

중학교 과학에서 서술형 평가의 실태 분석: 문항유형, 단원, 평가목표를 중심으로

노은실 · 김용진*

경상대학교

Analysis of the Descriptive Evaluation's Status in Middle School Science: Focused on the Item's Type, Subject Unit, and Evaluation Object

Eun Sill Noh · Yong-Jin Kim*

Gyeongsang National University

Abstract : In recent years, descriptive evaluation has been gaining importance at school. The purpose of this study is to analyse descriptive evaluation items which have been made by school teachers. 300 items for regular examinations are collected from 7 middle schools (grades 1 through 3) in Gyeongnam Province. Distribution and types of items are examined to compare educational objectives of standard achievements and those of evaluation according to Bloom's revised taxonomy of educational objectives. The result shows that there exist a majority of shared subject units for descriptive evaluation. Huge difference of frequency is also found among subject units and standard achievements. Less than 65% of evaluation items are descriptive and the rest is simple completion or short answers. In addition, it reveals only 40.3% of agreement between educational objectives of standard achievements and those of evaluation items. The interviews with 27 teachers indicate that lack of ability to develop proper items and grading are the major obstacles. In conclusion, systematic training courses are to be provided in order to resolve issues over descriptive evaluation.

keywords : descriptive evaluation, middle school science, item's type, subject unit, assessment objective, taxonomy of educational objectives.

I. 서론

최근 학교의 수업활동은 글로벌 지식 기반 사회가 필요로 하는 창의성과 인성을 고루 갖춘 인재 양성에 초점을 두고, 지식 전달 중심에서 지식 창조의 방향으로 변화를 추구하고 있다. 평가 영역에서도 학생의 지

적 수준이나 학습 능력에 대한 결과 중심 평가에서 학생의 다양한 학습 과정에서 일어난 경험과 체험활동을 통한 지식이나 정보를 재구성하는 등의 성장 중심 평가로 변화시키려는 정책적 의도들도 지속적으로 제시되었다(KICE, 2016; Sin, An & Kim, 2017). 학생평가는 학습활동을 통해 학생들이 이룬 성과를 측

*교신저자: 김용진 (gnuedu@nate.com)

**이 논문은 노은실의 2017년도 석사 학위논문의 데이터를 활용하여 재구성하였음.

***2018년 6월 21일 접수, 2018년 8월 27일 수정원고 접수, 2018년 8월 27일 채택

<http://dx.doi.org/10.21796/jse.2018.42.2.198>

정하는 것으로서 교과별 교육과정에 따른 교수·학습 활동의 효과를 확인할 수 있는 수단이면서 학생의 성장을 지원하는 교사의 피드백 활동에 필요한 정보를 제공하는 과정이다(Kim, 2015; Paik & Ryu, 2014). 기존의 평가제도는 학생들에게 서열에 대한 부담을 가중시키고 경쟁심을 조장하여 협동학습을 통한 나눔과 배려의 학습경험을 공유하지 못하게 함으로써 적성과 소질에 따른 다양한 교육과정을 선택하고 운영하는 데 어려움이 많았다. 교사는 학생이 무엇을 얼마만큼 알고 있는지와 무엇을 할 수 있는가를 판정할 수 있어야 하는데, 이는 학생평가가 교과 성취기준에 근거한 평가, 즉 준거참조평가(절대평가)에 근거해야 한다는 것을 의미한다(Kang, 2014).

학교 현장에서 평가 영역의 본격적인 변화는 2012년부터 성취평가제의 도입과 함께 서술형 평가가 확대 실시된 것을 가장 큰 핵심으로 꼽을 수 있다(Ji, 2011; Kim, 2015). 전국 17개 시·도 교육청에서는 제7차 교육과정 적용 이후부터 수행평가의 점진적 확대를 지속적으로 추진하였는데, 지난 2012년부터는 ‘학업성적관리지침’을 통해 서술형 평가의 의무 실시도 명료화 하였다. 2012년에는 국어, 도덕, 사회(역사 포함), 수학, 과학, 영어 등 7개 교과의 지필평가 또는 수행평가에 서술형평가를 반드시 30% 이상 포함하도록 하는 것에서 출발하여, 2014년에는 지필평가에 환산총점의 30% 이상 반영하도록 하였다. 2015년에는 지필평가에 환산총점의 50% 이상을 반영하게 하였으며 최근에는 80% 이상을 반영하도록 의무화 하는 지역도 많아졌다. 또한 국가수준의 학업성취도 평가에서도 서술형 평가 문항이 출제되고 있으며, 이에 대한 결과 분석으로 학생들의 학습 능력 발달의 변화를 추정하고 있다(Dong *et al.*, 2017).

서술형 평가를 성취평가제의 도입과 함께 지침에 반영하여 수행평가와 함께 서술형 평가를 강조하여 왔지만 단위 학교별 서술형 평가 실시 현황은 다양하며, 그 질적 향상을 위해 기술인 노력도 지속성은 물론 효율성 면에서도 많이 부족하다(Kang, 2014; Lee, 2012). 학교 현장에서 서술형 평가와 수행평가가 실질

적인 효과를 거두지 못하고 있는 것은 지금까지 서술형 평가나 수행평가가 본질적으로 갖추어야 할 특성에 초점을 두지 않은 채, 실시 비율 확대와 같은 형식적인 부분에만 치중해 왔기 때문으로 보인다. 최근 성취기준에 부합하는 타당도 높은 문항 개발과 객관적이고 공정한 채점을 위한 교사들의 평가 전문성 신장은 교육계에서 가장 시급하고 중요한 해결 과제로 부상하였다(Lee & Jeong, 2014). 물론, 외형적으로는 각 시·도 교육청별로 ‘교과별 평가 문항 예시 자료집’을 개발·보급하거나 평가 관련 교원 연수를 부분적으로 실시하는 등 교사들의 평가 전문성 신장을 위한 지속적인 노력을 계속해 왔다(BMCOE, 2010; KICE, 2016; Noh *et al.*, 2010). 그러나 자료집들의 대부분이 장학자료의 형태로 학년별·단원별로 3-4개의 평가 문항을 예시로 제시할 뿐이고, 문항 출제의 방법과 절차에 대한 실질적이고 구체적인 내용적 접근은 부족하다.

그동안 학교 현장에서의 서술형 평가에 대한 인식이나 문제점들에 대한 연구는 국어 교과(Kim, 2013; Lee, 2012), 수학 교과(Lee *et al.*, 2014; Kim, Cho & Joo, 2012) 및 지리 교과(Shin & Cho, 2008) 등에서 보고되어 왔다. 과학 교과에서 학교 현장에서 실행한 서술형 평가 문항의 유형과 특성 분석에 대한 연구는 고등학교(Lee & Jeong, 2014)와 초등학교(Choi & Paik, 2016)를 대상으로 일부 있으며, 중학교를 대상으로 한 연구는 국가수준 학업성취도 평가를 이용한 분석 연구(Lee *et al.*, 2016; Lee *et al.*, 2017) 이외에 학교 현장에서의 실태에 대한 연구도 매우 적다(Kim, Choi & Paek, 2015). 따라서 학교의 과학교육 현장에서 점차 강화되고 있는 서술형 평가에 대해서 관련 실태와 문제점을 파악하고 개선하기 위한 학문적 연구가 많이 이루어질 필요가 있다.

본 연구에서는 중학교 현장에서 2009개정 교육과정(MEST, 2009)에 따라 그동안 실제 시행된 과학교과의 서술형 평가 문항을 수집하고 분석하여 서술형 평가 문항에 대한 특성을 분석하고, 서술형 평가의 시행에 대한 개선 방안을 제시하여 과학교사들의 학생

평가 역량 제고에 기여하고자 한다. 이를 위해 서술형 평가 문항의 성취기준별 출제 빈도, 문항 유형별 출제 빈도, 성취기준의 교육목표와 비교한 평가목표의 적합성 등을 분석하고 출제자들과의 면담을 통해 분석 결과에 대한 원인을 조사하였다.

중학교 현장에서 출제된 과학교과의 서술형 평가 문항에 대한 구체적인 연구 문제는 다음과 같다.

첫째, 서술형 평가 문항의 유형별 분포는 어떠한가?

둘째, 단위(성취기준)별 서술형 문항의 출제 빈도는 어느 정도인가?

셋째, 서술형 평가 문항의 평가목표는 성취기준의 교육목표에 적합한가?

II. 연구 방법

1. 연구 대상

본 연구는 K도에 소재한 중학교에서 2009개정 교육과정(MEST, 2009)에 따라 최근 3년(2014-2016) 동안에 실시한 1, 2, 3학년 정기고사의 기출문제에서 서술형 평가문항을 대상으로 하였다. 연구를 위한 문항 사용에 동의한 교사들을 대상으로 고사 문항들을 수집하였다. 수집한 자료 중에서 중간고사 및 기말고사의 교사원안과 함께 이원목적분류표와 채점기준표

를 모두 제공한 7개교 27명의 과학교사가 출제한 고사 원안 중에서 서답형 평가 문항 300개를 연구 대상으로 하였다. 학년별 문항 수는 각각 100개씩이었다. 과학 교과의 단위 및 성취기준별 출제 빈도, 서술형 문항의 유형, 각 문항에 관련된 성취기준의 목표 분류와 각 문항에 대한 평가목표 분류 등을 통해 성취기준에 대한 평가문항의 적합 정도를 분석하였다. 분석 결과에 따른 구체적인 원인 조사를 위해 27명의 출제교사를 대상으로 설문조사와 면담을 진행하였다.

2. 연구 방법

학교 현장에서 시행된 정기고사의 서술형 평가 문항을 대상으로 단위(성취기준)에 대한 출제 빈도, 문항 유형, 문항별 성취기준에 대한 평가문항의 적합 정도를 분석하였다.

1) 문항 유형별 출제 빈도 분석

학교 현장에서 수집한 서술형 평가 문항들의 유형을 조사한 선행연구 결과(Noh *et al.*, 2010; Lee & Jeong, 2014)를 기초로 Table 1과 같이 문항 제시 방법과 응답 형태의 두 기준으로 나누어 문항 유형을 구분하였다. 최근 몇 년 동안 전국 시·도 교육청들이 고시한 학업성적관리지침에 의하면 단답형이나 완성형은 서술형 문항이 아니라고 명확하게 규정하고 있

Table 1. Types of descriptive evaluation items.

구분	문항 유형	내용
문항 제시	문항 단독형	추가로 제시하는 자료 없이 문항만 제시
	자료 제시형	답안 작성에 참고할 수 있는 자료를 문항과 함께 제시하는 문항 : 표/그래프, 그림/사진, 사례/설명, 지도, 복합 자료(2개 이상 결합)
응답 형태	단답형	간단한 어구나 문장, 기호, 숫자 등을 사용하여 비교적 간단한 형태의 반응을 요구하는 형식
	완성형	진술문의 일부분을 빼놓고 적합한 단어, 어구, 기호 등을 삽입하여 완전한 문장을 만들도록 하는 형식
	서술(논술)형	학생들이 직접 서술(구성)하는 형식

다. 그러나 실제 학교 현장에서는 여전히 혼재한 평가 문항들이 출제되고 있으며, 선다형 문항 형태가 아닌 문항 유형들을 서술형 평가에 포함하여 온 경우가 많다. 따라서 본 연구에서는 학교 현장의 실태를 기반으로 서술형 문항의 출제 형태를 분석하기에 적합하도록 서답형에 해당하는 단답형과 완성형을 서술형 문항 유형에 포함하여 분석하였다.

2) 단원 및 성취기준에 따른 출제 빈도 분석

수집한 서술형 평가 문항을 학년별로 분류하여 각각 100개의 문항(총 300개)에 대해 출제 단원 및 성취기준에 대한 출제 빈도를 분석하였다. 2009 개정 교육과정을 학교 현장에 적용함에 있어서는 교육과정에서 문서로 제시된 성취기준을 그대로 사용하는 것이 아니라 출제 교사가 학기 초에 수업과 평가를 계획하여 성취기준을 재구성하도록 되어 있다. 따라서 평가 문항에 대한 성취기준은 출제 교사가 직접 성취기준을 제시하고 학교장에게 결제를 받은 것이다.

3) 문항별 성취기준에 대한 평가목표의 적합성 분석

학교 현장에서는 교육과정에 제시된 성취기준을 수정하여 교육목적에 맞게 성취기준을 재구성하여 이원 목적분류표에 제시하고 해당 성취기준을 평가하기 위한 문항을 개발한다. 성취기준의 교육목표와 평가목표의 분석을 위해서 Bloom의 신 교육목표분류학(Anderson, Krathwohl & Bloom, 2001)에 근거한 분류 체계를 사용하였다. 이 분류 체계는 오랫동안 학교 현장에서 활용하던 Bloom의 교육목표 분류체계(Bloom *et al.*, 1956)에 대한 한계점을 보완하여 4개의 지식 차원(A. 사실적 지식, B. 개념적 지식, C. 절차적 지식, D. 메타인지 지식)과 6개의 인지과정 차원(1. 기억하다, 2. 이해하다, 3. 적용하다, 4. 분석하다, 5. 평가하다, 6. 창안하다)으로 분류한 이차원적 분류체계이다. 이는 교사의 수업목표, 수업활동, 평가가 어떻게 의미를 가지며 어느 정도 일관성을 가지고 있는가를 쉽게 파악하는데 유용하여 교육목표나 평가목표

의 분류에 관련된 다양한 교육연구에서 활용되어 왔다(Dong, Ha & Kim, 2015; Kim, 2013; Shin & Cho, 2008).

분석 과정을 살펴보면, Bloom의 신 교육목표 분류 체계에 따라 출제자가 제시한 문항별 성취기준의 지식 차원과 인지과정 차원에서의 교육목표를 먼저 분석하였다. 이후 같은 방법으로 실제 서술형 출제 문항에 들어있는 평가목표를 분석하였다. 이에 기초하여 서술형 평가 문항의 교육목표와 평가목표를 비교하여 문항의 적합성 정도를 알아보았다.

성취기준의 교육목표와 서술형 문항의 평가목표를 분류하기 위한 구체적 분석 과정은 다음과 같다.

- ① 출제자가 문항을 통해 측정하고자 하는 2009개정 교육과정 상의 성취기준을 명사 부분과 동사 부분으로 구분하였다.
- ② 명사 부분은 지식 차원(A. 사실적 지식, B. 개념적 지식, C. 절차적 지식, D. 메타인지 지식)에서의 해당 지식으로 분류하고, 동사 부분은 인지과정 차원(1. 기억하다, 2. 이해하다, 3. 적용하다, 4. 분석하다, 5. 평가하다, 6. 창안하다)에서의 해당 인지과정으로 분류하였다.
- ③ 평가 도구인 서술형 문항에서 요구하는 평가 목표를 추출하여 하나의 문장으로 재구성하였다.
- ④ 명사 부분과 동사 부분으로 나누어 각각 지식과 인지과정의 해당 차원으로 분류하였다.
- ⑤ 문항의 성취기준에 따른 교육목표와 실제 문항에서의 평가목표 분류가 일치하는지 여부를 비교하였다.

분석 결과의 객관성과 신뢰성 확보를 위해 신교육목표 분류체계를 이용하여 목표분류와 관련된 연구 논문으로 학회지에 게재된 경험이 있는 생물교육전문가 3인이 각자 실시한 후 분석 결과의 일치도를 조사하였다. 1차 분석에서 분석자 간의 분석 결과가 서로 다른 경우는 12.3% (300개의 문항 중 37개 불일치)로 나타났다. 이에 대해서는 모든 문항의 분석 결과가 일치할 때까지 각자 분석한 결과에 대한 이유를 설명하

Table 2. Examples of match analysis between Educational Objectives in Achievement Standards and Evaluation items.

평가문항에 대한 성취기준의 교육목표 분류					평가문항의 평가목표 분류					교육목표와
성취기준	명사 부분	동사 부분	지식 차원	인지과정 차원	평가문항	명사 부분	동사 부분	지식 차원	인지과정 차원	평가목표의 일치 여부
과9012-2. 과학이 우리 생활에 많은 영향을 미치고 있음을 설명할 수 있다.	과학의 영향	설명하다	A	1	과학의 부정적인 영향을 구체적으로 서술하시오	과학의 영향	서술하다	A	1	일치
과9023-3. 광물과 암석을 특징에 따라 분류할 수 있다.	광물과 암석의 특징	분류하다	B	2	석영과 방해석을 구별하는 방법을 서술하시오	광물의 구별 방법	서술하다	C	3	불일치
과9023-1. 지권은 다양한 암석과 광물로 구성되어 있음을 설명할 수 있다.	암석의 특징	설명하다	B	2	위치에 따라 화성암의 결정크기를 비교하시오	화성암의 결정 크기	비교하다	B	2	일치
과9025-1. 지진파를 이용하여 지구내부의 층상구조를 탐사하는 방법을 설명할 수 있다.	지구내부 탐사 방법	설명하다	C	2	지구내부를 조사하는 방법을 설명하시오	지구내부 탐사 방법	설명하다	C	2	일치
과9026-2. 판의 운동과 판의 운동, 지진, 화산 활동을 연계하여 설명할 수 있다.	판의 운동, 지진, 화산	설명하다	B	2	화산과 지진활동이 가장 활발한 2곳을 서술하시오	화산과 지진활동 이 활발한 곳	서술하다	A	1	불일치

고 논의하면서 재분류하였다. 실제 문항 분석 결과의 예시는 Table 2와 같다.

4) 분석 결과의 원인 조사를 위한 출제자와의 면담

분석 결과에 대한 보다 깊은 이해와 원인 탐색을 위해 모든 출제 교사(27명)를 대상으로 서술형 문항을 어떻게 출제하고 있는지에 대해 반구조화된 질문으로 개별 면담을 실시하였다. 주요 질문 내용은 성취기준

을 자체적으로 작성하고 있는지, 수업과 평가 계획은 어떻게 수립하고 있는지, 서술형 평가 문항을 어떤 방법으로 출제하는지, 문항 출제 시 중점적으로 고려하는 사항과 그 이유는 무엇인지 등으로 구성하였다. 면담은 개별 방문 또는 전화를 활용하여 실시했으며, 면담자와의 합의 과정을 거쳐 녹취한 음성 파일을 전사하고 자료를 정리하여 분석하였다.

Ⅲ. 연구 결과 및 논의

1. 서술형 평가의 문항 유형별 출제 빈도

서술형 평가 문항 300개(학년별 100개)에 대해 문항 제시 형태로 구분한 결과는 Table 3과 같다. 질문만 제시한 문항 단독형, 다양한 형태의 자료를 질문과 함께 제시한 자료 제시형으로 구분할 수 있다. 자료 제시형은 제시된 자료의 종류에 따라 세분화하였다.

학교 현장에서 실제 서술형 평가 도구로 사용한 문항 중에는 149개(49.7%)가 아무런 자료 없이 질문 내용만 단독으로 제시하는 문항이었으며, 151개(50.3%)가 자료 제시형 문항이었다. 자료 제시형에 해당하는 그림/사진 제시형이 37%이고 나머지 유형은 소수에 불과했다. 이는 고등학교 생명과학1의 서답형 문항을 분석한 선행연구(Lee & Jeong, 2014)에서 나타난 단독과제형(34.6%)보다 더 많은 빈도를 보이는

것이다. 이러한 결과를 볼 때, 중학교 현장에서는 학생들의 사고 과정을 이끌어 내기에 적합한 서술형 문항의 구성이 매우 부족하다는 것을 알 수 있다.

서술형 평가 문항의 유형을 응답 요구 형태에 따라 구분한 결과는 Table 4와 같다. 학교 현장의 출제 실태를 바르게 반영하기 위하여 본래는 서술형 문항으로 구분되지 않는 단답형과 완성형의 서답형 문항도 서술형에 포함하여 분석하였다.

응답 요구 형태 측면에서 보면, 서술형 문항이 65%로 나타났지만 고등학교 생명과학1의 서답형 문항을 분석한 선행연구(Lee & Jeong, 2014)의 80.4%보다 낮은 것이다. 또한 학교 현장에서 실제 평가도구로 사용하고 있는 서술형 평가 문항 중에는 서술형 평가 문항으로 분류하기 힘든 단답형, 완성형 문항의 출제 비율이 35%나 된다. 답안이 문장 형태로 기술되기는 하지만 실상은 중요 단어 하나에 술어가 첨가된 형태인 단답형이 29%나 나타나고 있었으며, 전체적인 답안을 학생 스스로 서술하는 것이 아니라 주어진 단어

Table 3. Analysis of evaluation items by the representing types

학년	문항 단독형	자료제시형					합계
		표/그래프	그림/사진	사례/설명	지도	복합자료	
1학년	52	11	31	3	3	0	100
2학년	63	7	26	0	2	2	100
3학년	34	5	54	6	0	1	100
계(%)	149(49.7%)	23(7.7%)	111(37.0%)	9(3.0%)	5(1.6%)	3(1.0%)	300(100%)

Table 4. Analysis of evaluation items by the response types.

학년	문항 유형			합계
	단답형	완성형	서술형	
1학년	30	7	63	100
2학년	23	7	70	100
3학년	34	4	62	100
합계(%)	87(29%)	18(6%)	195(65%)	300(100%)

를 중심으로 문장을 구성하여 답안을 작성하는 완성형도 6%로 나타났다. 이를 통해 학교 현장의 교사들이 ‘서술형 문항’으로 출제는 하였으나 실제로는 서술형 문항이 아니라 단순 지식의 암기 수준을 묻는 ‘단답형’, 또는 ‘완성형’이 의외로 많다는 점을 확인할 수 있었다.

이것은 많은 교사들이 ‘~을 서술하시오’, ‘~을 쓰시오’와 같이 단순히 문장 형태의 답안을 요구하기만 하면 서술형 문항이라는 왜곡된 인식을 가지고 있음을 추정할 수 있다. 그러나 출제자와의 면담을 통해, 대부분의 교사들이 서술형 평가 문항이 갖추어야 할 기본 형태를 알고 있지만 채점의 용이성 확보 및 교과 점수의 평균 상승을 위한 부담감 때문에 개선 노력을 하지 못하고 있는 것으로 확인할 수 있었다(유사 응답자 22명/27명중).

2. 단원(성취기준)별 서술형 문항의 출제 빈도

중학교 과학교과의 서술형 평가 문항 300개(1학년 100개, 2학년 100개, 3학년 100개)에 대한 단원별 및 성취기준별 출제 빈도는 Table 5와 같다. 다만 중학교 1학년은 교육과정 정책에 따른 자유학기제 운영으로 인해 2학기에는 정기고사를 실시하지 않았으므로 2학기 학습 주제에 대한 서술형 평가 문항을 분석할 수 없는 것은 본 연구의 제한점이다.

중학교 1학년의 단원별 서술형 문항의 출제 빈도는 현저한 차이를 보였다. 총 100개의 서술형 문항 중에서 ‘지구계와 지권의 변화’ 단원에서 46개, ‘힘과 운동’ 단원에서 44개가 출제되었으며, 해당 단원의 성취기준 중에서 출제되지 않은 성취기준 수는 2개 뿐이다. 반면, ‘과학이란’ 단원에서 3개만이 출제되었고, ‘광합성’ 단원에서 7개만이 출제되었다. 광합성 단원에 해당하는 12개의 성취기준 중에서 5개에 대한 것만이 출제되었다. 첫 단원인 ‘과학이란’은 과학의 전반적인 내용을 학습하는 단원이므로 그 중요도 면에서 정기고사의 서술형 평가보다는 과학 글쓰기 형태의 수행

평가에 활용되고 있었다(유사 응답자 21명/27명중).

‘광합성’ 단원은 교육과정의 학교급 간 연계성이나 내용의 중요도와 난이도를 고려할 때, 매우 중요한 단원임에도 불구하고 출제 빈도가 매우 낮게 나타났다. 이는 대부분의 학교들에서 ‘광합성’ 단원을 모두 학습하지 못한 상태로 1학기를 종료하고, 2학기에는 자유학기제 운영으로 인해 연속적인 학습이 어려워지기 때문으로 보인다. 자유학기제는 중학교 교육과정 중 한 학기 동안 학생들이 시험의 부담에서 벗어나 꿈과 끼를 찾을 수 있도록 수업을 개선하고 다양한 진로체험 활동을 운영하는 제도이다(Hwang *et al.*, 2013). 그러나 교수·학습 활동과 평가에 있어서 체계적인 계획과 관리가 이어지지 않는다면 식물에 대한 학습에서 가장 중요한 ‘광합성’ 단원이 학습 공백으로 남거나 평가에서 누락될 위험성이 커지게 된다.

중학교 2학년의 문항 출제 빈도는 전체 단원에 고루 분포하나 앞쪽 단원부터의 순서에 따라 높게 나타나는 경향을 보였다. 총 100개의 서술형 문항 중에서 ‘물질의 구성’ 단원에서 27개, ‘빛과 파동’ 단원에서 18개로서 이 두 개의 단원에서 45%의 문항이 출제되었다. 이들 단원에서는 서술형 평가에 누락된 성취기준이 적었지만, ‘기원과 우리 생활’, ‘소화 순환 호흡 배설’ 및 ‘물질의 특성’ 단원에서는 서술형 평가에서 누락된 성취기준이 약 50% 정도를 차지하였다.

앞쪽 단원에서 문항이 편중되어 출제되고 있는 원인을 알아보기 위해 출제자들을 면담한 결과, 1차 고사의 평가 문항 출제 시기에는 대부분의 학교에서 2단원 초반부까지 학습 진도를 나갈 수 있기 때문인 것으로 나타났다(유사 응답자 27명/27명중). 한편, ‘일과 에너지’ 단원은 학생들이 많이 어려워하는 단원이므로 교사들이 선택형 문항으로 쉽게 출제하여 학년 전체의 교과 평점을 높이려는 경향이 높기 때문에 서술형 문항 출제 빈도가 낮은 것으로 조사되었다(유사 응답자 25명/27명중).

중학교 3학년의 단원별 서술형 문항의 출제 빈도는 ‘전기와 자기’ 단원에서 27개, ‘생식과 발생’ 단원에서 23개, ‘유전과 진화’ 단원에서 20개로서 첫 단원과 생

Table 5. The number of evaluation items in each unit and achievement standard of middle school science.

학년	단원명	출제된 문항 수	단원별 성취기준 수		
			전체 수	출제된 수	미출제된 수
1	과학이란	3	3	2	1
	지구계와 지권의 변화	46	15	13	2
	힘과 운동	44	6	6	0
	광합성	7	12	5	7
	(소계)	100	36	26	10
2	물질의 구성	27	11	10	1
	빛과 파동	18	10	8	2
	기권과 우리 생활	14	18	8	10
	소화, 순환, 호흡, 배설	11	14	8	6
	물질의 특성	10	8	4	4
	일과 에너지	9	10	6	4
	자극과 반응	11	8	6	2
(소계)	100	79	50	29	
3	전기와 자기	27	9	9	0
	화학반응의 규칙성	13	8	5	3
	태양계	9	13	5	8
	생식과 발생	23	8	7	1
	여러 가지 화학반응	6	8	2	6
	유전과 진화	20	8	8	0
	외권과 우주 개발	2	14	2	12
	과학과 인류 문명	0	1	0	1
(소계)	100	69	38	31	
합계	300	184	114	70	

물 영역 단원에서 70%로 높게 나타났다. 학기별 첫 단원인 ‘전기와 자기’ 단원에서 출제 빈도가 높게 나타나는 것은 1, 2학년과 마찬가지로 학기별 1차 고사 실시 시기가 그 원인이었다. 반면, 물질 영역인 ‘화학반응의 규칙성’과 ‘여러 가지 화학반응’ 단원에서는

출제 빈도가 낮았고 누락된 성취기준이 많았다. 이러한 원인은 물질 영역의 내용이 1, 2학년에 비해 3학년에 있어서 갑자기 어려워지므로 교사들도 서술형 평가 문항 출제를 꺼리는 경향이 있기 때문으로 나타났다(유사 응답자 23명/27명중).

Table 6. Achievement standard with high frequency in descriptive evaluation items in middle school science.

학년	성취기준 번호	성취기준 내용	서술형 문항 수
1	과9032 (힘의 종류와 특성)	중력, 탄성력, 마찰력, 전기력, 자기력 등 여러 가지 힘의 특징을 알고, 이들이 주변의 현상을 이해하는데 어떻게 활용되는지 이해한다.	16
	과9034 (운동의 해석)	거리-시간, 속력-시간 그래프를 해석하여 물체의 운동을 설명할 수 있다.	14
	과9023-3 (암석의 분류)	광물과 암석을 특징에 따라 분류할 수 있다.	10
	과9023-1 (광물의 특징)	지권은 다양한 암석과 광물로 구성되어 있음을 설명할 수 있다.	6
2	과9083 (이온)	전자를 잃고 얻음으로 양이온, 음이온의 형성을 설명하고, 이를 원소기호를 사용하여 이온식으로 나타낼 수 있다.	9
	과9096-2 (파동)	파동의 진폭, 진동수, 파형으로부터 소리의 세기, 높낮이, 맵시를 알고, 설명할 수 있다.	5
	과9104-1 (대기 중의 물)	대기 중의 수증기량과 이슬점, 포화 수증기량, 상대습도를 설명할 수 있다.	5
	과9121 (순물질과 혼합물)	혼합물과 순물질의 차이를 이해하여 우리 주변에서 볼 수 있는 여러 물질들을 순물질과 혼합물로 구분할 수 있다.	5
	과9182-1 (체세포 분열)	체세포 분열의 특징을 염색체의 행동을 중심으로 설명할 수 있다.	8
3	과9151-2 (정전기 유도)	정전기 유도에 의해 물체가 대전되는 과정을 설명할 수 있다.	5
	과9181 (무성생식과 성생식의 비교)	무성 생식과 유성 생식의 특징을 파악하고 그 차이점을 비교할 수 있다.	5
	과9201-2 (유전의 기본 원리)	유전의 기본 원리인 독립의 법칙을 설명할 수 있다.	5

중학교 3학년의 지구과학 영역은 서술형 평가에서 거의 배제되고 있는 실정이었다. ‘태양계’ 단원은 사실적인 지식이 주류를 이루고 있는 내용이라 서술형 평가 문항 보다는 선택형 평가 문항으로 출제하는 경향을 보이면서(유사 응답자 25명/27명중), 서술형 평가 문항 수가 적었고 서술형 평가에 누락된 성취기준도 많았다. 또한, ‘외권과 우주 개발’ 단원과 ‘과학과 인류문명’ 단원에서 평가 문항이 거의 출제되지 않은 것은, 내신 성적 산출을 위한 성적의 조기 마감으로

인해 2학기 2차 고사 실시 시기가 너무 이르기 때문인 것으로 조사되었다(유사 응답자 27명/27명중).

각 단원별로 해당하는 성취기준이 있으므로 중학교 과학에서 서술형 평가문항의 출제 빈도가 높은 성취기준에 대하여 조사한 결과는 Table 6과 같다.

서술형 평가문항의 출제 빈도가 특히 높은 성취기준은 중학교 1학년 과학에서 ‘힘의 종류와 특성’, ‘운동의 해석’, ‘암석의 분류’, ‘광물의 특징’에 대한 내용 순서로 나타났다. ‘힘의 종류와 특성’이나 ‘운동의 해

석'에 관련된 성취기준에서 출제 빈도가 특히 높았지만 계산 과정, 그래프 및 표 완성, 주어진 결과 해석하기 등의 형태로 간단하거나 명확하게 답안 작성이 가능한 형태의 문항이 대부분이었다. 중학교 2학년 과학에서는 '이온' 내용의 성취기준이 가장 많이 출제되었고, 다음으로 '파동', '대기 중의 물', '순물질과 혼합물' 등이 고르게 많이 출제되었다. 중학교 3학년 과학에서는 '체세포 분열'에 대한 성취기준에서 가장 많이 출제되었고, '무성생식과 유성생식의 비교', '유전의 기본 원리', '정전기 유도' 등에서 비슷한 수의 문항들이 출제되었다. 그러나 출제 빈도가 높은 성취기준에 대한 서술형 문항은 주로 간단한 기호로 이루어진 식이나 숫자로 응답이 가능하거나 단순 분류 결과를 묻는 경우의 문항들이 많았다. 이러한 문항들은 학생들이 응답하기 용이하며 답안의 다양성 범위가 넓지 않아서 채점에 있어서도 부담이 적게 된다. 반면에 출제에서 누락된 성취기준을 살펴보면, 주로 각 영역별 기본 원리나 개념에 대한 내용들이어서 서술형 문항으로 출제할 경우, 설명하는 내용을 문장으로 서술하는 형태의 답안을 작성해야 하는 경우가 많다.

서술형 평가 문항의 출제와 관련하여 출제 교사들을 면담한 결과, 대부분의 교사들은 서술형 문항 출제시 고려해야 할 사항에는 타당도, 신뢰도, 객관도, 변별도, 난이도 등이 있으며, 이를 높이기 위한 방법도 알고 있었다(유사 응답자 24명/27명중). 그러나 서술형 문항을 출제할 때 가장 중요하게 생각하는 것은 측정하고자 하는 성취기준에 대한 적합성보다는 학생들의 응답 용이성이나 교사들의 채점 편이성인 경향이 많았다(유사 응답자 22명/27명중).

따라서 서술형 평가는 학생들의 고차원적인 사고 능력과 문제해결력 함양에 목적을 두어야 하는데, 오히려 학교 현장에서는 서술형 평가라는 명목 하에 학생들의 단순 지식 암기 능력을 향상시키거나 서술형 평가의 질적 수준을 저하시키는 결과를 초래하고 있는 측면도 있음을 비판적으로 생각해 볼 수 있다.

또한 실제로 학생들은 서술형 평가의 답안 작성 자체에 매우 큰 부담을 느끼고 있으며, 잘 알고 있는 내

용도 글쓰기 능력이 부족하여 제대로 표현하지 못하는 경우가 많아서 쉽게 출제해야 한다는 의견들도 있었다(유사 응답자 20명/27명중).

이와 유사하게 학생들에게 서술형 답안의 작성 방식에 대한 안내가 필요하다는 의견이 국어교과의 서술형 평가 사례 연구에서도 제시된 바 있다(Lee, 2012). 국가수준의 학업성취도 평가에서도 서술형 답안에 대한 응답률이나 정답률이 낮은 것도 학생들의 서술형 답안 작성 능력과 무관하지 않다(Lee et al., 2016; Lee et al., 2017). 따라서 서술형 평가나 과정 중심의 수행평가 실행이 증가하는 현실에서 이와 관련된 기능에 대해 학생들에게 사전 교육을 실시하고 평가가 이루어져야 한다고 생각한다.

3. 서술형 평가 문항별 성취기준과 평가목표의 적합성

출제자가 이원목적분류표에 제시한 문항별 성취기준이 포함하고 있는 교육목표와 실제 출제된 서술형 평가 문항이 포함하고 있는 평가목표를 Bloom의 신 교육목표 분류에 따라 분석한 결과는 Table 7과 같다.

서술형 평가 문항별로 제시된 성취기준의 교육목표('성취목표'로 칭하였음)는 '개념적 지식'과 '이해하다'가 결합된 형태가 74%로 가장 높은 비율을 차지하였다. 또한 '절차적 지식'이 9.0%에 불과하고, '분석하다' 이상의 상위차원에 해당하는 인지행동은 5.7% 이내로 매우 적었다. 따라서 과학이 탐구 활동을 강조하는 교과임에도 불구하고, 교사들이 선호하는 교육목표는 '개념적 지식을 이해시키는 것'에 편중되어 있음을 알 수 있었다.

서술형 평가 문항들에 나타난 평가목표에서는 '사실적 지식'과 '기억하다'가 결합된 단순 사고 형태가 성취목표(5.3%)에서보다 높게 나타났다(20%). 이것은 일선 학교 교사들이 서술형 평가 문항을 출제할 때, 학생들이 보다 쉽게 답할 수 있도록 문항을 단순화시키는 경향이 많기 때문으로 해석된다.

'개념적 지식'은 성취목표에서는 '이해하다'에 편중

Table 7. Analysis of the dimensions of cognitive processing and knowledge on the educational objectives in standard achievements and evaluation objects in evaluation items.

지식 \ 인지과정		1. 기억하다	2. 이해하다	3. 적용하다	4. 분석하다	5. 평가하다	6. 창안하다	합계
A. 사실적 지식	성취목표*	16(5.3%)	-	-	-	-	-	16(5.3%)
	평가목표**	60(20.0%)	-	-	-	-	-	60(20.0%)
B. 개념적 지식	성취목표	-	223(74.3%)	17(5.7%)	17(5.7%)	-	-	257(85.7%)
	평가목표	17(5.7%)	103(34.3%)	16(5.3%)	39(13.0%)	1(0.3%)	6(2.0%)	182(60.7%)
C. 절차적 지식	성취목표	-	5(1.7%)	18(6.0%)	-	-	4(1.3%)	27(9.0%)
	평가목표	-	2(0.7%)	30(10.0%)	21(7.0%)	-	5(1.7%)	58(19.3%)
D. 메타 인지적 지식	성취목표	-	-	-	-	-	-	-
	평가목표	-	-	-	-	-	-	-
합계	성취목표	16(5.3%)	228(76.0%)	35(11.7%)	17(5.7%)	0	4(1.3%)	300(100%)
	평가목표	77(25.7%)	105(35.0%)	46(15.3%)	60(20.0%)	1(0.3%)	11(3.7%)	300(100%)

* 성취목표 : 평가 문항에 해당하는 성취기준에 담긴 교육목표

** 평가목표 : 실제 평가문항에 담긴 교육목표

되어 있었지만(74.3%), 평가 문항에서는 34.3%로 낮아졌으며 ‘분석하다’의 인지과정에 결합된 문항이 13%로 증가하였다. ‘절차적 지식’은 성취목표에서 ‘적용하다’에 결합된 비율(6%)보다 평가목표에서 다소 높아졌고(10%), 성취목표에서는 ‘분석하다’와 연계된 것이 없었지만 평가목표에서는 21개(7%)로 증가되었다. 이와 같이 서술형 평가 문항의 성취기준에 담긴 교육목표보다 평가목표에서 다양한 인지과정으로 출제되는 것은 상위의 고차원적인 인지과정으로 사고의 영역을 확장하는 능력을 평가한다는 측면에서는 바람직한 일이다. 그러나 성취기준의 교육목표와 다른 평가 목표의 문항을 출제한다는 것은 성취기준에 따른 수업과 평가의 불일치를 보이는 것으로 문항의 타당도 측면에서는 적절하지 않은 것이다(Choi & Paik, 2016). 또한 ‘메타인지적 지식’과 ‘평가하다’의 인지과정이 성취목표나 평가목표에서 거의 반영되지 않은 것은 출제의 어려움이나 답안 채점의 용이성을

고려한 현실적 어려움 때문일 수 있다.

Table 7의 결과를 구체적으로 확인하기 위하여 300개의 서술형 문항들 각각에 대해 성취기준에서의 교육목표와 평가목표 사이의 일치 정도를 분석한 결과는 Table 8과 같다.

평가목표가 성취기준의 교육목표와 일치하는 문항은 121개(40.3%)에 불과하였다. ‘개념적 지식을 이해하다(B2)’의 교육목표를 가지는 성취기준은 223개로 평가문항의 대부분을 차지하였으며 이 중에서 90개(40.4%)만이 평가목표와 일치하였다. 이 밖에 소수의 문항이지만 성취기준 교육목표와 평가목표의 일치율이 높은 것은 ‘사실적 지식을 기억하는 것(A1)’으로 16개 문항 중에서 12개의 일치율(75%)을 보였으며, ‘개념적 지식을 분석하다(B4)’는 17개의 문항 중에서 7개의 일치율(41.2%)을 보였다. 나머지 문항들은 30% 미만의 낮은 일치율을 보였다.

평가 문항 출제에 있어서 학습내용에 대한 성취기

Table 8. Coherency between educational objectives and evaluation objectives in the evaluation items.

목표분류*	평가목표의 분류																성취목표수 문항	성취목표와 평가목표의 일치 문항 수(%)		
	A1	A2	A3	A4	A5	A6	B1	B2	B3	B4	B5	B6	C1	C2	C3	C4			C5	C6
성취기준의	A1	12						2	1							1			16	12(75.0)
	A2																			
	A3																			
	A4																			
	A5																			
	A6																			
교육목표	B1																			
	B2	37					17	90	10	29		4			19	13		4	223	90(40.4)
	B3	5						2	5	1	1				2	1			17	5(29.4)
	B4	3						2		7		1			2	2			17	7(41.2)
	B5																			
	B6																			
분류	C1																			
	C2	2						1						1	1				5	1(20.0)
	C3	1						5	1	1		1			5	4			18	5(27.8)
	C4																			
	C5																			
	C6								1						1	1			1	4
문항수	60						17	103	16	39	1	6		2	30	21		5	300	121(40.3)

* 지식 차원 분류 기호 : A-사실적 지식, B-개념적 지식, C-절차적 지식, D-메타인지적 지식,
인지과정 차원 분류 기호 : 1-기억하다. 2-이해하다. 3-적용하다. 4-분석하다. 5-평가하다. 6-창안하다.

준에 적합하도록 하는 것은 평가 문항의 타당도를 위해 매우 중요한 요인이다(Choi & Paik, 2016; Kim, 2013). 평가 문항에서 지식이나 인지과정 차원의 목표 수준이 성취기준에서보다 높거나 낮게 조정되어 평가 목표가 성취기준의 교육목표와 일치하는 비율이 낮은 원인을 찾기 위해 출제자들과 면담을 실시한 대표적인 내용은 다음과 같다.

- 연구자 : 선생님이 서술형 평가 문항별로 제시한 성취기준에서의 교육목표와 실제 문항에서의 평가목표를 비교하면, 지식 차원과 인지과정 차원이 일치하지 않는 경우가 있어요. 그 이유가 무엇 일까요?
- 출제자 A : 원래는 단위학교 차원에서 학생들의 수준과 특성을 반영하여 교육과정 내용별 성취기준을 자체적으로 마련해야 하지만, 학기 초는 많이 바쁘다 보니 기존에 교육부에서 제공하는 것을 그대로 제출하는 경우가 많아요. 하지만 서술형 시험문제를 낼 때에는 채점의 어려움까지도 고려하게 되어 성취기준과 별개로 제출하는 경우도 있어요. (유사 응답자 22명/27명).
- 출제자 B : 저는 문항을 먼저 출제하고 나서 제가 출제한 문항 내용과 비슷한 성취기준을 골라서 이원목적분류표에 기록해 왔습니다. 이런 상황 이니까 평가 문항의 평가목표와 성취기준이 일치하지 않는 경우가 발생한 것 같아요. 그런데

이렇게 대화를 나누다 보니 제가 지금까지 평가에서 가장 중요한 것을 놓치고 있었다는 생각이 드네요. (유사 응답자 5명/27명).

- 출제자 C : 학생들을 가르치다 보면 처음 계획과는 다르게 학습 내용에 따라 지식이나 활동 수준이 변화되는 경우가 많아요. 시험 문제는 학생들이 수업 시간을 통해 학습한 내용을 바탕으로 출제하니까 성취기준과는 일치하지 않게 되는 겁니다. 결국, 절차를 지키고 있지는 않지만 수업과 평가는 연계되어 시행되고 있는 셈이죠. (유사 응답자 9명/27명중).
- 출제자 D : 사실 학교에서 학생들을 지도하다 보면 해야 할 일들이 너무 많아 항상 시간에 쫓기면서 출제하게 되죠. 또한 서술형 평가 문항은 채점도 어렵고, 출제하는 전문성이 부족하다 보니 지식이나 인지과정 수준까지 생각하면서 문항을 만들지는 못하고 있어요. 그래서 사실 현장의 교사들은 평가와 관련된 전문성을 신장할 수 있는 기회를 절실히 원하고 있습니다(유사 응답자 18명/27명중).

면담 내용으로 볼 때, 평가 계획을 제출해야 하는 학년 초는 학교에서 매우 바쁜 시기이므로 체계적인 평가 계획을 수립하기 어려우며, 실제 평가 문항을 제작할 때에는 수업한 내용과 채점의 용이성을 고려하게 되므로 문항의 평가목표가 성취기준 교육목표와 일치하기 어려운 것으로 해석할 수 있다. 이와 같이 서술형 평가의 내용 타당도가 떨어지는 이유로 선행 연구들(Choi & Paik, 2016; Kim, Cho & Joo, 2012; Lee, 2012)에서도 채점의 용이성과 함께 이원목적분류표 작성 방법과 문항 개발에 대한 교사들의 평가 전문성 부족을 제시하고 있다. 따라서 수업과 평가에 관련된 단순 자료의 제공보다는 교사들의 평가 전문성을 실질적으로 신장시킬 수 있는 체계적인 연수가 필요하다.

IV. 결론 및 제언

서술형 평가는 선다형 평가의 한계를 보완하고 학생들의 창의성과 문제 해결력 등 고등 사고 기능을 평가하는데 유용하다. 최근에는 정책적으로 지필 평가에 서술형 평가의 도입을 필수화 했고 성적 반영 비율도 큰 폭으로 높였다. 그러나 학교 현장에서는 서술형 평가의 문항 출제에서부터 채점까지 일련의 과정이 안정적으로 정착되기 어려운 면이 있다. 이에 본 연구는 최근 3년간 K도의 7개 중학교에서 시행한 서술형 평가 문항 300개를 표집하여, 학교 현장에서 출제하고 있는 과학교과의 서술형 문항들에 대해 문항 유형 분류, 단위별 출제 분포, 성취기준 상의 교육목표와 평가목표 사이의 일치도 등을 분석하고 출제 교사와의 면담을 통해 그 원인을 탐색하였다. 연구 결과, 도출된 결론은 다음과 같다.

첫째, 학교 현장에서는 서술형 평가의 기본 취지와 목적에 부합하지 않는 문항 유형의 출제 비율이 매우 높다. 학교 현장에서 교사들이 직면하는 채점 기준 객관화의 어려움과 채점에 소모되는 시간 등을 해결하기 위한 방안으로 단순 기억으로 서술하는 문항들이나 간단한 단답형 문항들이 서술형 평가로 많이 사용되고 있다. 제시된 자료의 활용 없이 단순 기억에 의존하는 ‘문항 단독형’ 문항이 49.7%, 한 개의 단어를 요구하는 단답형 문항이나 주어진 단어를 연결하여 문장을 만드는 완성형 문항이 35%나 출제되었다. 이러한 문항은 채점 기준이 간단하여 서술형 문항의 채점 공정성을 확보하기 좋을 뿐만 아니라 채점이 용이하지만 학생들의 사고력을 평가할 수 있는 바람직한 서술형 문항은 아니다.

둘째, 서술형 평가 문항의 출제 빈도가 특정 단위과 성취기준에 편중되어 있다. 학교나 출제자가 다름에도 불구하고 시험 시기까지의 수업 진도 상황이 유사한 면이 있지만 학생들이 응답하기 용이한 학습 내용 중심으로 출제함으로써 성적을 높일 뿐만 아니라 답안 채점도 용이해지기 때문이라는 것을 교사들의 면담으로 알 수 있었다. 학생들에게 쉬운 내용에 해당하는 특정

단원이나 성취기준에 서술형 평가 문항이 집중되어 있는 것은 오히려 학생들의 학습 능력과 사고력을 제대로 함양시키고 평가하는데 장애가 된다고 볼 수 있다. 서술형 평가는 학생들의 문제해결 능력과 사고력을 평가하기 위한 것이므로 학습한 주제들에 대해 다양하고 통합적인 평가를 위한 문항을 개발하는 것이 필요하다.

셋째, 성취기준에서 추구하는 교육목표와 서술형 평가문항에 포함된 평가목표의 일치율(40.3%)이 낮아 서술형 평가문항의 타당도를 높이는 노력이 요구된다. 학교 현장에서 평가 문항의 타당도를 높이기 위해서는 교육과정 내용의 특성에 따라 성취기준이 담고 있는 교육목표에 부합되는 수업 활동이 이루어지고 이에 맞는 평가 목표를 가진 문항을 개발하는 것이 중요하다. 그런데 서술형 평가 문항들에 대해 출제 교사가 제시하는 성취기준은 개념적 지식을 이해하는 차원의 교육목표에 집중되어 있으며(74.3%), 평가목표는 성취기준과 다르게 다양한 지식과 인지과정 차원에 분포하였다. 교사들이 학기 초의 바쁜 시기에 설정하는 성취기준과 수업 실행 및 행정 업무로 바쁜 시기의 평가 문항 출제로 인해 '성취기준-수업-평가'의 일체화에 적합한 타당도 높은 평가 문항을 개발하기에는 어려움이 있다. 또한 학교 현장에서는 내신 성적 산출로 인해 평가의 공정성이 더욱 문제가 되고 있다. 특히 과정평가를 비롯한 서술형 평가의 채점에 대한 민원으로 어려움을 겪기 때문에 교사들 사이에서는 성취기준과 연계된 타당도 높은 평가 문항을 출제하기보다는 채점 기준이 간단하고 명확하도록 문항을 구조화 하는 것에 중점을 두고 있는 것으로 조사되었다. 이러한 요인들로 인해 학교 현장에서는 학생들의 사고 능력을 평가하고자 하는 서술형 평가의 본질적인 목적이 왜곡되어 있다고 볼 수 있다.

본 연구를 바탕으로 서술형 평가의 개선과 효율적 시행을 위해 다음과 같은 제언을 할 수 있다.

첫째, 동일한 학습내용에 대해 선다형 문항으로 평가하면 학생들이 답을 찾을 수 있지만, 서술형 문항으로 평가하면 알고 있는 내용을 제대로 작성하지 못하는 경우가 많다. 이에 대비하여 교사들은 서술형 문항

을 쉽고 단순하게 구성하여 기억에 의한 답안 작성을 요구하고 있다. 이로 인해 서술형 평가 문항의 유형이 적절하지 않을 뿐만 아니라 평가목표와 성취기준 목표가 일치하지 않게 된다. 따라서 학교 수업에서 서술형 문항의 답안 작성에 필요한 기초적인 역량을 사전에 지도하고 실습하는 시간을 가질 필요가 있다.

둘째, 본 연구에 참여한 대부분의 출제 교사들은 학생들의 창의력과 사고력 함양을 위해 제대로 된 서술형 평가가 꼭 필요하다는 공통 인식을 가지고 있었다. 그러나 올바른 서술형 평가 문항의 개발과 채점에 어려움을 느끼고 있음을 토로하였다. 따라서 서술형 평가를 위한 문항 개발 방법과 성취기준에 따른 평가목표의 도출, 채점 기준의 설정 방법 등에 관련된 체계적이고 전문적인 자료 지원과 교원 연수가 필요하다.

최근 학교 현장에서는 서술형 평가와 과정중심 평가의 성적 반영 비율을 더욱 높이도록 요구받고 있다. 이러한 교육정책은 평가와 피드백을 통해 학생들의 성장에 도움을 주려는 의도일 것이다. 따라서 학교 현장의 서술형 평가 실태와 원인에 대한 본 연구의 결과를 바탕으로 평가방법 개선을 위한 연구와 현장의 노력이 확산되기를 기대한다.

참 고 문 헌

- Anderson, L. W., Krathwohl, D. R., & Bloom, B. S. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's Taxonomy of educational objectives*. New York: Addison Wesley Longman.
- Bloom, B. S. (Ed.), Engelhart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (1956). *Taxonomy of educational objectives: The classification of educational goals. Handbook 1: Cognitive domain*. New York: David McKay.
- Busan Metropolitan City Office of Education [BMCOE]. (2010). *Tutorial for A descriptive*

- evaluation in Middle school and High school* (Busan Education, 2010-122). Busan: Author.
- Choi, J.-I., & Paik, S.-H. (2016). An analysis of content validity of behavioral domain of descriptive tests and factors that affect content validity : Focus on the fifth and sixth grade science. *Journal of the Korean Association for Science Education*, 36(1), 87-101.
- Dong, H. K., Kim, K. J., Kang, M., Jang, E.-S., Sung, K. H., Rim, H., Kim, S.-K., Lee, J., Pae, J.-K., Kim, S., Choi, B. T., Choi, W., Kim, Y.-J., & Lee, K. (2017). *A study on the development of the 2017 national assessment of educational achievement* (RRE 2017-2). Seoul: Korea Institute for Curriculum and Evaluation.
- Dong, H.-K., Ha, S.-H., & Kim, Y.-J. (2015). A comparative analysis of achievement standards of Korean science curriculum and performance expectation of Next Generation Science Standards (NGSS) in the United States: Content elements and objectives for middle school life science. *Educational Research*, 64, 95-125.
- Hwang, K.-H., Kim, K.-J., So, K.-H., Hong, J.-P., & On, J. (2013). *A study about operating model for free semester system in middle school*. Seoul: Ministry of Education.
- Ji, E. L. (2011). *Directions and tasks about introduction of absolute evaluation in middle school and high school*. Seoul: Korea Educational Development Institute.
- Kang, S. C. (2014). A study on the improvement for building the taxonomy table of two dimensional objectives based on it's actual condition survey in the standards based assessment of achievement. *The Journal of Curriculum and Evaluation*, 17(2), 21-48.
- Korea Institute for Curriculum and Evaluation [KICE]. (2016). *An interactive guide to understand the achievement assessment system-middle school* (PIM 2013-12). Seoul: Author.
- Kim, M. K., Cho, M. K., & Joo, Y. R. (2012). A survey of perception and status about descriptive assessment: Focused on elementary school teachers in Seoul area. *Journal of Elementary Mathematics Education in Korea*, 16(1), 63-95.
- Kim, S. (2013). Analysis on current status of descriptive test items in the Korean language curriculum : On the base of descriptive items in academic high schools. *Korean Language Education*, 140, 479-508.
- Kim, S. K., Choi, E.-J., & Paek, S.-H. (2015). An analysis of characteristic and factor about middle school science descriptive assessment items. *Bulletin of the Korean Chemical Society*, 59(5), 445-453.
- Kim, S. (2015). A study on implementation and improvement of classroom assessment. *The Journal of Curriculum and Evaluation*, 18(3), 257-282.
- Lee, K., Dong, H., Choi, W., Kwon, G., Lee, I., & Kim, Y.-J. (2017). Exploring a learning progression for eight core concepts of middle school science using constructed response items in the national assessment of educational achievement (NAEA). *Journal of Science Education*, 41(3), 382-401.
- Lee, I., Dong, H. K., Kwon, K. P., Choi, W., Lee, K., & Park, H. (2016). *Analysis on constructed response items in the 2015 national assessment of educational achievement (NAEA): science* (RRE 2016-9-4). Seoul: Korea Institute for

Curriculum and Evaluation.

Lee, Y.-J. (2012). A study on the development of test tools and scoring methods for Korean language essay test. *Korean Language Education Research*, 45, 413-449.

Lee, D., & Jeong, E. (2014). An analysis of paper and pencil test items of life science I in high school. *Journal of Science Education*, 38(3), 670-690.

Lee, S., Kim, G., Noh, S., Kim, M. K., & Kim, R. Y. (2014). Mathematics teachers' perceptions about and implementation of constructed-response assessment. *Journal of the Korean School Mathematics Society*, 17(2), 275-290.

Ministry of Education, Science and Technology [MEST]. (2009). *2009 revised curriculum (Elementary, Middle and High School): Summary* (Notice 2009-41, 2009. 12. 23). Seoul: Author.

Noh, E. S., Mun, K. C., Choi, J. A., Park, S. H., Min, S. U., Go, J. Y., Sim, W. H., Lee, J. H., Park, M. K., Park, T. J., Hong, S. J. Jeong, S. G. Park, J. M., Jeong, J. A., Yun, J. P., Lee, S. J., Song, Y. M., Kim, K. H., & Yun, H. D. (2010). *Theory and practice in questioning for assessing academic achievement* (Kyungnam Education 2010-056). Changwon: Gyeongsangnam-do Office of Education.

Paik, S. H., & Ryu, H. J. (2014). High school students' perceptions on descriptive assessment activity experiences by teacher or by peer. *Journal of the Korean Association for Science Education*, 34(6), 593-599.

Shin, J-G., & Cho, C-K. (2008). The statement of geography instruction objectives and the creation of evaluation questions based on revision of Bloom's Taxnomy of

educational objective. *The Journal of the Korean Association of Geographic and Environmental Education*, 16(2), 129-144.

Sin, H. J., An, S. Y., & Kim, Y. W. (2017). A policy analysis on the process-based evaluation-focusing on middle school teachers in Seoul. *The Journal of Curriculum and Evaluation*, 20(2), 135-162.

국 문 요 약

최근 들어 학생 평가 방법에서 서술형 평가의 중요성이 매우 강조되고 왔다. 본 연구는 학교 현장에서 출제해 온 서술형 평가문항의 실태와 원인을 분석하고자 하였다. 이를 위해 경남 지역의 중학교 7개 학교에서 1학년부터 3학년까지의 정기고사에 출제된 서술형 문항 300개를 수집하여 문항 유형과 단원별 문항 분포를 분석하고, 문항에 관련된 성취기준의 교육목표와 평가목표를 Bloom의 신 교육목표분류에 따라 분석하여 비교였다. 연구 결과, 대부분의 교사들이 공통된 단원에서 서술형 문항을 출제하는 경향을 보이며, 단원 및 성취기준별로 출제 빈도에 큰 차이를 보였다. 문항 유형으로는 서술형이 65%에 불과하였으며 나머지는 단답형이나 완성형 문항으로 구성되었다. 또한 성취기준에서의 교육목표와 문항에서의 평가목표 사이에 일치하는 비율이 40.3%에 불과하였다. 서술형 문항의 출제교사 27명을 면담한 결과, 서술형 문항의 개발 능력 부족과 채점에서의 어려움이 올바른 서술형 평가의 주요 장애 요인인 것으로 나타났다. 따라서 평가에 대한 체계적 연수를 통해 서술형 평가의 문제점들을 해결할 필요가 있다.

주제어: 서술형 평가, 중학교 과학, 문항 유형, 단원, 평가목표, 교육목표 분류