

수행평가에 대한 고등학생의 인식: 화학 교사의 수행평가를 중심으로

김현정 · 김성기^{†,*}

공주대학교 화학교육과

[†]한국교육과정평가원

(접수 2022. 7. 8; 게재확정 2022. 9. 12)

High School Students' Perception of Performance Assessment: Focusing on the Performance Assessment of Chemistry Teacher

Hyunjung Kim and Sungki Kim^{†,*}

Department of Chemistry Education, Kongju National University, Chungnam 32588, Korea.

[†]Korea Institute for Curriculum and Evaluation, Chungbuk 27873, Korea. *E-mail: mcarey2000@kice.re.kr

(Received July 8, 2022; Accepted September 12, 2022)

요 약. 이 연구는 설명적 순차 설계를 이용하여 화학 교사의 수업에서 수행되는 8가지 수행평가 방식에 대한 고등학교 학생들의 인식을 알아보았다. 이를 위해 10개 고등학교의 439명 학생을 대상으로 설문조사를 진행하였으며, 이후 5개 학교의 28명 학생을 대상으로 면담을 진행하여 설문 결과를 해석하였다. 연구 결과는 다음과 같다. 첫째, 학생들은 다양한 수행평가 방식보다는 편중된 평가 방식을 경험하고 있었다. 둘째, 학생들은 가장 재미있게 느꼈던 수행평가 방식에서는 직관적 확인 가능 여부, 진로와의 관계, 내적 성장 등을 언급하였으며, 어려움을 느꼈던 수행평가 방식에서는 준비도와 불명확한 평가 기준을 언급하였다. 셋째, 학생들은 화학 과목과 관련이 적은 개인의 능력에 영향을 받을 수 있다고 생각되거나 평가 기준이 상대적으로 불명확하다고 생각할수록 적절한 수행평가 방식이 아니라고 인식하고 있었다.

주제어: 수행평가, 인식, 화학 교사, 고등학생

ABSTRACT. This study investigated the perceptions of high school students for 8 performance assessment methods in a chemistry teacher's class by using an explanatory sequential design. To this end, a survey was conducted with 439 students from 10 high schools, and after that, 28 students from 5 schools were interviewed to interpret the survey results. The findings were as follows: First, students were experiencing a biased assessment method rather than various performance methods. Second, intuitive confirmation, relevance to a career path, and internal growth were mentioned in the performance assessment method that felt fun, and readiness and unclear evaluation criteria were mentioned in the performance assessment method that felt difficult. Third, the students recognized that it was not an appropriate performance method if the assessment result was affected by individual abilities less relevant to the chemistry subject or the assessment criteria were unclear.

Key words: Performance assessment, Perception, Chemistry teacher, High school student

서 론

학교에서 실행되는 평가는 학생의 학습을 진단하고 교사에게 교수의 방향을 제시한다. 최근 평가 패러다임이 학습 결과의 측정을 위한 평가에서 학습을 위한 평가로 옮겨감에 따라 결과 중심 평가에서 과정 중심 평가로의 변화가 강조되고 있다.^{1,2} 기존의 결과 중심 평가에서는 학습의 결과로 습득된 지식의 유무를 평가하였다면, 과정 중심 평가는 주어진 과제를 학생이 어떠한 문제 해결 과정을 통해 해답을 도출하는지에 대한 과정 자체와 그 안

에서 보이는 학생의 성장에 관심을 둔다. 이러한 과정에 대한 평가는 미래 사회에서 요구되는 역량을 기르는 교육과 밀접한 연관이 있어, 역량 함양 과정을 측정하는 평가의 실행과 관련하여 여러 연구가 수행되었다.³⁻⁵ 2015 개정 교육과정 총론에서는 '학습의 결과뿐 아니라 학생의 과정을 평가하여 모든 학생이 교육 목표에 성공적으로 도달할 수 있도록 한다.'라고 기술함으로써 교육과정에서 공식적으로 과정 중심 평가를 더욱 강조하였으며,⁶ 2022 개정 교육과정에서는 고교학점제 및 개정 교육과정에 부합하는 성장 중심 평가 체제의 구축을 목표로 하고 있다.⁷

한편 결과 중심의 평가에 대한 비판은 이전부터 있었으며, 이에 대한 대안적 평가로 수행평가가 현장에서 활용되고 있다. 수행평가는 주어진 과제를 학생이 해결하는 과정뿐만 아니라 해결 결과물인 산출물도 평가한다.⁸ 이러한 측면에서 과정 중심 평가와 수행평가는 유사하지만 동일한 개념으로 볼 수는 없다.⁹ 그럼에도 다른 평가에 비해 과정 중심 평가가 강조하는 측면이 수행평가와 일치하는 면이 있어 수행평가는 학교 현장에서 과정 중심 평가로 많이 활용되고 있다. 따라서 학교에서 과정 중심 평가가 자리 잡기 위해서는 수행평가가 보다 활성화될 필요가 있으며, 성취평가제의 확대 도입에 따라 도입되는 미수 예방을 위하여 수행평가를 통한 학생들의 학습 도달도 점검의 필요성이 강조되고 있다.¹⁰

수행평가와 관련된 기존의 선행연구를 고찰해보면 먼저, 수행평가가 기존의 결과 중심의 평가에서 벗어나 과정도 평가함에 따른 효과를 알아보는 연구로, 수행평가를 통해 학생들에게 더 실제적인 과제를 제공하고 수행하는 과정을 평가함에 따라 인지적 영역뿐만 아니라 정의적 영역에서도 효과가 있음을 보고하고 있다.^{11,12} 수행평가는 교사가 방식을 자유롭게 선택하여 진행하므로 교사의 인식에 관한 연구도 진행되었는데, 수행평가가 도입된 지 10년이 지났음에도 수행평가에 대한 교사의 인식 수준이 낮고 수행평가의 활용이 제대로 이루어지지 않음을 보고하고 있다.^{13,14}

학생을 대상으로 수행평가에 대한 인식을 알아본 연구는 다른 연구에 비해 비교적 적었다. 수학 교과에서 진행된 학생들의 수행평가에 관한 인식 연구 결과, 학생들은 수행평가를 다른 평가에 비해 타당성과 공정성이 결여된다고 인식하는 비율이 높게 나타났으며, 학생들은 입시와 직접적으로 관련성이 높은 수행평가 또는 실생활과 관련성이 높은 수행평가를 원하는 것으로 나타났다.¹⁵ 사회 교과에서 수행된 연구에서는 학생들이 수행평가에 대한 가치를 높게 인식하고 있었으나, 객관적인 평가 척도와 결과물에 대한 환류가 필요함을 언급하였다.¹⁶ 영어 교과에서 수행된 연구에서는 교사와 학생의 수행평가에 대한 인식을 조사하고 이를 비교하였는데, 교사들은 수행평가로 선호하는 유형이 몇 가지로 집중되는 데 반해, 학생들은 상대적으로 다양한 수행평가 유형을 선호하는 것으로 보고하였다.^{17,18} 과학 교과에서 수행평가 관련 학생의 인식을 살펴보았던 연구에서는 수행평가 과정에서 갖는 학생의 어려움을 보고하였다.¹⁹ 이처럼 기존의 수행평가와 관련한 학생의 인식을 알아본 연구는 다양한 교과에서 이루어졌으나, 전반적인 수행평가의 횟수, 효과, 어려움 등의 인식을 살펴보고 있어 과학 교과의 수행평가에 대한 구체적인 방안 도출에는 한계가 있었다.

또한, 학교에서 결과 중심 평가로 주로 지필 평가 방식을

활용하지만, 수행평가의 경우 다양한 방식의 평가가 가능하다. 측정학적 측면에서 같은 내용이어도 어떠한 방식으로 학생의 능력을 측정하여 평가하느냐에 따라 학생의 평가 결과는 달라질 수 있다.²⁰ 그렇기 때문에 측정하고자 하는 구인을 제대로 측정하기 위해서는 다양한 방법을 활용하여 측정하는 것이 요구된다. 유사하게 수행평가도 학생들이 자신의 능력을 최대한 발휘할 기회를 제공할 수 있도록 다양한 수행평가 방식이 제공될 필요가 있으나, 수행평가 방식의 선택은 해당 과목의 교사에 의해 결정되므로 교사는 수행평가 방식에 대한 학생들의 인식과 요구를 이해할 필요가 있으며, 이를 토대로 수업에 활용할 수행평가 방식을 결정할 필요가 있다. 하지만 다양한 수행평가 방식에 대한 학생의 인식과 요구에 관한 연구는 다른 수행평가 관련 연구에 비해 미미한 실정이다.

따라서 본 연구는 수행평가에서 상대적으로 덜 고려된 학생의 인식과 요구를 파악하기 위하여 수행평가 방식에 초점을 두고자 한다. 이 연구의 수행평가 방식에 따른 학생의 인식은 향후 화학 교사들이 수업에서 수행평가를 활용하는 방향을 제시해 줄 수 있을 것으로 기대된다. 이 연구의 연구 문제는 다음과 같다.

첫째, 화학 교사의 수업 시간에 학생들이 경험하였다고 인식한 수행평가 방식과 빈도는 어떠한가?

둘째, 학생들이 어떠한 수행평가 방식에서 재미 또는 어려움을 느끼는가?

셋째, 학생들이 수행평가 방식 중 적절하다고 인식되는 수행평가 방식은 무엇이며, 왜 그러한 인식을 하는가?

연구 방법

연구 설계

이 연구는 양적 연구 방법과 질적 연구 방법을 통합한 방법을 통해 화학 교사가 수업 시간에 진행하는 수행평가 방식에 대한 학생들의 인식을 분석하고 해석하고자 하였다. 일반적으로 혼합연구 방법은 단일한 연구 방법에 의존하지 않기 때문에 연구 결과를 보다 타당성 있고, 심도 있게 해석할 수 있다는 장점이 있다.^{21,22} 혼합연구는 그 목적에 따라 크게 수렴적 평행 설계, 탐색적 순차 설계, 설명적 순차 설계로 나뉜다.²³ 이 연구는 화학 교사가 진행하는 수업에서 여러 수행평가 방식에 대한 학생들의 인식을 분석하기 위해 먼저 설문지에 기반하여 항목별로 수행평가 방식에 대한 빈도를 조사하였다. 이후 얻어진 양적 결과에 대한 해석을 설문에 참여한 초점집단을 중심으로 시도하였다.

학생들이 화학 교사의 수업을 1년 경험한 후 인식을 조사하기 위하여 2021년 12월에 학교별로 기말고사를 진행한 후 양적 연구 설문을 진행하였으며, 질적 연구는 양적

연구의 분석이 완료된 후 2022년 3~4월 사이에 학교별로 진행되었다.

연구 대상

이 연구는 혼합연구 방법을 채택하였기 때문에 연구 대상이 2가지로 나뉜다. 먼저 학생들의 수행평가 방식에 대한 인식을 정량적으로 알아보기 위해 17개 시도 중 시도가 중복되지 않게 10개의 지역에서 각 1개의 고등학교를 편의 표집하였다. 표집된 학교별로 화학 교사 1인이 선정되었으며, 해당 화학 교사의 수업에서 시행되는 수행평가에 기반하여 설문에 응답하도록 학생들에게 요청하였다. 표집된 학교의 화학 교사 10인은 1학년부터 3학년까지 다양한 학년을 담당하고 있었으며, 1학년은 통합과학과 과학탐구실험을, 2학년은 화학 I을, 3학년은 화학 II를 담당하고 있었다. 10명의 화학 교사 중 8명은 남교사, 2명은 여교사였으며, 교육경력은 1년부터 25년으로 다양하게 분포하였다. 학교별로 2개 반에 대한 설문이 진행되었으며, 학교별, 학년별에 따라 1개 반의 인원수가 차이가 있어 학교별 참여 학생 수는 29명에서 63명으로 다양하다. 총 439명의 학생이 설문에 참여하였으며, 설문에 참여한 집단의 특성은 Table 1과 같다.

이후 양적 연구 결과를 해석하기 위해 심층 면담이 시행되었으며, 면담에 참여한 학생들은 설문조사에 참여한 10개교 중 5개교의 학생들을 대상으로 하였다. 본 연구의 설문에 협조한 10명의 화학 교사가 담당하는 학년을 대상으로 설문조사가 진행되었기 때문에 학교마다 설문에 참여한 학생들의 학년이 다소 차이가 있다. 따라서 학교마다 면담에 참여한 학년이 다르다. 면담 대상 학교를 선정할

때 가능한 학년이 고루 참여할 수 있도록 면담학교를 선정하였으나, 3학년의 경우 면담이 이루어지는 시점이 학교를 졸업한 시점이었기 때문에 면담 참여에 어려움이 있어 제외하였다. 설문을 진행하였을 때와 면담이 진행될 때의 시점이 차이가 있어 동일 대상의 학년이 달라졌으며, 이 논문에서 면담 대상의 학년 기술은 설문조사를 행했던 시점을 기준으로 기술하였다. 면담에 선정된 5개 학교 중 3개 학교는 1학년이, 2개 학교는 2학년이 면담 대상이다 (Table 2). 학교별로 2~13명이 면담 학생이 선정되었으며, 면담에 참여한 학생들은 수행평가에 관해 관심이 있는 학생들로 성적은 중상위권에 해당하며 면담자인 화학 교사와 래포가 형성되어 있어 진솔하게 자신의 의견을 말할 수 있는 학생으로 편의 표집하였다. 면담 방법은 전반적인 양적 결과에 대한 해석이기 때문에 집단 면담을 진행하는 것이 적절하다고 판단하였으며, 면담에 참여한 교사가 학교 상황을 고려하여 개별 또는 집단 면담을 진행하였다. 총 면담에 참여한 학생은 28명으로 이중 남학생은 9명, 여학생은 19명이며, 학년별로 1학년은 13명, 2학년은 15명이었다.

자료 수집 및 분석

학생들이 화학 교사의 수업에서 경험한 수행평가 방식에 대한 인식을 알아보기 위해 수행평가 방식과 관련된 선행연구²⁴⁻²⁶와 한국교육과정평가원 학생평가지원 포털²⁷에서 제시한 수행평가 방식을 탐색한 결과, 10개 정도의 수행평가 방식이 제안되었다. 이후 설문에 참여하는 화학 교사 10인과 적절성 검토를 온라인 화상 회의를 통해 실시하였으며, 이를 통해 이 연구에서는 수행평가 방식을 8가지(실험 평가, 보고서 평가, 관찰 평가, 토의 또는 토론 평가, 구술 발표 평가, 프로젝트 평가, 독서 또는 글쓰기 평가, 자기 또는 동료 평가)로 확정하였다.

확정된 8가지 수행평가 방식에 대해 크게 4가지 문항을 통해 화학 교사의 수업에 대한 수행평가의 인식을 알아보게 하였다. 먼저 설문에 표집된 학교의 화학 교사 수업 시간에서 경험하였다고 인식한 수행평가 방식의 유무를 수행평가 방식별로 답하고(Q1), 실제 경험하였다고 인식한 경우 수행평가 방식별로 어느 정도의 빈도로 해당 수행평가 방식을 경험하는지 답하도록 하였다(Q1-1). 두 번

Table 1. The characteristics of survey participants

Demographic variable	Category	Frequency(%)
Gender	Male	252(57.4)
	Female	187(42.6)
Region	Metropolis	222(50.6)
	Medium-sized city	130(29.6)
	Countryside	87(19.8)
Grade	First	248(56.5)
	Second	162(36.9)
	Third	29(6.6)

Table 2. The characteristics of interviewee by school

School code	Number of interviewee	Grade of interviewee	Type of interview
A	Male 1, Female 1	2	Group
B	Female 3	1	Group
C	Female 5	1	Group
D	Male 2, Female 3	1	Group
E	Male 6, Female 7	2	Individual

째 문항은 제시된 8가지 수행평가 방식 중 가장 재미있게 참여하였다고 느낀 화학(과학) 수업의 수행평가 방식을 1가지 작성하고 그 이유를 기술하도록 하였다(Q2). 세 번째 문항은 두 번째 문항과 대조적으로 화학(과학) 수행평가 중 참여할 때 가장 어렵다고 느낀 화학(과학) 수업의 수행평가 방식 1가지를 작성하고 그 이유를 기술하도록 하였다(Q3). 마지막, 네 번째 문항에서는 8가지 수행평가 방식 중 가장 적절한 평가라고 생각하는 수행평가(Q4)를 1순위부터 3순위까지 작성하도록 하였다.

실제 수업 시간에 진행되는 수행평가는 이 연구에서 제시한 8개의 방식이 분리되어 시행되거나 몇 가지 방식을 결합하여 진행하며, 교사에 따라서 결합하는 수행평가 방식은 다를 수 있다. 이 연구는 화학 교사들의 수업 시간에 학생들이 주로 경험하는 수행평가 방식에 따른 학생의 인식을 알아보는 것이 연구의 목적이기 때문에 연구에 참여한 10인의 화학 교사들과 온라인 화상 회의에서 결정된 8가지 수행평가 방식을 독립적으로 비교하도록 학생들에게 요구하였다.

설문조사가 시행된 후 수집된 자료에 대해 데이터 스크리닝과 코딩이 진행되었다. Q1인 수행평가 방식에 따른 경험 여부와 수행 횟수에 대한 학생들의 인식은 빈도 분석을 진행하였다. 성별에 따라 수행평가에 대한 인식이 다르게 나타난 선행연구를 참고하여²⁸⁻³⁰ Q2와 Q3은 빈도 분석 외에도 추가로 남녀 간 차이를 알아보기 위해 SPSS 27를 이용하여 교차 분석을 수행하였다. Q4는 순위형으로 적절한 수행평가를 질문하였기 때문에 순위별 가중치를 부여하여 분석하였다. 순위형 설문 분석 방법³¹을 참고하여 1순위는 3의 가중치를, 2순위는 2의 가중치를, 3순위는 1의 가중치를 부여한 후, 전체 가중치가 부여된 합계에 따른 성별과 전체 비율을 계산하였다. 또한 Q2와 Q4를 비교하여 재미있지만 적절한 수행평가로 선정되지 않은 수행평가 방식과, Q3과 Q4를 비교하여 어렵지만 적절한 수행평가로 선정된 수행평가 방식의 빈도와 비율을 조사하였다. 수집된 설문지별로 문항에 따라 결측값이 있었으며, 결측값을 제외한 자료에 대해서만 자료 분석이 이루어졌기 때문에 4가지 문항에 대한 성별 및 전체 합계가 다소 상이하다.

이렇게 해서 얻어진 양적 결과의 해석을 위하여 설문에 참여한 학생들을 대상으로 심층 면담을 진행하였다. 이 연구에서 심층 면담의 질문은 설명적 순차 설계에 따라 설문조사에서 얻어진 양적 자료에 기반하여 질문하였다. 예를 들어, ‘○○ 수행평가 방식이 가장 학생들이 재미있다고 조사되었는데 그 이유는 무엇인가요?, 남학생의 경우 ○○ 수행평가 방식이 어렵다고 많이 응답한 반면, 여학생의 경우 ○○ 수행평가 방식이 어렵다는 응답이 많은데 그 이유는 무엇이라고 생각하나요?’ 등과 같이 얻어진 양

적 자료 중 높은 빈도나 남녀 간 차이가 있는 결과에 기반하여 질문하였다. 1회의 면담은 45~60분 정도가 소요되었으며, 면담 초기에 연구에 취지와 동의를 얻은 후 모든 면담은 녹음되었다. 이후 녹음된 면담은 전사되었으며, 전사된 면담 자료는 학교별(A~E)로 코드를 부여한 후 각 학교에서 학생별로 숫자와 해당 학생의 성별(M 또는 F)로 최종 코드(학교 코드+학생 번호+성별)를 부여하였다. 전사된 자료는 과학 교육 박사 2인에 의해 개별적으로 1차 분석한 후, 면담 참여 교사들과 온라인 화상 회의를 통해 분석 결과에 대해 논의를 거치면서 양적 결과 해석에 타당성을 확보하고자 하였다. 면담을 통해 얻은 양적 결과에 대한 해석의 편향을 피하고자, 면담에서 표집된 5개 학교에서 공통으로 도출된 해석을 중심으로 양적 결과에 대한 해석을 시도하였다. 다만, 성별에 따른 특징과 관련된 논의는 해당 성별과 관련된 면담 내용을 중심으로 의견을 도출하였다.

연구 결과 및 논의

화학 교사의 수행평가 방식

학생들이 경험하였다고 인식한 수행평가 방식별 빈도와 비율은 Table 3과 같다. 실험 평가가 가장 높은 비율로 학생들이 경험하였으며(93.2%), 다음으로 관찰 평가(82.6%), 보고서 평가(82.1%), 프로젝트 평가(79.5%), 구술 평가(73.4%) 순으로 경험하였다. 토의 또는 토론 평가와 독서 또는 글쓰기 평가를 경험한 학생의 비율은 44~45% 정도로 상대적으로 낮게 나타났다. 더불어, 최근 강조되고 있는 자기 또는 동료 평가는 63.5%의 학생이 화학 교사의 수업에서 경험하는 것으로 나타났다.

Fig. 1은 수행평가 방식별로 경험하였다고 응답한 학생들을 대상으로 구체적으로 한 학기에 어느 정도의 빈도로 경험하는지를 조사한 결과이다. 전반적으로 8가지 수행평가 방식 모두 한 학기에 1회 또는 2회를 화학 교사의 수업을 통해 경험한다고 응답하였다.

Table 3. Students' experience by assessment method(Q1)

Assessment method	Frequency(%)	
	Experienced	Did not experience
Experiment	399(93.2)	29(6.8)
Report	352(82.1)	77(17.9)
Observation	354(82.6)	75(17.4)
Discussion or debate	192(44.2)	242(55.8)
Oral presentation	320(73.4)	116(26.6)
Project	338(79.5)	87(20.5)
Reading or writing	196(45.4)	236(54.6)
Self or peer	275(63.5)	158(36.5)

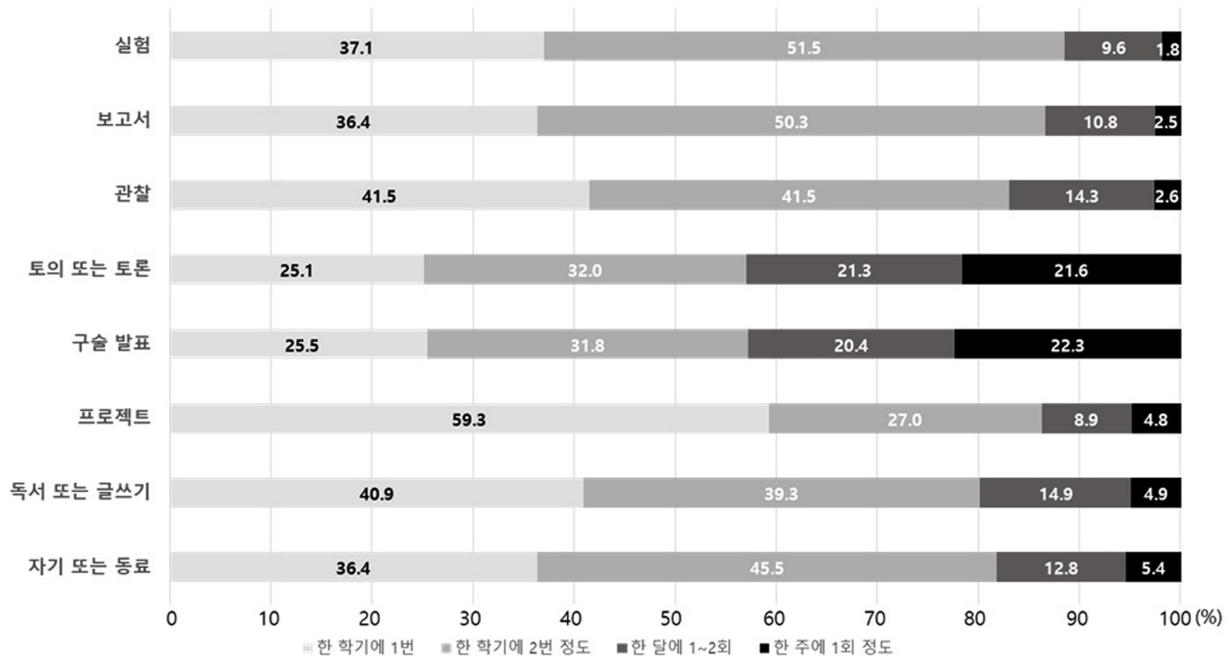


Figure 1. Execution frequency by assessment method (Q1-1).

실험 평가의 경우 51.5%로 절반 이상의 화학 교사가 한 학기에 2번 정도 학생들에게 평가의 기회를 제공하고 있었다. 눈여겨 볼만한 것이 토의 또는 토론 평가나 구술 발표 평가는 다른 평가에 비해 학생들이 경험하는 비율은 낮았으나, 이 평가가 실행되는 경우 한 주에 1회 정도 수행하는 빈도가 상대적으로 높게 나타났다.

재미 또는 어려움을 느끼는 수행평가 방식

학생들이 화학 교사의 수업에서 경험한 수행평가 방식 중에서 가장 재미가 있다고 생각한 수행평가 방식은 Table 4와 같다.

절반에 가까운 학생들이 실험 평가를 가장 많이 재미가 있다고 응답하였으며(N=177, 42.7%), 다음으로 보고서 평

가(17.3%), 프로젝트 평가(14.7%) 순이었다. 이 3가지 평가를 제외한 나머지 평가 방식에 대해서는 모두 10% 이하의 응답을 보였다. 남녀에서 2순위와 3순위의 순서에 다소 차이가 있으나, 이 3가지 방식이 학생들이 가장 재미있다는 느끼는 수행평가였다.

면담 결과 학생들이 3가지 평가 방식을 선정하는 이유는 조금씩 달랐다. 실험 평가의 경우 여러 학생들이 실제로 배운 내용을 직관적으로 확인할 수 있다는 점을 언급하였다.

자기가 직접 배운 내용을 직관적으로 볼 수 있어서 재미있어요. (A1-M)

직관적이고 실제로 관찰할 수 있어서 좋아요. (E12-F)

Table 4. Interesting performance assessment method by gender(Q2)

Assessment method	Frequency by gender(%)		Total(%)	$\chi^2(p)$
	Male	Female		
Experiment	112(47.1)	65(36.7)	177(42.7)	17.979 (.012*)
Report	42(17.6)	30(16.9)	72(17.3)	
Observation	13(5.5)	2(1.1)	15(3.6)	
Discussion or debate	6(2.5)	13(7.3)	19(4.6)	
Oral presentation	12(5.0)	16(9.0)	28(6.7)	
Project	30(12.6)	31(17.5)	61(14.7)	
Reading or writing	11(4.6)	12(6.8)	23(5.5)	
Self or peer	12(5.0)	8(4.5)	20(4.8)	
Total	238(100.0)	165(100.0)	415(100.0)	

*p<.05

반면 보고서 평가의 경우 공통으로 자신의 진로와 연관된 활동을 할 수 있음을 언급하였다. 자신이 평소 관심이 있는 주제와 관련된 주제를 탐색하는 측면에서 학생들은 재미를 느낀 것으로 보인다.

보고서는 평소 자신이 관심이 있는 주제로 작성할 수 있어 진로와 연관될 수 있어 좋아요. (B2-F)

프로젝트 평가에 대해 학생들은 내적 성장, 비교적 장시간의 평가임에 따른 교정 기회의 제공, 학문에 대한 심화 학습을 언급하면서 앞의 2가지 수행평가와 다른 이유를 언급하였다.

프로젝트는 다른 평가에 비해 장기간으로 진행되기 때문에 다른 평가에 비해 평가를 통해 얻어가는 것이 많아요. 그리고 기간이 기니깐, 다른 평가에 비해 수정할 수 있는 기회도 많아서 좋아요. (B3-F)

하나의 주제에 대해서 깊게 탐구하면서 과학에 대한 깊은 지식을 알게 되고, 심화된 과학 내용을 배워서 흥미를 갖게 돼요. (D1-F)

장기간으로 운영되니 끝나고 나서 산출되는 결과물에 대한 뿌듯함이 느껴져요. (E3-M)

재미있다고 느낀 수행평가 방식에 대한 빈도를 교차 분석을 진행한 결과, 남학생과 여학생 간의 유의미한 차이가 있었다($\chi^2=17.979, p<.05$). 실험 평가가 남학생과 여학생 모두 가장 높은 빈도로 언급한 평가 방식임에도 남학생이 여학생보다 10.4%p 더 높게 재미를 느낀 특징이 보였으며, 반면에 프로젝트 평가는 여학생이 4.9%p 더 재미있다는 응답이 높게 나타났다. 또한, 관찰 평가는 남학생이 여학생에 비해, 토의 또는 토론 평가와 구술 발표 평가는 여학생이 남학생에 비해 재미있다는 비율이 4%p 이상 높게

나타났다.

면담에서 남학생과 여학생 간의 차이가 발생하는 이유에 대해 학생들은 수행평가를 준비하는 성향과 부담에 따른 차이가 있을 수 있다고 응답하였다.

구술 평가의 경우 여학생들은 발표 대본도 써오고 피피티도 정성껏 만드는 데 반해 남자들은 뭔가 상대적으로 그런 거에 대한 걱정을 안 하고 그냥 발표 자료 같은 거 준비 안하고 바로 나와서 약간 즉흥적으로 하는 식이에요. 그래서 여학생들이 그런 보여지는 부분에 더 신경을 써서 부담감을 느낄 수 있다. (C1-F)

학생들이 화학 교사의 수업에서 경험한 수행평가 방식 중 어렵거나 참여가 힘들었다고 응답한 수행평가 방식은 Table 5와 같다.

흥미롭게도 Table 4에서 재미있다고 인식한 수행평가에서 빈도가 가장 높게 나타났던 실험 평가가 학생들이 가장 어려워하는 수행평가 방식이기도 하였다(N=88, 23.0%). 이러한 결과는 영어 교과에서 수행평가에 대한 학생의 인식을 알아본 연구와 일치한 것으로 이 연구에서도 가장 학생들이 선호하는 수행평가 유형과 가장 선호하지 않는 수행평가 유형이 서로 겹쳐 나타났다.³² 이는 학생들이 가장 많이 경험하는 방식의 경우 선호, 비선호가 모두 많을 수밖에 없다는 선행연구의 주장과 동일선에 있다고 판단되며, 본 연구에서는 이러한 이유 이외에도 추가적인 어려움을 발견하였다. 면담을 통해 학생들이 실험 평가가 어려운 이유로 실험을 수행하기 위해 관련 개념이나 원리에 대한 이해가 선행되어야 결론 도출을 적절히 할 수 있어야 하는 어려움이 있을 수 있으며, 예상과 다른 결과를 얻었을 때 대처와 관련하여 어려움을 가지는 것으로 설명하였다.

Table 5. Difficult performance assessment method by gender(Q3)

Assessment method	Frequency by gender(%)		Total (%)	$\chi^2(p)$
	Male	Female		
Experiment	69(31.7)	19(11.5)	88(23.0)	39.531 (.000 ^{***})
Report	51(23.4)	34(20.6)	85(22.2)	
Observation	11(5.0)	3(1.8)	14(3.7)	
Discussion or debate	5(2.3)	15(9.1)	20(5.2)	
Oral presentation	17(7.8)	25(15.2)	42(11.0)	
Project	30(13.8)	22(13.3)	52(13.6)	
Reading or writing	26(11.9)	37(22.4)	63(16.4)	
Self or peer	9(4.1)	10(6.1)	19(5.0)	
Total	218(100.0)	165(100.0)	383(100.0)	

^{***} $p<.001$

실험의 경우 실험 결과를 도출하기 위해서 먼저 원리에 대한 이해를 한 다음에 해야 하기 때문에 이에 대한 준비가 되어야 해서 부담이 됩니다. (D4-M)

실험할 때 고려해야 할 사항도 많고 실험 결과가 예상했던 것처럼 안 나올 수도 있어 그럴 경우 어떻게 해야 하는지 잘 몰라 두려움이 있어요. (E7-M)

이와 같은 실험에 대한 흥미와 어려움은 학교에서 진행되고 있는 실험이 주로 배운 이론을 확인하는 용도로 실험이 이루어지고 있다는 여러 선행연구 결과와 연결된다.^{33,34} 학생들이 실험 평가에 대한 즐거움을 단순히 이론을 확인할 수 있는 유용성에서 찾거나 실험 결과가 이론값과 다를 수 있는 것에 대한 두려움을 갖는 것 등은 아직 학교 현장에서 실제 세계에 기반하여 자신만의 모델을 만들고 만들어진 모델의 적합성을 확인하기 위한 도구로 실험을 활용하는 등 능동적인 탐구의 도구로서 실험이 활용되지 못하는 한계를 보여준다.

비슷한 빈도로 보고서 평가가 22.2%로 두 번째로 나타났으며, 독서 또는 글쓰기 평가가 16.4%로 그다음이었다. 보고서 평가의 경우 학생들은 실험 평가와 유사하게 보고서를 작성하기 위한 학생의 높은 준비도를, 독서 또는 글쓰기 평가의 경우 명확하지 않은 평가 기준에 따른 준비의 어려움을 언급하였다.

독서나 글쓰기 평가의 경우 약간 평가 기준이 주관적이고 추상적일 수 있어서 학생들이 이것을 이 평가를 준비하는 데 어려움이 있어요. (D2-M)

재미있는 수행평가 방식으로 많이 언급된 프로젝트 평가는 어렵게 느꼈다는 응답이 13.6%로 나타났다. 프로젝트 평가는 앞에서 언급한 학생의 준비도와 관련된 어려움과 달리, 모듈 단위로 운영됨에 따라 모듈원 간의 활동에서 갖는 어려움을 많이 언급하였다. 선행연구에서도 학생들은 수행평가 과정에서 무임승차 등과 같이 모듈원의 협력 부족을 어려움으로 인식하고 있었다.¹⁹ 교사가 모듈을 구성하여 운영하는 수행평가를 진행할 경우 학생들의 이러한 어려움을 고려하여 필요에 따라서 모듈 활동에 대한 평가 점수를 차등 배정 하는 등 다양한 방안 모색이 요구됨을 알 수 있다.

의견을 조율하는 과정에서 서로의 의견을 고집하다 보면 진전이 안 되기도 하고, 무임승차하는 친구가 있어서 불만이 발생할 수 있어요. (B2-F)

어려워하는 수행평가 방식에 대한 남녀차이를 교차 분

석한 결과 남녀 간에 유의미한 차이가 있었다($\chi^2=39.531$, $p<.05$). 실험 평가가 남학생의 31.7%가 어렵다고 응답한 반면, 여학생은 11.5%로 응답하여 20.2%p차이가 났다. 남학생은 실험 평가를 가장 어려워하는 반면, 여학생은 남학생에 비해 구술 발표와 독서 또는 글쓰기 평가에 대한 어려움의 비율이 높았으며, 남녀 간 차이가 각각 7.4%p, 10.5%p이었다. 이 결과는 일반적으로 수행이 강조된 영역인 실험의 경우 남학생들이 덜 어려움을 가질 것으로,³⁵ 반대로 여학생들이 구술이나 독서 또는 글쓰기 평가에서 남학생에 더 높은 평가를 받는 것으로³⁶ 간주되기 때문에 오히려 어려움이 더 적을 것으로 예상한 것과 다른 결과였다. 실험 평가의 경우 남학생들은 실험 과정을 설계하거나 예측되지 않는 결과를 얻은 후 이를 해결하기 위한 후속 행동에서 여학생에 비해 더 어려움을 느끼는 것으로 면담에서 공통으로 응답하였다.

실험을 할 때 실험 과정을 숙지하는 것이 어려워요. 그리고 실험을 할 때 고려할 것도 많고, 결과가 다르게 나올 수 있어서 두려움이 있어요. (D4-M)

막상 해보면 반응이 일어나는 게 재미있긴 한데, 뭔가 실험 과정을 설계해야 하는 것이 어려워요. (E1-M)

반면 구술 발표나 독서 또는 글쓰기 평가의 경우 여학생들은 남학생에 비해 관련 평가에 대한 준비도를 언급하였으며, 이러한 준비도가 많이 요구되는 수행평가를 어려운 평가로 간주하고 있었다.

구술은 여러 사람 앞에서 발표하기 때문에 발표 이전에 미리 준비해야 합니다. (E4-F)

글쓰기의 경우 글로 잘 다듬어서 표현해서 잘 정리를 해야 하는데, 이르기 위해서는 미리 관련 글 쓰는 주제에 대해서 알아봐야 하는 준비가 필요해요. (C3-F)

적절한 수행평가 방식에 대한 인식

적절한 수행평가 방식에 대한 학생의 인식은 Table 6과 같다. 학생들은 적절한 수행평가로 실험 평가를 응답한 비율이 가장 높았으며(28.4%), 그다음으로 보고서 평가(23.3%), 프로젝트 평가(14.3%) 순이었다.

적절한 수행평가 방식에 관하여 심층 면담을 실시한 결과, 3가지 이외의 다른 수행평가 방식이 상대적으로 학생들에게 ‘교과와 관련이 적은 개인의 능력에 따라 달라지는 평가’, ‘평가 기준이 불명확한 평가’로 인식되고 있었다. ‘교과와 관련이 적은 개인의 능력에 따라 달라지는 평가’란 예를 들어, 구술 발표에서 화학 교과의 학습과 관련된 역량이 아니라 개인별 발표 능력이 평가 결과에 영향을 미

Table 6. Performance assessment method recognized as appropriate (Q4)

Assessment method	Ratio by gender		Total
	Male	Female	
Experiment	31.2	24.8	28.4
Report	26.5	19.1	23.3
Observation	10.5	7.6	9.2
Discussion or debate	3.4	3.8	3.6
Oral presentation	6.6	9.6	7.9
Project	11.2	18.2	14.3
Reading or writing	6.2	12.5	8.9
Self or peer	4.4	4.4	4.4

치는 것으로, 학생들은 이와 관련된 요인은 배제가 되고 교과 관련 내용으로 평가받아야 한다고 생각하고 있었다. 유사한 의미에서 모둠으로 평가되는 평가 방식도 다른 친구들에 의해 평가 결과가 영향을 받기 때문에 자신이 제대로 평가받고 있지 않다고 언급하였다. 또한, 평가에 대한 기준이 학생들 관점에서 명확하지 않아 받은 점수가 이해되지 않는 것도 적절하지 않은 평가로 언급하였다.

말을 잘하는 학생이 더 좋은 평가를 받는 것은 부당해요. 그리고 과학 관련 내용 말고 개인의 발표 실력에 따라서 평가 결과가 달라지는 것도 적절하지 않다고 생각해요. (B3-F)

모둠으로 진행되는 수행평가에서 참여하지 않는 학생 때문에 제 점수가 영향을 받기도 합니다. 이건 부당하다고 생각해요. (C5-F)

토의 토론의 경우 어떤 기준으로 점수를 부여해서 채점이 되었는지 의심이 가요. (C2-F)

관찰 평가, 자기, 동료 평가, 토론 평가와 같은 평가는 모두 객관성이 부족해요. 특히, 자기 평가의 경우 자기 자신에게 높게 주거나 아니면 동료 평가의 경우 친한 학생들에게 무조건 높게 점수를 줘야 하니 객관성이 없어요. (A1-M)

이와 같은 결과는 수학이나 사회 교과의 학생 인식 연구에서 보여준 타당성과 공정성의 결여나 객관적인 평가 척도의 필요와 유사한 것이다.^{15,16,19} 다만, 이 연구에서는 수행평가 방식에 따른 인식을 알아봄으로써, 구체적으로 평가에 대한 타당성과 공정성 또는 객관적 평가 척도의 필요가 요구되는 수행평가 방식으로 구술 평가, 토의 또는 토론 평가, 관찰 평가, 자기 또는 동료 평가가 언급된 것이 차이점이다. 수행평가 중 이와 같은 평가를 운영할 때는 학생들이 상대적으로 평가의 공정성에 관해 더 우려가 깊다는 점을 고려하여 운영하는 것이 요구된다.

전체 학생 중 75명(17.1%)의 학생이 재미있는 수행평가 방식으로 언급한 수행평가를 적절한 수행평가로는 언급

하지 않았다. 많은 학생이 단순히 수행평가에서 느끼는 재미가 적절한 수행평가로 선정할 때 중요한 요소로 작용한 것이 아님을 알 수 있다.

반대로 전체 학생 중 156명(35.5%)의 학생이 어렵다고 인식하는 수행평가 방식과 적절한 수행평가 방식이 일치하였다(Table 7). 보고서 평가가 38.5%로 가장 높았으며, 실험 평가가 34.0%로 그다음으로 나타났다. 두 평가가 가장 높은 비율을 보인 이유로 학생들은 두 가지 평가 방식이 학생 관점에서 가장 익숙하다는 점을 언급하였다. 이는 편안하고 익숙한 평가가 학생들이 평가에 대한 선호를 결정할 때 중요한 요소임을 보인 선행연구의 결과와 일치한다.³²

두 평가가 제일 익숙한 평가라 좋아요. 익숙한 평가는 평가를 받을 때 다른 방식보다 편안함을 줘요. (E9-M)

특히, 보고서 평가가 어렵지만 적절한 수행평가로 인식되는 이유로 앞에서 언급한 적절한 평가의 2가지 요건 이외에도, 진로와 연관되어 평가를 받을 수 있기 때문에 관련 평가 내용이 생활기록부에 기록되어 대입 진학 때 도움이 된다는 점이 여러 학교에서 언급되었다. 이러한 결과는 수행평가 과제가 입시와 관련이 되어야 한다는 학생들의 요구와 일치한다고 할 수 있다.¹⁵ 또한, 보고서를 작성하면서 수업 시간에 배운 내용을 적용하는 과정에서 다시 학습할 수 있는 기회를 제공하여 적절하다고 인식하고 있었다.

본인이 노력한 만큼 평가를 받는 평가라 생각해요. 그리고 자신의 진로에 맞게 작성할 수 있으니 생기부에 기록되고, 대학에서 대입시 전공 적합성에서 이 내용을 많이 봐요. (B2-F)
배웠던 내용을 적용해서 학습할 수 있어요. (E4-M)

흥미롭게도 남학생의 경우 실험 평가를 어렵지만 적절한 평가로 생각하는 비율이 42.4%로 높았으며, 여학생은 7.9%로 34.5%p 차이로 성별에 따른 큰 차이를 보였다. 반

Table 7. Performance assessment method with difficulty and appropriate

Assessment method	Frequency by gender(%)		Total
	Male	Female	
Experiment	50(42.4)	3(7.9)	53(34.0)
Report	40(33.9)	20(52.6)	60(38.5)
Observation	7(5.9)	0(0)	7(4.5)
Discussion or debate	0(0)	1(2.6)	1(0.6)
Oral presentation	2(1.7)	3(7.9)	5(3.2)
Project	5(4.2)	5(13.2)	10(6.4)
Reading or writing	9(7.6)	6(15.8)	15(9.6)
Self or peer	5(4.2)	0(0)	5(3.2)
Total	118(100)	38(100)	156(100)

대로 여학생의 경우 보고서 평가가 52.6%로 가장 높았으며 남학생에 비해 18.7%p 더 높게 나타났다. 면담에서 적절한 평가 방식에 대해 남학생의 경우 익숙한 평가가 주는 편안함을 언급한 반면, 여학생의 경우 평가의 유용성을 더 많이 언급하였다. 이러한 점이 남녀 간에 차이로 나타난 것으로 판단된다.

결론 및 제언

이 연구는 학교 현장에서 화학 교사들이 주로 사용하는 수행평가에 관한 것으로, 8가지 수행평가 방식에 따른 학생들의 인식 차이와 그 이유를 분석하였다. 이를 통해 화학 교사의 수행평가 운영 시 요구되는 사항을 도출하고자 하였다. 이를 위하여 설문적 순차 설계로 먼저 시도의 중복 없이 10개 학교에서 총 439명의 고등학생을 대상으로 수행평가의 경험 정도, 재미와 어려움을 느꼈던 수행평가 방식, 적절하다고 인식되는 수행평가 방식에 대해 양적으로 조사하였다. 이후 설문에 응답한 5개 학교의 28명 학생을 대상으로 면담을 진행하여 양적으로 얻은 결과를 해석하였다.

첫째, 화학 교사의 수업 시간에 학생들이 경험하였다고 인식한 수행평가 방식에 대한 조사에서 실험 평가, 보고서 평가, 관찰 평가, 구술 발표, 프로젝트 평가가 많이 실행되는 것으로 나타났으며, 반대로 토의 또는 토론 평가나 독서 또는 글쓰기 평가는 절반 정도의 학생이 경험하는 것으로 나타났으며, 자기 또는 동료 평가도 학생들의 경험한 비율은 높지 않았다. 이를 통해 학생들이 화학 수업에서 수행평가를 다양하게 경험하지 않음을 확인할 수 있었다. 또한, 수행평가 방식별 경험 빈도 조사 결과 보통 한 학기에 1번 또는 2번을 경험하였으며, 가장 많이 활용되는 실험 평가는 한 학기에 2번 정도의 경험 기회를 제공하는 경우가 가장 많은 것으로 나타났다.

둘째, 학생들이 수행과정에서 재미와 어려움을 느꼈던 수행평가 방식을 살펴보면, 재미있는 수행평가 방식으로 실험 평가, 보고서 평가, 프로젝트 평가 순으로 가장 많았으며, 나머지 평가 방식은 적은 응답을 보였다. 실험 평가는 배운 내용의 직관적 확인 가능성을, 보고서 평가는 진로와 연관됨을 그 이유로 언급하였다. 한편, 학생들은 프로젝트 평가가 학생의 내적 성장 자극, 장기간 평가임에 따른 교정 기회의 제공, 학문에 대한 심화 학습을 제공하는 것을 그 이유로 제시하였다. 재미있는 수행평가 방식에서 남녀 간 차이가 통계적으로 유의미하였으며($\chi^2=17.979, p<.05$), 이에 대한 이유로 면담을 통해 수행평가를 준비하는 남녀 간 성향 차이를 확인할 수 있었다. 어려운 수행평가 방식에서는 실험 평가, 보고서 평가, 독서 또는 글쓰기 평가가 가장 많았다. 학생들은 실험 평가의 경우 실험 결과 도출에서 필요한 과학 개념이나 원리에 대한 학습 등 학생의 준비도를 언급한 반면, 독서 또는 글쓰기 평가는 평가 기준의 불명확성으로 평가를 준비하는데 어려움을 언급하였다.

셋째, 학생들은 적절한 수행평가 방식으로 실험 평가, 보고서 평가, 프로젝트 평가를 높은 비율로 선택하였으며, 그 이유로 제시된 3가지 평가를 제외한 평가 방식이 상대적으로 과목과 관련된 내용으로 평가되지 않고 교과와 관련이 적은 개인의 능력이 평가에 영향을 미치거나 평가 기준이 불명확하다고 인식하고 있었다. 재미있는 평가로 언급된 평가 방식을 적절한 평가로는 언급하지 않는 것에서 '평가 기준이 불명확한 평가'에 대한 학생의 인식을 보다 명확하게 확인할 수 있었다. 어렵지만 적절한 수행평가 방식에 관한 인식에서는 평가에 대한 익숙함을 학생들이 주로 언급하였고, 보고서 평가와 실험 평가가 가장 높은 빈도를 보였다. 흥미롭게도 남학생의 경우 실험 평가를 편안한 평가로 언급하는 경우가 있었고, 여학생은 보고서 평가를 평가의 유용성을 이유로 언급하며 남녀 간 차이를

보였다.

이 연구로부터 화학 교사들이 사용하는 수행평가의 방식, 학생들이 재미있거나 어려운 수행평가 방식에 관한 인식과 이유를 살펴볼 수 있었다. 더불어 남학생과 여학생의 수행평가 방식에 관한 어려움의 차이를 확인할 수 있었다. 또한, 적절한 수행평가에서 제시된 학생의 의견을 반영하여 수행평가가 진행될 때 공정한 평가라고 여길 수 있는 합의된 평가 기준과 이에 따른 수행평가의 진행이 요구됨을 알 수 있다.

이 연구는 10개 고등학교를 중심으로 진행된 연구로서 다른 시도의 학교를 표집하여 연구의 타당성을 확보하고자 하였으나 전체 화학 교사를 대표하기에는 연구 대상 수가 부족할 수 있으며, 학교 급에 따른 차이 등이 다르게 나타날 수 있으므로 연구 결과의 해석에 유의할 필요가 있다. 더불어 실제 수업 시간에 시행되는 수행평가는 8가지 수행평가 방식이 분리되어 운영되지 않을 수 있으며, 몇 가지 방식이 결합하여 운영될 수 있다. 그러나 이 연구에서는 교사들에 따라 다양하게 적용될 수 있는 여러 가지 수행평가 결합 방식을 모두 고려하여 조사할 수 없는 한계가 있어, 8가지의 수행평가 방식을 분리하여 학생들의 인식을 조사하였다. 따라서 수업에서 진행되는 수행평가에 대한 심도 있는 분석을 위해서는 실제 교사들의 수행평가 방식의 결합 등을 고려한 후속 연구가 진행될 필요가 있다.

이 연구 결과를 바탕으로 향후 현장의 수행평가 실행을 위한 제언을 하면 다음과 같다.

첫째, 현장에서 이루어지는 수행평가 방식 및 빈도에 관한 보다 상세한 연구가 필요하다. 이 연구는 10개 학교를 대상으로 화학 교사들이 수행하는 수행평가 방식에 대한 경험 정도를 조사하여 분석하였다. 이를 통해 상대적으로 수행평가 방식으로 잘 활용되지 않는 수행평가 방식을 확인할 수 있었다. 이러한 경향이 대상 학교의 특징인지 아니면 전반적인 화학 교사의 수업에서 나타나는 상황인지를 확인할 필요가 있으며, 학교 급별 차이가 있는지도 분석할 필요가 있다. 전반적인 화학 교사의 수행평가 방식 운영에 대한 조사는 운영 실태 및 이유를 탐색함으로써 특정 평가 방식의 적절한 운영을 위한 방안을 마련할 수 있으며, 과정 중심 평가의 성공적 실행을 위한 개선점을 도출하는 것이 가능하다. 이러한 연구 결과는 교사 교육에 활용될 수 있으며, 보다 다각적으로 학생들에게 평가의 기회를 제공하고 적절히 활용할 수 있도록 도움을 줄 것으로 기대된다.

둘째, 학교에서 운영되는 수행평가 방식에 따른 평가 기준의 타당성 확보를 위한 연구 및 개선 방향에 관한 교사 연구가 필요하다. 학생들은 평가 기준의 명확성이 적절한 평가 선정에 주요한 요인이었다. 실제로 명확성이

떨어진다는 수행평가 방식이 학생들에게 왜 불명확성으로 인식되는지를 면밀히 조사하고 이를 개선할 필요가 있다. 이러한 연구를 통해 얻은 결과는 학생들에게 상대적으로 평가 결과의 신뢰성에 의심을 받을 수 있는 수행평가 방식이 보다 학생들에게 신뢰를 주는 방안으로 자리 잡고 실행되는 데 도움을 줄 것으로 보인다. 또한 현장의 사례를 바탕으로 평가 기준의 개선 및 실행을 위한 연수를 진행하는 것은 교사들의 평가 전문성을 함양하고 다양한 평가의 적절한 실행으로 이어져 과정 중심 평가의 성공적 정착으로 이어질 수 있다.

셋째, 학생의 성장을 위한 수행평가 실행을 지원하는 연구가 진행될 필요가 있다. 2022 개정 교육과정에서는 새롭게 미이수를 도입하였으며, 이수 요건으로 40% 이상의 학업성취율을 예고하고 있다.⁷ 학생의 미이수를 예방하기 위해서는 진단평가와 형성평가가 수시로 진행되어 학생들에게 적절한 피드백이 제공되어야 하며, 이러한 의미에서 과정 중심 평가로 수행평가는 새로운 교육과정에서 더 많이 활용될 것으로 보인다. 하지만 이 연구에서 보인 바와 같이 실제 학교 현장은 소수의 수행평가 방식을 낮은 빈도로 활용하고 있어서, 보다 교사들이 다양한 수행평가를 적극적으로 활용할 수 있는 현장 지원과 관련된 연구가 필요하다. 여러 선행연구에서^{37,38} 교사들의 학습공동체를 통해 교사의 평가나 수업의 실행이 달라진 점을 참고해볼 때, 수행평가를 통한 평가의 개선을 중점으로 하는 연구를 교사들의 학습공동체를 중심으로 시작하는 방안도 가능할 것이다.

넷째, 화학 교사들이 운영하는 실험 평가의 개선이 요구된다. 이 연구에서 학생들이 가장 재미있게 경험하였거나 어려운 수행평가로 실험 평가가 많이 거론되었다. 실험은 과학 교육에서 탐구의 도구로 중요한 방법이다. 하지만, 이 연구에서 도출된 재미나 어려움에 대한 이유를 살펴보면, 학교 현장에서 단순히 실험이 수업 시간에 배운 이론을 확인하는 용도로 활용되고 있음을 알 수 있으며, 학생들이 실험 결과의 해석과 실험 결과와 이론값의 차이 등에서 어려움을 겪는 것을 알 수 있었다. 능동적인 탐구를 위한 방안으로 2015 개정 과학과 교육과정에서는 ‘모형의 개발과 사용’, ‘증거에 기초한 토론과 논증’ 등 8가지 기능을 제시하고 있다. ‘모형의 개발과 사용’의 의미는 학생 수준에서 나름의 학생 모델을 생산하고 이를 실험 등과 같은 탐구활동을 통해 검증하는 모델링 과정을 통해 능동적 과학자의 모습을 경험하도록 하는 것이며, ‘증거에 기초한 토론과 논증’에서는 탐구 과정에서 얻어진 자료를 바탕으로 과학적 증거에 근거한 주장과 근거에 바탕하지 않은 의견을 구별하고, 자료와 증거에 기초하여 논증이나 반론을 구성함으로써 과학적 소양을 기르는 것을 목적으

로 하고 있다. 학생들이 실험 평가에서 거론한 재미나 어려움은 실험이 교과서에서 제시된 법칙 등을 확인하기 위한 용도로 주로 쓰이고 있는 것과 연관된다. 이러한 문제를 해결하기 위해서는 능동적인 탐구 도구로 실험을 활용할 수 있도록 교사가 지도할 수 있어야 하며, 이러한 지도가 결실을 맺기 위해서는 정답을 찾는 실험 평가보다는 답을 찾기 위한 과정으로서의 실험 평가가 진행되어야 할 것이다.

REFERENCES

- Lee, J. W.; Noh, J. *East Asian Mathematical Journal* **2020**, *36*, 515.
- Shin, H.; Ahn, S. -Y.; Kim, Y. *The Journal of Curriculum and Evaluation* **2017**, *20*, 135.
- Cho, S. Y. *Secondary Education Research* **2017**, *65*, 255.
- Jeon, Y.; Lee, Y.-H. *Journal of the Korean School Mathematics* **2019**, *22*, 199.
- Lim, Y.; Kim, S. A.; Kim, S.-H. *The Korean Journal of Technology Education* **2022**, *22*, 54.
- Ministry of Education. *2015 Revised National Curriculum*; 2015.
- Ministry of Education. *Full-fledged Discussion of the Future Curriculum with the Public: Announcement of "2022 Revised National Curriculum Plan"*; 2021.
- Lane, S.; Stone, C. A. Performance Assessment. In *Educational Measurement*; Brennan, R. L. Ed.; Praeger Publishers: CT, 2006; pp 387-431.
- Kim, J. *The Journal of Learner-Centered Curriculum and Instruction* **2018**, *18*, 839.
- Gyeonggi-do Office of Education. *2022 Common Subject 'Integrated Science' to Ensure the Minimum Academic Achievement Level*; 2022.
- Lee, J.; Nam, J.; Moon, S. *Journal of the Korean Association for Science Education* **2003**, *23*, 66.
- Uhm, H. J.; Yang, J. H. *Journal of Korean Practical Arts Education* **2008**, *21*, 255.
- Oh, S.; Kwon, N. Y.; Kim, S. *The Journal of Educational Research in Mathematics* **2019**, *29*, 507.
- Lee, H.; Beak, J.; Kwak, Y. *Journal of the Korean Association for Science Education* **2020**, *44*, 133.
- Kim, J.; Kim S. *Journal of Educational Research* **2008**, *18*, 1.
- Lee, K. Y. *Social Studies Education* **2007**, *46*, 95.
- Lim, E.-H.; Ahn, K. *Journal of Learner-Centered Curriculum and Instruction* **2015**, *15*, 107.
- Park, T. *Secondary English Education* **2016**, *9*, 77.
- Ko, K. S.; Kim, Y.-J. *Korean Journal of Contemporary Education* **2020**, *32*, 87.
- Shadish, W. R.; Cook, T. D.; Campbell, D. T. *Experimental and Quasi-experimental Designs for Generalized Causal Inference*; Mifflin Company: Houghton, 2002.
- Kim, S. *Korean Journal of Teacher Education* **2019**, *35*, 1.
- Ivankova, N. V.; Creswell, J. W.; Stick, S. L. *Field Methods* **2006**, *18*, 3.
- Creswell, J. W.; Clark, V. L.; Gutmann, M. L.; Hanson, W. E. Advanced Mixed Methods Research Designs. In *Handbook of Mixed Methods in Social and Behavioral Research*; Tashakkori, A.; Teddlie, C. Eds.; Sage: CA, 2003; pp 209-240.
- Kim, H. K.; Ko, K. C.; Kim, D. J. *Journal of Elementary Mathematics Education in Korea* **1999**, *3*, 61.
- Kim, S.-W. *Journal of Educational Evaluation* **2007**, *20*, 53.
- Lee, K. Y.; An, H. S. *Journal of the Korean Association for Science Education* **2005**, *25*, 173.
- Korea Institute for Curriculum and Evaluation. "Student Assessment Support Portal"; <https://stas.moe.go.kr/cmnl/main> (accessed 2022-06-29)
- Hwang, S. W.; Kim H. K.; Yoo, H. J.; Park, S. J. *Journal of the Korean Association for Science Education* **2001**, *21*, 506.
- Kim, H.-J. *Asia-pacific Journal of Multimedia Services Convergent with Art, Humanities, and Sociology* **2019**, *9*, 131.
- Kim, S.; Kim, H. *Journal of the Korean Chemical Society* **2022**, *66*, 107.
- Seong, T. J. *Easy-to-understand Statistical Analysis Using SPSS/AMOS*; Hakjisa: Seoul, 2012.
- Yu, I.; Lim, H.-W. *Foreign Languages Education* **2021**, *28*, 85.
- Kim, K.; Paik, S. *Journal of Chemical Education* **2021**, *98*, 2381.
- Kim, J.; Kim, S.; Paik, S. *Journal of the Korean Chemical Society* **2020**, *64*, 111.
- Choi, O.; Yeau, S. *Journal of Research in Curriculum Instruction* **2004**, *8*, 437.
- Kim, J. H. *Journal of Curriculum Evaluation* **2017**, *20*, 271.
- Kim, S.; Min, H.; Paik, S. H. *Korean Journal of Teacher Education* **2020**, *36*, 1.
- Kim, S.; Min, H.; Ha, M. S.; Paik, S. H. *The Journal of Curriculum and Evaluation* **2021**, *24*, 231.