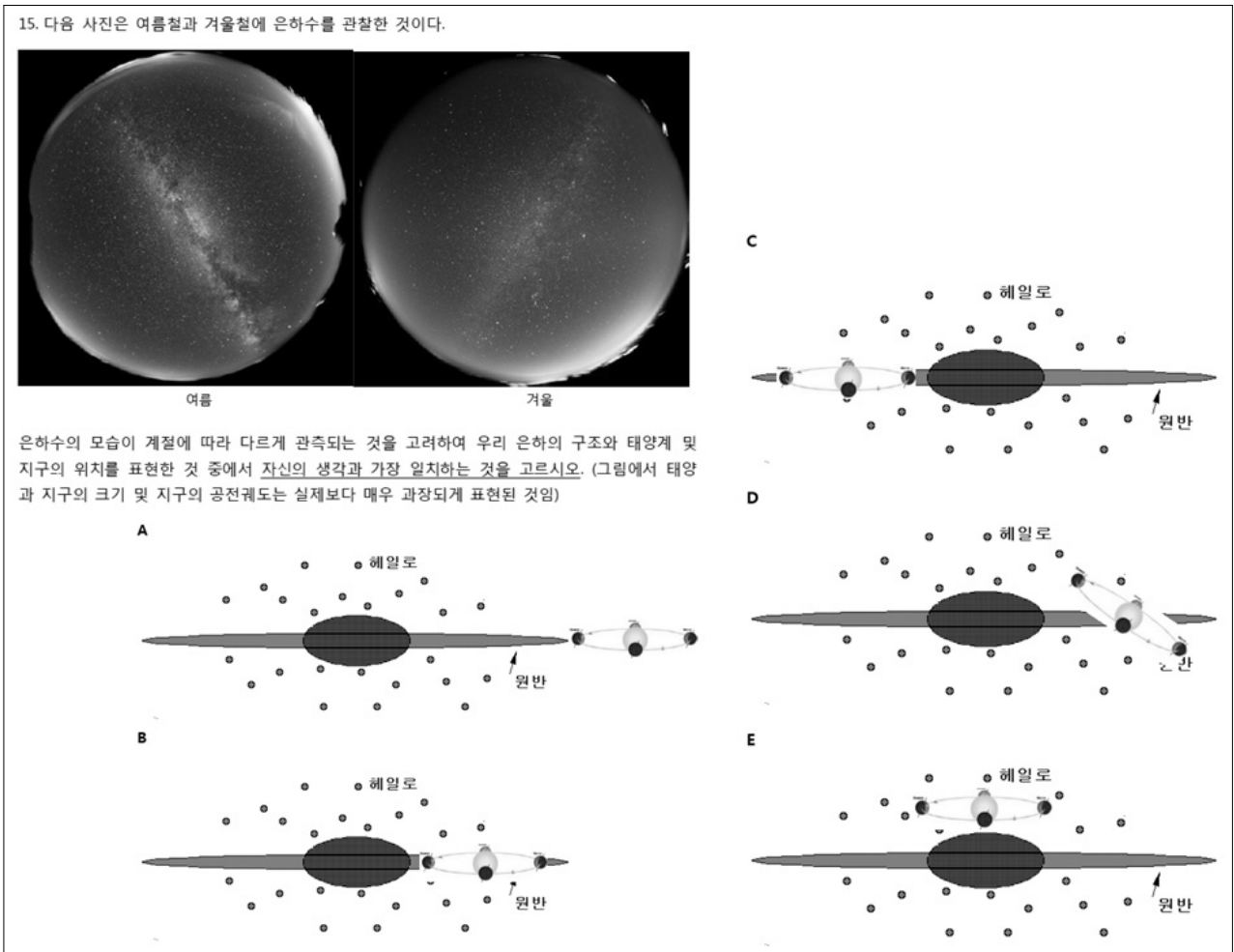


Figure 4. An example of the ordered multiple-choice items in this study: Galaxies and the universe

천체의 궤도에 표시하여 설명하는 것이다. 천문 시스템의 내용적 평가 구인은 달의 위상 변화이며, 천문학적 탐구실행의 평가 구인은 달의 위상 변화에 대한 패턴을 파악하고 이를 설명할 수 있는 지구-달 계의 시스템 모델을 구성하는 과정을 나타내는 시스템 사고로 선정하였다. 이 문항에서 선택지 D와 E는 각각 지구에서 보는 관측자 중심으로 지구-달 계를 이해하며 달의 위상이 변한다는 패턴만을 인식하는 수준 1에 해당하며, 하위 정착점을 나타낸다. 선택지 C는 지구-달 계를 구성하는 천체들을 파악하고 있지만, 각 구성요소들 간의 관계를 정확히 인식하지 못하고, 달의 위상이 제시되는 순서대로 달의 운동 방향을 설명하는 수준 2에 해당한다. 선택지 B는 지구-달 계의 구성 요소들 간의 상호관계를 천문학적 지식에 근거하여 궤도 관계를 이해하는 수준을 말한다. 선택지 A는 상위 정착점(수준 4)으로서 지구-달 계의 구성 요소들의 관측 정보를 종합하여 지구의 공전 궤도면과 달의 공전 궤도면이 서로 어긋나 있는 천문학적 모델을 이용하여 달의 위상 변화를 설명할 수 있는 수준을 말한다.

다. 태양계의 운동과 구조에 대한 문항 예시

태양계의 운동과 구조에 대한 순위 선다형 평가 문항 중 예시 문항(11S)을 Figure 3에 제시하였다.

이 문항의 맥락은 3 개월 동안 별자리를 배경으로 화성과 토성의 위치 변화를 관찰한 자료를 바탕으로 행성의 공전 운동을 파악하는

것이다. 천문 시스템의 내용적 평가 구인으로는 행성의 공전 운동과 공전 속도에 대한 것이 선정되었다. 또한, 천체의 위치 변화와 방향의 변화를 판단하고, 2차원적 평면 자료를 3차원적 공간 자료로 변환하여 사고하는 공간적 사고가 천문학적 탐구실행의 평가 구인으로 선정되었다. 이 문항에서 선택지 A와 선택지 D는 화성과 토성의 위치 변화와 그 방향을 2차원 평면의 운동으로만 이해하는 수준 1로서 하위 정착점에 해당한다. 선택지 C는 지구에서 바라보는 화성과 토성의 위치 변화와 운동 방향을 공간적인 방향, 즉 방위를 이용하여 표현하는 수준 2에 해당한다. 선택지 B는 행성들이 태양을 중심으로 서쪽에서 동쪽으로 공전한다는 천문학적 지식에 근거하여 관측 자료를 이해하지만, 우주에서 보는 관점의 모델을 그대로 도입하는 수준 3에 해당한다. 끝으로 선택지 E는 2차원 평면에 표시된 화성과 토성의 위치 변화와 운동 방향을 배경 별자리와 상대적인 위치 관계로 인식하여 3차원 공간적인 운동으로 파악할 수 있다(수준 4). 그래서 우주에서 바라보는 관점의 천문학적 설명 모델을 적용하여 화성과 토성의 운동 속도를 파악하는 상위 정착점에 해당한다.

라. 은하와 우주의 운동과 구조에 대한 문항 사례

은하와 우주의 운동과 구조에 대한 순위 선다형 평가 문항의 예시 문항(15GU)을 Figure 4에 제시하였다. 이 문항의 맥락은 여름과 겨울에 은하수의 모양과 방향, 별 개수의 차이에 대한 정보를 종합하여

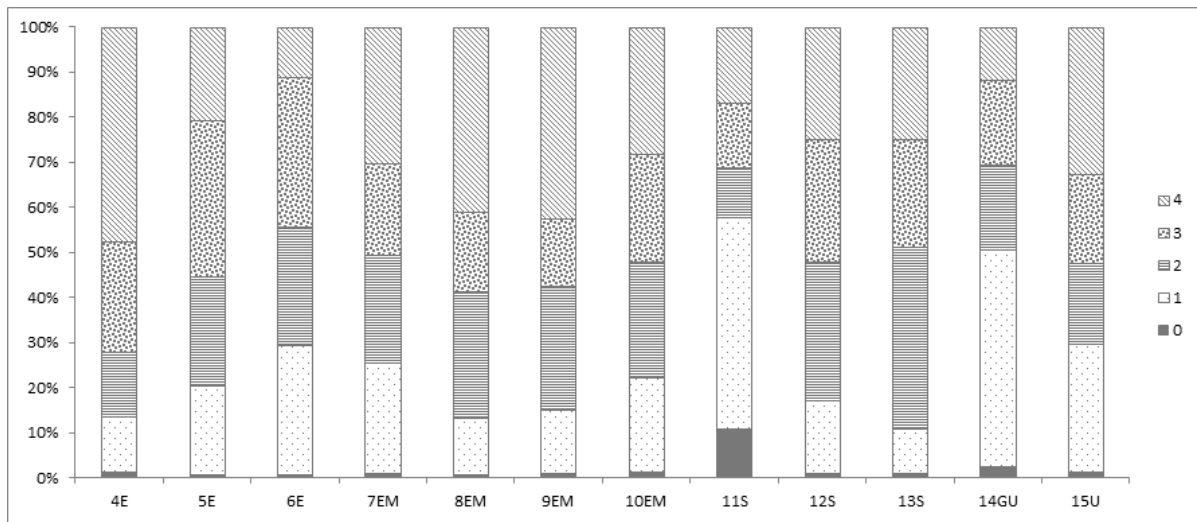


Figure 5. Outcome spaces of the assessment in this study (4E: Item 4 about Earth, 7EM: Item 7 about Earth-Moon system, 11S: Item 11 about solar system, 14GU: Item 14 about galaxy and the universe)

우리 은하의 구조와 태양계의 위치를 파악하는 것이다.

이 문항의 과학 내용적 평가 구인은 우리 은하의 구조와 우리 은하에서 태양계의 위치이며, 천문학적 탐구실행으로서 평가 구인은 천체의 위치와 방향을 판단하는 공간적 사고와 관측 현상에서 천체의 구조에 대한 패턴을 이해하고, 그 구조를 구성하는 하위 요소들 간의 상호관계를 종합하여 시스템 모델을 수립하는 시스템 사고의 복합적인 사고 능력이 선정되었다. 선택지 E는 이 문항을 해결할 때 공간적 사고와 시스템 사고를 적용하지 못하는 수준 0에 해당한다. 선택지 A는 은하수의 계절별 차이에 대한 패턴을 단순히 2차원적으로 파악하여 별이 많고 적음만을 이해하는 수준 1을 나타내며, 이 문항에서 하위 정착점에 해당한다. 선택지 C는 계절에 따른 지구와 태양 간의 위치와 공간적 방향을 이해하고, 우리 은하의 구조 및 구성 요소를 파악하는 수준 2에 해당한다. 선택지 B는 우리 은하의 구조 및 계절별 지구와 태양의 위치에 대한 상호연관성을 파악하여 은하수의 방향과 계절에 따른 차이를 이해하는 수준 3에 해당한다. 선택지 D는 은하 평면과 지구의 공전 궤도에 대한 이해를 포함하는 우리 은하의 구조에 대한 천문학적 모델을 이용하여 2차원으로 표현된 계절별 은하수의 관측 자료를 3차원 공간적 자료로 변환하여 설명하는 수준 4(상위 정착점)에 해당한다.

2. 검사 결과에 기반한 타당성의 근거

가. 문항별 평가 결과

개발된 순위 선다형 평가문항을 초등학교 5학년과 중학교 2학년 학생들에게 적용한 평가 결과를 그래프로 표시하였다(Figure 5). 막대의 표시는 각 문항별로 각각 수준 1, 수준 2, 수준 3, 수준 4에 해당하는 선택지를 선택한 학생들의 백분율을 나타낸 것이다. 수준 0은 무응답한 경우에 해당된다.

Figure 5를 보면, 문항 4E, 8EM, 9EM에서 수준 4의 선택지를 선택한 학생들의 비율이 다른 문항보다 높고, 문항 11S와 14GU에서 수준 1의 선택지를 선택한 학생들의 비율이 다른 문항보다 높다는 것을 대략적으로 비교할 수 있다. 그러나 이와 같은 평가 결과의 기술만으로는 문항의 곤란도와 학생들의 천문 시스템 이해 수준 및 천문학적

사고의 실천 수준에 따른 학습 발달과정을 조사하는데 필요한 구체적 정보를 얻을 수 없다. 그리고 이 연구에 적용된 순위 선다형 평가 문항은 선형적인 발달 경로에 해당하는 구인구성도에 기반하여 개발된 것이므로 실제 학생들의 응답이 반영된 경험적 연구에 의해 그 타당성이 검증되어야 한다. 따라서 검사의 결과가 구인구성도에서 선정한 구인들의 수준별 전개와 얼마나 일치하는지, 그리고 천문 시스템의 학습 발달과정을 제시하려는 목적에 얼마나 부합하는지를 해석할 수 있어야 한다. 이를 위해 측정 모델로서 Rasch 모델을 적용한 분석 결과를 아래에 제시한다.

나. Rasch 모델 측정 결과

Figure 6은 Rasch 모델에 맞추어 순위 선다형 평가 문항에 대한 학생들의 응답 반응을 측정된 결과를 나타낸 person-item map이다. 이 맵은 문항의 곤란도와 피험자의 능력을 동일한 스케일의 로지트(logit, log-odds-units, Bond & Fox, 2007) 단위로 환산하여 하나의 맵에 표시한 것으로, 이 맵을 고안한 시카고 대학(the university of Chicago)의 Benjamin Wright의 이름을 따서 흔히 Wright map이라고 부른다. Wright map은 문항의 곤란도와 피험자의 능력을 같은 스케일의 맵에 표시함으로써 문항의 곤란도 수준에 따라 피험자의 능력과 수준을 비교하고 구별할 수 있는 장점이 있다. Figure 6의 가장 왼쪽의 “MEASURE” 아래에 있는 숫자들(3, 2, 1, 0, -1, -2, -3)은 피험자의 능력과 문항의 곤란도를 고려하여 계산한 로지트 값을 나타낸 것이다. 그 오른쪽에 세로선 왼쪽의 “PERSON” 아래에 있는 # 표시들은 그 위치의 로지트 값에 해당하는 능력 수준을 가진 피험자들의 인원수를 나타낸 것이다. 즉, “#” 표시 하나가 20명을 의미하며, “-” 표시는 1~19명을 구분 없이 나타낸 것이다. 로지트 값은 조사된 속성(traits)의 측정 값을 선형적으로 나타내므로, 위쪽에 있을수록 더 높은 수준의 능력을 가진 피험자를 표시한 것이고, 아래로 내려갈수록 피험자의 능력 수준이 낮아지는 것이다. 여기서 피험자의 능력은 각 검사 문항에서 피험자가 획득한 총점수를 기준으로 환산된 것이다. 세로선 왼쪽의 “M”은 학생들의 응답 점수를 Rasch 모델로 측정하여 상대적인 능력을 로지트 단위로 환산한 평균값을 나타낸 것이다. 즉, 학생들의 평균적인

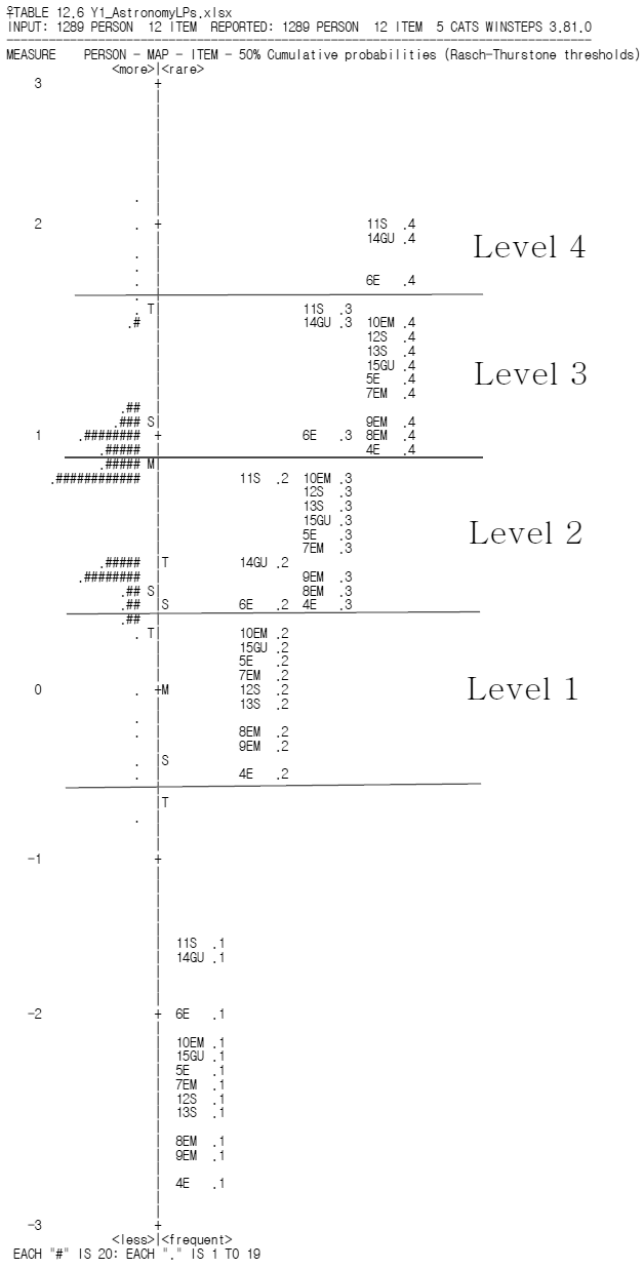


Figure 6. Wright map from Rasch measurement of the assessment in this study

능력값은 약 0.8 로지트 정도에 해당한다고 볼 수 있다.

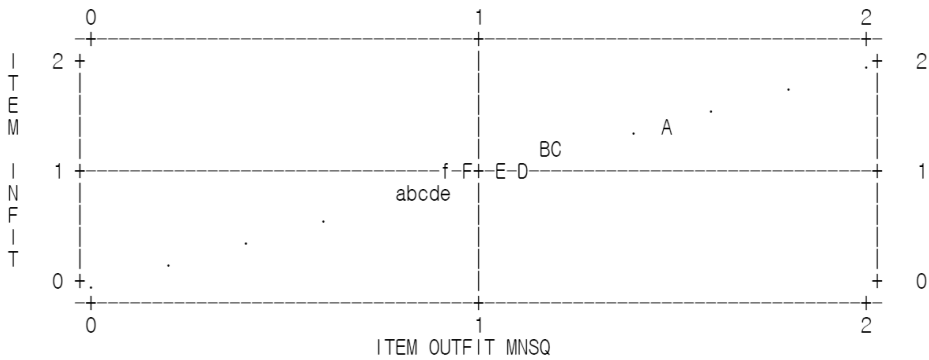
Figure 6에서 중앙의 세로선 오른쪽에는 각 문항별 선택지들의 수준에 대한 Rasch-Thurstone 임계값을 로지트로 환산한 위치가 그 문항의 선택지 수준별로 표시되어 있다. Rasch-Thurstone 임계값은 응답자들이 그 선택지를 선택할 확률과 그보다 낮은 수준의 선택지를 선택할 확률의 누적값이 0.5 또는 50%가 되게 하는 응답자의 능력값을 말한다. 문항 반응 이론에서는 그 문항을 맞혀서 해당하는 점수를 얻을 확률이 0.5가 되게 하는 응답자의 능력을 그 문항의 곤란도로 정의하므로, 이 연구에 적용한 Rasch 모델의 부분점수모형에서 각 문항별 선택지의 Rasch-Thurstone 임계값은 그 선택지의 곤란도를 나타낸다고 할 수 있다. 예를 들면, 6E. 3 은 6번 문항(지구의 운동)의 수준 3에 해당하는 선택지를 의미하는데, 이 선택지의 로지트 값은 1.0 이므로, 학생들 중 천문 시스템의 이해에 대한 능력의 로지트 값이 1.0보다 높은 학생들은 이 선택지를 선택하거나 이보다 더 높은 수준의 선택지

를 선택하여 점수를 획득할 것이고, 능력의 로지트 값이 1.0보다 낮은 학생들은 이보다 더 낮은 선택지를 선택하거나 이 선택지를 선택하지 않아 해당하는 점수를 획득하지 못한다는 것이다. 따라서 Rasch-Thurstone 임계값이 높은 문항 선택지들은 피험자 학생들에게 더 어려운 것이고, 이 임계값이 낮은 문항 선택지들은 학생들에게 쉬운 것에 해당한다.

한편, Figure 6에서 중앙 세로선 오른쪽에 있는 “M”은 문항별 곤란도의 평균값으로서 Rasch 모델에서 자동적으로 0에 맞추어져 있다. 즉, 평균적인 곤란도를 나타내는 문항의 로지트 값이 0이 되게 환산되어 다른 문항의 곤란도를 상대적인 로지트 값으로 표시하는 것이다. 따라서 12번 문항(12S: 태양계)의 두 번째 수준(12S. 2)에 해당하는 선택지가 평균적인 곤란도에 해당하며 이보다 더 어려운 선택지들은 그 위쪽에, 더 쉬운 선택지들은 그 아래쪽에 배치된다. 피험자의 능력 평균의 로지트 값과 문항 곤란도 평균의 로지트 값을 비교하면, 학생들은 이 연구에 사용한 순위 선다형 평가 문항의 평균적인 곤란도보다 조금 높은 수준의 평균적인 능력을 가지고 있음을 알 수 있다. 따라서 평균적으로 이 검사문항들이 천문 시스템에 대한 학생들의 이해 수준을 측정하는데 어려운 문항은 아니었음을 알 수 있다.

Wright map에서 천문 시스템의 이해에 대한 학생들의 과제 수행 능력과 각 문항 선택지들의 수준별 배치를 보면 다음과 같은 특징을 발견할 수 있다. 첫째, 수준 1의 곤란도에 해당하는 로지트 값(예: 12S. 1, 6E. 1 등)들은 이 연구의 검사에 응답한 학생들 중 그 위치에 해당하는 학생들이 표시되어 있지 않다. 그러므로 평가 결과에 의하면 학생들의 능력 수준은 모든 문항들의 수준 1에 해당하는 선택지의 수준보다는 높았으며, 천문 시스템의 학습 발달과정에서 하위 정착점은 구인구성도에서 예상한 것과 달리 수준 1에 해당하는 과제 곤란도보다 높은 위치에 있음을 알 수 있다. 이러한 결과는 검사 결과에 기반한 타당성의 근거로서 제시되어 이후 검사 문항을 수정할 때 수준 1에 해당하는 선택지를 재수정하는 근거가 될 수 있다. 둘째, 사용된 순위 선다형 평가 문항 중에서 문항 4E, 8EM, 및 9EM은 각각 상대적으로 낮은 로지트 값을 보이므로 가장 쉬운 문항들에 해당하고, 문항 11S, 14GU, 6E는 상대적으로 높은 로지트 값을 보이므로 가장 어려운 문항들에 해당함을 알 수 있다. 이것은 하루 동안의 지구의 운동(4E), 지구-달계의 구조(8EM)와 지구 및 달의 상대적 위치 변화(9EM)는 학습 발달과정에서 쉬운 수준에 해당하며, 태양계 행성의 운동을 공간적으로 이해하기(11S)와 팽창하는 우주에 대한 시스템적 이해(14GU), 계절 변화와 관련된 지구의 운동 이해하기(6E)는 학습 발달과정에서 어려운 수준에 해당함을 말해 준다. 셋째, 문항 11S, 14GU, 6E를 제외한 나머지 문항들의 수준2, 수준3, 수준4 선택지들은 구인구성도에서 예상했던 학습 발달의 경로와 대체로 부합한다는 것을 알 수 있다. 이것은 학생들의 발달이 낮은 수준에서 높은 수준의 방향으로 전개된다고 하면, 천문 시스템의 이해와 천문학적 사고 수행 능력은 각 문항들의 수준 2에서 수준 3을 거쳐 수준 4에 이르는 경로를 따라 발달한다고 주장할 수 있는 근거가 될 수 있다. 넷째, 문항 11S, 14GU, 그리고 6E는 다른 문항들과 비교할 때 선택지들의 수준을 한 단계 씩 상향 수정할 필요성이 제기되었다. 즉, 세 문항의 수준 2에 해당하는 선택지 11S. 2는 10EM. 3과, 6E. 2는 4E. 3과 동일한 곤란도의 로지트 값을, 14GU. 2는 9EM. 3과 7EM. 3의 중간 정도에 해당하는 곤란도의 로지트 값을 나타낸다. 마찬가지로 14GU. 3은 10EM. 4와, 6E. 3은 8EM.

TABLE 9.2 Y1_AstronomyLPs.xlsx ZOU309WS.TXT Aug 18 23:50 2014
 INPUT: 1289 PERSON 12 ITEM REPORTED: 1289 PERSON 12 ITEM 5 CATS WINSTEPS 3.81.0



MNSQ	A: 11S	B: 15GU	C: 7EM	D: 14GU	E: 4E	F: 9EM	f: 8EM	e: 12S	d: 10EM	c: 6E	b: 5E	a: 13S
infit	1.48	1.18	1.15	1.08	1.09	.97	.92	.89	.88	.84	.83	.79
outfit	1.50	1.18	1.16	1.10	1.07	.96	.91	.90	.87	.85	.83	.80

Figure 7. MNSQ value of item infit and outfit

4와 동일한 곤란도의 로지트 값을, 그리고 11S, 3은 10EM, 4보다 약간 높은 곤란도의 로지트 값을 나타낸다. 따라서 이들 세 문항에서 수준 2로 작성된 선택지들은 다른 문항의 수준 3과, 세 문항에서 수준 3으로 작성된 선택지들은 다른 문항의 수준 4에 해당하는 곤란도 수준을 가진다고 할 수 있다. 한편, 세 문항의 수준 4로 작성된 선택지들의 곤란도와 일치하는 로지트 값을 가진 학생들은 높은 능력을 가진 것이므로 비록 그 숫자가 많지 않더라도 상위 정착점으로 재조정되는 것이 필요함을 알 수 있다.

이러한 분석 결과와 함께 Rasch 모델에 기반한 문항 반응 분석 프로그램인 Winsteps에서는 각 문항들에 대한 학생들의 응답 결과가 Rasch 모델에서 계산된 것과 부합하는 정도를 표현하기 위하여 item fit에 대한 결과를 제공해 준다. 즉, Rasch 모델이 계산한 피험자의 능력과 문항 반응에 대한 결과가 예상한 것과 일치하는 정도를 나타내는 것이다. Fit 과 관련해서 응답자, 검사도구, 그리고 평가 구인에 대한 검사 문항과 응답자의 관계를 예상할 때 사용된 이론을 함께 고려해야 한다 (Boone, Staver, & Yale, 2014). 어려운 문항인데 낮은 수준의 능력을 가진 학생이 높은 수준의 선택지를 선택하거나, 쉬운 문항의 낮은 수준 선택지를 높은 수준의 능력을 가진 학생이 실수하여 선택한 경우는 올바른지 않은 fit, 즉 misfit에 해당한다. 이와 같이 예상되는 결과에서 벗어난 사례(outlier)에 대한 misfit을 outfit 이라고 한다. 이에 반해 infit은 주어진 문항의 곤란도와 피험자의 능력 수준에 맞게 배치되는 사례(inlier)가 발생하는 정도를 나타낸 것이다. 따라서 문항과 피험자의 적절성을 조사할 때 outfit을 중심으로 판단하는 경우가 일반적이다. Boone et al. (2014)에 의하면, 일반적으로 fit 통계에서 주로 비교하는 값은 outfit MNSQ (mean square value)이다. MNSQ는 outfit과 infit에 대한 chi-square 계산 값이다. MNSQ의 계산은 1.0을 평균값으로 환산하기 때문에 일반적으로 수용 가능한 MNSQ 값의 범위를 0.5 ~ 1.5로 보며(Wright & Linacre, 1994), MNSQ 값이 1.0보다 크면 설명되지 않는 변수가 너무 많은 경우(underfit)으로 판단하고, 1.0보다 작으면 모델이 데이터를 과도하게 예측하는 경우(overfit)로 판단한다.

이 연구에 적용한 순위 선다형 평가 문항의 응답 반응 결과를 Rasch 모델로 측정한 후 도출된 문항의 fit 값을 Figure 7에 제시하였다. 분석 결과 연구에 적용된 12 문항이 모두 적합한 범위 내의 문항 fit 값을

보였다. Figure 7에서는 F 또는 f에 해당하는 문항들이 가장 적절한 item fit 값을 갖는 것으로 표시된다. F에서 대문자 A로 갈수록 설명되지 않는 변수가 많은 under fit에 가까워지며, f에서 a로 갈수록 과도하게 예측하는 overfit에 가까워지는 문항에 해당된다. A로 표시된 문항 11S의 경우 outfit 값이 1.50으로 적합한 outfit 범위에 있고 infit 값도 1.48이므로 적절한 fit 값을 나타낸다. 그러나 outfit의 경계값에 가까운 데다 앞서 Wright map을 분석한 결과에도 선택지의 재구성이 필요하다고 판정되었으므로, 2차 검사를 위한 문항 수정 과정에서 주의 깊은 검토가 필요한 것을 알 수 있다.

3. 가설적 학습 발달과정

Figure 6에 표시된 Level 1(4E, 2 ~ 10EM, 2), Level 2(4E, 3 / 6E, 2 ~ 10EM, 3 / 11S, 2), Level 3(4E, 4 ~ 11S, 3), 그리고 Level 4(6E, 4 ~ 11S, 4)는 구인구성도에 맞추어 작성된 순위 선다형 평가 문항의 수준별 선택지들의 수준 구분(문항별 .1, .2, .3, .4) Wright map을 분석한 결과를 바탕으로 재구성하여 전체적으로 네 범위로 다시 나눈 것이다. 초기 구인구성도에서 구분한 수준 1(문항별 .1)은 학생들의 학습 발달과정에서 제외되었고, 구인구성도의 수준 2(문항별 .2)가 대부분 수정된 수준 1(Level 1)에 배치되었다. 마찬가지로 구인구성도의 수준 3(문항별 .3)은 수정된 수준 2(Level 2), 구인구성도의 수준 4(문항별 .4)는 수정된 수준 3(Level 3), 그리고 문항 11S, 4, 14GU, 4, 6E, 4 는 수정된 수준 4(Level 4)에 배치되었다. 이렇게 나눈 네 범위는 천문 시스템에 대한 학생들의 이해 수준의 구분을 수정하는 기준으로 활용될 수 있다. 학습의 발달 경로가 낮은 과제 곤란도에서 어려운 곤란도로, 그리고 학생들의 낮은 능력에서 높은 능력을 향하는 방향으로 전개됨을 전제로 한다면, Wright map을 분석하여 수정된 수준 구분을 근거로 천문 시스템의 이해에 대한 학생들의 가설적인 학습 발달과정을 제시할 수 있다. Table 3에 제시한 천문 시스템의 가설적 학습 발달과정은 지구의 운동, 지구-달 계, 태양계 및 은하와 우주에 대하여 이 연구의 평가 문항에 포함된 소주제들을 이해하는데 필요한 천문학적 사고로서 공간적 사고와 시스템 사고의 수행 수준을 기준으로 정리되었다.

Table 3. Hypothetical learning progressions for astronomical systems with spatial and systems thinking

	수준 1	수준 2	수준 3	수준 4
지구의 운동	<ul style="list-style-type: none"> • 4E: 지구가 자전할 때 태양과 지구의 운동 방향을 서->동으로 기술 • 5E: 지구가 공전할 때 별자리의 운동 방향을 동->서로 기술함 	<ul style="list-style-type: none"> • 4E/5E: 우주에서 본 관점에서 지구의 자전(반시계방향), 태양의 일주운동(동->서) 및 지구의 공전(반시계방향), 별자리의 연주운동(동->서)을 서술 • 6E: 계절에 따라 태양의 일주권이 길이와 높이가 변하는 패턴을 인식하고, 낮의 길이가 변화함을 인식한다. 	<ul style="list-style-type: none"> • 4E/5E: 우주에서 내려다 본 관점에서 지구의 자전과 공전 모델을 적용하여 지구에서 바라본 태양의 일주운동 및 별자리의 연주운동을 관점 전환하여 설명함 • 6E: 계절에 따라 태양 일주운동 시 남중고도의 변화에 따라 낮의 길이가 달라지는 상호 연관성을 파악할 수 있음 	<ul style="list-style-type: none"> • 6E: 우주에서 내려다 본 관점에서 모델을 적용하여 태양의 일주권의 변화와 낮의 길이 변화의 관계를 설명할 수 있음
지구 - 달 계	<ul style="list-style-type: none"> • 9EM: 태양 중심의 지구-달 계 모델에서 달의 위치 변화와 운동을 오른쪽에서 왼쪽으로 이동하는 것으로 표시함 • 7EM: 하루 동안 지구에서 본 태양과 달, 별의 위치 변화와 운동을 부정확하게 이해함 • 8EM: 태양 중심의 지구-달 계 모델로 일주운동을 표현하지만, 태양, 달, 별의 상대적 위치 관계를 인식하지 못함 • 10EM: 달의 위상 변화를 나타내는 태양 중심의 지구-달 계 모델에서 시간에 따른 달의 위치 변화를 달의 위상과 연결시키지 못함 	<ul style="list-style-type: none"> • 9EM: 태양 중심의 지구-달 계 모델에서 지구의 위치 변화와 운동을 2차원 평면적으로 설명함 • 7EM: 하루 동안 지구에서 본 태양, 달, 별의 위치 변화를 본질적으로 인식함 • 8EM: 하루 동안 태양, 달, 별의 상대적 위치 관계를 우주에서 내려다 본 2차원적 태양 중심 모델로 설명 • 10EM: 2차원 평면으로 표시된 태양 중심의 지구-달 계 모델에서 달의 위치와 달의 위상을 연결시켜 설명할 수 있음 	<ul style="list-style-type: none"> • 9EM: 우주에서 내려다 본 지구-달 계 모델에서 지구의 위치변화와 운동을 3차원적 모델로 묘사할 수 있음 • 7EM: 하루 동안 지구에서 본 태양, 달, 별의 위치 변화를 연속적인 일주권으로 인식함 • 8EM: 하루 동안 태양, 달, 별의 상대적 위치 관계를 우주에서 내려다 본 3차원적 모델로 묘사하여 지구에서 관측한 자료와 연관시킬 수 있음 • 10EM: 3차원적으로 표현된 태양 중심의 지구-달 계 모델에서 달의 위치와 달의 위상을 연결시켜 설명할 수 있음 	
태양계	<ul style="list-style-type: none"> 12S: 행성이 공전하는 동안 화성, 토성의 위치 변화를 인식하지만, 배경 별자리에 대한 상대적인 운동 방향을 파악하지 못하고, 2차원 평면에서 보이는 그대로 서술함 	<ul style="list-style-type: none"> • 11S: 2차원 평면 자료에서 화성과 토성의 운동 방향을 정확한 방위로 나타낼 수 있으나, 지구에서 바라본 관점으로만 이해함 • 12S: 우주에서 내려다 본 관점에서 작성된 태양 중심의 태양계 모델에 화성, 토성의 운동 방향을 나타낼 수 있으나, 두 행성의 정확한 상대적인 위치를 파악하지 못함 • 13S: 서로 다른 비율로 표시된 행성들의 크기를 비교하여 지구형 행성과 목성형 행성을 구분할 수 있음 	<ul style="list-style-type: none"> • 11S: 태양 중심의 태양계 모델에 대한 지식을 바탕으로 화성과 토성의 공전 방향을 서술할 수 있음 • 12S: 화성과 토성의 운동을 배경 별자리에 대한 상대적 위치 변화로 파악하여, 우주에서 내려다 본 관점으로 작성된 태양계 모델에 두 행성의 운동 방향과 상대적인 위치를 정확히 표현할 수 있음 • 13S: 행성의 크기와 무게가 서로 다른 비율로 표시된 자료에서 크기와 무게의 상호연관성을 파악하여 작고 가벼운 행성들과 크고 무거운 행성들을 분류함 	<ul style="list-style-type: none"> • 11S: 2차원 평면 자료에서 화성과 토성의 배경 별자리에 대한 상대적인 위치 변화를 근거로 화성과 토성의 공전 속도와 운동 방향을 설명할 수 있음
은하와 우주	<ul style="list-style-type: none"> • 15GU: 여름/겨울에 은하수의 두께가 다름을 인식하고, 지구와 태양의 위치를 정확히 나타내지만, 태양계를 은하 외부에 표시함 	<ul style="list-style-type: none"> • 14GU: 모든 은하들이 멀어짐을 인식함 • 15GU: 여름/겨울에 은하수의 두께와 지구와 태양의 위치가 관련되어 있음을 인식하지만, 태양계를 은하 원반면에 포함시켜 표현함 	<ul style="list-style-type: none"> • 14GU: 우주 공간의 팽창을 인식함 • 15GU: 태양계를 은하 원반면에 포함시키며, 여름/겨울의 은하수의 두께와 방향, 은하수의 각도 및 지구/태양의 위치의 상호연관성을 이해함 	<ul style="list-style-type: none"> • 14GU: 멀리 있는 은하가 더 빠른 속도로 멀어짐을 인식함

※ 음영으로 표시된 항목은 시스템 사고, 나머지 항목은 공간적 사고에 해당함

학습 발달과정의 하위 정착점에 해당하는 Level 1의 학생들은 태양의 일주운동, 지구 - 달 계의 일주운동과 별자리의 연주운동, 달의 공전 운동, 행성의 공전 운동에서 천체의 위치 변화와 운동의 방향을 평면 자료에서 보이는 그대로 기술하는 공간적 사고를 지녔다. 또한, 이 수준의 학생들은 태양 중심의 지구-달 계의 모델에서 천체의 위치를 부정확하게 연결하거나, 달의 위상에 따른 달의 위치를 연결하지 못하며, 태양계 행성들의 물리량을 지구와 단순 비교하는 정도의 정량적 사고와 계절에 따른 은하수의 차이를 인식하는 수준의 시스템 사고 수준을 보였다.

학습 발달과정의 중간 단계에 해당하는 Level 2와 Level 3은 구체적 인 발달의 경로를 보여준다. 즉, 우주에서 내려다 본 관점에서 지구의 자전, 태양의 일주운동, 지구의 공전, 별자리의 연주운동의 방향을 서술하는 공간적 사고 수준(Level 2)에서 이것을 지구에서 관측되는 모습과 연관시켜 관점의 전환 능력을 보이는 공간적 사고 수준(Level 3)으로 발달한다. 그리고 낮의 길이가 변화함을 인식하고 일주권의 길이와 높이가 변하는 패턴을 파악하는 시스템 사고 수준(Level 2)에서 남중고도의 변화와 낮의 길이 사이의 상호연관성을 파악하는 시스템 사고 수준(Level 3)으로 발달한다. 한편, 지구 - 달 계에서 달의

공전은 우주에서 내려다 본 2차원 평면적 이해를 나타내는 공간적 수준(Level 2)에서 지구-달 계의 3차원적 모델로 묘사하여 지구에서 관측한 자료와 연결시키는 공간적 사고 수준(Level 3)으로 발달한다. 그리고 2차원적 지구-달 계 모델에서 달의 위치와 달의 위상을 연결시킬 수 있는 시스템 사고 수준(Level 2)에서 이것을 3차원적 모델로 설명할 수 있는 시스템 사고 수준(Level 3)으로 발달한다. 태양계에서 행성의 공전 운동은 행성의 운동 방향을 지구에서 바라본 관점에서 방위로 나타내고 이 방향을 우주에서 내려다 본 태양계 모델에 표시하는 정도의 공간적 사고 수준(Level 2)에서 우주에서 내려다 본 관점으로 작성된 태양 중심의 태양계 모델에 대한 지식을 바탕으로 화성과 토성의 공전 방향과 궤도 상의 상대적인 위치를 정확히 표현할 수 있는 공간적 사고 수준(Level 3)으로 발달한다. 태양계 행성의 물리량에 대한 시스템 사고의 발달은 서로 다른 비율로 표시된 행성의 크기를 근거로 태양계 행성을 분류하는 수준(Level 2)에서 서로 다른 비율로 표시된 행성의 크기와 무게의 상호연관성을 파악하여 작고 가벼운 행성과 크고 무거운 행성을 분류하는 수준(Level 3)으로 발달한다.

학습 발달과정의 상위 정착점에 해당하는 Level 4의 학생들은 우주에서 내려다 본 관점에서 형성된 천체의 운동 모델을 지구에서 바라본 관점에서 얻은 자료와 서로 연관시키는 관점 전환의 공간적 사고를 지구의 운동, 지구-달 계의 운동, 태양계 행성의 운동, 은하와 우주의 운동에 적용할 수 있었다. 또한 상위 정착점에 도달한 학생들은 천문 시스템의 설명 모델을 구성하는 하위 요소들 간의 상호 연관성을 파악하고 이를 설명 모델에 적용할 수 있는 시스템 사고를 나타냈다. 상위 정착점에 해당하는 천문 시스템의 내용 요소들은 태양의 일주권과 낮의 길이 변화, 행성의 공전 속도, 우주의 팽창 등이었다.

IV. 결론 및 논의

이 연구에서는 구인 모델링 방식의 4 단계 평가 시스템 구성단위를 거쳐 천문 시스템의 이해에 대한 평가 문항을 개발하고, 이를 초중등 학생들에게 적용하여, 천문학적 사고(공간적 사고와 시스템 사고)에 기반한 천문 시스템의 가설적 학습 발달과정을 제시하였다. 또한, 학습 발달과정을 도출하는 동안 검사 결과에 근거한 학습 발달과정의 타당화 과정에 대한 구체적 근거들을 확보할 수 있었다.

연구 결과, 학생들의 천문 시스템에 대한 이해 수준을 조사하기 위한 순위 선다형 평가 문항들은 천체의 운동과 구조를 중심으로 천문 시스템의 핵심 개념과 과학적 탐구실행으로서 천문학적 사고를 복합적으로 포함할 수 있도록 제작되었다. 이는 구인특화 단계에서 작성된 천문 시스템의 개념에 대한 구인구성도와 천문학적 사고에 대한 구인구성도에 근거하여 문항 개발 단계에서 순위 선다형 평가 문항의 선택지를 구성함으로써 가능하였다. 또한, 평가결과 기술 단계와 측정 모델 적용 단계의 연구 결과는 이 연구의 순위 선다형 평가 문항들이 천문 시스템에 대한 학습 발달과정을 조사하기 위한 검사 도구로서 그 특성을 잘 구현하고 있음을 보여주었다. Item fit 분석 결과는 문항 반응 결과들이 Rasch 모델로 예측한 결과에 부합하였음을 말해주며, Wright map 분석 결과는 이 연구의 평가 문항들을 통해 천문 시스템에 대한 학습 발달과정의 가설적 발달 경로를 나타낼 수 있음을 보여주었다. 학습 발달과정의 하위 정착점에서 학생들은 태양의 일주운동과 별자리의 연주운동, 단기간의 달의 위치 변화, 몇 달 동안 행성의 위치

변화 등에 대하여 지구에서 보는 관점에서 2차원 평면으로 해석하는 공간적 사고 수준을 보였으며, 지구-달 계의 모델에서 천체의 위치와 변화를 부정확하게 연결하거나, 태양계 행성들의 물리량과 은하수에 대한 단순한 패턴을 인식하는 시스템 사고 수준을 보였다. 학습 발달과정의 중간 단계에서 학생들은 우주에서 내려다 본 지구 및 지구-달 계, 태양계의 모델을 근거로 천체의 운동을 해석하는 공간적 사고의 수준을 나타냈으며, 천체의 구조를 구성하는 요소들을 이 모델들과 연관시키는 시스템 사고를 보였다. 학습 발달과정의 상위 정착점에서 학생들은 우주에서 내려다 본 관점에서 형성된 천체의 운동 모델을 지구에서 바라본 관점에서 얻은 자료와 서로 연관시키는 관점 전환의 공간적 사고를 지구의 운동, 지구-달 계의 운동, 태양계 행성의 운동, 은하와 우주의 운동에 적용할 수 있었다. 또한 상위 정착점에 도달한 학생들은 천문 시스템의 설명 모델을 구성하는 하위 요소들 간의 상호 연관성을 파악하고 이를 설명 모델에 적용할 수 있는 시스템 사고를 나타냈다.

한편, 이 연구의 결과로 얻은 천문 시스템에 대한 가설적 학습 발달과정을 토대로 지구의 운동, 지구-달 계의 운동과 구조, 태양계의 운동과 구조, 은하와 우주에 대한 성취 수준을 학습 발달의 단계별로 제시할 수 있다. 그리고 천문 시스템을 구성하는 하위 시스템으로서 지구, 지구-달 계, 태양계, 및 은하와 우주의 학습 과제 곤란도의 위계를 구분할 수 있다. 즉, Figure 6에 제시된 Wright map을 보면, 지구의 자전과 공전 및 일주운동, 연주운동, 지구-달 계에서 달의 위치와 운동 및 달의 위상, 태양계 행성의 구분 및 행성의 운동에 대한 기초적 이해 등은 비교적 저학년 또는 낮은 학교급에서 학습될 수 있지만, 지구의 공전과 낮의 길이 변화, 행성의 시운동 속도 차이, 은하의 구조, 우주의 팽창 등은 높은 학교급에서 학습되는 것이 적절함을 알 수 있다. 이러한 결과들은 지구과학 교육과정에서 학년별로 구성되는 학습 내용 체계의 세부 항목의 수준을 학생들이 수행한 학습 평가의 결과에 근거하여 설정하기 위한 근거로서 기초 자료가 될 수 있다. 즉, 천문 시스템의 학습 발달과정 연구를 통해 학생들의 학습 성취에 대한 평가의 결과를 교육과정의 개발에 연계시켜 줄 수 있는 방법론적 틀(Maeng et al., 2013)을 마련할 수 있는 것이다.

학습 발달과정이 국내 과학교육계에 소개된 이래 학습 발달과정 연구의 학술적, 교수학적 및 정책적 의의를 인식하고 이 분야의 연구에 참여하는 연구자들이 국내에서도 점차 늘어나고 있다. 학습 발달과정이 처음 제안되었던 미국에서 초기에 교육 연구자들 사이에 학습 발달과정의 의미나 접근 방식, 연구 방향에 대하여 여러 가지 견해들이 난립하기도 했던 것처럼 국내의 과학교육계에서도 학습 발달과정에 대한 부분적인 이해나 오해들이 발견되고 있다. 이러한 상황에서 이 연구는 학습 발달과정을 연구하고 조사하는데 필요한 다음과 같은 몇 가지 시사점을 제공해 준다. 첫째, 학습 발달과정 연구는 과학적인 평가 시스템을 바탕으로 설계되고 수행되어야 한다. 이 연구에서 학습 발달과정을 조사하기 위한 평가 시스템으로 구인 모델링 방식을 적용하였다. 구인 모델링 방식은 미국의 학습 발달과정 연구에서 많이 활용되는 평가 시스템이지만, 학습 발달과정 연구에 유일한 평가 시스템은 당연히 아니다. 연구자의 관점과 연구의 목적에 따라 적절한 평가 시스템을 선정하여 연구를 설계하는 것이 필요할 것이다. 평가 시스템에 대해서는 Knowing what students know: The science and design of educational assessment (NRC, 2001)을 참고할 수 있다. 둘째, 학습

발달과정은 특정한 과학 내용 지식을 개념적 위계에 따라 나열한 것이 아니라, 과학 개념을 이해할 때 활용되는 과학적 탐구실행 능력의 수행 정도에 따라 도달하는 개념 이해 수준의 차이를 바탕으로 제시되어야 한다. 이를 위해서 학습 발달과정의 주제는 과학적 탐구실행과 과학 내용이 서로 증첩된 형태로 선정되어야 하며(Maeng *et al.*, 2013), 학습 발달과정을 조사하기 위한 평가 문항 역시 과학 개념과 과학적 탐구실행이 통합된 형태로 개발되어야 한다.

이 연구는 구인 모델링 방식의 평가 시스템 구성단위의 1차 순환을 거쳐 도출된 결과를 중심으로 제시하였다. 그러나 학습 발달과정 연구는 여러 차례 반복적인 평가 시스템을 거쳐 확립되는 것이다. 이 연구의 결과도 이후 연차적으로 반복적인 검사를 거쳐서 가설적 학습 발달 과정을 수정하여 발전시킬 계획이다. 현재 도출된 천문 시스템의 학습 발달과정은 평가 결과에 근거하였지만 여전히 가설적이다. 왜냐하면, 학습 발달과정의 본질적 의미는 “과학의 핵심 개념에 대한 이해 및 그와 관련된 과학 탐구의 실행 능력이 적절한 과학 교수활동에 의해 더 정교해지는 과정에 대한 경험적 근거에 기반한 검증 가능한 가설”(Corcoran *et al.*, 2009, p. 15)이기 때문이다. 즉, 학습 발달과정의 본질적 특성을 갖추기 위해서는 학생들의 학습 발달을 유도하는 적절한 교수활동을 제시하고, 그 수업활동의 결과로 형성된 학습의 발전 양상을 제시할 수 있어야 하는 것이다. 이 때 적용되는 교수활동은 이전 연구의 평가 결과에서 도출된 학생들의 발달 수준 및 발달의 가설적 경로를 반영하여 학생들의 학습 발전을 촉진시키는 이른바 평가 적응적 교수활동(adaptive instruction, Corcoran & Silander, 2009)을 말한다. 천문 시스템의 학습 발달과정을 위한 후속 연구는 이 연구에서 도출된 가설적 학습 발달과정에 근거하여 평가적응적 교수활동을 개발하고 이를 실제 수업에 적용하고, 수정된 학습 평가를 실시하여 더욱 경험적이고 구체적인 학습 발달의 과정을 탐색하게 될 것이다.

국문요약

이 연구에서는 지구, 지구 - 달 계, 태양계, 그리고 은하와 우주를 포함하는 천체의 구조와 운동을 종합한 천문 시스템(astronomical systems)의 학습 발달과정(learning progressions)을 조사하기 위하여, 순위 선다형 평가 문항을 개발하고 이를 초등/중학생들에게 적용한 검사 결과에 기반하여 학습 발달과정 해석의 타당성 검증의 근거를 제시하였다. 연구의 과정은 ‘구인특화, 평가문항 개발, 평가결과 기술, 측정 모델’로 이루어진 4 단계의 구인 모델링 방식(construct modeling approach)에 기반하여 진행되었다. 천문 시스템의 내용적 평가 구인으로 천체의 운동과 구조를 선정하였고, 탐구실행적 평가 구인으로 공간적 사고와 시스템 사고와 같은 천문학적 사고를 선정하였다. 이 연구에서 개발된 순위 선다형 평가 문항들은 천문 시스템에 대한 학습 발달과정을 조사하기 위한 검사 도구로서 적절하였다. 즉, item fit 분석 결과는 학생들의 문항 반응 결과가 Rasch 모델로 측정된 결과와 부합하였다. 그리고 Wright map 분석 결과는 이 연구의 평가 문항들이 천문 시스템에 대한 학생들의 가설적 발달 경로를 조사하는데 매우 효과적임을 보여주었다. 학습 발달과정의 하위 정착점에서 학생들은 천체들의 위치와 운동 방향의 변화를 지구에서 보는 관점에서 2차원 평면으로 해석하는 공간적 사고 수준을 보였으며, 지구-달 계의 모델에서 천체의 위치와 변화를 부정확하게 연결하거나, 태양계 행성들과 은하

수에 대한 단순한 패턴을 인식하는 시스템 사고 수준을 보였다. 학습 발달과정의 중간 단계에서 학생들은 우주에서 내려다 본 지구 및 지구-달 계, 태양계의 모델을 근거로 천체의 운동을 해석하는 공간적 사고의 수준을 나타냈으며, 천체의 구조를 구성하는 요소들을 이 모델들과 연관시키는 시스템 사고를 보였다. 학습 발달과정의 상위 정착점에서 학생들은 우주에서 내려다 본 관점에서 형성된 천체의 운동 모델을 지구에서 바라 본 관점에서 얻은 자료와 서로 연관시키는 관점 전환의 공간적 사고를 지구의 운동, 지구-달 계의 운동, 태양계 행성의 운동, 은하와 우주의 운동에 적용할 수 있었다. 또한 상위 정착점에 도달한 학생들은 천문 시스템의 설명 모델을 구성하는 하위 요소들 간의 상호 연관성을 파악하고 이를 설명 모델에 적용할 수 있는 시스템 사고를 나타냈다.

주요어 : 학습 발달과정, 천문 시스템, 천문학적 사고, Rasch 모델, 타당성의 근거

References

- Achieve, Inc. (2013). Next Generation Science Standards. Achieve Inc. On behalf of the twenty-six states and partners that collaborated on the NGSS.
- AERA, APA, & NCME. (1999). Standards for educational and psychological testing. Washington, DC: AERA.
- Alonzo, A. C., & Steedle J. T. (2009). Developing and assessing a force and motion learning progression. *Science Education*, 93(3), 389-421.
- Ben-Zvi Assaraf, O., & Orion, N. (2005). Development of system thinking skills in the context of Earth System education. *Journal of Research in Science Teaching*, 42, 518-560.
- Ben-Zvi Assaraf, O., & Orion, N. (2010). System thinking skills at the elementary school level. *Journal of Research in Science Teaching*, 47, 540-563.
- Berland, L.K., & McNeill, K.L. (2010). A learning progression for scientific argumentation: Understanding student work and designing supportive instructional contexts. *Science Education*, 94, 765-793.
- Black, P. & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education*, 5, 7-74.
- Black, P., Wilson, M., & Yao, S. Y. (2011). Road maps for learning: A guide to the navigation of learning progressions. *Measurement: Interdisciplinary Research & Perspective*, 9, 71-123.
- Bond, T. G., & Fox, C. M. (2007). *Applying the Rasch model: Fundamental measurement in the human sciences* (2nd edition). New York, NY: Routledge.
- Boone, W. J., Staver, J. R., & Yale, M. S. (2014). *Rasch analysis in the human sciences*. New York, NY: Springer.
- Briggs, D. C., & Alonzo, A. C. (2012). The psychometric modeling of ordered multiple-choice item responses for diagnostic assessment with a learning progression. In A. C. Alonzo & A. W. Gotwals (Eds.), *Learning progressions in science: Current challenges and future directions* (pp. 293-316). Rotterdam, The Netherlands: Sense Publishers.
- Briggs, D. C., Alonzo, A. C., Schwab, C., & Wilson, M. (2006). Diagnostic assessment with ordered multiple-choice items. *Educational Assessment*, 11(1), 33-63.
- Corcoran, T., & Silander, M. (2009). *Instruction in high schools: The evidence and the challenge*. The Future of Children: America's High Schools, 19, 157-183.
- Corcoran, T., Mosher, F. A., & Rogat, A. (2009). *Learning progressions in science: An evidence based approach to reform*. Consortium for Policy Research in Education Report #RR-63. Philadelphia, PA: Consortium for Policy Research in Education.
- Duncan, R. G., Rogat, A. D., & Yarden, A. (2009). *A learning progression*

- for deepening students' understandings of modern genetics across the 5th-10th grades. *Journal of Research in Science Teaching*, 46, 655-674.
- Duschl, R. A., & Grandy, R. (Eds.) (2008). *Teaching scientific inquiry: Recommendations for research and implementation*. Rotterdam, The Netherlands: Sense Publishers.
- Duschl, R., Maeng, S., & Sezen, A. (2011). Learning progressions and teaching sequences: A review and analysis. *Studies in Science Education*, 47, 123-182.
- Furtak, E. M. (2012). Linking a learning progression for natural selection to teachers' enactment of formative assessment. *Journal of Research in Science Teaching*, 49(9), 1181-1210.
- Gotwals, A. W., & Songer, N. B. (2013). Validity evidence for learning progression-based assessment items that fuse core disciplinary ideas and science practices. *Journal of Research in Science Teaching*, 50, 597-626.
- Ji, E., & Chae, S. (2000). *Theory and practices of Rasch model*. Seoul: Kyoyook-Kwahak-Sa.
- Jin, H., & Anderson, C. W. (2012). A learning progression for energy in socio-ecological systems. *Journal of Research in Science Teaching*, 49, 1149-1180.
- Maeng, S., Seong, Y., & Jang, S. (2013). Present states, methodological features, and an exemplar study of the research on learning progressions. *Journal of the Korean Association for Science Education*, 33, 161-180.
- Masters, G. N. (1982). A Rasch model for partial credit scoring. *Psychometrika*, 47(2), 149-174.
- National Research Council (NRC). (2006). *Learning to think spatially*. Washington, DC: National Academies Press.
- National Research Council. (2007). *Taking science to school: Learning and teaching science in grades K-8*. Washington, DC: National Academies Press.
- National Research Council. (2012). *A framework for k-12 science education: Practices, crosscutting concepts, and core ideas*. Washington, DC: National Academies Press.
- Neumann, K., Viering, T., Boone, W. J., & Fischer, H. E. (2013). Towards a learning progression of energy. *Journal of Research in Science Teaching*, 50, 162-188.
- Orion, N., & Basis, T. (2008). Characterization of High School Students' System Thinking Skills in the Context of Earth Systems. Presented in the 2008 NARST Annual Meeting. March, 2008. Baltimore, U.S.A.
- Plummer, J. D. (2014). Spatial thinking as the dimension of progress in an astronomy learning progression. *Studies in Science Education*, 50(1), 1-45.
- Plummer, J. D., & Krajcik, J. (2010). Building a learning progression for celestial motion: Elementary levels from an Earth-based perspective. *Journal of Research in Science Teaching*, 47(7), 768-787.
- Schwarz, C. V., Reiser, B. J., Davis, E. A., Kenyon, L., Achér, A., Fortus, D., Shwartz, Y., Hug, B., & Krajcik, J. (2009). Developing a learning progression for scientific modeling: Making scientific modeling accessible and meaningful for learners. *Journal of Research in Science Teaching*, 46, 632-654.
- Seong, T. (2002). *Validity and Reliability*. Seoul: Hakjisa.
- Seong, T., & Si, K. (2006). *Research Methodology*. Seoul: Hakjisa.
- Shea, N. A., & Duncan, R. G. (2013). From theory to data: The process of refining learning progressions. *Journal of the Learning Sciences*, 22, 7-32.
- Songer, N. B., & Gotwals, A. W. (2012). Guiding explanation construction by children at the entry points of learning progressions. *Journal of Research in Science Teaching*, 49, 141-165.
- Stevens, S., Delgado, C., & Krajcik, J. S. (2010). Developing a hypothetical multi-dimensional learning progression for the nature of matter. *Journal of Research in Science Teaching*, 47, 687-715.
- Wilson, M. (2005). *Constructing measures: An item response modeling approach*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Wilson, M. (2009). Measuring progressions: Assessment structures underlying a learning progression. *Journal of Research in Science Teaching*, 46, 716-730.
- Wilson, M. (2012). Responding to a challenge that learning progressions pose to measurement practice. In A. C. Alonzo & A. W. Gotwals (Eds.), *Learning progressions in science: Current challenges and future directions* (pp. 317-343). Rotterdam, The Netherlands: Sense Publishers.
- Wright, B. D., & Linacre, J. M. (1994). Reasonable mean-square fit value. *Rasch Measurement Transactions*, 8(3), 370.