

## 2015 개정 교육과정 ‘과학탐구실험’ 평가 도구 및 평가 현황 탐색

백종호, 변태진, 이동원, 심현표\*  
한국교육과정평가원

## An Investigation on the Assessment Tool and Status of Assessment in the ‘Scientific Inquiry Experiment’ of the 2015 Revised Curriculum

Jongho Baek, Taejin Byun, Dongwon Lee, Hyeon-Pyo Shim\*  
Korea Institute for Curriculum and Evaluation

## ARTICLE INFO

## Article history:

Received 9 September 2020

Received in revised form

2 October 2020

27 October 2020

Accepted 28 October 2020

## Keywords:

Scientific inquiry experiments,  
student assessment, scientific  
inquiry, 2015 revised curriculum

## ABSTRACT

‘Scientific inquiry experiments’, which was newly created subjects in the 2015 revised curriculum, was expected in the aspect of learning science and developing core competences through science practices. Based on changed view of evaluation, assessments of a practice-centered subject ‘Scientific inquiry experiments’ should be try to conducted in various ways, but many challenges were reported. In this study, through analysis of current status of assessment of the subject, we intended to find the way of conducting and supporting ‘Scientific inquiry experiments’. We collected assessment materials and explanatory description about them from 25 teachers who taught ‘Scientific inquiry experiments’ in 2018 and 2019. And we analyzed the cases with framework which were consisted with three main categories: elements, standards, methods of assessments. Also, we investigated how the results of assessment were utilized. For the validity, we requested verification of the results of our data analysis to experts of science education and science teachers. From them, we also collected their opinions about our analysis. As a result of the study, teachers assessed some elements of inquiry skills such as ‘analysis and interpreting the data’, ‘conducting inquiry’ more than others which were closely related to what subject-matter the teachers used to organized inquiry program with. In the aspect of domain of assessments, though cognitive domain and affective domain as well as skills were evaluated, we also found that the assessment of those domains had some limitation. In terms of standard of assessment, the goals of assessment were presented in most cases, but there were relatively few cases which had the specific criteria and the stepwise statements of expected performance of students. The time and subject of the assessment were mainly post-class and teachers, and others such as in-class assessments, peer-assessments were used only in specific contexts. In all cases, the results of assessments used for calculating students’ grade, but in some cases, we could observe that the results used for improving teaching and feedback for students. Based on these results, we discussed how to support the assessments of ‘Scientific inquiry experiments’.

## 1. 서론

학교 교육에서 평가는 학생들이 교과 수업에서 설정한 목표에 얼마만큼 도달하였는지를 판단하고, 목표에 도달하기 위해서 어떠한 교육이 필요한지에 대한 정보를 제공받을 수 있다는 점에서 매우 중요한 요소이다. 그러나 평가는 성적이나 등급을 산출하는 기능 또한 담당하고 있기 때문에, 학교 현장에서 평가 본래의 목적이 희석되기 쉽고 이를 달성하기 위해서 교사들의 많은 노력이 요구되는 것이 현실이다. 한편, 과학 교육에서 과학 탐구는 오랫동안 중요한 요소로 간주되어 왔으며, 과학의 본성(Nature of science)의 관점에서 과학적 소양을 기르고 과학을 수행하는 방법을 학생들에게 가르치는 데 있어서 중요한 요소이다(Song, 2006; Lederman *et al.*, 2014). 과학 탐구의 실천을 통해서 학생들은 과학 개념에 대해 이해할 수 있고(Fogleman, McNeill, & Krajcik, 2011), 과학적 사고 및 탐구 능력을 신장시켜 나갈 수 있다. 학생들의 과학 탐구 과정에서의 평가는 이러한 성장의

촉진제로서 기능할 수 있으며, 이를 위해 적절한 평가가 무엇이고 평가 실천의 방안을 탐색해 보는 것은 과학 교육 연구의 핵심적인 과제 중 하나로 볼 수 있다.

과학 탐구의 학습 및 평가 요소를 구분하고 개념화하는 연구들은 지속적으로 수행되어 왔다. 국제 비교 연구인 PISA에서는 과학과 평가들을 지식, 태도, 맥락 차원으로 구성한 후, 이를 토대로 현상에 대한 과학적 설명, 과학 탐구의 평가 및 설계, 자료 및 증거의 과학적 해석 역량을 평가하도록 하였다(OECD, 2017). TIMSS에서도 과학과 평가 영역을 인지 영역을 알기, 적용하기, 추론하기로 구분하여 각 영역의 세부 요소들을 제안하였고(Victoria & Lee, 2017), 미국의 국가 연구 위원회의 ‘A framework for K-12 science education’에서는 과학과 공학의 8가지 실천(practices) 방안을 과학 교육의 학습 및 평가 요소들로 제안하였다(NRC, 2012). 구체적인 평가의 장면과 관련된 연구로는 과학 탐구의 요소 중 하나인 논변에 대한 평가들을 개발한 Sampson, Enderle, & Walker(2011)의 연구를 들 수 있다. 연구자

\* 교신저자 : 심현표 (hyunpyo@kice.re.kr)

\*\* 본 연구는 한국교육과정평가원의 ‘2015 개정 교육과정 ‘과학탐구실험’ 평가 현황 조사 및 평가 방안 탐색(ORM 2020-40-1)’의 일부 내용을 발췌하여 새롭게 구성하였음.  
<http://dx.doi.org/10.14697/jkase.2020.40.5.515>

들은 ASAC(Assessment of Scientific Argumentation in the Classroom) 관찰 프로토콜을 개발하고 검증하는 연구를 수행하였는데, 이 연구에서 개발한 평가틀은 활용하는 데 있어서 숙련의 과정이 필요하다는 단점은 있으나 구체적인 평가의 틀을 제안하였다는 점에서 시사하는 바가 크다고 볼 수 있다.

국내에서도 2015 개정 교육과정에서 과학과 교과 역량과 기능을 문서화 하면서 과학 탐구 능력 및 탐구 과정을 평가하기 위해 관련 연구들이 수행되었다(Ha *et al.*, 2018; Lee, Lee, & Lee, 2017; Song & Shim, 2019; Yun & Choi, 2019). 교과 역량을 보다 세밀하게 조작적으로 정의하고 측정할 수 있는 문항지를 개발하여 학생들에게 적용하여 타당성을 점검하는 연구가 수행되었고(Ha *et al.*, 2018), 교과서 분석을 통해 과학 교과의 역량 평가가 어떻게 이루어지고 있는지를 확인하는 연구도 진행되었다(Song & Shim, 2019; Yun & Choi, 2019). 실제 평가의 실행 측면에서는 과학 탐구 과정을 평가하기 위해 루브릭을 개발하는 연구가 진행되었는데(Lee, Lee, & Lee, 2017), 이 연구에서 개발된 루브릭은 크게 보고서 평가와 관찰 평가로 영역을 구분하고, 보고서 평가에서는 기초탐구과정(관찰, 분류, 측정, 예상, 추리)과 통합탐구과정(가설설정, 변인통제, 자료해석, 결론도출, 일반화)을, 관찰평가에서는 실험기능(장치설치, 기구조작, 실험수행)을 평가 준거로 설정하고 각 평가준거별로 3단계의 수준을 설정하였다. 이외에도 과정 중심 평가 및 핵심 역량의 평가 측면에서 실제 평가 사례를 살펴본 연구(Lee *et al.*, 2017)가 있었다.

한편, 2015 개정 교육과정이 도입됨에 따라 공통이수과목으로 통합사회, 통합과학과 더불어 과학탐구실험 과목이 신설되었다(MOE, 2015a). 2015 개정 과학과 교육과정 체제 안에서 과학탐구실험은 통합과학을 비롯한 개념 중심의 과학 교과목과 상호 보완적인 역할을 맡을 것으로 기대된다(Jang *et al.*, 2015). 또한 2015 개정 교육과정에서 새로 도입된 5가지의 과학과 교과역량(과학적 사고력, 과학적 탐구 능력, 과학적 문제 해결력, 과학적 의사소통 능력, 과학적 참여와 평생 학습 능력)을 기르는 데 있어 과학탐구실험 과목이 학생들에게 과학적 실천(science practices) 및 탐구 경험을 제공할 수 있을 것이다. 과학탐구실험은 과학적 실천을 수업으로 구현하는 과정에서 실험, 조사, 토의·토론 등 다양한 형태의 교수·학습 방법의 활용이 가능하며, 이를 통해 학습자 참여를 끌어낼 수 있는 토대를 마련해줄 수 있을 것으로 보인다(Kim & Kang, 2018; Jho, 2018). 이러한 과학 탐구 활동은 개념 이해 중심의 성취 이외에도 탐구 기능의 신장 및 과학에 대한 흥미와 태도 함양에 효과가 있음이 보고되고 있어(Kim *et al.*, 2019), 결과적으로 과학탐구실험을 통해 다양한 영역에서의 고른 역량 발달이 가능할 것으로 기대된다.

그러나 과학탐구실험이 처음 도입된 2018년을 포함하여 지난 2년간의 운영 실태를 살펴본 연구들에 의하면, 실행 상의 어려움이 여러 측면에서 나타나고 있었다. 과학탐구실험 과목을 실천 중심의 수업으로 구성하기 위해 요구되는 충분한 수업 시수나 학교 교육과정 편성 시 블럭타임과 같은 유연한 방식의 교육과정 구성이 어렵다는 문제가 제기되었다(Byun *et al.*, 2019). 또한 과학 개념 학습과 과학 문제 해결을 중심으로 운영되는 타 과학 과목과 비교하여 과학적 실천과 관련된 정의적, 기능적 측면이 인지적 측면과 함께 강조되어야 하나, 지필평가 중심의 평가 도구 개발에 익숙한 교사들은 이러한 정의적, 기능적 측면에서의 평가를 실천하는 데 있어 어려움을 겪고 있었다

(Byun *et al.*, 2019). 이에 더해 아직 운영 기간이 짧은 과목이고, 과학 탐구실험을 담당하는 교사들 중 저경력 교사의 비율이 높다는 점에서 경험 부족으로 인한 평가 부담이 가중되고 있었다(Byun *et al.*, 2019; Kwak, 2020).

과학탐구실험 과목의 평가와 관련된 선행 연구들을 살펴보면 교육 과정 및 교과서 분석 연구, 과학탐구실험 운영 실태를 조사한 연구, 과학 탐구 능력 및 탐구 과정 평가를 탐색한 연구 등이 있었다. Kim, Jang, & Hong(2019)은 텍스트 네트워크 분석을 통해서 교육과정에 제시한 평가 관련 내용을 분석하는 연구를 실시하였고, Song & Shim(2019)은 과학탐구실험 교과서 분석을 통해 과학과 교과역량 중 과학적 참여와 평생학습능력이 다른 역량에 비해 다소 적게 반영된 것으로 보고하였다. 또한, 중학교 과학 교과서의 활동을 분석한 연구에서도 과학적 의사소통능력이나 과학적 참여와 평생학습능력에 대한 평가 비중이 다른 역량에 비해 작았고, 하위요소 수준에서 분석할 때도 특정요소에 치우치는 경향이 나타났다(Yun & Choi, 2019). Shin *et al.*(2018)과 Park *et al.*(2019)은 과학탐구실험이 도입된 첫 해 2015 개정 교육과정의 신설과목을 중심으로 교과목 운영 실태와 교과목에 대한 교사와 학생의 인식을 다양한 측면에서 살펴보고, 과학탐구실험이 도입된 2차 년도를 대상으로 전반적이고 광범위하게 운영 실태를 살펴본 연구에는 Byun *et al.*(2019)과 Kwak(2020)의 연구도 있었다. 이 연구들에서 과학탐구실험의 평가 체계가 2018학년도에 9등급 상대평가 체제였다가 2019학년도부터 3등급 절대평가로 전환됨에 따라 학생 변별에 대한 교사들의 부담이 상대적으로 감소하였고, 이러한 변화로 인해 과학탐구실험 과목의 개설 취지를 달성할 수 있는 기반이 조금은 마련되었으며, 교사들 역시 과정 중심 평가를 실천하는 데 적합한 과목으로 인식하고 있음을 보고하였다(Byun *et al.*, 2019; Kwak, 2020). 그러나 이러한 변화에도 불구하고 많은 교사들이 과학탐구실험을 운영함에 있어서 여전히 어려움을 겪고 있는 것으로 나타났는데, 특히, 수업을 준비하는 데 필요한 시간과 교사의 노력이 많이 요구되고 있었다(Byun *et al.*, 2019; Kwak, 2020).

위와 같이 교육과정이나 교과서 분석 및 모니터링의 측면에서 과학 탐구실험을 살펴본 연구들은 수행되었으나, 구체적으로 수업 및 평가 운영 사례들을 자세하게 들여다본 연구들은 많지 않았다. Shin *et al.*(2018)이 과학탐구실험 교과목의 운영 사례를 구체적으로 분석하는 연구를 수행하였으나 한 명의 교사의 실행을 관찰하였다는 점에서 한계가 있다. 연구의 범위를 과학 탐구에 대한 평가로 확장하면, 과학 영재학교 학생을 대상으로 모의 실험 상황에서의 과정 중심 탐구형 수행평가도구를 개발한 Kim & Ha(2020)의 연구를 참고할 만한 하나, 보다 근본적으로 실제 과학탐구실험의 평가 맥락에서 발생하는 문제점을 진단하고 이를 개선할 수 있는 방안에 대한 연구가 수행될 필요가 있다.

이에 본 연구에서는 과학탐구실험의 평가 실태를 사례 연구를 통해서 살펴보고자 하였다. 이를 위해 다양한 과학탐구실험의 평가를 수행한 교사들의 다양한 사례들을 수집하여 분석하였고, 분석을 통해 평가의 현황과 개선 방안을 탐색해 보았다. 구체적인 연구 문제는 다음과 같다. 첫째, 과학탐구실험 과목에서 교사들은 어떠한 탐구 요소를 평가하고 있는가? 둘째, 과학탐구실험 과목의 평가 도구는 어떻게 구성되어 있는가? 셋째, 과학탐구실험 과목의 평가는 어떻게 활용되고 있는가? 이와 같은 연구 문제를 중심으로 과학탐구실험의 평가

도구를 종합적으로 분석하여 과학탐구실험의 평가 방안과 평가를 지원하기 위한 방법들을 고찰하고자 한다.

## II. 연구 방법

본 연구는 과학탐구실험 과목에서의 평가가 어떻게 이루어졌는지 살펴보고, 교육과정 및 과목의 목표에 부합하는 평가 실행 방안을 탐색하는 데 목적이 있다. 이를 위해 25명의 교사로부터 과학탐구실험 평가 도구 및 관련 자료를 수집하고 교사 및 전문가들을 대상으로 전문가 협의회를 실시하였다.

### 1. 자료 수집

본 연구에서는 교사 25명으로부터 과학탐구실험 수업에서 활용한 평가 도구와 이에 대한 교사의 설명 자료를 수집하였다(Table 1). 본 연구에 참여한 교사들은 2018년 또는 2019년에 ‘과학탐구실험’ 수업을 운영한 경험이 있었으며, 전공별로 물리학 4명, 화학 8명, 생명과학 9명, 지구과학 4명이었다. 수집된 사례 중 2018년에 활용된 자료는 8건, 2019년에 활용된 자료는 17건이었다. 평가 도구는 ‘과학탐구실험’ 수업에서 교사가 직접 활용하였던 자료들 중 대표적인 사례 1건에 대한 자료를 수집하였다. 수집한 평가 도구에는 학생 활동지와 평가 기준표 등이 포함된다. 학생 활동지를 통해 학생들이 수업의 맥락과 실제 수행한 활동에 대한 정보를 파악하고자 하였다. 평가 기준표에는 해당 차시에서의 평가의 목적, 방법, 요소, 기준 등이 기술되어 있어 교사가 어떠한 계획과 기준으로 학생들을 평가하고자 하였는지에 대해 알아보하고자 하였다. 평가 도구에 대한 설명 자료는 교사들이 제공하는 평가 도구가 사용된 맥락과 의미를 해석하기 위한 참고자료로써 활용하기 위해 수집하였다. 평가 도구 설명 자료는 평가 도구에 대한 기본 정보(적용 시기, 교과서, 해당 단원), 평가 요소, 평가 주제별 평가 결과 반영 비율, 평가 도구의 활용 방식 등을 묻는 내용으로 구성되어 있었으며, 교사들은 설명 자료에서 제시하는 질문들을 바탕으로 평가 도구에 대한 정보와 의견을 작성하였다.

Table 1. Data collected from research participants who taught ‘Science inquiry experiment’ in 2018, 2019

구분	세부 내용
평가 도구	2018, 2019년에 과학탐구실험 수업에서 활용한 학생 활동지, 평가 기준표 등
설명 자료	평가 도구에 대한 기본 정보(적용 시기, 교과서, 해당 단원), 평가 요소, 평가 주제별 반영 비율, 평가 도구의 활용 방식, 평가 도구의 장단점, 기타 의견

### 2. 자료 분석

#### 가. 평가 도구 분석틀

수집된 평가 도구의 분석을 위한 분석틀은 교육과정 문서 및 선행 연구들을 참고하여 개발하였으며(MOE, 2015b; Yun, Ko, & Choi, 2018; Ha et al., 2018), 평가 요소, 평가 기준의 기술 수준, 평가 방법의

대범주로 구분한 후 각 범주별로 세부 항목으로 나누어 구성하였다. 이 과정에서 개발된 분석틀 초안에 대해 과학교육 전문가 8명에게 검토를 의뢰하였고, 이를 반영하여 Table 2와 같이 분석틀을 확정하였다.

구체적으로, 평가 요소는 개념 이해, 탐구 기능, 정의적 영역의 측면으로 세분화하여 분석하였다. 개념 이해는 평가 요소 중 과학적 개념을 이해하고 있는지를 평가하는 내용이 포함되어 있는지를 파악하기 위한 항목이고, 탐구 기능은 2015 개정 과학과 교육과정에서 제시하고 있는 탐구 기능의 하위 요소들이 평가 요소에 포함되어 있는지를 파악하기 위한 항목이다. 본 연구에서는 교육과정에서 제시하고 있는 8개의 탐구 기능 항목 중 ‘문제 인식’을 ‘문제 인식’과 ‘가설 설정’으로, ‘탐구 설계와 수행’을 ‘탐구 설계’와 ‘탐구 수행’으로 구분하여 총 10개의 세부 요소로 구분하여 분석틀을 구성하였다. 마지막으로 정의적 영역에 해당하는 내용은 태도와 같은 요소와 더불어, 기능적, 인지적 영역이 아닌 타 요소를 평가 기준에 포함하고 있는지 살펴보기 위한 목적에서 분석하였다. 한편, 평가 도구를 분석하는 과정에서 단순히 학생 활동지에 관련 평가 요소 항목이 있는지를 기준으로 판단하지 않았고, 평가 기준표를 참고하여 실제로 해당 요소에 대한 ‘평가’가 이루어지고 있는지를 기준으로 하여 포함 여부를 판단하였다.

다음으로, 평가 기준의 기술 수준은 평가 문항 또는 도구가 어떠한 기준으로 배점이 이루어지고, 어떻게 평가 요소들이 구분되고, 세분화되는지 확인하기 위한 목적에서 분석하였다. 이를 위하여 최종적인 학습 목표의 진술 여부와 평가 기준표에 세부적인 평가 항목과 해당 항목에 대해 구체적인 평가 기준이 제시되어 있는지 살펴보았다. 이 항목의 분석을 통해서 평가 목표와 평가 요소가 수업 전에 체계적으로 설정되어 있고, 이를 구체적으로 평가할 수 있는 기준이 마련되어 있는지를 확인하고자 하였다.

마지막으로, 평가 방법은 평가 주제, 평가 시점, 피드백 명시 여부를 하위 요소로 두었다. 평가 주체는 교사 평가, 동료 평가, 자기 평가 등 다양한 주체들이 평가에 참여하였는지를 확인하기 위하여 포함하였고, 평가 시점은 수업이 진행되는 과정 중에 평가가 진행되었는지, 과정을 마친 후에 평가가 진행되었는지를 파악하기 위한 목적에서 분석하였다. 마지막으로 피드백 명시 여부는 평가의 결과가 학생들에게 유의미한 정보를 제공할 수 있는지를 살펴보기 위해 하위 요소로 포함시켰다. 다만, 평가 도구 분석에 초점을 두는 연구의 특성 상 과정 중에 일어나는 피드백을 포착하는 데에는 한계가 있었다. 이러한 평가 방법에 대한 분석을 통해 교육과정에서 강조하고 있는 과정 중심 평가가 목적에 맞게 진행되고 있는지를 점검하고자 하였다. 더불어 평가 도구의 활용 방식은 작성한 내용을 유목화하여 구분하였다.

#### 나. 자료 분석

자료의 분석은 연구자 4인의 교차 검토를 통해서 이루어졌으며, 각 연구자가 분석한 결과에 대해 논의하며 최종적인 합의를 거쳐 분석 결과를 확정하였다. 분석 과정에서는 학생 활동지, 평가 기준표, 평가 도구에 대한 교사의 설명 자료를 총체적으로 활용하였고, 문서에서 드러나는 정보를 중심으로 분석을 진행하여 연구자의 주관적인 추측에 따른 해석을 최대한 배제하기 위해 노력하였다. 이러한 기준

Table 2. Analysis framework for assessment tool

구분	세부 내용
기본 정보	적용 시기, 사용 교과서 및 단원, 운영 차시 외
평가 요소	개념 이해 : 과학 개념 포함 여부
	탐구 기능 : 2015 개정 과학과 교육과정 '기능' 요소 포함 여부
	정의적 영역 : 태도 등의 정의적 영역 평가 기준 포함 여부 및 평가 내용
평가 기준의 기술 수준	최종 목표 진술 여부
	평가 세부 항목 구성 및 단계별 수준 진술 여부
평가 방법	평가 주제 : 교사 평가, 동료 평가, 자기 평가
	평가 시점 : 과정 중, 과정 후
	피드백 명시 여부

에서 평가 도구 상에서 드러나지 않는 맥락에 대한 추측을 최대한 자제하여 명시적으로 드러나지 않는 부분들은 평가 도구에 포함되지 않은 것으로 분석하였다. 다만, 이는 사례를 분석함에 있어서 객관성을 확보할 수 있다는 장점은 있으나, 수업에서 실제 일어난 구체적인 장면을 포착하기 어려운 부분이 있으며, 연속적이고 누적적으로 수집된 사례가 아니라 연구 참여자가 선택하여 제공한 일부 사례만을 분석한 부분은 본 연구의 한계로 볼 수 있다.

추가적으로, 분석 결과에 대한 타당성을 확보하고 분석 결과가 의미하는 바를 보다 구체적으로 파악하기 위하여 과학교육 전문가 4인, 과학탐구실험을 담당하였던 교사 16인을 대상으로 온라인으로 검토를 받았다. 자문 양식은 자료 분석 결과에 대한 요약된 내용을 제시하고, 해당 결과에 대한 의견을 묻는 형식으로 구성하였다. 세부적으로는 탐구 기능 관련 평가 요소 비율에 대한 해석, 과학탐구실험 수업에서 과학 개념, 정의적 영역, 교과 역량의 평가, 과정 중심 평가 적용의 어려움과 활성화 방안, 평가 결과를 활용한 피드백 활성화 방안, 평가 도구 개발에서의 어려움과 지원 방안, 과학탐구실험 수업에서 평가 측면에서의 지원 방안에 대한 내용을 다루었다. 이를 통해 수집된 자료들을 연구진이 사례 분석 결과를 정리하고 지원 방안을 탐색하는데 참고 자료로 활용하였다. 연구 결과에 대한 설명을 위하여 평가 자료를 제공한 교사들은 T 1~T 25로 코드를 부여하였고, 분석 결과에 대한 자문 요청에 응답한 과학교육 전문가와 교사들은 S 1~S 20으로 코드를 부여하였다.

### III. 연구 결과 및 논의

#### 1. 과학탐구실험에서의 평가 요소

수집된 25건의 사례에서 나타난 평가의 요소를 개념 이해, 탐구

기능, 정의적 영역(태도)의 항목으로 구분하여 각 평가 요소의 포함 여부를 분석한 결과는 Table 3과 같다.

평가 요소에 개념 이해가 포함된 평가 도구는 15사례, 하나 이상의 탐구 기능 요소를 포함하고 있는 평가 도구는 23사례, 태도가 포함되어 있는 사례의 수는 13이었다. 탐구 기능의 하위 요소별로 살펴보면, 자료 분석 및 해석이 18사례로 가장 많았고, 그 다음으로 탐구 수행이 15사례였다. 탐구 설계, 결론 도출, 의사소통은 각각 9사례를 찾아볼 수 있었고, 문제 인식, 토론 및 논증, 모형 개발은 각각 4사례, 4사례, 2사례로 다른 기능 요소들에 비해 해당 요소를 평가하고 있는 사례의 수가 적었다. 이러한 결과는 과학탐구실험 교과서를 분석했던 연구에서 자료에 대한 조사, 변인에 대한 관찰과 측정, 결과에 대한 변환과 설명에 관한 탐구가 많았다는 것과 무관하지 않음을 알 수 있다(Lee, Lee, & An, 2019).

#### 가. 과학 개념의 평가

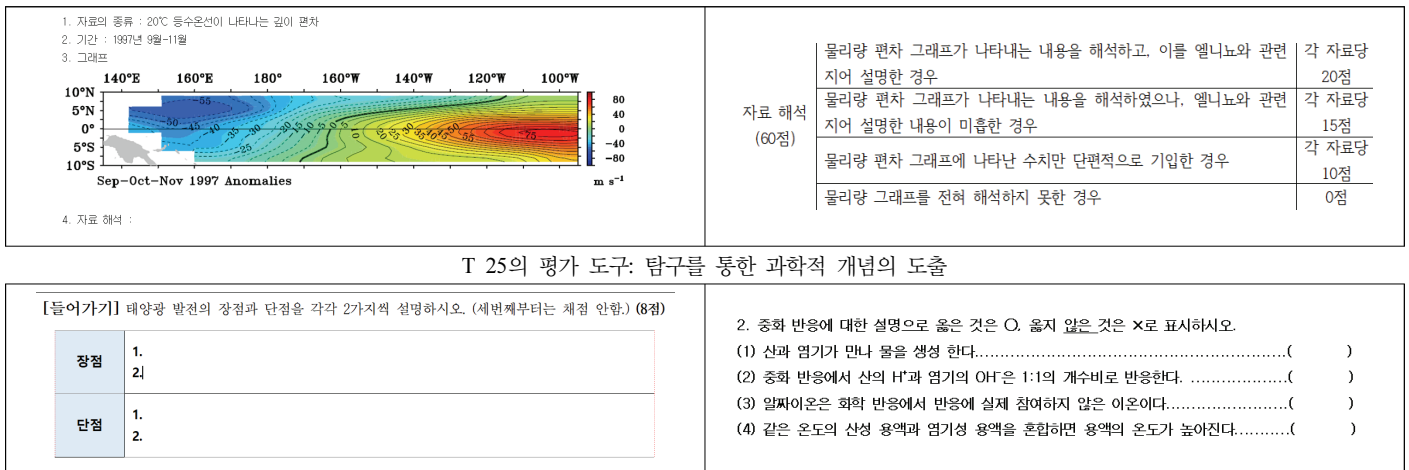
평가 요소 중 개념 이해가 포함되어 있는 각 사례에서 과학 개념을 어떻게 평가하고 있는지 분석하였다. 각 사례들에서 과학 개념을 평가하는 방법은 크게 탐구의 맥락 속에서 이루어지는 경우와 탐구 내용과 관련은 있으나 탐구 과정과는 별개로 단순한 지식을 확인하는 경우로 구분할 수 있었다.

먼저, 과학 개념을 탐구의 맥락 속에서 평가하는 방식은 크게 탐구의 과정 속에서 과학적 개념을 정확하게 활용할 수 있는지 평가하는 경우와 탐구를 수행하기 위해 필요한 선행 지식을 평가하는 경우로 구분할 수 있었다(Figure 1). Figure 1에 제시한 T 25의 사례에서는 학생들이 직접 엘니뇨와 라니냐에 대한 자료를 찾아보도록 하고, 이렇게 조사한 자료를 분석 및 해석하는 활동이 진행되었다. 평가 기준표를 살펴보면, 이러한 활동에서 자료의 해석이 과학적으로 타당하게 되었는지를 평가하고 있었는데, 결과적으로는 자신이 학습한 과학적 개념을 활용하여 조사한 자료를 정확하게 해석할 수 있는지를 평가하는 과정으로 볼 수 있었다. 한편, Figure 1에서 T 24의 사례는 탐구 활동 수행을 위한 과학 개념 평가가 이루어진 예시로, 탐구를 수행하기 이전에 필요한 선행 지식에 대한 점검의 성격을 지니고 있는 것으로 볼 수 있었다. 반면, Figure 1에 제시한 T 22의 사례는 해당 문항이 탐구 주제와 관련된 내용이라는 점에서 다른 유형과 구분되었다.

탐구를 수행하는 과정에서 탐구와 관련되어 있는 과학적 개념의 이해는 필요한 요소이다. 또한, 탐구를 통해서 새로운 과학적 지식을 생성하는 것도 의미있는 것으로 볼 수 있다. 특히, 과학적 실천의 관점에서 탐구를 바라볼 때, 지식과 행동이 서로 얽혀있어 구분되기 어렵다는 점을 고려한다면(Oh, 2020), 아래 S 19의 언급과 같이 탐구의 맥락에서 도출되는 과학적 개념의 평가는 자연스러울 수 있다. 이러

Table 3. The number of cases including each factor of assessment

평가 요소	개념 이해	탐구 기능										태도
		문제 인식	가설 설정	탐구 설계	탐구 수행	자료 분석	수학적 사고	모형 개발	토론 논증	결론 도출	의사 소통	
빈도	15	4	6	9	15	18	6	2	4	9	9	13
비율	60%	16%	24%	36%	60%	72%	24%	8%	16%	36%	36%	52%



T 24의 평가 도구: 탐구 보조형 과학 개념 평가

T 22의 평가 도구: 단순 지식 평가

Figure 1. The types of assessment of scientific concept in collected cases

한 관점에서 ‘과학탐구실험’의 평가 요소에 개념의 이해가 포함될 필요가 있다.

[S 19] 과학탐구실험은 역량기반 과학과 교육과정 목표 달성을 위해 특별히 기획된 교과목입니다. 그래서 교육과정 성취기준도 통합과학과 달리 역량을 기반으로 구성되어 있습니다. 그래서 개념에 대한 이해를 따로 평가할 필요는 없다고 생각하지만, 실제 과학탐구 수행 과정 중 자료해석, 결론도출, 선행 연구 조사 등에서 개념에 대한 이해가 나타나기 때문에 탐구 기능 위주로 평가해도 무방하다고 생각합니다.

다만, 탐구의 주제와 단순히 연관되어 있는 지식을 평가하는 것은 지양할 필요가 있어 보인다. 과학적 개념은 과학적 실천을 통해서 능동적으로 구성될 수 있으며(Fogleman et al., 2011), 결과적으로 탐구 실험 수행 속에서의 과학적 개념의 이해에 대한 평가도 이러한 맥락 속에서 이루어질 필요가 있다. 그리고 이를 위해서는 S 1의 의견처럼 과학탐구실험에서는 탐구의 맥락 속에서 활용되는 과학적 지식이나, 탐구의 결과로서 도출되고 생성되는 지식을 평가하는 것이 보다 적합해 보인다.

[S 1] 과학탐구실험에서 과학의 사실 개념은 가설로부터 이끌어지는 실험 설계와 관련하여, 그리고 결과로부터 이끌어지는 결론 도출 과정에 한하여 평가 장면이 구성되어야 할 것으로 생각한다.

### 나. 탐구 기능 요소의 평가

수집한 전체 평가 도구 중, 2사례를 제외한 나머지 23사례에서는 적어도 1가지 이상의 탐구 기능 요소에 대한 평가가 이루어졌으며, 가장 많은 탐구 기능 요소를 포함한 사례에서는 총 7가지의 요소에 대한 평가가 이루어졌다. 탐구 기능 요소가 가장 많았던 사례는 Figure 2의 T 12의 사례로, 이 평가 도구는 천연 항생 물질의 효과를 알아보는 실험을 먼저 수행한 후, 실생활 속에서 천연 항생 물질을 찾기 위한 탐구 과정을 설계하도록 구성되어 있었다. 즉, 한 번의 탐구 실험을 수행하는 일반적인 수업의 모습과는 달리 먼저 실험 방법을 익히도록 한 후, 이와 관계된 새로운 탐구를 학생들이 직접 설계하도록 한 것이다. 결과적으로 이러한 교수·학습 전략을 활용함으로써

다양한 탐구 기능 요소에 대한 평가를 가능케 한 것으로 볼 수 있다.

한편, 탐구 기능 요소가 발견되지 않은 2사례 중 1사례는 지필평가를 실시하였고, 다른 1사례는 게임 형태의 활동을 수행하고 개념 이해 위주로 평가를 실시하였다. 과학탐구실험이 처음 도입되었을 때 평가를 통해 9등급으로 나누어야 했기 때문에 학교 현장에서는 지필평가 형식의 총괄평가가 불가피하게 시행되는 경우가 많았는데, 현재의 3등급 절대 평가 체제로 전환되면서 수행평가의 실행 비율이 증가하였다(Byun et al., 2019). 그러나 T 16의 언급처럼 평가의 객관성과 공정성의 문제로 인해 불가피하게 지필평가가 병행되고 있는 사례도 있는 것으로 보인다.

[T 16] 과학탐구실험 교과는 학생들 활동중심으로 평가하는 것이 가장 좋은 방법이지만 학생들의 불만을 고려하여 지필서술형 평가를 실시하였다. 보다 객관적인 과정형평가가 이루어진다면 지필평가는 실시하지 않아도 될 것 같다.

지필평가는 아니지만 개념 이해에 대한 평가 위주로 평가 기준을 설정하여, 탐구 기능에 대한 평가 요소를 찾아볼 수 없었던 다른 사례인 T 18의 평가 도구는 게임 활동이라는 과제의 특성 상 탐구와 관련된 평가 기준을 설정하기가 어려웠던 것이 영향을 끼친 것으로 보인다(Figure 2). 이 사례는 학생들이 흥미로워하는 소재를 바탕으로 활동을 구성하였다는 점에서 의미가 있지만, 구체적인 탐구 관련 요소들의 반영이 어렵고 이로 인해 개념 중심으로만 평가가 이루어졌다는 한계점을 가지고 있었다.

과학탐구실험의 목표 중 하나를 탐구 기능의 신장으로 보았을 때, 다양한 탐구 관련 기능들에 대한 평가를 실행하는 것은 교과 목표 달성에 효과적이다. 다양한 요소들을 동시에 평가하는 게 항상 좋은 것이라고 할 수는 없으나, T 12는 학생들로 하여금 다양한 탐구 기능을 활용하여 탐구를 수행하도록 유기적인 형태로 과제를 구성하였고, 자연스럽게 다양한 기능을 평가할 수 있었다는 측면에서 긍정적이라고 볼 수 있다. 반면, 평가의 형식이나 과제의 특성 상 탐구 기능에 대한 평가 요소가 반영되기 어려운 사례들도 찾아볼 수 있었다. 평가의 객관성과 공정성의 측면에서 불가피한 선택일 수도 있으나, 과제의 맥락에 따라 인지적 영역 중심의 평가 문항을 개발하고, 탐구 기능을 포함시키려는 노력을 지속할 필요가 있다.



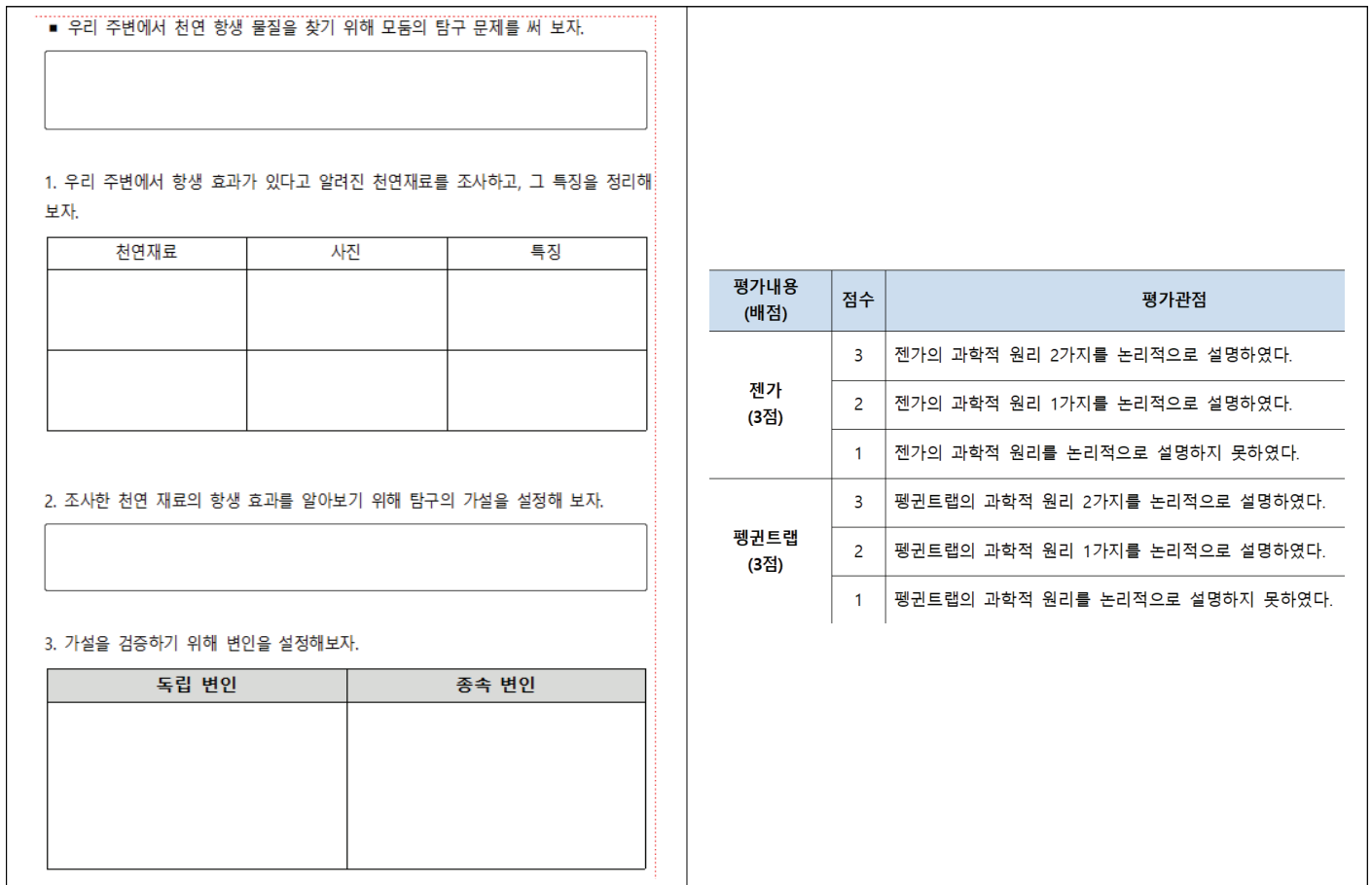


Figure 2. A Case including the most elements of assessment of scientific inquiry skills(T 12) and a case including no elements of assessment of the skills(T 18)

다. 탐구 기능 요소 평가의 특징

Table 4는 수집된 사례에서 탐구 기능 요소가 어떠한 방식으로 평가되고 있는지를 요약한 것이다. 예를 들어, 평가 도구에서 탐구의 방법을 제시한 후 해당 탐구 방법에 적합한 가설을 설정하도록 하고 해당 내용에 대한 평가가 이루어진 경우, 이 평가 도구에 ‘가설 설정’에 대한 평가 요소가 포함되어 있는 것으로 분류하였다. 탐구 기능 요소 중 가장 많은 평가 도구에서 활용된 요소는 ‘자료 수집·분석 및 해석’, ‘탐구 수행’으로, ‘자료의 수집·분석 및 해석’을 포함하는 평가 도구에는 탐구에 필요한 정보들을 조사하고 분석·해석하는 활동이 마련되어 있었고, ‘탐구 수행’은 탐구의 수행을 통해서 얻은 결과를 보고서에 기록하도록 한 후 평가에 활용하였다. 이 요소들은 전형적인 실험 탐구 과정뿐만 아니라 자료 조사형 탐구에서도 일반적으로 수행되는 활동으로, 교사들이 해당 활동의 결과를 평가하기에도 다른 요소들보다 수월하다는 장점이 있으며, 실제로 탐구 실험을 수행할 때 평가에 가장 많이 활용되는 요소이기도 하다(Byun et al., 2019).

반면, 수집된 사례 중에서 평가 요소로서의 활용이 많지 않았던 요소로는 ‘문제 인식’, ‘증거에 기초한 토론과 논증’, ‘모형의 개발과 사용’이 있었다. 먼저, ‘문제 인식’을 평가 요소로 활용한 자료들을 살펴보면, 평가 도구에서는 구체적인 상황, 영상 자료 등을 제시하고

관련 상황에서의 문제를 찾도록 하는 방식으로 구성되어 있었다 (Figure 3). 또한, 한 과제는 ‘자유 탐구’ 형식을 빌려 직접 탐구를 설계하도록 하는 방식으로 구성된 것을 살펴볼 수 있었다. 다음으로 ‘증거에 기초한 토론과 논증’을 평가 요소로 포함하고 있는 평가 도구들은 수업 주제와 관련된 자료들을 직접 조사 및 수집하고, 수집된 자료를 근거로 과학적 개념에 대해 설명하도록 하거나 모둠 활동 속에서 모둠원들과 토의하는 활동이 수업의 과정으로 마련되어 있었다 (Figure 3, T 21). Figure 3의 T 21의 사례에서는 일종의 비동시적 의사소통 도구를 활용하여 동료들과 토의가 자연스럽게 이루어지도록 하였으며(de Vries, Iund, & Baker, 2002), 제한적이거나 교사가 평가를 목적으로 토론을 관찰하고 참여하는 방식이 탐구 과정 내에 이루어졌다고 볼 수 있다. 한편, ‘모형과 개발과 사용’의 평가 요소들이 활용된 2과제는 모두 ‘모형’과 관련된 내용이 과제의 주요 활동 과정으로 이미 구조화되어 있던 평가 도구였다. 이 과제의 주제는 생체 모방 기술을 활용하여 새로운 상품을 개발하는 것과 외계인 모형을 활용하여 원자 모형의 의미를 살펴보는 것으로, 수업의 주제가 ‘모형’과 관련되어 해당 평가 요소를 포함하는 것으로 분류되었다.

이러한 요소들의 활용도가 높지 않은 이유로는 해당 요소를 평가하기 위한 별도의 수업의 방식이나 과제가 마련되어 있어야 하고, 결과적으로, S 2의 언급처럼, 해당 요소를 평가할 수 있는 탐구의 주제가 다소 제한적이기 때문으로 볼 수 있다.

6) 요소별 빈도는 Table 3에 제시함.

Table 4. The way of utilization and the frequency of each element of scientific inquiry skills in collected cases

탐구 기능 요소	평가 도구에서의 활용 방식
문제 인식	· 탐구 주제를 직접 선정하여 탐구를 수행하도록 함 · 실생활과 관련된 문제 상황을 제시하여 문제를 발견하도록 함
가설 설정	· 탐구의 방법을 제시하고 해당 탐구 방법에 적합한 가설을 설정하도록 함
탐구 설계	· 수행한 탐구 활동과 관련된 새로운 탐구 활동을 설계하도록 함
탐구 수행	· 탐구 수행을 통해 도출된 결과를 보고서 기록하도록 함
자료의 수집·분석 및 해석	· 탐구에 필요한 정보들을 조사하도록 함 · 수집된 탐구 결과를 분석 및 해석하도록 함
수학적 사고와 컴퓨터 활용	· 컴퓨터를 활용하여 필요한 정보를 수집하는 방법을 안내하고, 이를 수행하도록 구성됨
모형의 개발과 사용	· 생체 모방과 같이 실생활 속에서 발견할 수 있는 과학적 원리를 적용하여 모형을 개발하도록 구성됨 · 전체적인 탐구의 활동 과정에서 개발된 모형을 적극적으로 활용하고 새로운 모형을 구성하도록 함
증거에 기초한 토론과 논증	· 데이터를 수집하고 수집된 데이터를 근거로 설명·토의하는 활동으로 구성됨
결론 도출 및 평가	· 도출 및 분석된 결과의 과학적 의미를 파악하도록 함
의사소통	· 모둠별 탐구 수행 결과를 발표하도록 함

[S 2] 모형 개발과 사용의 경우에도 모형 개발과 활용에 관한 탐구 활동이 포함된 매우 특이적인 탐구활동에 국한될 수 있을 것입니다.

또한 이 요소들이 다른 요소에 비해 객관적이고 정량적인 평가의 기준을 설정하기 어렵고, 교사들이 해당 요소를 수행하거나 평가해본 경험이 상대적으로 부족한 것도 이 요소들의 활용 빈도가 높지 않은 이유로 볼 수 있었다.

[S 3] 자료의 수집은 정량적으로, 분석 및 해석은 일반적으로 정답이 존재하기 때문에 평가의 객관성을 확보할 수 있습니다. 하지만 문제 인식, 모형의 개발과 사용은 구체적인 채점 기준을 확보하기 어렵습니다. 교육 현장에서는 평가의 경우, 채점 기준표를 구체적으로 작성하여 객관성을 확보하는 것이 중요하기 때문에 위와 같은 경향이 나타나는 것으로 생각합니다.

[S 1] 교사 연수 교육에서도 교육 현장에서 탐구 기능과 관련하여 교사 스스로가 수행에 참여한 경험이 없기 때문에, 그리고 일부 과학교사들을 제외하고는 과제 연구 경험이 없기 때문에 문제 인식과 문제 인식으로부터 가설 설정에 대한 평가 장면을 구성하기 어렵다.

한편, 해당 요소들을 활용한 평가 도구들의 구체적인 채점 기준을 살펴보면, ‘증거에 기초한 토론과 논증’을 포함한 평가 도구의 경우 수집된 증거와 주장 사이의 관계나 토론의 과정에 대한 평가보다는 보고서에 작성된 결과 및 결론의 타당성에 초점(3사례)을 두거나, 토의 활동에의 참여 여부만을 평가(2사례)하고 있었다. 이러한 평가 기준의 설정은 T 21의 의견처럼 교사 한 명이 학생들의 토의 활동을 관찰하면서 구체적으로 평가하는 데 어려움이 있기 때문으로 볼 수 있다.

<p>1. 탐구단계 수행하기</p> <p>[의문 구체화] 다음에 제시된 상황을 읽고 떠오르는 의문을 적어보자.(1점)</p> <p>철수는 부모님과 함께 오랜만에 중화요리를 먹으러 갔다. 여러 요리를 먹은 후 마지막으로 온 가족이 좋아하는 짜장면을 먹으면서 철수는 신기한 것을 발견했다. 아버지와 형의 짜장면 그릇에는 짜장 소스가 골여져서 국물처럼 보였지만 어머니와 철수의 짜장면 그릇은 전혀 그렇지 않은 것이었다. 어머니는 형이 아버지를 달아서 짜장면을 끊어먹으니깐 질이 섞여서 그런 것이라고 했다. 비빔냉면을 먹을 때는 관찰하지 못한 것을 짜장면 먹을 때는 볼 수 있는 이유가 무엇일까? 철수는 궁금해졌다.</p> <p>우리의 의문은</p> <p>입니다.</p>	<p>4</p> <ul style="list-style-type: none"> <li>· [동료 평가] 프로토타입을 1분정도 발표하고 나머지 모둠에서는 포스트잇에 + (잘된 점), -(개선할 점), ?(의문점), ! (새로운 제안)를 작성하여 피드백에 참여한다.</li> <li>★ tip : 프로토타입 발표는 1분 내외로 제한하며 어떤 문제를 해결하고자 개발되었는가, 제품(아이디어)의 중요한 점 등을 회사의 홍보 팀장의 발표 또는 광고처럼 재미있게 발표하도록 한다.</li> <li>★ tip : 설의응답 시간을 2-5분 정도 갖으면서 포스트잇에 피드백(동료 평가)을 작성한다. 피드백 내용은 잘된 점(+), 개선할 점(-), 의문점(?), 새로운 제안(!)이 있으면 작성하고 피드백 내용에 자신의 이름을 작성하게 하는 것이 장난스럽거나 단순한 피드백 내용을 줄일 수 있다. 피드백 내용 중 좋은 것은 과제특에 적어준다고 하고 좋은 피드백을 유도하고 좋은 의견은 과제특에 반영한다.</li> <li>· [자기 평가] 참여도와 교과 이해 정도 등을 점검하고 자신이 한 역할, 칭찬할 모둠원과 그 내용, 느낀 점을 작성한다.</li> <li>★ tip : 생기부 과제특 작성 시 참고하여 작성한다.</li> </ul> <p>활동지 8-9면 항목 포스트잇, 자석(프로토타입 고정용 사용)</p>																										
<p>T 2의 평가 도구: 문제 인식</p>	<p>T 21의 평가 도구: 증거에 기초한 토론과 논증</p>																										
<p>◆ 주제 : 생체 모방, 적정기술 또는 신소재를 활용하여 새로운 상품 개발하기</p> <ul style="list-style-type: none"> <li>- 생체 모방, 적정기술 또는 신소재 중 하나를 선택</li> <li>- 선택한 재료(기술)의 특성을 살려서 자신만의 새로운 상품을 개발하여 작성 (선택한 재료(기술)의 특성을 2가지 이상 활용해야 하며, 특성을 구체적으로 설명해야 함 같은 특성을 설명할 경우 1가지로 간주)</li> </ul>	<p>2. 위의 의제인 모형에서 각각의 요소들이 무엇을 의미하는지 적으시오.</p> <table border="1"> <thead> <tr> <th>요소</th> <th>의미하는 것</th> <th>요소</th> <th>의미하는 것</th> </tr> </thead> <tbody> <tr> <td>머리카락의 수(1~8개)</td> <td></td> <td>손가락 수(1~2개/1~8개)</td> <td></td> </tr> <tr> <td>표정(웃는 / 찡그린)</td> <td></td> <td>날씬한 정도(날씬한 / 뚱뚱한)</td> <td></td> </tr> <tr> <td>발의 수(1~3개)</td> <td></td> <td>몸의 무늬</td> <td></td> </tr> </tbody> </table> <table border="1"> <thead> <tr> <th>평가 요소</th> <th>평가 기준</th> <th>점수</th> </tr> </thead> <tbody> <tr> <td rowspan="3">외제인 모형의 정확성</td> <td>외제인 모형의 요소가 원자 모형에서 무엇을 의미하는지 5개 이상 답변한 경우</td> <td>5점</td> </tr> <tr> <td>외제인 모형의 요소가 원자 모형에서 무엇을 의미하는지 3개 이상 답변한 경우</td> <td>4점</td> </tr> <tr> <td>외제인 모형의 요소가 원자 모형에서 무엇을 의미하는지 2개 이하 답변한 경우</td> <td>3점</td> </tr> </tbody> </table>	요소	의미하는 것	요소	의미하는 것	머리카락의 수(1~8개)		손가락 수(1~2개/1~8개)		표정(웃는 / 찡그린)		날씬한 정도(날씬한 / 뚱뚱한)		발의 수(1~3개)		몸의 무늬		평가 요소	평가 기준	점수	외제인 모형의 정확성	외제인 모형의 요소가 원자 모형에서 무엇을 의미하는지 5개 이상 답변한 경우	5점	외제인 모형의 요소가 원자 모형에서 무엇을 의미하는지 3개 이상 답변한 경우	4점	외제인 모형의 요소가 원자 모형에서 무엇을 의미하는지 2개 이하 답변한 경우	3점
요소	의미하는 것	요소	의미하는 것																								
머리카락의 수(1~8개)		손가락 수(1~2개/1~8개)																									
표정(웃는 / 찡그린)		날씬한 정도(날씬한 / 뚱뚱한)																									
발의 수(1~3개)		몸의 무늬																									
평가 요소	평가 기준	점수																									
외제인 모형의 정확성	외제인 모형의 요소가 원자 모형에서 무엇을 의미하는지 5개 이상 답변한 경우	5점																									
	외제인 모형의 요소가 원자 모형에서 무엇을 의미하는지 3개 이상 답변한 경우	4점																									
	외제인 모형의 요소가 원자 모형에서 무엇을 의미하는지 2개 이하 답변한 경우	3점																									
<p>T 10의 평가 도구: 모형의 개발과 사용</p>	<p>T 17의 평가 도구: 모형의 개발과 사용</p>																										

Figure 3. Examples of assessment materials including and assessment element which were less frequently utilized

**[T 21] 각 모듈 안에서 아이디어를 제안하고 발전하는 과정을 모두 관찰하기 어려움**

실제로 학교 현장에서 학생들의 활동을 하나씩 관찰하면서 평가하는 것은 매우 어려운 일이다(Lee et al., 2017). 다만, 과학적 논증이 증거와 주장 사이의 논리적인 관계나 지식의 구성 과정에 초점을 두고 있다는 점에서(Sampson et al., 2011), 논증 과정과 관련된 평가 요소 및 평가 기준을 설정하여(Yang et al., 2009) 이에 대한 평가를 시도하는 것이 학생들의 활동을 직접 관찰하는 것이 제한되는 경우에 활용될 수 있을 것으로 보인다. 그러나 수집된 자료들에서 이러한 평가 기준에 초점을 두고 있는 평가 도구를 찾기에는 한계가 있었다. 비슷한 맥락에서, ‘모형의 개발과 사용’도 모형을 개발하고 활용하는 과정, 모형에 대한 이해와 같은 요소보다는 모형을 활용하여 관련 개념을 잘 이해하고 있는지를 평가하도록 구성되어 있다는 점에서 해당 기능 자체에 대한 평가는 다소 부족한 것을 볼 수 있었다(Figure 3).

**[T 17] 직접 학생들이 카드를 배열하면서 외계인 모형의 각 요소의 특징의 경향성을 파악하고 나아가 외계인 모형의 각 요소가 원자 모형에서 어떤 요소를 나타내는지 알아내게 함. ... (중략)... 결국 학생이 외계인 모형을 바탕으로 주기율표의 특징(가로줄과 세로줄의 특징)을 잘 이해했느냐에 대한 내용으로 평가를 진행하였음.**

종합하면, 수집된 사례에서 탐구 기능 요소들의 활용 빈도에 차이가 있었고, 특히 활용 빈도가 낮은 평가 요소의 경우 해당 요소에 대한 평가에서 각 기능 요소의 특징들이 평가의 장면 속에서 포착되기 보다는, 객관성과 수월성에 초점을 둔 평가가 이루어지고 있는 것을 관찰할 수 있었다. 이러한 현상은 시간 부족 등과 같이 외적인 요인에 의한 영향도 있겠지만(Lee et al., 2017), 관련 소재들을 활용한 평가 맥락의 구성과 구체적인 평가 기준의 설정과 실행이 어렵다는

점 또한 영향을 미친 것으로 보인다. 따라서 이를 개선하기 위해서 관련 평가 사례의 개발과 제공, 평가 역량의 신장 등이 요청된다.

**라. 정의적 영역의 평가**

정의적 영역에 해당하는 평가 요소는 총 13사례에 포함되어 있었다. 이 범주에 포함된 평가 내용들을 구체적으로 살펴보면, Figure 4의 사례에서 볼 수 있듯이 ‘수업 참여도’와 관련된 내용이 10사례로 가장 많았고, ‘실험 후 정리’, ‘제출 기한’, ‘과제를 통해서 느낀 점’과 같은 내용들도 있었다. 일부 사례에서는 수행 평가 과정에 대한 소감, 모듈에서의 자기 역할, 수업을 통해 성장한 점들을 기록하도록 한 평가 도구도 있었으나, 점수에는 반영되지 않았다.

교육 과정에 따르면 인지적, 기능적, 정의적 영역의 고른 발달을 위해서 각 영역의 평가가 균형적으로 이루어져야 하는데(MOE, 2015a), 수집된 사례에서 정의적 영역에 대한 평가는 수업 참여와 같이 매우 제한적인 범위 내에서만 이루어지고 있었으며, 구체적인 평가 기준도 발견되지 않았다. 이는 ‘태도’에 대한 평가가 객관적으로 이루어지기 어렵고, 실제로 측정이 쉽지 않기 때문으로 볼 수 있다. 그나마 ‘수업 참여도’가 반영 가능한 요인으로 교사들이 인식하고 있었으며, 그럼에도 일부 교사들은 직접적으로 점수를 반영하기 보다는 감점 형식으로만 활용하는 사례들도 있었고, 자기 평가, 동료 평가 등과 같이 평가의 주체를 학생들에게 일임하는 사례도 있었다.

**[S 16] 평가가 객관적으로 이루어져야 하므로 평소 학생들이 가지고 있는 과학에 대한 생각 등은 채점이 어렵기 때문에 대부분 수업 참여도로 정의적 평가를 수행할 것이라 판단됩니다.**

**[S 10] 정의적 영역은 수치화하기 힘든 요소라 생각한다. 수업참여도는 수치화할 수 있는 좋은 요소이기 때문에 대부분 교사들이 평가 요소로 선정한다고 생각한다.**

그리기 평가 기준!!		점수
• 그림을 성의 있게 그리고, 특징과 이용사례가 적절하게 제시되었다.		감점 없음
• 그림에 생명체의 특징이 잘 드러나지 않거나, 이용사례가 적절하지 않다.		-5점
• 그림체가 성의 없거나(ex. 외곽선 하나만으로 그림을 표현) 특징과 이용 사례 모두가 적절하지 않다.		-10점
• 그림을 아예 그리지 않았다.		-20점

T 15의 평가 도구: 감점 형식으로 태도를 평가

[자기 평가]		
구분	평가 항목	도달 정도(낮음→높음)
지식	연역적탐구방법을 설명할 수 있다.	1 2 3 4 5
탐구	탐구를 통해 가설을 설정하고 실험을 설계하여 결과를 도출해낼 수 있다.	1 2 3 4 5
태도	모둠원과 협동하여 탐구 활동에 적극적으로 참여하였다.	1 2 3 4 5
	탐구활동 뒷정리를 잘 하였다.	1 2 3 4 5
	탐구 활동 후 자기 평가를 객관적으로 실시하였다.	1 2 3 4 5

T 2의 평가 도구: 자기 평가의 활용

Figure 4. Examples of materials including assessment elements related with affective domain



결과적으로, 다양한 정의적 요소의 평가는 객관성과 정량화 등의 요인을 극복해야만 가능할 것으로 보인다. 즉, 점수 부여나 등급 구분을 목적으로 한다면 정의적 영역에 대한 평가는 어려울 것이고, 이보다는 실제 정의적 영역의 변화 과정을 관찰하고 기록하는 데 초점을 두는 것이 적절해 보인다. 학습 동기의 발현으로서 수업에 적극적으로 참여하는 태도도 중요한 요인이나, 과학 교과라는 점을 고려하여 학생들의 과학에 대한 자기 효능감, 과학의 가치에 대한 이해, 과학에 대한 태도 등도 중요하게 다뤄질 필요가 있다(Osborne, Simon, & Collins, 2003). 다른 과학 교과목에 비해 평가에 대한 자유도가 높은 과학탐구실험은 이러한 시도가 가능한 과목으로 보이며, 정의적 영역의 특성을 고려하였을 때 학생들의 변화 과정을 지속적으로 점검할 수 있는 실제적인 방안들의 마련이 요구된다.

## 2. 평가 도구의 구성

### 가. 평가 기준의 구성 및 정합성

평가 기준은 학습 목표에 따라 진술된 평가 목표와 평가 내용, 평가 기준의 세부 항목을 종합적으로 분석하여 살펴보았다. 학습 목표는 교사가 수업을 통해서 의도하거나 계획한 성취 수준에 대한 진술이며 (Kim et al., 2017), 평가 도구는 학습 목표에의 도달 여부를 진단할 수 있도록 구성되어야 한다. 또한, 수업의 학습 목표를 달성하는 과정에서 필요한 구체적인 지식, 기능, 태도 등을 확인할 수 있는 세부 평가 요소를 구분하여 진술하는 것이 요구되며, 각 요소는 종합적인 평가 목표와 맥을 같이하도록 설정되어야 함은 자명하다. 즉, 평가와 관련된 각 사항들은 서로 수업의 목표와 일련의 정합성을 갖는 형태로 조직되어야 한다. 이러한 관점에서 교사들이 제공한 평가 도구가 최종 목표를 진술하고 있는지, 세부 평가 요소를 구분하고 정의하는지, 세부 요소별로 평가 또는 채점의 단계가 구분되며, 구분한 단계를 확인할 수 있도록 특성을 진술하고 있는지의 여부를 살펴보았다. 평가 기준의 구성의 측면에서 분석한 결과는 Table 5에 제시하였다.

평가 목표 진술의 측면에서 대다수의 참여자들은 평가 목표와 그에 따른 세부 목표들을 진술하고 있었으나, 일부의 경우에는 평가 도구나 평가 기준을 수립함에 있어 목표 진술을 생략하기도 하였다. 구체

적으로는 수집한 평가 도구 중, 평가의 목표 또는 요소별로 세부 평가 목표를 진술한 사례는 총 22사례이고, 세부 평가 요소를 정의한 경우는 19사례로 나타났다. 세부 평가 요소를 정의한 19건의 사례에서도 5건의 경우에는 평가의 최종 목표는 진술하지 않아, 평가 도구 구성의 측면에서 목표에 대한 진술의 필요성이 간과되고 있다고 여겨진다. 이는, 학생들이 활동 중 기록하는 내용들을 이용하여 평가가 이루어진다는 측면에서 활동지 또는 보고서와 평가 도구를 동일한 것으로 간주하여 수업 내용의 결정과 조직에 더 많은 시간 할애를 하고 있기 때문에 나타나는 것으로 추측할 수 있다. 이는 자신의 평가 도구가 갖는 장점, 한계 또는 개선점에 대한 설명에서 자신이 계획한 ‘탐구 활동의 장점’에 대한 설명과 ‘탐구 활동의 한계’에 대한 반성이 주를 이루고 있다는 측면에서 확인할 수 있었다.

평가 도구를 구성할 때, 세부 평가 요소들을 수준별로 구분하여 단계를 진술한 비율도 높지 않았다. 총 6사례에서 세부 평가 요소를 정의함과 동시에 각 요소의 성취 수준별 단계를 진술하였다. 그리고 7사례는 세부 요소에 따른 평가 점수 산출을 단계별로 구분하는 방식으로 진술하였으나, 구체적인 성취 수준이나 행동 특성을 기술하는 것은 미비하였다. T 24의 경우에는 ‘태양전지 작동 원리’와 관련된 활동에서 가설 설정, 변인 설정 및 탐구 설계, 보고서 작성 등의 세부 요소로 평가 항목을 구분하면서 Figure 5와 같은 형태로 평가 기준을 제시하였다. 이와 같은 경우는 앞서 기술한 바와 같이 총 6건의 평가 도구에서 확인할 수 있었으며, 다수의 경우에서 보고서의 항목 또는 활동지의 문항별로 배점을 결정하고, 수준을 구분하여 차등 배점하는 방안을 기록하고 있었다. 이와 관련하여 연구에 참여한 교사들 중 적지 않은 수가 평가 기준과 관련하여 자신의 도구를 “평가 기준이 모호”하다고 평가하였으며, 일부는 “모든 평가 요소를 평가에 반영하였을 때 한계”가 발생함을 토로하였다. 아래 T12의 진술과 같이, 일부의 경우에는 이를 극복하기 위해 평가 목표에 따라 세부적으로 요소와 단계를 구분하고, 적절한 평가를 계획하는 과정이 필요함은 인지하고 있었다. 그럼에도 불구하고, 과학적 지식의 평가가 주를 이루지 않는 과목의 특성 상 기능, 태도 등의 기준 수립 과정에는 여전히 “에매함”이 존재한다는 갈등에 놓여 있었다. ‘실험 평가와 관련된 표준화 된 방법’이 필요하다는 언급에서도 이러한 갈등을 찾아볼 수 있다.

Table 5. The level of organization of assessment standards and its frequency

항목	빈도 (비율)
최종 평가 목표 진술	22 (88%)
세부 평가 요소 구분 및 정의	19 (76%)
세부 평가 요소별 평가 단계 구분	13 (52%)
세부 평가 요소별 평가 단계 구분 및 성취 수준별 단계 진술	6 (24%)

[T 12] 평가 계획 수립 시, 성취 기준과의 연계성을 먼저 고려한 뒤에 도달하기 위한 과정에서 필요한 능력을 평가 요소로 구체화하고 이를 가장 효과적으로 평가할 수 있는 수행평가 방법을 선정하고 채점 기준을 수립하는 등의 작업을 거치지 못했다.

또한, 평가 기준의 구성 측면에서 평가 문항의 내용과 평가 목표 또는 평가 기준이 일치하지 않는 경우가 3건 있었다. 이는 평가 문항이 파악하고자 하는 학생의 개념에 대한 이해 또는 역량과 사전에

변인 설정 및 탐구 설계하기	변인 등을 올바르게 제시하고 실험 절차를 타당하게 구상함.	8
	변인만을 옳게 제시함.	6
	변인의 제시에 오류가 있음.	4
	탐구 과정을 제시하지 못함.	2

Figure 5. Assessment standard of ‘Principle of the solar cells’ planned by T 24

수립한 평가의 기준이 서로 부합하지 않은 경우와 평가 기준이 너무 포괄적으로 제시된 사례들이 해당하였다. 이로 인해 학생들이 해당 탐구 활동의 경험으로부터 학습의 목표에 도달하였는지 여부를 명확하게 파악하기 어려울 것으로 예상되었다. 예를 들면, ‘토양 및 호수의 산성화 연구 결과 보고서’라는 탐구 활동을 계획한 한 교사는 탐구 내용으로 토양 또는 호수가 산성화 되는 원인, 피해 그에 대한 대책 등을 조사하고 정리하도록 수업을 구성하였으나, 평가 기준으로 산과 염기의 정의, 중화 반응에 대한 이해도와 같이 교과 지식을 중심으로 제시하여 해당 내용에 대한 구체적 평가를 확인하기 어려웠다. 다른 경우에는 평가 기준에 교육과정에서 제시된 성취기준을 그대로 제시하는 사례들도 존재하였다.

평가 기준 수립에서 드러난 두 특징은 교사들이 평가 기준을 설정함에 있어 다양한 한계를 느끼고 있음을 드러낸다. 무엇보다 대다수의 경우에는 학기 초에 평가에 대한 계획과 기준을 미리 결정해두어야 하므로 실제 수업 내용을 구성할 때 평가의 기준의 보완과 수정에 어려움이 있다. 또한 ‘정해진 공간’과 ‘제한된 시간’ 내에 평가가 이루어져야 하는 상황의 특성 상, 과거와 비슷한 형태로 평가의 틀을 구성할 수밖에 없음을 토로한다. 교사들이 평가와 수업을 동일하게 여기고 있다는 앞의 진술처럼, 과학탐구실험 과목의 특성상 평가 도구의 구성과 내용은 활동 내용과 밀접하게 연관되어 있다. 그렇기 때문에 수업의 계획과 평가 계획의 수립이 분절되어 있는 상황에서는 학습 목표, 수업 내용, 평가 내용 및 기준 간의 불일치가 일어날 수 있으며, 학생들에게 유의미한 피드백을 제공하거나 평가 결과를 환류하는 것에 어려움을 초래할 수 있다. 따라서 평가 도구를 구성하는 과정에서 교육과정과 수업, 평가 간의 정합성과 일관성을 강조하는 것이 필요하다. 이를 위해서는 평가 기준을 수립하는 과정에서 특정 주제에 따라 과학 개념, 역량, 기능, 참여 등의 다양한 과학 탐구 관련 요소 (Song et al., 2019)들과 같이 세부 항목이 필요한 경우를 구분할 수 있도록 안내하는 것이 필요하다. 다만, 다양한 수업 상황에 유연하게 평가 도구를 변화시키며 적용할 수 있도록, 가능한 구체적이고 다양한 사례들을 개발하고 제시하여 평가 기준에서 채점의 단계에 이르는 일련의 과정을 안내하는 지원책이 필요하다. 또한 아래 발췌한 T3의 의견과 같이, 다양한 탐구 주제와 활동 내용들에 대한 교사들 간의 공유를 활성화하는 노력도 필요하다.

**[T 3] 교과서의 제시된 탐구 실험뿐만 아니라 다양하게 개발되어 있는 학교 현장의 탐구, 실험에 대한 자료 공유가 활발해졌으면 함.**

#### 나. 평가 방법: 평가의 시점

과학탐구실험은 탐구를 강조하는 과목이니만큼, 다양한 종류의 활동을 중심으로 교과서들이 구성되어 있고(Jho, 2018), 개념 중심의 다른 과학 과목에 비해 과정 중심 평가의 활용 가능성이 높다고 볼 수 있다. 특히, 학생들의 다양한 탐구 수행 과정을 살펴볼 기회가 제공된다는 측면에서 수행에 대한 기능적 측면이나 역량에 대한 평가가 가능할 것으로 기대된다. 이렇듯 지속적인 학습을 안내하기 위한 목적에서 과정 중심 평가를 운영하기 위해서는 학생들의 수행을 살펴보는 평가의 시점을 수업 중 다양하게 배치하고, 다음의 학습으로 안내하는 것이 요구된다(Lee et al., 2017).

평가 시점에 따라 수집한 자료들을 분석한 결과, 모든 교사들은 수업 후에 평가를 실시하였다고 응답하였으며, 수업 중간에 이루어진 평가를 포함시킨 경우는 7사례(28%)로 상대적으로 소수에 해당하였다. 이는 구체적으로 학생들이 수업 시간 혹은 과제로 보고서 또는 활동지와 같은 학습의 결과물을 작성하고, 이후에 해당 자료들을 토대로 학습의 결과를 평가하였음을 의미한다. 실제 많은 평가 도구들이 ‘변인을 구분하여 적고’, ‘변인을 고려하여 실험방법을 설계’하는 방식으로 탐구 과정을 학생들이 기록하도록 구성되어 있으나, 이러한 활동들은 주로 수업 시간이 마친 이후에 평가가 이루어졌다. 실제 앞선 설명과 같이 탐구 활동을 계획한 교사의 설명에 따르면, 해당 평가 도구는 ‘성적 산출을 주 목적으로’ 활용되었다.

탐구 과정 중의 평가가 이루어진 사례는 몇 가지로 구분할 수 있다. 그 중 첫 번째는 동료 평가와의 연관시킨 사례들이다. T 5는 ‘비행 원리와 드론 조작 활동’을 주제로 한 탐구에서 토론 참여도, 협력, 드론 조종 능력과 같은 사항들을 같은 모둠의 구성원들 간 동료 평가를 실시함으로써 평가를 진행하였다. 유사하게 T 12의 경우에도 천연 항생 물질과 관련된 탐구 중 천연 재료들의 항생 효과를 알아보기 위한 탐구를 계획하여 결론을 끌어내는 과정을 모둠 발표로 정리하도록 계획하고, 발표에 대해 동료 간의 평가를 실시하도록 계획하였다. 두 번째 사례는 교사가 탐구 중의 학생들의 활동 내용을 관찰하여 기록하는 방식들이었다. 다만, 이러한 사례들은 탐구 참여에 대한 ‘적극성’, ‘성실성’과 같이 학생들의 참여를 고취하기 위한 목적에서 이루어졌다. 즉, 학생 참여와 관련하여 일부 정의적 영역의 평가를 대변한다고 볼 수 있으나, 이러한 방식은 탐구의 실행과 직접적으로 연관되지는 못하였다는 측면에서 한계가 있다.

과정 중 평가의 다른 사례는 장기간의 탐구를 수행한 사례에서 확인할 수 있었다. T 14의 사례는 장기간에 걸쳐 여러 차시에 걸친 자유 탐구를 진행하는 과정 중에 진행 상황을 점검하고 지원하기 위한 목적에서 평가를 도입한 경우이다. T 14는 자유 탐구를 과학탐구 실험 수업에서 도입한 이유를 “과학탐구실험 교과서가 워크북 형태로 구성됨에 따라 구조화된 탐구 설계를 경험”하는 것에만 머물러 있어 “자신의 과학적 호기심이나 문제인식을 바탕으로 하여 탐구를 계획하는 경험을 제공할 필요”에 초점을 두고 있기 때문으로 설명하면서, 자연스럽게 탐구가 완료되지 않은 시점에도 평가가 이루어지게 수업을 운영하게 됨을 강조하였다. 다만 이러한 사례들은 전체 25건의 사례에서 일부에 지나지 않았다.

이렇듯, 결과로부터 살펴볼 수 있는 사례들은 향후 학습의 방향을 안내하고 학습을 증진하기 위한 목적의 평가 달성의 어려움을 보여준다. 탐구를 중심으로 한 과목이니만큼, 협력과 토론 및 토의, 도구 활용의 개선과 탐구 수행 과정 점검 등의 목적에서 이루어지는 평가를 기대하나 실제 두드러지지 않았다. 제한적으로 여러 차시를 관통하여 이루어지는 수업에서는 이러한 목적의 평가를 확인할 수 있었으나, 한 차시의 수업에서 이루어지는 탐구 활동 내에서는 찾아보기 어려웠다. 이에 대하여 교사들은 관련한 다양한 어려움이 존재한다면 서 다음과 같이 말한다.

**[T 21] 각 모둠 안에서 아이디어를 제안하고 발전하는 과정을 모두 관찰하기 어려움. (중략) 1주일에 1회 수업인 과학탐구실험에서 50분의 수업 안에 관련 이론을 학습하고 실험 수행하고 결과 피드백까지 받을 수 있는 평가 사례가 필요함.**

**[T 16]** 과학탐구실험 교과는 학생들 활동중심으로 평가하는 것이 가장 좋은 방법이지만 학생들의 불만을 고려하여 지필서술형 평가를 실시하였다. 보다 객관적인 과정형평가가 이루어진다면 지필평가는 실시하지 않아도 될 것 같다.

앞선 T 21과 T 16의 언급과 같이 교사들은 제한된 시간 내에 활동을 관찰해야 한다는 부담과 현실적 제약, 고등학교 평가의 맥락에서 요구된다고 생각하는 객관성 측면을 언급하였다. 이외에도 과정 중심 평가 운영에 따른 수업 준비와 업무의 부담, 기존 평가와 다른 형태의 평가에 대한 낯설, 활동을 중심의 평가에 대한 교사 개인의 역량 부족 등의 이유를 꼽았다. 이와 관련하여 교사들은 ‘학생의 실험 과정을 평가하고, 그 자리에서 바로 평가하여 피드백을 줄 수 있는 평가도구’를 지향한다는 언급과 같이 과정 중심 평가의 방향에 대해 공감하고 실천에 옮기고 싶어 하나, 현실적인 제약과의 괴리를 강조하고 있다. 이러한 한계는 평가의 주체가 한정적이라는 측면과도 밀접하게 연관 된다.

**다. 평가 방법: 평가의 주체**

평가의 주체를 확인하기 위한 목적에서 점수 산출을 하는 대상자로 구분하였을 때, 평가의 주체는 교사, 타 학생(동료), 학생 본인으로 구분할 수 있었다. 이에 따라 평가의 주체에 따라 교사 평가, 동료 평가, 자기 평가로 구분하여 분석한 결과는 Figure 6에 제시한 바와 같다. 수집한 평가 도구 중 교사 평가로만 평가가 진행된 사례가 총 18사례이었으며, 7사례에서는 동료 평가가 이루어졌으며, 5사례에서는 자기 평가가 이루어졌다. 자기 평가가 이루어진 사례는 모두 동료 평가를 같이 진행하였다.

앞서 평가의 시점에 대한 분석 결과와 유사하게, 교사들은 평가의 공정성과 타당성의 저해를 우려하여 평가의 주체를 확장하지 못하고 있었다. 특히 평가의 주요 목적으로 성적의 산출을 꼽으면서, 평가 주체의 확장은 ‘개인적 감정에 의한 평가 오류’, ‘학생들의 평가 기준에 대한 부정확성’ 등의 이유로 평가의 주요 목적과 상충되는 전략이라고 인지하고 있었다. 이와 동시에 아래 T 1이나 T 11이 언급하는 바와 같이 평가 주체의 확장에 대한 필요성도 느끼고 있었다. 특히 그 이유로 평가 방식의 다각화, 피드백의 제공 등을 꼽으면서, 반대로 평가를 통한 성적 산출과 평가를 활용한 학습 안내 및 참여 고취 측면

간의 딜레마 상황에 놓여 있는 것으로 보인다. 즉, 최근의 과정 중심 평가와 역량에 대한 강조는 교사들로 하여금 주체를 다양화 하여 평가를 계획해야 한다(Lee et al., 2017)는 인식을 쌓게 하는 것으로 보이나, 이에 대해 전문가들은 실례가 부족하거나 교사들이 직접 접하지 못하여 아직 현장에서 선부르게 시도하지 못하고 있는 것으로 예상하였다.

**[T 1]** 조별 활동을 보고서 형태가 아닌 다른 형태로 평가할 수 있는 도구를 개발해야 함. 조별 활동의 내용을 ‘보고서’라는 형식에 온전히 담을 수 없음. 발표, 동료평가를 추가해 보완하는 것을 생각할 수 있음.

**[T 11]** 평가 방법에 동료 평가와 자기 평가가 포함되면 좀 더 다양한 평가가 가능하고, 학생들에게 좀 더 긍정적인 피드백이 될 것으로 생각됨. 다양한 평가 방법에 대한 논의가 필요함.

타 과목에 비해 평가에 대한 부담이 적다는 측면에서(Byun et al., 2019) 평가의 주체를 확장시키는 노력이 요구되며, 본 연구에서 수집한 자료에서도 다양한 평가 주체들이 평가에 참여하도록 한 사례를 찾을 수 있었다. Figure 7에 제시한 평가 기준표는 교사 평가와 동료 평가, 자기 평가를 모두 골고루 사용한 두 교사의 사례이다. T 12는 각각 동료 평가와 자기 평가에 18%, 30%의 비중을 부여하였으며, T 19의 경우에도 동료 평가에 42%, 자기 평가에 18%로 평가 비율을 구성하였다. 두 교사 모두 학생들로 하여금 활동 내용이나 발표를 토대로 평가해야 하는 요소와 그 척도를 명시하였다는 점이 주목할 만하다. 이 사례들에서는 평가의 주체를 다양화하기 위해 학생들을 평가에 참여시키기 위한 목적과 구체적인 방안이 사전에 마련되어 있었고, 이를 고려하여 교수·학습 전략이 설계되어 있었다. 즉, 학생 스스로 평가에 활용하기 위한 평가 틀이나 기준이 활동 내용에 명시 해두었다.

**3. 평가 결과의 활용**

평가 결과는 학생들의 학습 상황을 파악하고, 추후 학습을 안내하고 교사 본인의 수업 개선 등의 목적으로 활용될 수 있다. 이에 따라 교사들에게 평가 결과를 어떻게 활용하였는지 묻은 결과를 Table 6에 제시하였다.

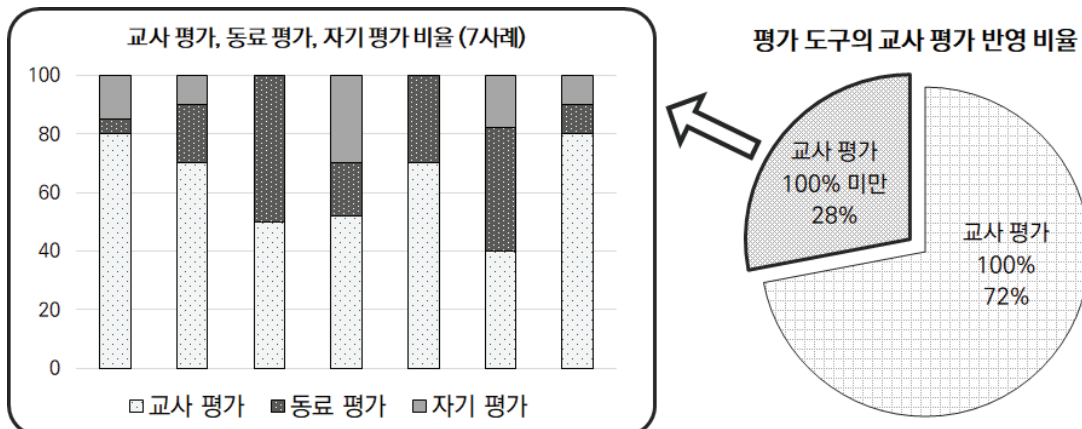
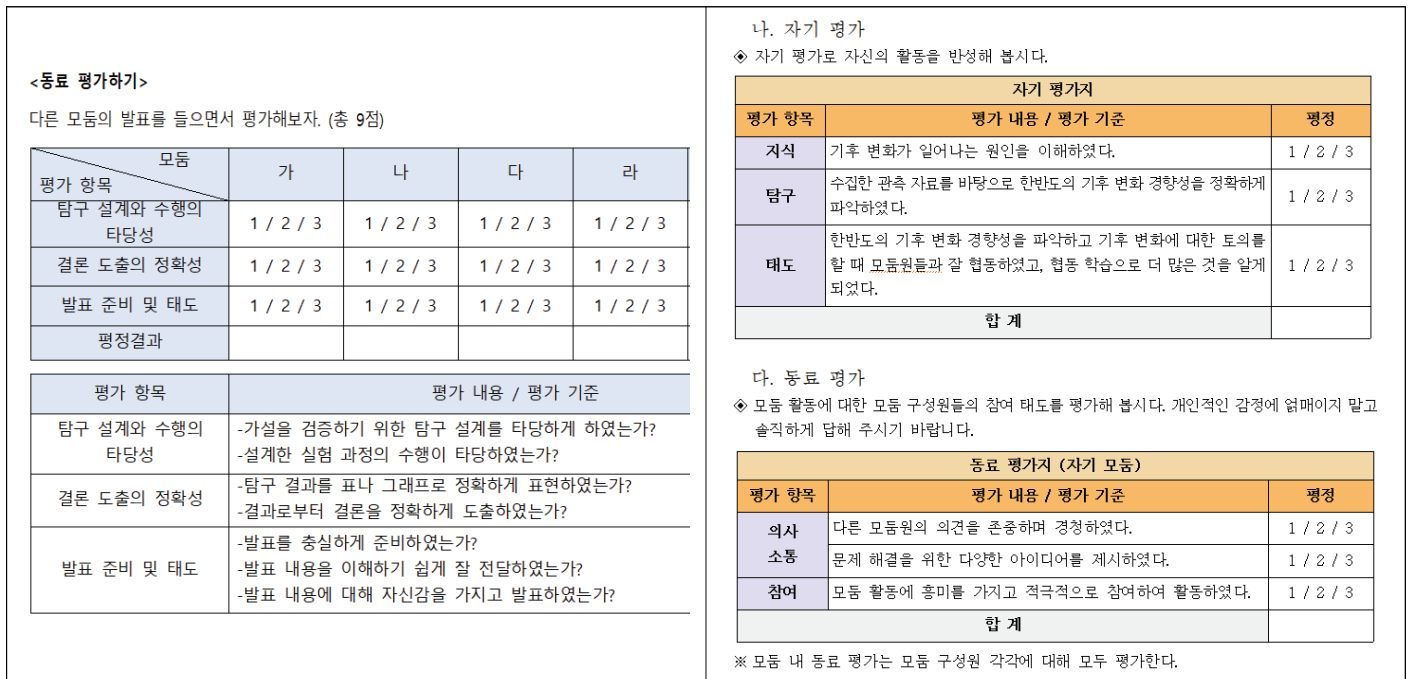


Figure 6. The ratio of assessment in terms of subject of assessment in collected cases



T 12의 평가 도구: 동료 평가 운영 사례

T 19의 평가 도구: 동료 평가 및 자기 평가 운영 사례

Figure 7. Two cases of assessment plan including peer- and self-assessment

Table 6. Results on the utilization of results of assessment

항목	빈도 (비율)
성적 산출	25 (100%)
수업 개선 - 과학탐구실험 수업 개선	7 (28%)
수업 개선 - 타 과목과의 연계성 확보	5 (20%)
학생 피드백	8 (32%)
기타	8 (32%)

\* 중복 응답을 허용하였음.

우선, 모든 교사들은 자신이 수합한 평가 결과를 성적 산출에 활용하였다. 이어 많은 비중을 차지한 것은 수업 개선으로 총 12명의 교사들이 응답하였다. 수업 개선은 크게 두 가지로 구분할 수 있었는데, 하나는 과학탐구실험 과목의 개선이었고, 다른 하나는 통합과학과 같은 타 과목과의 연계성 확보였다. 과학탐구실험 과목 자체에 대한 개선은 ‘학생들의 역량 파악 및 부족한 역량을 감안한 수업 설계’, ‘해당 차시 및 다음 차시 수업의 재구조화’ 등의 방식으로 이루어졌다. 이러한 사례들은 평가 결과를 자신의 수업에 대한 피드백으로 활용하는 대표적인 사례라고 할 수 있다. 다른 하나의 수업 개선 방식은 타 과목과의 연계성 확보이다. 이는 주로 통합과학과의 연계로 이루어졌는데, 아래 발췌한 T 25의 진술은 엘니뇨를 중심으로 자신이 같이 담당하고 있는 두 과목인 과학탐구실험과 통합과학을 연계한 대표적 사례라고 볼 수 있다.

**[T 25] 평가 도구의 사용 결과 많은 학생들이 수은 이외에 엘니뇨의 특징을 나타내는 물리량에 대한 이해가 부족하다는 것을 판단하여, 통합과학 시간에 이 부분에 대한 설명을 추가하여 수업함.**

평가의 결과를 피드백에 활용하였다고 응답한 교사들은 총 8명으로 32%에 해당하였다. 평가 도구를 분석한 결과에서는 피드백을 평

가 도구 내에 명시적으로 드러낸 사례를 확인하기 어려웠으나, 이는 대다수의 피드백이 계획적이라기보다는 학생 활동의 양상에 따라 맥락을 반영하여 이루어지기 때문으로 예상된다. 교사들은 ‘학생들이 우수한 영역과 부족한 영역’을 구분하여 설명하거나, ‘개선 방향을 안내 또는 아이디어 제공’의 방식을 택하기도 하고, 탐구 수행 중 참여 독려 및 격려의 방식으로 피드백을 제공하였다고 설명하였다. 다만, ‘교사 혼자 짧은 시간 내에 학생들의 특정 장면을 포착하여 피드백을 제공하는 것’은 한계가 있으며, ‘한 교사가 여러 반을 동시에 맡아 학생들을 관찰하고 꼼꼼히 기록’하는 것이 어려운 시간과 자원의 부족을 감안할 필요가 있음을 전문가들은 피력하였다. 즉, 교사가 모든 맥락에 따른 적절한 피드백을 제공하는 것에는 한계가 있음을 감안하여(Baek et al., 2015), 피드백을 효과적으로 제공하기 위한 방안 마련이 필요하다.

한편, 평가 결과의 활용에 대한 기타 사례로는 동료 평가와 관련하여 학생들의 수업 참여 고취를 위해 활용하였거나, 생활 기록부의 과목별 세부 능력 및 특기 사항의 기록에 평가의 결과를 활용한 경우가 있었다. 하단의 발췌문은 생활 기록부와 관련하여 평가 결과를 활용한 사례이다.

**[T 19] 학교생활기록부 교과 세부능력 특기사항 기재 : 장기간의 기상 데이터 분석 과정을 주어진 상황에 맞게 효과적으로 계획하고 실행하는 능력이 뛰어나. 모듈원들에게 균등한 역할을 배분하고 과정 전체를 점검하며 활동 목표에 도달하는 집단 창의 지성을 끌어내는 모습이 인상적임.**

이러한 접근 방식은 특히 학생들의 특징이나 성장을 기록하고 제시하기 위한 목적에서 활용될 수 있을 것으로 기대된다. 특히, 정량적으로 수치화하거나, 구체적으로 평가를 내리기 어렵다고 판단하는 경우에 적절한 방안이 될 수 있다. 즉, 학생의 학습 양상이나 변화에 대한 전체 맥락을 포괄적으로 기술할 필요가 있을 때 활용 가능성이 크다고 볼 수 있다.



#### IV. 결론 및 제언

본 연구에서는 25명의 교사들로부터 과학탐구실험에 적용하였던 평가 도구를 수집하여 분석을 실시하였다. 연구의 목적 상 평가가 이루어지는 전체적인 맥락 속에서 다양한 측면에서의 분석이 필요하였고, 이에 평가 요소, 평가 도구의 구성, 평가 결과의 활용의 범주로 구분된 분석틀을 개발하여 수집한 자료들을 구체적으로 살펴보았다. 분석한 결과에 의하면, 탐구 기능과 관련하여 기존의 과학 교과와 수행 평가에서도 많이 사용되었던 자료 분석 및 해석, 탐구 수행 등과 같은 평가 요소가 주로 활용되고 있었으며, 수행 과제 및 수업 방식의 영향을 많이 받는 문제 인식, 모형의 개발 및 사용 등과 같은 요소들은 활용 빈도가 낮은 것을 확인할 수 있었다. 이와 함께 과학탐구실험에서 과학 개념에 대한 평가도 많은 사례에서 관찰되었는데, 이 중 일부는 탐구의 맥락 속에서 평가가 이루어지기 보다는 수행의 맥락에서 벗어난 단순 과학 지식에 대한 평가가 이루어지기도 하였다. 정의적 영역과 관련된 평가도 절반 이상의 사례에서 관찰되기는 하였으나, 수업 참여도와 같은 측면 이외의 정의적인 요소에 대한 평가를 수행한 사례는 찾아보기 어려웠다.

한편, 평가 도구에서 평가의 목표는 충실하게 진술되어 있었으나, 세부 평가 요소별로 단계를 구분하거나 각 요소 별로 수준별 단계 진술이 이루어진 사례는 소수에 해당하였다. 이와 함께 평가가 이루어지는 방법의 측면에서 평가의 시점과 주체는 제한적인 것으로 조사되었다. 평가 결과는 모든 사례에서 성적 산출을 목적으로 활용되었고, 일부 교사들은 수업 개선과 학생에 대한 피드백을 위하여 활용하였다고 진술하였다. 결과적으로, 평가의 공정성과 객관성, 측정의 정량화 등의 이유로 인해 교과목의 핵심적인 목표들이 고르게 평가되는 데에는 한계점을 나타내었다.

과학탐구실험의 평가 운영은 3단계 절대평가로 체제가 변화하면서 많은 부분 개선이 가능할 것으로 기대되고 있다(Byun *et al.*, 2019; Kwak, 2020). 실제로 본 연구에 참여한 교사들도 이러한 전환에 따라 ‘상대평가 실시와 1등급 변별을 위한 지필고사’를 실시하지 않아도 된다는 점과 ‘학생들의 부담이 줄어드는 측면’이 교사에게 큰 도움이 된다고 설명하였다. 그럼에도 불구하고, 다양한 측면에서 평가의 어려움은 여전히 존재하고 있었다. 특히, 연구에 참여한 교사들은 학생의 성장 지원과 교사의 수업 개선을 꾀하는 이상적인 평가의 기능에 대해 인지하고는 있었으나, 이러한 지향점과 실제 수업에서 이루어지는 평가 사이에는 간극이 존재한다고 보고 있었으며, 결과적으로 일종의 딜레마를 경험하는 것을 은연중에 드러내고 있었다. 이러한 딜레마는 교육과정 문서에서 강조하는 과정 중심 평가나 평가의 본래 목적에서 의도하는 선순환적인 환류 체계에 대한 희망과 기대, 그리고 객관적이고 공정한 평가를 시행해야 한다는 관점이 서로 충돌하는 가운데 발생하였다.

구체적으로, 연구에 참여한 교사들은 성취 수준을 판단하여 성적 산출에 활용하기 위한 목적의 평가가 이루어져야 한다고 봄과 동시에, 다각화된 방식의 평가 도구를 활용하여 보다 나은 학습을 위한 목적의 평가 도구 개발을 희망하고 있었다. 특히, 활동 중심의 수업으로 운영되는 과학탐구실험의 특성 상, 수업에서 학생이 조사하여 기록하고 발표하는 일련의 과정과 평가를 동일시하여 평가를 개선함으로써 자신의 수업을 개선하고자 하는 의욕도 드러내었고, 연구 결과로부터

이러한 가능성을 일부 살펴볼 수 있었다. 하지만, 수업 개선과는 별개로 일종의 정량화된 산출 결과를 통해 공정성과 객관성도 담보하고자 함을 확인하였다. 이는 고등학교에서 이루어지는 대부분의 평가가 수치로 드러낼 수 있는 결과물로 산출되어야 한다는 생각은 과학탐구 실험에도 여전히 유효하였음을 드러내는 결과이며, 변화하는 맥락을 반영한 관점에서의 평가를 시도하는 것에 제약으로 작용하였다. 더불어 학기가 시작되기 이전에 평가의 계획을 모두 수립해야 하며, 학기가 진행되는 중에는 많은 업무로 인해 평가 방안에 대해 고민할 수 있는 시간적 여유가 부족한 상황에서 지향하는 바를 실천으로 옮기기에는 현실적 어려움이 많은 것으로 보인다. 이를 극복하기 위해 앞서 발췌한 참여자 T3의 의견과 같이 다양한 평가 사례를 접하거나 논의할 수 있도록 공유와 협력을 위한 제도적 지원이 필요함을 많은 교사들은 피력하였다.

교사들은 이러한 맥락 속에서 제한적이거나 변화를 꾀하고 있었다. T 12는 “올해(2019년)부터는 담당 교사 간에 수업의 재구성과 평가에 대한 필요성을 인식”하여 ‘점진적으로 변화를 꾀하고 있다’고 설명하였고, T 3 역시 “2019학년도부터 절대평가로 평가가 이루어지고 있으므로 교과목의 개설 목적에 부합한 평가가 실시되고 있음”을 강조하고 있었다. 또한, 학생들의 탐구 과정과 결과를 학교생활기록부의 기록에 반영하는 등, 다양한 방식으로 평가 결과를 활용하려는 접근도 살펴볼 수 있었다. 즉, 아직은 일부이나 변화하는 맥락을 반영한 평가 방안의 모색이 시작되고 있었다. 이는 평가에서 중요하게 살펴야 하는 측면에 대한 관점 변화와 과학탐구실험에서 학생들의 성취를 어떠한 측면에서 살펴야 하는지 대변한다고 볼 수 있다. 즉, 탐구를 중심으로 하여 핵심역량을 함양할 수 있도록 안내하고(Jang *et al.*, 2015), 이 과정에서 그 동안 상대적으로 소홀했던 평가 결과를 토대로 이루어지는 교수학습의 개선이 그것이다(Park *et al.*, 2018). 본 연구에 참여한 교사들이나 전문가들 역시 이와 같은 방향에 대해 공감하나, 학교 현장의 복잡한 맥락을 감안하였을 때 보다 명료한 안내와 연수, 지원이 필요하고, 과정 중심 평가의 목적과 방향에 대한 안내가 요구된다고 생각하였다. 특히, 교육과정 문서나 학술적 연구 수준에서 논의되는 사항들이 학교 현장에서 뿌리내리고 적합한 형태로 안착되기 위해서는 실천적 차원에서의 운영 방안을 제안하고 적용 결과에서 나타나는 사항들을 검토하여 현장에 다시 적용할 수 있는 순환적 절차가 요구된다.

본 연구의 결과로부터 고려할 수 있는 하나의 방안은 ‘교사의 평가 결과로부터 즉각적으로 피드백을 제공할 수 있는 자료’를 개발하고 활용하는 것이다. Park *et al.*(2018)의 연구에서 언급한 바와 같이 ‘구체적인 피드백의 시점과 내용’을 미리 마련하는 것이 자료를 개발하는 단초의 역할을 할 수 있다. 이는 평가 도구의 구성 측면과도 연관될 수 있는데, 평가 도구의 개발 과정에서 평가 요소를 세분화함과 동시에 각 요소의 성취 수준에 따른 행동 특성을 구분하고 이를 확인할 수 있는 시점을 수업 시간 내에 마련하는 것이다. 즉, 구분된 단계에 따라 현재의 행동 특성과 상위 성취 수준의 행동 특성 간 차이를 확인하여 상위 성취 수준에 도달하기 위한 전략과 방안을 교사가 제공하는 형태로 피드백을 계획할 수 있다. 이는 세부 평가 기준과 수업 시간 내에 명시적 피드백 제공 시점 등이 유기적으로 상호 연관됨으로써 과정 중심의 평가를 실현하는 데에도 기여할 수 있다. 다만, 이와 같은 접근을 위하여 교사들에게 사전에 많은 실례와 활용도와 확장성

이 높은 평가 기준 등이 제공될 필요성이 있다.

더불어 앞선 현실적 지원 방안을 마련함과 동시에 과학탐구실험의 목표가 무엇이고, 학생들의 탐구 역량 신장을 위해서 어떠한 평가 요소들이 고려되어야 하는지에 대한 논의가 활발히 이루어질 필요가 있다. 즉, 현실적 지원 방안이 평가 도구의 구성 측면에서 수업의 진행과 평가의 활용이 유기적으로 연계되는 시스템 구성을 강조하였다면, 이와 함께 평가하고자 하는 요소를 타당하게 평가하는 방식과 사례를 논의하고 공유할 필요가 있다. 본 연구에서 수집한 자료들에 담겨있는 평가 요소들이 어떻게 평가되고 있는지를 구체적으로 살펴보면, 평가 기준들이 해당 평가 요소들의 수행을 통해서 발휘되고 점검되어야 하는 특성들에 대해서 명확하게 짚어내지 못하고 있음을 확인할 수 있었다. 예를 들어, ‘증거에 기초한 토론과 논증’에 대해 평가하기 위해서 수업 속에 토론, 논의, 발표 등의 활동은 포함되어 있으나, 활동 속에서 학생들의 논의가 타당한 근거에 기초하고 있는지, 근거에 기반 하여 논의가 이루어지고 있는지에 대해서 평가하기 보다는 수집된 근거를 토대로 주장하고 있는 과학 지식이 타당한지, 모둠원들과의 토의 활동에 적극적으로 참여하였는지에 초점을 두고 평가 기준이 설정되어 있었다. 즉, 탐구 과정에 대한 평가는 탐구 주제와 관련된 내용들을 잘 이해하고 있는지도 중요하지만 탐구의 결과를 어떻게 획득하였고, 그 자료가 어떻게 과학 개념과 상호 연관되어 설명이 이루어지고 있는지에 대해서도 함께 평가될 필요가 있는데 수집된 자료에서는 이와 관련된 좋은 사례를 찾아보는 데 어려움이 있었다. 결과적으로, 과학탐구실험에서의 평가가 과목의 목표에 부합되기 위해서는, 교사들 스스로 평가하고자 하는 요소가 ‘탐구를 통해 학습’하는 데 있어서 어떠한 의미가 있는지를 먼저 살피고 이에 근거하여 평가의 요소와 장면을 구성하는 역량을 신장시켜 나가야 할 것으로 생각된다.

다음으로, 교과목의 목표에 초점을 둔 평가가 이루어지기 위한 정책적 차원에서의 변화를 제안하고자 한다. 학교 현장에서 평가의 가장 중요한 기능은 성적 처리와 등급의 부여임은 부인할 수 없는 사실이나, 이로 인해 다양하고 자유로운 평가의 실행 또한 제한되고 있다. 본 연구에서 수집한 사례에서도 ‘측정 가능한 요인’ 중심으로 평가 요소들이 선정되고, 실제 학습자들이 탐구에 대해 얼마나 이해하고 있으며, 탐구 역량이 어떻게 변화하고 있는지에 대한 점검은 거의 이루어지고 있지 못하였다. 또한 탐구 프로그램의 효과를 관찰하는 많은 연구들에서는 과학의 본성에 대한 이해, 과학 탐구에 대한 이해, 정의적 영역의 측면에서의 학생들의 변화를 살펴보고 있으나, 실제 학교 현장에서 이러한 부분들이 활발하게 적용되지는 못하고 있다. 현재의 다소 완화된 평가 지침을 좀 더 과감하게 변화시켜, 과학탐구실험에서의 평가 목표를 성적이나 등급 산출이 아닌 역량 점검에 초점을 둘 수 있도록 하는 것은 시도해 볼직하다. 그리고 이러한 변화를 바탕으로 다양한 검사 도구들을 활용하여 학습자들의 탐구 역량 및 과학 학습에 있어서의 정의적 영역을 거시적인 관점에서 점검해 나가는 방식으로 평가의 실천 방향을 전환시킨다면, 과학탐구실험의 본래 목표를 달성할 수 있을 것으로 보인다.

본 연구에서는 과학탐구실험의 평가 현황을 살펴보기 위하여 다양한 평가 도구 사례를 수집하여 분석하였다. 다만, 수집된 사례들은 각 교사들이 과학탐구실험에서 이용하였던 평가 도구 중 하나이기 때문에, 보통 한 학기 동안 운영되는 과학탐구실험 수업 전체를 조망

하는 데에는 한계가 있다. 또한 평가 도구와 교사의 설명 자료만을 토대로 분석을 시도하였기 때문에, 평가의 맥락을 정확하게 짚어내는 데에도 어려움이 있었다. 따라서 향후 본 연구를 토대로 과학탐구실험의 실제 수업 장면을 살피고 순환적인 평가를 위해 요구되는 사항은 어떠한 것들이 있는지 보다 구체적으로 파악할 필요가 있다. 특히, 실제 수업의 맥락에서 일어날 수 있는 피드백의 형태와 그 결과의 환류가 어떻게 이루어지는지에 대해서 심층적으로 분석할 필요도 있다. 이러한 접근은 역동적으로 이루어지는 학생 활동 중심의 과목 운영에서 일어나는 구체적인 학습과 평가를 확인하고, 실제적인 개선 방안을 탐색하는 데 활용될 수 있을 것이다. 더불어 본 연구가 제안하는 평가 방안과 개선 사항들에 대한 활용 가능성을 탐색하는 연구가 이루어진다면, 평가의 개선을 통한 탐구 중심 과학 교육의 질 제고에 도움이 될 것이다.

## 국문요약

2015 개정 교육과정의 도입됨에 따라 신설된 과학탐구실험은 과학적 실천을 통한 과학의 학습과 핵심역량의 함양 등의 측면에서 기대를 받고 있는 과목이다. 특히, 변화하는 평가의 맥락에 기초하여 실천 중심의 교과목인 과학탐구실험에서의 다각화된 평가가 운영될 필요가 있으나, 이에 대한 다양한 어려움과 한계가 보고되고 있다. 이에 본 연구는 과학탐구실험의 평가 현황을 살피고, 이로부터 평가의 운영과 지원 방안을 고찰하고자 하였다. 이를 위해 25명의 현직 과학 교사로부터 과학탐구실험 수업에서 사용되었던 평가 도구와 이에 대한 설명 자료를 수집하여 분석하였다. 평가 도구의 분석은 평가 요소, 평가 기준, 평가 방법의 측면에서 이루어졌고, 평가 결과가 어떻게 활용되는지도 살펴보았다. 결과에 대한 타당성 확보를 위해 과학교육 전문가와 교사들로부터 점검을 받았고, 추가적으로 분석 결과에 대한 의견도 청취하여 연구에 활용하였다. 연구 결과, 교사들은 탐구 기능 요소 중 특정 요소만을 중점적으로 평가하고 있었고, 일부 기능 요소는 특정한 주제 및 맥락과 연계하여 평가되었으나 활용 빈도는 낮았다. 또한, 인지적 영역과 정의적 영역에 대한 평가도 이루어지기는 했으나, 각 영역을 평가하는 데 있어서 한계점도 관찰할 수 있었다. 평가 기준의 측면에서 평가 목표는 대체적으로 제시되어 있으나, 평가 목표에 대한 세부 평가 요소의 구분 및 각 요소별 행동 특성을 단계적으로 진술하는 등의 구체적 기준 수립이 이루어진 사례는 상대적으로 적었다. 평가 방법에서 평가의 시점과 주체는 각각 ‘수업 후 평가’와 ‘교사 평가’가 주로 사용되었고, 다른 방법은 특정 내용, 특정 상황으로 한정되어 활용되었다. 평가 결과는 모든 사례에서 성적 산출을 위해서 활용되었고, 일부만이 수업 개선과 학생 피드백을 위한 목적으로 활용하고 있었다. 이러한 결과를 바탕으로 과학탐구실험에서의 학생 평가를 지원하기 위한 방안에 대해 논의하였다.

**주제어 :** 과학탐구실험, 학생 평가, 과학 탐구, 2015 개정 교육과정

## References

- Baek, J., Koh, E., Cho, Y. H., & Jeong, D. H. (2015). Authentic thinking with argumentation: putting on the thinking caps of scientists and designers. In Cho, Y. H., Caleon, I. S., & Kapur, M. (Eds.). *Authentic*

- Problem Solving and Learning in the 21st Century: Perspectives from Singapore and Beyond (pp. 173-191). Springer.
- Byun, T., Baek, J., Shim, H.-P., & Lee, D. (2019). An investigation on the implementation of the ‘scientific inquiry experiment’ of the 2015 revised curriculum. *Journal of the Korean Association for Science Education*, 39(5), 669-679.
- de Vries, E., Lund, K., & Baker, M. (2002). Computer mediated epistemic dialogue: Explanation and argumentation as vehicles for understanding scientific notions. *Journal of the Learning Sciences*, 11 (1), 63-103.
- Fogleman, J., McNeill, K. L., & Krajcik, J. (2011). Examining the effect of teachers’ adaptations of a middle school science inquiry-oriented curriculum unit on student learning. *Journal of Research in Science Teaching*, 48(2), 149-169.
- Ha, M., Park, H., Kim, Y.-J., Kang, N.-H., Oh, P. S., Kim, M.-J., Min, J.-S., Lee, Y., Han, H.-J., Kim, M., Ko, S.-W., & Son, M.-H. (2018). Developing and applying the questionnaire to measure science core competencies based on the 2015 revised national science curriculum. *Journal of the Korean Association for Science Education*, 38(4), 495-504.
- Jang, Y.-R., Kang, N.-H., Kang, S., Ko, S., Kwak, Y., Kwon, H., Kim, K., Kim, M., Kim, S., Kim, Y., Kim, J., Kim, H., Kim, H.-K., Kim, H., Na, J., Min, B., Park, K., Park, B., Park, S., Park, S., Park, C., Park, H., Bang, D., Byun, T., Son, J., Song, J., Shin, Y., Sim, K., Ahn, J., Ahn, P., Oh, S., Oh, W., Yoon, H., Lee, K., Lee, M., Lee, M., Lee, M., Lee, Y., Lee, Y., Lee, I., Lee, J., Lee, H., Lee, H., Lim, H., Jang, S., Jeon, H., Jeong, W., Jwa, Y., Choi, I., Cho, J., Choi, H., Choi, H., Hyeon, J., Hong, J., Hwang, U., & Hwang, I. (2015). A study on development prototype of 2015 science national curriculum. BD15070002, Seoul: Korea Foundation for the Advancement of Science & Creativity.
- Jho, H. (2018). An analysis of elements of scientific inquiry presented in 2015 revised national science curriculum: focusing on scientific inquiry experiment. *Journal of Research in Curriculum & Instruction*, 22(3), 208-218.
- Kim, H.-K., Lee-M., Lee, J., Lee, S., Lee, Y.-R., Shin, Y., Ahn, J., Lee, S., Cho, S., Ji, J., J, H., Ihm, H., Park, C., & Ko, S. (2017). Research and development of the assessment standards of the 2015 revised national science curriculum at the high school levels. CRC 2017-5-7, Korea Institute for Curriculum and Evaluation.
- Kim, M. & Kang, N.-H. (2018). Analysis of inquiry activities in high school science inquiry laboratory textbooks. *Chungnam Science Education Journal*, 24(2), 60-69.
- Kim, M., & Ha, S. (2020). Development of process-oriented inquiry-based science performance assessment method and application to science gifted high school students. *New Physics: Sae Mulli*, 70(5), 454-465.
- Kim, M., Kim, S., Hoh, T., & Choi, S. (2019). The influences of integrated science and science inquiry experiment developed under the 2015 revised national curriculum on students’ interest in science, scientific attitude, views on science-technology-society relationship, and views on nature of science. *Journal of the Korean Association for Science Education*, 39(6), 791-797.
- Kim, Y.-J., Jang, W., & Hong, H.-G. (2019). The analysis and implication of student evaluation in 2015 science curriculum using text network analysis (TNA) method: focused on the connection with process centered evaluation. *The Journal of Curriculum and Evaluation*, 22(3), 225-250.
- Kwak, Y. (2020). Trend analysis of curriculum application status of 2015 revised integrated science and scientific laboratory experiment curriculum. *Journal of Korean Society of Earth Science Education*, 13(1), 53-63.
- Lederman, J. S., Lederman, N. G., Bartos, S. A., Bartels, S. L., Meyer, A. A., & Schwartz, R. S. (2014). Meaningful assessment of learners’ understandings about scientific inquiry: The views about scientific inquiry (VASI) questionnaire. *Journal of Research in Science Teaching*, 51(1), 65-83.
- Lee, H.-N., Lee, Y.-J., & Lee, H.-D. (2017). Development and application of rubric for assessing scientific inquiry process. *Secondary Education Research*, 65(1), 145-171.
- Lee, I., Park, S., Shim, H.-P., & Lee, J. (2017). Development and application of an assessment procedure to enhance core competencies: focusing on communication and social competencies. RRE 2017-7, Korea Institute for Curriculum and Evaluation.
- Lee, J., Lee, K., & An, J. (2019). An analysis of the authentic inquiry components in science inquiry experiments textbooks developed under the 2015 revised national curriculum. *Journal of the Korean Chemical Society*, 63(3), 183-195
- Ministry of Education(MOE). (2015a). The general explanation of 2015 revised national curriculum. Notification No. 2015-74 [issue 1]. Sejong: Ministry of Education.
- Ministry of Education(MOE). (2015b). Science curriculum. Notification No. 2015-74 [issue 9]. Sejong: Ministry of Education.
- National Research Council(NRC) (2012). A framework for K-12 science education: practices, crosscutting concepts, and core idea. Washington, DC: National Academy Press.
- OECD (2017). Pisa 2015 assessment and analytical framework: science, reading, mathematic, financial literacy and collaborative problem solving. OECD Publishing.
- Oh, P. S. (2020). A critical review of the skill-based approach to scientific inquiry in science education. *Journal of the Korean Association for Science Education*, 40(2), 141-150.
- Osborne J., Simon, S., & Collins, S. (2003). Attitudes towards science: A review of the literature and its implications. *International Journal of Science Education*, 25(9), 1049-1079.
- Park, H., Sim, J.-H., Choi, H., Lim, H., Park, J., Ahn, H., & Yang, S. (2019). Study on the monitoring and implementation of the 2015 science national curriculum. BD19030006, Seoul: Korea Foundation for the Advancement of Science & Creativity.
- Park, J., Jin, K.-A., Kin, S., & Lee, S.-A. (2018). A Study for enhancing the teacher’ expertise in student assessment for stabilizing process-fortified assessment policy. RRE 2018-5, Korea Institute for Curriculum and Evaluation.
- Sampson, V., Enderle, P. J., & Walker, J. P. (2011). The development and validation of the assessment of scientific argumentation in the classroom (asac) observation protocol: a tool for evaluating how students participate in scientific argumentation. *Perspectives on Scientific Argumentation*, 235-264.
- Shin, S., Park, C., Lee, C. & Hong, H.-G. (2018). A case study on the practice of ‘science inquiry experiment’ in the 2015 revised national curriculum: an understanding in the perspective of cultural-historical activity theory(CHAT). *Journal of the Korean Association for Science Education*, 38(6), 885-899.
- Shin, Y., Kwak, Y., Sung, J., Lee, J., & Lim, E. (2018). Research on the quality control for the implementation of new courses according to the 2015 revised curriculum. Sejong: Ministry of Education.
- Song, J. (2006). J. J. Schwab’s life and his ideas of science education. *Journal of the Korean Association for Science Education*, 26(7), 856-869.
- Song, J., Kang, S.-J., Kwak, Y., Kim, D., Na, J., Do, J.-H., Park, S. C., Son, Y.-A., Son, J. W., Oh, P. S., Lee, J.-K., Lee, H. J., Ihm, H., Jeong, D. H., Joung, Y. J., & Kim, J. (2019). Developing performance expectations, school implementation strategies, evaluation indicators of the Korean science education standards (kses) for the next generation. BD19080002, Seoul: Korea Foundation for the Advancement of Science & Creativity.
- Song, S., & Shim, K. (2019). Analysis of key competencies in inquiry activity of science inquiry experiment textbooks for high school students. *Journal of Learner-Centered Curriculum and Instruction*, 19(22), 363-383.
- Victoria, A. S. C., & Lee, R. J. (2017). Timss 2019 science framework, In I. V. S., Mullis & M. O. Martin (Eds.). *Timss 2019 assessment frameworks*. Retrieved from Boston College, IEA TIMSS & PIRLS International Study Center. Retrieved 2020. 10. 5. from <http://timssandpirls.bc.edu/timss2019/frameworks/>
- Yang, I.-H, Lee, H.-J., Lee, H., Cho, H. (2009). The development of rubrics to assess scientific argumentation. *Journal of the Korean Association for Science Education*, 29(2), 203-220.
- Yun, D., & Choi, A. (2019). Analysis of achievement standards, activities, and assessment items in the 2015 revised science curriculum and grade 7 science textbooks: focusing on science core competencies. *Journal of the Korean Chemical Society*, 63(3), 196-208.
- Yun, D., Ko, E., & Choi, A. (2018). Identifying and applying components of five scientific core competencies in the 2015 science curriculum. *Journal of Learner-Centered Curriculum and Instruction*, 18(24), 1301-1319.

## 저자정보

백중호(한국교육과정평가원 부연구위원)  
 변태진(한국교육과정평가원 부연구위원)  
 이동원(한국교육과정평가원 부연구위원)  
 심현표(한국교육과정평가원 부연구위원)