

## 초등과학영재학생의 과학창의성에 대한 자기 평가, 교사 평가, 객관적 평가의 비교 분석

김민주 · 임채성<sup>†</sup>

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

Kim, Min-Ju · Lim, Chae-Seong<sup>†</sup>

#### ABSTRACT

This study aims to compare student self-, teacher-, and objective assessments of elementary science-gifted students' scientific creativity. A science-gifted program on the topic of Hydraulic Machine was implemented to 40 fifth-graders in the Science-Gifted Education Center of an education office in Seoul, Korea for four weeks. The products of the students' activities were assessed by three types of 'Student Self-Assessment', 'Teacher-Assessment', and 'Objective Assessment using Formula'. Based on two essential components of creativity, the scientific creativity is divided into two parts of originality and usefulness. Ideas that satisfy both components can be counted as scientifically creative. The main results of this study are as follows: First, the scores of each week and the average of the overall four-week scores on scientific creativity were significantly correlated. Student self-assessment ( $r=.687$ ), teacher-assessment ( $r=.715$ ), and objective assessment ( $r=.724$ ) appeared consistently over instructional periods. Second, the average scores of student self-, objective, and teacher-assessments were 73.15, 35.72, and 26.60, respectively. The result of student self-assessment on scientific creativity tended to be higher than those of formula and teacher. Third, among the three types of assessment on scientific creativity, a strong correlation appeared between teacher- and objective assessment ( $r=.974$ ), but neither between student self- and objective ( $r=.161$ ) nor between student self- and teacher- ( $r=.213$ ). Fourth, the scores on originality component had a positive correlation between teacher- and objective assessment ( $r=.713$ ). The scores of student self- and teacher-assessments had a significant correlation too ( $r=.315$ ), but not between student self- and objective assessment ( $r=.279$ ). Fifth, the scores on usefulness component did not have a significant correlation between student self- and teacher-assessment ( $r=.155$ ). Sixth, there was no significant difference on scientific creativity between student self- and objective assessment [ $\chi^2(1, n=40)=1.667, p<.197$ ]. Not between student self- and teacher-assessment either [ $\chi^2(1, n=40)=1.616, p<.204$ ]. On the contrary, there was a significant difference between teacher- and objective assessment [ $\chi^2(1, n=40)=32.593, p<.000$ ]. Seventh, the students were categorized into four groups according to the levels of their scores by student self- and teacher-assessment. The result showed that factors influencing student self-assessment are inherent in the personality traits of gifted individuals, such as self-esteem and perfectionism. The findings suggested that there are challenges for the educators to make efforts to construct consistent assessment methods for scientific creativity.

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

## I. 서 론

21세기는 지식정보사회로서 급격하게 변화하고 있으며, 앞으로도 지속될 이러한 변화에 능동적으로 대처하기 위해 창의성 개발은 세계적 관심사이자 국가적 시책이 되었다. 이렇듯 현대 교육의 중심점이라고 할 수 있는 창의성은 오늘날 학교 교육과정에서 주요 목표로 제시되고 있으며(Newton, 2010; Walker & Gleaves, 2008), 우리나라 과학 교과 또한 교육과정(교육부, 2015)에서 과학적이고 창의적으로 문제를 해결하기 위한 과학적 소양 함양에 핵심 목표를 두고 있다. 이러한 시책의 일환으로 2004년에 교육인적자원부는 창의적 인재 양성을 위한 「수월성 교육 중합대책」을 발표하였고, 2005년부터 본격적으로 창의적 인재를 양성하고자 영재교육을 위한 예산을 편성하였다.

창의성은 매우 복잡한 개념으로 이를 정의하는데 많은 요소들이 관여하고 있으며(Weisberg, 1993), 창의성으로 간주되는 특징 또한 300가지 이상으로 다양하다(Treffinger, 2009). 그러나 대부분의 연구는 공통적으로 독창성(참신성)과 유용성(적절성)을 창의성의 요소로 포함하고 있다(Sternberg, 1998). 그래서 최근 창의성에 대한 정의는 새로우면서도 유용한 산물을 만들어내는 성향이나 능력(Mayer, 1999; Mumford, 2003; Sternberg, 1998)으로 합의가 이루어지고 있다. 영역 특수적인 지식과 기능들이 창의성 발휘에 중요한 영향을 미치고(Hu & Adey, 2002), 학교 현장에서는 교과를 중심으로 수업이 이루어지는 경우가 많다. 이러한 영역 특수적인 접근 방식으로 볼 때, 창의성을 정의하는 두 가지 핵심 속성에 창의성이 발현되는 영역을 반영하면 과학창의성을 ‘과학적으로 새로우면서도 유용한 산물을 만들어내는 능력’이라고 할 수 있다.

Renzulli (2003)는 평균 이상의 지능, 과제 집착력, 창의성의 세 가지 요소를 영재의 기본 조건으로 정의하고 있다. 신지은(2002)에 의하면 과학 영재 학생은 과학 분야에서 창의성을 발휘할 잠재력을 갖춘 학생이라고 볼 수 있으며, 이들 집단에서 과학적이고 창의적인 아이디어가 다양하게 생성될 가능성이 높다(Callahan & Miller, 2005). 그들로부터 산출된 창의적 아이디어의 다양한 사례를 분석한다면 영재 학생을 포함하여 일반 학생의 과학창의성을 효과적으로 발휘하도록 하는 평가 방법에

대한 시사점을 이끌어낼 수 있을 것이다.

지금까지 과학창의성을 평가하기 위한 다양한 방법들이 시도되었는데(Treffinger, 2002), Bouzidi and Jaillet (2009)에 따르면 다면적 평가 방식이 평가의 신뢰도와 타당도 향상에 기여한다고 하였다. 현재 교육현장의 평가는 대부분 교사의 관점에서 이루어지고 있지만 실제적인 평가를 위해서는 학생의 실질적인 학습 과정과 사고를 반영하도록 학생의 관점을 고려할 필요가 있다(Barbera, 2009; Barrett, 2007; Plake & Impara, 1996). 임채성(2012)은 창의성을 발휘하는 주체도 고려하여 학생 과학창의성(student's scientific creativity)을 학생 수준에서 과학적으로 새로우면서 유용한 것을 만들어내는 능력으로(little-c 창의성), 과학자 과학창의성(scientist's scientific creativity)을 과학계에서 새로우면서 유용한 것을 만들어내는 능력으로(Big-C 창의성) 정의하였다. 본 연구에서는 초등학생 수준에서의 과학 창의성 평가 방법을 연구하므로 Big-C 창의성과 구분되는 little-c 창의성의 관점으로 접근하였다.

과학창의성은 발산 및 수렴적 사고를 포함한 고차원적 사고 과정을 통해 나타나므로(yager, 2000) 학생의 인식과 사고 과정을 정확하게 알 필요가 있지만, 그들이 스스로 인식하는 과학창의성을 규명한 연구가 부족한 실정이다(Chan, 2005). Little-c 창의성 접근 방식은 학생이 자신의 과학창의성을 어떻게 인식하는지와 깊은 관련이 있고, 자기 평가와 교사 평가의 비교에 따른 시사점에 방향을 제시해 준다. 많은 연구에서 자기 평가와 교사 평가 결과가 일관적이라고 밝힌 반면(Bouzidi & Jaillet, 2009; Cho *et al.*, 2006; Liu, 2002; Sadler & Good, 2006; Sung *et al.*, 2005; Tsai & Liang, 2009; Tseng & Tsai, 2007), 자기 평가와 교사 평가 결과가 일관적이지 않음을 발견한 연구도 있다(Knowles *et al.*, 2005). 이처럼 창의성 평가 결과의 일관성이 일치되지 않게 나타난 데에는 평가받는 학생의 수준, 채점 기준, 평가 방식 등 다양한 요인이 영향을 미쳤을 가능성이 있다(Chang *et al.*, 2012). 과학창의성에 대한 자기 평가, 교사 평가, 객관적 평가 결과의 상관성은 과학창의성 평가의 주체와 그 비중에 대한 의사결정에 영향을 줄 것이다.

따라서 본 연구에서는 효과적인 과학창의성 평가 방법을 마련하고자 교육주체인 교사와 학생으로 평가의 관점을 나누고, 그 평가 결과가 어떻게

달라지는지를 과학창의성 평가 공식(임채성, 2014)을 통해 산출한 자료와 함께 비교하였다. 또한 현재 영재 교육을 받는 학생을 자기 평가, 교사 평가 결과 수준에 따라 어떠한 특징을 갖고 있는지를 분석하였다.

## II. 연구 방법 및 절차

### 1. 연구 설계

본 연구는 초등과학영재학생을 대상으로 유압 기계에 대한 과학영재 프로그램을 4주 동안 실시한 후 학생 결과물을 통해 과학창의성을 평가한 정량적 데이터로 학생의 자기 평가, 교사 평가, 과학창의성 평가 공식을 이용한 객관적 평가의 세 가지 평가 결과 간 관련성을 분석하였다. 정량적 데이터로 나타난 현상의 원인을 알아보기 위해 관찰의 정성적 데이터도 함께 수집하여 분석하였다. 모든 연구 과정은 초등 과학 교육 및 초등 과학 영재 교육 석·박사 전공교사 및 전공교수의 자문과 협의 과정을 거쳐 진행되었다.

### 2. 연구 대상 및 적용 주제

본 연구에 참여한 학생들은 서울특별시 K교육지원청에 소속된 각자의 학교에서 담임교사 및 학교 추천위원회의 추천을 거친 후, GED(Gifted Education Database) 기관심사 1단계인 창의적 문제해결력 평가와 2단계인 면접 평가 그리고 각 학교에서 수집된 학생 정보를 종합적으로 고려하여 과학 영재로 선발된 학생들이다. 이처럼 여러 단계를 거치고 교사 관찰, 자기 평가 체크리스트 등 다양한 정보를 수집하는 영재 선발 과정이 학생의 다각적인 측면을 고려한다고 판단하여 교육지원청 소속 영재 학생을 연구 대상으로 선정하였다. 또한 과학을 학습하는 초등 3~6학년 학생 중에서 과학 탐구 능력과 구체적 조작능력이 일정 수준 이상 갖추어져 있고, 관찰이나 면담이 비교적 용이한 5학년 학생을 연구 대상으로 선정하였다.

연구 대상의 남녀 구성은 남학생 34명과 여학생 6명이고, 20명씩 2개 학급으로 이루어져 있다. 40명의 연구 대상자 중 수업에 참석하지 않아 결과물이 누락되거나 자신이 산출한 결과물에 대한 자기 평가를 하지 않은 학생이 있어 자료 집계 방식에 따

라 표본 수가 29명, 34명 혹은 40명으로 달라진다.

과학창의성의 연구 대상으로 초등과학영재학생을 선정한 까닭은 첫째, 과학영재학생이 자신의 과학창의성을 어떻게 인식하고 있는지를 검토하여 일반 교육과정에서 과학창의성 함양을 위한 목표와 평가 체계를 마련하기 위한 시사점을 얻기 위해서이다. 둘째, 과학영재학생으로부터 창의적 아이디어의 사례를 최대한 다양하게 수집할 수 있기 때문이다. 셋째, 아이디어가 다양한 만큼 그들의 과학창의성 점수 또한 다양하게 분포할 것이라고 추론했기 때문이다.

연구에 적용한 주제는 학생의 과학창의성 발현 가능성과 활동 관찰의 용이성을 고려하여 선정하였다. 이에 따라 학생 중심의 다양한 조작 활동과 모둠 활동으로 수업을 구성할 수 있고, 활동한 것을 바탕으로 과학창의성을 평가할 수 있는 결과물을 도출할 수 있는 주제인 ‘유압 기계 장치’를 선정하였다.

### 3. 자료 수집 및 절차

본 연구는 신뢰도와 타당도를 확보하기 위해 다양한 방법으로 자료를 수집하였다. 먼저 학생의 과학창의성을 평가할 수 있는 교수학습 프로그램 및 과제를 개발하였다. 자기 평가는 학생이 자신의 수업 활동 결과물에 대해 독창성과 유용성 측면에서 각각 100점 만점으로 평정하는 방식으로 이루어졌다. 교사 평가는 수업 후 과학영재 전공 교사 4인이 학생결과물에 대해 학생 자기 평가에서와 마찬가지로 창의성의 두 가지 요소 각각에 대해 100점 만점으로 평정하는 방식으로 이루어졌다. 자기 평가, 교사 평가 모두 독창성과 유용성 각각에 대해 직관적으로 점수를 매기도록 하였으며, 급간에 대한 평가기준을 정하지는 않았다. 객관적 평가는 임채성(2014)의 과학창의성 평가 공식에 따라 창의성의 두 가지 요소별로 점수를 산출하는 방식으로 이루어졌다. 이 평가 공식에서, 독창성 요소는 그 정의대로 산물의 희귀성 정도가 객관적으로 산출되지만, 유용성 요소는 불가피하게 산물의 과학성에 대한 평정자의 주관에 따라 채점된다. 따라서 객관적 평가에서 유용성 요소에 대한 점수는 교사 평가 결과로 산출된 점수를 수집하였다. 이러한 정량적 자료 외에도 수업 참여 관찰 기록과 같은 정성적 자료를 수집하고 정량적 자료와 연계하여 분석하였다.

### 1) 과학창의성 평가 과제

본 연구에 참여한 학생들은 2개월간 4회의 영재 수업 프로그램에 참여하여 과학창의성을 요구하는 과제를 해결하였다. 이 기간 동안 매 수업에서 학생은 파스칼의 원리에 대해 학습하고, 이를 이용한 장치를 직접 조립한 다음 그 장치를 실생활에 활용할 아이디어를 구상하는 과학창의성 과제를 수행하였다. 프로그램은 유압 기계라는 공통된 주제 하에 유압 가제트 집게 팔, 유압 프레스, 유압 리프트, 유압 2관절 로봇 팔과 같이 생김새와 기능이 다양한 장치를 만드는 것으로 구성되었다.

각 장치를 실생활에서 활용할 아이디어는 1인당 1개 이상 제시할 수 있고, 글이나 그림으로 표현하게 하였다. 동시에 자신이 생각한 아이디어에 대한 독창성과 유용성 점수를 매기도록 하였는데, 과제를 제시할 때 창의성을 평가하는 과제라는 암시를 주지 않도록 창의성, 독창성, 유용성과 같은 용어는 사용하지 않았다. 또한, little-c 창의성 측면에서 독창성과 유용성은 각각 자신에게 ‘새롭다고 생각하는 정도’, ‘유용하거나 타당하다고 생각하는 정도’와 같이 초등학생 수준에서 이해하기 쉬운 말로 대체하였다.

### 2) 과학창의성 평가

Treffinger (2002)에 따르면 창의성 평가 방법은 100여 가지로 다양하다. 본 연구에서는 창의성의 두 가지 핵심 요소인 독창성(참신성)과 유용성(적절성, 타당성, 과학성) 점수의 곱을 과학창의성 점수로 산출하였다. 합이 아닌 곱으로 계산되는 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(독창성, 참신성), 유용한(적절성) 아이디어나 산물을 만들어내는 능력이라고 정의하고, 이 정의에 따라 다음과 같은 과학창의성 평가 공식을 개발하였다.

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

이 공식에 의하면, 과학창의성 점수는 독창성과

유용성 요소 각각에 대해 10점 만점으로 평정한 후, 두 점수를 곱하여 100점 만점으로 산출된다. 위 식을 기본으로 하여 독창성과 유용성 점수를 매기는 주체에 따라 학생의 자기 평가, 교사 평가, 객관적 평가로 평가의 관점을 나누었다. 교사와 학생이 독창성과 유용성 각각에 대해 100점 만점으로 평가한 자료를 수집하였고, 이를 10점 만점으로 환산하여 각 점수는 0~10점 사이 소수점 첫째 자리까지로 책정되었다. 교사, 학생, 객관적 평가 모두 독창성 점수와 유용성 점수를 곱한 것을 과학창의성 점수로 산출하였다.

학생의 자기 평가는 창의적 과제를 수행한 후 자신의 아이디어에 대해 독창성과 유용성 점수를 각각 매기도록 하였다. 이때, 독창성과 유용성에 대해 쉬운 말로 풀이하여 평가 요소에 대한 이해를 도왔다. 교사 평가는 수업 후 학생의 각 산출물에 대하여 교사 4인이 독창성과 유용성을 각각 평가한 점수를 평균하였다. 채점에 참여한 교사는 모두 교육지원청 소속 초등과학영재원에서 강의한 이력이 있으며, S교육대학교 교육전문대학원에서 초등영재교육을 전공하였다. 채점에 참여한 교사들은 독창성과 유용성의 의미를 숙지하고 평가에 임하였다. 채점자 4인의 신뢰도인 Cohen's Kappa 계수는 .86이었다. 마지막으로 객관적 평가는 위의 공식에 따라 이루어졌다.

위 식에서 SC는 과학창의성(scientific creativity)을, 왼쪽의  $\left[1 - \frac{(n-1)}{(N-1)}\right]$ 항은 아이디어의 고유한 정도, 즉 독창성을 의미하고, 오른쪽의 U는 아이디어의 유용한 정도, 즉 과학적 타당성을 의미한다. N은 집단의 전체 구성원 수이고, n은 그 집단에서 특정 아이디어를 제시한 학생의 수이다. 왼쪽 항은 0~10의 값을 가지며, 10에 가까울수록 그 아이디어가 독창적임을 나타낸다. 오른쪽의 U 역시 0~10의 값을 가지며, 10에 가까울수록 그 아이디어나 산물이 과학적으로 적절하거나 타당함을 뜻한다.

유용성 항은 독창성 항과 달리 평가자의 전문적인 판단에 의해 채점된다. 본 연구에서는 초등교육의 측면에서 교사를 평가자로 상정하였고, 교사 평가의 유용성 점수를 활용하였다. 즉, 초등과학영재를 지도하는, 관련 전공 교사 4인의 유용성 점수를 평균한 값이 U항이 된다. 신뢰도 확보를 위해 다수에 대한 평균값을 사용하였지만 유용성 점수에 있어 평가자의 개인적 주관에 개입할 가능성을 완전히

줄이지는 못하는 것이 객관적 평가의 제한점이다.

### 3) 관찰

Patton (2005)은 어느 하나의 방법만으로는 복잡한 현상을 해석하기 어렵기 때문에 참여관찰과 심층면담을 병행하여 사용할 필요가 있다고 하였다. 본 연구에서는 연구자가 수업을 하며 관찰하는 참여 관찰(Musante & DeWalt, 2010)의 형태로 학생의 발표 횟수, 창의적 발표 횟수, 개별 행동 및 모둠별 활동에서 나타내는 행동을 관찰하고 기록하는 방식을 사용하였다. 연구자는 1개 학급을 대상으로 수업을 진행하면서 영재 학생 20명을 관찰하였지만, 학생의 과학창의성에 대한 평가 과정과 별개로 관찰 결과를 기록하였기 때문에 과학창의성 평가 결과의 영향을 받지 않은 관찰을 했다고 할 수 있다.

### 4. 자료 분석 방법

자기 평가, 교사 평가, 객관적 평가 결과는 과학창의성 및 그 하위요소인 독창성과 유용성에 대하여 평균과 표준편차를 비교 분석하였으며, 시간에 따른 일관성 및 상관관계와 빈도분포에 대해서는 I-STATistics 프로그램으로 분석하였다. 또한 자기 평가와 교사 평가 결과 받은 점수 각각에 대하여 중앙값을 기준으로 수준을 나누어 상상, 상하, 하상, 하하의 네 집단으로 구성하고, 수집된 정량적 데이터와 정성적 데이터에서 나타나는 각 집단별 특성

을 분석하였다. 이때 사용된 정량적 데이터는 과학창의성의 하위요소인 독창성과 유용성 점수이며, 정성적 데이터는 학생의 수업 행동 관찰 내용이다.

## III. 연구 결과 및 논의

### 1. 초등과학영재학생의 과학창의성 및 그 하위 요소에 대한 학생 자기 평가, 교사 평가, 객관적 평가의 비교

#### 1) 세 평가 방식에 대한 평가의 일관성 분석

초등과학영재학생의 과학창의성에 대한 학생 자기 평가, 교사 평가, 객관적 평가가 2개월 동안 4회에 걸쳐 일관성 있게 측정되었는지를 각각 분석하였다. 독창성 점수와 유용성 점수를 곱하여 산출한 과학창의성, 독창성, 유용성에 따라 각 평가 방식의 주차별 점수와 전체 평균값 간 상관관계를 분석한 결과는 Table 1과 같다. 이 결과를 통해 각각의 평가 방식이 한 학생에 대한 과학창의성 그 하위요소인 독창성, 유용성을 4회에 걸쳐 일관성 있게 측정하였음을 알 수 있다. 그러므로 세 평가 방식은 평가의 두 요건인 타당성과 신뢰성 중 시간이 흐름에도 안정적으로 측정하는지를 의미하는 신뢰성을 만족시킨다고 할 수 있다(Carmines & Zeller, 1979; Salvia *et al.*, 2012).

**Table 1.** Correlation between weekly and overall average scores of three types of assessment on scientific creativity, originality, and usefulness N=34

변인		과학창의성			
		1주차	2주차	3주차	4주차
자기 평가	과학창의성	0.65***	0.69***	0.74***	0.67***
	독창성	0.72***	0.48**	0.79***	0.73***
	유용성	0.66***	0.84***	0.69***	0.60***
교사 평가	과학창의성	0.77***	0.75***	0.72***	0.63***
	독창성	0.73***	0.59***	0.65***	0.67***
	유용성	0.71***	0.74***	0.65***	0.58***
객관적 평가	과학창의성	0.72***	0.71***	0.79***	0.67***
	독창성	0.45**	0.39*	0.73***	0.60***
	유용성	0.71***	0.74***	0.65***	0.58***

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ .

이에 따라 앞으로 논의할 연구 결과에서 과학창의성을 포함한 유용성, 독창성 점수는 4회에 걸쳐 평균한 값을 사용하기로 하였다. 이는 평가 결과가 일정 기간에 걸쳐 일관성 있게 나타났을 뿐만 아니라, 가변적이고 우연의 영향을 받는 1회의 평가보다 4회의 평가를 평균한 값이 과학창의성에 대한 서로 다른 평가 방식을 비교하기에 더 용이하다고 보았기 때문이다.

2) 과학창의성 및 그 하위요소에 대한 세 가지 평가 결과 비교

초등과학영재 학생의 과학창의성에 대한 학생 자기 평가, 교사 평가, 객관적 평가의 평균 및 표준편차는 Table 2와 같다. 100점 만점으로 측정되는 과학창의성의 평균값은 자기 평가 73.2점, 객관적 평가 35.7점, 교사 평가 26.6점이었다. 여기에서 자기 평가와 객관적 평가 점수의 차이는 37.5점, 객관적 평가와 교사 평가 점수의 차이는 9.1점이므로 학생의 자기 평가가 다른 평가에 비해 높은 경향을 보였다. 반면, 교사 평가는 객관적 평가보다도 더 낮았다. 이는 자기 평가 점수가 교사 평가 점수보다 높은 점을 들어 교사는 엄격한 평가 기준을 갖고 있는 반면, 학생은 자신에 대해 관대한 평가 기준을 적용한다고 밝힌 선행 연구와 일치하는 결과이다(Ross, 2006; Boud & Falchikov, 1989). 연구에 따르면 어린 학생일수록 자신의 능력에 대한 정보를 통합하는 메타 인지 사고 기술이 부족하고, 단지 자신이 희망하는 결과에 대한 생각에 기초해서 평가할 수 있기 때문에 과대평가할 가능성이 높다고 한다.

과학창의성의 두 하위 요소인 독창성과 유용성 점수에 대해서도 세 평가 관점의 평균값을 비교 분석하였다. 10점 만점으로 측정되는 독창성 점수의 평균값은 자기 평가 8.3점, 객관적 평가 7.5점, 교사 평가 5.3점이었다. 여기에서 자기 평가와 객관적 평

가 점수의 차이가 0.8점, 객관적 평가와 교사 평가 점수의 차이가 2.2점으로 공식에 의해 측정되는 독창성 점수와 학생이 인식한 독창성이 비슷하였다. 교사의 독창성 평가는 과학창의성 평가와 마찬가지로 낮게 나왔다.

10점 만점으로 측정되는 유용성 점수의 경우, 과학창의성 평가 공식의 정의에 따라 객관적 평가에서 유용성 점수는 교사를 포함한 전문가가 평가한 결과를 사용하므로 자기 평가와 교사 평가만을 비교하였다. 그 결과, 자기 평가 8.8점, 교사 평가 4.8점으로 자기 평가가 교사 평가보다 4점 높았다. 이는 독창성 점수에서 차이가 얼마 나지 않았던 자기 평가와 객관적 평가 점수의 차가 많이 나는 유용성 점수로 인하여 독창성과 유용성 점수의 곱으로 산출되는 과학창의성 점수 차가 벌어졌다는 것을 유추할 수 있다.

과학창의성에 대한 표준편차의 경우, 자기 평가 15.47점, 객관적 평가 15.45점, 교사 평가 11.67점이었다. 독창성 점수의 표준편차 또한 자기 평가 1.09점, 객관적 평가 0.89점, 교사 평가 0.76점이었다. 여기에서 교사 평가의 표준편차가 다른 두 평가와 비교했을 때 상대적으로 낮다는 것은 교사가 매긴 학생 간 과학창의성 점수 차가 크지 않음을 보여준다. 반대로 유용성 점수의 표준편차는 자기 평가 1.26점, 교사 평가 1.53점으로 교사 평가가 자기 평가보다 0.27점 높았다. 세 평가 방식에서 유용성의 점수 분포가 독창성보다 다양하게 나타났는데, 특히 교사 평가에서 편차가 컸다. 이를 통해 가치 판단의 문제가 개입되는 유용성은 교사가 평가할 때 점수 차이를 두드러지게 하는 요인임을 알 수 있다.

각 평가 방법에서 독창성과 유용성 점수 간 차이를 비교 분석하였다. 자기 평가에서는 유용성 점수가 독창성 점수보다 0.5점 더 높았다. 반대로 객관적 평가와 교사 평가에서는 독창성점수가 유용성 점수보다 각각 2.7점, 0.5점 더 높았다. 객관적 평가

Table 2. The average scores and standard deviations of three types of assessment on scientific creativity and its two essential components N=40

변인	자기 평가	객관적 평가	교사 평가	
과학창의성	73.2±15.47	35.7±15.45	26.6±11.67	
과학창의성 하위요소	독창성	8.3±1.09	7.5±0.89	5.3±0.76
	유용성	8.8±1.26		4.8±1.53

평균±표준편차.

와 교사 평가에 사용된 유용성 점수는 모두 교사에 의해 평가된 점수이다. 객관적 평가에 의한 독창성 점수와 교사의 유용성 점수의 차이가 큰 반면, 교사가 평가한 독창성 점수와 유용성 점수의 차이는 작다. 또한 학생 자신이 평가한 독창성 점수와 유용성 점수의 차이도 작은 편에 속해 있다. 이를 통해 평가 주체는 독창성 평가와 유용성 평가에 대해 비슷한 점수대를 부여하는 것을 알 수 있다.

**3) 과학창의성 및 그 하위요소에 대한 세 평가 방식의 상관관계**

초등과학영재 학생의 과학창의성에 대한 학생 자기 평가, 교사 평가, 객관적 평가 결과 간 상관관계를 분석한 결과는 Table 3과 같다. 과학창의성에 대한 세 평가 방식 중 객관적 평가와 교사 평가 결과만이 강한 상관관계를 나타냈다( $r=.974$ ). 이는 교사 평가와 객관적 평가가 유용성 점수를 공유하기 때문에 독창성과 유용성 점수의 곱으로 산출되는 과학창의성 또한 밀접한 상관관계를 맺고 있다고 해석할 수 있다. 반면, 자기 평가와 객관적 평가( $r=.161$ ), 자기 평가와 교사 평가( $r=.213$ )는 상관성이 유의하게 나타나지 않았다. 이는 평가 관점에 따른 평가 결과, 비교 연구(Lin *et al.*, 2001a, b; Sadler & Good, 2006)에서 자기 평가가 교사 평가나 상대 평가에 비해 훨씬 높은 점수대라는 점과 관련 지을 수 있다.

과학창의성의 두 하위 요소인 독창성과 유용성 점수에 대해서도 세 평가 방식의 상관관계를 각각 분석하였다. 독창성 점수에 대한 학생 자기 평가, 교사 평가, 객관적 평가 간 상관관계를 분석한 결과는 Table 4와 같다. 여기에서도 세 평가 방식 중 객관적 평가와 교사 평가가 과학창의성에서 강한 상관관계를 나타냈다( $r=.713$ ). 과학창의성, 유용성과 달리 독창성 점수에 있어 객관적 평가와 교사 평가는 부분적으로도 점수를 공유하고 있지 않다. 그러므로 교사와 객관적 평가 방식이 비슷한 경향

**Table 3.** Correlation among three types of assessment on scientific creativity N=40

변인	자기 평가	객관적 평가
객관적 평가	0.161	
교사 평가	0.213	0.974***

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

**Table 4.** Correlation among three types of assessment on originality N=40

변인	자기 평가	객관적 평가
객관적 평가	0.279	
교사 평가	0.315*	0.713***

\*  $p<0.05$ , \*\*\*  $p<0.001$ .

성을 갖고 있다고 해석할 수 있다. 자기 평가와 교사 평가는  $r=.315$ 의 유의한 상관관계를 보인 반면, 자기 평가와 객관적 평가는  $r=.279$ 로 상관성이 유의하게 나타나지 않았다. 이는 독창성에 대한 교사와 학생의 평가 방식이 크게 다르지 않으나, 교사 평가에 비해 자기 평가는 그 객관성이 떨어짐을 시사한다. 이 결과는 Knowles *et al.* (2005)의 연구 결과와 일치하지만, 학생의 평가가 객관적이고 믿을 만하다는 결론을 내린 연구(Bouzidi & Jaillet, 2009; Cho *et al.*, 2006; Liu, 2002; Sadler & Good, 2006; Sung *et al.*, 2005; Tsai & Liang, 2009; Tseng & Tsai, 2007)와는 반대된다.

교사 평가와 학생의 자기 평가의 상관성이 유의미하게 높은 경우는 학생이 자기 평가가 동료 평가나 교사 평가와 비교될 것이라는 사실을 인지하거나(Fox & Dinur, 1988), 자기 평가하는 방법과 채점 기준을 명확하게 인식하고 있을 때(Ross *et al.*, 1999; Sung *et al.*, 2005)라고 한다. 반면, 자기 평가가 학생의 성적에 반영될 경우 과대평가할 가능성이 더 높은 경향을 보인다(Boud & Falchikov, 1989). 또한 Butler (1990)는 5세, 7세, 10세의 학생을 대상으로 한 연구에서 학생의 나이가 증가할수록 학생의 자기 평가와 교사 평가의 상관성이 점점 증가한다는 것을 밝혔다. 이렇듯 각 연구가 다양한 연령층의 학생을 대상으로 서로 다른 평가 내용과 방법이라는 다양한 맥락을 토대로 시행하였기 때문에 학생의 자기 평가와 교사 평가의 상관관계에 대해 상반되는 연구 결과가 나타났다(Ross, 2006).

본 연구에서는 영재 학생이 자기 평가하기에 충분한 인지 능력을 갖췄으나, 과학창의성의 의미 및 구성 요소에 대한 교육을 받아본 적이 없고, 과학창의성에 대한 객관적이고 명확한 채점 기준이 없었으며, 학생이 자기 평가가 다른 평가 방식과 비교될 것이라는 사실을 알지 못하였다. 그러므로 관련 연구를 종합해 볼 때 위 세 가지 가능성으로 인해 학생 자기 평가와 교사 평가의 상관성이 떨어졌

다고 볼 수 있다.

객관적 평가와 교사 평가는 유용성 점수를 공유하고 있으므로 자기 평가와 교사 평가의 상관성을 분석한 결과,  $r=.155$ 로 그 결과가 유의하게 나타나지 않았다. 이는 교사와 학생이 과학창의성 요소인 유용성에 대하여 크게 다른 평가 방식을 갖고 있음을 의미한다. 또한 앞선 연구 결과에서 교사가 유용성에 적용하는 기준이 높았다는 것과 선행 연구 (Sowden & Dawson, 2011)에서 엄격한 평가자는 독창성보다 유용성에 대한 기준이 더 높았다는 것과 연관된다.

4) 과학창의성에 대한 세 평가 방식 간 빈도분포

학생 자기 평가, 교사 평가, 객관적 평가의 각 평가 방식에 대하여 각 점수 분포의 중앙값을 기준으로 상·하위를 나누어 상상상, 상상하, 상하하, 하상상 등으로 그룹화하였다. 초등과학영재 학생의 과학창의성에 대하여 학생 자기 평가에 따른 객관적 평가와 교사 평가의 빈도분포 차이를 분석한 결과는 Table 5와 같다. 자기 평가에 따른 객관적 평가와 교사 평가의 빈도분포는 통계적으로 유의한 차이가 없었다.

Table 5에서 8개의 집단 중 학생 자기 평가, 객관적 평가, 교사 평가에서 모두 하위인 집단이 가장

높은 빈도 분포를 보이고 있다(13명, 32.5%). 그리고 학생 자기 평가, 객관적 평가, 교사 평가에서 모두 상위인 집단이 그 다음으로 높은 빈도 분포를 보이고 있다(10명, 25.0%). 객관적 평가가 상위, 교사 평가가 하위인 집단은 1명도 없으며, 객관적 평가가 하위, 교사 평가가 상위인 집단은 2명으로, 그 2명이 자기 평가 상위, 하위에 각각 1명씩 분포하고 있다.

위 결과에서 빈도분포에 유의한 차이는 없었으나, 각 학생의 점수에 대해 크게 상·하위로 나누어 인식하였을 때 학생, 교사, 객관적 평가 주체는 하위 집단에 대한 인식이 거의 일치하는 것으로 나타났다. 또한 상위 집단에 대한 인식도 비슷한 수준으로 일치하였다.

Salvia *et al.* (2012)에 따르면 실제적인 평가는 일관되고 정확하게 측정하는지의 신뢰도 측면과 측정하고자 하는 것을 제대로 측정하고 있는지의 타당도 측면으로 나뉜다. 지금까지의 결과로 볼 때, 평가의 신뢰도 측면에서 세 평가 방식은 유효하지만 평가의 타당도 측면에서는 세 평가 방식의 양상이 다르게 나타난다. 학생의 자기 평가가 교사 및 객관적 평가와 일치하지 않았다는 결과는 기존의 자기 평가와 교사 평가의 일관성을 입증한 연구 (Bouzidi & Jaillet, 2009; Cho *et al.*, 2006; Liu, 2002;

Table 5. Frequency differences of scientific creativity for objective assessment on student self-assessment

		객관적 평가				합계	카이제곱	p
		상		하				
		교사상	교사하	교사상	교사하			
상	빈도	10	—	1	9	20		
	객관적, 교사 평가 (%)	50.0	—	5.0	45.0	100.0		
	자기 평가 (%)	62.5	—	50.0	40.9	50.0		
자기 평가	합계 (%)	25.0	—	2.5	22.5	50.0		
	빈도	6	—	1	13	20		
	객관적, 교사 평가 (%)	30.0	—	5.0	65.0	100.0	1.727	0.422
자기 평가 (%)	37.5	—	50.0	59.1	50.0			
합계 (%)	15.0	—	2.50	32.5	50.0			
합계	빈도	16	—	2	22	40		
	객관적, 교사 평가 (%)	40.0	—	5.0	55.0	100.0		
	자기 평가 (%)	100.0	—	100.0	100.0	100.0		
	합계 (%)	40.0	—	5.0	55.0	100.0		



Sadler & Good, 2006; Sung *et al.*, 2005; Tsai & Liang, 2009; Tseng & Tsai, 2007)와 반대된다. Knowles *et al.* (2005)은 자기 평가와 교사 평가는 상관성이 낮고 학생의 자기 평가는 교사 평가에 비해 그 객관성과 타당도가 떨어진다고 진술하였다. 따라서 본 연구는 과학창의성 평가에서 학생의 자기 평가보다는 교사 및 객관적 평가가 타당도 측면에서 유효하다고 보고, 과학창의성의 핵심 요소인 독창성과 유용성에 대한 세부적인 결과를 토대로 그 양상 및 원인을 심층적으로 분석하였다.

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

### 1) 자기 평가와 교사 평가 결과 수준에 따른 그룹화

본 연구에서는 평가 방식에 따라 결과가 달라지는 원인을 규명하기 위하여 과학창의성 평가 결과의 수준에 따라 네 집단으로 나누고, 정량적 데이터 및 관찰을 통한 정성적 데이터로 집단별 특징을 분석하였다. 앞선 자기 평가, 교사 평가, 객관적 평가의 세 평가 방식 비교에서 객관적 평가를 기준으로 직관에 의지하는 자기 평가와 교사 평가 방식을 비교하였다. 집단별 심층 분석에서는 평가자 주관의 영향을 받는 자기 평가와 교사 평가만을 비교함으로써 평가의 주체에 따라 달라지는 평가 결과의 원인을 다각도로 심층 분석하였다. 이를 통해 교사와 학생 간 과학창의성 인식의 차이를 줄이는 방법, 평가의 과정이 내재된 교수·학습 방법, 학습자 중심의 과학창의성 계발 등의 여러 분야에 적용하여 교육적 함의를 이끌어낼 수 있기 때문이다.

본 연구에 참여한 40명의 학생을 대상으로 자기 평가 점수와 교사 평가 점수의 중앙값을 기준으로 상·하 수준에 따라 학생들을 네 집단으로 분류한 결과는 Table 6과 같다. 자기 평가의 상·하 수준을 나눈 중앙값은 100점 만점 중 72.95점이며, 교사 평가의 상·하 수준을 나눈 중앙값은 100점 만점 중 24.68점이다. 표에서 알 수 있는 바와 같이, 자신이 인식하는 과학창의성의 수준과 교사가 인식하는 과학창의성의 수준이 일치하는 집단은 HH, LL집단으로 각각 13명이다. 반면, 자신이 인식하는 과학창의성의 수준과 교사가 인식하는 과학창의성의 수준이 불일치하는 집단은 HL, LH집단으로 각각 7명

**Table 6.** Type distribution according to the scores' levels of student self-assessment and teacher-assessment on scientific creativity [인원수(%)]

		교사 평가	
		하위(L)	상위(H)
학생 자기 평가	상위(H)	HL(자기상-교사하) 7(17.5)	HH(자기상-교사상) 13(32.5)
	하위(L)	LL(자기하-교사하) 13(32.5)	LH(자기하-교사상) 7(17.5)

이다.

본 연구에서는 각 집단의 특징을 규명하기 위해 다음과 같은 정량적, 정성적 자료를 분석하고자 한다. 첫째, 과학창의성의 하위요소인 독창성과 유용성 점수 간 크기 비교, 둘째, 과학창의성의 하위요소인 독창성과 유용성 점수 차, 셋째, 학생의 수업 행동 관찰 결과이다.

### 2) 과학창의성의 하위요소인 독창성과 유용성 점수 차이 비교

과학창의성 평가에 있어 두 핵심 요소인 독창성과 유용성 중 어느 것의 점수가 더 큰지 비교하여 독창성 점수가 유용성 점수보다 높았던 학생의 수를 집단에 따라 분석한 결과는 Table 7과 같다. 자기 평가와 교사 평가의 비교를 용이하게 하기 위해 독창성 점수가 유용성 점수보다 더 높은 학생의 수에 초점을 두고 집단에 따라 분석하였다. 참고로 자기 평가에서는 독창성 점수가 유용성 점수보다 더 높은 학생이 9명, 독창성 점수와 유용성 점수가 같은 학생이 1명, 독창성 점수가 유용성 점수보다 더 낮은 학생이 29명으로 합산하면 40명이다. 교사 평가에서는 독창성 점수가 유용성 점수보다 더 높은 학생이 26명, 독창성 점수와 유용성 점수가 같은 학생이 1명, 독창성 점수가 유용성 점수보다 더 낮은 학생이 13명으로 합산하면 40명이다.

표에서 확인할 수 있듯이 자기 평가에서는 독창성 점수가 유용성 점수보다 더 높은 학생이 LL집단에 가장 많이 분포하였다(13명 중 5명, 38.5%). 교사 평가에서는 독창성 점수가 유용성 점수보다 더 높은 학생이 HL집단(7명 중 7명, 100%), LL집단(13명 중 11명, 84.6%)에 주로 분포하였다. 표와 앞선 연구 결과에서 확인할 수 있듯 학생은 대체로 유용성의 점수를 높게 주는 경향이 있었으나, 독창성

**Table 7.** The number of students who got the higher scores of originality than those of usefulness in student self-assessment and teacher-assessment [인원수]

		교사 평가			
		하위(L)		상위(H)	
학생 자기 평가	상위(H)	7명 중		13명 중	
		자기 평가에서 2	교사 평가에서 7	자기 평가에서 1	교사 평가에서 5
	하위(L)	13명 중		7명 중	
		자기 평가에서 5	교사 평가에서 11	자기 평가에서 1	교사 평가에서 3

점수를 높게 주는 학생은 LL집단에 주로 분포하였다. 달리 말해 자신의 과학창의성 수준을 낮게 인식하고, 교사로부터 낮은 평가를 받은 학생은 자신의 유용성 점수를 낮게 평가할 가능성이 크다.

또한 표와 앞선 연구 결과에서 확인할 수 있듯 교사는 대체로 유용성의 점수를 낮게 주는 경향이 있는데, HL, LL집단에서는 그 경향이 특히 두드러졌다. 즉, 교사는 자신의 과학창의성 수준을 높게 인식하나, 교사로부터 낮은 평가를 받은 학생은 모두 유용성 점수를 낮게 평정하였고, 두 평가 모두 낮은 평가를 받은 학생에게는 대부분 유용성 영역을 낮게 평정하였다.

**3) 과학창의성의 하위요소인 독창성과 유용성 점수 차이 비교**

과학창의성의 두 핵심 요소인 독창성과 유용성 간 점수 차이가 얼마나 크게 나는지 비교하여 그 차이가 1점 이상인 학생의 수를 분석한 결과는 Table 8과 같다. 자기 평가에서 총 40명의 학생 중 24명의 학생이 독창성과 유용성 점수의 차가 0에서

1 사이의 값을 가진 반면, 16명의 학생은 1 이상 최대 4.05의 값을 가졌다. 교사 평가에서는 총 40명의 학생 중 25명의 학생이 독창성과 유용성 점수의 차가 0에서 1 사이의 값을 가진 반면, 15명의 학생은 1 이상 최대 3.77이었다.

표에서 확인할 수 있듯이 자기 평가에서는 |독창성 - 유용성| ≥ 1인 학생이 LL집단에 가장 많이 분포하였다(13명 중 11명, 84.6%). 이는 다른 집단(LH집단 28.6%, HH집단 15.4%, HL집단 14.3%)과 비교하여 볼 때 높은 수치이다. 교사 평가에서는 |독창성 - 유용성| ≥ 1인 학생이 HL집단에 가장 많이 분포하였다(7명 중 6명, 85.7%). 이 또한 그 다음으로 많이 분포하였던 LL집단(13명 중 5명, 38.5%)을 고려하면 높은 수치라고 할 수 있다.

이러한 현상의 원인을 심층적으로 해석하기 위해 자기 평가와 교사 평가에서 각각 1명씩 독창성과 유용성 점수의 차이가 가장 큰 학생의 응답을 분석한 결과는 Table 9와 같다. 자기 평가의 분석 대상 학생은 LL집단에 속해 있으며, 교사 평가의 분석 대상 학생은 HL집단에 속해 있다.

**Table 8.** The number of students who had more than 1 point of the difference between originality and usefulness scores in student self-assessment and teacher-assessment [인원수]

		교사 평가			
		하위(L)		상위(H)	
학생 자기 평가	상위(H)	7명 중		13명 중	
		자기 평가에서 1	교사 평가에서 6	자기 평가에서 2	교사 평가에서 3
	하위(L)	13명 중		7명 중	
		자기 평가에서 11	교사 평가에서 5	자기 평가에서 2	교사 평가에서 1

**Table 9.** The responses of students with the greatest difference between originality and usefulness scores in student self-assessment and teacher-assessment

학생	응답	자기평가 점수		교사평가 점수	
		독창성	유용성	독창성	유용성
A (LL)	유압 집게 팔로 TV보다 귀찮을 때 멀리 있는 리모컨을 집어 온다.	7	2.5	4.4	2.5
	유압 리프트로 사람 키에 맞게 컵받침의 높이를 조절한다.	7.5	3.9	4.9	3.0
	유압 프레스로 쓰레기를 압축시켜 버린다.	평가하지 않음		3.6	5.5
	팔이 닿지 않는 곳을 유압 로봇 팔로 파스를 붙일 수 있다.			5.1	6.5
B (HL)	유압 집게 팔로 분리수거한다.	9	10	4.9	1.8
	유압 리프트로 장기 자량용 무대를 만든다.	10	10	4.7	2.5
	유압 프레스 리모컨을 누르면 가구들이 나오고 방이 열린다.	10	10	5.4	0.8
	유압 로봇 팔로 물건을 옮긴다.	9	10	3.6	1.3

학생 A는 수업 행동 관찰 결과, 학습 활동이나 과제에 임하는 태도가 대체로 소극적이고 학습에 자신감과 열정이 부족한 편이었으며, 창의적인 아이디어를 내려는 의지가 없었다. 이는 자기 아이디어의 가치를 판단하는 유용성 점수에 영향을 준 것으로 나타났다. 반면, 학생 B는 수업 활동에 대체로 성실하게 임하는 편이나 아이디어가 자세하지 않아, 교사가 평가할 때 유용성 점수를 현저하게 낮게 평정하였다.

#### 4) 영재 학생들의 자기 평가-교사 평가 집단에 따른 수업 행동 관찰 결과 분석

과학 영재 학생의 개별 수업 모습을 관찰한 결과, 각 학생의 과학창의성에 대한 자기 평가 점수-교사 평가 점수 유형에 따른 수업 중 개인 활동의 특징은 Table 10과 같다.

Table 10에서 HH형과 LH형의 행동 패턴이 거의 일치한다는 것을 알 수 있다. 수업에서 대체로 긍정적이고 적극적인 태도를 보이는 학생이 학생 결과물을 통해 교사가 평가한 과학창의성 점수가 높게 나온 경향을 보인다. 하지만 자신이 인식하는 과학창의성에는 개인별로 차이가 있다. 수업 행동 관찰로 파악할 수는 없지만 개인의 성격적 특성 중 완벽주의 성향이 이를 좌우한 것으로 해석된다. LoCicero and Ashby (2000)는 영재 학생의 특징 중 하나가 완벽주의 성향이지만 개인에 따라 정도의 차이가 있으며, 완벽주의 성향이 두드러질수록 자신에 대해 보다 엄격한 기준을 적용함을 밝혔다.

HL형은 학생의 수업 행동 패턴이 다양하게 드러

나 일반화하기 어려워 크게 두 유형으로 나누었다. 발표 태도가 적극적인 학생, 발표 태도가 소극적인 학생으로 나눈 결과, 발표 태도가 적극적인 학생은 수업 태도가 불량하거나 수업 집중도가 떨어지는 경향을 보였다. 발표 태도가 소극적인 학생은 수업 태도가 양호하고 활동지도 성실하게 작성하지만, 아이디어가 평범한 것에 그쳤다. 두 유형 모두 발표 태도가 적극적이면서 바람직한 학습 태도를 동시에 갖추지는 못했다는 공통점이 있다.

LL형은 수업에서 문제를 일으키는 것은 아니지만 대체로 소극적이고 학습에 대한 자신감과 열정이 부족했다. 이러한 태도가 자신이 인식하는 과학창의성 점수와 학생 결과물을 통해 교사가 평가한 과학창의성 점수에 영향을 미친 것으로 해석된다. 자신이 인식하는 과학창의성 점수는 자신에 대해 긍정적으로 생각하는 척도인 자아존중감과도 깊은 연관이 있다. 김원경과 우남희(2002)는 자아존중감이 자기효능감을 매개로 창의성에 간접적인 영향을 미친다는 보고를 하였는데, 자아존중감이 높은 사람은 자신에 대해 긍정적으로 평가하며 자신감 있게 행동할 뿐만 아니라, 창의적 과업을 달성하기 위해 끈기 있게 노력하는 적극적인 태도를 갖추고 있다는 것이다(Cropley, 2001). 이러한 견지에서 볼 때, 적극적인 태도를 갖춘 LH형 학생은 타인의 기준에서 높은 과학창의성 점수임에도 불구하고, 자신의 과학창의성 점수를 낮게 평가한 것은 낮은 자기효능감 때문이라기보다는 영재 특유의 완벽주의 성향 때문이라고 해석할 수 있다.

HL형과 LL형은 모두 교사로부터 상대적으로 낮

**Table 10.** Characteristics of individual students' activities by the scores' levels of student self-assessment and teacher-assessment on scientific creativity

유형	활동의 주요 특징
HH형	<ul style="list-style-type: none"> <li>수업에 발표하고 모둠별 활동에 적극적으로 참여하는 태도가 매우 바람직함. 자기 주도적으로 학습하는 탐구심이 돋보이며, 종종 창의적인 아이디어를 낸.</li> </ul>
HL형	<ul style="list-style-type: none"> <li>창의적으로 발표하려는 태도가 돋보이나, 수업 태도가 불량한 경향이 있어 성실하고 바른 학습 태도를 갖출 필요가 있음. 만들기 수행에 있어 뒤떨어지는 경향이 있어 수업에 좀 더 집중할 필요가 있음.</li> <li>수업에 참여하는 태도가 양호하나, 발표 횟수가 거의 없으므로 수업 참여에 있어 좀 더 성실성과 적극성을 가지기를 바람. 활동지를 정성껏 작성하나 일반적인 아이디어에 그침.</li> </ul>
LH형	<ul style="list-style-type: none"> <li>수업에서 발표하고 모둠별 활동에 적극적으로 참여하는 태도가 바람직함. 자기 주도적으로 학습하는 탐구심이 돋보이며 종종 창의적인 아이디어를 낸.</li> </ul>
LL형	<ul style="list-style-type: none"> <li>학습 활동이나 과제에 임하는 태도가 대체로 소극적이고, 학습에 자신감과 열정이 부족한 편임. 창의적인 아이디어를 내려는 의지가 없음.</li> </ul>

은 과학창의성 점수를 받았다. LL형은 대체로 소극적이고 자신감이 부족했던 것으로 보아 낮은 자아 존중감을 갖고 있는 것으로 파악되며, HL형은 자신의 과학창의성을 높게 평가하였지만 태도가 적극적인 학생, 소극적인 학생이 혼재하는 다양한 패턴으로 나타났다.

Table 10은 학생의 유형별 행동을 패턴화한 것으로, 패턴을 벗어나는 특이한 경우의 학생도 있었다.

<HH형 특이한 경우>

- 김\*\* : 창의적 아이디어를 잘 내지만 집중력이 떨어지고 지시 사항을 종종 이해 못하는 등 수업 참여 태도가 약간 불량하므로 성실성을 갖출 필요가 있음.
- 김\*\* : 모둠 활동에 협력하는 태도가 양호함. 발표 횟수가 거의 없어서 창의적으로 사고하는지 알 수 없음.

<LH형 특이한 경우>

- 김\*\* : 발표 횟수는 많지 않으나 한 번 발표할 때마다 매우 창의적인 아이디어를 낸. 학습에 좀 더 성실성과 적극성을 가지고 참여한다면 창의성을 더 계발할 수 있을 것임.

Table 10에서 HH형과 LH형은 발표 태도가 적극적인 경향을 보인다고 하였으나, 모두가 그런 것은 아니며 비교적 소극적이고 발표를 거의 하지 않는 학생이 2명 있었다. 또한 HH형은 대체로 바람직한 수업 태도를 갖추었으나 수업 집중도가 떨어지고, 수업에 참여하는 태도가 성실하지 않은 학생이 1명 있었다.

본 연구에서 초등과학영재학생의 과학창의성 및 그 하위요소에 대한 학생 자기 평가, 교사 평가, 객관적 평가를 비교 분석한 결과는 다음과 같다.

첫째, 세 평가 방식이 신뢰할 만한 것인지를 확인하기 위해 각 주차별 과학창의성과 1~4주차의 과학창의성 점수를 평균한 값의 상관관계를 분석한 결과, 학생의 과학창의성이 일관성 있게 평가되고 있음을 확인할 수 있었다. 자기, 교사, 객관적 평가 모두 평가한 시기 동안 일관성 있게 나타나므로 평가의 주요 요건인 신뢰성을 만족시킨다고 할 수 있다.

둘째, 과학창의성 및 그 하위요소인 독창성, 유용성 점수의 평균값은 자기 평가, 객관적 평가, 교사 평가의 순으로 높았다. 그러므로 객관적 평가와 비교하여 볼 때 교사는 엄격한 평가 기준을 적용하는 반면, 학생은 자기 자신에 대해 관대한 평가 기준을 적용한다. 특히 유용성 점수에서 자기 평가와 교사 평가 간 차이가 두드러졌는데, 엄격한 평가자는 창의성의 요소 중 독창성보다 유용성에 대한 기준이 더 높은 것으로 나타났다. 하지만 과학적 타당도를 나타내기도 하는 유용성 점수는 가치 판단이 개입될 수 있으므로, 이를 더 객관적으로 평가할 수 있는 구체적인 방안을 모색할 필요가 있다.

셋째, 상관관계 및 빈도 분석을 통해 교사 평가와 객관적 평가는 강한 상관관계를 맺으며 통계적으로 유의한 차이가 있었다. 다시 말해, 교사 평가와 객관적 평가는 비슷한 경향성을 띠는 반면 자기 평가는 다른 두 평가와 관련성이 적었다. 평가의 타당성 측면에서 교사와 객관적 평가가 상대적으로 공정하고 신뢰할 만한 방법인 것으로 나타났다

**IV. 결론 및 제언**

므로, 두 평가가 평가의 비중에서 많은 부분을 차지하는 것이 바람직할 것이다. 하지만 자기 평가가 학생의 메타 인지적 사고를 자극하고 스스로 학습 성취를 조정하게 하는 데 효과적이기 때문에 자기 평가를 평가 방식에서 배제할 수는 없다. 자기 평가는 학생이 학습 과정에 책임감을 가지고 몰입할 수 있도록 도와주며, 학습 동기를 부여하므로 교수·학습 활동에 활용해 볼 수 있다.

넷째, 독창성 점수의 경우 교사와 객관적 평가, 교사와 자기 평가가 유의한 상관성을 나타냈으며, 유용성 점수의 경우 교사와 자기 평가는 상관성이 유의하게 나타나지 않았다. 이는 과학창의성의 평가 요소인 독창성이 객관적으로 파악이 가능하나, 다른 요소인 유용성은 평가의 주체에 따라 결과가 달라지므로 평가자 주관의 영향을 받고 있음을 나타낸다. 유용성의 정의대로 얼마나 유용하고 가치가 있는가, 타당한가를 판단하는 문제는 가치 판단의 문제와 연결된다. 본 연구에서 학생들은 독창성, 유용성의 의미와 그것이 과학창의성과 갖는 관계를 알지 못한 채 자기 평가에 임했으므로 평가 기준에 따른 정확한 채점이 어려웠을 수 있다. 따라서 이러한 주관성의 문제를 해결하기 위해 학습자가 채점기준, 평가 요소를 정확하게 이해하도록 과학창의성의 두 요소인 독창성, 유용성에 대한 정의와 특징을 배우는 선행 학습을 제공할 필요가 있다.

다섯째, 학생들의 산물에 대한 창의성 요소들을 자신이 평가하게 했을 때 교사나 공식에 따른 평가보다 더 높게 나타났는데, 자신의 산물을 평가하는 방식 이외에 다른 학생의 산물을 동료 학생으로서 평가하게 했을 때 어떠한 패턴을 보이는지를 조사·분석할 필요가 있다. 과학창의성 평가에서 학생의 자기 평가보다는 교사 및 객관적 평가가 타당도 측면에서 유효할 수 있다. 하지만 이러한 방식은 학생 수준에서 과학적으로 새로우면서 유용한 것을 만들어내는 능력을 학생 과학 창의성이라고 보는 littl-c 창의성에 근거한 것이 아니라, 과학자 수준 창의성인 Big-C 창의성에 근거하여 학생의 과학창의성을 평가할 가능성이 있다. 후속 연구로써 동료 평가와 자기 평가, 교사 평가, 객관적 평가의 비교 분석이 이루어진다면 다른 평가와 맺는 관련성을 통해 동료 평가의 교육적 활용 방법과 littl-c 창의성의 본질을 탐구할 수 있을 것이다.

여섯째, 학생이 스스로 평가한 과학창의성 점수

는 자아존중감과 연관될 수 있다. 자아존중감이 낮은 학생은 자신에 대해 부정적으로 평가할 가능성이 크기 때문이다. 반면에 자아존중감이 높더라도 스스로를 낮게 평가하는 경우도 있었다. 이러한 유형의 학생은 관찰 결과, 적극적이고 긍정적인 태도를 지녔으므로 자아존중감의 문제라기보다는 영재 특유의 완벽주의 성향으로 인해 자신에게 높은 기준을 적용했을 가능성이 있다. 이렇듯 과학창의성에 대한 자기 평가에 영향을 미치는 요소는 영재 개인의 성격적 특성에 내재하고 있으며, 자아존중감이 낮은 학생이라면 높은 자기효능감을 가질 수 있도록 격려하여 과학창의성에 좋은 영향을 미칠 수 있도록 해야 한다.

본 연구에서는 초등과학영재학생이 스스로 인식하는 과학창의성과 교사가 평가한 과학창의성 및 객관적 평가에 의한 과학창의성을 각각도로 탐구·분석하였다. 실시한 평가 방식 중 공정하고 타당한 평가 방식은 객관적 평가와 교사 평가였는데, 교사 평가가 Big-C 창의성에 근거하여 평가했을 가능성을 감안하면 객관적 평가가 창의성 평가에 있어 가장 효과적이었다고 볼 수 있다. 더욱이 객관적 평가는 교사 평가에 비해 학생 자기 평가와 크게 차이 나지 않았다.

과학창의성은 고차원적 사고 과정을 통한 결과물이므로 학생의 진짜 생각이 무엇인지에 근거하여 평가해야 실제적 평가라고 할 수 있을 것이다. 그러므로 학생, 교사가 객관적 평가에서 차이가 나타나는 원인이나 메커니즘을 규명하고, 나아가 이러한 격차를 줄일 수 있는 방안을 모색할 필요가 있다. 또한, 창의성에서 많은 아이디어를 산출하는 능력인 유창성 요소가 강조되는 경우가 많은데, 유창성과 과학창의성의 관계를 체계적으로 규명할 필요가 있다.

## 참고문헌

- 곽금주, 정운경, 김민화(2010). 아동발달심리학. 서울: 박학사.
- 교육부(2015). 초·중등학교 교육과정 총론. 교육부 고시 제 2015-74호 [별책 1].
- 김원경, 우남희(2002). 아동의 창의성에 대한 심리적 관련 변인 연구. 아동학회지, 23(2), 1-16.
- 신지은, 한기순, 정현철, 박병건, 최승언(2002). 과학 영재 학생과 일반 학생은 창의성에서 어떻게 다른가? -

- 서울대학교 과학영재교육센터 학생들을 중심으로. 한국과학교육학회지, 22(1), 158-175.
- 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.
- 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, H. (2007). Researching electronic portfolios and learner engagement: the reflective initiative. *Adolescent and Adult Literacy*, 50(6), 436-449.
- Boud, D. & Falchikov, N. (1989). Quantitative studies of student self-assessment in higher education: A critical analysis of findings. *Higher Education*, 18, 529-549.
- Bouzidi, L. & Jaillet, A. (2009). Can online peer assessment be trusted? *Educational Technology & Society*, 12(4), 257-268.
- Butler, R. (1990). The effects of mastery and competitive conditions on self-assessment at different ages. *Child Development*, 61, 201-210.
- Callahan, C. M. & Miller, E. M. (2005). A child-responsive model of giftedness. *Conceptions of Giftedness*, 2, 38-51.
- Carmines, E. G. & Zeller, R. A. (1979). Reliability and validity assessment (Vol. 17). Sage publications.
- Chan, D. W. (2005). Self-perceived creativity, family hardiness, and emotional intelligence of Chinese gifted students in Hong Kong. *Journal of Secondary Gifted Education*, 16(2-3), 47-56.
- Chang, C. C., Tseng, K. H. & Lou, S. J. (2012). A comparative analysis of the consistency and difference among teacher-assessment, student self-assessment and peer-assessment in a Web-based portfolio assessment environment for high school students. *Computers & Education*, 58(1), 303-320.
- Cho, K., Schunn, C. & Wilson, R. (2006). Validity and reliability of scaffolded peer assessment of writing from instructor and student perspectives. *Journal of Educational Psychology*, 98(4), 891-901.
- Cropley, A. J. (2001). Creativity in education & learning: A guide for teachers and educators. Psychology Press.
- Fox, S. & Dinur, Y. (1988). Validity of self-assessment: A field evaluation. *Personnel Psychology*, 41, 581-592.
- Hu, W. & Adey, P. (2002). A scientific creativity test for secondary school students. *International Journal of Science Education*, 24(4), 389-403.
- Knowles, M. S., Holton, E. F. III. & Swanson, R. A. (2005). The adult learner (6th Ed.). Burlington, MA: Elsevier.
- Lin, S.-J., Liu, Z.-F. & Yuan, S.-M. (2001a). Web-based peer assessment: Attitude and achievement. *IEEE Transactions on Education*, 44(2), 13.
- Lin, S.-J., Liu, Z.-F. & Yuan, S.-M. (2001b). Web-based peer assessment: Feedback for students with various thinking-styles. *Journal of Computer Assisted Learning*, 17(4), 420-432.
- Liu, Z.-F. (2002). *Studies of networked peer assessment*. Unpublished doctoral dissertation. Hsinchu: National Chiao Tung University.
- 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.
- Mayer, R. E. (1999). Fifty years of creativity research. In R. J. Sternberg (ed.), *Handbook of creativity*. New York: Cambridge University Press.
- Mumford, M. D. (2003). Where have we been, where are we going? Taking stock in creativity research. *Creativity Research Journal*, 15(2-3), 107-120.
- Musante, K. & DeWalt, B. R. (2010). Participant observation: A guide for fieldworkers. Rowman Altamira.
- Newton, L. & Newton, D. (2010). Creative thinking and teaching for creativity in elementary school science. *Gifted and Talented International*, 25(2), 111-124.
- Patton, M. Q. (2005). Qualitative research. Encyclopedia of statistics in behavioral science.
- Plake, B. S. & Impara, J. C. (1996). Teacher assessment literacy: What do teachers know about assessment?. In *Handbook of classroom assessment* (pp. 53-68).
- Renzulli, J. S. (2003). Conception of giftedness and its relationship to the development of social capital. *Handbook of Gifted Education*, 3, 75-87.
- Ross, J. A. (2006). The reliability, validity, and utility of self-assessment. *Practical Assessment, Research and Evaluation*, 11(10). Retrieved January 31, 2009 from <http://pareonline.net/getvn.asp?v=11&n=10>
- Ross, J. A., Rolheiser, C. & Hogaboam-Gray, A. (1999). Effect of self-evaluation on narrative writing. *Assessing Writing*, 6(1), 107-132.
- Sadler, P. & Good, E. (2006). The impact of self- and peer-grading on student learning. *Educational Assessment*, 11(1), 1-31.
- Salvia, J., Ysseldyke, J. & Witmer, S. (2012). Assessment: In special and inclusive education. Cengage Learning.
- 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.
- Sternberg, R. J. (1998, Ed.). *Handbook of human creativity*. Cambridge: Cambridge University Press.
- 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.
- Treffinger, D., Young, G., Shelby, E. & Shepardson, C. (2002). *Assessing creativity: A guide for educators (RM02170)*. Storrs: University of Connecticut, The National Research Center on the Gifted and Talented.
- Tsai, C.-C. & Liang, J.-C. (2009). The development of science activities via on-line peer assessment: The role of scientific epistemological views. *Instructional Science*, 37, 293-310.
- Tseng, S.-C. & Tsai, C.-C. (2007). On-line peer assessment and the role of the peer feedback: A study of high school computer course. *Computers & Education*, 49(4), 1161-1174.
- Walker, C. & Gleaves, A. (2008). An exploration of students' perceptions and understandings of creativity as an assessment criterion in undergraduate-level studies within higher education. *Irish Educational Studies*, 27(1), 41-54.
- Weisberg, R. W. (1993). *Creativity: Beyond the myth of genius*. WH Freeman New York.
- Yager, R. E. (2000). A vision for what science education should be like for the first 25 years of a new millennium. *School Science and Mathematics*, 100(6), 327-341.

---

김민주, 서울문현초등학교 교사(Kim, Min-Ju; Teacher, Seoul Munhyun Elementary School).

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