

## 중학교 과학과 서술형 평가의 문항 특성 및 요인 분석

김성기 · 최은주<sup>†</sup> · 백성혜<sup>‡,\*</sup>

전남과학고등학교

<sup>†</sup>창용중학교

<sup>‡</sup>한국교원대학교

(접수 2015. 5. 18; 게재확정 2015. 8. 3)

## An Analysis of Characteristic and Factor about Middle School Science Descriptive Assessment Items

Sungki Kim, Eunju Choi<sup>†</sup>, and Seounghey Paik<sup>‡,\*</sup>

Jeonnam Science Highschool, Naju 520-824, Korea

<sup>†</sup>Changyoung Middleschool, Suwon 440-813, Korea

<sup>‡</sup>Department of Chemistry Education, Korea National University of Education, Cheongju 363-791, Korea.

\*E-mail: shpaik@knue.ac.kr

(Received May 18, 2015; Accepted August 3, 2015)

**요 약.** 문제해결력 및 창의력 등과 같은 고등정신능력 배양을 위한 평가체제 구축의 일환으로 2005년 도입된 서술형 평가는 매년 평가비율이 증가하면서 단순 지식을 요구하는 선다형 평가의 대안적 평가로 떠오르고 있다. 정부의 서술형 평가 비율 확대에 학교현장에서 표면적으로는 어느 정도 정착이 되어 운영되고 있으나 실제 서술형 평가가 의도한 목적을 달성했다고 확인할 수 없다. 그래서 본 연구에서는 경기도 5개 중학교에서 실시된 과학과 서술형 평가 문항을 대상으로 서술형 평가 문항의 특성을 분석하였으며, 이러한 특성을 갖는 요인을 분석하고자 문항을 출제한 교사를 대상으로 심층면접을 실시하였다. 연구결과 서술형 평가 문항들은 ① 출제되는 단원의 편중, ② 단위별 문제 형식의 차이, ③ ‘고등정신능력’ 측정의 미흡이라는 문항의 특성을 갖고 있었다. 이러한 문항의 특성을 갖는 요인은 교사들의 평가 문항 제작, 평가 도구의 이해 등에 대한 능력이 결여되었기 때문이었다. 이러한 요소들은 모두 평가 전문성의 결여로 설명될 수 있다. 이를 통해 교사들의 평가 전문성을 신장을 위한 방안이 요구됨을 알 수 있었다.

**주제어:** 서술형 평가, 평가 전문성, 교사 전문성, 학생평가

**ABSTRACT.** In 2005, descriptive assessment was introduced to increase students' higher mental ability like problem-solving ability and creativity. Every year the ratio of descriptive assessment increases and it is regarded as an alternative evaluation to a multiple-choice item which measures simplicity knowledge. Externally the descriptive assessment took a root in school, but we can't say definitely that it meets its original goal. In this paper, science descriptive assessment items of 5 middle schools in Gyeonggi-do were analysed; examiners was interviewed about how well they understood the characteristic of the assessment items. According to the analysis, characteristic of the items are ① unequal distribution of unit, ② difference of item's type by unit, and ③ disappearance of measuring higher mental ability. It is considered that there are several factors of these characteristic - the lack of teachers' ability to make assessment item; understanding of assessment instrument. These factors can be explained by the lack of assessment expertise. So the society's effort is needed to raise teacher's ability for the descriptive assessment.

**Key words:** Descriptive assessment, Assessment expertise, Teachers' expertise, Student assessment

### 서 론

학교교육의 내실화와 교실 안의 문제를 효율적으로 해결하기 위해 평가에 관한 요구가 점차 증가하고 있으며, 지식정보화 시대에 맞추어 다양한 능력을 갖춘 인재 양성이 중요시 되고 있다. 학교교육에서도 학생 개개인의 소

질과 특성을 신장시키고, 단편적, 사실적 지식의 암기 능력보다는 창의적 문제해결능력과 같은 고등정신능력을 신장시키는 것이 절실히 요구되고 있다.<sup>1</sup> 이에 사고력, 문제해결력 및 창의력 등과 같은 고등정신능력 배양을 위한 평가체제 구축의 노력을 기울이고 있다. 그 일환으로 서술형·논술형 평가가 도입되어 2005년 2학기부터 서울 시

내 모든 초등학교 3학년에서 6학년까지의 학업성취도 평가가 서술형 문항으로 구성되고, 중·고등학교의 경우 처음 30%에서 시작하여 2006년 40%, 2007년 모든 학년에서 50%까지 확대되었다.<sup>2</sup> 또한 경기도 교육청은 초·중·고의 서술형 평가 비율을 2010년 20%에서 30%로, 2014년에는 40%까지 확대하였으며, 다른 시도교육청에서도 서술형 평가 비율을 점차확대하고 있다.<sup>3,4</sup>

이처럼 고등능력신장이란 의도로 실시된 서술형·논술형 평가는 학교현장에서 매년 그 반영비율이 증가하여 표면적으로 정부가 실시한 서술형·논술형 평가가 매우 안정적으로 정착되어 운영되고 있는 것으로 보이나 실제 서술형·논술형 평가가 의도한 목표를 달성하고 있는지는 미지수이다. 일부 연구에서는 서술형·논술형 평가가 정부의 떠밀기식 일방적 추진으로 이루어져 왔으며, 그렇기 때문에 평가 방식에 따른 교사의 수업 방식의 변화, 문항개발이나 채점과 관련된 여러 연수의 부재 등 교육 여건이 준비가 되지 않은 채 실시되어 교사에게 부담을 주고 있으며, 실제 이 평가의 목표가 제대로 달성되고 있지 않음을 지적하고 있다.<sup>5-9</sup> 이는 교육 프로그램의 성패 여부는 해당 교육 프로그램 그 자체도 중요하지만, 가장 중요한 것은 실제 그 교육 프로그램을 학생에게 투입할 최전선에 있는 교사의 역할을 강조한 Fullan의 생각과 일치한다.<sup>10</sup> 그러므로 서술형 평가의 성공적 정착에서 교육프로그램의 최종 실행자인 교사가 중요한 변인이 된다.

교육 프로그램의 실행자인 교사에게 필요한 능력을 흔히들 교사 전문성이라 한다. 이러한 교사 전문성은 수업에 초점을 두고 바라볼 때 수업 전문성, 전반적인 학생 생활에 초점을 두고 바라볼 때 교직 전문성 등 핵심 초점이 무엇이냐에 따라 다양한 갈래로 나뉘며, 이러한 전문성은 학자에 따라 조금씩 상이하게 정의가 내려지기도 한다. 일반적으로 흔히 말하는 교사의 전문성은 보통 수업에 대한 전문성 또는 교과 내용에 대한 전문성을 지칭하기 때문에 교사의 전문성 개발은 주로 교과 내용과 관련된 지식기반에 관심을 두며, 교사 양성이나 이후 연수과정 또한 이러한 전문적인 교과 지식 획득에 많은 시간을 투입하고 있다.<sup>11,12</sup>

하지만 교실에서 행해지는 수업에서 학생들이 교육목표를 달성하였다고 판단하는 것은 “그것을 아는 것(knowing that)”이라기보다는 “어떻게 아는가(know how)”에서 포착될 수 있다. 그렇기 때문에 교사의 전문성은 수업의 대상이 되는 교과 지식에 대한 내용의 전문성뿐만 아니라 그것을 어떻게 알고 있는지 측정하고 판단하며, 평가 결과를 해석하여 그 결과에 따라 행동(feedback)할 수 있는 평가 전문성이 중요하다고 할 수 있다.<sup>13,14</sup> 여기서의 평가 전문성은 한국교육과정평가원에서 정의한 ‘학생의 학습과 성취

에 관한 평가 정보를 수집하고 해석하여 활용할 수 있는 능력’에 평가에 대한 교사의 생각과 신념까지를 포함한다. 이에 대해 Edward는 설계(design), 활용(use), 공평성(equity) 차원으로 나누어 개념화를 시도하기도 하였다.<sup>11,15-18</sup>

이처럼 교사의 전문성이 주로 교과 내용과 관련된 지식기반에 관심을 두기 때문에 예비교사를 양성하는 기관이나 교사 연수기관을 통해서 평가 전문성을 획득한 기회가 상대적으로 충분하지 못하며, 학교 현장의 교사조차도 스스로 자신이 평가 전문성을 갖추고 있지 못하다고 인식하고 있었다.<sup>19</sup> 과학 서술형 평가에 대한 국내 선행연구는 경우 주로 평가 도구 개발에 초점을 두거나 또는 단순히 서술형 평가에 대한 학생과 교사의 인식에 그치고 있으며 교사의 평가 전문성 신장에 초점을 두지 않고 있다.<sup>20-22</sup>

이에 따라 본 연구에서는 먼저 학교 현장에서 시행되고 있는 중학교 2학년 과학과 서술형 평가 문항의 특성을 분석하여 현재 서술형 평가가 어떻게 학교 현장에 진행되는지를 살펴보았다. 이후 서술형 평가 문항의 특성에 대한 요인을 현상학적으로 알아보았다. 이를 통해 교사의 서술형 평가에 대한 평가 전문성을 향상시킬 수 있는 실제적인 방안을 모색하고자 한다.

## 이론적 배경

### 서술형 평가

21세기 지식·정보화 사회에 맞는 수업의 질 제고를 위한 평가 방법 개선의 필요성이 부각되었다. 이와 더불어 학교평가 문화를 개선하고 평가 방법의 개선을 통한 수업 방법의 변화를 도모하여 지식 정보화 사회에 필요한 고등 지식능력과 창의적 문제해결력을 배양할 목적으로 서술형 평가를 실시하게 되었다.

서술형 평가란 학생이 알고 있는 정보나 지식 등을 활용하여 문제에 대해 논리적으로 분석·설명·해석하거나, 창의적으로 문제를 해결할 수 있는 능력을 측정하는 평가 유형이다. 서술형 평가 문항의 기본 구조는 평가요소와 반응 지시어, 배점 등으로 간략하게 나타낼 수 있다. 평가 요소는 학생이 응답하기를 요구하는 내용요소이며, 반응 지시어는 학생이 문제를 해결한 결과를 어떤 결과를 어떤 형태로 작성해야 하는지를 지시하는 일종의 형식 요소라고 할 수 있다. 그리고 배점은 해당 문항을 해결했을 때 취득할 수 있는 점수이다. 확장구조는 서술형 평가 문항의 기본 구조에 조건 첨부, 자료(제시문, 그래프, 그림, 표) 제시 등이 더 포함된다.

서술형 평가 문항의 유형은 학생의 반응 허용 정도가 어떠한가에 따라 응답 제한형과 응답 자유형으로 구분되기도 하고, 문항에 대한 작성과 자료나 정보를 포함시켜

제시하는가의 여부에 따라 단독 과제형과 자료 제시형으로 구분하기도 하며, 문항을 통해 측정하고자 하는 평가 내용이 무엇인가에 따라 일반 쓰기 능력과 교과관련 능력으로 구분하기도 한다.

서술형 평가는 학생의 고등사고 기능과 복합적인 학습 결과 뿐만 아니라 정의적 측면까지 측정할 수 있으며, 이를 통해 보다 학생의 학습수준을 정확하게 측정할 수 있는 장점이 있다. 뿐만 아니라 학생에게 보다 좋은 학습 경험을 제공하여 고등사고 기능을 증진시킬 수 있으며, 교수·학습 활동의 개선 측면에서 교수·학습 활동 개선에 대한 정보를 얻을 수 있고 피드백이 유용하여 학생의 학습태도를 개선시킬 수 있다. 하지만 서술형 평가는 채점에 많은 시간과 노력이 필요하며, 학생이 작성한 답안에 대해 다양한 해석이 가능하여 공정한 채점이 어려워 객관성과 신뢰도가 낮을 수 있다는 단점이 있다.<sup>23,24</sup>

### 평가 전문성

전통적으로 학교에서 행해지는 평가의 목적은 학생에게 점수를 부여하는 일(grading)이었다. 하지만 이제 평가란 더 이상 단순하게 시험을 치러 점수를 매기고 성적을 기록하여 학생의 서열을 결정하여 기록하는 것이 아니며, 학생 학습목표에 대한 진척 상황에 관한 지속적인 분석 정보를 수업과 통합시키는 일을 강조하고 있다. 그러면서 교사에게는 보다 높은 차원의 평가 전문성을 요구하고 있다.

교사가 학교에서 행하는 업무(task)에 대한 연구 결과에 의하면 교사들은 업무시간의 1/3에서 1/2을 평가와 관련된 업무에 쓰고 있었으며, 교사들은 평가가 자신의 전문적인 업무에서 중요한 부분이며 평가 관련 능력이 필요함을 인식하고 있었으나, 교사들은 평가와 관련된 전문 훈련을 받은 경험이 거의 없거나 적으며 평가 관련 행동을 수행하는데 준비가 덜 되어 있다고 교사 스스로 고백하기도 한다. 또한 교원 양성기관의 교직 프로그램에서 제공하는 평가 관련 연수나 교육이 평가와 관련하여 요구되는 지식과 기술과는 거리가 멀다는 지적도 있다.<sup>25-27</sup>

학생평가의 본질적 기능은 학생의 학습(learning)과 성취(achievement)를 평가하여 평가 정보를 학생, 교사, 학부모 등 관련 이해당사자에게 제공함으로써 그들의 교육적 노력 및 의사결정을 도움을 주는데 있다고 할 수 있다. 그렇기 때문에 교사는 평가를 통해 학생의 학습 성취에 대한 정보를 수집할 수 있으며, 이를 반영하여 수업에 활용하는 활동을 전개할 수 있다. 그러므로 교사의 평가 전문성이란 학생의 학습과 성취에 관한 평가 정보를 수집하고 해석하여 활용할 수 있는 능력이라 할 수 있으며, 교사에게 요구되는 평가 전문성은 평가 활동을 하기 위한 핵심적인 필수 능력이다.<sup>28,29</sup> 한국교육과정평가원에서 제시한

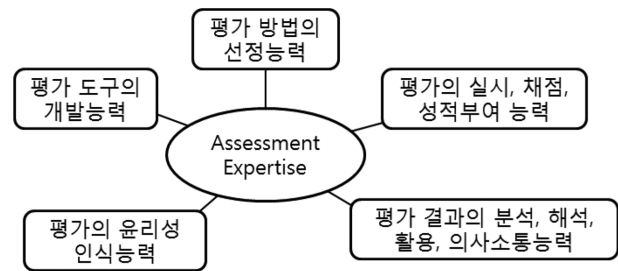


Figure 1. Assessment Expertise (KICE, 2007).

교사의 평가전문성 요소는 Fig. 1과 같다.

교사의 평가 전문성의 5가지 구성 요소는 교사가 학생을 대상으로 평가 활동을 수행할 때 발휘해야 하는 전문적인 능력이다. 먼저 평가 방법의 선정능력은 교사가 학습목표와 평가의 목적을 확인하여 명료화할 수 있어야 하며, 학습목표 및 평가의 목적에 적합한 평가 방법을 선정해야 할 수 있는 능력을 말한다. 평가도구의 개발 능력은 평가의 목적과 내용에 적합한 평가 도구를 개발하거나 선택하여 사용할 수 있으며, 평가 도구의 질을 스스로 점검하고 개선할 수 있음을 의미한다. 평가의 실시, 채점, 성적부여 능력은 평가 계획에 부합하도록 평가를 실시하며, 교사가 정확하게 채점하고 평가의 목적에 부합하도록 성적을 부여하는 것을 의미한다. 평가 결과의 분석, 해석, 활용, 의사소통 능력은 평가 결과를 정확하고 타당하게 분석 및 해석할 수 있으며, 평가 결과를 수업이나 학생에 대한 교육적 의사 결정에 활용할 수 있으며 교사가 학생이나 학부모 또는 교육 관련자와 평가 결과에 대해 의사소통할 수 있는 능력을 지칭한다. 끝으로 평가의 윤리성 인식 능력은 교사는 학생의 인격을 존중하며 모든 평가 활동 시 윤리적, 법적 책임을 준수하며, 학생의 특성 및 배경을 고려하여 공정하게 평가를 실시하고, 평가의 적절성을 판단하는 것을 의미한다.

이러한 평가전문성에 대한 정의는 과학 교사의 평가 전문성을 개념화한 Edward의 개념과 유사하다. Edward는 학생이 무엇을 알고 무엇을 할 수 있는지를 발견하는 설계차원(design), 평가의 시기를 평가 목적이나 교육과정 맥락에서 고려하며, 평가 결과를 학생의 학습을 지원하기 위한 방식에서 고려하는 이용차원(use), 평가에서 학생의 사회 문화적 영향 차이를 고려하여 적절한 형태의 평가를 제공하는 공평성 차원(equity)으로 나누어 평가 전문성을 개념화하였다. 다만 공평성 차원에서 미국의 문화적 특성인 언어 소수자들에 대한 배려를 추가한 것이 그 차이점이라 할 수 있다.

## 연구 방법

### 연구 절차

본 연구에서는 2011년 경기도 지역 중학교 2학년을 대

상으로 실시된 1학기 과학과 지필고사(중간, 기말고사)를 수집하였다. 이때 평가문항 원안만 뿐만 아니라 이원목적 분류표 및 채점기준표 등 평가와 관련된 전반적 문서화된 서류를 수집하였다. 수집된 서술형 문항을 특성을 분석하기 위해서 문항의 출제 형식 분류틀을 개발하였으며, 이에 대한 타당성을 확보하기 위해 과학교육관련 석사학위 이상의 현직교사 5명과 과학교육 전공 교수 1명의 검토를 통해 최종 수정·보완하였다.

최종 수정·보완된 분류 틀을 이용하여 출제된 서술형 문항을 분석하였다. 이때 서술형 문항뿐만 아니라 선다형 문항을 함께 비교분석하여 서술형 문항이 갖는 고유의 특성을 찾도록 하였다. 또한 단순히 출제된 문항뿐만 아니라 평가이전의 계획과정이나 평가이후 채점과정까지 평가의 전반적인 과정도 분석되었으며, 이러한 부분은 주로 면담을 통해 분석되었다. 해당 내용을 토대로 분류된 서술형 문항의 특성을 토대로 현상에 대한 요인을 분석하기 위해 반구조화된 면담을 실시하여 종합적인 분석을 실시하였다. 심층 면담을 통해 얻어진 서술형 문항 특성의 요인에 대한 결론의 타당화를 위해 삼각화 검증(triangulation) 방법을 이용하였다. 질적연구 특성상 분류된 특성들이 완전 분리된 특성이 아니었으며, 그렇기 때문에 면담자료 또한 명확히 구분되지 않는 자료도 존재하였다. 이때는 가장 해당 특성을 보여주는 특성 요인으로 분류하였다.

## 연구 대상

**서술형 평가 문항 분석 대상:** 2011학년도 1학기에 시행된 경기도 지역의 5개 중학교의 서술형 평가의 문항을 분석하였다. 대상 학교의 중학교 2학년 과학 지필고사(중간고사, 기말고사)를 대상으로 하였다. 5개 학교 모두 서술

형 평가 반영비율은 20%이었으며, 출제교사는 학교에 따라서 2~3명 이었다. 자세한 서술형 평가 문항 대상 학교의 특성은 Table 1과 같다.

**면담 대상:** 서술형 평가 문항을 분석한 후 문항제작의 의도, 출제할 때의 유의점, 이원목적분류표 작성 및 채점 과정, 서술형 평가 문항이 갖는 특성의 요인 등의 평가의 전반적인 과정을 알아보기 위해 서술형 평가 문항을 제출한 6개 학교(교사 13명)에 각 1명씩 면담을 추천 의뢰하였다. 이중 3명의 교사가 면담을 승낙하였으며, 나머지 학교에서는 면담에 응하지 않음을 통보하였다. 본 연구는 교사들이 공개를 꺼려하는 평가 문항을 제공받아야 하며, 이후 진행되는 평가의 전반적 과정에 대한 면담이 문항을 출제한 교사에게 부담이 되었을 것으로 판단되었다. 그리하여 면담자를 추천하지 않는 학교에 면담 추천을 재의뢰하지 않았으며, 면담에 응한 3명의 교사로 면담을 진행하였다. 면담 대상 교사의 배경은 Table 2와 같다.

## 문항 분석틀

**단원별 출제 빈도:** 전반적인 단원별 출제 빈도를 알아보기 위해 중학교 2학년 과학과 단원 중에서 1학기에 해당하는 4개 대단원과 18개의 소단원의 목록을 작성하였으며, 이 목록을 토대로 선다형 문항과 서술형 문항의 빈도를 기록하였다. 일부 문항 중에는 18개의 소단원 중 정확히 한 개의 소단원으로 분류가 되지 않는 것도 존재하였으며, 이때는 해당하는 소단원 모두에 빈도에 포함하도록 하여 전체 빈도수는 분석된 문항 수와 상이할 수 있다.

**이원목적분류표:** Bloom이 제시한 인지적 영역의 분류가 과학지식의 평가 틀로 많이 활용된다. Bloom은 지적

**Table 1.** Characteristics of the schools

School	A	B	C	D	E
Location	Suwon	Ichoen	Suwon	Kwangju	Suwon
Number of class	6	7	12	14	10
Number of item	58	56	58	60	55
Number of descriptive assessment item	8	8	8	8	9
Number of science teacher	2	2	3	3	3

**Table 2.** Background of subject

	Teacher a	Teacher b	Teacher c
Major	chemistry	chemistry	physics
Career of education	10 years	7 years	9 years
Sex	female	female	male
Highest degree	master	master	BS
Undertaking class' number	5	6	6
Feature	-	in charge of gifted class (4 years)	in charge of gifted class (1 years)

행동과 교수목표를 크게 지식 그 자체와 지식에 대한 기  
능으로 분류하였으며, 일반적으로 지식, 이해까지를 단순  
정신 능력이라 하고, 적용, 분석, 종합, 평가를 고등정신이라  
고 하였다. 이원목적분류표의 행동영역 활용은 학교마다  
상이하였으나 보통 지식, 이해, 적용의 3단계로 시행하  
고 있었다. 본 연구는 서술형 평가 문항이 실제 의도한 고  
등정신을 측정하고 있는지 판단하기 위해, 문항을 출제한  
교사에게 Bloom의 6단계 행동영역으로 이원목적분류표  
를 재작성 하도록 요구하였다. 교사가 재작성한 이원목적  
분류표에 대해 석사학위이상의 5인의 현직교사 검토를  
받았으며 일부 행동영역 분류를 수정하였다.

**문항 출제 형식:** 서술형 평가 문항의 유형은 학생의 반  
응 허용 정도가 어떠한가에 따라 응답 제한형과 응답 자  
유형으로 구분하기도 하고, 문항에 답한 작성과 관련한  
자료나 정보를 포함시켜 제시하는가의 여부에 따라 단독  
과제형과 자료 제시형으로 구분되기도 한다. 또한, 문항  
을 통해 측정하고자 하는 평가 내용이 무엇인가에 따라  
일반쓰기 능력형과 교과관련 능력형으로 구분된다.<sup>29</sup> 1차  
적으로 이 틀을 이용하여 분석한 결과 반응의 허용정도에  
따른 분류에서는 응답 제한형만, 자료 제시여부에 따른  
분류에서는 자료 제시형만, 평가 내용에 따른 분류에서는  
교과관련 쓰기 능력형만으로 출제되었다. 이틀에서 특별  
한 의미를 발견할 수 없었기 때문에 서술형 41개 문항을  
토대로 과학교육 전문가 1인과 현직교사 3인의 자문을 받  
아 출제 형식 분류를 재조직 및 수정하였으며, 최종 4가  
지 유형으로 분류를 마련하였다. 분석된 41개 문항을  
토대로 재조직 및 수정된 분류는 지식 서술형, 계산형, 자  
료 제시형, 이유 설명형이었다. 일부 문항의 경우 엄밀하  
게 하나의 형식으로 분류되지 않는 것도 있었으며, 그때  
의 경우 중복된 유형 모두를 해당 유형으로 간주하였다.  
동일한 문항일지라도 사람에 따라 분류가 일부 상이할 수  
있기 때문에 분석된 자료를 토대로 현장교사 2인의 검토  
를 받아 신뢰도를 확보하였다. 분석에 사용한 서술형 평  
가 문항 형식별 의미는 다음과 같다.

#### 서술형 문항 형식별 의미:

- 지식 서술형: 이미 배운 과학지식을 회상하여 차례대로  
적은 문항 형식
- 계산형: 주어진 정보를 이용해 어떤 값이나 결과를 구  
하는 형태의 문항 형식
- 자료 제시형: 수, 영상, 단어를 통해서 얻어지는 정보를  
어떠한 목적을 위해 쉽게 풀어내거나 나누는 형태의  
문항 유형
- 이유 설명형: 어떠한 결론이나 결과에 이르는 까닭이나  
근거를 밝혀 말하는 형태의 문항 형식

## 연구 결과 및 논의

### 출제되는 단원의 편중

서술형 평가 문항은 ‘지구의 모양과 크기’, ‘용해도’, ‘잎의  
기능’, ‘운동의 표현’ 순으로 많이 출제되어지고 있었다.  
분석한 서술형 41문항 중에서 ‘용해도’는 5문항(12.2%)으로  
선다형 평가 문항 분석 결과와 유사하게 높은 출제 빈도를  
나타내었다. 그러나 ‘운동의 표현’, ‘지구의 모양과 크기’는  
선다형 문항에서는 출제빈도가 높지 않았으나 서술형 평  
가에서는 각각 4문항(9.8%), 6문항(14.6%)으로 높은 출제  
빈도를 나타내었다. 또한, 선다형 평가에서 높은 출제 빈  
도를 차지하였던 ‘속력과 방향이 모두 변하는 운동’은 서  
술형 평가에서는 전혀 출제되지 않기도 하였다. Table 3, 4는  
단원별 출제 빈도를 나타낸 것이다. 이처럼 특정 단원의  
서술형 평가가 출제에서 배제되거나 특정한 단원이 과다  
하게 출제되는 원인으로 교사들은 채점의 용이성을 들었다.

연구자: ‘속력과 방향이 모두 변하는 운동’을 출제하지  
않는 이유는 무엇인가요?

a교사: 운동에서는 계산 문제 내는 것이 가장 쉽죠. 관련된  
예를 물으면 답이 너무 많아요. 아이들은 정말  
다양한 답을 하거든요. 이러한 경우에는 선다형  
문제로 출제하는 것이 훨씬 편하죠.

b교사: 운동 단원에서 답으로 나올 수 있는 경우의 수가  
많고, 어느 시점에서 보느냐에 따라 답이 달라질  
수도 있어서 출제가 곤란해요.

연구자: 서술형 평가 문항을 출제할 때 고려하는 사항이  
주로 어떤 것인가요?

a교사: 처음에는 중요한 학습내용을 우선으로 했었는데요,  
몇 번의 시험을 치르면서 채점의 용이성이 가장 우  
선시 되더군요. 채점이 명확하게 될 수 있는 학  
습 내용을 선정해서 출제하고 있어요.

b교사: 답안의 명료성이 우선이고, 학습 내용의 중요도,  
서술 답안의 길이, 평균을 고려한 난이도도 고려  
하죠.

c교사: 첫 번째는 채점의 용이성이고요, 그 밖의 채점 기준  
의 명확성, 채점 기준의 당위성을 고려해요. 그렇지  
만 결과적으로 채점결과에 이의를 제기하지 못  
하게 하는 게 가장 궁극적인 목적이죠.

면담결과 교사는 서술형 평가 문항은 채점이 용이할 수  
있도록 답안이 명료하게 제작되어지는 단원을 중심으로  
출제하고 있었다. 그렇기 때문에 출제 단원이 선다형 평  
가 문항에 비해 다소 편중되는 경향을 보였다. 이는 앞서  
알아본 평가 전문성 중에서 평가의 실시, 채점, 성적부여

**Table 3.** Frequency of multiple-choice items by unit

Unit		School					Total
		A	B	C	D	E	
Various motion	Representative of motion	2	0	3	1	2	8(3.2%)
	Motion of object with force	3	3	1	3	2	12(4.8%)
	Motion of object changing velocity by force	3	4	4	3	2	16(6.5%)
	Motion of object changing direction by force	2	1	1	2	2	8(3.2%)
	Motion of object changing velocity and direction by force	2	5	5	4	3	19(7.7%)
Characters of material	Melting point and boiling point	2	5	3	4	2	16(6.5%)
	Density	6	3	3	5	5	22(8.9%)
	Solubility	5	4	7	4	5	25(10.1%)
The Earth and stars	Shape and size of the earth	1	2	1	2	2	8(3.2%)
	Solar system	6	4	6	5	6	27(10.9%)
	Stars	5	3	4	3	3	18(7.3%)
	Galaxy	2	2	2	3	1	10(4.0%)
Structure and function of plant	Structure of plant	2	2	2	1	1	8(3.2%)
	Structure and function of root	1	2	2	2	2	9(3.6%)
	Structure and function of stem	2	4	1	2	2	11(4.4%)
	Structure of leaves	2	0	2	2	1	7(2.8%)
	Function of leaves	3	0	4	5	4	16(6.5%)
	Flower and fruit	1	4	1	1	1	8(3.2%)
Number of item		50	48	52	52	46	248(100%)

**Table 4.** Frequency of descriptive assessment items by unit

Unit		School					Total
		A	B	C	D	E	
Various motion	Representative of motion	1	2	0	1	0	4(9.8%)
	Motion of object with force	1	0	1	1	0	3(7.3%)
	Motion of object changing velocity by force	1	0	0	0	2	3(7.3%)
	Motion of object changing direction by force	0	0	1	0	0	1(2.4%)
	Motion of object changing velocity and direction by force	0	0	0	0	0	0(0%)
Characters of material	Melting point and boiling point	0	1	1	0	1	3(7.3%)
	Density	1	0	0	1	0	2(4.9%)
	Solubility	0	1	1	1	2	5(12.2%)
The Earth and stars	Shape and size of the earth	2	0	1	1	2	6(14.6%)
	Solar system	1	1	0	0	0	2(4.9%)
	Stars	0	1	0	1	0	2(4.9%)
	Galaxy	0	0	0	0	0	0(0%)
Structure and function of plant	Structure of plant	0	0	0	0	1	1(2.4%)
	Structure and function of root	1	1	0	1	0	3(7.3%)
	Structure and function of stem	0	0	0	0	0	0(0%)
	Structure of leaves	0	0	0	0	0	0(0%)
	Function of leaves	0	0	3	1	1	5(12.2%)
	Flower and fruit	0	1	0	0	0	1(2.4%)
Number of item		8	8	8	8	9	41(100%)

능력 중에서 교사가 정확하게 채점하는 것에 부합하려는 교사의 노력을 엿볼 수 있다고도 할 수 있다. 하지만 이때 정확한 채점이라는 것은 평가의 목적에 부합하도록 성적을 부여하는 것이다. 서술형 평가 문항 제작이 채점이 용

이한 단원의 선정은 평가의 목적에 어긋나는 것으로 이는 평가 전문성이 결여된 형태이다.

더군다나 서술형 평가는 다양한 답이 충분히 가능하다는 평가도구의 특성이 있음에 채점의 용이성을 먼저 고려하는

**Table 5.** Type of descriptive assessment items by unit

Type	Description	Calculation	Data analysis	Explanation of reason
Unit	Various motion	2	4	3
	Characters of material	-	4	1
	The Earth and stars	1	4	4
	Structure and function of plant	5	-	5
Total		8(16%)	12(27%)	8(18%)
				17(37%)

것은 평가도구에 대한 이해가 부족함을 알 수 있다. 평가에 사용되는 도구의 특성에 대한 이해가 필요하며, 유일한 하나의 정답이 아니라 다양한 정답이 존재할 수 있음을 인정하는 교사의 평가에 대한 태도의 변화가 필요하다.

### 단원별 문제 형식의 차이

Table 5는 서술형 평가 문항의 단원별 출제 형식을 분석한 결과이다. 다양한 운동 단원의 경우 4가지 문제 형식의 큰 차이가 없었다. 반면 물질의 특성 단원은 ‘계산형’과 ‘이유 설명형’으로 편중되어 출제되고 있었다. 지구와 별 단원은 ‘지식 서술형’이 상대적으로 적게 출제되었으며, 식물의 구조와 기능 단원은 ‘계산형’, ‘자료 분석형’ 문항은 전혀 출제되지 않았다.

식물의 구조와 기능 단원에서 ‘계산형’이 출제되지 않는 것은 단원 특성상 특정한 형태의 문제 형식이 제작되기 어렵다는 요인이 있음을 부인할 수 없다. 하지만 이러한 요인으로 단원별 문제 형식의 차이를 전부 설명하기는 어렵다. 예를 들어 식물의 구조와 기능 단원에서 ‘자료 분석형’이 나오지 않는 것은 이로 설명할 수 없기 때문이다. 또한 ‘계산형’이 식물의 구조와 기능 단원처럼 출제되기 어려운 단원이 있음에도 불구하고 전체 27%를 차지하는 것은 결코 작은 수치라 할 수 없다. 결국 단원별 문제 형식의 차이는 단원의 특성의 기인되는 요인도 있으나 또 다른 요인이 있다고 판단되었으며, 이를 알아본 면담은 다음과 같다.

연구자: ‘이유설명형’과 ‘계산형’으로 주로 많이 출제하는 이유는 무엇인가요?

b교사: 채점의 용이성 때문이지 않을까요? 이 두 가지의 경우 정답이 보다 객관적이잖아요.

c교사: 간단하면서 정확한 답을 쓸 수 있기 때문에 채점하기 쉬워요. 학생들의 인정 답안도 간추려지는 편이고요.

일부 단원의 특성상 차이는 있으나 전체적으로 문항의 출제 형식이 ‘이유 설명형’과 ‘계산형’으로 집중되는 이유는 편중된 단원의 출제와 유사하게 채점의 용이성이었다. 교사에게 객관적인 채점은 평가에 대한 의사 결정에서 주요한 요인이었다. 결국 교사의 문항 제작은 서술형 평가가 이루고자 하는 교육목표의 측정이기 보다는 누가 채점해도 동일한 결과를 얻을 수 있는 채점의 객관성이 교사에게 더 중요한 것이었다. 이러한 면은 교사가 평가에 대한 전문성이 낮은 수준임을 여실히 보여준다고 할 수 있다.

### ‘고등정신능력’ 측정의 미흡

선다형 평가 문항의 이원목적분류표에 근거하여 서술형 평가 문항의 이원목적분류표를 작성한 결과는 Table 6과 같다.

Table 6은 Bloom이 제시한 인지적 영역의 분류에 따라 서술형 평가 문항을 분석한 결과이다. 학교별로 인지적 영역의 분류 항목의 수에 다소 차이는 있었으나 대부분이 지식, 이해, 적용으로 문항을 분류하였다. 또한 같은 유형의 출제 문항일지라도 교사의 개인적인 분류 기준에 따라 다르게 분류되기도 하였다. Bloom은 지식, 이해를 단순정신능력이라 하고 적용, 분석, 종합, 평가를 고등정신능력이라고 하였다. 이에 따라 서술형 평가 문항을 분석하면 단순정신능력에 초점을 둔 문항은 23문항이고 고등정신능력에 초점을 둔 문항은 18문항으로 절반 이상의 서술형 평가 문항이 인지적 영역의 단순정신능력에 초점을 두어

**Table 6.** 2-Dimension of bloom's taxonomy

Type	Knowledge	Understanding	Application	Analysis	Synthesis	Evaluation
Unit	Various motion	3	2	4	2	-
	Characters of material	2	4	2	2	-
	The Earth and stars	5	-	4	1	-
	Structure and function of plant	2	5	3	-	-
Total		12	11	13	5	0

출제되어지고 있었다. 이는 학생들이 고등정신능력과 창의적 문제해결력을 측정하고자 하는 서술형 평가의 본연의 목적과 다르게 시행되고 있음을 보여준다.

다음은 교사의 이원목적분류표 작성에 대한 면담 내용이다.

연구자: 출제 문항을 제작하기 전에 이원목적분류표를 먼저 작성하시나요?

a교사: 전 나중에 작성하긴 하는데요. 문항 내기 전에 시험 범위 내에서 문항 수를 우선 나누고 문항을 제작해요.

b교사: 거의 대부분의 교사들이 나중에 작성하지 않을까요? (웃음) 문항 출제하는 게 더 급하다 보니깐 때로는 나중에 분류별 문항 수가 비슷하게 조절하기도 하죠.

c교사: 미리 작성하고 출제하지는 않지만, 이런 문제를 왜 내야하는지 정도는 생각하면서 문제를 내고 있습니다. 전체적으로 훑어보면서 어떤 내용을 출제할지 미리 선택한 후에 빠진 내용은 없는지 검토해요.

세 교사는 시험 범위 내에서 적절하게 문항 수를 조절하고 대략적인 출제 내용을 선정하는 정도로 이원목적분류표를 하였다. 평가 문항을 출제한 후에 실질적인 이원목적분류표를 작성하였다. 특히 B교사는 인지적 행동 영역 분류가 편중되면 임의대로 다른 분류로 수정하기도 하였다. 이처럼 교사들은 서술형 평가 문항을 출제할 때 서술형 평가가 측정하고자 하는 고등정신능력이라는 목표를 충실하게 반영하기 보다는 선다형 문항과 유사하게 단순정신 측정에 초점을 두고 있었다. 이는 서술형 평가라는 평가도구가 지향하고자 하는 평가도구의 목적에 대한 이해가 부족함을 보여준 것으로 평가 전문성이 부족함을 보여준다고 할 수 있다.

## 결론 및 제언

본 연구는 현재 학교에서 실시되고 있는 서술형 평가의 문항 특성을 알아보았으며, 이를 토대로 이러한 문항의 특성을 갖는 요인을 분석하였다. 이에 대한 연구결과는 다음과 같다.

첫째로, 서술형 평가 문항들은 출제 범위 중에서 고르게 출제된 것이 아니라 특정한 단원에서 편중되어 출제되었다. 이러한 요인으로 특정 단원이 채점하기에 용이한 답을 요구하기 유리하기 때문이었다.

둘째로, 서술형 평가 문항들은 단원에 대한 편중과 유

사하게 출제되는 형식이 ‘계산형’과 ‘이유 설명형’으로 집중되는 경향을 보였다. 이는 단답형 주관식이 아닌 서술형 문항을 출제하는 교사의 입장에서 이 두 가지 유형이 학생들의 다양한 응답을 제한할 수 있기 때문이었으며 이는 위에서 언급한 채점의 용이성과 연결되었다.

셋째로, 서술형 평가가 의도한 목적인 고등정신능력의 측정이 실제 학교에서는 제대로 이루어지고 있지 않았으며, 교사들이 출제한 서술형 평가는 단순정신능력 측정에 대한 문항이었다. 이는 평가 도구의 대한 이해의 부족이 기인한 것으로 보인다.

이러한 본 연구의 결과를 토대로 얻을 수 있는 결론은 다음과 같다. 위의 서술형 평가 문항의 특성이 갖는 요인을 분석하면 교사들은 평가 문항 제작, 평가 도구의 이해 등에 대한 능력이 다소 부족함을 알 수 있다. 이러한 요소들은 모두 평가 전문성의 결여로 설명될 수 있다. 이를 통해 교사들의 평가 전문성 신장을 위한 방안이 요구됨을 알 수 있다.

이 연구의 결론을 토대로 교사들의 평가 전문성 신장을 위한 제언은 다음과 같다. 첫째로, 우리가 평가의 결과를 매우 중요시 여기며, 이를 통해 여러 의사결정을 하는데 반해 교원양성이나 교원연수과정에서 실제적인 평가와 관련된 연구와 교육이 체계적으로 실행되지 않았다. 그렇기 때문에 예비교사나 현직교사들이 평가 전문성을 신장할 수 있는 기회가 교과 내용에 대한 전문성을 신장하는 기회에 비해 상대적으로 더 적었다고 할 수 있다. 교사의 평가 전문성 신장을 위해 이러한 기관에서 체계적 연구와 교육이 선행되어야 할 것이다.

둘째로, 서술형 평가 문항 특성의 요인 분석에서 교사들에게 일관되게 나온 응답은 채점의 용이성이었다. 점점 평가가 중요시 되며, 평가 결과에 따른 영향이 크기 때문에 채점의 공정성이 교사에게 강조되고 있다. 교사는 이러한 상황에서 어느 누가 채점해도 동일한 결과를 얻을 수 있는 객관적인 정답이 요구되는 문항을 제작할 수밖에 없으며, 이는 교사의 평가 전문성 신장에 위협을 주는 요인이다. 같은 단원을 수업하더라도 교사마다 다양한 수업이 이루어질 수 있으며, 이러한 수업전문성에 대해 이의를 제기하는 사람은 없다. 하지만 유독 평가에 대한 통일성의 강조는 크다고 할 수 있다. 성취평가제가 도입되는 등 학교에서의 평가가 서열 판별에서 학습목표 도달의 측정으로 변화하고 있다. 이러한 환경변화는 교사가 서열 판별을 위해 평가의 채점용이성을 강요받게 되고, 이에 따라서 교사는 학습목표의 도달 및 평가도구의 측정하고자 하는 능력이 아닌 채점이 용이한 문항을 제작하는 기존의 환경에서 다시 벗어나 실제 측정하고자 하는 평가의 목적을 달성할 수 있는 여건을 마련하였다고 할 수 있다. 서술형



평가 문항을 채점의 용이성에서 벗어나 관련 단원의 교육 목표에 맞게 제작하고, 이에 따라 행해진 서술형 평가 채점의 결과를 신뢰하는 등의 사회적 분위기 역시 교사의 평가 전문성을 위해 필요한 요소라 할 수 있다.

**Acknowledgments.** Publication cost of this paper was supported by the Korean Chemical Society.

## REFERENCES

1. Beak, S. G. *Performance Test*; Kyoyookbook: Gyeonggi-do, 2002.
2. Choi, S. W. The Effect of Method in an Essay Formative Testing Science Studies. Thesis, Ewha Woman University, 2001.
3. Hong, A. N. Effects of Science Writing Instruction using Visual Materials on High School Student's Achievement and Awareness toward Science Essay Test. Thesis, Kwangwon University, 2013.
4. Yang, G. S. *A Direction of Student Evaluation and Manufacture of Descriptive Assessment*; Pusan Education Office: Pusan, 2006.
5. Do, J. W.; Oh, J. Y.; Kong, J. I.; Joo, M. J.; Kim, M. Y.; Lee, D. H.; Bak, M. G. *Edu. Primary School Math.* **2009**, 12, 63.
6. Kim, L. Y.; Lee, M. H. *J. Edu. Res. Math.* **2013**, 23, 533.
7. Kim, S. H. *J. Korean Edu.* **2013**, 40, 479.
8. Sim, A. S. Study on the Recognition of Teachers and Students & Actual Conditions about Middle School Science Essay Test. Thesis, Korea National University of Education, 2008.
9. Choi, S. W. The Effect of Method in an Essay Formative Testing Science Studies. Thesis, Kwangwon University, 2011.
10. Fullan, M. *The New Meaning of Educational Change*; Columbia University Press: London, 2001.
11. Edward, G. L. *J. Sci. Edu.* **2013**, 35, 1208.
12. Glaser, R.; Chi, M. T. *Nature of Expertise*; Erlbam: Hillsdale, 1988.
13. Dreyfus, H.; Dreyfus, S. *Mind Over Machine: The Power of Human Intuition and Expertise in the Era of the Computer*; The Free Press: New York, 1988.
14. Kun, T. S. *The Structure of Scientific Revolutions*; University of Chicago Press: Chicago, 1970.
15. Black, P.; Wiliam, D. *Phi. Delt. Kap.* **1998**, 80, 139.
16. Hill, M.; Cowie, B.; Gilmore, A.; Smith, L. F. *Ass. Matt.* **2010**, 2, 43.
17. Popham, W. J. *Assessment for Education Leaders*; Allyn & Bacon: Boston, 2009.
18. National Research Council. *National Science Education Standards*; National Academy Press: New York, 2001.
19. Kim, S. Y. *J. Edu. Eval.* **2007**, 20, 1.
20. Lee, J. K.; Rim, N. R. Whang, S. S. *J. Sci. Edu.* **2004**, 29, 115.
21. Kim, Y. H.; Kim, Y. S. *Bio. Edu.* **2012**, 40, 167.
22. Hong, M. J.; Chung, H. S. *J. Sci. Edu.* **2006**, 30, 65.
23. Hong, S. Y.; Kim, M. K.; No, S. S.; Kwon, J. R. *J. Math. Edu.* **2008**, 18, 335.
24. Kim, S. H.; Kim, S. Y.; Kim, J. C.; Ban, J. C.; Bak, S. G.; Seo, M. W. *Education Evaluation for Teacher*; Hwaksisa: Seoul, 2010.
25. Stiggins, R. J.; Conklin, N. F. *In Teachers' Hands: Investigating the Practices of Classroom Assessment*; SUNY Press: New York, 1992.
26. Bell, B.; Cowie, B. *Formative Assessment and Science Education*; Kluwer Academic Publishers: Dordrecht, 2001.
27. Fusco, D.; Barton, A. C. *J. Res. Science Teaching.* **2001**, 38, 337.
28. Bak, G. R.; Bang, J. S. *Math. Edu.* **2008**, 47, 181.
29. Yu, J. I.; Shin, H. J. *J. Teacher Edu.* **2008**, 23, 79.
30. Bak, D. S. *Education Evaluation: Understanding and Application*; Education Science Press: Seoul, 2007.