



과학탐구실험의 ‘역사 속의 과학 탐구’에서 과학교사의 평가 실태와 평가 지향 조사

김민환, 박다해, 노태희*
서울대학교

Science Teachers' Actual and Preferred Cases of Assessment in 'Scientific Inquiries in History' of Science Inquiry Experiment

Minhwan Kim, Dahae Park, Taehee Noh*
Seoul National University

ARTICLE INFO

Article history:

Received 28 September 2022

Received in revised form

17 November 2022

Accepted 28 November 2022

Keywords:

the 2015 Revised Curriculum, Scientific Inquiry Experiment, NOS, assessment

ABSTRACT

In this study, we investigated actual cases of assessment science teachers conducted and the cases they preferred in a typical situation based on the curriculum in the context of 'Scientific Inquiries in the History' of Scientific Inquiry Experiment. A questionnaire composed of descriptive questions was developed and a survey was conducted with 70 science teachers with experience in teaching 'Scientific Inquiries in History'. Interviews were conducted with eight of them. The assessment cases were analyzed in terms of the assessment areas and assessment methods, and the results were compared. The analyses of the results revealed that 'scientific inquiry ability' accounted for the highest ratio of the assessment areas in the actual cases of assessment. There were few cases that assessed the core concepts presented in the curriculum, 'the nature of science' and 'scientists' inquiry methods'. The assessment methods were greatly biased toward the report method and various assessment methods were not used. In preferred cases of assessment, the ratio of cases that assessed the core concept increased slightly, however the frequencies remained at a low. As for the assessment methods in preferred cases of assessment, the measurement methods decreased, the performance methods increased, and the informal methods which were not shown in the actual cases appeared. However various assessment methods were still not used. The causes of the survey results were analyzed based on the opinions of the teachers who participated in the interviews. Based on above results, plans to actively conduct NOS assessments in Scientific Inquiry Experiment are discussed.

1. 서론

과학의 본성(nature of science; 이하 NOS)은 그동안 과학적 소양의 함양을 위해 과학교육의 중요한 목표 중 하나로 여겨져 왔다(AAAS, 1993; McComas & Nouri, 2016; NRC, 2011; Song *et al.*, 2019). 2015 개정 과학과 교육과정에서 새로 개설된 공통 과목인 과학탐구실험은 이전의 다른 과학 과목과 달리 NOS를 명시적인 학습 내용으로 교육 과정에 포함하고 있다. 과학탐구실험은 총 3개의 대단원으로 구성되는데, 교육과정에서 1단원 '역사 속의 과학 탐구'의 성취기준을 "과학자들이 행했던 역사적인 실험들을 과학의 탐구 과정을 따라 수행함으로써 과학의 본성을 깨닫도록 한다."로 제시하고 있다(MOE, 2015). 또 교육과정의 내용 체계에서는 핵심 개념으로 '과학의 본성'과 '과학자의 탐구 방법'을 제시했으며, 두 핵심 개념에 대응하는 일반화된 지식을 각각 "과학자들의 탐구실험에서 과학의 다양한 본성이 발견되며, 과학 탐구 수행 과정에서 과학의 본성을 경험한다."와 "주제에 따라 다양한 과학 탐구 방법이 활용된다."로 설명하고 있다(MOE, 2015). NOS를 정의하는 방법은 연구자에 따라 다양하지만 핵심 개념을 설명한 일반화된 지식을 살펴볼 때 두 핵심 개념 모두 과학교육

분야에서 일반적으로 일컫는 NOS의 범위에 포함된다고 할 수 있다. 이처럼 우리나라의 교육과정에서 특정 과목의 내용 체계에 NOS를 명시한 것은 과학탐구실험에서 거의 처음 이루어진 시도이다. 따라서 과학탐구실험의 1단원 '역사 속의 과학 탐구'에서 NOS 수업을 적극적으로 실행할 기회가 마련된 것이라고 할 수 있다.

한편, 최근의 연구들에서는 과학교사들의 NOS 수업에서 평가가 부족하다는 점을 일관되게 지적하고 있다(Bell *et al.*, 2000; Hanuscin, 2013; Hanuscin *et al.*, 2011; Supprakob *et al.*, 2016). 예를 들어 과학 교사들은 NOS를 가르치기 위해 다양한 교수전략을 활용하였으나 NOS에 대한 학생들의 이해를 평가하기 위해서는 특별한 전략을 갖고 있지 않았다(Hanuscin *et al.*, 2011). 또한 NOS 수업에서 퀴즈나 활동지, 발표 등을 이용하여 평가를 하더라도 과학 내용 지식만을 평가하고 NOS는 평가하지 않았다(Supprakob *et al.*, 2016). 과학교사들이 과학탐구실험에서 실행한 NOS 수업을 분석한 Kim *et al.*(2020)의 연구에서도 네 명의 교사 중 세 명은 NOS가 인지적 학습의 영역이 아니라거나 NOS에 대한 견해가 개인의 주관이라는 등의 이유로 NOS에 대한 평가를 하지 않았다. 그러므로 학교 현장의 과학탐구실험 수업에서도 NOS에 대한 평가가 제대로 이루어지고 있지 않을 가능성이 있다. 특히나 일반 과학 과목과 달라 교사들에게 낯선 과학탐구실

* 이 논문은 2021년 대한민국 교육부와 한국연구재단의 인문사회분야 중견연구지원사업의 지원을 받아 수행된 연구임(NRF-2021S1A5A2A01061452)

교신저자 : 노태희 (nohth@snu.ac.kr)

<http://dx.doi.org/10.14697/jkase.2022.42.6.597>

험의 평가 지침은 더욱 높은 수준의 평가 전문성을 요구한다. 2019학년도부터 성취도를 3단계로 산출하는 절대평가 방식의 평가 지침이 과학탐구실험에 적용되었고, 이에 따라 이 과목에서는 교사들에게 상대적으로 익숙한 지필평가가 아니라 수행평가를 중심으로 평가가 이루어지고 있다(Byun *et al.*, 2019). 따라서 과학탐구실험에서 NOS를 명시한 교육과정의 취지에 맞게 NOS에 대한 평가가 이루어지고 있는지 그 실태를 조사할 필요가 있다.

NOS 평가의 실태를 알아보는 것과 더불어 교사들이 지향하는 평가를 살펴보는 것도 중요하다. 현실적인 환경적 제약으로 인해 평가 실태는 교사들이 지향하는 평가를 온전히 반영하지 못했을 가능성이 있기 때문이다. 예를 들어 구성주의적 평가를 충분히 알고 있는 교사들도 실제 수업에서는 전통적인 평가를 실시할 수 있다(Thomas *et al.*, 2011). 유사한 관점으로 Noh *et al.*(2015)은 과학교사들이 과학수업에서 실제로 수행했던 평가 사례와 전형적인 가상의 상황에서 실행할 것이라고 지향한 평가 사례를 조사하였다. 그 결과 교사들이 지향한 평가는 실행한 평가보다 다양한 평가 방법을 포함하였으나 평가 목적 측면에서는 실태와 마찬가지로 여전히 전통적인 평가 목적을 벗어나지 못하는 것으로 나타났다. 과학탐구실험을 담당하는 교사들은 ‘수업 준비 시간과 자료의 부족’ 같은 외적 요인과 ‘탐구 교육에 대한 전문성 부족’과 같은 내적 요인으로 현실적인 어려움을 겪고 있다(Byun *et al.*, 2019). 따라서 교사들이 과학탐구실험 수업에서 실행한 평가 사례와 전형적인 상황에서 지향하는 평가 사례가 적지 않은 차이를 보일 가능성이 있다. 이때 평가의 실태와 지향에서 나타나는 구체적인 차이와 그 원인을 분석한다면, NOS 평가에 대한 교사들의 인식과 전문성을 더욱 심층적으로 조사할 수 있을 것이다. 예컨대 실제로 교사들이 NOS를 평가하지 않고 있을 때 지향에서는 NOS를 평가하고자 한다면 교사들이 지향하는 평가를 실행하지 못하도록 하는 현실적인 어려움은 무엇인지 조사할 수 있으며 지향에서도 NOS를 평가하고자 하지 않는다면 그 이유를 조사할 필요가 있다.

과학탐구실험에서의 평가와 관련된 선행 연구들을 살펴보면, 우선 과학탐구실험에서 평가 도구와 현황을 분석한 Baek *et al.*(2020)의 연구가 있었다. 이 연구에서는 25명의 교사로부터 과학탐구실험의 대표적인 평가 사례를 1인당 1건씩 총 25건의 평가 사례를 수집하였고, 이를 평가 요소, 평가 기준의 기술 수준, 평가 방법의 세 가지 측면에서 분석하였다. Shim & Yoo(2020)는 경기도에 소재한 고등학교 10개교를 대상으로 학교알리미에 공시된 통합과학과 과학탐구실험의 수행평가 채점기준을 분석하여 두 과목의 수행평가에 반영된 과학과 핵심 역량을 조사하였다. 또 과학탐구실험의 운영 실태를 조사한 Byun *et al.*(2019)의 연구에서는 설문지의 일부 항목으로 평가 방식 및 평가 반영 비율을 조사하기도 하였다. 그러나 이상의 연구 모두 과학탐구실험의 전반을 대상으로 하였으며 주제나 내용에 따른 분석이 이루어지지 않았으므로 1단원의 특성을 고려한 분석은 이루어지지 못했다. 다시 말해 NOS를 명시한 교육과정의 취지에 맞게 ‘역사 속의 과학 탐구’ 수업에서 NOS에 대한 평가가 이루어지고 있는지

파악할 수 없었다.

이에 이 연구에서는 과학탐구실험의 1단원 ‘역사 속의 과학 탐구’ 수업에서 과학교사들이 실행한 평가 사례와 가상의 상황에서 지향한 평가 사례를 조사하여 비교하였다. 이로써 과학탐구실험에서 이루어지는 평가의 실태를 파악하고, 평가 실태와 지향의 차이로부터 평가 실태를 개선하는 방법을 탐색하였다.

II. 연구 방법

1. 연구 참여자 및 연구 절차

본 연구의 목적에 맞게 설문을 개발하고 2021년 1학기 말에 온라인 설문을 실시하였다. 2018학년도부터 2021학년도까지 과학탐구실험의 1단원 ‘역사 속의 과학 탐구’를 가르친 경험이 있는 과학교사 70명(교사 1~70)이 설문문에 참여하였다. 한 학교에서 여러 명의 교사가 학급을 나누어서 가르친 경우에는 담당 교사가 다르더라도 동일한 평가를 실행했을 가능성이 높으므로, 1명의 교사만 설문문에 참여하도록 요청하였다. 같은 학교에 근무 중인 교사라도 다른 학년도에 과목을 담당했던 교사들은 평가를 다르게 실행했을 수 있으므로 모두 설문문에 참여할 수 있도록 하였다. 또한 교사들이 설문에서 사용한 용어의 의미를 이해하고 설문 결과의 의미를 구체적으로 파악하기 위해 설문문에 참여한 교사 중 면담 참여 의사를 밝힌 8명(교사 A~H)을 대상으로 면담을 실시하였다. 면담에서는 설문에 응답한 평가 사례를 바탕으로 수업 상황, 평가 문항과 평가 기준 등의 구체적인 정보를 자세히 설명하도록 요청하였고, 과학탐구실험에서의 평가, NOS 평가에 대한 의견도 질문하였다. 연구자가 속한 기관의 생명윤리위원회(IRB)에서 승인을 받아 모든 연구 절차를 진행하였으며, 연구에 참여한 교사들에게는 상응하는 보상을 지급하였다.

2. 설문 도구

교사들이 과학탐구실험 1단원에서 실제로 실행한 평가 사례와 가상으로 제시된 예시 상황에서 지향하는 평가 사례를 조사하는 설문지를 개발하였다. 교사들이 자신의 응답을 자유롭게 기술하도록 서술형 문항으로 설문지를 구성하였으며, Noh *et al.*(2015)이 과학 교사의 평가 실태와 지향을 조사하기 위해 개발한 설문 도구에 기초하여 평가 실태와 평가 지향의 두 가지 영역으로 구성하였다(Table 1).

먼저 평가 실태 영역에서는 교사가 과학탐구실험 1단원에서 실시했던 모든 평가 사례를 평가 영역과 평가 방법에 초점을 두고 가능한 자세히 기술하도록 하였다. 이때 성적 산출을 위해 점수화하지 않은 사례도 포함하도록 하였다. 교사들이 꼭 응답해야 하는 평가 영역과 평가 방법을 강조하기 위해 예시 문구를 4개 제시하고, 평가 영역과 평가 방법에 해당하는 부분을 각각의 예시 문구에서 표시하였다. 예를 들어 “학생들이 자신만의 주기율표를 만들고, 여기서 찾은 원소의

Table 1. Composition of the questionnaire

| 영역 | 서술 내용 | 응답 개수 |
|-------|--|-----------------|
| 평가 실태 | 평가 영역과 평가 방법을 포함하도록 자신의 평가 사례를 자유롭게 서술 | 최대 5개 응답 가능 |
| 평가 지향 | 주어진 예시 상황에서 지향하는 평가 사례를 평가 영역과 평가 방법을 포함하도록 서술 | 예시 상황당 1개 응답 가능 |

주기적 성질을 활동지에 작성했는지 여부(평가 방법)로 수업 참여도 (평가 영역)를 확인하여 수평평가 점수에 반영함.”과 같이 예시 문구를 제시하였다.

평가 지향 영역에서는 교육과정을 바탕으로 전형적인 수업 상황을 제시하고 이 상황에서 교사들이 지향하는 평가 사례를 응답하도록 하였다. 교육과정에서는 1단원 ‘역사 속의 과학 탐구’의 핵심 개념으로 ‘과학의 본성’과 ‘과학자의 탐구 방법’을 포함하고, 각 핵심 개념에 두 개씩 총 4개의 성취기준을 제시하였으며 각 성취기준에 대응하는 탐구 활동을 하나씩, 총 4가지 제시하고 있다. 이에 성취기준과 탐구 활동을 바탕으로 1단원에서 평가를 실행할 수 있는 전형적인 수업 상황으로 ‘자유 낙하와 수평으로 던진 물체의 운동 비교하기’(이하 ‘자유낙하’), ‘멘델레예프의 주기율표 만들기’(이하 ‘주기율표’), ‘파스퇴르의 생물 속생설 도출 과정 검토하기’(이하 ‘파스퇴르’), ‘지질 시대 동안 생물 대멸종의 원인과 그 후 변화 조사하기’(이하 ‘대멸종’) 4가지를 제시하였다. 교사들은 수업 상황을 묘사하는 각각의 제시문을 읽고 이 상황에서 실행할 수 있는 이상적인 평가 사례를 생각하여 평가 영역과 평가 방법을 중심으로 기술하였다(Figure 1).

3-1. 예시 상황 1

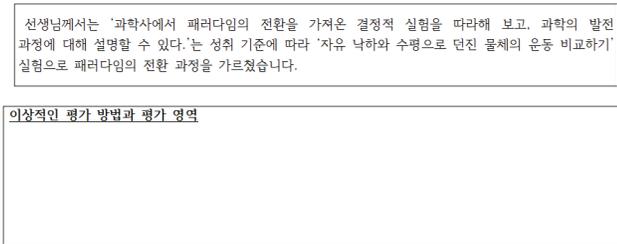


Figure 1. An example of the questionnaire for preferred cases of assessment

평가 실태는 최대 5개의 사례를 응답하도록 하였고 평가 지향은 4가지 상황별로 평가 사례를 하나씩 응답하도록 하였다. 설문지를 개발하는 과정에서 세미나를 실시하여 타당도를 검토하였다. 세미나에는 과학교육 전문가와 현직 교사, 과학교육 전공 대학원생 등 10명이 참여하였고, 세미나에서 논의된 내용을 수정하고 다시 검토하는 과정을 4회 이상 거쳤다. 완성한 설문지는 현직 과학 교사 2인을 대상으로 1회의 예비 설문을 실시하여 문항에 대한 이해도 등을 점검한 후 수정하였다.

3. 분석 방법

평가 지향을 묻는 설문에서는 수업 상황을 제시하고 상황별로 하나의 평가 사례를 응답하도록 하였으므로 모든 응답이 4개의 주제로 구분되었다. 그러나 평가 실태를 묻는 설문에서는 교사들이 자유롭게 평가 사례를 응답하였으므로 우선 교사들이 응답한 평가 사례를 주제에 따라 구분하였다. 평가 지향과 마찬가지로 교육과정을 기준으로 한 4가지 주제와 이외에 교사들의 응답 결과로부터 도출한 ‘귀납&연역’과 ‘기타’를 추가하여 6가지 주제로 평가 사례를 구분하였다. ‘귀납&연역’은 귀납적 탐구 방법과 연역적 탐구 방법을 통합해서 다른 수업에서 이루어진 평가이고, ‘기타’는 학생들이 자유롭게 주제를 정하는 자유 탐구 수업에서의 평가 사례, 과학탐구실험의 2, 3단원에

해당하는 주제를 이용하여 1단원의 핵심 개념을 가르치고 평가한 사례 등 앞서 설명한 5가지 주제에 포함되지 않는 평가 사례이다.

다음으로 평가 실태와 평가 지향의 사례를 평가 영역과 평가 방법의 두 가지 측면에서 분석하였다(Table 2). 먼저 평가 영역은 선행 연구(Baek et al., 2020; Kim & Kim, 2002)를 기반으로 지식, 역량, 태도로 하위 영역을 나누어 분석하였다. 지식 영역에는 교육과정에서 1단원의 핵심 개념으로 제시한 ‘과학의 본성’, ‘과학자의 탐구 방법’을 포함하였고, 교사들의 응답 결과에서 귀납적으로 도출된 ‘과학 개념’과 ‘과학사’를 추가하였다. 역량 영역에서는 2015 개정 교육과정의 과학과 핵심 역량 다섯 가지를 포함하되, 교사들의 응답 결과에서 나타나지 않은 ‘과학적 참여와 평생 학습 능력’을 제외하여 4개의 역량을 포함하였다. 각각의 역량이 구체적으로 어떤 기능을 포함하는지는 Yun et al.(2018)의 핵심 역량 하위요소 정의를 따랐다. 태도 영역에는 지식과 역량에 포함되지 않은 수업 참여도, 실험실 안전 수칙 준수 여부 등을 포함하였다. 평가 방법은 Noh et al.(2015)의 연구에서 Wang et al.(2010)의 분류 방법을 적용하여 만든 분석틀을 활용하였다. 평가 방법을 크게 측정, 수행, 비형식의 3가지 유형으로 분류한 뒤 각 유형에 해당하는 구체적인 평가 방법은 본 연구의 자료에서 귀납적으로 도출하여 선행 연구에서 나타나지 않았던 평가 방법도 추가하였다. 예를 들어 Noh et al.(2015)에서 사용한 분석틀에는 수행 유형에 발표 평가나 산출물 평가가 포함되지 않았으나, 과학탐구실험의 특성상 탐구 결과에 대한 발표를 평가하거나 학생들이 만든 입체 주기율표에 대한 평가 등이 다수 나타났다. 따라서 이러한 평가 방법을 각각 발표 평가, 산출물 평가로 수행 유형에 포함하였다.

Table 2. Assessment areas and methods

| 평가 영역 | | 평가 방법 | |
|-------|-------------|-------|----------|
| 지식 | 과학 개념 | 측정 | 선택형 평가 |
| | 과학의 본성 | | 서답형 평가 |
| | 과학자의 탐구 방법 | | 보고서 평가 |
| 역량 | 과학사 | 수행 | 논술형 평가 |
| | 과학적 사고력 | | 포트폴리오 평가 |
| | 과학적 탐구 능력 | | 토의 평가 |
| | 과학적 문제 해결력 | | 발표 평가 |
| 태도 | 과학적 의사소통 능력 | 비형식 | 산출물 평가 |
| | 태도 및 참여도 | | 관찰 평가 |
| | | | 면담 평가 |
| | | | 자기 평가 |
| | | | 동료 평가 |

평가 방법이나 평가 영역이 누락되어 분석이 어려운 사례나 한 참여자가 중복으로 참여하여 완전히 동일한 사례를 제출한 것 등은 분석 대상에서 제외하였다. 평가 지향을 묻는 문항에서 하나의 상황에 2개 이상의 평가 사례를 응답한 경우도 있었는데 이러한 경우 각각의 사례를 별도의 구분된 사례로 분석하였다. 평가 실태는 주제의 구분 없이 총 223개, 평가 지향은 주제별로 최대 74개의 사례를 분석 대상으로 하였다. 2인의 연구자가 이 사례들을 분석틀에 따라 각자 분석한 뒤 분석 결과를 협의하였다. 협의 과정에는 또 다른 연구자 1인을 추가하여 3인의 연구자가 참여했으며, 의견이 일치하지 않는

사례는 논의를 통하여 분류 기준을 수정하였다. 수정한 기준을 바탕으로 모든 사례를 분석한 결과가 일치할 때까지 분석을 반복하였다. 이렇게 분석한 결과를 바탕으로 각 평가 사례가 포함하는 평가 영역과 평가 방법의 빈도 그리고 백분율을 분석하였다. 백분율은 주제별로 전체 평가 사례의 빈도 중 해당 평가 영역이나 평가 방법을 포함한 사례가 차지하는 비율을 구하였다. 예를 들어 평가 실태에서 ‘자유낙하’에 해당하는 사례는 총 44개이고, 이 사례 중에서 평가 영역으로 과학적 탐구 능력을 포함한 사례는 28개였다. 따라서 이 주제에서 과학적 탐구 능력을 포함한 사례의 백분율은 63.6%로 계산하였다. 이때 하나의 평가 사례가 여러 평가 영역과 평가 방법을 포함하는 경우가 많았으므로, 각 주제에서 평가 영역과 평가 방법의 백분율 계는 100%를 넘었다.

III. 연구 결과 및 논의

1. 과학탐구실험 1단원에서의 평가 실태

연구 참여자 70명 중 3명은 1단원에 대한 학생들의 흥미가 낮고 정량적인 평가를 하기 어렵다는 이유로 1단원에서 평가를 실시하지 않았다. 67명의 교사가 실행한 223개의 평가 사례를 주제별로 분류한 결과는 Table 3과 같다.

교육과정에서 제시한 탐구 활동 4가지 중 ‘주기율표’와 관련된 수업에서 실행한 평가 사례가 50개로 가장 많았고, ‘자유낙하’ 44개, ‘대멸종’ 36개, ‘파스퇴르’ 31개의 순서로 많은 사례가 나타났다. 이때 ‘귀납&연역’이 연역적 탐구 방법을 다루는 ‘파스퇴르’와 귀납적 탐구 방법을 다루는 ‘대멸종’을 통합하여 실시한 평가라는 점을 고려하면 4가지 주제의 수업에서 비교적 고르게 평가가 이루어진 것으로 볼 수 있다. 한편 ‘기타’의 사례가 49개로 교육과정에서 제시한 주제로 실시한 평가 사례와 비슷한 수의 평가가 이루어졌다. 따라서 적지

않은 교사가 교과서나 교육과정이 제시하는 탐구 활동 외에 자율적으로 수업을 재구성하고 그 수업에서 평가를 실행하였음을 알 수 있다.

가. 평가 영역

평가 실태의 평가 영역을 분석한 결과를 Table 4에 나타내었다. 전체 223개의 평가 사례에서 총 292개의 평가 영역이 나타나 평가 사례당 평균 1.3개의 평가 영역을 포함하였다. 평가 영역의 세 가지 하위 영역을 비교해보면, 역량을 평가한 사례가 전체 사례 중 78.0%로 가장 많았고 지식과 태도를 평가한 사례는 각각 39.0%, 13.9%를 차지하였다. 이 결과는 교사들이 과학탐구실험의 교수학습 목표로 지식의 습득보다 역량의 함양을 중요하게 생각하고 있으며, 2015 개정 교육과정에서 강조하는 핵심 역량을 이 과목의 평가에서 적극적으로 반영하고 있다는 것을 보여준다.

평가 영역을 주제별로 살펴보면, ‘자유낙하’와 ‘주기율표’에서는 역량 영역 중 탐구 능력을 평가한 사례가 각각 63.6%, 56.0%로 가장 큰 비중을 차지하고 있었다. 탐구 능력을 평가 영역으로 한 사례는 다음의 교사 67 사례와 같이 다른 과학 과목에서도 보편적으로 실행되는 실험 수업에서의 평가와 유사하였다. 또한 두 주제에서는 사고력을 평가한 사례도 각각 20.5%, 26.0%로 적지 않았는데, 교사 46의 사례와 같이 실험을 수행하고 결과를 정리할 때 필요한 역량으로서 탐구 능력과 함께 사고력을 평가하는 경우가 많았다.

교사 67 사례: 자유낙하 운동과 수평 방향으로 던진 물체의 운동을 스마트폰으로 영상을 찍고, 트랙커를 이용하여 운동을 분석하여 활동지에 그래프를 완성하고, 운동을 설명하는 글쓰기를 하도록 하여 과학적 탐구 능력을 수행평가에 반영함.

교사 46 사례: 아리스토텔레스의 생각에서 모순점을 찾아내는 사고실험 및 미니 실험 진행 후 정리 및 실험을 계획하는 활동지를

Table 3. Actual assessment cases by topic

| | 과학의 본성 | | 과학자의 탐구 방법 | | 귀납 & 연역 | 기타 | 계 |
|----------|--------|------|------------|-----|---------|----|-----|
| | 자유낙하 | 주기율표 | 파스퇴르 | 대멸종 | | | |
| 평가 사례(개) | 44 | 50 | 31 | 36 | 13 | 49 | 223 |

Table 4. The assessment areas of actual cases of assessment by topic

단위: 개(%)

| | 과학의 본성 | | 과학자의 탐구 방법 | | 귀납&연역 (13개) | 기타 (49개) | 계(223개) | |
|----|------------|-----------|------------|-----------|----------------|-------------|-----------|------------|
| | 자유낙하(44개) | 주기율표(50개) | 파스퇴르(31개) | 대멸종(36개) | | | | |
| 지식 | 과학 개념 | 3(6.8) | 5(10.0) | 0 | 5(13.9) | 0 | 2(4.1) | 15(6.7) |
| | 과학의 본성 | 4(9.1) | 4(8.0) | 0 | 1(2.8) | 0 | 2(4.1) | 11(4.9) |
| | 과학자의 탐구 방법 | 0 | 0 | 19(61.3) | 9(25.0) | 13(100.0) | 13(26.5) | 54(24.2) |
| 역량 | 과학사 | 2(4.5) | 4(8.0) | 1(3.2) | 0 | 0 | 0 | 7(3.1) |
| | 사고력 | 9(20.5) | 13(26.0) | 3(9.7) | 2(5.6) | 1(7.7) | 8(16.3) | 36(16.1) |
| | 탐구 능력 | 28(63.6) | 28(56.0) | 8(25.8) | 9(25.0) | 0 | 18(36.7) | 91(40.8) |
| | 의사소통 능력 | 3(6.8) | 10(20.0) | 3(9.7) | 15(41.7) | 1(7.7) | 10(20.4) | 42(18.8) |
| | 문제 해결력 | 1(2.3) | 0 | 1(3.2) | 0 | 1(7.7) | 2(4.1) | 5(2.2) |
| 태도 | 태도 및 참여도 | 7(15.9) | 14(28.0) | 2(6.5) | 4(11.1) | 0 | 4(8.2) | 31(13.9) |
| | 계 | 57(129.5) | 78(156.0) | 37(119.4) | 45(125.0) | 16(123.1) | 59(120.4) | 292(130.9) |

완성하게 함으로써 과학적 사고력과 과학적 탐구 능력을 평가하였다.

반면 두 주제의 핵심 개념인 '과학의 본성'을 평가한 사례는 '자유낙하'에서 9.1%, '주기율표'에서 8.0%로 매우 적었고, 사례 수 또한 각각 4개에 그쳤다. 평가 영역으로 과학의 본성을 포함한 평가 사례의 대표적인 예시는 다음과 같다.

교사 10 사례: 아리스토텔레스, 갈릴레이, 뉴턴에 의하여 자유낙하에 대한 패러다임의 변화를 학습한 후 아리스토텔레스의 생각에서 논리적 모순이 무엇이었는지 쓰고 우주론의 역사 지문을 읽고 글에 나타난 과학의 본성을 찾아 근거를 설명하는 서술형 활동지를 작성하도록 함.

교사 57 사례: 현대의 주기율표가 만들어지기까지의 과정을 배우고 미래의 주기율표는 오늘날과 같을 것인지 달라질 것인지에 대해 자신의 견해를 논술하도록 하여 과학의 본성과 추론 능력을 평가함.

두 주제에서 핵심 개념인 '과학의 본성'을 평가 영역으로 한 사례가 매우 적었던 것은 부정적인 결과라고 할 수 있다. 교사들이 두 주제에서 '과학의 본성'을 평가하지 않은 가장 큰 이유는 다음 교사 A의 면담 내용과 같이 수업 자체가 실험을 중심으로 진행되어 '과학의 본성'을 학습할 기회를 제공하지 못하였기 때문으로 보인다. 교사 A는 교육과정에서 제시한 실험을 '과학의 발전 과정을 가르치려는 목적보다는 통합과학에서 물리 개념을 가르치려는 목적과 유사하게 진행하였다'고 설명하였다.

교사 A: 자유낙하 실험도 사실은 이제 통합과학이랑 같이하면서 그냥 연직 운동이랑 수평 운동을 합쳐서 이렇게 운동을 분석할 수 있다 약간 이런 식으로 수업을 했던 것 같아요.

연구자: 평소 물리 실험처럼 하신 느낌인가요?

교사 A: 약간 그 느낌이네요. 지금 보니까 이 취지랑은 그렇게 맞지 않았던... 맞아요. 지금 보니까 이거랑 그렇게 맞진 않네요. (교사 A의 면담)

수업에서 '과학의 본성'을 가르친 경우에도 교사 D의 면담 내용에서 볼 수 있는 바와 같이 교사들이 스스로 '과학의 본성'에 대한 평가 경험과 전문성이 부족하다고 느끼고 있으며 이 때문에 '과학의 본성'을 평가 영역에 포함하지 못한 것으로 보인다.

연구자: 수업에서 과학의 본성에 대해서 비교적 명시적으로 설명하신 것 같은데, 평가에는 반영하지 않으신 이유가 있으신가요?

교사 D: 근데 학생들이 그렇게 쓴 의견은 평가를 하기에는 뭐라고 쓰면 뭐라고 평가할지 제가 기준이 없기 때문에 그런 거는 활동 중에 답을 쓰라고 하진 않았죠. 평가하는 문항은 제가 뭔가 기준을 근거로 해야 되는데, 그냥 평가하기가 그래서 제가 강의를 하는 걸로만 끝내는 것도 있어요.

(교사 D의 면담)

'주기율표'에서는 태도 영역에 대한 평가가 28.0%로 다른 주제에 서보다 많이 나타났다. 이는 멘델레예프의 주기율표 만들기 활동이

정해진 답 없이 학생들에게 자유롭게 주기율표 만들어보도록 하는 경우가 많았고, 따라서 평가 또한 엄격한 채점 기준을 적용하기보다는 완성한 주기율표가 규칙성을 갖지만을 확인하여 수업 참여도를 평가한 사례가 많았기 때문으로 보인다. 다음은 이러한 평가 사례의 예시이다.

교사 30 사례: 원소 배열 방식의 규칙성을 활동지를 모둠별로 작성하여 제출함. 성적에 반영하지 않고 팀별 활동 참여도 등을 고려하여 생기부에 그 활동 내용만을 반영함.

교사 34 사례: 학생들에게 원소 카드를 제공하고 각자가 세운 기준에 따라 원소를 분류하는 활동을 통해 수업 참여도를 평가함.

'파스퇴르'와 '대멸종'에서는 두 주제의 핵심 개념인 '과학자의 탐구 방법'을 평가 영역으로 한 사례가 각각 61.3%, 25.0%로 나타났다. 앞서 '자유낙하'와 '주기율표'에서 '과학의 본성'을 평가한 사례가 10%에도 이르지 못했던 것과 비교하면 '파스퇴르'와 '대멸종'에서는 핵심 개념을 평가한 사례가 상대적으로 많았다. 따라서 교사들이 '과학의 본성'에 비해 '과학자의 탐구 방법'을 평가하는 것을 더욱 수월하게 여긴 것으로도 볼 수 있다. 특히 두 주제를 함께 다루는 '귀납&연역'에서는 13개의 모든 사례에서 '과학자의 탐구 방법'을 평가한 것으로 볼 때, 이러한 경향은 더욱 두드러진다고 할 수 있다. 그러나 '귀납&연역'에서 이루어진 평가 사례까지 고려하더라도 교육과정에서 핵심 개념으로 제시한 학습 영역을 평가한 사례의 절대적인 빈도가 높다고 보기는 어려우므로, 핵심 개념에 대한 교수와 평가가 더욱 적극적으로 이루어질 필요가 있을 것이다.

한편, '과학자의 탐구 방법'을 평가 영역으로 한 사례가 연역적 탐구 방법을 다루는 '파스퇴르'에서는 61.3%를 차지했지만, 귀납적 탐구 방법을 다루는 '대멸종'에서는 25.0%로 큰 차이가 있었다. 이는 교사들이 귀납적 탐구 방법보다 연역적 탐구 방법을 학습과 평가에서 더욱 중요하게 여긴 결과로 해석할 수 있다. 과학 탐구의 기본적인 방법으로서 연역적 탐구 방법만을 강조하는 모습은 교사 H의 면담 내용에서도 살펴볼 수 있다.

교사 H: 과학탐구실험을 체계화 시켜주셨던 저희 학교의 다른 선생님들께서도 다른 부분보다도 가설 설정을 하고 그 다음에 결론까지 어떻게 가야되는지 이 부분을 좀 강조해서 가르치시려고 하신 것 같고, 그런 부분들을 학생들한테 훈련시키는 데 조금 더 많은 시간을 할애하시는 걸로 커리큘럼을 짜셨던 것 같습니다.

(교사 H의 면담)

'대멸종'에서는 의사소통 능력을 평가한 비율이 41.7%로 다른 주제에서와 달리 탐구 능력을 평가한 비율보다 더 높았다. 이 주제의 탐구 활동은 주어진 여러 관찰 증거로부터 생물 대멸종의 원인을 귀납적으로 찾아보는 것으로, 다른 주제들과 달리 직접 실험을 수행하는 탐구 활동이 아니었다. 따라서 아래 교사 54의 사례처럼 조별 토의와 발표를 수업 활동으로 구성하고, 이에 상응하게 의사소통 능력을 평가 영역으로 한 사례가 많이 나타난 것으로 보인다.

교사 54 사례: 지질시대 생물 대멸종의 증거자료들을 카드로 만들어 모둠원들과 함께 읽고 자료를 바탕으로 모둠원들과 토

의하여 보고서를 작성하도록 하여 의사소통 능력과 과학적 사고력을 평가함.

‘기타’에서는 학생들이 스스로 주제를 찾아 탐구하는 자유 탐구를 진행한 사례가 많았다. 따라서 지식 영역 중 ‘과학자의 탐구 방법’(26.5%)과 역량 영역에서 탐구 능력(36.7%)이 평가의 중심이 되었다. 이러한 자유 탐구 형태의 평가 사례는 교사들이 교육과정에서 제시한 탐구 활동을 확장해서 과학탐구실험 과목 전반의 목표인 탐구 능력과 1단원의 핵심 개념인 ‘과학자의 탐구 방법’을 함께 평가하고자 한 것이므로 긍정적인 사례라고 할 수 있다.

교사 11 사례: 1단원 내용을 포함하여 자신이 하고 싶은 자유 탐구 주제를 선정 및 탐구하고 보고서를 제출하도록 함으로써 과학적 탐구 능력을 평가하고 그 내용을 세부능력 특기사항에 반영함.

마지막으로 지식 영역의 과학 개념이나 과학사를 평가한 사례가 모든 주제에서 일부 나타났다. 다음에 제시한 교사 50의 사례는 ‘대멸종’을 주제로 수업하였으나 핵심 개념 대신 교사들에게 익숙한 과학 개념을 평가한 것이며, 교사 5의 평가 사례는 여러 과학자가 제시한 원소 분류 방법의 변천 과정을 알고 있는지 과학사에 대한 지식을 평가한 것이다.

교사 50 사례: 지질 시대 생물의 화석을 제시하고, 화석으로 알 수 있는 지질 시대의 특징에 대해 객관식 문항으로 평가함.

교사 5 사례: 주기율표가 발전되는 과정에서 각 과학자들의 의견과 주기율표의 구성에 대해 평가하는 문제.

과학사를 평가한 사례가 나타난 것은 ‘역사 속의 과학 탐구’라는 제목에서 알 수 있는 바와 같이 1단원이 ‘역사적인 실험’으로 ‘과학의 본성’과 ‘과학자의 탐구 방법’을 학습하도록 구성된 것의 영향으로 보인다. 그러나 과학탐구실험 과목의 성격과 1단원의 핵심 개념, 성취

기준 등을 고려하면 지식 영역에서 과학 개념을 평가한 것은 바람직하다고 보기 어렵다고 할 수 있다.

나. 평가 방법

평가 실태의 전체 223개 평가 사례에서 총 251개, 사례당 평균 1.1개의 평가 방법이 나타났다(Table 5). 수행 방법이 90.1%로 가장 많이 나타났고, 측정 방법의 평가가 20.6%를 차지하였으며 비형식 방법은 1.8%로 거의 나타나지 않았다.

수행 방법의 평가 사례 중 가장 많이 활용된 방법은 54.7%를 차지한 보고서 평가였다. 이는 교사 E, F와의 면담 내용에서 살펴볼 수 있는 바와 같이 과학탐구실험 과목의 특성상 수업에서 직접 실험을 수행하는 경우가 많은데 교사들이 실험의 결과물로서 보고서를 가장 보편적인 것으로 인식했기 때문으로 보인다. 또한 평가 영역을 탐구 능력으로 한 사례가 많았기 때문에 탐구 능력을 평가하기 위한 방법으로 보고서 평가를 가장 많이 활용한 것이라고도 할 수 있다.

교사 E: 과학탐구실험 과목이니까 좀 실험이랑 보고서가 메인으로 들어와야 된다고 생각을 했어요. (교사 E의 면담)

교사 F: 그냥 실험을 하고 나면 보고서 쓰는 게 당연하니까 그래서 당연히 보고서를 쓰라고 했고, 거기에 이왕이면 이제 이런 과정이 들어가게끔 쓰라고 한 거예요. 실험을 하면 객관식 문제를 내는 게 더 힘든 것 같아요. (교사 F의 면담)

특히 ‘자유낙하’에서 보고서 평가의 비율이 75.0%로 매우 높았는데, 이는 교육과정에서 제시한 핵심 개념인 ‘과학의 본성’을 다루는 것보다는 교사 14의 사례처럼 ‘자유 낙하와 수평으로 던진 물체의 운동 비교하기’라는 실험 활동 자체를 중요하게 다룬 결과로 보인다.

Table 5. the assessment methods of actual cases of assessment by topic

단위: 개(%)

| | | 과학의 본성 | | 과학자의 탐구 방법 | | 귀납&연역(13개) | 기타(49개) | 계(223개) | |
|-----|-------|-----------|-----------|------------|-----------|------------|-----------|------------|-----------|
| | | 자유낙하(44개) | 주기율표(50개) | 파스퇴르(31개) | 대멸종(36개) | | | | |
| 측정 | 선택형 | 5(11.4) | 5(10.0) | 5(16.1) | 6(16.7) | 4(30.8) | 3(6.1) | 28(12.6) | 46(20.6) |
| | 서답형 | 0 | 0 | 7(22.6) | 2(5.6) | 2(15.4) | 7(14.3) | 18(8.1) | |
| 수행 | 보고서 | 33(75.0) | 30(60.0) | 14(45.2) | 19(52.8) | 3(23.1) | 23(46.9) | 122(54.7) | 201(90.1) |
| | 논술형 | 5(11.4) | 6(12.0) | 4(12.9) | 2(5.6) | 3(23.1) | 3(6.1) | 23(10.3) | |
| | 포트폴리오 | 0 | 0 | 1(3.2) | 0 | 0 | 7(14.3) | 8(3.6) | |
| | 토의 | 0 | 1(2.0) | 1(3.2) | 2(5.6) | 0 | 0 | 4(1.8) | |
| | 발표 | 4(9.1) | 7(14.0) | 3(9.7) | 7(19.4) | 1(7.7) | 8(16.3) | 30(13.5) | |
| | 산출물 | 0 | 7(14.0) | 0 | 4(11.1) | 0 | 3(6.1) | 14(6.3) | |
| | 관찰 | 2(4.5) | 0 | 1(3.2) | 1(2.8) | 0 | 0 | 4(1.8) | |
| 비형식 | 면담 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 4(1.8) |
| | 자기 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| | 동료 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| | 계 | 49(111.4) | 56(112.0) | 36(116.1) | 43(119.4) | 13(100.0) | 54(110.2) | 251(112.6) | |

교사 14 사례: 자유낙하하는 공과 수평 방향으로 던진 공의 운동을 촬영한 후, 단위 시간 동안 이동거리를 측정하여 시간에 대한 속도 그래프로 나타내 보게 함. 비상 교과서에 나와 있는 방법대로 포스트잇을 이용하여 좌표평면에 붙여보게 한 후, 2가지 경우에서 그래프의 특징을 서술해보도록 함. 자료 변환, 자료 분석 및 해석 능력을 평가하고, 그 내용을 세부능력 특기사항에 반영함.

보고서 평가 다음으로는 발표 평가가 13.5%, 논술형 평가가 10.3%로 많았고 일부 주제에서는 산출물 평가나 포트폴리오 평가도 나타났다. 먼저 발표 평가는 토의나 보고서 작성 등 조별 활동의 결과물을 발표하도록 하고 이를 평가하는 사례가 많았다. 아래 교사 55의 사례는 '주기율표'에서 조별로 작성한 주기율표를 발표하도록 하고 이를 평가한 사례이며, 교사 60의 사례는 '대멸종'에서 조별로 대멸종의 원인을 조사하고 보고서를 작성하여 이를 발표하도록 한 평가 사례이다.

교사 55 사례: 조별 활동으로 4절지에 학생들이 주기율표를 만들도록 한 후 각자 발표하게 하여 수업 참여도 및 의사소통 능력을 평가함.

교사 60 사례: 지질 시대 5대 생물 대멸종의 원인을 모둠별로 조사하고 전시장 관람구조로 발표하고 보고서를 작성하도록 하여 의사소통 능력을 신장시킬 수 있도록 하였고 발표 진행 과정을 관찰하여 세부능력특기사항에 반영함.

논술형 평가는 '과학사(과학자들의 논쟁)를 조사한 후 이에 반박하는 글쓰기 활동'을 하는 사례처럼 각 주제와 관련된 과학사를 다루는 평가에서 주로 활용되었는데, 1단원이 '역사적인 실험'을 중심으로 구성된 단원이므로 이에 어울리는 평가 방법으로 보인다. '자유낙하'와 '주기율표'에서는 총 11개의 논술형 평가 사례 중 5개가 '과학의 본성'을 평가 영역으로 하였으며, '파스퇴르', '대멸종', '귀납&연역'에서는 총 9개의 논술형 평가 사례 중 4개가 '과학자의 탐구 방법'을 평가 영역으로 하고 있었다. 즉 논술형 평가는 교육과정의 취지에 맞게 과학사를 중심으로 핵심 개념을 평가할 때 자주 활용된 방법이라고 할 수 있다. 다음은 '과학의 본성'과 '과학자의 탐구 방법'을 평가 영역으로 하고 논술형 평가 방법을 활용한 사례들이다.

교사 10 사례: 아리스토텔레스, 갈릴레이, 뉴턴에 의하여 자유낙하에 대한 패러다임의 변화를 학습한 후 아리스토텔레스의 생각에서 논리적 모순이 무엇이었는지 쓰고 우주론의 역사 지문을 읽고 글에 나타난 과학의 본성을 찾아 근거를 설명하는 논술형 활동지를 작성하도록 하여 평가함.

교사 47 사례: 귀납적 탐구 방법과 연역적 탐구 방법을 구분하고 각 탐구 방법의 대표적인 예시에 대해 과정을 분석하는 글 쓰기를 하도록 하여 평가함.

교사 38 사례: 주기율표의 발견 과정을 조사하고 과학자의 우연한 발견과 이때 나타난 과학의 본성에 대해 글을 써보는 방법을 통해 평가하고 세특에 반영함.

이처럼 논술형 평가는 교육과정의 취지에 맞게 이루어진 경우가

많았으므로 과학탐구실험의 1단원 수업에서 더욱 적극적으로 시도될 필요가 있을 것이다. 한편 교사들은 '과학의 본성'을 평가하기에 논술형 평가가 적절하다고 생각하였으나 교사 C와 같이 평가의 객관성에 문제가 생길 수 있다고 우려하기도 하였다. 논술형 평가를 실시할 때 생길 수 있는 평가의 객관성에 대한 우려는 앞서 교사 38의 사례와 같이 교사들이 '과학의 본성'에 대한 논술형 평가의 결과를 점수화된 성적으로 반영하기보다는 생활기록부에 정성적으로 반영하는 원인이 되었을 수도 있다.

교사 C: 이런 얘기(과학사)를 해주면 학생들이 흥미롭게 듣는데 그러고 나서 이제 '너희들이라면 어떤 것 같냐' 이런 식으로 의견을 물어보는 거죠. 그래서 이제 평가는 과학 철학적인 개념을 적용해서 글을 쓰게 해본 다음에 그 글 속에 어떤 의견이라도 말하도록, 주어진 글을 바탕으로 '네 생각은 뭐냐?' 이걸 말하라고 하는 논술 형식으로...

연구자: 그렇게 평가하는 게 어렵다고 생각하세요?

교사 C: 어렵진 않을 것 같아요. 근데 이게 논리적으로 추론이 되었느냐, 그거는 글쓰기의 영역에 해당할 것 같아서 함부로 채점하기는 어렵고, 이의 제기가 많이 들어올 수 있어서 디펜스 하기가 좀 어렵죠.

(교사 C의 면담)

산출물 평가는 '주기율표'에서 가장 많았는데, 이 방법은 교사 62의 사례처럼 학생들이 자신만의 규칙성을 반영하여 창의적으로 표현할 수 있도록 열린 과제를 제시한 경우에 주로 사용되었다.

교사 62 사례: 자신만의 입체 주기율표를 만들고, 이때 각 원소의 특징과 주기성을 보이는 특징들을 주기율표에 나타내었는지, 입체적인 구조를 잘 활용해 나타내었는지를 평가함.

포트폴리오 평가는 총 8개의 사례 중 '기타'에서 7개의 사례가 나타났다. 포트폴리오 평가는 일회성으로 이루어지는 평가가 아니라 장기간에 걸친 누적된 결과로 학생의 학습 과정을 평가하는 방법이다 (Kim, 2012). '기타'에서 많은 사례를 차지한 자유 탐구는 학생들이 주제를 정하여 탐구를 설계하고 실행하는 활동을 상대적으로 긴 시간 동안 하게 되므로 포트폴리오 평가가 적합했던 것으로 보인다. 다음 사례는 과제연구 활동에서의 탐구일지를 활용해 포트폴리오 평가가 이루어진 사례이다.

교사 4 사례: 그룹별 과제연구 활동을 수행하고 탐구일지를 작성하도록 하여, 과학적 탐구 능력과 과제 수행 능력(연역적 탐구 중 가설 설정 능력)을 평가하고 그 내용을 세부능력 특기사항에 반영함.

토의 평가의 비율은 1.8%로 매우 적었다. 평가 사례에 나타난 수업 과정을 살펴보면 수업에서 토의가 이루어진 사례가 적지 않았음에도 불구하고 토의를 평가에 반영한 경우는 거의 없었다. 특히 '대멸종'에서는 학생들이 대멸종의 근거를 직접 조사하고 조별 토의 과정을 거침으로써 가장 적절한 대멸종의 원인을 도출해내는 등 수업 과정에서 토의를 포함하는 경우가 많았으나 평가는 다음 사례와 같이 보고서만으로 이루어지는 경우가 대부분이었다.

교사 33 사례: 지질 시대 동안 생물의 대멸종의 원인을 추정할 수 있는 자료를 활용하여 조별 토론을 통해 생물의 대멸종 원인을 정리하여 보고서를 작성하여 과학적 탐구 능력과 수업 참여도를 평가함.

토의 평가는 학습과 평가가 동시에 이루어지는 평가 방법으로 학습 과정에 초점을 둔 평가이다(Back, 2000). 즉 학습을 위한 토의 과정에서 교사는 학생들의 참여도나 태도, 토의 내용에 대한 이해도 등을 평가하며 학생들에게 피드백을 제공할 수 있다. 따라서 과정 중심 평가가 강조되는 교육과정의 흐름을 고려하여 수업 중에 이루어지는 토의 활동을 교사들이 평가에도 적극적으로 활용할 수 있도록 하는 개선이 필요할 것이다.

다음으로 측정 방법의 평가를 살펴보면 전체에서 차지하는 비율은 20.6%로 적지 않았다. 서답형보다 선택형 평가가 많았으며 특히 ‘자유낙하’와 ‘주기율표’에서는 교사 50, 5의 사례처럼 모든 측정 방법 평가가 선택형 평가였다. 두 주제에서의 평가 영역에 대한 결과를 함께 고려하면 핵심 개념인 ‘과학의 본성’보다는 과학 개념과 과학사를 지식 영역으로 평가하기 위해 선택형 평가라는 전통적인 방법이 활용된 것으로 보인다.

교사 50 사례: 같은 높이에서 가만히 놓은 물체 A와 수평 방향으로 던진 물체 B가 수평면과 나란한 기준선 P를 동시에 지나는 모습을 나타낸 그림을 제시하고, A와 B의 자유낙하 운동에 대해 객관식 문항으로 평가함.

교사 5 사례: 주기율표가 발전되는 과정에서 각 과학자들의 의견과 주기율표의 구성에 대해 객관식 문항 평가

나머지 ‘파스퇴르’, ‘대멸종’, ‘귀납&연역’, ‘기타’에서는 선택형 평가와 함께 서답형 평가도 나타났다. 앞선 두 주제에서와 달리 이 네 가지 주제에서는 선택형 평가와 서답형 평가 모두 주제와 관련된 핵심 개념인 ‘과학자의 탐구 방법’을 평가할 때 사용된 경우가 많았다. 이는 귀납적 탐구 방법과 연역적 탐구 방법을 가르칠 때 탐구 방법의 과정이나 단계를 묻는 형태로 측정 가능한 문항을 만들기가 비교적 수월하기 때문으로 보인다. 서답형 평가는 특히 ‘파스퇴르’에서 많았는데, ‘파스퇴르 실험에 대한 그림을 주고 연역적 탐구 방법의 과정에 맞추어 쓰기’와 같이 연역적 탐구 방법에 대한 이해를 직접 서술하도록 하여 평가하는 경우가 대부분이었다. 마찬가지로 ‘대멸종’, ‘귀납&연역’에서도 서답형 평가는 교사 65의 사례처럼 귀납적 탐구 방법과 연역적 탐구 방법에 대한 이해를 서술하도록 한 사례였다. ‘기타’에서는 교사 3의 사례와 같이 새로운 과학사나 실험 주제를 제시하고 ‘과학자의 탐구 방법’에 대한 이해를 평가하는 사례가 많았다.

교사 65 사례: 귀납적 탐구 방법의 예시와 연역적 탐구 방법의 예시를 분석해 설명하고 두 방법의 차이점을 설명하는 서술형 문항으로 과학자의 탐구 방법을 평가해 수행평가 점수에 반영함.

교사 3 사례: 통합과학 중 생명과학 시간에 다룬 세포막을 통한 물질 이동 실험, 카탈레이스 실험(호소 실험) 등을 제시문으로 주고 연역적 탐구 과정의 단계, 실험군/대조군, 변인 등을 찾는 객관식/주관식 문항으로 과학적 탐구 능력을 평가함.

마지막으로 비형식 방법의 평가는 1.8%에 불과하여 일반적인 과학 수업에서의 평가 실태를 조사한 Noh *et al.*(2015)의 연구 결과와 마찬가지로 매우 적었다. 4개의 비형식 평가 사례는 모두 관찰 평가였으며, 교사 65의 사례처럼 평가 결과를 성적에 반영하기보다는 학생의 활동 과정 등을 관찰하고 이를 생활기록부의 세부능력 및 특기사항으로 반영하는 형태였다. 교사 G는 면담에서 “학생의 발언이 (점수로 나타내는 성적보다) 의미가 있으니까 그런 것들을 잘 관찰하고 기록했다가 생활기록부에 적어주는 정성적인 평가에 신경을 썼다.”고 응답하며 정성적인 평가 방법으로서 관찰 평가의 장점과 관찰 평가의 결과를 생활기록부에 기록하는 것의 긍정적인 측면을 언급하기도 하였다.

교사 65 사례: 모둠별로 실험을 수행하는 과정을 관찰함으로써 탐구 수행 능력 및 정의적 영역을 세부능력 특기사항에 반영함.

관찰 평가의 결과를 세부능력 및 특기사항으로 생활기록부에 기록하는 것은 점수 혹은 단계로 표현되는 성적보다 학생의 학습 과정과 결과를 더욱 구체적으로 반영할 수 있다는 점에서 바람직하다고 할 수 있다. 따라서 관찰 평가를 비롯한 비형식 방법의 평가를 더욱 적극적으로 활용할 필요가 있으며 정성적인 평가 결과를 반영할 수 있는 기회 또한 확대될 필요가 있을 것이다.

한편 면담 평가와 동료 평가, 자기 평가 사례는 하나도 없었다. 특히 동료 평가와 자기 평가에 주목해 볼 필요가 있는데, 과학탐구실험의 모든 교과서에서 차시 혹은 소단원을 마칠 때마다 자기 점검의 형태로 간단한 자기 평가를 포함하고 있으며 동료 평가 또한 마찬가지로 일부 교과서에서 포함하고 있음에도 사례가 하나도 나타나지 않았기 때문이다. 자기 평가와 동료 평가는 학생이 학습 과정뿐 아니라 평가 과정에도 직접 참여하도록 할 수 있는 구성주의적 평가 방법의 하나이며, 학생 수가 많을 때 교사가 관찰 평가나 면담 평가를 실시할 수 없다는 단점을 극복할 수 있는 유용한 평가 방법이다(Back, 2000; Cho, 2001). 특히 과학탐구실험에서 NOS를 가르치는 교사 중 일부는 NOS에 대한 견해가 개인의 주관적인 것이므로 채점에 부담을 느끼는 경우도 있었다(Kim *et al.*, 2020). 따라서 절대평가의 도입으로 평가의 객관성의 확보와 변별의 부담이 적은 과학탐구실험에서 자기 평가와 동료 평가를 더욱 적극적으로 실행할 필요가 있을 것이다.

2. 과학탐구실험 1단원에서의 평가 지향

가. 평가 영역

평가 지향 사례의 평가 영역을 분석한 결과를 Table 6에 나타내었다. 전체 272개의 평가 사례에서 총 371개, 사례당 평균 1.4개의 평가 영역이 나타났다. 세 가지 평가 영역이 차지한 비율은 역량, 지식, 태도의 순서로 각각 84.2%, 37.1%, 15.1%로 역량을 평가한 사례의 비율이 실태의 사례에서 78.0%를 차지한 것 비해 5% 포인트 이상 다소 증가한 것을 제외하면 지식과 태도 영역에서는 큰 차이가 없었다.

‘자유낙하’와 ‘주기율표’에서의 결과를 평가 실태의 결과와 비교하면 대부분의 평가 영역에서 5% 포인트 미만의 차이를 보여 실태와 지향에서의 결과가 크게 다르지 않았다. 5% 포인트 이상의 차이를 보인 결과를 중심으로 살펴보면 먼저 ‘주기율표’에서는 의사소통 능력을 평가한 사례가 실태에서는 20.0%였으나 지향에서는 27.5%로

증가하였다. 그리고 두 주제 모두에서 '과학의 본성'을 평가한 사례는 증가하였고, 사고력을 평가한 사례는 감소하였다. '과학의 본성'을 평가한 사례는 '자유낙하'에서 14.9%(실태 9.1%)로 증가하였고, '주기율표'에서는 26.1%(실태 8.0%)로 크게 증가하였다. 이처럼 평가 지향에서 '과학의 본성'을 평가한 사례가 증가한 것은 교육과정에서 제시한 두 주제의 핵심 개념을 평가한 사례가 늘어난 것이므로 긍정적인 결과로 해석할 수 있다. 반면 사고력을 평가한 사례는 '자유낙하'에서 13.5%(실태 20.5%)로, '주기율표'에서 17.4%(실태 26.0%)로 감소하였다. 사고력을 평가한 사례가 감소한 것은 '과학의 본성'을 평가한 사례가 상대적으로 증가한 것의 영향으로 보인다.

평가 지향에서 '과학의 본성'을 평가 영역에 포함한 사례가 증가한 것은 '과학의 본성'에 대한 평가를 지향하지만 실제로는 이를 평가하지 못한 교사들이 있었기 때문이다. 이들이 '과학의 본성'을 평가하지 못했던 이유로 가장 먼저 꼽은 것은 수업의 어려움이었다. 다시 말해 '과학의 본성'을 평가하기 위해서는 수업부터 '과학의 본성'을 중심으로 이루어져야 하지만 '과학의 본성'을 강조하면 수업이 지나치게 지식 중심으로 이루어지거나 수준이 높아지고 학생들의 흥미를 떨어뜨릴 수 있다는 우려 때문에 수업을 제대로 하지 못했고 이에 따라 평가도 '과학의 본성'을 중심으로 하지 못한 것이다. 다음 교사 E의 면담은 이러한 의견의 대표적인 사례이다.

교사 E: 일단은 이 기준으로 평가를 하려면 수업이 굉장히 좀 지식 중심으로 흘러가야 되고... 애들의 능력을 하나하나 평가를 해야되기 때문에 애들도 힘들고 저도 힘들죠. 그 다음에 이거는 좀 잘하는 학생들에게만 진짜로 의미가 있지, 못 따라오는 애들한테는 전혀 재미도 없고 의미도 없고 어떻게 하는지도 모르고... 약간 힘든 수업이 될 수 있기 때문에...

(교사 E의 면담)

한편 평가 지향에서 '과학의 본성'을 평가한 사례가 증가하긴 했지만, 절대적인 빈도는 '자유낙하'와 '주기율표'에서 각각 14.9%와 26.1%로 여전히 높지 않았다. 이러한 결과는 현실적인 제약에서 벗어나 교육과정의 성취기준에 따라 수업을 한 경우에도 교사들이 '과학의 본성'을 평가 영역으로 고려하지 않는 경우가 많았음을 의미한다. Kim et al.(2020)의 연구에서도 교사들은 NOS가 인지적 학습의 영역이 아니거나 NOS에 대한 견해가 개인의 주관이라는 등의 이유로

NOS를 평가의 대상으로 생각하지 않았다. 본 연구에 참여한 교사들도 유사한 견해를 보이는 경우가 있었는데, '과학의 본성'에 대한 평가를 지향하지 않았던 교사 A는 '과학의 본성'을 평가 영역에 포함해야 한다는 인식조차 하지 못하였다고 응답하였다.

교사 A: 과학의 본성을 평가하지 못하는 게 제 생각에는 이런 게 선생님들도 좀 많이 낳은 것 같아요. 일반적으로 봤을 때 그런 것들을 평가 영역으로 고려해야 된다는 인식 자체가 얼마 없을 것 같고, 과학탐구실험에서처럼 (과학의 본성을) 가르치려는 시도도 그 전에 많이 없었고요. 물화생지 각각의 내용을 평가하는 것에만 선생님들이 익숙했던 탓도 있지 않을까.

(교사 A의 면담)

이처럼 교사들이 '과학의 본성'을 평가 영역으로 고려하지 않는 것에는 다양한 이유가 있을 수 있으므로 구체적인 이유를 밝히고 교사들의 인식을 개선하기 위한 후속 연구가 필요할 것이다. 본 연구에서 면담에 참여한 일부 교사들을 중심으로 그 이유를 살펴보면 다음과 같다. 먼저 교사 F는 교육과정의 문제와 교사 자신의 전문성 부족을 이야기하였다. 즉, 교육과정에서 제시한 탐구 활동만으로는 '과학의 본성'의 가르치고자 하는 것인지, '과학의 본성'의 어떤 내용을 가르쳐야 하는지 알기 어려웠으며 이에 따라 '과학의 본성'의 어떤 내용을 어떻게 평가해야 하는지도 파악하기 어려웠다고 말하였다. 또한 '과학의 본성'을 평가하고자 하더라도 이에 대한 경험과 전문성이 부족하기 때문에 평가를 지향하지 않았다고 하였다.

교사 F: 저는 제가 수업을 할 때 이 두 실험이 과학적 발전 과정을 설명할 수 있게끔 하는 그런 내용이라는 걸 잘 몰랐어요. 2번(주기율표)도 과학의 본성에 대해서 정확히 어떤 부분을 가르치려고 하는 실험인지를 모르겠더라고요. 그래서 이거 두 개만 봤을 때는 과학의 본성에 어떤 부분을 평가를 해야 될지, 과학의 발전 과정을 어떻게 설명할 수 있는지 모르겠더라고요. 그러니까 이것만 가지고는 어떻게 설명을 해야 옳은 설명이라고 내가 평가를 해야 되고 이런 걸 잘 모르겠더라고요. (과학의 본성을 평가하려고 해도) 평가에서 어긋나는 사례, 그러니까 과학의 본성을 잘못 설명한다는 기준도 제가 본 적이 없고 평가를 하는 방법을 모르겠어서 함부로 어떻게 평가를 해야 될지 잘 모르겠어요.

(교사 F의 면담)

Table 6. The assessment areas of preferred cases of assessment by topic

단위: 개(%)

| | 과학의 본성 | | 과학자의 탐구 방법 | | 계(272개) | | |
|----|------------|-----------|------------|-----------|-----------|------------|-----------|
| | 자유낙하(74개) | 주기율표(69개) | 파스퇴르(66개) | 대멸종(63개) | | | |
| 지식 | 과학 개념 | 8(10.8) | 6(8.7) | 0 | 1(1.6) | 15(5.5) | 101(37.1) |
| | 과학의 본성 | 11(14.9) | 18(26.1) | 0 | 0 | 29(10.7) | |
| | 과학자의 탐구 방법 | 1(1.4) | 0 | 34(51.5) | 19(30.2) | 54(19.9) | |
| | 과학사 | 0 | 3(4.3) | 0 | 0 | 3(1.1) | |
| 역량 | 사고력 | 10(13.5) | 12(17.4) | 3(4.5) | 0 | 24(8.8) | 229(84.2) |
| | 탐구 능력 | 50(67.6) | 36(52.2) | 32(48.5) | 37(58.7) | 155(57.0) | |
| | 의사소통 능력 | 7(9.5) | 19(27.5) | 4(6.1) | 18(28.6) | 49(18.0) | |
| | 문제 해결력 | 1(1.4) | 0 | 0 | 0 | 1(0.4) | |
| 태도 | 태도 및 참여도 | 9(12.2) | 20(29.0) | 7(10.6) | 5(7.9) | 41(15.1) | |
| 계 | | 97(131.1) | 114(165.2) | 80(121.2) | 80(127.0) | 371(136.4) | |

교사 H의 면담 내용에서는 교육과정의 또 다른 문제가 드러나기도 하였다. 교사 H는 평가 지향에서 교육과정이 제시한 탐구 상황에 맞게 지향하는 평가 사례를 응답하기 위해 교육과정 문서를 참고하였다고 하였다. 즉, 교육과정의 ‘평가 방법 및 유의사항’에서 “뉴턴의 중력 실험과 멘델레예프의 실험 등을 재현하는 과정에서 작성한 실험 보고서를 통해 과학적 사고력을 평가할 수 있다.”는 부분(MOE, 2015)을 보고 과학적 사고력을 평가하겠다고 응답하였다. 교사 H는 이와 같은 설명을 하면서 교육과정조차 핵심 개념인 ‘과학의 본성’을 평가 영역으로 다루지 않고 있다는 사실을 지적하였다. 이처럼 교육과정에서도 성취기준과 핵심 개념에 따른 평가 영역을 안내하지 않고 있는 점은 개선이 필요하다고 할 수 있다.

다음으로 ‘파스퇴르’와 ‘대멸종’에서는 두 주제의 핵심 개념인 ‘과학자의 탐구 방법’이 평가 지향에서도 적지 않은 비중을 차지하고 있었다. 구체적으로 살펴보면 ‘파스퇴르’에서는 ‘과학자의 탐구 방법’을 평가한 사례가 51.5%로, 실태의 61.3%보다 비율은 다소 감소하였으나 여전히 과반을 차지하였으며 ‘대멸종’에서는 30.2%로 실태의 25.0%에 비해 5% 포인트 이상 증가하였다. 다만 실태에서는 교사들이 연역적 탐구 방법을 더 중요하게 생각하고 평가하는 경향이 있었지만, 지향에서는 귀납적 탐구 방법과 연역적 탐구 방법을 상대적으로 고르게 평가하려는 경향이 있었다.

두 주제의 평가 지향에서 나타난 또 다른 특징은 탐구 능력을 평가한 사례가 크게 증가한 것이다. 탐구 능력을 평가 영역에 포함한 사례는 ‘파스퇴르’에서 48.5%(실태 25.8%)로, ‘대멸종’에서 58.7%(실태 25.0%)로 증가하였다. 탐구 능력을 평가하는 사례가 증가한 것은 교사들이 ‘과학자의 탐구 방법’을 가르칠 때 지식으로서 ‘과학자의 탐구 방법’을 이해하는 것뿐 아니라 이를 실제 탐구에 적용하는 역량으로서 탐구 능력을 함양하는 것 또한 중요하게 생각하여 함께 평가한 것으로 해석할 수 있다.

교사 E: 성취기준이 직접적으로 ‘탐구 실험을 수행하고 방법을 설명할 수 있다’ 이렇게 되어 있으니까 어쨌든 수행을 해야 되잖아요. 그런데 교과서에 있는 거 그대로 하면은 애들이 수행을 직접

하는 게 좀 아닌 것 같아서 비슷한 탐구 자료를 제공하고 직접 수행하는 그 과정과 결과를 관찰해 가지고 평가를 하는 게 이상적이라고 생각을 한 거고, 그다음에 ‘탐구 방법을 설명할 수 있다’니까 설명할 수 있는지도 봐야 되고, 이상적으로 평가하려면 이렇게 해야 된다고 생각을 했어요.

(교사 E의 면담)

탐구 능력을 평가한 사례가 증가함에 따라 다른 영역을 평가한 사례는 상대적으로 감소하였다. ‘파스퇴르’에서는 사고력을 평가한 사례가 9.7%에서 4.5%로 감소하였으며 특히 ‘대멸종’에서 여러 평가 영역의 감소가 나타났다. 과학 개념, 사고력, 의사소통 능력을 평가한 사례가 각각 13.9%에서 1.6%로, 5.6%에서 0%로, 41.7%에서 28.6%로 감소하였다. 이 중 과학 개념을 평가하는 것은 과학탐구실험 과목의 성격과 1단원의 핵심 개념, 성취기준 등에 부합한다고 보기 어려우므로 과학 개념을 평가한 사례가 감소하고 탐구 능력과 ‘과학자의 탐구 방법’을 평가한 사례가 증가한 것은 긍정적인 결과로 해석할 수 있다. 교사 F는 ‘탐구 과정에서 평가하는 역량 안에는 자연스럽게 지식도 알게 되기 때문에 따로 과학 개념의 이해를 확인하는 평가가 필요하지 않다’고 설명하였다.

교사 F: 탐구 과정 속에 이런 식으로 발견됐다는 과학적 지식이 녹아 있으니까 평가는 탐구 과정을 중심으로 평가를 했어요. 평가는 안 했지만 당연히 지식이 포함되어 있으니까 지식도 같이 알게 되는 거라서...

(교사 F의 면담)

나. 평가 방법

평가 지향의 전체 274개의 평가 사례에서 총 332개, 사례당 평균 1.2개의 평가 방법이 나타났다(Table 7). 수행, 측정, 비형식의 순서로 많은 비율을 차지한 것은 평가 실태의 결과와 비슷한 경향성을 보였으나 각 영역이 차지한 비율은 다소 차이가 있었다. 즉 비형식 방법의 평가는 평가 지향에서도 5.1%로 평가 실태에서의 1.8%와 큰 차이가

Table 7. The assessment methods of preferred cases of assessment by topic

단위: 개(%)

| | | 과학의 본성 | | 과학자의 탐구 방법 | | 계(272개) | |
|-----|-------|-----------|-----------|------------|-----------|------------|------------|
| | | 자유낙하(74개) | 주기율표(70개) | 파스퇴르(67개) | 대멸종(63개) | | |
| 측정 | 선택형 | 1(1.4) | 2(2.9) | 5(7.5) | 2(3.2) | 10(3.6) | 32(11.7) |
| | 서답형 | 4(5.4) | 3(4.3) | 8(11.9) | 7(11.1) | 22(8.0) | |
| 수행 | 보고서 | 64(86.5) | 51(72.9) | 47(70.1) | 47(74.6) | 209(76.3) | 286(104.4) |
| | 논술형 | 3(4.1) | 3(5.7) | 1(1.5) | 3(4.8) | 10(3.6) | |
| | 포트폴리오 | 0 | 0 | 6(9.0) | 1(1.6) | 7(2.6) | |
| | 토의 | 0 | 1(1.4) | 1(1.5) | 1(1.6) | 3(1.1) | |
| | 발표 | 7(9.5) | 20(28.6) | 6(9.0) | 15(23.8) | 48(17.5) | |
| | 산출물 | 1(1.4) | 5(7.1) | 0 | 3(4.8) | 9(3.3) | |
| | 관찰 | 2(2.7) | 3(4.3) | 3(4.5) | 1(1.6) | 9(3.3) | |
| 비형식 | 면담 | 0 | 0 | 0 | 0 | 0 | 14(5.1) |
| | 자기 | 0 | 0 | 1(1.5) | 0 | 1(0.4) | |
| | 동료 | 1(1.4) | 2(2.9) | 1(1.5) | 0 | 4(1.5) | |
| 계 | | 83(112.2) | 91(130.0) | 79(117.9) | 80(127.0) | 332(121.2) | |

없었고 그 비율 또한 여전히 매우 적었다. 그러나 측정 방법의 평가는 11.7%로 실태에서 20.6%에 비해 감소하였으며 수행 방법의 평가가 104.4%로 90.1%보다 적지 않게 증가하였다.

측정 방법의 평가부터 구체적으로 살펴보면, 사례가 감소하긴 하였으나 이 감소량은 대부분 선택형 평가였다. 선택형 평가는 총 10개의 사례가 3.6%를 차지하여 실태와 비교하여 네 주제 모두에서 5% 포인트 이상 크게 감소하였다. 주제별로 살펴보면 '자유낙하'는 11.4%에서 1.4%로, '주기율표'는 10.0%에서 2.9%로, '파스퇴르'와 '대멸종'은 각각 16.1%에서 7.5%, 16.7%에서 3.2%로 감소하였다. 반면 서답형 평가는 8.0%를 차지하여 실태의 8.1%와 거의 유사하였다. 측정 방법의 평가 사례가 선택형 평가를 중심으로 감소한 결과는 서답형 평가가 학생들의 이해를 심층적으로 평가할 때 자주 활용된다는 점에서 바람직한 결과로 볼 수 있다(Noh *et al.*, 2015). 다음은 각각 사고력과 '과학의 본성'을 평가한 서답형 평가의 지향 사례이다.

교사 45 사례: 해당 실험 결과가 물체의 운동에 대한 과학자들의 해석(아리스토텔레스, 갈릴레이 등) 중 어떤 의견을 지지하는 것인지 서술형 문항으로 서술하도록 하여 과학적 증거와 이론을 토대로 논리적으로 추론하는 과학적 사고력을 평가함.

교사 59 사례: 주기율표를 통해 알 수 있는 과학의 본성을 설명하는 서술형 문항을 출제한다. 이를 통해 학생들의 탐구 능력과 과학의 본성에 대한 이해도를 평가한다.

측정 방법의 평가 사례가 선택형 평가를 중심으로 감소한 결과는 평가 지향에서 '과학의 본성'을 평가 영역으로 포함한 사례가 증가하면서 '과학의 본성' 평가에 어울리지 않는 선택형 평가가 감소했기 때문으로도 볼 수 있다. 교사 B는 '일반적인 과학 내용 지식의 경우 합의된 지식을 중심으로 교수와 평가가 이루어지기 때문에 지식을 측정하는 방법의 평가가 어울리지만 과학의 본성은 수업과 평가 또한 학생들이 생각을 자유롭게 표현하도록 하는 서답형 평가나 논술형 평가가 적절하다'고 설명하였다.

교사 B: 과학의 본성에 대해서는 다양한 의견이 존재하기 때문에... 아직까지는 과학의 본성은 다섯 가지야라고 딱 잘라서 얘기할 수 있나요? 그게 안 되기 때문에 그렇다고 생각을 하거든요. 교과서에서 과학 지식을 측정하는 건 굉장히 명시적이예요. 엄청나게 많은 과학자들이 합의를 한 지식을 중심으로 지식을 명시적으로 내가 가르쳐주잖아요. 그러고나서 이걸 이해했는지를 평가하는 방향으로 수업을 진행하거든요. 그런데 과학의 본성을 수업할 때는 그렇게 수업하지 않고, 과학자들이 이런 식으로 했는데 너는 이 과정을 통해서 과학의 본성이 무엇이라고 생각해? 라고 질문을 나중에 던지거든요. 만약에 내가 과학의 본성은 이러이러한 거라고 다섯 가지를 알려주고, 나중에 이걸 이해했는지를 물어보는 평가를 한다면 맞고 틀리고를 채점했을 것 같아요. 근데 그렇게 질문하지 않았기 때문에...

(교사 B의 면담)

수행 방법의 평가는 실태와 마찬가지로 보고서 평가가 가장 많았으며 평가 영역에서 탐구 능력을 평가한 사례의 대부분이 평가 방법에서는 보고서 평가를 활용한 것 역시 실태와 유사한 경향을 보였다.

전체 평가 사례에서의 비율은 76.3%로 실태의 54.7%보다 크게 증가하였으며 이러한 경향성은 네 주제 모두에서 동일하였다. '자유낙하'와 '주기율표'는 각각 75.0%에서 86.5%, 60.0%에서 72.9%로 증가하였으며 '파스퇴르'와 '대멸종' 또한 45.2%에서 70.1%, 52.8%에서 74.6%로 증가하였다.

발표 평가는 전체 평가 사례의 17.5%를 차지하여 수행 방법의 평가에서 두 번째로 높은 빈도로 활용되었다. 특히 발표 평가는 '주기율표'에서 28.6%로 실태의 14.0%에 비하여 크게 증가하였다. 이 주제와 관련된 교육과정의 성취기준은 "과학사에서 우연한 발견으로 이루어진 탐구 실험을 수행하고, 그 과정에서 발견되는 과학의 본성을 설명할 수 있다."로 제시되어 있으나 '과학의 본성'의 구체적인 예시는 언급되지 않고 있다. 따라서 교사 11의 사례처럼 학생들이 조별 활동을 통해 스스로 '과학의 본성'을 찾도록 하는 활동을 지향한 사례가 많았으며, 교사들은 이러한 수업에 어울리는 평가 방법으로 발표 평가를 지향하게 된 것으로 보인다.

교사 11 사례: 주기율표 만들기 활동을 통해 과학의 본성에 대해 모둠별로 토의하고 발표하여 평가함.

교사 9 사례: 서로 다른 친구들이 만든 주기율표의 규칙에 따른다면 다음에 올 원소의 특징은 어떤 것인지 예측하게 하는 활동을 한 후 이처럼 기준을 세워 정리하고 예측하는 것이 어떤 과학적 의의가 있을지에 대한 발표 자료를 만들고 발표시켜 평가한다.

포트폴리오 평가는 '파스퇴르' 주제에서 3.2%에서 9.0%로 5% 포인트 이상 증가하였다. 이 주제에서의 포트폴리오 평가는 자유 탐구 형태의 평가 사례가 많았다. 즉, 평가 실태에서는 '기타'에서 많이 이루어진 포트폴리오 평가 사례가 평가 지향에서는 연역적 탐구 방법을 다루는 '파스퇴르'에서도 조금씩 시도된 것으로 보인다. 포트폴리오 평가는 학생들의 탐구 과정을 평가할 수 있는 이상적인 방법이라고 지향에서 그 사례가 증가한 것은 긍정적이라고 할 수 있다.

수행 방법의 평가 사례가 평가 지향에서 전체적으로는 증가하였으나 오히려 평가 사례가 감소한 평가 방법도 있었다. 먼저 논술형 평가는 '대멸종'을 제외한 세 주제에서 감소하였다. '자유낙하'에서 4.1%(실태 11.4%), '주기율표'에서 5.7%(실태 12.0%), '파스퇴르'에서 1.5%(실태 12.9%)로 실태에 비해 5% 포인트 이상 감소하였다. 산출물 평가 또한 '주기율표'와 '대멸종'에서 5% 포인트 이상 감소하였는데 각각 7.1%(실태 14.0%), 4.8%(실태 11.1%)였다.

평가 지향에서 측정 방법의 평가 사례가 감소하고 수행 방법의 평가 사례가 증가한 것은 고무적인 결과라고 할 수 있다. 그러나 수행 방법의 평가 중 보고서 평가 사례만이 크게 증가하였고 논술형 평가와 산출물 평가는 일부 주제에서 오히려 감소하는 등 다양한 방법의 평가가 활용되지 못했다. 평가 지향에서도 교사들이 다양한 평가 방법을 활용하지 못한 것은 과학탐구실험의 1단원 수업에서 활용할 수 있는 다양한 평가 방법에 대한 교사들의 인식 자체가 부족한 것이라고 할 수 있다. 따라서 교사의 전문성을 중심으로 이러한 상황을 개선하기 위한 노력이 필요하다. 예컨대 1단원 수업 상황에서 논술형 평가나 토의 평가 등 다양한 평가 방법을 활용한 평가 사례를 개발하고 이를 교사들에게 제공할 수 있다. 이때 보고서 평가를 활용한 사례의

대부분이 탐구 능력을 평가한 것이었으므로 평가 방법 측면에서 보고서 평가 사례만이 크게 증가하였고 다양한 방법의 평가가 활용되지 못한 결과는 평가 영역에서 탐구 능력을 평가한 사례가 증가함에 따라 다른 영역을 평가한 사례가 감소한 결과와도 관련된다고 할 수 있다. 따라서 다양한 평가 방법만을 강조하는 것이 아니라 다양한 평가 영역과 함께 다양한 평가 방법을 소개할 필요가 있다.

보고서 평가만이 주로 활용된 이유는 교사의 전문성뿐 아니라 교육과정에서도 찾아볼 수도 있다. 앞서 언급한 교육과정의 ‘평가 방법 및 유의사항’에서는 “이 단원에서는 다양한 탐구 활동 수행에 대한 보고서 평가, 수행평가 등을 활용하여 학생의 성취 수준 및 학습 과정을 평가할 수 있다.”고 명시하고 있다(MOE, 2015, p. 115). 다시 말해 교육과정에서 1단원의 평가 방법을 제시할 때에도 이미 교사들이 많이 활용하고 있는 보고서 평가 이외에 다양한 종류의 평가 방법을 제시하지 못하고 있었으며 이것이 교사들의 평가 지향을 단순화시키는 원인이 되었을 수 있다. 따라서 교사들에게 익숙하지 않고 학교 현장에서 많이 활용되지 않는 다양한 평가 방법을 교육과정에서 적극적으로 제시하는 방향으로의 개선이 필요할 것이다.

마지막으로 비형식 방법의 평가는 5.1%로 여전히 적었다. 그러나 실태에서 관찰 평가가 없었던 ‘주기율표’에서 3개(4.3%)의 관찰 평가 사례가 나타난 것을 포함해 모든 주제에서 관찰 평가가 나타났다. 또한 실태에서는 등장하지 않았던 동료 평가와 자기 평가 사례가 각각 4개(1.5%)와 1개(0.4%)가 나타났다. 다음 교사 3의 사례는 모둠 활동을 한 수업에서 동료 평가를 활용한 사례이다. 평가 사례에 드러난 수업 과정을 살펴보면 모둠 활동을 이상적인 수업 활동으로 생각한 교사가 많았음에도 동료 평가를 지향한 사례가 많지 않았던 것은 아쉬운 결과라고 할 수 있다.

교사 3 사례: 학생들이 모둠별로 특색 있는 주기율표를 만들고 발표하도록 하여 의사소통 능력을 평가할 때 동료 평가를 활용할 수 있음.

면담 평가는 실태와 지향 모두에서 등장하지 않았다. 비형식 방법의 평가 결과를 정리하면, 동료 평가와 자기 평가를 지향한 사례가 나타난 것은 긍정적이었으나 일부 주제에서만 아주 적은 사례가 나타나는 데 그쳤으며 비형식 방법의 평가를 사용한 사례가 전체 평가 사례에서 차지하는 비율은 여전히 적었다. 이러한 연구 결과는 현실적인 어려움과 제약을 극복한 상황에서도 교사들이 비형식 방법의 평가를 지향하지 못한 것을 의미한다. 또한 Noh *et al.*(2015)의 연구를 참고해 결과를 해석해 볼 수도 있는데, 평가 지향을 조사하기 위한 가상의 사례로 구성주의적 평가를 실행하기 용이한 상황을 제시했다는 점에서 본 연구의 결과와 직접적인 비교는 어려울 수 있으나 Noh *et al.*(2015)의 연구에서는 본 연구의 결과와 달리 비형식 방법의 평가를 지향한 사례가 적지 않았다. 다시 말해 이러한 결과의 원인이 NOS를 명시한 1단원의 특성에 의한 것인지는 명확히 알기 어려우나 과학 교사들이 과학탐구실험의 1단원 수업에서 특히나 비형식 방법의 평가를 지향하지 못하였다고 할 수 있다. 그러므로 교사들이 비형식 방법의 평가를 지향할 수 있도록 교육할 필요가 있을 것이다. 이때 비형식 방법의 평가가 교사들에게 상대적으로 낯설다는 점에서 이를 활용하는 것이 다소 도전적일 수는 있으나 과학탐구실험의 절대평가 방식은 평가의 부담이 적다는 점(Byun *et al.*, 2019)에서 평가에서

중요한 현실적 어려움 중 하나인 객관성 확보와 변별의 문제 등을 쉽게 극복할 수 있게 한다. 따라서 과학탐구실험에서는 교사들이 비형식 방법의 평가를 지향할 수 있도록 한다면 실제 평가에서도 교사들이 이를 실천할 가능성이 더욱 클 것이다. 실제로 면담에 참여한 교사 F는 과학탐구실험의 절대평가 방식이 평가의 부담을 적게 지우므로 평가를 도전적으로 시도할 수 있는 기회를 제공한다고 설명하였다.

교사 F: 그게 학생들에게 도움이 되니까, 평가가 힘든 건 사실이지만, 이런 활동들이 학생들에게 도움이 되는 건 사실이고, 이런 활동들을 통해서 학생들의 사고가 발달하고 깨우치는 것도 맞는 거니까 이거를 평가를 하는 거죠. 근데 만약에 상대평가였으면 이거를 평가한다고 안 적었을 수도 있을 것 같은데. 절대평가기 때문에 아무래도 그 평가에 좀 부담이 좀 덜하잖아요 상대평가 보다는. 이런 평가를 한다고 하면은 그래도 한 마디라도 더 하니까 그냥 그런 생각이었던 것 같아요.

연구자: 수업 참여도도 높일 수 있다는 기대도 하신다는 건가요?

교사 F: 네. 그리고 다 점수 잘 주면 되니까 열심히 한 친구들은.

(교사 F의 면담)

IV. 결론 및 제언

이 연구에서는 과학탐구실험의 1단원 ‘역사 속의 과학 탐구’를 가르친 경험이 있는 교사를 대상으로 이 수업에서 실행한 평가 사례와 전형적인 수업 상황에서 지향하는 평가 사례를 조사하였다. 연구 결과, 이 단원에서 실행한 평가 사례는 교육과정에서 제시한 네 가지 주제에 비교적 고르게 분포되어 있었다. 평가 실태로 응답한 사례를 평가 영역 측면에서 분석한 결과, 역량을 평가한 사례가 가장 큰 비율을 차지하였고 역량 중에는 특히 탐구 능력을 평가 영역으로 한 사례가 주제 전반에서 많이 나타났다. ‘자유낙하’와 ‘주기율표’에서 핵심 개념인 ‘과학의 본성’을 평가 영역으로 한 사례는 매우 적었고, 또 다른 핵심 개념인 ‘과학자의 탐구 방법’을 평가한 사례는 ‘파스퇴르’에서만 절반을 조금 넘는 비율로 나타났다. 평가 방법은 수행 방법의 평가가 가장 큰 비중을 차지하였다. 수행 방법 중에는 보고서 평가가 가장 많이 활용되었고 발표 평가와 논술형 평가도 각각 10%정도를 차지하였다. 측정 방법의 평가에서는 서답형 평가보다 선택형 평가가 많았으며, 비형식 방법의 평가는 관찰 평가만 매우 적은 사례가 나타났다. 평가 지향의 사례를 분석한 결과 ‘자유낙하’와 ‘주기율표’에서 ‘과학의 본성’을 평가한 사례가 소폭 증가하였으나 절대적인 빈도는 여전히 높지 않았다. ‘파스퇴르’와 ‘대멸종’에서는 탐구 능력을 평가한 사례가 증가함에 따라 다른 영역을 평가한 사례는 상대적으로 감소하였다. 평가 방법 측면에서는 측정 방법의 평가가 선택형 평가를 중심으로 감소하였다. 수행 방법의 평가는 증가하였으나 보고서 평가만 크게 증가하여 다양한 방법의 평가가 활용되지 못했다. 또 비형식 방법의 평가는 실태에서 나타나지 않았던 동료 평가와 자기 평가가 나타나기도 하였다. 본 연구의 결과에서 나타난 평가 실태를 개선하기 위한 방안을 제안하면 다음과 같다.

먼저 학교 현장에서 NOS 수업이 적극적으로 이루어지도록 할 필요가 있다. 교사들이 평가 실태에서 핵심 개념을 평가하지 않았던 이유 중 하나는 수업 자체가 ‘과학의 본성’을 학습할 기회를 제공하지 못하였기 때문이었다. 또 일부 교사들은 ‘과학의 본성’에 대한 평가를

지향했음에도 '과학의 본성'을 가르치기 위한 수업을 실행하기 어려워 NOS에 대한 평가를 실천하지 못하였다. 따라서 NOS에 대한 평가가 적극적으로 이루어지기 위해서는 NOS에 대한 수업부터 제대로 이루어질 필요가 있다. NOS 수업은 NOS를 별도의 학습 영역으로 여기고 수업에서도 이를 직접 다루는 명시적인(explicit) 방법으로 이루어질 때 효과적이다(Kim *et al.*, 2008; Akerson & Abd-El-Khalick, 2003; Khishfe & Abd-El-Khalick, 2002). 따라서 효과적인 NOS 수업 방법이 적용된 구체적인 수업 사례를 소개하고 직접 수업을 설계해 볼 수 있는 기회를 제공하는 등의 방법으로 NOS 수업에 대한 교사들의 전문성을 향상할 필요가 있을 것이다. 이때 일부 교사들은 NOS 수업이 지식 전달이 중심이 되므로 학생들에게 어렵고 흥미롭지 않다고 생각하는 경우가 있었으므로 학생들이 흥미를 갖고 참여할 수 있는 다양한 NOS 활동을 소개할 필요도 있다. 그리고 교육과정에서 제시한 탐구 활동이 '과학의 본성'을 가르치기에 적합하지 않다고 생각하는 교사의 의견도 있었으므로 교육과정을 개선하기 위한 노력도 필요하다.

다음으로 NOS에 대한 과학교사의 평가 전문성을 향상하고자 할 때는 구체적으로 아래의 몇 가지 사항을 고려할 필요가 있다. 우선 교사들은 NOS를 명시적으로 가르친 경우에도 NOS를 평가하지 않거나 과학 개념이나 사고력 등 다른 영역만을 평가하는 경우가 있었으며 지향하는 평가 사례를 응답하라고 한 경우에도 NOS에 대한 평가를 지향하지 않았다. 따라서 교사들이 NOS가 명시적인 교수와 평가가 필요한 영역이라고 생각할 수 있도록 인식 개선이 필요하다. 그리고 NOS에 대한 평가를 지향하고 평가를 실천하고자 할 때처럼 NOS를 평가 영역으로 인식한 경우에도 평가 기준이 모호하다거나 평가 경험이 부족하다는 등의 이유로 NOS 평가 실천을 망설였다. 따라서 교사들이 NOS 평가에 익숙해질 수 있도록 다양한 평가 사례를 개발하고 이를 교과서에 포함하거나 별도의 자료로 제작하여 학교 현장에 배포할 필요가 있다. 그리고 이때 교사들이 쉽게 활용할 수 있는 평가 준거 또한 함께 포함할 필요가 있을 것이다. 평가 방법 측면에서는 수행 방법의 평가 중 토의 평가나 발표 평가, 논술형 평가 등 다양한 평가 방법이 활용될 수 있도록 해야 한다. 특히 교사들은 논술형 평가가 NOS 평가에 적합하다고 생각하였으며 실제 평가 사례도 교육과정의 취지에 맞게 이루어진 경우가 많았으므로 논술형 평가를 중심으로 NOS 평가를 확대할 필요가 있다. 또 토의 활동이 수업 중에 이루어지는 경우가 많았으므로 토의 활동을 평가에도 적극적으로 활용할 수 있도록 하는 개선이 필요할 것이다. 비형식 방법의 평가는 교사들이 지향조차 하지 않는 경우가 많았으므로 교사들이 비형식 방법의 평가를 지향하도록 교육할 필요가 있을 것이다. 마지막으로 수행 방법이나 비형식 방법의 평가를 활용함에 있어 측정 방법의 평가에 비해 공정성이나 객관성이 부족한 것을 우려하는 경우가 있었다. 그러나 과학탐구실험의 성적 산출 방식이 절대평가 체제로 전환되면서 학생을 변별해야 하는 교사들의 부담은 상대적으로 감소하였다고 할 수 있다(Byun *et al.*, 2019; Kwak, 2020). 또한 본 연구에서 일부 교사의 평가 사례와 면담 내용에서 드러난 바와 같이 생활기록부에 정성적으로 반영하는 방식으로도 수행 방법이나 비형식 방법의 평가를 활용할 수 있을 것이다.

NOS 평가의 개선을 위해서는 교육과정 해설 또한 개선할 필요가

있다. 현재 과학탐구실험 교육과정에서는 1단원의 내용 체계와 성취 기준에서 모두 NOS를 강조하고 있음에도 불구하고, '평가 방법 및 유의사항'에서는 NOS 평가와 관련된 어떠한 안내도 하지 않고 있다. 그러므로 교육과정에서 '탐구 활동', '성취기준 해설', '교수·학습 방법 및 유의사항', '평가 방법 및 유의사항' 등이 일관된 방향성을 가질 수 있도록 개선할 필요가 있다. 그리고 평가 방법 측면에서도 다양한 평가 방법을 제시하지 않고 있었으므로 교사들에게 이미 익숙한 수행평가와 보고서를 넘어서서 2015 개정 교육과정에서 강조하는 과정 중심 평가에 어울리는 더욱 다양한 평가 방법을 제안할 수 있어야 한다.

이 연구는 서술형 문항으로 구성된 설문을 이용하여 과학교사의 평가 실태와 지향을 정량적으로 조사하였다. 그러므로 과학교사의 평가 실태와 전문성을 심층적으로 조사하는 연구가 필요하다. 예를 들어 과학교사가 평가에 활용한 구체적인 문항이나 활동지 등을 수집하고 이와 관련된 면담을 실시하여 평가 실태를 구체적으로 조사할 수 있을 것이다. 또한 과학교사의 평가 전문성을 더욱 심층적으로 조사하기 위해서는 교사가 평가를 계획하는 단계부터 평가를 실행하고 평가 결과를 평정하여 성적에 반영하는 등 평가 전반의 과정을 관찰하고 분석할 수 있을 것이다. 그리고 본 연구에서는 교사들의 평가 지향을 조사하기 위해 교육과정을 바탕으로 구성된 전형적인 수업 상황을 제시하고 이 상황에서의 평가 사례를 응답하도록 하였다. 그러나 이러한 연구 설계가 다양한 평가 사례의 응답을 오히려 제한하였을 수도 있다. 따라서 교육과정의 내용 체계와 성취기준만을 제시한 후 교사들이 수업을 직접 구성하도록 하고 이 상황에서 지향하는 평가를 조사하는 방법도 생각해 볼 수 있을 것이다.

국문요약

이 연구에서는 과학탐구실험 1단원 '역사 속의 과학 탐구'에서 과학교사들이 실행한 평가 사례와 교육과정에 따라 구성된 전형적인 가상의 상황에서 지향하는 평가 사례를 조사하였다. 서술형 문항으로 구성된 설문지를 개발하여 '역사 속의 과학 탐구'를 가르친 경험이 있는 70명의 과학교사들을 대상으로 설문을 실시하였다. 설문에 참여한 교사 중 8명과 면담을 실시하였다. 교사들이 응답한 평가 사례를 평가 영역과 평가 방법의 측면에서 분석하였고, 평가 실태와 평가 지향의 결과를 비교하였다. 연구 결과, 평가 실태에서 평가 영역은 탐구 능력이 가장 많았고 교육과정에서 제시한 핵심 개념인 '과학의 본성'과 '과학자의 탐구 방법'을 평가한 사례는 적었다. 평가 방법은 보고서에 크게 치우치고 다양한 평가 방법이 활용되지 못했다. 평가 지향에서는 핵심 개념을 평가한 사례가 다소 증가하였으나 절대적인 빈도는 여전히 적은 수준에 머물렀다. 평가 지향의 평가 방법은 측정 방법의 평가가 줄고 수행 방법의 평가가 증가하였고 실태에서는 나타나지 않았던 비형식 방법의 평가가 나타나기도 하였으나 여전히 다양한 평가 방법이 활용되지 못하였다. 면담에 참여한 교사들의 의견을 바탕으로 설문 결과에 대한 원인을 분석하였다. 연구 결과를 바탕으로 과학탐구실험에서 NOS를 적극적으로 할 수 있는 방안을 논의하였다.

주제어 : 2015 개정 교육과정, 과학탐구실험, 과학의 본성(NOS), 평가

References

- Akerson, V. L., & Abd-El-Khalick, F. (2003). Teaching elements of nature of science: A yearlong case study of a fourth-grade teacher. *Journal of Research in Science Teaching*, 40(10), 1025-1049.
- American Association for the Advancement of Science (AAAS) (1993). *Benchmarks for science literacy: A project 2061 report*. New York, NY: Oxford University Press.
- Baek, J., Byun, T., Lee, D., & Shim, H.-P. (2020). An investigation on the assessment tool and status of assessment in the 'Scientific Inquiry Experiment' of the 2015 Revised Curriculum. *Journal of the Korean Association for Science Education*, 40(5), 515~529.
- Baek, S.-G. (2000). *Principles of performance assessment*. Seoul: Kyoyookbook.
- Bell, R. L., Lederman, N. G., & Abd-El-Khalick, F. (2000). Developing and acting upon one's conception of the nature of science: A follow-up study. *Journal of Research in Science Teaching*, 37(6), 563-581.
- 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.
- Cho, H.-M. (2001). An assessment tool from the view point of constructivism. *The Journal of Education*, 18, 183-197.
- Hanuscin, D. L. (2013). Critical incidents in the development of pedagogical content knowledge for teaching the nature of science: A prospective elementary teacher's journey. *Journal of Science Teacher Education*, 24(6), 933-956.
- Hanuscin, D. L., Lee, M. H., & Akerson, V. L. (2011). Elementary teachers' pedagogical content knowledge for teaching the nature of science. *Science Education*, 95(1), 145-167.
- Khishfe, R., & Abd-El-Khalick, F. (2002). Influence of explicit and reflective versus implicit inquiry-oriented instruction on sixth graders' views of nature of science. *Journal of Research in Science Teaching: The Official Journal of the National Association for Research in Science Teaching*, 39(7), 551-578.
- Kim, C.-J. (2012). *Portfolio instruction and portfolio assessment*. Seoul: Kyoyookbook.
- Kim, K., Noh, J.-a., Seo, I., & Noh, T. (2008). The effects of explicit and reflective instruction about nature of science using episodes from the history of science in 'Composition of Material' unit of middle school science. *Journal of the Korean Association for Science Education*, 28(1), 89-99.
- Kim, K.-M., & Kim S.-W. (2002). A study of weight of assessment domains in science education focused on the teacher's view points. *Journal of the Korean Association for Science Education*, 22(3), 540-549.
- Kim, M., Shin, H., & Noh, T. (2020). An exploration of science teachers' NOS-PCK: Focus on Science Inquiry Experiment. *Journal of the Korean Association for Science Education*, 40(4), 399-413.
- 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, N. G., Abd-El-Khalick, F., Bell, R. L., & Schwartz, R. S. (2002). Views of Nature of Science Questionnaire: Toward Valid and Meaningful Assessment of Learners' Conceptions of Nature of Science. *Journal of Research in Science Teaching*, 39(6), 497-521.
- 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.
- McComas, W. F., & Nouri, N. (2016). The nature of science and the next generation science standards: Analysis and critique. *Journal of Science Teacher Education*, 27(5), 555-576.
- Ministry of Education(MOE). (2015). *2015 Revised Science National Curriculum*. Seoul: Ministry of Education.
- National Research Council (NRC) (2011). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. Washington, DC: National Academies Press
- Noh, T., Lee, J., Kang, S., & Kang, H. (2015). Secondary school science teachers' actual and preferred types of assessment. *Journal of the Korean Association for Science Education*, 35(4), 725-733.
- OECD. (2017). *Pisa 2015 assessment and analytical framework: Science, reading, mathematics, financial literacy and collaborative problem solving*. OECD Publishing.
- Shim, B. K., & Yoo, M. H. (2020). An analysis of science core competencies reflected in the performance assessment of 2015 revised curriculum Integrated Science and Scientific Inquiry Experiments. *School Science Journal*, 14(4), 481-500.
- Song, J., Kang, S., Kwak, Y., Kim, D., Kim, S., Na, J., ... Joung, Y. (2019). Contents and features of 'Korean Science Education Standards (KSES)' for the next generation. *Journal of the Korean Association for Science Education*, 39(3), 465-478.
- Supprakob, S., Faikhamta, C., & Suwanruji, P. (2016). Using the lens of pedagogical content knowledge for teaching the nature of science to portray novice chemistry teachers' transforming NOS in early years of teaching profession. *Chemistry Education Research and Practice*, 17(4), 1067-1080.
- Thomas, L., Deaudelin, C., Desjardins, J., & Dezutter, O. (2011). Elementary teachers' formative evaluation practices in an era of curricular reform in Quebec, Canada. *Assessment in Education: Principles, Policy & Practice*, 18(4), 381-398.
- Wang, J.-R., Kao, H.-L., & Lin, S.-W. (2010). Preservice teachers' initial conceptions about assessment of science learning: The coherence with their views of learning science. *Teaching and Teacher Education*, 26, 522-529.
- 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

김민환(서울대학교 교육종합연구원 객원연구원)
박다혜(서울대학교 학생)
노태희(서울대학교 교수)