

## 우리나라 초·중·고등학교 과학과의 평가 실태

정은영 · 최원호<sup>1\*</sup>

전남대학교 · <sup>1</sup>순천대학교

### A Survey on Evaluation in Science Education at Primary and Secondary School in Korea

Eunyoung Jeong · Wonho Choi<sup>1\*</sup>

Chonnam National University · <sup>1</sup>Sunchon National University

**Abstract** : The purpose of this study was to monitor evaluation in science education at primary and secondary school in Korea. In order to do this, the survey was administered. The subject was 518 teachers; 292 elementary school teachers, 110 middle school science teachers and 116 general high school science teachers. It was found that the ratio of paper examination increased while that of performance evaluation decreased according to the grade of school. In the science paper examination, the ratio of selection type item was higher than that of supply type item, and the ratio of item belonging to knowledge area was higher than that of item belonging to the inquiry area regardless of the grade of school, teaching career, and teacher's gender. The teachers recognized the necessity of supply type item, but they suffered from the reliability of the result of that item. And it was hard for them to do performance evaluation because it took much time and effort. This study suggested that the high-quality in-service teacher training be provided and laboratory assistants who help the science experiment and grading test papers be employed in order to fulfil the performance evaluation and increase the ratio of supply type item and items belonging to the inquiry area.

keywords : evaluation in science education, performance evaluation, paper examination, supply type item

## I. 서론

교육에 대하여 공통적으로 가지는 일반적인 정의는 ‘인간이 지니고 있는 무엇인가를 변화시키려는 활동’이다(성태제, 2010). 학교에서 다루는 교과마다 교육과정이라는 특정 목표를 가지기 때문에 학교에서 이뤄지는 각 교과에서의 교육은 각 교과에서 규정한 목표에 도달하기 위해 학생들을 수업 중 변화시키는 활동이라고 표현할 수 있다. 그래서 교육목표를 설정하여 교수·학습이 이루어지면 그 목표를 달성하였는지를 점검하여 교수·학습의 효과를

확인하고 부족한 점을 개선할 수 있다. 즉, 평가를 통하여 학생의 학습에 피드백을 제공할 수 있으므로 학생의 학습 향상에 도움을 줄 수 있고(AFT, NCME, & NEA, 1990), 수업 중 평가 활동을 통해 학생의 능력 향상에 도움을 주어 학생의 학습을 촉진할 수 있다(김신영, 2002; Stiggins, 2002).

학생들의 학습 능력을 촉진하는 방향은 21세기 지식 기반 사회에 대비하여 지식의 양보다는 질이 더 중요하다는 인식 하에 이미 영국, 미국 등 교육 선진국에서는 교육에서의 평가 방식을 절대 평가 체제를 실시하는 형태로 교육 개혁을 위한 노력을

\*교신저자 : 최원호(stensil@sunchon.ac.kr)

\*\*이 논문은 한국과학창의재단 과학교육과정 연구인 ‘우리나라 초·중등 과학교육 문제점 진단 연구’의 일환으로 실시된 설문 조사 중 과학과 평가에 대한 내용임.

\*\*\*2014년 2월 24일 접수, 2014년 4월 5일 수정원고 접수, 2014년 4월 7일 채택

기울이고 있으며, 우리나라도 성취평가제라는 이름으로 절대 평가 체제를 도입하며(교육과학기술부, 2011; 김신영, 2012) 학생들의 학습 능력 신장을 위한 노력을 기울이고 있다.

특히 평가는 교육의 진행 정도를 확인하고 피드백할 수 있는 장치로, 교사의 평가 전문성은 학교 교육의 발전을 위해 매우 중요한 요소이다. 우리나라 과학 교사의 평가 능력에 초점을 맞춰 기존에 진행된 연구를 살펴보면, 과학과 지필평가에서 탐구 영역보다는 지식 영역에 한정하여 출제하는 경향을 보였으며(양일호 외, 2008; 오현석, 이기영, 2006), 지식 영역에서도 특히 기억을 요구하는 문항을 주로 출제하고 있었다(김운희 외, 2010). 또한 서답형 문항의 실태에 관해 조사해보면, 국가수준 학업성취도 평가에서 서답형 문항이 차지하는 비율이 외국의 주요 국가들에 비해 낮고(심재호 외, 2012), 교사들은 학생들의 사고력 함양을 위해 서술형 평가가 필요하다고 인식하지만(심애순, 2008), 출제 및 채점의 편의성을 위해 특정 단원에서만 출제하거나 난이도를 다양하게 조절하지 못하는 문제점(이미화, 2013)과 고등정신 능력이나 창의적 문제 해결력을 측정하는 대신 지식이나 단순한 기능에 초점을 맞춰 명확하게 답이 나타나는 문항을 출제하는 경향(최은주, 2011)이 있었다. 수행평가와 관련하여 조사해보면, 수행평가가 차지하는 비율이 지필평가보다 낮으며, 수행평가를 실시하더라도 신뢰도 높은 성적 산출을 목적으로 주로 실험 보고서 중심의 평가를 실시하는 경향이 있었다(김석우, 2007; 서정관, 2008; 손정숙, 2011).

우리나라 과학과 교육과정은 ‘과학의 기본 개념의 이해, 과학의 탐구 능력, 과학적 태도 등을 균형 있게 평가’할 것을 강조하고 있으며, ‘선다형, 서술형, 논술형, 관찰, 보고서, 실기 검사, 면담, 포트폴리오 등의 다양한 방법’을 사용할 것을 명시하고 있다(교육과학기술부, 2009). 교사가 교육과정에 근거한 타당한 문항을 출제하고 채점할 수 있는 평가 전문성을 가지는 것은 학생들의 학습 능력 향상과 학교 교육 개혁을 위해 매우 중요함에도 불구하고 선행 연구에서 보듯이 교사의 평가 전문성이 부족함을 알 수 있다.

선행 연구에서는 연구 대상 선정의 어려움으로 다양한 학년과 지역에 대한 연구를 진행하지 못한 한계를 가지고 있었는데, 본 연구에서는 우리나라 과학교육 평가의 개선 방향 도출을 목적으로 우리나라 초·중·일반계 고등학교에서 과학과 평가의 실태를 다양한 학년과 지역을 대상으로 조사하고자 하였다. 이를 위해 지필평가와 수행평가의 비율, 지필평가에서 문항 유형 및 행동 영역별 비율, 서술형 평가의 실시 현황, 수행평가의 실시 현황 등에 대해 교사 대상으로 설문 조사를 하였다. 이를 통해 과학과 평가의 실태를 파악함으로써, 현장 과학교육에서 평가와 관련된 문제점을 진단하고, 과학과 평가의 개선 방안을 모색하고자 하였다.

## II. 연구 방법

### 1. 설문 대상

본 연구에서는 초등학교 교사, 중학교와 일반계 고등학교 과학 교사를 대상으로 설문을 실시하였다. 고등학교의 경우 설립 목적에 따라 여러 유형으로 구분되는데, 본 연구에서는 일반계 고등학교로 한정하였다. 설문 조사를 위해 지역의 규모, 시도교육청별 학교 수, 학교 유형 등을 고려하여 학교급별 모집단 학교 수의 약 1%를 추출하였는데, 표집 대상 학교는 초등학교 60개교, 중학교 40개교, 일반계 고등학교 20개교였다. 초등학교의 경우 4, 6학년 담임교사 및 과학 전담 교사, 중학교와 일반계 고등학교의 경우 표집학교의 전체 과학 교사가 설문에 응하도록 하였다. 설문지 회수 결과, 초등학교 교사 292명, 중학교 교사 110명, 고등학교 교사 116명, 총 518명이 응답하였다.

### 2. 설문지의 개발

과학과 평가의 실태를 조사하기 위해 표 1과 같이 설문 항목과 설문 내용을 설정하였다. 과학교육 전공 교수 2명이 설문 문항의 초안을 작성하고, 과

표 1. 설문지의 구성

설문 항목	설문 내용	문항 번호
평가 방법별 비율	- 지필평가와 수행평가의 비율	1
지필평가에서 문항 유형 및 행동 영역별 비율	- 지필평가에서 문항 유형별 문항 비율	2
	- 지필평가에서 행동 영역별 문항 비율	3
지필평가의 서술형 문항	- 서술형 문항의 개발 및 실시 수준	4
	- 서술형 문항을 개발하지 않거나 단답형 문항에 가까운 유형으로 개발하여 실시하는 이유	4-1
	- 서술형 문항을 개발하여 실시하려고 노력하는 이유	4-2
수행평가	- 수행평가의 실시 방법	5
	- 수행평가를 실시하지 않거나 형식적으로 실시하는 이유	5-1
	- 수행평가를 실시 목적에 맞게 실시하려고 노력하는 이유	5-2
	- 수행평가 실시의 어려운 점	5-3

학교육 전공 교수 1인, 과학교육 전공 대학원생 2인, 과학 교사 3인이 논의하면서 수정·보완하였다.

시 현황, 수행평가의 실시 현황 등에 대해 교사 대상으로 설문 조사를 한 결과는 다음과 같다.

### 3. 설문 결과 분석 방법

설문 문항별 빈도와 백분율을 산출하여 전반적인 경향을 살펴보고, 각 문항에 대하여 학교급별, 성별, 교직 경력별 빈도와 백분율을 산출하여 설문 대상의 변인별 특징을 고찰하고자 하였다.

## Ⅲ. 결과 및 논의

본 연구에서 우리나라 초·중·일반계 고등학교 과학과의 지필평가와 수행평가의 비율, 지필평가에서 문항 유형 및 행동 영역별 비율, 서술형 평가의 실

### 1. 지필평가와 수행평가의 비율

과학과 지필평가와 수행평가의 비율에 대한 교사들의 응답 결과는 표 2와 같다.

지필평가와 수행평가의 실시 비율 중 수행평가 비율을 학교급별로 나타내면 초등학교 42.7%, 중학교 32.1%, 고등학교 28.2%이었으며, 학교급이 올라갈수록 지필평가의 비율이 높고 수행평가의 비율이 줄어들고 있었다. 수행평가 비율을 교사의 성별로 나타내면 남자 36.3%, 여자 36.8%이었으며, 성별 차이는 거의 없었다. 수행평가 비율을 교직 경력별로 나타내면 14년 이하에서는 37.2%, 15년 이상에서는 35.7%이었으며, 경력이 많은 경우 수행평가 비율이 다소 낮았다.

표 2. 과학과 지필평가와 수행평가의 비율

단위 : %

	학교급			성별		교직 경력		전체 (N=518)
	초 (n=292)	중 (n=110)	고 (n=116)	남 (n=164)	여 (n=354)	14년 이하 (n=331)	15년 이상 (n=187)	
지필평가	57.3	67.9	71.8	63.7	63.2	62.8	64.3	63.3
수행평가	42.7	32.1	28.2	36.3	36.8	37.2	35.7	36.7

요약하면, 학교급이 올라갈수록 수행평가의 실시 비율이 감소하는데, 특히 초등학교에서 중학교로 올라가면서 약 10% 정도의 실시 비율이 줄었고 중학교에서 고등학교로 올라가면서 약 4% 정도 줄었다. 2009 개정 과학과 교육과정(교육과학기술부, 2009)의 목표에서는 기본 개념의 이해와 함께 탐구 능력의 함양과 과학적 태도 함양을 중요하게 다루고 있다. 수행평가는 지필평가에서 평가하기 어려운 내용을 중심으로 학생들의 실제적인 행위를 직접 관찰하며 검사하는 평가로(성태제, 2010), 학생들의 지식이나 기능을 직접 관찰하며 평가하기 때문에 과학과 교육과정의 목표에서 지향하는 탐구 능력의 파악을 위해 효과적으로 이용할 수 있다. 학교급이 올라가더라도 수행평가의 실시 비율이 낮아지는 현상은 교육과정의 목표 구현 측면에서 바람직하지 않으므로 각 학교급에서는 수행평가의 최소 실시 비율을 권장할 필요가 있다.

## 2. 지필평가에서 문항 유형별 비율

과학과 지필평가에서 문항 유형별 비율에 대한 교사들의 응답 결과는 표 3과 같다.

지필평가에서 문항 유형별 비율을 학교급별로 나타내면 선다형 문항의 경우 초등학교 56.9%, 중학교 70.3%, 고등학교 71.4%이었고, 단답형 문항의 경우 초등학교 22.4%, 중학교 8.2%, 고등학교 5.3%이었으며, 서술형 문항의 경우 초등학교 20.7%, 중학교 21.8%, 고등학교 23.5%이었다. 선다형 문항의 경우 그 비율이 초등학교에서 중학교로 가면서 10.4%, 중학교에서 고등학교로 가면서

1.1% 증가한 반면, 단답형 문항의 경우 초등학교에서 중학교로 학교급이 높아지면서 14.2% 감소하였고, 중학교에서 고등학교로 가면서 2.9% 낮아졌다. 즉, 학교급이 높아지면서 선다형 문항의 비율은 증가하고 단답형 문항의 비율은 낮아졌다. 서술형 문항의 비율은 학교급이 높아질수록 다소 증가하고 있는 것으로 나타났다. 문항 유형별 비율을 교사의 성별로 비교하면, 선다형 문항의 경우 남교사와 여교사 모두 약 63% 정도 비율로 출제하고 있고, 단답형 문항의 경우 남교사 16.0%, 여교사 14.9%, 서술형 문항의 경우 남교사 20.7%, 여교사 22.1%로, 여교사가 남교사보다 서술형 문항을 좀 더 출제하는 것으로 나타났는데, 큰 차이로 보기는 어렵다. 문항 유형별 비율을 교직 경력별로 비교하면, 선다형 문항의 경우 교직 경력에 관계없이 약 63% 정도의 비율로 출제하고 있고, 단답형 문항의 경우 교직 경력 14년 이하인 교사들은 15.7%, 교직 경력 15년 이상인 교사들은 14.4%, 서술형 문항의 경우 교직 경력 14년 이하인 교사들은 21.1%, 교직 경력 15년 이상에서는 22.6%로 나타나, 교직 경력에 따라 문항 유형별 출제 비율이 달라진다고 보기는 어렵다.

요약하면, 중·고등학교는 초등학교와 비교할 때 과학과 지필평가에서 선다형 문항의 비율이 높다. 평가 목적에 따라 평가 문항의 형태가 달라질 수 있지만 선다형 문항은 단답형 문항에 비해 학생들이 답을 직접 작성할 수 없기 때문에 평가에서 학생들의 수준 파악이 상대적으로 어렵고 학생들은 평가 대비 학습을 단편적으로 수행할 가능성도 커진다. 따라서 중·고등학교에서 선다형 문항의 비율

표 3. 과학과 지필평가에서 문항 유형별 비율

단위: %

	학교급			성별		교직 경력		전체 (N=518)
	초 (n=292)	중 (n=110)	고 (n=116)	남 (n=164)	여 (n=354)	14년 이하 (n=331)	15년 이상 (n=187)	
선다형	56.9	70.3	71.4	63.6	63.1	63.2	63.5	63.3
단답형	22.4	8.2	5.3	16.0	14.9	15.7	14.4	15.3
서술형	20.7	21.8	23.5	20.7	22.1	21.1	22.6	21.6

을 낮추고 단답형 문항이나 서술형 문항의 비율을 높일 필요가 있다. 설문 조사 결과, 중·고등학교에서 단답형 문항보다 서술형 문항의 비율이 더 높은 것은 바람직한 현상이지만, 선다형 문항의 비율이 학교급이 높아질수록 증가하고 있는 것은 개선이 필요한 부분이다. 초등학교에서는 단답형 문항의 비율을 줄이고 서술형 문항의 비율을 늘리는 방향으로, 중·고등학교에서는 선다형 문항의 비율을 줄이고 서술형 문항의 비율을 높이는 방향으로 노력할 필요가 있다.

### 3. 지필평가에서 행동 영역별 비율

과학과 지필평가에서 지식 영역과 탐구 영역의 비율에 대한 교사들의 응답 결과는 표 4와 같다.

지필평가에서 탐구 영역의 출제 비율을 학교급별로 비교하면, 초등학교 42.6%, 중학교 48.2%, 고등학교 47.2%로, 초등학교에 비해 중·고등학교의 탐구 영역 출제 비율이 높았지만, 중·고등학교의 경우 그 차이가 1%로 크지 않았다. 탐구 영역의 출제 비율을 교사의 성별로 비교하면, 남교사 46.1%, 여교사 44.4%로, 남교사가 여교사보다 1.7% 높았으나 큰 차이로 보기는 어렵다. 탐구 영역의 출제 비율을 교직 경력별로 비교하면, 교직 경력 14년 이하인 교사들은 43.2%, 교직 경력 15년 이상인 교사들은 48.0%로 나타나, 경력이 많을수록 탐구 영역 출제 비율이 높게 나타났다.

요약하면, 과학과 지필평가에서 지식 영역과 탐구 영역의 출제 비율 비교에서 학교급, 성별, 교직 경력과 상관없이 모두 지식 영역이 탐구 영역보다 출제 비율이 더 높았다. 초등학교보다 중·고등학교

에서 지식 영역보다 탐구 영역의 비율이 다소 높은 것은 고무적인 일이다. 지필평가 문항의 행동 영역별 분포를 조사한 선행 연구에 의하면, 전체 문항 중 기억에 해당하는 문항의 비율이 76.1%(김윤희 등, 2010)이거나 69.6%(박현주, 2012)로 지식 영역 중 기억에 의존하는 문항 출제 비율이 본 연구 결과보다 높았으며, 탐구 영역의 출제 비율은 24~26%로(오현석, 이기영, 2006) 본 연구 결과보다 낮았다. 이는 문항이 해당되는 행동 영역에 대한 교사들의 판단과 연구자의 판단의 차이가 원인일 수 있는데, 교사들의 행동 영역별 문항 분류와 제3자인 연구자의 문항 분류 결과의 차이를 분석하는 연구를 추후 수행할 필요가 있다.

### 4. 서술형 평가의 실시 현황

과학과 지필평가에서 서술형 문항의 개발 및 실시 수준에 대한 교사들의 응답 결과는 표 5와 같다.

서술형 문항의 개발 및 실시를 채점 시간을 고려하여 제한된 범위에서 하는 비율(54.4%)이 가장 높았고, 서술형 문항으로 형식적으로 구분하기는 하지만 실제로는 서술형이 아니라 단답형에 가까운 응답을 요구하는 문항을 개발하여 실시하는 비율이 26.3%로 나타났으며, 채점 시간이 오래 걸리더라도 학생의 심층적 사고를 평가할 수 있는 형태로 개발하여 실시하는 비율이 19.3%로 나타났다. 응답 결과를 학교급별로 비교하면, 중·고등학교의 경우 채점 시간을 고려하여 제한된 범위에서 실시하는 비율이 60% 이상으로 초등학교의 경우(48.8%)보다 높게 나타났고, 초등학교에서는 학생의 심층적 사고를 평가할 수 있는 형태로 서술형 문항을

표 4. 과학과 지필평가에서 지식 영역과 탐구 영역의 비율

단위: %

	학교급			성별		교직 경력		전체 (N=518)
	초 (n=292)	중 (n=110)	고 (n=116)	남 (n=164)	여 (n=354)	14년 이하 (n=331)	15년 이상 (n=187)	
지식 영역	56.6	51.8	52.7	53.4	55.1	56.2	51.8	54.6
탐구 영역	42.6	48.2	47.2	46.1	44.4	43.2	48.0	44.9

표 5. 과학과 서술형 문항의 개발 및 실시 수준

단위 : 명(%)

	학교급			성별		교직 경력		전체
	초	중	고	남	여	14년 이하	15년 이상	
형식적 구분만, 실제로는 단답형 문항에 가까운 유형 개발	69 (27.2)	25 (22.9)	31 (27.4)	47 (30.3)	78 (24.3)	73 (24.0)	52 (30.2)	125 (26.3)
채점 시간을 고려하여 제한된 범위에서 실시	124 (48.8)	66 (60.6)	69 (61.1)	74 (47.7)	185 (57.6)	170 (55.9)	89 (51.7)	259 (54.4)
학생의 심층적 사고를 평가할 수 있는 형태로 개발하여 실시	61 (24.0)	18 (16.5)	13 (11.5)	34 (21.9)	58 (18.1)	61 (20.1)	31 (18.0)	92 (19.3)
합계	254 (100.0)	109 (100.0)	113 (100.0)	155 (100.0)	321 (100.0)	304 (100.0)	172 (100.0)	476 (100.0)

개발하여 실시하는 비율이 24%로 중·고등학교의 경우보다 높게 나타났다. 응답 결과를 교사의 성별로 비교하면, 남교사의 경우 형식적 구분만 하는 비율이 여교사의 경우보다 다소 높게 나타났고 제한된 범위에서 실시하는 비율이 여교사의 경우보다 10% 정도 낮게 나타났다. 응답 결과를 교직 경력 별로 비교하면, 경력 15년 이상인 교사들의 경우 형식적 구분만 하는 비율이 경력 14년 이하인 교사들의 경우보다 다소 높게 나타났다.

학교 현장에서는 채점 시간을 고려하여 서술형 문항을 제한된 범위에서 실시하는 경우가 많은데, 이러한 경향은 초등학교보다 중·고등학교에서 더 많이 나타났다. 그리고 서술형의 응답이 아니라 단

답형의 응답을 요구하는 문항으로 개발하는 경우도 그 다음으로 많았다. 이는 학생들에게 서술형의 응답을 요구함으로써 고등정신능력을 함양하고자 한다는 취지를 고려할 때 개선이 필요한 부분이라고 생각한다.

서술형 문항을 개발하지 않거나 단답형 문항에 가까운 유형으로 개발하여 실시한다고 응답한 교사를 대상으로 그 이유를 질문한 결과는 표 6과 같다.

그 이유로, ‘채점 결과의 신뢰도가 낮아서’라고 응답한 비율이 45.5%로 가장 높았고, ‘학생들이 서술형 문항을 너무 어려워하여 과학을 기피할 우려 때문’이라고 응답한 비율이 27.6%로 나타났으며,

표 6. 서술형 문항을 개발하지 않거나 단답형 문항으로 개발하여 실시하는 이유

단위 : 명(%)

	학교급			성별		교직 경력		전체
	초	중	고	남	여	14년 이하	15년 이상	
채점 결과의 신뢰도가 낮아서	33 (48.5)	10 (41.7)	13 (41.9)	25 (53.2)	31 (40.8)	33 (45.2)	23 (46.0)	56 (45.5)
채점에 시간과 노력이 많이 필요해서	14 (20.6)	5 (20.8)	7 (22.6)	10 (21.3)	16 (21.1)	17 (23.3)	9 (18.0)	26 (21.1)
수업에 이용할만한 자료를 얻지 못해서	3 (4.4)	1 (4.2)	2 (6.5)	2 (4.3)	4 (5.3)	2 (2.7)	4 (8.0)	6 (4.9)
너무 어려워하여 학생들이 과학을 기피할까봐	17 (25.0)	8 (33.3)	9 (29.0)	10 (21.3)	24 (31.6)	20 (27.4)	14 (28.0)	34 (27.6)
기타	1 (1.5)	0 (0.0)	0 (0.0)	0 (0.0)	1 (1.3)	1 (1.4)	0 (0.0)	1 (0.8)
합계	68 (100.0)	24 (100.0)	31 (100.0)	47 (100.0)	76 (100.0)	73 (100.0)	50 (100.0)	123 (100.0)

‘채점에 시간과 노력이 많이 필요하기 때문’이라고 응답한 비율이 21.1%로 나타났다. 학교급별로 비교하면, 초등학교의 경우 ‘채점 결과의 신뢰도가 낮아서’라고 응답한 비율이 48.5%로, 중·고등학교의 경우보다 다소 높게 나타났고, 중학교의 경우 ‘학생들이 서술형 문항을 너무 어려워하여 과학을 기피할 우려 때문’이라고 응답한 비율이 다른 학교급의 경우보다 다소 높게 나타났다. 교사의 성별로 비교하면, 남교사의 경우 ‘채점 결과의 신뢰도가 낮아서’라고 응답한 비율이 여교사의 경우보다 10% 이상 높게 나타났고, ‘학생들이 서술형 문항을 너무 어려워하여 과학을 기피할 우려 때문’이라고 응답한 비율은 여교사의 경우보다 약 10% 낮게 나타났다. 교사의 교직 경력별로 비교하면, 교직 경력 14년 이하인 교사들의 경우 ‘채점에 시간과 노력이 많이 필요하기 때문’이라고 응답한 비율이 교직 경력 15년 이상이 교사들의 경우보다 다소 높게 나타났는데, 교직 경력에 관계없이 ‘채점 결과의 신뢰도가 낮아서’라고 응답한 비율은 45~46%로 가장 높았다.

요약하면, 교사들은 서술형 문항에 대한 채점 결과를 신뢰하기 어려워져 서술형 문항을 제대로 개발하지 않는 것으로 나타났다. 채점에 시간과 노력이 많이 필요할 뿐만 아니라, 학생들에게 서술형

문항이 어려워져 오히려 과학을 기피할 수 있다는 우려도 있는 것으로 나타났다. 선행 연구에서 교사들은 서술형 문항이 직접 답을 문장으로 구성하기 때문에 학생들의 이해 수준이나 어려워하는 점을 구체적으로 파악할 수 있는 장점이 있다고 인식하고 있지만 채점의 편의성을 위해 다양하게 출제하지 못하거나(이미화, 2013), 고등정신 능력이나 창의적 문제 해결력을 측정하는 대신 단순정신능력에 초점을 두어 명확하게 답이 나타나는 문항을 출제하는 경향(최은주, 2011)이 있다고 보고된 바 있다. 교수학습과 평가의 유기적 관련성을 고려할 때, 평가에서 서술형 응답에 대한 어려움을 겪지 않도록 수업에서도 변화가 필요하다. 과학 수업에서 서술형 응답을 요구하는 기회가 많아지면 학생들이 말과 글로 응답하는 경험이 늘어 서술형 문항의 평가에서도 어려움이 줄어들 것이다. 서술형 문항의 출제에서 그 취지를 살려 고차적인 능력을 평가하는 문항을 출제할 수 있도록 수업과 연계하여 노력할 필요가 있다.

서술형 문항을 개발하여 실시하려고 노력한다고 응답한 교사를 대상으로 그렇게 하는 이유를 질문한 결과는 표 7과 같다.

서술형 문항을 개발하여 실시하는 이유로, ‘학생

표 7. 서술형 문항을 개발하여 실시하는 이유

단위: 명(%)

	학교급			성별		교직 경력		전체
	초	중	고	남	여	14년 이하	15년 이상	
학생들의 이해 수준을 파악할 수 있어서	89 (47.6)	43 (51.2)	41 (50.0)	59 (54.6)	114 (46.5)	115 (49.6)	58 (47.9)	173 (49.0)
학생들의 다양한 응답을 파악할 수 있어서	33 (17.6)	6 (7.1)	12 (14.6)	13 (12.0)	38 (15.5)	37 (15.9)	14 (11.6)	51 (14.4)
학생들의 글쓰기 능력을 파악할 수 있어서	5 (2.7)	0 (0.0)	0 (0.0)	1 (0.9)	4 (1.6)	4 (1.7)	1 (0.8)	5 (1.4)
교육청과 학교의 요구 사항이기 때문에	25 (13.4)	19 (22.6)	22 (26.8)	21 (19.4)	45 (18.4)	47 (20.3)	19 (15.7)	66 (18.7)
학생들의 학습 방법에 변화를 가져올 수 있어서	32 (17.1)	16 (19.0)	7 (8.5)	14 (13.0)	41 (16.7)	26 (11.2)	29 (24.0)	55 (15.6)
기타	3 (1.6)	0 (0.0)	0 (0.0)	0 (0.0)	3 (1.2)	3 (1.3)	0 (0.0)	3 (0.8)
합계	187 (100.0)	84 (100.0)	82 (100.0)	108 (100.0)	245 (100.0)	232 (100.0)	121 (100.0)	353 (100.0)

들의 이해 수준을 파악할 수 있어서'라고 응답한 비율이 49.0%로 가장 많았고, '교육청과 학교의 요구 사항이기 때문'이라고 응답한 비율이 18.7%로 그 다음으로 많았으며, '학생들의 학습 방법에 변화를 가져올 수 있어서', '학생들의 다양한 응답을 파악할 수 있어서' 등으로 응답한 비율은 약 15%였다. 한편 중·고등학교 교사들은 '학생들의 글쓰기 능력을 파악할 수 있어서'라고 응답한 경우가 없었고, 5명의 초등학교 교사들만 그렇게 응답하였다. 학교급별로 비교하면 중·고등학교의 경우 '학생들의 이해 수준을 파악할 수 있어서'라고 응답한 비율이 초등학교의 경우보다 다소 높게 나타났고, '교육청과 학교의 요구 사항이기 때문'이라고 응답한 비율은 학교급이 올라감에 따라 다소 증가하는 경향이 있다. 고등학교의 경우 '학생들의 학습 방법에 변화를 가져올 수 있어서'라고 응답한 비율은 초·중학교의 경우보다 낮게 나타났는데, 이는 대학 입시를 대비한 학습 중심으로 고등학교 과학 수업이 운영되는 것과 관련 있다고 볼 수 있다. 교사의 성별로 비교하면 남교사의 경우 '학생들의 이해 수준을 파악할 수 있어서'라고 응답한 비율이 여교사의 경우보다 다소 높았고, '학생들의 다양한 응답을 파악할 수 있어서', '학생들의 학습 방법에 변화를 가져올 수 있어서' 등의 응답 비율은 다소 낮았다. 교직 경력별로 비교하면 교직 경력 14년 이하인 교사들의 경우 '교육청과 학교의 요구 사항이기 때문'이라고 응답한 비율이 교직 경력 15년 이상인 교사들의 경우보다 높게 나타났고, '학생들의 학습 방법에 변

화를 가져올 수 있어서'라고 응답한 비율은 10% 이상 낮게 나타났다.

교사들은 서술형 문항을 개발하여 실시함으로써 학생들의 이해 수준을 파악할 수 있음을 잘 인식하고 있다. 그런데 고등학교에서는 대학 입시라는 당면한 목표와 서술형 문항이 가지는 장점 사이에 갈등하고 있음을 짐작해볼 수 있다. 서술형 평가가 학생들의 학습에 끼치는 장점이 결국 대학 입시에 긍정적 영향을 줄 수 있다는 인식이 학교 현장에 정착될 수 있도록 여러 측면에서의 노력이 필요하다.

### 5. 수행평가의 실시 현황

수행평가 실시 방법을 두 가지 형태로 제시하고 응답을 요구하였다. 수행평가를 형식적으로 실시하여 실제로는 지필평가를 통해서 학생의 능력을 대부분 변별하는 경우와 수행평가의 실시 목적에 맞게 실시하여 지필평가를 통하여 평가하지 못하는 능력을 변별하려고 노력하는 경우로 구분하여 응답하게 하였다. 수행평가 실시 방법에 대한 교사들의 응답 결과는 표 8과 같다.

수행평가 실시 목적에 맞게 수행하는 비율(70.0%)이 형식적으로 실시하는 비율(30.0%)보다 더 높게 나타났다. 수행평가 실시 방법 중 수행평가의 실시 목적에 맞게 실시하는 비율을 학교급별로 비교하면 초등학교 69.3%, 중학교 79.6%, 고등학교 62.5%이었다. 학교급에 따른 경향성은 나타나지 않았으나 중학교가 수행평가 실시 목적에 맞

표 8. 과학과 수행평가 실시 방법

단위: 명(%)

	학교급			성별		교직 경력		전체
	초	중	고	남	여	14년 이하	15년 이상	
형식적으로 실시. 실제로는 지필평가를 통해서 학생의 능력을 변별	81 (30.7)	22 (20.4)	42 (37.5)	56 (36.1)	89 (27.1)	103 (33.4)	42 (23.9)	145 (30.0)
실시 목적에 맞게 실시. 지필평가를 통하여 평가하지 못하는 능력을 변별하려고 노력	183 (69.3)	86 (79.6)	70 (62.5)	99 (63.9)	240 (72.9)	205 (66.6)	134 (76.1)	339 (70.0)
합계	264 (100.0)	108 (100.0)	112 (100.0)	155 (100.0)	329 (100.0)	308 (100.0)	176 (100.0)	484 (100.0)

게 실시하는 비율이 가장 높았으며, 고등학교가 가장 낮았다. 수행평가의 실시 목적에 맞게 실시하는 비율을 교사의 성별로 비교하면, 남교사 63.9%, 여교사 72.9%로, 여교사가 남교사에 비해 수행평가를 실시 목적에 맞게 실시하는 비율이 더 높았다. 수행평가의 실시 목적에 맞게 실시하는 비율을 교직 경력별로 비교하면 교직 경력 14년 이하인 교사들은 66.6%, 교직 경력 15년 이상인 교사들은 76.1%로, 교직 경력이 많은 교사들이 수행평가의 실시 목적에 맞게 실시하는 비율이 더 높았다.

지필평가와 수행평가 실시 비율을 조사한 결과(표 2)에서 고등학교는 수행평가 실시 비율이 가장 낮았는데, 실시 목적에 맞게 실시하는 비율도 가장 낮게 나타났다. 이는 대학교 입시 준비에 대한 영향과 관련지어 해석된다. 하지만 대학 입시에서 내신 성적 위주의 수시 모집 전형이나 대학수학능력시험 위주의 정시 모집 전형 모두 내실 있는 학교 교육에 근거하여 학생을 선발하는 것이며, 수행평가는 학생들의 지식이나 기능을 직접 관찰하며 평가하여 과학과 교육과정의 목표에서 지향하는 탐구능력의 파악을 위해 효과적으로 이용할 수 있다는 점에서 고등학교에서도 수행평가를 실시 목적에 맞게 실시하기 위한 노력을 더 기울일 필요가 있다.

수행평가를 실시하지 않거나 형식적으로 실시한다고 응답한 교사를 대상으로 수행평가를 실시하지 않거나 형식적으로 실시하는 이유를 조사하였으며,

응답 결과는 표 9와 같다.

수행평가를 실시하지 않거나 형식적으로 실시하는 이유로, '다양한 영역을 평가할 수 없고 수행과 관련되어서만 평가하기 때문'이라고 응답한 비율은 46.2%로 가장 많았고, '계획에 맞춰 실시하려면 시간과 노력이 많이 필요하기 때문'이라고 응답한 비율이 39.3%이었다. '평가 목표에 정확하게 맞춰 출제하는 능력이 부족하기 때문'이라고 응답한 비율이 6.9%로 가장 적었다. 학교급별로 나타내면, '다양한 영역을 평가할 수 없고 수행과 관련되어서만 평가하기 때문'이라고 응답한 비율은 초등학교가 48.1%, 중학교가 31.8%, 고등학교가 50.0%이었다. '계획에 맞춰 실시하려면 시간과 노력이 많이 필요하기 때문'이라고 응답한 비율은 초등학교가 42.0%, 중학교가 40.9%, 고등학교가 33.3%이었다. '평가 목표에 정확하게 맞춰 출제하는 능력이 부족하기 때문'이라고 응답한 비율은 초등학교가 7.4%, 중학교가 9.1%, 고등학교가 4.8%이었다. 초등학교와 고등학교에서는 '다양한 영역을 평가할 수 없고 수행과 관련되어서만 평가하기 때문'이라고 응답한 비율이 가장 많았으며, 중학교에서는 '계획에 맞춰 실시하려면 시간과 노력이 많이 필요하기 때문'이라고 응답한 비율이 가장 많았다. 교사의 성별로 비교하면, 남교사의 경우 '다양한 영역을 평가할 수 없고 수행과 관련되어서만 평가하기 때문'이라고 응답한 비율이 가장 높았고 여교사의 경우 이 이유뿐만 아니라 '계획에 맞춰 실시하려면 시간과 노력

표 9. 과학과 수행평가를 실시하지 않거나 형식적으로 실시하는 이유

단위: 명(%)

	학교급			성별		교직 경력		전체
	초	중	고	남	여	14년 이하	15년 이상	
다양한 영역을 평가할 수 없고 수행과 관련되어서만 평가하므로	39 (48.1)	7 (31.8)	21 (50.0)	29 (51.8)	38 (42.7)	47 (45.6)	20 (47.6)	67 (46.2)
계획에 맞춰 실시하려면 시간과 노력이 많이 필요해서	34 (42.0)	9 (40.9)	14 (33.3)	19 (33.9)	38 (42.7)	43 (41.7)	14 (33.3)	57 (39.3)
평가 목표에 정확하게 맞춰 출제하는 능력이 부족해서	6 (7.4)	2 (9.1)	2 (4.8)	2 (3.6)	8 (9.0)	6 (5.8)	4 (9.5)	10 (6.9)
기타	2 (2.5)	4 (18.2)	5 (11.9)	6 (10.7)	5 (5.6)	7 (6.8)	4 (9.5)	11 (7.6)
합계	81 (100.0)	22 (100.0)	42 (100.0)	56 (100.0)	89 (100.0)	103 (100.0)	42 (100.0)	145 (100.0)

이 많이 필요하기 때문'이라고 응답한 비율이 모두 42.7%로 나타났다. 교직 경력별로 나타내면 교직 경력과 상관없이 '다양한 영역을 평가할 수 없고 수행과 관련되어서만 평가하기 때문'이라고 응답한 비율이 가장 높았다.

수행평가를 실시하지 않거나 형식적으로 실시하는 이유를 요약하면, 다양한 영역을 평가할 수 없고 수행과 관련되어서만 평가하기 때문이라고 응답한 경우는 수행평가가 지필평가에서 평가하기 어려운 내용을 중심으로 학생들의 실제적인 행위를 직접 관찰하며 검사하는 평가(성태제, 2010)라는 측면을 고려하지 못하고 있는 것으로 보인다. 수행평가가 지필평가를 보완하고 교육과정 목표에 도달하기 위해 필요한 평가라는 점을 모든 학교급의 교사들이 고려할 필요가 있다. 그리고 수행평가를 계획에 맞춰 실시하려면 시간과 노력이 많이 필요하기 때문에 실시하지 않거나 형식적으로 실시한다는 응답은 교사 측면이 아니라 학교 행정이나 교육 당국에서 고려해야 할 사항이다. 교육의 본질적 목적을 달성하기 위해 교사들이 교육 평가에 좀 더 시간을 투자할 수 있도록 행정적 배려를 할 필요가 있다.

수행평가를 실시 목적에 맞게 실시하며, 지필평가를 통하여 평가하지 못하는 능력을 변별하려고 노력한다고 응답한 교사를 대상으로 수행평가를 실

시 목적에 맞게 실시하는 이유를 조사하였으며, 응답 결과는 표 10과 같다.

수행평가를 실시 목적에 맞게 실시하는 이유로, '학생들의 다양한 능력별로 평가하기 위해서'라고 응답한 비율이 67.7%로 가장 많았고, '평가 결과에 근거하여 학생 면담이나 수업 자료로 활용하기 위해'라고 응답한 비율이 26.1%로 그 다음으로 많았다. '성적 분포를 다양하게 하기 위해'라고 응답한 비율과 '장학이나 감사 자료에 대비하기 위해'라고 응답한 비율은 각각 2.1%, 3.5%로 적었다. 학교급별로 비교하면 초·중·고등학교 모두 '학생들의 다양한 능력별로 평가하기 위해서'라고 응답한 비율이 가장 많았는데, 고등학교가 세 학교급 중에서 비율이 가장 높았다(78.9%). '평가 결과에 근거하여 학생 면담이나 수업 자료로 활용하기 위해'라고 응답한 비율은 세 학교급 중 초등학교가 가장 높았다(32.2%). 교사의 성별로 비교하면 '학생들의 다양한 능력별로 평가하기 위해서'라고 응답한 비율이 남교사의 경우 69.7%, 여교사의 경우 66.9%로 가장 많았고, '평가 결과에 근거하여 학생 면담이나 수업 자료로 활용하기 위해'라고 응답한 비율이 남교사 25.3%, 여교사 26.4%로 그 다음으로 많았다. 교사의 성별에 따른 응답 결과에 별다른 차이가 나타나지 않았다. 교직 경력별로 비교하면 '학생들의 다양한 능력별로 평가하기 위해서'라고 응답한 비

표 10. 과학과 수행평가를 실시 목적에 맞게 실시하는 이유

단위: 명(%)

	학교급			성별		교직 경력		전체
	초	중	고	남	여	14년 이하	15년 이상	
평가 결과에 근거하여 학생 면담이나 수업 자료로 활용하기 위해	59 (32.2)	18 (20.7)	12 (16.9)	25 (25.3)	64 (26.4)	48 (23.3)	41 (30.4)	89 (26.1)
성적 분포를 다양하게 하기 위해	4 (2.2)	2 (2.3)	1 (1.4)	1 (1.0)	6 (2.5)	5 (2.4)	2 (1.5)	7 (2.1)
장학이나 감사 자료에 대비하기 위해	4 (2.2)	6 (6.9)	2 (2.8)	4 (4.0)	8 (3.3)	9 (4.4)	3 (2.2)	12 (3.5)
학생들의 다양한 능력별로 평가하기 위해	115 (62.8)	60 (69.0)	56 (78.9)	69 (69.7)	162 (66.9)	143 (69.4)	88 (65.2)	231 (67.7)
기타	1 (0.5)	1 (1.1)	0 (0.0)	0 (0.0)	2 (0.8)	1 (0.5)	1 (0.7)	2 (0.6)
합계	183 (100.0)	87 (100.0)	71 (100.0)	99 (100.0)	242 (100.0)	206 (100.0)	135 (100.0)	341 (100.0)

율이 14년 이하 경력에서 69.4%, 15년 이상 경력에서 65.2%로 가장 많았고, ‘평가 결과에 근거하여 학생 면담이나 수업 자료로 활용하기 위해’라고 응답한 비율이 14년 이하 경력에서 23.3%, 15년 이상 경력에서 30.4%로 그 다음으로 많았다.

수행평가를 실시 목적에 맞게 실시하는 이유를 종합하면, 교사들은 수행평가를 목적에 맞게 실시하는 가장 중요한 이유가 학생들의 다양한 능력을 평가하기 위함이었다. 수행평가는 지필평가와 달리 수행 과정에서 평가하기 때문에 지필평가에서 측정할 수 없는 능력을 측정할 수 있어서 다양한 능력의 평가 목적으로 적절하다. 그런데 고등학교에서는 학생들의 다양한 능력을 평가하는 데 수행평가를 주로 이용하는 반면, 이를 학생 면담이나 수업 자료로 활용하는 비율은 세 학교급에서 가장 낮았다. 고등학교에서는 학생들의 부족한 능력을 파악하고 이를 다시 수업에 피드백하여 적용하는 노력을 기울일 필요가 있다.

수행평가를 실시할 때의 어려운 점에 대한 교사들의 응답 결과는 표 11과 같다.

전체적으로 수행평가를 실시할 때의 어려운 점 중 ‘학생들이 제출한 자료를 보는 데 시간과 노력이 많이 든다’라고 응답한 비율이 가장 높았다(40.3%). 그 다음으로 ‘평가의 타당도나 신뢰도를 확보하기 어렵다’는 응답이 29.0%, ‘다양한 형태의

평가를 설계하기 어렵다’는 응답이 22.7%이었으며, ‘별다른 어려움이 없다’는 응답은 7.3%로 낮았다. 학교급별로 비교하면 ‘학생들이 제출한 자료를 보는 데 시간과 노력이 많이 든다’가 모든 학교급에서 가장 높았는데, 초등학교가 31.9%, 중학교가 49.5%, 고등학교가 50.0%이었다. 초등학교와 중학교는 ‘다양한 형태의 평가를 설계하기 어렵다’와 ‘평가의 타당도나 신뢰도를 확보하기 어렵다’에 대한 응답 비율이 유사한 반면, 고등학교의 경우는 ‘다양한 형태의 평가를 실시하기 어렵다’는 12.5%이고 ‘평가의 타당도나 신뢰도를 확보하기 어렵다’가 35.6%로 차이가 컸으며, 특히 평가의 타당도나 신뢰도 확보의 어려움은 세 학교급에서 그 비율이 가장 높았다. 그리고 ‘별다른 어려움이 없다’고 응답한 비율이 초등학교가 11.9%, 중학교가 2.1%, 고등학교가 1.9%로 초등학교가 가장 높았다. 표 2의 지필평가와 수행평가 비율에서 세 학교급 중 초등학교가 수행평가 실시 비율이 가장 높았는데, 수행평가를 가장 많이 실시하는 초등학교에서는 수행평가 실시와 관련된 경험적 지식이 많이 축적되어 별다른 어려움이 없다고 응답한 비율도 세 학교급에서 가장 높게 나타난 것으로 추측된다. 교사의 성별로 비교하면, ‘학생들이 제출한 자료를 보는 데 시간과 노력이 많이 든다’라는 응답의 비율이 남·여 교사 모두 가장 높았는데, 남교사가 49.6%, 여교사가 35.9%로, 그 차이가 가장 컸다. 즉, 남교사는

표 11. 과학과 수행평가를 실시할 때 어려운 점

단위: 명(%)

	학교급			성별		교직 경력		전체
	초	중	고	남	여	14년 이하	15년 이상	
학생들이 제출한 자료를 보는 데 시간과 노력이 많이 든다	72 (31.9)	48 (49.5)	52 (50.0)	68 (49.6)	104 (35.9)	104 (38.5)	68 (43.3)	172 (40.3)
다양한 형태의 평가를 설계하기 어렵다	60 (26.5)	24 (24.7)	13 (12.5)	23 (16.8)	74 (25.5)	66 (24.4)	31 (19.7)	97 (22.7)
평가의 타당도나 신뢰도를 확보하기 어렵다	66 (29.2)	21 (21.6)	37 (35.6)	35 (25.5)	89 (30.7)	78 (28.9)	46 (29.3)	124 (29.0)
별다른 어려움이 없다	27 (11.9)	2 (2.1)	2 (1.9)	10 (7.3)	21 (7.2)	20 (7.4)	11 (7.0)	31 (7.3)
기타	1 (.4)	2 (2.1)	0 (.0)	1 (.7)	2 (.7)	2 (.7)	1 (.6)	3 (.7)
합계	226 (100.0)	97 (100.0)	104 (100.0)	137 (100.0)	290 (100.0)	270 (100.0)	157 (100.0)	427 (100.0)

여교사에 비해 학생들이 제출한 자료를 보며 평가하는 데 더 어려움을 겪는 것으로 보인다. 반면 ‘다양한 형태의 평가를 설계하기 어렵다’는 응답 비율이 남교사가 16.8%, 여교사가 25.5%로 여교사의 응답 비율이 남교사의 경우보다 높았다. 교직 경력 별로 나타내면 학교급이나 성별 비교와 달리 차이가 크게 나타나지 않았다.

요약하면, 수행평가를 실시할 때 가장 어려운 점으로 교사들은 ‘학생들이 제출한 자료를 보는 데 시간과 노력이 많이 든다’를 꼽았다. 수행평가는 지필평가에서 측정하기 어려운 점을 수행을 통해 측정이 가능하다는 장점이 있지만 그만큼 평가에 시간과 노력이 많이 들어야 하는 어려운 점이 있다. 이러한 어려움을 해결하기 위한 방안을 모색할 필요가 있는데, 컴퓨터 기반 과학 수행평가 문항을 활용함으로써 즉각적 피드백이 가능한 시스템(최혁준, 2013)의 도움을 받는 것도 하나의 방안이다. 한편 고등학교에서는 ‘평가의 타당도나 신뢰도를 확보하기 어렵다’라고 응답한 비율이 초·중학교에 비해 높았다. 이는 고등학교에서는 다른 학교급에 비해 대학 입시를 위한 공정한 내신 성적 산출을 더 중요하게 여기기 때문에 나타나는 현상으로 추측된다. 수행평가는 수행 과정과 수행을 통해 작성한 결과물을 평가하기 때문에 지필평가에 비해 상대적으로 평가의 신뢰도를 확보하기 어려운 점은 있지만 성적 산출이라는 일부 목적을 위해 과학과 교육과정의 중요한 목표 달성에 필요한 수행평가의 비율을 낮추는 것은 적절하지 않은 교육적 판단일 것이다.

#### IV. 결론 및 제언

본 연구에서는 초·중·고등학교 과학과 평가의 실태를 교사 대상의 설문 조사를 통해 고찰하고자 하였다. 설문 조사 결과, 학교급이 올라갈수록 지필평가의 비율이 늘고 수행평가의 비율이 줄어들고 있었다. 교사들은 학생들의 다양한 능력을 평가하기 위해 수행평가를 실시 목적에 맞게 실시하는 경우

도 있지만, 수행과 관련되어서만 평가하고 시간과 노력이 많이 필요해서 수행평가를 형식적으로 실시하는 경우도 있었으며, 학생들이 제출한 자료를 보는 데 시간과 노력이 많이 들고, 평가의 타당도나 신뢰도를 확보하기 어려워서 수행평가 실시에 어려움을 겪는다고 하였다. 과학과 지필 평가에서 서술형 문항의 비율이 선다형 문항에 비해 낮게 나타났다. 교사들은 서술형 문항을 개발하여 실시함으로써 학생들의 이해 수준을 파악할 수 있음을 잘 인식하고 있는 경우도 있지만, 서술형 문항에 대한 채점 결과를 신뢰하기 어려워서 서술형 문항을 제대로 개발하지 않는 것으로 나타났다. 그리고 과학과 지필평가에서 지식 영역의 출제 비율이 탐구 영역의 경우보다 높게 나타났다.

설문 조사 결과에 근거하여, 우리나라 초·중·고등학교 과학과 평가의 개선 방안을 제안하면 다음과 같다.

첫째, 수행평가 최소 실시 비율을 준수하고 수행평가의 목적에 맞게 실시하여 과학과 교육과정의 목표에 해당하는 탐구 능력의 신장, 창의적·비판적 사고력 향상을 성취해야 할 것이다. 우리나라에서는 1990년대부터 학생들의 문제 해결 능력, 비판적 사고력 등을 함양하기 위해 수행평가를 도입하였는데, 과학과에서는 학생들의 과학 탐구 능력, 실험 기능, 의사소통 능력, 과학적 태도, 과학의 본성에 대한 인식 등을 평가하기 위해 수행평가를 활용할 필요가 있다. 그리고 2007년 개정 과학과 교육과정에서부터 과학 글쓰기와 토론을 강조하고 있는데, 이를 평가하기 위해서도 수행평가를 활용할 필요가 있다. 그런데 교사 대상 설문 조사 결과에서 수행평가 실시 비율이 초등학교에서 중학교로 올라가면서 약 10%, 중학교에서 고등학교로 올라가면서 약 4% 감소하는 등 학교급이 올라갈수록 지필평가의 비율이 늘고 수행평가의 비율이 줄어들고 있는 것으로 나타났다. 이러한 현상은 교육과정의 목표 구현 측면에서 바람직하지 않으므로 각 학교급에서는 수행평가의 최소 실시 비율을 권장하거나 의무화할 필요가 있다. 그리고 수행평가가 그 목적에 맞게 실시되도록 실험 및 평가 업무 협조를 위한 보조 교사를 확보하는 등 행정적 지원을 할 필요가

있다. 한편 대학교 입시 준비로 고등학교에서 수행평가 실시 비율이 줄어드는 현상을 막기 위해 대학 입시에 수행평가를 반영하는 방안을 고려할 필요가 있고, 학교 현장에서 실시할 수 있는 다양한 수행평가 방법에 대한 교사 연수, 수행평가 관련 자료의 개발 및 공유 등이 필요하다.

둘째, 지필평가에서 서술형 문항의 출제 비율을 확대하여 학생들의 다양한 능력을 판별하는 데 도움을 받고, 이를 바탕으로 학생들의 고차원적 사고력의 향상을 성취해야 할 것이다. 서술형 문항의 경우 학생들이 답을 문장으로 직접 구성하기 때문에 학생들의 이해 수준이나 어려워하는 점을 구체적으로 파악할 수 있는 장점이 있고, 선다형 문항에 비해 상대적으로 학생들의 문제 해결력을 함양하고 논리적 표현력을 향상시킬 수 있다. 설문 조사 결과를 살펴보면 과학과 지필평가에서 서술형 문항의 비율이 선다형 문항에 비해 낮고, 교사의 성별이나 경력에 따라 선다형과 서술형 문항 출제 비율에 차이가 크지 않았다. 그리고 교육청과 학교의 요구 사항 때문에 서술형 문항을 개발한다는 응답이 중·고등학교에서 높게 나타났다. 또한 과학 교사들은 채점 결과에 대한 신뢰 부족, 채점 시간과 노력의 필요, 학생들이 과학이 어렵다는 인식 확산에 대한 우려 등을 이유로 서술형 문항 출제를 꺼리고 있었다. 과학과에서 서술형 평가를 강조할 필요가 있으므로 서술형 문항 출제 비율이 안정적으로 유지될 때까지 선다형, 단답형, 서술형으로 구분하여 문항 출제 비율을 권장할 필요가 있다. 그런데 단순히 출제 비율을 권장하거나 강제하는 것으로는 부족하고 실제 서술형 문항 출제 비율을 늘릴 수 있는 행·재정적 협조가 필요하다. 교사들 대상으로 과학과 서술형 문항의 개발 및 채점에 대한 실제적인 연수를 실시하고, 과학과 실험 조교를 학교당 1명씩 배정하여 실험 보조뿐만 아니라 과학과 평가 관련 업무를 지원하게 하는 것도 좋은 방안이라고 생각된다.

셋째, 지필평가에서 탐구 영역의 문항 출제 비율을 확대하여 학생들의 과학 개념 이해뿐만 아니라 과학 탐구 능력의 향상을 성취해야 할 것이다. 설문 조사 결과를 살펴보면 지필평가의 탐구 영역 출

제 비율이 40% 이상으로 나타났는데, 선행 연구에서는 이보다 낮았다. 문항의 행동 영역별 분류에서 교사와 연구자의 판단에 차이가 있다는 결과를 고려할 때, 교사들이 실제 출제하는 문항이 타당도 높은 탐구 영역의 문항이 되도록 교사 연수를 확대할 필요가 있다. 한편 탐구 영역의 문항 비중을 높이기 위해 교실 수업보다 과학실에서 탐구 중심 수업 및 실험 수업을 하도록 권장하고, 이를 위해 과학실을 확충하고 실험 지원을 하는 조교를 배치하는 등 실험실 수업의 활성화를 위한 교수·학습 지원 환경을 개선해야 할 것이다.

## 참고 문헌

- 교육과학기술부 (2009). 2009 개정 교육과정-과학과 교육과정. 교육과학기술부 고시 제2009-41호.
- 교육과학기술부 (2011). 창의·인성교육 강화를 위한 중등학교 학사관리 선진화 방안. 교과부 보도 자료.
- 김석우 (2007). 고등학교 과학과 수행평가 실태분석 및 개선 방안. *교육평가연구*, 20(4), 53-73.
- 김신영 (2002). 현장교사의 평가전문성 연구. *교육평가연구*, 15(1), 67-85.
- 김신영 (2012). 교사의 학생평가와 성취평가제. *교육평가연구*, 25(4), 655-677.
- 김윤희, 윤기순, 권덕기 (2010). Bloom의 신 교육 목표분류에 기초한 중학교 생물 영역 총괄평가 문항의 목표 분석. *과학교육연구지*, 34(1), 164-174.
- 박현주 (2012). 중학교 과학과 총괄 지필평가 문항의 목표 분석. *과학교육연구지*, 36(2), 293-302.
- 서정관 (2008). 고등학교 과학 과목의 수행평가 운영 실태 조사 : 실험·실습법을 중심으로. 경상대학교 교육대학원 석사학위 논문.
- 성태제 (2010). *현대교육평가*(3판). 서울: 학지사.
- 손정숙 (2011). 중등학교 실험실습 수행평가에 대한 과학 교사들의 인식 조사. *한국교원대학교 교*

## 국문 요약

이 연구에서는 초·중·일반계 고등학교 과학과 평가의 실태를 고찰하고자 하였다. 이를 위해 초등학교 60개교, 중학교 40개교, 일반계 고등학교 20개교를 표집하여 설문지를 발송하였고, 초등학교 교사 292명, 중학교 교사 110명, 고등학교 교사 116명, 총 518명의 응답 결과를 얻었다. 설문 조사 결과, 학교급이 올라갈수록 지필평가의 비율이 높고 수행평가의 비율이 줄어들고 있었다. 그리고 과학과 지필평가에서 서술형 문항의 비율이 선다형 문항에 비해 낮았고, 중·고등학교는 초등학교와 비교할 때 선다형 문항의 비율이 높았다. 또한 지필평가에서 지식 영역과 탐구 영역의 출제 비율을 살펴본 결과, 학교급, 성별, 교직 경력과 상관없이 모두 지식 영역이 탐구 영역보다 출제 비율이 더 높았다. 채점 시간을 고려하여 서술형 문항을 제한된 범위에서 실시하는 경우가 많은데, 이러한 경향은 초등학교보다 중·고등학교에서 더 많이 나타났다. 교사들은 서술형 문항을 개발하여 실시함으로써 학생들의 이해 수준을 파악할 수 있음을 잘 인식하고 있는 경우도 있지만, 서술형 문항에 대한 채점 결과를 신뢰하기 어려워서 서술형 문항을 제대로 개발하지 않는 것으로 나타났다. 한편 고등학교의 경우 수행평가를 실시 목적에 맞게 실시하는 비율이 가장 낮게 나타났다. 교사들은 학생들의 다양한 능력을 평가하기 위해 수행평가를 실시 목적에 맞게 실시하는 경우도 있지만, 수행과 관련되어서만 평가하고 시간과 노력이 많이 필요해서 수행평가를 형식적으로 실시하는 경우도 있었다. 학교 현장에서 수행평가가 실시 목적에 맞게 실시되고 서술형 평가가 정착되며 탐구 영역의 출제가 원활하게 이루어지도록 교사 연수, 실험과 평가 지원을 위한 조교 배치 등 행·재정적 지원이 필요하다.

주요어: 과학교육 평가, 수행평가, 지필평가, 서답형 문항

- 육대학원 석사학위 논문.
- 심애순 (2008). 중학교 과학과 서술형 평가에 대한 교사와 학생의 인식 조사 및 실태 분석. 한국교원대학교 교육대학원 석사학위 논문.
- 심재호, 노은희, 김명화 (2012). 국내외 학업성취도 평가의 서답형 문항 특성과 채점 체제 고찰. 열린교육연구, 20(4), 159-185.
- 양일호, 나중철, 임성만, 임재근, 최현동 (2008). Klopfer의 교육 목표 분류 체계에 의한 초등학교 과학과 지필평가 문항 분석: 5학년 1 학기를 중심으로. 초등과학교육, 27(3), 221-232.
- 오현석, 이기영 (2006). 현행 중등학교 과학 교과 지필평가 문항분석. 교육과정평가연구, 9(1), 405-424.
- 이미화 (2013). 고등학교 화학과 서술형 평가에 대한 현황 분석 연구. 한국교원대학교 석사학위 논문.
- 최은주 (2011). 중학교 과학과 서술형 평가에 대한 문항의 특성과 교사와 학생 인식 조사. 한국교원대학교 석사학위논문.
- 최혁준 (2013). 컴퓨터 기반 과학 수행평가 문항 개발 : 전류계와 전압계 회로 연결 실험. 과학교육연구지, 37(2), 348-358.
- American Federation of Teachers, National Council on Measurement in Education, & National Education Association(AFT, NCME, & NEA) (1990). Standards for teacher competence in educational assessment of students. Educational Measurement: Issues and Practice, 9(4), 30-32.
- Stiggins, R. J. (2002). Assessment crisis: The absence of assessment for learning. Phi Delta Kappan, 83(10), 758-765.