

창의적 산출물 평가에서 평정자의 지식 및 평가 도구 유형에 따른 일치도 분석

이 수 진
월영초등학교

최 호 성
경남대학교

박 경 희
경남대학교

본 연구에서는 창의적 산출물 평가의 대표적 도구인 O'Quin과 Besemer(1989)의 '창의적 산출물 분석행렬(CPAM)'과 '창의적 산출물 어의 척도(CPSS)'의 두 가지의 산출물 평정 척도를 적용하여 동일한 산출물을 평가하는 데 있어 평가 결과의 차이를 분석하였다. 연구 결과, 창의적 산출물 평가에서 평가자의 암묵적 지식은 존재하며, 암묵적 지식과 명시적 지식에 기반한 평가의 평정 결과는 차이가 있는 것으로 나타났다. 평가자에게 명시적 지식이 쌓여 갈수록 산출물에 대한 전반적인 평가 점수는 낮게 나타나 더 엄격해지는 경향을 보였다. 상대적으로 세분화된 채점기준을 가진 CPSS로 평정하였을 때는 창의성의 하위 차원에서 점수가 모두 낮게 나타났으며, 평가자가 산출물에 대한 첫인상이나 개인적 평가 기준이 너무 확고한 경우에는 평정도구에 영향을 받지 않을 수도 있다는 것을 살펴볼 수 있었다. 영재교육 담당교사들은 창의적 산출물 평가에서 전문가들과 유사한 평가를 하고 있었으며 산출물 평가 도구를 사전에 충분히 숙지하고 산출물을 지도하거나 평가한다면 보다 새롭고, 유용하고 적절하며, 해결성이 높은 창의적인 산출물이 생성될 수 있을 것이다.

주제어: 창의적 산출물 평가, 평가도구 유형, 평가 일치도, 평정자의 지식, CPAM, CPSS

I. 서 론

21세기 지식정보사회가 도래하면서 창의성 개발에 대한 요구는 세계 여러 나라의 교육 개혁을 견인하고 있다. 미래사회를 대비하기 위해 지식과 기능을 통합하여, 다양한 맥락과 상황에서 이를 유연하게 적용하고, 상상력과 창의력을 발휘하여 새로운 가치를 창조해 내는 능력이 공통적으로 강조되고 있다. 우리 정부 또한 이에 대응하여 2015년 9월 23일, 2015 개정교육과정을 발표하고, 인문학적 상상력과 과학기술 창조력을 갖춘 '창의융합형 인재 양성'으로 개정의 중점을 밝혔다. 창의성에 대한 초기 연구에서는 특별한 사람만이 지니고 있는 능력이라고 생각하던 것에서 최근에는 누구에게나 있고, 길러질 수 있다는 논리가 많아지고 있다는 연구(Runco,

Johnson & Baer, 1993)를 국가 수준 교육과정에서도 쉽게 확인할 수 있는 것이다.

이와 같이 국가 교육의 목표이자 방향인 ‘창의성’은 학교 안팎에서의 다양한 아이디어, 과제, 각종 활동 및 산출물 등 통해 발현되고 또 발견되어진다. 산출물은 어느 분야에나 있으며, 영재교육 또는 예술 영역과 같이 창의적 산출을 요구하는 교육장면에서는 더욱 유용할 수 있다. 산출물은 그것을 창조한 사람, 생성하기까지의 과정 및 환경을 포함하여 창의성에 영향을 주는 제 요소에 대한 정보를 제공하기 때문에 MacKinnon(1987)은 “창의성 연구의 기본 원리는 창의적 산출물을 분석하는 것에서 시작되며 이는 창의적 산출물과 다른 평범한 산출물의 차이점을 파악하는 것이다”라고 강조하였다.

뿐만 아니라 창의성을 평가하는 장면에 있어서도 창의적 산출물은 주목받기 시작하였다. 기존의 창의성 검사들의 타당도에 문제가 제기 되면서 창의성이 발현된 결과로서의 산출물 자체에 주목하고, 이에 대한 직접적인 평가를 통해서 창의성 수준을 판단해 보려는 검사들이 새로운 대안으로 제시되고 있는 것이다(Amabile, 1982; Besemer, 1998).

사람들이 창의적 행동이나 산출물을 평가할 때, 조직화되어있지는 않지만 일상생활 속에서 형성된 창의성에 대한 자신들의 생각은 평가의 심리적 기초가 된다(Sternberg, 1985). 이러한 창의성에 대한 암묵적 이론(implicit theories)은 Sternberg, Conway, Kerton과 Bernstein(1981)의 연구 이후로 증가하기 시작하였다. 창의적 산출물에 대해 가지고 있는 암묵적 이론을 살펴보는 것은 창의적 산출물에 대한 사람들의 전반적인 생각을 파악해 볼 수 있고, 중요하게 여기는 관점을 도출할 수 있으며, 그 내용들을 분석함으로써 창의적 산출물을 교육하기 위한 방향을 정립할 수 있다. 이와 같이, 창의적 산출물을 평가할 때 사람들의 머릿속에 내재되어 있는 암묵적인 지식은 영향을 끼칠 수 있기에 이에 대한 논의의 의미가 있다.

창의적 산출물 평가와 같은 주관적 평가 상황에서는 머릿속에 내재된 암묵적인 지식에 근거하여 평가하는 경우와 명시적인 평가 기준을 인지한 후 평가하는 경우, 그 결과는 달라질 수 있다. 즉 평가 기준에 대한 인지 정도는 평가 결과에 영향을 미칠 수 있을 것이다. 만약, 양자의 평가 결과가 일치한다면 창의적 산출물에 대한 평가 기준의 보편적 합의를 가정해볼 수 있고 주관적 평가에 있어서 제기되는 객관성이나 신뢰도의 문제 역시 해결될 수 있을 것이기에 본 연구에서는 이들 간에 어떠한 차이가 있는지를 살펴보고자 하였다.

산출물 평가 도구로는 O'Quin과 Besemer(1989)가 개발한 ‘창의적 산출물 분석행렬(CPAM)’과 ‘창의적 산출물 어의 척도(CPSS)’를 적용하여 동일한 산출물을 평가하는 데 있어 평가 결과의 차이를 분석해보고자 한다. 이들은 동일한 평가 차원을 지녔으나 채점기준의 세분화 정도에는 차이가 있다. 평정 척도의 명시적, 세분화된 정도에 따른 평정 결과를 분석해 보는 것은 창의적 산출물 평정 도구 선택에 있어서 시사점을 제공해 줄 수 있을 것이다.

이와 같은 목적에 따라 설정된 연구문제는 다음과 같다.

- 1) 창의적 산출물 평가에 있어서 평정자의 암묵적 지식과 명시적 지식에 근거한 평가결과가 일치하는가?
- 2) 창의적 산출물 평가에서 활용하는 평정 도구에 따라 평정 결과에 어떤 차이가 있는가?

II. 이론적 배경

1. 창의성의 평가에서의 암묵적 지식 활용

창의성에 대한 암묵적 이론 연구는 Sternberg(1985)가 지혜, 지능, 창의성에 대한 암묵적 이론 연구를 시작하면서부터 활발하게 진행되었다. 지능이 무엇인지 알기 위한 역사적으로 중요한 접근 중의 하나는 전문가 의견에 의존하는 것으로, Sternberg(1985, 2004)는 이것은 ‘암묵적 이론(implicit theories)’이라고 설명하고 과학적 가설과 자료 수집을 위한 근거로 사용되는 ‘명시적 이론(explicit theories)’과 구별하였다. 명시적 이론은 심리학자나 과학자들이 구성한 것으로서 사람의 심리작용을 측정하는 데 유용하다고 여겨지는 다양한 과제 수행 결과에 기초한 것이며 무수한 논문들의 근거가 된다. 즉, 자료에 기초하여 학자들에 의해 제시된 공식적인 이론이며, 문헌을 통하여 제시되고 검증되는 것이다.

이와는 대조적으로 나타나는 암묵적 이론에 대해 Sternberg(1985)는 지적인 사람이 어떠한 특성을 가질 것이라고 사람들이 생각할 때는 이미 자신의 머릿속에 지능에 대한 암묵적인 개념이 형성되어 있고, 지적인 사람의 원형(prototype)을 떠올린다고 주장하였다. Sternberg는 암묵적 이론에 대해 한 개인이 성장하면서 겪은 경험과 사회·문화적 가치에 영향을 받아 형성된 심리학적 구성개념이라고 정의하였고, 공식적 또는 형식적으로 표출된 것이 아니라, 어떠한 형태든 사람들의 머릿속에 이미 존재하는 것이기에 ‘발명’의 대상이 아닌 질문과 추론 등을 통해 ‘발견’되는 대상으로 보았다.

암묵적 이론에 대해 Wegner와 Vallacher(1977)는 학자들처럼 의식적으로 이론적인 틀을 구성하지는 않지만 일반인도 어느 정도 조직화된 틀에 가지고 있으며, 이를 사회적 지식의 도식으로 사용한다고 하였다. Romo와 Alfonso(2003)는 사회적 지식이 조직화된 것으로 보고, 개인이 세상을 나름대로 통제하고 예측하고자 하며, 이해하려는 요구에 의해 개인적으로 형성해 나가는 것이라고 보았다. 이처럼 암묵적 이론은 사회나 문화에서 공통적으로 인식하고 있는 생각과 견해를 포함하고 있으며, 개개인이 가지고 있는 견해나 개념을 함께 포함한다고 볼 수 있다. 암묵적 이론에 대한 연구는 어떤 심리학적 구성개념에 대한 사람들의 지배적인 보편 문화적 관점(common-cultural views)과 까닭을 명확히 하는데 유용하고, 명시적 이론을 이해하는데 토대를 제공할 수 있다. 또한 체계적으로 조직화 되어있지는 않지만 사람들에게 내재되어 있으며 사회·문화의 영향을 받아 삶의 전체 기간 동안 형성되며, 직·간접적으로 후세와 타인에게 교육되어지기 때문에 자신과 타인을 평가하는 심리적 기초가 될 수 있다(최인수, 임미정, 표정민, 2011).

2. 창의적 산출물의 개념과 주요 평가 도구

창의적 산출을 수행하는 것은 학습자가 배워야 할 수단이면서 동시에 학습한 결과를 증명하는 창의성 교육의 목적으로서도 의미가 크다.

창의적 산출물에 대해 Bruner(1962)는 놀랍고, 새롭고, 유용한 것이라고 설명하며, 창의적

산출물에 대한 기준을 제시한 바 있다. Rhodes(1987)는 ‘사고의 인공물(the artifacts of thoughts)’이라고 표현하며, 새로운 개념이 만들어지는 순간 개인의 통찰 기록이 산출물에 나타날 수 있다고 설명하였다. Reis와 Renzulli(1991)는 추상적이거나 구체적인 형태를 모두 포함하지만 그 결과는 반드시 구체적으로 표현되어야 유의미하다고 주장하였다. 영재가 구체적인 산출물을 만들어냈다는 것은 그 수행 과정에서 추상적인 산출물을 경험했다는 것을 의미하며, 이는 다른 학습상황에 전이될 수 있기 때문에 높은 가치를 지닌다고 하였다. Renzulli, Leppin과 Hays(2000)는 학생의 학습 경험에 대한 결과물로, 다른 학생들과 공유하기 위해 알고 있는 것을 구체적으로 표현한 것을 산출물이라고 정의하였다.

이를 종합해 보면 창의적 산출물은 인간의 창의적 사고 과정의 결과물을 총칭하는 것으로 타인과 사회의 영역 안에서 인정받고 공유될 수 있도록 개인의 잠재력이 밖으로 드러나야 한다. 이는 산출물의 개념 뿐만 아니라 창의성을 정의하는 데에 있어서 중요한 전제가 될 수 있다.

그 동안 창의성에 대한 심리 측정적 접근은 인지적 및 정서적 특성에 초점을 두었으나 많은 창의성 연구자들은 이러한 검사의 타당도에 대해 문제를 제기하며, 기존의 검사들이 예측 변수는 측정하지만 창의성을 직접적으로 평가하는지는 불분명하다고 주장하였다(Amabile, 1982). 이러한 문제는 창의성의 외적 준거의 필요성이 제기되는 배경이 되었고, 많은 창의성 연구자들은 창의적 산출물이 외적 준거로서 적절함을 제안하였다(Amabile, 1983, 1996; Besemer & Treffinger, 1981; Jackson & Messick, 1967; Treffinger & Poggio, 1972). 즉, 창의성이 발현된 결과로서 창의적 산출물 자체에 주목하고, 이에 대한 평가를 통해서 창의성의 보유 여부를 판단하는 검사들이 대안으로 제시되고 있는 것이다(Amabile, 1982; Besemer, 1998).

1980년대 이후 창의적 산출물에 대한 구체적인 연구가 시작되었고, 측정 도구가 개발되기 시작하였다(Horn & Salvendy, 2006). Amabile(1982)은 해당 영역의 전문가들이 창의적이라고 합의할 수 있는 암묵적 준거에 의해 산출물을 평가할 수 있다는 논리로 합의적 평가 기법(consensual assessment technique)을 개발하였다. Besemer와 Treffinger(1981)는 창의적 산출물과 관련된 문헌 연구를 바탕으로 평가 준거를 도출하여 3가지 차원을 설명하는 창의적 산출물 분석 행렬(The creative product analysis matrix)을 제시하였고, Besemer와 O'Quin(1986)은 창의적 산출물 분석 행렬의 요인 구조를 바탕으로 창의적 산출물 어의 척도(creative product semantic scale)를 개발하였다. Reis와 Renzulli(1991)는 산출물의 전반적인 우수성과 개별성을 평가하는 15개 항목으로 구성된 학생 산출물 평가 양식(student product assessment form)을 개발하였다.

국내 창의성 교육 연구들이 창의성에 대해 어떠한 연구관점을 가지고 있는지를 분석한 조연순과 정지은(2012)에 따르면 창의성 범주-개인, 과정, 산출물, 환경-중 개인적인 특성으로 조망한 연구가 전체 399편 중 242편(60.66%)로 가장 높은 비율로 나타났으며, 반대로 산출물에 관심을 두었던 연구는 전체 399편 중 19편(4.76%)로 가장 낮은 비율로 분석되었다. 이처럼 산출물과 산출물에 대한 평가가 창의성 연구에서 지니는 중요성에도 불구하고 관련 연구는 턱없이 부족하다. 한기순(2005) 역시 아동의 행동이나 행동 산출물에 근거하여 창의성을 측정

하거나 평가한 연구는 거의 없으며 수행평가를 사용한 경우라도 한 두 영역의 창의성 측정에만 그치는 것이 현실이며 전문가 평정을 활용한 연구들은 미비한 편이라고 밝힌 바 있다.

가. 창의적 산출물 분석 행렬(CPAM)

Besemer와 Treffinger(1981)는 창의성의 준거(criteria)가 나타난 연구물을 참고하여, 125개가 넘는 형용사 준거들을 추출하고 유사성을 고려하여 14개로 범주화한 후, ‘새로움(novelty), 해결성(resolution), 정교성 및 종합성(elaboration & synthesis)’의 3개 차원으로 위계화한 ‘창의적 산출물 분석 행렬(The creative product analysis matrix, 이하 CPAM)’을 개발하였다. Besemer(1998), Besemer와 Treffinger(1981), Besemer와 O’Quin(1986), O’Quin과 Besemer(1989) 등의 연구를 통해 CPAM은 20여 년간 지속적인 타당화 작업을 거쳤고, 3개 차원 및 9개의 하위 요소로 수정되었다(Besemer, 1998).

첫 번째 차원인 새로움(novelty)은 산출물이 새로운 정도를 의미한다. 이것은 산출물에 새로운 과정(process), 기술(technique), 재료(material), 개념(concept)이 얼마나 포함되어 있는지에 따라 판단할 수 있다. 하위 요소로는 ‘독창성(original), 놀라움(surprise)’이 포함되어 있다.

두 번째 차원인 해결성(resolution)은 산출물의 문제 해결 능력이 얼마나 적절한지, 산출물이 목적에 맞게 제 기능을 하는지에 대한 측면이라고 할 수 있다. 하위 요소로 ‘논리성(logical), 유용성(useful), 가치로움(valuable), 이해가능성(understandable)’이 구성되어 있다.

세 번째 차원인 정교성과 종합성(elaboration & synthesis)은 산출물의 형식미(stylistic)에 대한 것으로, 얼마나 세련되게(elegant) 제작되었는지, 문제에 대한 해결책이 어떠한 방식으로 나타나는지에 대하여 판단하는 차원이다. ‘유기적조직성(organic), 우아함(elegance), 완성도(well-craftedness)’가 하위 요소로 구성되어 있다.

나. 창의적 산출물 어의 척도(CPSS)

Besemer와 O’Quin(1986)은 CPAM을 바탕으로 산출물을 평가할 수 있도록 양극 형용사 평정 척도 형태인 ‘창의적 산출물 어의 척도(creative product semantic scale, 이하 CPSS)’를 개발하였으며, 지속적인 타당화 및 수정작업을 수행하였다.

초기의 Besemer와 Treffinger(1981)는 문헌에서 수집했던 110개의 준거 표현들로 형용사 문항을 구성하였다. 이 후 이 문항 리스트를 평정 척도로 만들어 CPAM의 3요인 모형을 경험적으로 검증하려는 시도를 하였다(Besemer & O’Quin, 1987). 그러나 100개가 넘는 문항은 너무 길고, 응답자의 피로 효과 등의 문제가 제기되었다. 이에 따라 Besemer와 O’Quin(1986)은 Osgoods 등(1957)에 의해 제안된 의미변별척도(semantic differential format)를 활용하여 양극 형용사 척도의 형태로 개발하였다. 이후로도 Besemer(1998), Besemer와 O’Quin(1999), O’Quin과 Besemer(1989)에 의해 타당화와 수정 축소 작업을 거쳐 현재는 3개 요인의 9개 하위 차원에 대한 43개 양극 형용사 척도로 이루어진 형태를 띠게 되었다.

CPSS는 CPAM에 기초하여 제작되었기 때문에 다른 평정척도에 비하여 산출물의 창의성에 내재된 이론적 차원구조를 명시적으로 반영하고, 보다 폭 넓은 상황에 적용할 수 있다(김영록,

이순목, 2004). Besemer(1998)의 연구를 통해 CPSS를 활용한 창의적 산출물에 대한 평가는 영역의 전문가 뿐만 아니라 관리자, 소비자 등 다양한 계층이 평가자가 될 수 있으며 3차원의 CPSS가 타당하다는 것이 검증되었다.

III. 연구 방법

1. 연구 대상

산출물의 창의성을 평가하는 데 있어서 창의성의 수준과 정도에 대한 판단은 평가자에게 의존하기 마련이다. 특히, 영재교육과정에서는 영재 학생들은 해당 영역에서 전문가가 수행하는 것과 같은 과정을 거쳐 질 높은 산출물을 만들어 내도록 격려되고 있고, 산출물을 만들어 내는 과정에서 학생들의 개성·상상력·창의력이 다양하고 충분히 반영되는 것이 중요하기에(교육부, 2008) 영재교육을 담당하는 교사들은 일반교사들보다 학생들의 창의적인 산출물을 교육하고, 평가해야 할 기회가 빈번하다고 판단된다. 그러므로 본 연구에서는 영재교육을 담당하는 초등학교 교사들을 대상으로 실시하였으며, 산출물 평가자의 특성은 <표 1>과 같다.

<표 1> 산출물 평가자 일반 현황

변수	구분	빈도	비율
영재교육담당	영재학급 강사	32	23.0
	영재교육원 강사	107	77.0
영재 교육분야	발명	51	36.7
	과학	34	24.5
	수학, 정보	28	20.1
	언어인문기타	26	18.7
영재연수이수	해당 없음	3	2.2
	180시간 미만	57	41.0
	180시간 이상	79	56.8
전체	139	100.0	

전체 139명의 교사를 대상으로 분석하였으며, 재직경력은 ‘5년 미만’이 6명(4.3%), ‘5년 이상-10년 미만’이 31명(22.3%), ‘10년 이상-15년 미만’이 44명(31.7%), ‘15년 이상-20년 미만’이 29명(20.9%), ‘20년 이상’이 20명(20.9%)으로 나타났다. 학력은 ‘대학 졸업’ 33명(23.7%), ‘석사과정 수료 및 졸업’ 85명(61.2%), ‘박사과정 수료 및 졸업’ 21명(15.1%)이다. 영재교육담당은 ‘영재학급 강사’는 32명(23.0%), ‘영재교육원 강사’는 107명(77.0%)이었으며, 영재교육분야는 ‘발명’이 51명(36.7%), ‘과학’이 34명(24.5%), 수학과 정보가 28명(20.1%), 언어인문기타가 26명(18.7%)이었다. 영재교육 연수이수는 ‘180시간 이상’이 79명(56.8%), ‘180시간 미만’이 57명(41.0%)로 대부분 연수를 이수한 것으로 나타났다.

2. 연구 도구

전문가의 암묵적 준거에 의존하고 있는 Amabile(1982)의 CAT는 창의적 산출물의 세부적인 특징에 관한 정보를 제공하지 못하지만 CPSS는 세부적으로 창의적인 산출물의 특징을 설명하고 있다는 장점이 있다. 그러나 CPSS는 형용사의 수가 너무 많고, 시간이 걸리는 제약이 있으며 미묘한 차이를 구별하기 어려운 비슷한 의미의 양극 형용사들이 반복되어 평가자들에게 지루함과 피로감을 줄 수도 있다(박선희, 2002; 최인수, 2000)는 한계점이 논의된 바 있다. 이에 평가 차원은 동일하나 상대적으로 간명한 CPAM으로 평가를 수행할 수도 있다. 따라서, 본 연구에서는 CPAM과 CPSS 평정 척도를 사용하여 평가 결과의 차이를 분석하였다.

가. 창의적 산출물 분석 행렬(CPAM)

O'Quin & Besemer(1989)가 개발한 CPAM에 대한 내용을 번안 후 전문가 3인에게 내용타당도를 의뢰하여 수정 작업을 거쳤다. 다시 창의성 전문가의 도움을 받아 수정·보완하였으며, 교사 5인을 대상으로 예비조사를 거쳐 이해 가능 여부를 파악하였다. 또한 이 과정에서 평가 소요시간을 측정하였으며 평가의 지침 등 세부사항들을 수정·보완하였다. CPAM에는 '새로움(novelty), 해결성(resolution), 정교성 및 종합성(elaboration & synthesis)'이라는 3개의 차원과 11개의 하위 요소들이 포함되어 있다(<표 2>참조).

<표 2> CPAM의 평가 내용

차원	하위 요소	평가 내용	보통 이하		보통		보통 이상	
			1	2	3	4	5	6
새로움	독창성(original)	비슷한 경험과 훈련을 받은 다른 사람들은 만들 수 없을 정도인가? 흔치않게 독특하고, 드물게 관찰되는가?						
	놀라움(surprising)	기발하거나 신기해서 놀라움을 주는가?						
	발전가능성(germinal)	이후 미래의 창의적 산출물을 제작하는 데 영향을 줄 수 있을 만한 것인가?						
해결성	가치로움(valuable)	사용자의 재정적, 물리적, 사회적, 심리적 요구를 충족시켜 가치롭다고 여겨지겠는가?						
	논리성(logical)	산출물이 해당하는 그 분야에서 수용되고 인정되는 규칙을 따르는가?						
	유용성(useful)	실제로 적용되는가? 쓸 만한가? 작동가능한가?						
정교성 및 종합성	유기적 조직성(organic)	전체적으로 총체성 또는 완결성이 있는가?						
	우아함(elegant)	해결책이 군더더기 없이 세련된 방식으로 잘 표현되었는가?						
	복합성(complex)	하나 이상의 다양한 요소 또는 수준들을 포함하는가?						
	이해가능성(understandable)	산출물은 분명하고 이해하기 쉬운가?						
	완성도(well-crafted)	정성, 열성적인 노력을 기울여 그 시점에서 가능한 가장 높은 수준으로 개발되었는가?						

이 외에도 추가적으로 응답자의 ‘교육경력, 영재교육 연수 이수’ 등 일반 현황 4개 문항과 산출물을 처음 접했을 때 느껴진 전반적인 창의성에 대한 첫 인상을 평가하는 문항과 CPAM 평가 후 산출물의 전반적인 창의성에 대해 최종 평가하는 문항을 각각 처음과 마지막에 제시하였다. 또한 서술형문항을 통해 첫인상평가와 달리 최종평가가 달라진 경우 또는 CPAM 평가 결과와 CPSS 평가 결과가 달라진 경우 그 이유를 응답하도록 하였다. 평가자가 창의적 산출물 평가 시 평가 기준을 고르게 반영하는지, 또는 평가 기준에 대한 체득과정 없이 개인적인 평가 기준에만 영향을 받는지에 대한 정보를 얻고자 하였다. 문항 구성 내용은 <표 3>과 같다.

<표 3> CPAM의 평가 문항 구성

대영역	문항내용	문항 수	문항유형
A. 응답자 일반현황	재직경력	1문항	선다형
	영재교육 연수 이수 여부	1문항	
	영재교육 담당 여부	1문항	
	영재교육 담당 분야	1문항	
B. 산출물의 창의성평가	암묵적 지식 기반 (첫인상 평가)	1문항	6점 리커트 척도
	새로움 차원	3문항	
	해결성 차원	3문항	
	정교성 및 종합 차원	5문항	
	명시적 지식 기반 (최종평가)	1문항	
	첫인상평가와 최종평가가 일치하는 않은 이유	1문항	
합계		19문항	

CPAM의 평가 도구 검사가 지니는 신뢰도를 분석해 본 결과 <표 4>와 같이 산출물 1, 산출물 2, 산출물 3의 Cronbach α 값은 각각 .942, .955, .950으로 매우 높게 나타났다.

<표 4> CPAM 평가 도구 신뢰도 분석 결과

차원	산출물1	산출물2	산출물3	문항 수
새로움	.846	.907	.898	3
해결성	.890	.864	.855	3
정교성과종합성	.903	.909	.902	5
전체	.942	.955	.950	11

나. 창의적 산출물 어의 척도(CPSS)

O'Quin과 Besemer(1989)의 CPSS는 평정자내 신뢰도 계수가 대부분 .60이상으로 보고되었 다. 또한 Besemer(1998)는 EQS 프로그램을 이용한 확인적 요인분석을 통해 2차원과 3차원 중 더 적합한 구조를 검증하였으며 2차원의 구조는 CFI 지수가 .90을 넘어서지 못했으나 3차원의 구조는 .90 이상의 적합도 지수를 밝혀 3차원 구조의 적합성을 검증하였다. CPSS 평가 도구의 55개 형용사 문항을 번역 시 단순한 영문 번역은 우리 문화적 맥락에서 의미가 변질될

수 있어 유의어와 반의어를 함께 검토하였으며, CPAM 평가와 마찬가지로 동일한 전문가 3인에게 내용 타당도를 의뢰하여 수정 작업을 거쳤다. 다시 창의성 전문가의 도움을 받아 적절한 어휘로 수정·보완 작업을 거쳤다.

<표 5> CPSS의 평가 문항 구성

구분	문항내용	문항 수	문항유형
A. 산출물의 창의성 평가	새로움 차원	15문항	의미변별척도 (7점 척도)
	해결성 차원	15문항	
	정교성 및 종합 차원	25문항	
	명시적 기반 평가(산출물에 대한 최종평가)	1문항	6점 리커트 척도
	CPAM 평가 시의 순위와 달라진 경우, 그 이유	1문항	서술형
계		57문항	

CPSS 평가 도구의 검사 신뢰도를 분석해본 결과 <표 6>과 같이 Cronbach α 계수가 대체적으로 매우 높게 나타났다.

<표 6> CPSS 평가 도구 신뢰도 분석 결과

차원	하위 요소	산출물1	산출물2	산출물3	문항 수
새로움	독창성	.807	.884	.899	5
	놀라움	.879	.919	.892	5
	발전가능성	.819	.886	.890	5
	전체	.935	.956	.954	15
해결성	가치로움	.848	.912	.886	5
	논리성	.852	.872	.848	5
	유용성	.843	.790	.865	5
	전체	.929	.922	.934	15
정교성과 종합성	유기적조직성	.875	.850	.892	5
	우아함	.888	.874	.883	5
	복합성	.751	.786	.806	5
	이해가능성	.883	.852	.888	5
완성도	완성도	.878	.878	.867	5
	전체	.952	.947	.951	25

3. 연구 절차

본 연구는 산출물의 창의성을 평가하기 위해 영재교육 담당 교사를 대상으로 3개의 산출물에 대해 CPAM 평가를 실시하고 일주일 후 CPSS 평가를 실시하게 하였다. 평가자에게는 산출물 A, B, C에 대한 제작 배경, 주제 및 제작 과정 등에 대한 안내서 각 1부, CPAM 평가지, CPSS 평가지가 제공되었다. 오프라인 설문지를 배부하여 직접 회수하는 방식으로 진행되었다. 설문조사는 2015. 7. 10.~7. 30.까지 이루어졌으며, 설문 응답은 평균적으로 1인당 30~40분 정도가 소요되었다.

가. 창의적 산출물의 선정

인간의 창의성이 결과물로 구현된 발명 활동을 통해서 인류의 삶은 더욱 편리하고 풍요로워졌다. 이러한 발명의 속성과 관련된 중요한 다섯 가지 차원은 ‘새로움(novelty), 유용성(utility), 비용 효과성(cost-effectiveness), 시장에 대한 영향력(impact on the marketplace), 특허 획득의 기회(opportunity for patent acquisition)’라고 볼 수 있다(Henderson, 2004). ‘새로움’은 기존의 발명품 또는 기술과 구별 지을 수 있는 속성이라 할 수 있다. 이는 CPAM에서의 ‘새로움(novelty)’ 차원과 관련이 깊다. ‘유용성’이란 발명의 실행 가능성, 비용 면에서의 효율성, 사회에서의 필요성 등으로 정의할 수 있다. 이는 CPAM에서 ‘해결성(resolution)’ 차원 하위 ‘유용성(useful)’에 대한 내용과 일맥상통한다. ‘비용 효과성’과 ‘시장에 대한 영향력’에 대한 척도는 사람들이 그 발명품을 얼마나 많이, 얼마나 자주, 얼마나 오래 사용하는가의 여부로 판단할 수 있을 것이다. 발명의 영향력 또한 그 발명이 사람들의 생각, 느낌 혹은 행동을 변화시키는 정도에 달려 있다고 할 수 있다. 이는 ‘해결성(Resolution)’ 차원의 ‘유용성(Useful)’ 요소와 관련이 깊다.

마지막으로 ‘특허 획득’ 속성은 특허 획득 자체가 창의적인 가치가 있다는 것을 보증하는 것은 아니며 매우 독특한 특징이라고 볼 수 있다(Henderson, 2004).

이와 같이, 발명은 구체적인 산출물로 발현되는 대표적인 분야이고, 발명의 창의적 속성들이 CPAM 과 CPSS 평가 도구의 3개 차원 및 하위요소와도 밀접히 관련되기에 본 연구에서 산출물의 적절한 분야라고 판단하였다.

나. 전국학생과학발명품 경진대회 수상작 선정

과학 발명 활동을 통하여 창의력을 계발하고 과학에 대한 탐구심을 길러주기 위한 목적으로 현재까지 총 38회가 개최된 미래창조과학부와 동아일보가 주최하는 ‘전국학생과학발명품 경진대회’ 수상작 중에서 산출물을 선정하였다.

이 대회의 심사기준은 크게 4가지로 작품과 아이디어의 독창성과 문제해결 노력 등에 관한 ‘창의성·탐구성’, 작품이 일상생활에서의 실제 응용 가능성, 기존제품에 비해 개선·발전된 정도 등에 대한 ‘실용성’, 작품 제작과정의 노력 및 참여정도에 대한 ‘노력도’, 마지막으로 작품 제작의 경비 절감 및 경제적 파급 효과 등을 고려하여 ‘경제성’의 측면이다.

평가 대상 산출물은 총 2차에 걸쳐 선정하였으며, 1차는 본 연구자가 다음의 원리에 근거하여 선정하였다.




첫째, 발명품의 특성상 최신의 작품으로 하되 산출물에 대한 설명이 지면으로 구성될 것임을 고려하여 발명품 정보가 잘 드러나 있어야 한다.

둘째, 같은 해의 수상작 중에서 선정한다. 같은 해의 수상작에서 선정해야 평가자의 동질성을 확보할 수 있으며 해마다 상대평가가 이루어지므로 당해 연도 출품작의 상대적 수준에 따라 평가 등위가 변동될 수 있는 여지를 차단하였다.

셋째, 기능이 복잡하고 고도로 정교한 산출물의 경우는 비전문가에게 평가될 때 정교성 차이는 드러내기는 어렵기 때문에 제외한다.

넷째, 출품 분야는 제한하지 않는다. 분야별 특성이 영향을 미칠 수도 있으나, 수상작을 분석해 본 결과 ‘생활 과학’ 분야나 ‘학습 용품’, ‘과학 완구’ 분야는 특성의 차이가 그다지 뚜렷하지 않았기 때문이다. 단, ‘자원 재활용 분야’의 경우 버려진 자원을 다시 활용하기 때문에 완성도와 성능에서 한계가 있기에 제외하였다.

등위별 상, 중, 하 작품을 5개씩 선정하여, 비전문가 교사 12명, 전문가 교사 5명에게 순위를 평정하게 하여 다음과 같이 최종 작품 3개를 선정하였다.

상 수준(국무총리상)	중 수준(은상)	하 수준(동상)
		
숨 쉬는 과일채소 저장박스	보쌈 캐리어	접고 접는 휴대용 쓰레기통

[그림1] 평가 대상 산출물

1) 산출물 1

산출물 1 ‘숨 쉬는 과일채소 저장박스’는 2013년도에 국무총리상을 수상한 ‘생활과학’ 부문의 작품으로 <상>수준의 대표 작품으로 선정하였다. 대통령상과 국무총리상은 부문에 상관없이 1개의 작품만을 시상하므로 충분한 가치를 인정받은 작품이라고 볼 수 있다.

본 발명품은 공기 순환이 잘 되어 과일이나 채소를 장기간 보관할 수 있는 콘테나에 관한 것이다. 기존의 콘테나에 보관되는 과일은 구멍이 뚫린 바깥쪽에 비해 중심부에는 공기 순환이 더디고 상부에 쌓이는 콘테나의 무게를 받기 때문에 중심부에 있는 과일은 쉽게 부패하였다. 이를 해소하기 위해 구멍이 뚫린 두 개의 스포츠 콘을 설치한 결과, 콘이 설치된 것은 공기 순환이 잘되고, 상부의 무게를 분산시켜 기존 콘테나에 비해 부패율이 약 5배나 낮아졌다. 온·습도를 조절하는 창고를 짓는 비용보다 훨씬 적은 비용으로 생산이 가능하므로 경제적이다. 내구성이 뛰어나다는 장점이 있다.

2) 산출물 2

‘접고 접는 휴대용 쓰레기통’은 2013년도에 동상을 수상한 ‘생활과학’ 부문의 작품으로 <하>수준의 대표 작품으로 선정하였으며 본 연구에서는 ‘산출물 2’로 설정하였다.

야외 활동에서 쓰레기통이 없을 때의 불편함을 극복하기 위해 접어서 보관하고, 펴서 쓰레

기통으로 사용할 수 있는 접이식 쓰레기통을 개발하였다. 바닥면을 열십자 모양으로 회전시킨 후, 경첩을 이용해 바닥면에 수직으로 세우면 쓰레기통이 완성된다.

3) 산출물 3

산출물 3 ‘보쌈캐리어’는 2013년도에 은상을 수상한 ‘생활과학’ 부문의 작품으로 <중>수준의 대표 작품으로 선정하였다. 본 발명품은 로프에 바퀴가 달려 물건을 쉽게 이동시킬 수 있는 휴대용 로프식 캐리어에 관한 것이다. 마트에서 물건을 구입하면 카트에서 박스로 옮겨 담아 차량까지 카트로 이동하거나 들고 옮기게 되는 번거로운 과정을 줄일 수 있도록 하였다. 아크릴 소재의 받침대에 경첩을 연결하여 고정시키고 바퀴를 달고 바퀴와 연결된 줄을 버클로 고정한다. 박스 사면의 각에 맞게 경첩 바퀴를 끼고 줄을 버클로 고정한 후 잡고 이동할 수 있다. 이 발명품을 활용하여 빠른 시간에 포장이 가능하며, 이동 시에도 별도의 캐리어를 사용하지 않고도 이동이 가능하며 휴대와 보관이 간편하다는 장점이 있다.

IV. 연구 결과

1. 산출물 평가에서 암묵적·명시적 지식 근거 평정 결과의 차이

암묵적 지식과 명시적 지식에 근거했을 때 영재교육 담당 교사의 산출물에 대한 평정 결과가 어떻게 달라지는지를 알아보기 위해 3개의 산출물 각각에 대해 반복측정분석을 실시했다. 먼저 암묵적 지식 기반 평가는 CPAM 평가 문항을 접하기 전에 가장 먼저 실시한 산출물에 대한 첫인상을 6점 척도로 평가하게 하였다. 다음은 CPAM에서 3가지 하위 차원에 대해서 응답을 한 후에 산출물의 전반적인 창의성에 대한 최종 평가를 6점 척도로 평가하게 하였다. 다시 1주일 후 CPAM 보다 상대적으로 세분화되어 있는 CPSS 평가 문항에 응답하게 한 후에 산출물의 전반적인 창의성에 대해 최종평가를 하도록 하였다. 이러한 과정을 3개의 산출물에 대해 수행하고 그 결과를 분석하였다. 3개의 산출물에 대한 평균과 표준편차를 제시하면 다음과 같다.

<표 7> 암묵적 지식에 근거한 평가와 명시적 지식에 근거한 평가($n=139$)

산출물 종류	암묵적 지식 기반 평가(첫인상 평가)		CPAM 명시적 지식 기반 평가(1차 평가)		CPSS 명시적 지식 기반 평가 (2차 평가)	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
산출물1	4.91	1.032	4.77	.980	4.65	.907
산출물2	4.32	1.044	4.21	1.086	4.19	1.011
산출물3	4.35	.946	4.35	.961	4.19	.970
합계	4.53	1.007	4.44	.982	4.34	.962

먼저 산출물1에 대한 교사들의 평점 점수를 보면, 산출물의 첫인상을 나타내는 암묵적 지식 기반평가($M=4.91, SD=1.032$)가 가장 높고, CPAM 평가를 경험한 후에 한 1차 평가

($M=4.77$, $SD=.980$)점수가 중간, 그리고 일주일 뒤에 실시한 CPSS 평가 후 실시한 2차 평가 ($M=4.65$, $SD=.907$)가 가장 낮게 나타났다. 산출물2와 산출물3에 대해서도 같은 결과가 도출되었는데, 전반적으로 평가가 거듭될수록 평가점수가 낮아지는 경향을 보이고 있는데, 첫인상으로 판단하는 암묵적 지식 기반 평가 보다 구체적인 평가 기준을 경험하고 나서 실시한 명시적 지식 기반 평가에서 더 엄격해 지는 경향을 보이고 있었다. 그리고 본 연구에서 <상>수준으로 보고 있는 산출물1에 대해서 영재교육 교사의 평가점수가 가장 높았으나, <하>수준인 산출물2에 대해서는 교사들 간의 편차가 컸으며, <중>수준인 산출물3은 첫인상 평가와 1차 평가간의 점수 차이가 없음을 알 수 있었다.

다음은 3종류의 산출물에 대해서 3번의 평가결과에 대해서는 평가자내 평정결과의 유의한 차이가 있는지를 통계적으로 검증하기 위하여 반복측정 일원분산분석을 실시하였다(<표 8> 참조).

<표 8> 암묵적(첫인상), 명시적 지식(1차 평가, 2차 평가)에 근거한 평가 결과의 차이

산출물	Source	SS	df	MS	F	p
산출물1	평가자 간	331.703	138	2.404		
	평가자 내	66.000	278			
	암묵적·명시적 지식	4.681	1.725	2.714	10.535	.000
	오차	61.319	238.038	.258		
	Total	397.703	416			
산출물2	평가자 간	388.686	138	3.817		
	평가자 내	67.333	278			
	암묵적·명시적 지식	1.501	1.721	.872	3.147	.052
	오차	65.832	237.479	.277		
	Total	456.019	416			
산출물3	평가자 간	320.719	138	2.324		
	평가자 내	62.000	278			
	암묵적·명시적 지식	2.115	1.787	1.184	4.874	.011
	오차	59.885	246.580	.243		
	Total	382.719	416			

먼저 산출물1에 대해서 평정자 내 평정결과의 차이를 분석하면, 3번의 평정결과에서 통계적으로 유의한 차이가 확인되었다($F=10.535$, $p<.001$). 따라서 산출물1은 평정결과가 다르게 나타났음을 의미한다. 산출물2는 유의수준 .05에서 통계적으로 유의한 차이가 나타나지 않았으나 유의확률이 .052로 근사한 차이가 있음을 알 수 있었고, 산출물3은 평정결과가 다르게 나타남을 확인할 수 있었다($F=4.874$, $p<.05$). 산출물1은 국무총리상을 받은 <상>수준의 산출물이긴 하지만 암묵적 지식 기반 평가와 명시적 지식 기반 평가 간에 가장 뚜렷한 점수 차이가 나타난 산출물이기도 하다.

2. 산출물 평가에서 활용하는 평정 도구의 평정척도에 따른 평정결과의 차이

가. CPAM과 CPSS의 평가 차원별 평균 차이 검증

창의적 산출물 평가에 있어서 평정 도구 간, 차원별 평가 결과는 매우 중요하다. 각기 다른 산출물에 대해서 어떤 차원에 높은 점수를 부여했는지, 그리고 같은 차원이라도 평정척도와 척도의 표현 형식이 달라짐에 따라 평정결과의 차이가 어떻게 달라지는지를 분석하고자 했다. 본 연구에서 산출물을 1차 평가한 CPAM 평가 도구는 6점 척도이고, 2차 평가한 CPSS 평가 도구는 7점 척도로 2개의 도구의 평정 척도가 달라서 100점 만점의 환산점수를 사용하였다. 즉, CPAM의 원점수 $\times \frac{100}{6}$ 그리고 CPSS의 원점수 $\times \frac{100}{7}$ 으로 계산하였다. 각 차원별 평균점수에 대해서 대응표본 t검정을 실시하였다. 산출물별, 차원별 CPAM과 CPSS 평균과 표준편차는 <표 9>와 같다.

<표 9> CPAM 평가 도구와 CPSS 평가 도구의 차원별 기술통계

차원	평가 도구	산출물1		산출물2		산출물3	
		M	SD	M	SD	M	SD
새로움	CPAM	74.02	16.027	67.79	16.709	71.58	15.528
	CPSS	71.59	14.046	63.95	15.270	66.05	14.812
해결성	CPAM	80.38	16.097	72.26	15.534	70.90	14.134
	CPSS	76.32	13.624	71.63	12.496	70.44	13.113
정교성과	CPAM	77.36	14.190	71.62	14.161	71.82	13.250
종합성	CPSS	71.98	12.572	68.96	10.949	67.93	11.829
합계	CPAM	77.08	15.432	70.38	15.468	71.43	14.303
	CPSS	73.39	13.413	68.16	12.736	67.81	13.251

세 개의 모든 산출물들에 대한 평가가 CPAM도구로 했을 때 CPSS보다 점수가 높게 나타났다. 특히 산출물1의 경우에는 다른 산출물에 비해서 모든 차원에서 점수가 높게 나타났다. 새로움 차원에서 CPAM(M=74.02), CPSS(M=71.59)로 높았으며, 해결성 차원에서도 CPAM(M=80.38), CPSS(M=76.32). 정교성과 종합성 차원에서도 CPAM(M=77.36), CPSS(M=71.98)로 나타났다. 산출물1은 국무총리상을 받은 작품으로 본 연구에서 산출물 선정 시 가장 <상>수준에 해당하는 작품이다.

산출물3은 <중> 수준에 해당하는 작품으로 새로움 차원의 점수가 산출물1 다음으로 높았는데, 새로움 차원에서 두 평가 도구인 CPAM(M=71.58)과 CPSS(M=66.05)의 평정 점수의 차이가 크게 나타났다. 산출물2는 새로움 차원에서 가장 점수가 낮게 나타났는데, 산출물 선정 시 <하> 수준에 해당하는 작품이다. 따라서 산출물의 새로움 차원의 점수가 산출물1(<상>수준), 산출물3(<중>수준), 산출물2(<하>수준)로 나타나, 발명품경진대회에서 수상을 결정한 전문가들의 의견과 일치하는 것으로 나타났다. 반면, 정교성과 종합성 차원에서 산출물2와 산출물3의 점수 차이는 미미하게 나타났다.

CPAM 평가도구와 CPSS 평가도구의 산출물별 대응차이 및 상관을 분석한 결과는 <표 10>과 같다.

<표 10> CPAM 평가 도구와 CPSS 평가 도구의 산출물별 대응차이 및 상관($n=139$)

산출물차원	t-paired					correlation		
	<i>M</i>	<i>S.E</i>	<i>t</i>	<i>df</i>	<i>p</i>	<i>r</i>	<i>p</i>	
산출물1	새로움 CPAM-CPSS	2.43	1.148	2.115	138	.036	.602	.000
	해결성 CPAM-CPSS	4.06	1.071	3.786	138	.000	.650	.000
	정교성과종합성 CPAM-CPSS	5.38	0.942	5.712	138	.000	.661	.000
산출물2	새로움 CPAM-CPSS	3.83	1.100	3.483	138	.001	.674	.000
	해결성 CPAM-CPSS	0.63	1.099	.578	138	.564	.592	.000
	정교성과종합성 CPAM-CPSS	2.66	0.961	2.765	137	.006	.622	.000
산출물3	새로움 CPAM-CPSS	5.53	1.214	4.558	138	.000	.556	.000
	해결성 CPAM-CPSS	0.47	0.988	.474	138	.636	.637	.000
	정교성과종합성 CPAM-CPSS	3.89	0.979	3.977	138	.000	.582	.000

창의성 차원은 같으나 점수를 평정 하는 방식이 다른 CPAM과 CPSS 도구의 점수 차이를 분석한 결과, 산출물1에서는 새로움($t=2.115, p<.05$), 해결성($t=3.786, p<.001$), 정교성과 종합성($t=5.712, p<.001$)에서 평가 도구별 차이가 있었으며, CPAM이 CPSS보다 평균점수가 높게 나타났다.

산출물2에서는 새로움($t=3.483, p<.01$), 정교성과 종합성($t=2.765, p<.01$)차원에서 유의한 차이가 나타났으나, 해결성 차원은 유의미한 차이가 나타나지 않았다. CPAM이 CPSS보다 평균 점수가 높게 나타났다. 상관은 해결성 차원은 CAPM과 CPSS간의 상관이 있기는 하나 높지 않았다. 산출물3에서는 새로움($t=4.558, p<.001$), 정교성과 종합성($t=3.977, p<.001$)에서 평가 도구별 차이가 있었으며, CPAM보다는 CPSS의 평균점수가 높게 나타났다. CPAM이 CPSS보다 평균점수가 높게 나타났다. CAPM과 CPSS간 상관은 모든 차원에서 $r=.556\sim.674$ 로 나타나 상관이 높은 편이었다.

평정자의 평가결과는 평가도구와 평정척도에 따라 점수가 다소 변화되는 것을 알 수 있었다. 즉, 같은 산출물에 대한 평가라 하더라도 평가가 진행되면 평정점수는 낮아지게 되고, 상대적으로 문항 수가 적은 CPAM 평가도구의 경우 주관이 개입될 여지가 있으며, 일부 측면만이 고려될 수 있다. 평가자들은 CPSS 평가의 많은 문항 수와 촘촘한 채점 기준을 접하면서

CPAM 평가 시 미처 인지하지 못했던 평가 기준에 대해 숙고하면서 더 낮은 점수를 부여했을 것이라 판단된다. 이러한 차이는 평정 도구의 차원별로도 차이가 나타나는데, 새로움 차원과 정교성과 종합성 차원은 평정도구의 문항이 세분화 될수록 점수가 낮아지는 경향을 보인 반면, 해결성 차원은 평정도구 간에 점수 차이가 거의 없는 것으로 나타났다. 산출물의 가치로움, 논리성, 유용성은 평정자가 비교적 일관되게 평정하는 창의성 영역인 것을 알 수 있었다.

나. 평가 결과가 달라진 이유

CPAM 평가와 CPSS 평가 말미에 자유응답형 문항을 추가하여 평가 결과가 달라진 이유를 기술하게 하였고, 이를 통해 영재교육 교사들이 평가의 세부적인 과정에 어떤 사고 과정이 일어나는지를 분석할 수 있었다.

첫 번째 유형은 창의적 산출물 평가 척도를 경험하면서 평가 결과가 이전에 비해 달라진 경우이다. 즉 평가 척도의 경험이 평가 결과에 영향을 미친 경우이다. 이 때, 평가 기준에 따라 고르고 균등하게 평가를 해야 함에도 불구하고 1-2개의 차원에만 집중하여 평가하는 사례들이 있었다.

캐리어를 사용하지 않고 바퀴와 연결 끈만을 활용하여 물체를 이동시킬 수 있는 발명품이라는 생각이 처음에는 다소 창의적이라고 생각하였지만, 실제로 활용하기(유용성)에서 무리가 있을 것 같은 생각이 들었다.

부피를 줄여 휴대하기 좋은 점은 크게 와 달았으나 실제 사용할 경우, 박스와 캐리어의 분리에 대한 염려, 길의 상태에 따른 끌기의 힘겨움 등이 염려되어 완성도에 대한 결과가 낮아졌다.

처음에는 새로운 것을 평가 기준으로 삼았으나 실제 창의성이란 약간의 변형과 조합을 통해 새로운 사물을 만들어 내는 것으로 판단했기 때문이다.

실용성 여부가 가장 중요한 기준으로 작용하였다.

그리고 산출물에 가장 중요하게 생각하던 부분이 평가 척도를 통해서 변하게 된 경우도 있었다. 즉, 평가 척도를 경험하고 난 후에는 이전에 생각하지 못했던 평가 기준이 생기게 된 경우이다. 다음의 예가 해당된다.

처음에는 새로움이 조금 부족하여 낮은 점수를 주었지만 유용성 측면을 다시 생각해보니 차량에 싣고 다니면 꽤 유용할 것 같습니다.

두 번째는 평가척도의 영향을 받지 않은 경우이다. 이 경우에도 두 가지 측면에서 정리해 볼 수 있는데, 하나는 암묵적 판단(첫인상) 지배형이라 볼 수 있다. 처음부터 좋았거나 처음부터 마음에 들지 않았기 때문에 마지막 최종평가에서도 결과가 별로 달라지지 않았다는 것이다.

이 산출물은 처음부터 진부한 느낌이 들었기 때문에 최하위 성적을 주었다.

또 하나는 개인적 평가 기준이 너무 확고해서 평가척도의 문항들이 영향을 미치지 않은 경우이다. 응답의 예를 들면 다음과 같다.

학생들이 만든 작품이기에 복잡한 것보다는 단순하지만 의미 있는 것이 더 중요하다.

V. 결 론

본 연구는 창의적 산출물 평가의 대표적 도구인 O'Quin과 Besemer(1989)가 개발한 '창의적 산출물 분석행렬(CPAM)'과 '창의적 산출물 어의 척도(CPSS)'의 두 가지의 산출물 평정 척도를 적용하여 동일한 산출물을 평가하는 데 있어 평가 결과의 차이를 분석하였다. 본 연구의 연구문제에 대한 결론은 다음과 같다.

첫째, 창의적 산출물 평가에서 암묵적 지식은 존재하며, 암묵적 지식 기반 평가와 명시적 기반 평가의 평정 결과에는 차이가 있는 것으로 나타났다. 산출물에 대한 3번의 반복측정평가가 이루어졌으며, 산출물1과 산출물3에서 차이가 유의미하게 나타났으며, 산출물2는 유의한 차이는 아니지만 평균차이가 나타났다. 그리고 평가가 거듭되어 명시적 지식이 쌓여갈수록 산출물에 대한 전반적인 평가 점수가 낮아지는 경향을 보이고 있었다. 특히 영재교육 교사들은 산출물1에서 대해서 가장 높은 창의성 점수를 부여하고 있는데, 산출물1은 국무총리상을 받은 작품으로 본 연구에서 가장 수준이 높은 산출물로 선정된 것이다. 따라서 영재교육 교사의 산출물에 대한 평가관점은 발명산출물 대회에서 심사위원으로 참여한 전문가들의 평가 결과와 일치한다고 볼 수 있다.

둘째, 창의적 산출물 평가에서 활용하는 평정 척도에 따른 평정 결과에도 차이가 있는 것으로 나타났다. CPAM과 CPSS로 평정했을 때 창의성의 새로움 차원, 해결성 차원, 정교성과 종합성 차원에서 모두 CPAM 평가 결과 점수가 높게 부여되었다. 상대적으로 동일 차원 내에서도 문항 수가 적은 CPAM 평가 도구의 경우 일부 측면만이 고려될 수 있다. 영재교육 교사들이 CPSS 평가와 같이 세분화되고 세밀한 채점 기준을 통해 평가하였을 때 그 전에 미처 인식하지 못했던 창의성 요소에 더 깊이 지각을 했을 수 있다. 또한 CPSS와 같이 양극단의 형용사로 이루어진 어의 척도의 경우, 부정 어미를 붙여서 표현된 명시적 부정문항(explicit negation items)과 의미적으로 반대의 의미를 가지는 암묵적 부정문항(implicit negation items)으로 구분될 수 있는데 본 연구의 CPSS에서는 암묵적 부정문항을 주로 사용하였다. 명시적으로 부정의 형태가 된 것보다 이와 같이 암묵적으로 부정의 형태가 사용된 것이 평가자들이 창의성의 각 차원에 대해서 부정적인 쪽으로 평가하게 한 이유가 될 수도 있다. 또한 차원별로 뚜렷한 차이를 나타내는데, 평정자가 새로움과 정교성과 종합성 차원은 평가 도구에 따라 점수를 다르게 평정하였지만, 해결성 차원에서는 평정자 간에 비교적 일치하는 결과를 보였다.

마지막으로 CPAM 평가와 CPSS 평가에서는 서술형 문항을 통해 영재교육 교사들이 평가

의 세부적인 과정에 어떤 사고 과정이 일어나는지를 분석할 수 있었다. 첫 번째는 평가척도를 전혀 경험하지 않은 첫인상에서의 평가와 창의성 평가를 경험하면서 최종평가의 일치도가 달라지거나, CPAM과 CPSS의 서로 다른 척도를 경험하면서 산출물에 대한 평가가 변한 경우이다. 이 경우에는 평가의 차원을 고루 고려하지 않고 1-2가지의 기준에만 집중을 한 경우도 있고, 창의성의 차원을 보면서 중시하는 기준이 변한 경우도 있었다. 두 번째는 평가척도의 영향을 받지 않은 경우이다. 이 경우에도 두 가지 측면에서 정리해 볼 수 있는데, 하나는 암묵적 판단(첫인상) 지배형이라 볼 수 있다. 처음부터 좋았거나 처음부터 마음에 들지 않았기 때문에 마지막 최종평가에서도 결과가 별로 달라지지 않았다는 것이다. 또 하나는 개인적 평가 기준이 너무 확고해서 평가척도의 문항들이 영향을 미치지 않은 경우 등으로 나타났다. 이는 평가결과가 평가자의 엄격성이나 관대성이 개인의 배경(체점 훈련, 전공, 성격)과도 관련이 있다는 것과도 일치한다.

결론적으로 본 연구에 참여한 영재교육 교사들은 그간 영재학급이나 영재교육원등에서 강사로 활동하고 있고, 영재교육 교원연수를 이수하였으며, 영재 학생들의 산출물을 지도하거나 평가한 경험이 있는 집단이다. 따라서 이미 영재교육과 창의성에 대한 암묵적 지식이 일반교사보다는 조직화된 상태라 볼 수 있다. 본 연구 결과에서도 알 수 있듯이 창의적 산출물 평가에서 평가 전문가들의 평가와 유사한 평가를 하고 있다. 따라서 향후 영재학생들의 산출물 평가에서도 채점자의 신뢰도를 고려하지 않아도 신뢰로운 판단을 한다고 볼 수 있다.

본 연구의 결과에서 알 수 있듯이 평가의 기준과 평가척도가 달라지면 동일한 산출물에 대한 평정 점수에서 약간의 변화가 있을 수 있으나, 평정결과가 크게 변화되지는 않는다는 사실을 발견했다. 따라서 시간이나 비용 등의 측면을 고려하면 종합적 채점 기법에 의한 산출물에 대한 평정도 무리가 없으나 영재교육 담당 교사들이 분석적 채점 기법을 충분히 숙지하고 산출물 평가를 지도하거나 평가활동에 참여한다면 학생들이 보다 새롭고, 유용하고 적절하며, 해결성이 높은 창의적 산출물을 생성할 수 있을 것이다. 본 연구에서 알 수 있듯, 산출물의 창의성을 평가하는데 있어서 창의성의 유무나 그 정도에 대해 판단하는 평가자의 문제는 매우 중요하기에 우리나라 영재교육을 담당하는 중등교원을 대상으로 하는 연구 등이 폭넓게 진행될 필요가 있다.

참 고 문 헌

교육부 홈페이지. 제2차영재교육진흥종합계획. <http://www.moe.go.kr> (검색일: 2016. 11. 4)

김아영, 김세영 (2003). 명시적 부정문항과 암묵적 부정문항이 심리척도의 요인구조에 미치는 영향. *교육평가연구*, 16(1), 39-52.

김영록, 이순복 (2004). 한국판 창의적 산물 평가도구의 개발 및 구조확인. *한국심리학회지: 산업 및 조직*, 17(3), 305-327.

김정섭 (2008). 한국 창의성 연구의 동향과 쟁점. *한국교육심리학회*, 22(4), 939-960.

박선희 (2002). 아동의 창의적 산물 평가 척도의 타당화 연구. 석사학위논문. 성균관대학교.

- 조연순, 정지은 (2012). 국내 창의성 교육 연구 동향분석: 창의성의 범주 및 수준을 중심으로. **영재교육연구**, 22(2), 333-352.
- 최인수 (2000). 유아의 창의성 측정도구에 관한 고찰. **유아교육연구**, 20(2), 139-166.
- 최인수, 임미정, 표정민 (2011). 한국·중국·일본의 창의성에 대한 암묵적 지식 비교. **창의력교육연구**, 11(2), 27-47.
- 한기순 (2005). 창의성 영역문제의 탐색 및 재접근. **영재교육연구**, 15(2), 1-34.
- Amabile, T. M. (1982). Social psychology of creativity: A consensual assessment technique. *Journal of Personality and Social Psychology*, 43(5), 997-1013.
- Amabile, T. M. (1983). *The social psychology of creativity*. New York: Springer-Verlag.
- Amabile, T. M. (1996). *Creativity in context: Update to the social psychology of creativity*. New York: Westview Press.
- Besemer, S. P. (1998). Creative Product Analysis Matrix: Testing the Model Structure and a Comparison Among Products—Three Novel Chairs. *Creativity Research Journal*, 11(4), 333-346.
- Besemer, S. P., & O'Quin, K. (1986). Analyzing Creative Products: Refinement and Test of a Judging Instrument. *The Journal of Creative Behavior*, 20(2), 115-126.
- Besemer, S. P., & O'Quin, K. (1987). Creative Product Analysis: Testing a model by Developing a Judging Instrument. In S. G. Isaksen (Ed.), *Frontiers of creativity research: Beyond the basics* (pp. 341-357). Buffalo, nY: Bearly Ltd.
- Besemer, S. P., & O'Quin, K. (1999). Confirming the Three-Factor Creative Product Analysis Matrix Model in an American Sample. *Creativity Research Journal*, 12(4), 287-296.
- Besemer, S. P., & Treffinger, D. J. (1981). Analysis of Creative Products: Review and Synthesis. *The Journal of Creative Behavior*, 15(3), 158-178.
- Bruner, J. (1962). The conditions of creativity. In J. Bruner (Ed.), *Beyond the information given: Studies in the psychology of knowing* (pp. 208-217). New York: Norton.
- Henderson, S. J. (2004). Product inventors and creativity: The finer dimensions of enjoyment. *Creativity Research Journal*, 16(4), 293-312.
- Horn, D. B., & Salvendy, G. (2006). Consumer-based assessment of product creativity: A review and reappraisal. *Human Factors and Ergonomics in Manufacturing & Service Industries*, 16(2), 155-175.
- Jackson, P. W., & Messick, S. (1967). The person, the product, and the response: Conceptual problem in the assessment of creativity. In J. kagang (Ed.). *Creativity and Learning* (pp.1-19). Boston: Houghton Mifflin.
- Mackinnon, D. W. (1987). Some critical issues for future research in creativity. In S. G. Isaksen (Ed.), *Frontiers of creativity research* (pp.120-130). Buffalo, NY: Bearly Limited
- O'Quin, K., & Besemer, S. P. (1989). The development, reliability, and validity of the revised creative product semantic scale. *Creativity Research Journal*, 2(4), 267-278.

- Reis, S. M., & Renzulli, J. S. (1991). The Assessment of Creative Products in Programs for Gifted and Talented Students. *Gifted Child Quarterly*, 35(3), 128-134.
- Renzulli, J. S., Leppin, J. H & Hays, T. S. (2000). *The Multiple Menu Model: A Practical Guide for Developing Differentiated Curriculum*. Mansfield Center, CT: Creative Learning Press.
- Rhodes, M. (1987). An analysis of creativity. In S. G. Isaksen (Ed.), *Frontiers of Creativity Research* (pp.216-222). Buffalo, NY: Beariy Limited.
- Romo, M., & Alfonso, V. (2003). Implicit Theories of Spanish Painters. *Creativity Research Journal*, 15(4), 409-415.
- Runco, M. A., Johnson, D., & Baer, P. (1993). Parents' and teachers' implicit theories of children's creativity. *Child Study Journal*, 23, 91-113.
- Sternberg, R. J. (1985). Implicit theories of intelligence, creativity, and wisdom. *Journal of Personality and Social Psychology*, 49(3), 607-627.
- Sternberg, R. J. (2004). **지혜, 지능 그리고 창의성의 종합** [김정희 역]. 서울: 시그마프레스. (원본출간년도: 2003).
- Sternberg, R. J. (2006). The Nature of Creativity. *Creativity Research Journal*, 18(1), 87-98.
- Sternberg, R. J. (2010). **지혜의 탄생** [최호영 역]. 서울 : 21세기북스. (원본전출간년도: 1990).
- Sternberg, R. J. (2012). The Assessment of Creativity: An Investment-Based Approach. *Creativity Research Journal*, 24(1), 3-12.
- Sternberg, R. J., Conway, B. E., Kerton, J. L., & Bernstein, M. (1981). People's conceptions of intelligence. *Journal of Personality and Social Psychology*, 41(1), 37-55.
- Treffinger, D. J., & Poggio, J. P. (1972). Needed Research on the Measurement of Creativity. *The Journal of Creative Behavior*, 6(4), 253-267.
- Wegner, D. M., & Vallacher, R. R. (1977). *Implicit psychology: An introduction to social cognition*. Oxford University Press.

= Abstract =

Consistency in Assessment of Creative Products in Terms of Evaluators' Knowledge of Creativity Assessment Criteria and the Type of Assessment Tools

Su Jin Lee

Wolyoung Elementary School

Ho Seong Choe

Kyungnam University

Kyung Hee Park

Kyungnam University

This study analyzed the difference in evaluation results in evaluating identical products by applying two different types of evaluating scales, Creative Product Analysis Matrix (CPAM) and Creative Product Semantic Scale (CPSS) by O'Quin and Bessemer (1989). As a result, evaluation based on explicit knowledge scored lower than evaluation based on implicit knowledge, implying that the evaluation becomes stricter. When evaluated with CPSS, which as relatively more segmentalized grading criteria, all sub-dimensions of creativity showed low scores, and it show that when evaluator's first impression or personal evaluation standard on the products is firm, they may not be evaluated by the evaluation tools. Gifted education teachers were giving similar evaluations as experts in creative product evaluation, and understanding the product evaluation tool fully in advance before teaching or evaluating products may lead to the generation of newer, more useful and appropriate, and highly creative product with high solvability.

Key Words: Consistency in Assessment, Evaluators' Knowledge, Creativity Assessment Criteria, Type of Assessment Tools, CPAM, CPSS

1차 원고접수: 2016년 11월 29일
수정원고접수: 2016년 12월 23일
최종게재결정: 2016년 12월 27일