

## 초등과학영재학생의 과학창의성에 대한 자기 평가, 동료 평가의 비교 분석

김민주 · 임채성<sup>†</sup>

### A Comparative Analysis of Student Self- and Peer-Assessments of Elementary Science-Gifted Students' Scientific Creativity

Kim, Minju · Lim, Chaeseong<sup>†</sup>

#### ABSTRACT

This study aims to compare self- and peer-assessments of science-gifted elementary students' scientific creativity. A science-gifted program on the Pascal's principle was implemented to 40 fifth-graders in the Science-Gifted Education Center for two weeks. After that, students presented their results from a scientific creativity task using the principle in class. The task was to devise a new and useful tool using the principle, and it included the students' self-assessment about their idea. During presentation, students were asked to assess the works of peers and write down the reasons that they gave the scores they gave. Shortly, student self- and peer-assessments about students' scientific creativity outcomes were compared. Based on two essential components of creativity, ideas that satisfy both originality and usefulness can be counted as scientifically creative. The main results of this study are as follows: First, the average scores of student self- and peer-assessments were 71.5 and 61.9. Second, the standard deviations of student self- and peer-assessments were 14.47 and 5.79. Third, among scientific creativity, originality, usefulness scores, only originality had a significant correlation between student self- and peer-assessment ( $r=.42$ ). Fourth, the students were categorized into four groups according to the levels of their scores by student self- and peer-assessment. And the frequencies of peer-assessment group had a significant difference at  $p<0.05$  level, according to self-assessment group (Chi Square=4.0000,  $df=1$ ,  $p=0.0455$ ). Fifth, through a case study by group, the results suggesting that self-assessment could be affected by the students' self-efficacy and perfectionism and such effect could also influence peer-assessment have been found. The result showed that how the student self- and peer-assessment of scientific creativity are different and what the students' thoughts on the evaluation of scientific creativity are. The findings suggested that there are several things to consider for the educators to make efforts to construct consistent assessment methods for scientific creativity.

**Key words:** science gifted, scientific creativity, student self-assessment, peer-assessment

#### I. 서 론

창의성은 개개인이 일상에서 마주하는 문제를 해결하는 데 필요한 귀중한 자산이며, 개인과 사회의 발전에 모두 기여한다(Besançon *et al.*, 2015). 21세기 핵심 기술 중 하나인 창의성은 많은 연구자들의 관심을 받고 있는 주제지만 그 메커니즘이 완전

히 규명되지 않았으며, 복잡한 개념인 만큼 이를 정의하는 데 많은 요소가 관여하고 있다(Weisberg, 1993). 그러나 대부분의 연구는 공통적으로 독창성과 유용성을 창의성의 요소로 포함하고 있으며, 교육학과 심리학 분야에서 창의성은 특정 과제나 영역에 있어 적절하고 가치 있으면서 독창적인 것을 생산하는 능력으로 정의된다(Runco & Jaeger, 2012;

이 논문은 2019년도 서울교육대학교 교내연구비에 의하여 연구되었음.

2019.8.25(접수), 2019.9.23(1심통과), 2019.9.25(2심통과), 2019.9.27(최종통과)

E-mail: cslim@snu.ac.kr(임채성)

Mumford, 2003). 여기서 특정 과제나 영역이 요구하는 조건을 만족시켜야 창의적이라고 보는 것은 영역 특수적인 지식과 기능들이 창의성 발휘에 중요한 영향을 미친다는 입장인 영역 특수적 입장을 반영한 것이다(Hu & Adey, 2002). 이상으로 미루어 볼 때, 창의성이 발휘되는 분야와 창의성을 정의하는 두 가지 속성을 반영하면 과학창의성을 ‘과학적으로 새로우면서도 유용한 산물을 만들어내는 능력’이라고 할 수 있다.

창의성의 본질을 탐구하기 위해서는 창의성의 정의뿐만 아니라, 창의성을 발휘하는 주체 또한 고려해야 한다(임채성, 2012). 창의성을 연구하는 학계에는 크게 두 흐름이 있는데, 하나는 저명한 창의적 인물과 그 업적을 연구하는 것이고, 다른 하나는 보통의 사람이 일상생활에서 발휘하는 창의성을 연구하는 것이다. 임채성(2012)은 과학계에서 새로우면서 유용한 것을 만들어내는 능력인 Big-C 창의성과 학생 수준에서 과학적으로 새로우면서 유용한 것을 만들어내는 능력인 little-c 창의성을 대조하고, 초등과학교육에서는 little-c 접근법이 타당하다고 하였다. 한편, 이 이분법적 관점을 확장하여 개인 내부에서, 발달적으로 나타나는 창의성을 설명하고자 나타난 것이 Four C 모델의 mini-c 개념이다(Kaufman & Beghetto, 2009).

Runco (2007b)는 little-c 관점이 학생의 창의성을 분류하는 데 있어 제한적이라 발견하고 양육하기 보다는 간과하게 된다고 주장하였다. 예를 들어 초등학교 4학년 학생이 왜 명왕성을 행성으로 분류해야 하는지 혹은 하지 말아야 하는지에 대해 독특하면서 개인적으로 의미 있는 아이디어를 떠올릴 수 있다. 만약 이것이 Big-C의 대안인 little-c 범주로 분류된다면 이에 대해 토론하는 천문학 전공 교수나 대학원생과 함께 분류되어야 한다. 이는 아직 지식과 경험이 부족한 학생에게 불공평한 기준이다(Runco, 2004). 또한 의사소통 시 아이디어를 더 효과적으로 나타낼 수 있는 학생과 비교할 때도 마찬가지다. mini-c는 실제적인 형태로 나타나지 않았지만 개인 내부에 존재하는 아이디어 수준의 것이라도 창의적인 것으로 간주할 수 있다. 아이디어를 얼마나 잘 표현하는지가 아닌, 아이디어 자체에 대한 평가가 mini-c 창의성을 나타내는 것이다(Beghetto & Kaufman, 2007).

자기 평가는 이러한 mini-c 창의성을 가장 잘 평

가할 수 있는 방법으로서(Kaufman & Beghetto, 2009), 세 가지 측면에서 유용하다. 첫째로 학생으로 하여금 자신의 창의성에 대해 반성적으로 사고하게 한다. 둘째, 교육자는 이를 통해 창의적 자기효능감 혹은 창의적 잠재력과 그 영역에 흥미를 가진 학생을 식별할 수 있다. 셋째, 발달적인 측면에서 학생의 창의성 수준을 파악하고, 적절하게 피드백해 줄 수 있다. 어느 한 아이디어를 떠올릴 때, 그것을 스스로 생각해낸 것인지, 비슷한 다른 아이디어에서 차용한 것인지는 학생 스스로가 가장 잘 알 것이다. 과학적 아이디어는 발견된 것이 이미 많이 존재하므로 비슷한 다른 아이디어가 존재하는 줄 모르고 스스로 새롭다고 생각하면서 아이디어를 떠올리는 경우도 있을 것이다. 타인이 보기에 새롭지 않지만 학생이 이전에는 스스로 생각하지 못했던 것을 떠올린다면, 발달적인 측면에서 볼 때 그 가치를 인정해 주고 격려할 필요성이 있는 것이다.

동료 평가는 자기 평가보다 객관적인 기준이면서 같은 초등학교 수준에서 평가하므로 교사나 부모와 같은 외부적 입장에서의 창의성 평가의 대안이 될 수 있다(Newton, 2010). 임채성(2012)이 강조한 초등학교 수준의 little-c 창의성을 탐구하기 위해서는 동년배 학생들의 창의성 인식을 조사할 필요가 있다. 지식과 사고 기능 수준이 비슷한 또래 학생들의 입장에서 한 아이디어에 대한 의견이 어떠한지 조사하는 것은 진정한 의미에서 초등학교 수준의 과학창의성을 평가하는 방법인 것이다.

자기 평가와 동료 평가 각각을 조사하는 것과 더불어 두 평가 간 결과가 어떻게 달라지는지 비교할 필요가 있다. 학생 창의성 발달의 측면을 확인할 수 있는 자기 평가는 그 자체로 의미가 있지만, 매우 주관적이라는 단점이 있다. 동료 평가는 같은 발달 수준에 있는 평가자의 평가이면서 좀 더 객관적이므로 자기 평가를 보완할 수 있다. 자기 평가와 동료 평가가 어떤 양상으로 달라지는지, 두 관점 간 차이가 어디에서 비롯하는지 알 수 있다면 학생의 과학창의성을 발달 수준에 맞게 평가하는 방식을 마련하는 데 시사점을 얻을 수 있을 것이다.

그러면 학생의 과학창의성은 어떻게 평가할 것인가? 창의성의 평가 방식에는 학생의 창의성에 대한 인지·정의·기능적 요소를 종합적으로 평가하는 방식, 독창성, 유창성, 정교성, 유연성 등 지금껏 밝혀진 다면적 요소로 평가하는 방식 등 다양한 방

식이 존재한다(박병기와 유경순, 2000). 앞서 밝힌 영역 특수적 입장에 따라 정의한 과학창의성을 평가하는 방식은 바로 과학 교과에서 학생이 산출한 산물에 대하여 평가 공식을 적용하여 평가하는 것이다(임채성, 2014). 이는 학생의 내적 특성, 환경과 같은 여러 요인이 아닌 산물 그 자체에 대한 단일한 요인에 대한 평가이므로 자기 평가와 동료 평가의 비교를 용이하게 한다. 또한 독창성과 유용성 요소에 대한 점수만 측정하므로 초등학생이 평가 방식을 쉽게 이해하고 사용할 수 있다.

Renzulli (2003)에 의하면 평균 이상의 지능, 과제 집착력, 창의성의 세 가지 요소가 영재성을 정의한다. 그러므로 과학 분야에서 창의적 잠재력을 갖춘 과학 영재 학생 집단에서 과학적이고 창의적인 아이디어가 다양하게 생성될 가능성이 높다(Callahan & Miller, 2005). 이들의 다양한 창의적 산물을 분석하고, 이들이 자신과 또래의 과학창의성을 어떻게 인식하고 있는지를 검토한다면 일반 학생의 과학 창의성 함양을 위한 목표와 평가 체계를 세우는 데 시사점을 얻을 수 있을 것이다.

현재 교육현장에서 평가는 대부분 교사에 의해 이루어지고 있지만, 학생의 개별적이고 발달적인 측면을 반영하기 위해서는 학생의 관점을 고려할 필요가 있다(Barbera, 2009; Barrett, 2007). 또한 창의성에 있어 교사는 학생의 과학창의성에 대해 보다 엄격한 기준을 적용하여 평가하는 경향이 있으므로 학생의 잠재적인 창의성이 간과될 가능성이 높다(김민주와 임채성, 2018). 과학창의성에 대한 학생의 자기 평가, 동료 평가를 연구하는 것은 학습 과정에 내재된 초등학생 수준의 창의성을 발견하고 양성할 수 있게 할 것이다. 그러므로 본 연구에서는 학생 자신이 인식하는 과학창의성과 또래 학생이 인식하는 과학창의성을 비교하고, 이를 통해 학생 발달 수준에 맞춘 평가 방식 마련에 시사하는 바를 얻고자 한다.

## II. 연구 방법 및 절차

### 1. 연구 설계

본 연구는 초등과학영재학생을 대상으로 파스칼의 원리에 대한 과학영재 프로그램을 2주 동안 실시한 후, 그 원리를 적용하는 과학창의성 과제에 대한 학생의 자기 평가와 동료 평가 결과 간 관련

성을 분석하였다. 3주차 수업에서 학생은 과학창의성 과제를 돌래 학생들에게 실물화상기로 제시하고 구두 발표하였으며, 또래 학생들은 각 발표에 대한 과학창의성 점수와 그렇게 평가한 이유를 함께 기록하였다. 점수로 표현되는 정량적 평가 결과와 더불어 그렇게 평가한 이유, 동료 평가 활동 소감과 같은 정성적 데이터도 수집하여 분석하였다. 또한 모든 연구 과정은 초등 과학 교육 및 초등 과학 영재 교육 석·박사 전공교사 및 전공교수의 자문과 협의 과정을 거쳐 진행되었다.

### 2. 연구 대상

자기 평가와 동료 평가를 비교 분석하는 본 연구에서 초등과학영재학생을 연구 참여자로 선정할 까닭은 첫째, 과학영재학생이 자신과 또래 학생의 과학창의성을 타당하고 신뢰성 있게 평가할 능력을 갖추었다고 판단하였기 때문이다. 둘째, 과학영재학생은 평가 활동을 하면서 그렇게 평가한 이유를 명확히 기술할 수 있고, 평가 활동 후 평가 방식의 개선해야 할 점과 같은 의견을 다양하게 제시할 수 있다고 보았기 때문이다. 셋째, 과학 지식과 탐구능력을 일정 수준 이상 갖춘 과학영재학생으로부터 창의적 아이디어의 사례를 최대한 다양하게 수집할 수 있기 때문이다. 넷째, 아이디어가 다양한 만큼 그들의 과학창의성 점수가 다양하게 분포할 것이라고 추론했기 때문이다.

본 연구에 참여한 학생들은 서울특별시 교육지원청 소속 과학영재원의 초등학교 5학년 학생들이다. 교육지원청 소속 과학영재원의 선발과정은 크게 세 단계로 나뉜다. 학생들은 먼저 각자의 학교에서 담임교사 및 학교추천위원회의 추천을 거친 후, 1단계 창의적 문제해결력 평가와 2단계 면접 평가를 통해 선발된다. 이처럼 여러 단계를 거치고 각 학교에서 수집된 교사 관찰 체크리스트, 자기 평가 체크리스트 등 다양한 정보를 종합적으로 고려하여 선발하는 과정이 학생의 다각적인 측면을 고려한다고 판단하여 교육지원청 소속 과학 영재 학생을 연구 대상으로 선정하였다. 또한 과학을 학습하는 초등 3~6학년 학생 중에서 과학적 지식과 탐구 능력이 일정 수준 이상 갖추어져 있고, 자신의 생각을 분명하게 표현하여 면담이 비교적 용이한 5학년 학생을 연구 대상으로 선정하였다.

연구 참여자의 남녀 구성은 남학생 29명과 여학

생 11명이었고, 20명씩 2개 학급으로 이루어져 있다. 40명의 연구 참여자 중 수업에 참석하지 않은 1명의 학생이 있어 총 39명의 활동 결과를 분석하였으며, 자신이 산출한 결과물에 대한 자기 평가를 하지 않은 3명의 학생이 있어 자료 집계 방식에 따라 표본 수가 36명 혹은 39명으로 달라진다.

동료 평가는 수업 중 학생들의 발표를 겸해서 진행되었으므로 1개 학급 내에서 이루어졌다. 즉, 20명으로 이루어진 1개 학급에서는 1명이 발표하고 나머지 19명이 그 학생의 아이디어를 채점했다는 뜻이다. 그러나 학급에 따라 결석하거나 지각한 학생, 이유는 알 수 없지만 특정 학생에 대하여 동료 평가를 누락한 학생이 있어 19명보다 적은 학생이 채점하였다. 그러므로 학생이 제시한 하나의 아이디어에 대해서 1개의 자기 평가 결과와 19개의, 혹은 그보다 적은 개수의 동료 평가 결과가 존재한다.

### 3. 자료 수집 및 절차

본 연구는 학생의 과학창의성 과제, 과제에 대한 학생의 자기 평가 점수, 동료 학생의 발표를 들으면서 기록한 동료 평가 점수, 각 평가에 대한 이유, 활동 후 학생 소감 등 정량적, 정성적 데이터를 아우르는 자료를 수집하였다.

#### 1) 과학창의성 평가를 위한 과제 및 활동

학생이 다양하게 아이디어를 제시할 수 있고, 생활과 연계되어 응용이 용이한 파스칼의 원리를 과제의 주제로 선정하였다. 파스칼의 원리는 포클레인, 프레스, 리프트와 같이 실생활에서 쉽게 접할 수 있는 유압 기계 장치의 기본 원리이기도 하다. 연구자는 파스칼의 원리를 학습하는 영재 수업 프로그램을 기획하고, 파스칼의 원리를 응용하여 새롭고 유용한 것을 제시하도록 하는 과학창의성 평가 과제를 개발하였다. 본 연구에 참여한 학생들은 3주의 프로그램 중 첫 2주는 파스칼의 원리를 학습하면서 이를 적용한 유압 프레스, 유압 2관절 로봇 팔과 같이 생김새와 기능이 다양한 장치를 만든 다음, 그 장치를 실생활에 활용할 아이디어를 구상하는 과제를 수행하였다. 2주의 수업 후 배운 것을 종합하여 파스칼의 원리를 실생활에 적용한 아이디어를 자유롭게 구상하는 과제를 주었고, 마지막 주에 그 결과물을 발표하는 시간을 가졌다.

학생은 과제에서 자신이 만든 도구나 장치의 이

름, 그것의 사용 목적 및 사용자를 적고, 그 설계 디자인을 글과 그림으로 제시한다. 동시에 자신의 아이디어에 대한 평가로 독창성과 유용성 점수를 매기는데, little-c 창의성 측면에서 독창성과 유용성은 각각 ‘새롭다고 생각하는 정도’, ‘유용하거나 타당하다고 생각하는 정도’와 같이 초등학생 수준에서 이해하기 쉬운 말로 대체하였다. 또한 과제를 제시할 때 창의성을 평가하는 과제라는 암시를 주지 않도록 창의성, 독창성, 유용성과 같은 용어는 사용하지 않았다.

#### 2) 과학창의성의 자기 평가, 동료 평가 산출 방식

Treffinger (2009)에 따르면 창의성 평가 방법은 100여 가지로 다양하다. 창의성에 대한 접근방법이 다양하기 때문에 그 방법에 기초한 검사 도구 또한 다양하나, 크게 나누어서 인지검사, 성향검사, 창의적 산물검사의 3가지 부류로 나눌 수 있다(최인수, 2000). 본 연구는 창의성을 영역 특수적인 것으로 보고, 과학창의성에 한정하여 연구문제를 수립하였다. 또한 학생이 가진 창의성의 잠재적인 요소를 종합적으로 보는 것이 아니라, 과학 교과에서 학생이 산출한 아이디어 자체에 대한 평가를 실시하므로 창의적 산물 검사에 기초한 임채성(2014)의 과학창의성 평가 공식을 사용하였다.

평가 공식은 창의성의 두 가지 핵심 요소인 독창성(참신성)과 유용성(적절성, 타당성, 과학성) 점수의 곱을 과학창의성 점수로 산출한다. 곱으로 계산하는 이유는 창의성이 두 요소의 교집합(‘original and useful’, not ‘original or useful’)으로 정의되기 때문이다. 즉, 이 두 요소 중 하나의 요소만 충족시키는 것이 아니라, 두 요소를 모두 충족시켜야 창의적이라고 할 수 있다. 임채성(2014)은 학생과학창의성 혹은 초등과학창의성을 초등학생 수준(little-c 창의성)에서 과학적으로 새롭고(독창적인, 참신한), 유용한(적절한, 과학적으로 타당한) 아이디어나 산물을 만들어내는 능력이라고 정의하고, 이 정의에 따라 다음과 같은 과학창의성 평가 공식을 개발하였다.

$$SC = 10 \times \left(1 - \frac{n-1}{N-1}\right) \times U$$

위 식에서 SC는 과학창의성(scientific creativity)을, 왼쪽의  $[10 \times \{1 - (n-1)/(N-1)\}]$ 항은 아이디어의

새로운 정도, 즉 독창성을 의미하고, 오른쪽의 U는 아이디어의 유용한 정도, 즉 과학적 타당성을 의미한다. N은 집단의 전체 아이디어 수이고, n은 그 집단에서 특정 아이디어가 제시된 수이다. 왼쪽 항은 0에서 10 사이의 값을 가지며, 10에 가까울수록 그 아이디어나 산물이 독창적임을 나타낸다. 오른쪽의 U 또한 0에서 10 사이의 값을 가지며, 10에 가까울수록 그 아이디어나 산물이 유용하면서 과학적으로 타당함을 뜻한다.

본래 창의성 공식은 전체 아이디어의 수라는 N과 아이디어를 비슷한 것끼리 범주화하면 얻을 수 있는 수인 n을 통해 보다 객관적인 독창성 점수를 얻을 수 있는 공식이나, 본 연구에서는 독창성 항과 유용성 항의 곱이라는 공식의 기본적인 아이디어만을 차용하였다. 주관성이 개입될 여지가 있지만 점수에 담긴 학생의 생각과 의도를 파악하는 것이 본 연구의 주제이기 때문이다.

따라서 본 연구는 학생이 자신의 아이디어와 또래 학생의 아이디어에 대해 독창성과 유용성 점수를 매긴 것을 수집하였고, 과학창의성 점수는 학생이 매긴 독창성과 유용성 점수를 곱하여 자동으로 산출하였다. 즉, 학생이 독창성과 유용성 각각에 대해 100점 만점으로 직접 평가한 자료를 수집하였고, 이를 10점 만점으로 환산하여 각 점수는 0~10점 사이 소수점 첫째 자리까지로 책정되었다. 이렇게 책정된 독창성 점수와 유용성 점수를 곱하면 0에서 100 사이의 값을 가지는 과학창의성 점수가 산출된다.

자기 평가는 자신의 창의적 과제를 수행한 후 기록하게 하였고, 동료 평가는 학생결과물을 발표하는 수업에서 다른 학생의 발표를 들으면서 각 결과물에 대해 점수를 기록하게 하였다. 평가에 참여한 학생들은 왜 두 요소가 과학창의성을 구성하는 요소인지, 두 요소 사이에 어떤 관계가 있는지, 과학창의성의 의미와 평가 방식을 미리 숙지하고 평가에 임하였다. 자기 평가와 동료 평가 모두 독창성과 유용성 각각에 대해 아이디어를 보고 바로 떠오르는 점수를 매기도록 하였으며, 급간에 대한 평가 기준을 정하지는 않았다.

이러한 직관적(intuitive) 평가는 의식적 추론 없이 이루어지는 것으로 총체적(holistic) 평가라고도 한다(Colman, 2003). 학생의 창의성을 대상으로 한 Cropley (2001), 학생의 작곡을 평가한 Hickey (2001)

등의 연구에서 이러한 평가가 쉽고 간편하며 어느 정도 신뢰성과 타당성이 있었음을 밝힌 바 있다. 그러나 평가의 내용, 평가의 주체, 평가 상황의 맥락에 따라 이와 다른 결과가 나타날 수 있음을 유념해야 한다. Newton (2010)은 학생의 과학창의성에 대한 교사의 총체적이고 직관적 평가가 타당성이 있었지만, 신뢰성은 매우 떨어졌음을 밝혔다. 그 연구에서 처음에 과학창의성을 하나의 전체로 보고 평가하게 하였다가 새로움, 타당성, 정교성의 특정한 평가로 나누어 평가하게 하자 신뢰도가 약간 증가하였다고 한다. 본 연구에서는 창의성의 두 요소인 독창성과 유용성으로 나누어 각 요소에 대한 평가를 하므로 신뢰도가 어느 정도 담보될 것이라는 전제 하에 총체적 평가를 실시하였다. 평가자간 신뢰도에 대한 논의는 연구 결과에서 할 것이다.

### 3) 면담 및 그 외 정성적 자료

위와 같은 정량적 자료 외에도 학생 면담 결과, 학생이 서술한 평가 이유와 같은 정성적 자료를 수집하고, 정량적 자료와 연계하여 분석하였다. Patton (2002)은 어느 하나의 방법만으로는 복잡한 현상을 해석하기 어렵기 때문에 면담과 관찰을 병행하여 사용할 필요가 있다고 하였다. 본 연구에서는 학생이 과학창의성 평가 방식에 대한 의견을 자연스럽게 말할 수 있도록 자유롭게 활동 소감을 나누는 면담과 학생이 자신과 동료의 창의성 점수를 기록할 때에 그 옆에 그렇게 평가한 이유를 자유롭게 하는 방식을 사용하였다.

## 4. 자료 분석 방법

자기 평가, 동료 평가 결과는 과학창의성 및 그 하위요소인 독창성과 유용성에 대하여 I-STATistics 프로그램으로 종속 표본 t검정을 실시하여 평균과 표준편차를 비교하였으며, 상관관계와 빈도분포, 동료 평가의 신뢰도 또한 분석하였다. 또한 자기 평가와 동료 평가 점수에 대하여 중앙값을 기준으로 수준을 나누어 상상, 상하, 하상, 하하의 네 집단으로 구성하고, 수집된 정량적 데이터와 정성적 데이터에서 나타나는 각 집단별 특성을 분석하였다. 이때 사용된 정량적 데이터는 과학창의성의 하위요소인 독창성과 유용성 점수이며, 정성적 데이터는 학생이 제시한 평가 점수의 이유, 면담을 통해 수집한 학생의 과학창의성 평가 활동에 대한 소감

이다. 이 데이터들을 종합적으로 분석하여 추후 과학창의성 평가 방식의 방향에 대한 시사점을 논의하고자 한다.

### III. 연구 결과 및 논의

#### 1. 초등과학영재학생의 과학창의성 및 그 하위 요소에 대한 자기 평가, 동료 평가의 비교

##### 1) 동료 평가의 신뢰도 분석

초등과학영재학생의 과학창의성에 대한 동료 평가는 채점자가 실질적으로 한 사람인 자기 평가와 달리 채점자 간 신뢰도를 분석할 수 있었다. 과학창의성 점수는 독창성 점수와 유용성 점수를 곱하여 산출되므로 학생이 직접 매긴 독창성, 유용성 점수를 대상으로 채점자 간 신뢰도를 파악하였다. 하나의 아이디어에 대한 평정자 간 일치도는 Table 1과 같다.

동료 평가는 20명씩 2개 반에서 진행되었고, 이날 결석한 학생이 각 반에 1명씩 있었으므로 표본은 한 반 당 19개이다. 이렇게 해서 채점자 간 신

**Table 1.** Inter-rater reliability of peer-assessment on originality, and usefulness

		(N=19)
변인		상관도
동료 평가 신뢰도	1반 독창성	.34
	2반 독창성	.41
	1반 유용성	.40
	2반 유용성	.50

**Table 2.** The result of paired sample *t*-test of student self-, and peer-assessment on scientific creativity and its two essential components

변인	평균	표준편차	사례수	<i>t</i>	<i>p</i>
자기평가 과학창의성	71.5	14.47	36	3.564**	0.001
동료평가 과학창의성	61.9	5.79			
자기평가 독창성	8.4	1.32	36	1.953	0.059
동료평가 독창성	8.0	0.51			
자기평가 유용성	8.5	1.04	36	3.638***	0.001
동료평가 유용성	7.7	0.64			

\*\*  $p < 0.01$  \*\*\*  $p < 0.001$ .

뢰도(Spearman-Brown formula; Rosenthal & Rosnow, 1991)는 독창성이 1반은  $R=.34$ , 2반은  $R=.41$ , 유용성은 1반이  $R=.40$ , 2반이  $R=.50$ 으로 충분히 신뢰할 만한 기준인  $R=.70$ 을 넘지 못한 값이다. 신뢰도가 높지 않은 이유로 채점자 수가 많다는 것, 평가 기준이 명확하지 않다는 것 등이 있을 것이다. 학생의 과학창의성에 대한 교사의 직관적인 평가가 타당성이 있었지만 신뢰성은 떨어졌음을 밝힌 연구(Newton, 2010)에서 처음에 과학창의성을 하나의 전체로서 평가하게 하였다가 새로움, 타당성, 정교성의 요소로 나누어 평가하게 하자 신뢰도가 약간 증가하였다. 본 연구는 과학창의성의 평가를 독창성과 유용성의 두 요소로 나누었으나, 급간 채점 기준을 설정한다면 신뢰도가 이보다 더 증가할 것으로 판단된다.

##### 2) 과학창의성 및 그 하위 요소에 대한 자기 평가와 동료 평가 결과 비교

초등과학영재 학생의 과학창의성에 대한 종속 표본 *t*검정을 실시하여 자기 평가, 동료 평가 결과 간 평균의 유의한 차이를 양방 검증한 결과는 Table 2와 같다. 먼저 과학창의성과 독창성 및 유용성 요소의 평균과 두 평가 방식 간 유의한 차이를 확인하고, 각 평가의 표준편차를 분석하고자 한다.

100점 만점으로 측정되는 과학창의성의 평균값은 자기 평가 71.5점, 동료 평가 61.9점으로 통계적으로  $p < 0.01$  수준에서 유의한 차이가 있었다( $t=3.5635$ ,  $df=35$ ,  $p=0.0011$ ). 자기 평가 점수가 동료 평가 점수보다 높다는 것은 학생은 자신에 대해 관대한 평가 기준을 적용한다고 밝힌 선행 연구와 일치하는 결과이다(Ross, 2006; Knowles *et al.*, 2005).

mini-c 관점에 따르면 창의성은 개인의 경험, 사건, 현상에 대한 새롭고 의미 있는 해석을 모두 포함하는 개념이다. 이미 밝혀진 지식이나 아이디어라 할지라도 개인이 나름대로 새롭게 재구성한 아이디어라면 창의적이라고 보는 것이다. 그래서 mini-c를 평가하는 대표적인 방법이 자기 평가다. 반면, little-c 관점은 학습자의 초보적 수준을 고려하지만 비교적 객관적인 관점에서 창의성을 바라보고 있다. 학생이 자신의 아이디어를 나름 새롭고 의미 있는 것으로 본다고 할지라도 타인이 보기에 그것은 그다지 새롭지 않거나, 의미 있지 않을 수 있다. 종합하면 위 결과는 자신의 창의성에 대한 해석에 있어 자신의 관점과 타인의 관점에 차이가 있고, mini-c 관점이 little-c 관점보다 관대하고 포괄적임을 나타낸다.

과학창의성의 두 하위 요소인 독창성과 유용성 점수에 대해서도 두 평가 관점의 평균값을 비교 분석하였다. 10점 만점으로 측정되는 독창성 점수의 평균값은 자기 평가 8.4점, 동료 평가 8.0점으로, 과학창의성 점수와 마찬가지로 자기 평가 점수가 더 높지만 통계적으로 유의한 차이는 없었다. 10점 만점으로 측정되는 유용성 점수의 평균값은 자기 평가 8.5점, 동료 평가 7.7점으로 역시 자기 평가 점수가 더 높았는데, 여기서는 통계적으로  $p < 0.001$  수준에서 유의한 차이가 있었다( $t = 3.6380, df = 35, p = 0.0009$ ).

이 결과를 통해 학생은 과학창의성의 유용성 요소에 있어 자신보다 동료에게 보다 엄격한 기준을 적용한다고 해석할 수 있다. 이는 Sowden and Dawson (2011)의 연구에서 엄격한 평가자는 독창성보다 유용성에 대한 기준을 더 높게 설정한 것과 관련지을 수 있다. 본 연구에서 자신보다 타인을 평가할 때 학생은 더 엄격해지는 경향이 있기 때문이다. 이는 학생의 과학창의성에 대한 교사 평가와 자기 평가를 비교한 연구에서 교사가 학생보다 유용성의 기준이 더 높았다는 연구 결과와 상통하는 부분이다 (김민주와 임채성, 2018).

위 결과를 종합하면 자기 평가와 동료 평가 과학창의성 점수의 차는 유용성 요소로 인해 벌어졌음을 알 수 있다. 즉, 차이가 많이 나는 유용성 점수로 인하여 독창성과 유용성 점수의 곱으로 산출되는 과학창의성 점수 차가 벌어진 것이다. 유용성은 가치 판단의 문제로 인해 평가자의 주관에 반영될 가능성이 높기 때문에 점수 차가 크게 나타날 수

있다.

다음으로 과학창의성과 독창성 및 유용성 요소의 표준편차를 분석하였다. 과학창의성, 독창성, 유용성 점수 모두 자기 평가의 표준편차가 더 컸다. 김민주와 임채성(2018)의 연구 역시 자기 평가의 과학창의성 표준편차가 교사 평가보다 컸는데, 타인에 대해서는 비교적 일정한 범위 내에서 평가하는 반면, 자기 자신에 대해서는 보다 큰 편차를 두고 평가함을 나타낸다. 이는 창의성에 대한 자기 평가가 자기효능감, 자아존중감 같은 개인의 내적 특성에 따라 쉽게 영향을 받으며, 편향될 가능성이 있다는 Sitzmann *et al.* (2010)의 지적과 상통하는 결과다.

### 3) 과학창의성 및 그 하위요소에 대한 자기 평가와 동료 평가의 상관관계

초등과학영재 학생의 과학창의성에 대한 학생 자기 평가, 동료 평가 결과 간 상관관계를 분석한 결과는 Table 3과 같다. 과학창의성과 그 하위요소 중 독창성 요소만이 유의한 상관관계를 나타냈다 ( $r = .42$ ). 이는 앞서 Table 2에서 독창성 요소의 자기 평가와 동료 평가 간 유의한 차이가 없었다는 결과와 연결된다. 또한 과학창의성 및 그 하위요소에 대한 자기 평가와 교사 평가 간 상관관계를 분석한 연구에서 독창성 요소에서만 상관성이 유의하게 나타나고( $r = .32$ ), 과학창의성과 유용성 요소에서는 나타나지 않았던 결과와 일치한다(김민주와 임채성, 2018).

이처럼 학생의 자기 평가와 동료 평가의 상관성이 유의미하게 나타나는 경우는 학생이 자기 평가가 동료 평가와 비교될 것이라는 사실을 인지하거나(Fox & Dinur, 1988), 자기 평가 시 평가 방법과 채점 기준을 명확하게 인식하고 있을 때(Ross *et al.*, 1999; Sung *et al.*, 2005)라고 한다. 반면, 자기 평가

**Table 3.** Correlation between student self-, and peer-assessment on scientific creativity and its two essential components (N=36)

변인	자기 평가		
	과학창의성	독창성	유용성
과학창의성	-.12		
동료 평가	독창성	.42*	
	유용성		-.12

\*  $p < 0.05$ .

가 학생의 성적에 반영될 경우, 과대평가할 가능성이 더 높아진다고 한다(Boud & Falchikov, 1989).

본 연구에서는 영재 학생이 자기 평가와 동료 평가가 비교될 것이라는 사실을 알지 못했다. 또한 학생의 자기 평가가 성적에 반영되지 않을 것이라는 점이 명확했다. 그럼에도 독창성 요소에 있어 자기 평가와 동료 평가가 관련성을 가진 것은 아이디어의 희소성을 뜻하는 독창성 요소가 비교적 객관적으로 평가될 수 있음을 시사한다. 이와 달리 유용성 요소는 그 자체로 객관적인 평가 방법과 채점 기준이 되지 못하기에 이에 대한 급간 채점 기준 설정이 필요하다.

4) 과학창의성에 대한 자기 평가와 동료 평가 간 빈도분포

자기 평가, 동료 평가에 대하여 각각의 중앙값을 기준으로 상·하위를 나누어 상상, 상하, 하상, 하하로 집단화하였다. 자기 평가의 상·하 수준을 나눈 중앙값은 100점 만점 중 72.98점이며, 동료 평가의 상·하 수준을 나눈 중앙값은 100점 만점 중 62.88점이다. 이를 통해 초등과학영재학생의 과학창의성에 대하여 자기 평가에 따른 동료 평가의 빈도분포 차이를 분석한 결과는 Table 4와 같다. 자기 평가 상·하 그룹에 따른 동료 평가 상·하 그룹의 빈도

분포는 통계적으로  $p < 0.05$  수준에서 유의한 차이가 있었다(카이제곱=4.0000, 자유도=1,  $p=0.0455$ ).

Table 4에서 4개의 집단 중, 자기 평가가 상위이고 동료 평가가 하위거나, 자기 평가가 하위이고 동료 평가가 상위인 집단이 가장 높은 빈도 분포를 보이고 있다(12명, 33.3%). 자기 평가와 동료 평가가 모두 상위 혹은 모두 하위인 집단이 그 다음 빈도 분포를 보이고 있다(6명, 16.7%). 그룹 분석에서 빈도 분포에 유의한 차이가 있었으므로 이를 토대로 그룹별 사례 분석을 더 진행하였다.

2. 초등과학영재학생의 과학창의성에 대한 자기 평가와 동료 평가 결과 수준에 따른 집단별 특징

1) 자기 평가와 동료 평가 결과 수준에 따른 그룹화

본 연구에서는 학생들이 평가하면서 가진 생각과 의견을 심층적으로 규명하기 위하여 과학창의성 평가 결과의 수준에 따라 네 집단으로 나누고, 사례를 통한 정성적 데이터로 집단별 특징을 분석하였다. 이를 통해 학생들이 가지고 있는 창의성에 대한 평가 기준을 파악하고 평가 주체 간 과학창의성 인식의 차이를 줄일 뿐만 아니라, 학생들의 창의적 메타인지능력을 향상시킬 교수·학습 방법에

Table 4. Frequency differences of scientific creativity for peer-assessment on self-assessment

		동료 평가		합계	카이제곱	p	
		하	상				
자기 평가	상	빈도	12	6	18	4.000*	0.046
		동료 평가(%)	66.7	33.3	100.0		
		자기 평가(%)	66.7	33.3	50.0		
		합계 %	33.3	16.7	50.0		
	하	빈도	6	12	18		
		동료 평가(%)	33.3	66.7	100.0		
		자기 평가(%)	33.3	66.7	50.0		
		합계(%)	16.7	33.3	50.0		
	합계	빈도	18	18	36		
		동료 평가(%)	50.0	50.0	100.0		
		자기 평가(%)	100.0	100.0	100.0		
		합계(%)	50.0	50.0	100.0		

\*  $p < 0.05$ .



대한 교육적 함의를 이끌어낼 수 있기 때문이다.

자기 평가와 동료 평가가 모두 상위 수준인 그룹을 HH, 모두 하위 수준인 그룹을 LL, 자기 평가는 상위지만 동료 평가는 하위인 그룹을 HL, 자기 평가는 하위지만 동료 평가는 상위인 그룹을 LH로 칭하였고, 각 집단에 대하여 정량적, 정성적 자료 분석을 진행하였다. 첫째로, 집단별 자기 평가와 동료 평가 점수의 평균은 Table 5와 같다. 이를 통해 어느 정도 점수대가 상위(High) 혹은 하위(Low) 수준을 이루고 있는지 확인할 수 있다. 각 집단별 평균 점수에서 HH, HL 집단은 자기 평가 점수가 동료 평가 점수보다 큰 차이로 높고, LH, LL 집단은 자기 평가 점수가 동료 평가 점수보다 약간 낮은 이유는 전체적으로 자기 평가 점수의 평균이 더 높으면서 표준편차가 컸기 때문이다.

2) 자기 평가 최상위 · 최하위 학생의 응답 분석

학생의 생각과 의견을 심층적으로 분석하기 위해 자기 평가와 동료 평가에서 각각 최상위 · 최하위 학생의 응답을 분석하였다. 자기 평가 최상위 학생 2명의 응답을 분석한 결과는 Table 6과 같다.

Table 5. The average scores of self- and peer-assessment for each group

		동료 평가	
		하	상
자기 평가	상	HL 자기: 81.4 동료: 56.4	HH 자기: 86.1 동료: 66.7
		LL 자기: 54.8 동료: 58.6	LH 자기: 62.7 동료: 66.6
	하	HL 자기: 81.4 동료: 56.4	HH 자기: 86.1 동료: 66.7
		LL 자기: 54.8 동료: 58.6	LH 자기: 62.7 동료: 66.6

학생 가는 자기 평가의 과학창의성 점수가 93.1점으로 표본 집단 학생 중 1위였으며, 학생 나 의 자기 평가 과학창의성 점수는 90.0점으로 2위였다. 두 학생 모두 HH집단에 속해 있다. 이 학생들의 경우, 자신의 아이디어에 대한 자신감이 발표하는 과정에서 동료 평가에 영향을 주었을 가능성이 있다. Bandura (1993)는 자신감이 타인의 평가에 영향을 받기도 하지만, 반대로 타인의 평가에 영향을 주기도 한다고 밝힌 바 있다.

자기 평가 최하위 학생 2명의 응답을 분석한 결과는 Table 7과 같다. 학생 다는 자기 평가의 과학창의성 점수가 32.0점으로 표본 집단 학생 중 최하 1위였으며, 학생 라의 자기 평가 과학창의성 점수는 46.4점으로 최하 2위였다. 두 학생 모두 LL집단에 속해 있다. 학생 가와 나 의 경우를 뒤집어 생각해 보면, 학생 다와 라는 자신의 아이디어에 대한 자신감 없음이 발표하는 과정에서 동료 평가에 영향을 주었을 가능성이 있다. 또한 이 학생들은 모두 독창성 점수가 현저하게 낮다는 특징이 있다.

3) 동료 평가 최상위 · 최하위 학생의 응답 분석

동료 평가 최상위 학생 2명의 응답을 분석한 결과는 Table 8과 같다. 학생 마는 동료 평가의 과학창의성 점수가 72.2점으로 표본 집단 학생 중 1위였으며, 학생 바의 동료 평가 과학창의성 점수는 70.5점으로 2위였다. 두 학생 모두 자기 평가는 낮은데, 동료 평가는 높은 LH집단에 속해 있다. 자기 평가가 낮은 것은 자기 효능감이 원인일 수 있다. 자신이 인식하는 과학창의성 점수는 자신에 대해 긍정적으로 생각하는 척도인 자아존중감, 창의적 자기 효능감과 깊은 연관이 있고, 이 척도가 높은 사람은 자신에 대해 긍정적으로 평가하며 자신감

Table 6. The responses of students with the highest scores in self-assessment and the reasons for their assessments

학생	응답				
가 (HH)	자기 평가 점수	독창성	9.5	평가 이유	파스칼의 원리를 이용한 수박커터기는 본 적이 없기 때문이다. 가정에서도 쉽고 안전하게 사용할 수 있기 때문이다.
		유용성	9.8		
나 (HH)	자기 평가 점수	독창성	10.0	평가 이유	이건 누구도 발명하지 못해 새롭지만, 병을 따는 용도라는 한정되어 있는 용도라서 유용성 점수는 좀 낮다.
		유용성	9.0		

**Table 7.** The responses of students with the lowest scores in self-assessment and the reasons for their assessments

학생	응답				
다 (LL)	공 던지는 수동 대포. 수동으로 피스톤을 누르면 대포 통 안에 있던 공이 날아감. 공을 잘 못 던지는 사람도 혼자 혹은 친구들과 공놀이할 수 있기 위해				
	자기 평가	독창성	4.0	평가 이유	자동 형은 벌써 나와 있어서
	점수	유용성	8.0		쓸 데가 있어서
라 (LL)	쓰레기 압축기. 유압프레스의 원리를 이용하여 쓰레기를 납작한 큐브 모양으로 만들어 부피를 줄여 보관				
	자기 평가	독창성	5.1	평가 이유	다른 친구들도 충분히 생각해 본 것일 것 같다.
	점수	유용성	9.1		요즘 환경문제가 논란이 되고 있고, 쓰레기 배출량도 그 문제들 중 하나다.

있게 행동할 뿐만 아니라, 창의적 과업을 달성하기 위해 끈기 있게 노력하는 적극적인 태도를 갖추고 있다고 한다(Cropley, 2001). 하지만 동료들이 높은 점수를 부여할 정도로 좋은 아이디어이고, 창의적 자기 효능감이 높은 학생이 자기 평가를 낮게 주었다면 영재 특유의 완벽주의 성향이 이에 영향을 줬을 가능성도 있다. LoCicero and Ashby (2000)는 영재 학생의 특징 중 하나인 완벽주의 성향이 두드러질수록 자신에 대해 보다 엄격한 기준을 적용함을 밝힌 바 있다.

이 학생들에 대한 동료 학생들의 평가 의견을 살펴보면, ‘가끔 엄마가 안 계실 때 소파나 책장 밑 청소를 하는 데 힘들었는데 청소를 할 때 편할 것 같다’의 의견처럼 학생들이 일상생활에서 필요함을 느꼈던 부분에 대해 공감하고 점수를 부여했다는 것을 확인할 수 있다.

동료 평가 최하위 학생 2명의 응답을 분석한 결과는 Table 9와 같다. 학생 사는 동료 평가의 과학 창의성 점수가 50.7점으로 표본 집단 학생 중 최하

1위였으며, 학생 아의 동료 평가 과학창의성 점수는 51.5점으로 최하 2위였다. 두 학생 모두 자기 평가는 높는데, 동료 평가는 낮은 HL집단에 속해 있다. 이 학생들에 대한 동료 학생들의 평가 의견을 살펴보면, ‘지하수가 아깝다’, ‘땅을 파야 하고 돈이 많이 든다’에서 볼 수 있는 것처럼 식수대와 같이 학생들의 일상생활에 친숙한 사례라도 효용 가치를 따져보고 평가하고 있음을 알 수 있다. 또한 학생 아의 상표 붙이는 도구는 공장이나 가게에서 쓰는 것으로 일상에서 접하지 못하는, 학생들에게 와 닿지 않는 사례로서 학생들에게 좋은 점수를 받지 못했음을 확인할 수 있다.

**4) 영재 학생들의 과학창의성 평가 활동에 대한 소감 분석**

다음은 면담 활동을 통해 수집한 영재 학생들의 활동 소감으로 동료 평가 활동에 대한 소감, 과학 창의성 평가 방식에 대한 의견으로 나누어 분석하고자 한다.

**Table 8.** The responses of students with the highest scores in peer-assessment and the reasons for their assessments

학생	응답				
마 (LH)	가구 밑 청소 도우미. 청소기로 잘 청소하지 못하는 가구 밑을 쉽게 청소하기 위해. 냉장고 같은 무거운 가구 밑에 유압 프레스를 끼우고 발로 밟으면 유체가 주입되어서 힘으로 냉장고를 들어 올림.				
	동료 평가	독창성	8.6	평가 이유	• 가끔 엄마가 안 계실 때 소파나 책장 밑 청소를 하는 데 힘들었는데 청소를 할 때 편할 것 같다. • 생활에 필요가 있는 도구라서
	평균	유용성	8.4		
바 (LH)	하수구 청소 로봇. 하수구 밑은 청소하기 힘들고 손으로 만져야 하므로 2관절 유압로봇 집게팔로 하수구 뚜껑을 올리거나 내릴 수 있게 함. 집게팔로 하수구 구멍 안에서 자유롭게 움직일 수 있고 쓰레기를 집어 올릴 수 있음.				
	동료 평가	독창성	8.6	평가 이유	• 노동을 줄여준다. • 실생활에 필요하다. • 쓰레기 치우는 것(위생)을 고려했다.
	평균	유용성	8.2		

**Table 9.** The responses of students with the lowest scores in peer-assessment and the reasons for their assessments

학생		응답	
유압 펌프. 적은 힘으로 물을 끌어올릴 수 있음. 놀이터에 이 시설을 설치하면 목이 마를 때 쉽게 지하수 마실 수 있음			
사 (HL)	동료 평가 평균	독창성	7.8
		유용성	6.5
		평가 이유	<ul style="list-style-type: none"> <li>창의적이지만 밑에 지하수가 꼭 있을 거란 보장이 없다.</li> <li>지하수가 아깝다./땅을 파야 하고 돈이 많이 든다.</li> <li>아리수 식 수도물은 너무 혼다.</li> </ul>
유압 상표. 공장이나 가게에서 자신들의 상표를 붙일 때 사용. 컨베이어 벨트로 상품들이 하나씩 이동해 오면 사람이 발판을 밟는 동시에 프레스가 내려와 상표 찍음 (전기 없이 편리하게).			
아 (HL)	동료 평가 평균	독창성	7.8
		유용성	6.6
		평가 이유	<ul style="list-style-type: none"> <li>내용이 구체적이지 않고 문제에 대한 해결방법이 부족하다.</li> <li>일상생활에는 필요 없다/수동이고 한정된 물건에만 사용할 수 있어 불편할 것 같다.</li> </ul>

<친구의 발표를 듣고 질문 및 평가하는 활동에 대한 학생의 생각>

- 학생 1: 좋은 점과 부족한 점을 알려주는 게 좋았고, 아이디어를 나누고 비교하면서 더 좋은 생각이 나왔다.
- 학생 2: 친구의 창의적 아이디어를 참고하여 내 공부에 도움이 되었고, 다른 친구의 생각을 듣는 것이 좋았다.
- 학생 3: 선생님보다 친구가 칭고해 주는 것이 더 편하고 좋았다.
- 학생 4: 서로 성장할 수 있어 좋았다.
- 학생 5: 친구가 보완해 주고 칭고해줘서 좋았다.

학생 1은 아이디어를 나누고 비교하면서 더 좋은 생각이 나왔다고 했는데, 이는 서로 다른 학생의 아이디어가 공유됨으로써 더 새롭고 높은 차원의 아이디어가 창발될 수 있다는 창발성의 개념(임채성, 1997)을 드러낸 것이다. 학생 3은 선생님보다 친구가 칭고해 주는 것이 더 편하고 좋았다고 답변함으로써 동료 평가 활동이 좋은 이유를 구체적으로 지적하였다.

- <과학창의성 공식에 따른 평가 방식에 대한 학생의 생각>
- 학생 6: 독창적이진 않는데 쓸모 있는 것이 있고, 독창적인데 유용하지 않은 것이 있다. 무엇이 창의적인지 파악할 수 있어서 좋았다.
  - 학생 7: 더 자세한 평가 기준이 필요하고, 독창성과 유용성에 대해 더 설명해 줬으면 좋겠다.
  - 학생 8: 점수보다는 단계별로 기준을 두고 분류했으면 좋겠다.
  - 학생 9: 전체적인 것을 합산하는, 창의성 전반에 대한 평가도 했으면 좋겠다.
  - 학생 10: 연령별로 생활 지능이 다르므로 어떤 연령대를

기준으로 한 것인지 정했으면 좋겠다.

학생 6은 과학창의성 공식에 따른 평가를 긍정적으로 보았고, 나머지 학생들은 이 평가의 보완할 점에 대해 여러 의견을 냈다. 학생 7은 더 자세한 평가 기준을 요구했는데, 연구자가 평가 전에 독창성과 유용성에 대해 설명해 주었지만, 보다 구체적 예시를 통한 자세한 설명을 요구하였다. 학생 8의 체점 기준에 따라 분류하지는 의견은 학생 7과 비슷한 의견이다. 학생 9는 점수를 직접 매기는 독창성, 유용성 항목에 더하여 창의성 자체에 대한 항목도 요구하였다.

학생 10의 의견은 어느 수준에 맞추어 창의성을 평가해야 하는지 정해야한다는 것으로, mini-c, little-c의 개념을 지적하고 있다. 과학창의성을 평가할 때 혼란스러운 부분은, 학생이 스스로 생각해 낸 아이디어지만 이미 밝혀지고 알려진 아이디어나 발명품이어서 독창성에 좋은 점수를 주기가 어렵다는 점에 있다. 그러나 학생이 이미 존재하는 아이디어를 기억하고, 그대로 사용하는 것이 아니라면, 사건, 경험, 현상에 대한 개인의 새롭고 의미 있는 해석인 mini-c 측면에서 창의성이라고 보는 것이 타당하다(Kaufman & Beghetto, 2009). 그렇기에 외부자인 교사보다는 비슷한 지식수준을 갖춘 또래 학생들의 평가가 과학창의성에 있어 더욱 유효하게 작용하는 것이다(Newton, 2010). 하지만 다른 학생은 이미 알고 있는 지식인데, 자신은 잘 몰랐던 지식을 스스로 발견하였다면, 이것이 개인 수준에서 창의적인지 가장 잘 평가할 수 있는 사람은 바로 학생 자신이다. 그러므로 학생 10의 의견을 반영하기 위해 본 연구에서 사용한 학생의 자기 평

가와 동료 평가를 보완·발전시켜 활용해 볼 수 있다. 그리고 이는 학생의 개인적이고 발달적인 측면을 중시하는 학습자 중심 창의성 교육을 가능하게 한다.

#### IV. 결론 및 제언

본 연구에서 초등과학영재학생의 과학창의성 및 그 하위요소에 대한 학생 자기 평가와 동료 평가 결과를 비교 분석한 결과는 다음과 같다.

첫째, 과학창의성 및 그 하위요소인 독창성, 유용성 점수의 평균값은 자기 평가가 동료 평가보다 높았다. 학생은 자기 자신을 평가할 때 관대한 평가 기준을 적용함을 알 수 있다. 표준편차 또한 과학창의성, 독창성, 유용성 요소에서 모두 자기 평가가 동료 평가보다 높았는데, 이것은 자기평가가 자기 효능감, 자기 존중감, 창의적 효능감 같은 개인의 내적 특성에 따라 쉽게 영향을 받으며, 편향될 가능성이 큼을 암시한다. 자기 평가는 mini-c 관점에 있어 중요하고, 학습 과정에 책임감을 가지고 몰입할 수 있도록 학습 동기를 부여하므로 이를 보완하여 교수·학습 활동에 활용하는 것이 필요하다. 그러므로 학생들이 보다 객관적으로 자신의 과학창의성을 파악하도록 구체적인 기준을 제시한다면 학생의 과학창의성에 대한 메타인지능력 또한 키워줄 수 있을 것이다.

둘째, 자기 평가와 동료 평가 간 상관관계 분석에서 독창성 요소만이 유의한 상관성을 가졌다. 이를 통해 독창성은 유용성에 비해 비교적 객관적인 기준이 될 수 있음을 알 수 있다. 무엇이 최소한지 객관적으로 판단할 수 있는 독창성 요소와 달리 유용성 요소에는 무엇이 유용하고 타당한지 판단하는 가치 판단의 요소가 내재하므로 평가자 간 구체적인 기준을 합의하지 않으면 서로 다른 결과가 나타날 수 있다. 그러므로 본 연구에서 사용한 과학창의성 평가 방식을 보완할 때, 특히 유용성 요소에 중점을 두고 채점자들이 합의할 수 있는 채점 기준표를 개발해야 한다.

셋째, 그룹별 사례 분석을 통해 학생들의 자기 아이디어에 대한 자신감이 발표하는 과정에서 동료 평가에 영향을 주었을 가능성, 자기 효능감 또는 완벽주의 성향이 학생의 자기 평가에 영향을 주었을 가능성을 확인하였다. 자기 효능감이나 자아

존중감이 낮은 학생은 스스로를 부정적으로 평가할 가능성이 크다. 반면, 자아존중감이 높더라도 스스로를 낮게 평가한 경우도 있는데, 이러한 학생들은 영재 특유의 완벽주의 성향으로 인해 자신에게 높은 기준을 적용했을 가능성이 있다. 자아존중감이나 자기 효능감은 타인의 평가에 영향을 받기도 하지만, 반대로 타인의 평가에 영향을 주기도 한다. 자신의 아이디어가 창의적이라는 자신감이 발표하는 과정에서 타인의 인식에 영향을 줄 수 있는 것이다. 그러므로 학생의 창의적 아이디어에 대해서 선부른 판단과 평가를 지양하고, 자기 효능감이 오히려 긍정적인 영향력을 발휘할 수 있도록 격려해야 한다.

넷째, 동료 평가에 대한 그룹별 사례 분석에서 학생들이 일상생활에서 필요를 느꼈던 부분이고, 이에 공감할수록 더 높은 점수를 부여했으며, 학생들에게 와 닿지 않고 일상생활과 동떨어진 사례는 더 낮은 점수를 부여한 것으로 나타났다. 그리고 학생들은 일상생활에 친숙한 사례라도 실현 가능성, 가격 대비 효율성 등 효용 가치를 따져보며 평가하였다. 각 평가에 대한 근거가 명확한 것을 보면 영재 학생들은 과학창의성을 평가할 능력을 어느 정도 갖추고 있다. 하지만 아이디어가 자신의 삶과 친숙한지에 따라 평가 결과가 달라지므로 과학창의성 과제를 내거나, 이를 학생이 평가하게 할 때, 누구를 위한 아이디어인가와 같이 활용의 초점을 명확히 하는 것이 필요하다.

다섯째, 평가 활동에 대한 학생들의 소감을 통해 동료 평가 방식에 대부분 긍정적이라는 것을 알 수 있었고, 과학창의성 평가에 있어서는 보완할 점에 대한 여러 의견을 모을 수 있었다. 학생들은 동료 평가에서 서로 보고 배울 수 있어 만족해 했으며, 선생님보다 친구가 충고해 주는 것이 더 편하고 좋았다고 했다. 과학창의성 평가의 경우, 세부적인 채점 기준과 추가적인 채점 항목을 요구하였는데, 이는 검토를 거쳐 추후 과학창의성 평가 연구에서 보완할 예정이다. 구체적인 채점 기준 확립도 중요하지만, 학습자가 평가 기준, 평가 요소를 정확하게 이해하도록 과학창의성의 두 요소인 독창성, 유용성에 대한 정의와 특징을 배우는 심화된 학습을 평가 전에 제공할 필요가 있다.

교사나 부모는 초등학생 수준이 아닌 외부적 입장에서 창의성을 평가할 가능성이 크다. 즉, 성인의

관점에서 과학자 수준 창의성인 빅-C 창의성에 근거하여 학생의 과학창의성을 평가할 가능성이 있는 것이다. 초등과학교육에서 학생의 과학창의성은 발달 중에 있다는 점, 창의적 자기 효능감이 타인의 평가에 영향을 받을 수 있는 시기라는 점을 인지하고, 그들의 창의성을 간과하거나 무시하는 일이 없도록 노력할 필요가 있다. 추후 연구에서 자기 평가와 동료 평가를 교사나 부모의 평가와 비교한다면 이러한 인식의 차이를 확인하고, 이 격차를 줄이는 초석을 마련할 수 있을 것이다.

자기 평가는 학생으로 하여금 자신의 창의성에 대해 반성적으로 사고하게 하고, 교육자는 이를 통해 창의적 자기 효능감 혹은 창의적 잠재력을 가진 학생을 식별할 수 있다. 동료 평가는 학생들이 또래 친구들의 다양한 아이디어를 접하면서 보고 배울 수 있고, 교사보다 친구들의 조언을 더 편하게 여긴다는 장점이 있다. 그리고 자기 평가와 동료 평가는 학습자의 발달 과정과 수준에 맞춰 평가할 수 있고, 교사가 쉽게 지나치거나 무시할 수 있는 학생의 과학창의성을 발견할 수 있다는 장점이 있다. 그러므로 교육 현장에서 학생의 자기 평가와 동료 평가를 활용하는 방안을 모색한다면 과학창의성 교육에 있어 학습자 중심의 교육을 실천하는 길을 다질 수 있을 것이다

## 참고문헌

- 김민주, 임채성(2018). 초등과학영재학생의 과학창의성에 대한 자기 평가, 교사 평가, 객관적 평가의 비교 분석. *초등과학교육*, 37(4), 440-454.
- 박병기, 유경순(2000). 창의성과 지능의 관계구조. *교육심리연구*, 14(2), 235-261.
- 임채성(1997). 협동학습의 대뇌생물학적 기초: 아이디어-공유 창출 모델. *생물교육 (구 생물교육학회지)*, 25(2), 143-155.
- 임채성(2012). 뇌기반 진화적 접근법에 따른 창의적 과학 문제해결 지도 모형 개발. *생물교육*, 40(4), 429-452.
- 임채성(2014). 과학창의성 평가 공식의 개발과 적용. *초등과학교육*, 33(2), 242-257.
- 최인수(2000). 유아용 창의성 측정도구에 관한 고찰. *유아교육연구*, 20(2), 139-166.
- Bandura, A. (1993). Perceived self-efficacy in cognitive development and functioning. *Educational Psychologist*, 28(2), 117-148.
- Barbera, E. (2009). Mutual feedback in e portfolio assessment: an approach to the netfolio system. *British Journal of Educational Technology*, 40(2), 342-357.
- Barrett, P. (2007). Structural equation modelling: Adjudging model fit. *Personality and Individual Differences*, 42(5), 815-824.
- Beghetto, R. A. & Kaufman, J. C. (2007). Toward a broader conception of creativity: A case for "mini-c" creativity. *Psychology of Aesthetics, Creativity, and the Arts*, 1(2), 73.
- Besançon, M., Fenouillet, F. & Shankland, R. (2015). Influence of school environment on adolescents' creative potential, motivation and well-being. *Learning and Individual Differences*, 43, 178-184.
- Boud, D. & Falchikov, N. (1989). Quantitative studies of student self-assessment in higher education: A critical analysis of findings. *Higher Education*, 18, 529-549.
- Callahan, C. M. & Miller, E. M. (2005). A child-responsive model of giftedness. *Conceptions of Giftedness*, 2, 38-51.
- Colman, A. M. (2003). *A dictionary of psychology* (2 ed.): Oxford University Press.
- Cropley, A. J. (2001). *Creativity in education & learning: A guide for teachers and educators*. NY: Psychology Press.
- Cropley, A. J. (2001). *Creativity in education and learning: A guide for teachers and educators*. London: Kogan Page.
- Fox, S. & Dinur, Y. (1988). Validity of self-assessment: A field evaluation. *Personnel Psychology*, 41, 581-592.
- Hickey, M. (2001). An application of Amabile's consensual assessment technique for rating the creativity of children's musical compositions. *Journal of Research in Music Education*, 49(3), 234-244.
- Hu, W. & Adey, P. (2002). A scientific creativity test for secondary school students. *International Journal of Science Education*, 24(4), 389-403.
- Kaufman, J. C. & Beghetto, R. A. (2009). Beyond big and little: The four c model of creativity. *Review of General Psychology*, 13(1), 1-12.
- Knowles, M. S., Holton, E. F., III. & Swanson, R. A. (2005). *The adult learner* (6th Ed.). Burlington, MA: Elsevier.
- Locicero, K. A. & Ashby, J. S. (2000). Multidimensional perfectionism in middle school age gifted students: A comparison to peers from the general cohort. *Roeper Review*, 22(3), 182-185.
- Mumford, M. D. (2003). Where have we been, where are we going? Taking stock in creativity research. *Creat-*

- ivity Research Journal*, 15(2-3), 107-120.
- Newton, D. P. (2010). Assessing the creativity of scientific explanations in elementary science: An insider-outsider view of intuitive assessment in the hypothesis space. *Research in Science & Technological Education*, 28(3), 187-201.
- Patton, M. Q. (2002). *Qualitative research and evaluation methods*. Thousand Oaks, CA: Sage.
- Renzulli, J. S. (2003). Conception of giftedness and its relationship to the development of social capital. *Handbook of Gifted Education*, 3, 75-87.
- Rosenthal, R. & Rosnow, R. L. (1991). *Essentials of behavioral research: Methods and data analysis*. New York: McGraw-Hill.
- Ross, J. A., Rolheiser, C. & Hogaboam-Gray, A. (1999). Effect of self-evaluation on narrative writing. *Assessing Writing*, 6(1), 107-132.
- Ross, J. A. (2006). The reliability, validity, and utility of self-assessment. *Practical Assessment, Research, and Evaluation*, 11(10), 1-13.
- Runco, M. A. (2004). Everyone has creative potential. In Sternberg, R. J., Grigorenko, E. L. & Singer, J. L. (Eds.), *Creativity: From potential to realization*. (pp. 21-30). Washington, DC: American Psychological Association.
- Runco, M. A. (2007b). *Creativity: Theories and themes: Research, development, and practice*. San Diego, CA: Academic Press.
- Runco, M. A. & Jaeger, G. J. (2012). The standard definition of creativity. *Creativity Research Journal*, 24(1), 92-96.
- Sitzmann, T., Ely, K., Brown, K. G. & Bauer, K. N. (2010). Self-assessment of knowledge: A cognitive learning or affective measure?. *Academy of Management Learning & Education*, 9(2), 169-191.
- Sowden, P. T. & Dawson, L. (2011, November). Creative feelings: The effect of mood on creative ideation and evaluation. In *Proceedings of the 8th ACM Conference on Creativity and Cognition* (pp. 393-394). ACM.
- Sung, Y.-T., Chang, K.-E., Chiou, S.-K. & Hou, H.-T. (2005). The design and application of a web-based self- and peer-assessment system. *Computers and Education*, 45(2), 187-202.
- Treffinger, D. J. (2009). Myth 5: Creativity is too difficult to measure. *Gifted Child Quarterly*, 53(4), 245-247.
- Weisberg, R. W. (1993). *Creativity: Beyond the myth of genius*. WH Freeman New York.

---

김민주, 서울문현초등학교 교사(Kim, Minju; Teacher, Seoul Munhyun Elementary School).

† 임채성, 서울교육대학교 교수(Lim, Chaeseong; Professor, Seoul National University of Education).