

## 평가자 유형에 따른 창의성 프로그램이 대학생의 창의성 증진에 미치는 효과\*

정은이(鄭殷伊)\*\*

### 논문 요약

본 연구에서는 창의성 프로그램을 적용하여 대학생들에게 창의성 훈련을 받게 한 후 자기평가, 동료평가, 전문가 평가 등 각기 다른 평가자에 의한 창의성 평가를 실시하였다. 이후 Torrance의 도형 창의성 검사를 사용하여 창의성 증진 효과가 어떻게 다른지 알아보았다.

첫째, 자기평가 집단이 나머지 두 집단(동료평가, 전문가평가)보다 창의성 평균과 유창성, 독창성, 정교성 및 종결에 대한 저항에서 높은 점수를 나타냈고 이는 통계적으로 유의한 결과였다. Schéffe 검증 결과 집단간 차이도 유의했다. 이는 자기평가를 할 때 스스로 자신의 수행과정 및 결과를 돌아보고 성찰해 보는 시간 자체가 창의성을 발휘하는 하나의 과정이 되기 때문인 것으로 보인다. 이 과정에서 개인은 자신이 스스로 문제를 해결하고 개선하려는 자발성을 키우며 자신의 평가에 대한 자신감과 만족감을 느끼게 되는 것이다. 즉, 자기평가를 통해서 자발성과 독자적인 판단력 및 평가에 대한 만족 등을 얻게 되고 이것은 다시 창의적 능력에 영향을 미치게 되어 창의성 증진의 효과를 가져오게 되는 것이다.

둘째, 동료평가 집단과 전문가 평가 집단간에는 통계적으로 유의한 차이가 나타나지 않았다. 이러한 결과는 창의성 증진의 차원에서 보면 같은 또래가 평가를 하든 전문가가 평가를 하든 별 차이가 없다는 것이다.

■ 주요어 : 평가자 유형, 창의성 프로그램

\* 이 연구는 2003학년도 청운대학교 교내 연구 지원에 의해 수행되었음

\*\*청운대학교 교양학부

## I. 서론

### 1. 연구의 필요성 및 목적

새로운 시대에 걸맞는 창의적이고 자율적인 인간 육성이 우리 교육의 중요한 목표가 되고 있는 현실에서 창의성 교육은 그 어느 때보다 강조되고 있다. 또한 열린 교육에서는 학생의 창의성과 흥미 및 적성을 고려한 다양한 교수 방법 및 학습 프로그램들을 학교 현장에 활용하고 있다. 이러한 상황에 발맞춰 새로운 창의성 프로그램 개발 및 그 효과에 대해서 그동안 많은 연구가 진행되어 왔다. 특히 최근에 우리나라에서는 유치원, 초등학교 수준에서 창의적 능력 및 인성에 관한 여러 프로그램들이 개발되었다(최미정, 2000; 호사라, 2001; 정황순, 2001; 조성수, 1996; 장좌욱, 2002; 부지영, 2000).

그러나, 창의성 훈련 프로그램에 비해 적절한 평가 방안에 대해서는 아직 연구가 미흡하다. 그동안 전통적인 평가는 한줄 세우기의 관점에서 점수 매기기에 준하여 학생들을 성공과 실패로 나누기에 급급했다. 전평국 등(2001)은 이러한 전통적인 평가는 창의성 연구 및 이론과 적어도 2가지 점에서 모순된다고 했다. 첫째, 전통적인 시험과 평정은 창의적인 활동을 하려는 내적인 동기 및 창의적 결과의 질 모두를 저하시키는 것으로 판명된 평가 보상, 경쟁이라는 요소들의 전형이다. 둘째, 창의성과 연관된 결과들(문제 해결하기, 독창적인 아이디어 표현하기)은 하나의 평어로 나타내거나 선다형 시험을 통해 부각되지 않는다. 실제 창의적 사고가 선다형 시험을 방해할 수도 있다.

이러한 창의성에 관한 평가의 문제에 있어서 최근 평가관이 크게 변화하면서 대안적인 평가 방식의 일환으로 도입된 수행평가는 하나의 시사점이 될 수 있을 것이다.

백순근(1998)은 수행평가의 필요성을 세계화, 정보화 시대를 맞이하여 사고의 다양성과 창의성을 신장하고 조장하기 위한 것이라고 언급하면서 여러 측면의 지식이나 능력을 지속적으로 평가할 수 있고 특히 문제해결 능력, 비판적 사고 능력 평가에 적합하다고 하였다.

이러한 관점에서 볼 때 창의성의 평가에 있어서 수행평가 방법을 도입하는 것은 고등정신기능의 평가와 평가 방식의 다양화 측면에서 볼 때 바람직한 현상일 것이다. 창의성을 키우려 한다면 창의성을 신장에 초점을 두는 평가 방안이 마련되어야 할 것이다. 따라서 창의성을 평가할 때는 실제적인 평가(authentic assessment)의 차원에서 수행이나 산출물을 평가하기 위한 준거와 기준을 명료화하여야 할 것이다. 수행평가에서 수행, 산출물 평가 방법으로 체크리스트와 포트폴리오를 많이 사용하는데 이는 창의성 평가에도 유용한 하나의 방법이 될 수 있다. 또한 자기평가와 동료 평가 등 기존의 교사 또는 전문가 평가와는 다른 방식의 평가도 필요할

것이다. 이러한 새로운 방식의 평가는 창의성 등 고등정신능력의 평가에 적합한 모델이 될 수 있을 것이다.

그러므로 본 연구에서는 대학생들을 위한 창의성 프로그램을 적용하여 평가자 유형에 따른 차이를 알아보고자 한다. 특히, 창의성 산출물 평가를 하는 평가자에 따른 차이 즉, 자기평가, 동료평가, 전문가 평가를 적용하여 창의성 증진 효과가 어떻게 다른가 알아보고자 한다.

위에서 서술한 연구 목적을 수행하기 위하여 다음과 같은 연구 문제를 설정하였다.

: 평가자 유형(자기평가, 동료평가, 전문가평가)에 따른 창의성 평가 결과 창의성 평균 및 하위 요인의 점수는 차이가 있을 것인가?

## II. 이론적 배경

### 1. 대학생을 대상으로 한 창의성 프로그램의 효과

대학생들에게 일련의 창의성 프로그램을 적용하여 창의성을 개발하고자 했던 선행연구를 보면 유머나 브레인스토밍이나 시네틱스 기법과 같이 한가지 변인(Getzel, & Jackson, 1962; Treadwell, 1970; Ziv, 1976, 1988)을 투입하는 훈련에서부터 체계적으로 구성된 패키지 프로그램(deBono, 1971, 1975, 1978; Feldhusen, Treffinger, & Bahlke, 1970; Noller, Parnes, & Biondi, 1976)에 이르기까지 다양한 수준의 창의성 훈련을 실시했는데, 이러한 훈련이 효과적인 것으로 나타났다.

Khatena(1971)는 그가 개발한 프로그램을 대학생들에게 4일에 걸쳐서 240분의 창의성 훈련을 실시한 결과 2개의 독창성 검사 모두에서 유의한 효과가 있는 것으로 나타났고, 대학생들을 대상으로 한가지 기법에 의해 훈련을 시키거나 지침서에 따른 학습만으로도 긍정적 효과가 나타났다. 또한 Korth(1973)는 창의성 기법 중에 널리 사용되고 있는 시네틱스 훈련결과 과제를 서로 연합시켜 얻는 유창성(Associational Fluency)에서 효과가 있었고, Burns(1983)는 브레인스토밍 훈련 결과 융통성과 독창성에서 각각 유의한 향상을 보였고, Wey(1983)는 창의적 사고 모형에 관한 지침서를 통해 매주 숙제로 이를 수행하도록 하였을 때 Torrance의 창의성 검사에서 유의한 향상을 발견하였다.

우리나라에서 수행한 대학생의 창의성 증진을 위한 몇몇 연구들을 살펴보면, 박동준, 허경조(1988)는 대학생들을 대상으로 Torrance(1979)의 'The Search for Satori and Creativity' 책에 제시되어 있는 연습 내용과 Osborn(1963)의 'Applied Imagination'의 연습문제를 참고하여 창의성 프로그램을 개발하였고, 이 프로그램을 토대로 박동준, 임성문, 정영수(1989)가 대학생들

에게 프로그램을 적용한 결과 창의성 함양에 효과가 있는 것으로 나타났다. 전경원(1997)은 창의성 프로그램이 대학생의 창의성 계발에 미치는 효과를 연구하여 창의성 훈련을 받은 집단이 창의성이 유의미하게 향상되었고 수업시간에만 창의성을 강조해도 그 효과가 큰 것으로 나타났다. 정은이(2003)는 대학생을 대상으로 개발한 창의성 훈련 프로그램을 사용하여 개별·협동 학습 집단별 효과를 알아보았다. Torrance의 도형검사를 사용한 이 연구 결과 비교집단에 비해 훈련 집단이 유의미한 창의성 증진 효과를 보였고, 협동학습 집단이 개별학습 집단에 비해 창의성 총점과 유창성, 독창성에서 높은 점수를 나타냈다. 박병기(2004)는 교양강좌를 이용한 대학생 창의성 교육의 효과 분석을 한 결과 창의성 동기와 태도 측면에서 효과가 있었다고 하였다. 특히 개방성, 유희성, 호기심이 실험집단과 비교집단을 구분하는 데 기여한 변인이라고 하였다.

이상의 연구결과들을 통해 볼 때, 훈련 요인, 훈련 시간, 훈련 방법의 차이 및 효과에 대한 해석의 문제점에도 불구하고 대학생들의 창의성 훈련이 상당한 정도로 창의성을 증진시키는데 효과적이었음을 알 수 있다.

## 2. 평가자 유형(자기평가, 동료평가, 전문가평가)에 따른 창의성 증진 효과

창의성을 신장시키기 위한 평가 기법으로 평가자 유형을 기준으로 살펴보면 자기평가, 동료평가, 전문가 평가 등을 들 수 있다. 각 유형의 특징과 이러한 평가자에 따른 창의성 평가가 창의성 증진에 어떠한 영향을 미치는 지 알아보자.

### 1) 자기 평가

자기 평가 방식은 흔히 체크리스트를 학생들에게 준다. 그들이 과제의 각 단계를 완성해 감으로써, 그들은 거기에 해당하는 특별한 항목을 체크해 나간다. 이 체크리스트는 보다 발산적인 사고 및 새로운 아이디어의 생성을 지향하는 지침이 될 수도 있다. 동시에 체크리스트는 몇 가지 아이디어들로부터 혁신적인 아이디어를 선택하는데 필요한 수렴적 사고를 촉진시킨다. 이러한 체크리스트는 창의적인 산물을 내게끔 학생들을 유도해 주는 준거를 주어 그것을 따르면서 완수해가게 하기 위해 사용한다. 이런식으로 학생들로 하여금 융통성 있게 그리고 발산적으로 사고하며 혁신을 감수하게 조장해 준다.

특히 자기 평가는 동료 등 관찰자가 접근할 수 없는 내적인 요인들을 반영하여 자신의 장점과 단점에 대한 정확한 인식을 할 수 있도록 도와준다. 그러나, 자기 평가는 편파가 개입할 여지가 있어서 자신에게 유리한 쪽으로 평가를 하거나 사실을 축소하거나 과장하는 등 왜곡될

가능성이 많이 이에 대한 보완책이 요구된다.

## 2) 동료 평가

동료 평가는 과제를 함께 수행하는 동료에 의해 평가가 이루어지는 방식이다. 과제를 이해한 동료들은 새로운 연결을 이끌어 낼 수 있고 또 다른 학생의 진전을 주시하면서 새로운 아이디어를 자극할 수도 있다(진평국 외, 2001). 이 방법은 자기 평가와는 달리 주어진 평가 대상에 대한 다수의 평가치를 이용할 수 있으며 좀 더 객관적인 자료를 얻을 수 있다. 동료평가는 좀 더 정확한 정보를 제공할 가능성이 높으며 또래들끼리 경쟁의식과 협동의식을 동시에 불러 일으킬 수 있는 장점을 가지고 있다.

그러나, 동료평가는 자기평가와는 반대로 개인의 주관적인 정보와 특성을 반영하기 어렵고 고정관념에 의한 평가를 하기 쉽고 타당하지 않은 평가를 내릴 수도 있다.

## 3) 전문가 평가

전문가 평가는 창의성 영역과 이론에 입각한 창의성 전문가와 해당 영역 전문가가 산출물 및 과정을 평가하는 방법이다. 그동안 전통적인 창의성 평가는 과정과 산물 평가 모두 전문가들에 의해 이루어졌다. 학교 차원에서는 교사가 학생의 창의적 산출물과 태도 등에 관한 평가를 하고 이를 학생에게 통보하는 형식으로 이루어졌다. 전문가 평가는 창의성과 각 영역별 전문가가 평가를 하기 때문에 가장 정확하고 객관적이며 타당한 평가를 할 수 있는 장점이 있다. 그러나, 이를 통해 학생들의 내적인 동기 유발을 시키는 데는 한계가 있을 수 있다. 특히 창의성 등 고차원적인 정신능력의 평가에 있어서는 평가가 긍정적인 피드백 역할을 해야 하는데, 전문가 평가를 할 경우 학생들은 평가를 외적인 압력으로 생각할 수도 있다.

이러한 평가자 유형에 따른 창의성 평가 방식은 각자 독특한 특성을 가지고 있으나 창의성 증진 측면에서는 자기 평가 방식이 가장 큰 효과가 있을 것으로 생각된다. Amabile(1998)은 창의성을 방해하는 요소로 평가를 지적하면서 다른 사람의 평가에 지나치게 관심을 두면 자신의 작품 세계를 제대로 표현하지 못하고, 다른 사람이 원하는 방향으로 작품을 만들게 되어 창의성이 저하된다고 하였다. 아동을 대상으로 한 Amabile(1998)의 실험결과 그림에 대한 평가를 하지 않고 콜라주를 했던 집단이 아동이 중간에 그림에 대한 평가를 한 집단보다 더 창의적인 콜라주를 만든 것으로 나타났다. 이렇듯 자신의 작품에 대한 타인의 평가(동료나 전문가)는 창의성을 저해하는 요소로 나타날 수 있는 것이다. 또한 전경원(2000)도 창의성을 저해하는 요인을 조사한 연구에서 대학생들은 '타인의 시선 의식'을 창의성을 저해하는 가장 큰 요인으로 꼽

있다고 밝혔다. 즉 타인에 의해 이루어지는 평가 상황은 호기심, 자발성 등의 내적 동기 유발을 저해시키며 결과적으로 창의적인 표현 및 독창성의 발달을 저해시키는 요소가 되는 것이다. 동료평가의 경우 지나친 경쟁심, 전문가 평가의 경우 좋은 점수를 받으려는 욕구, 평가에 대한 지나친 의식 등이 창의성을 저해시킬 수 있는 것이다.

### III. 연구 방법

#### 1. 연구 대상

본 연구의 대상은 충청남도에 위치하고 있는 S대학교 학생들로서 교양과정으로 '창의성 개발'을 듣는 학생 109명이다. 이중 사전, 사후 검사 중 누락된 학생이나 성실하게 응답하지 않은 16명은 제외하였다. 총 94명의 학생 중 남학생이 45명, 여학생이 49명이었으며, 1학년 21명, 2학년 27명, 3학년 25명, 4학년 21명이었다. 전공별로는 인문·사회 계열이 30명, 이공 계열이 19명, 예술 계열이 24명, 경상계열이 21명이었다.

#### 2. 측정 도구

Torrance(1974)에 의해 제작된 Torrance 창의적 사고력 검사의 도형검사 TTCT(Torrance Tests of Creative Thinking) A, B형을 사용하였다. 검사지침서를 참고하여 검사 실시에 주의를 기울여 사용했고, 훈련내용과 효과측정치간의 유사성을 가능한 줄이기 위해 서로 다른 2개의 검사를 사용하여 연구 방법상의 문제를 감소시켰다.

실험집단 모두에게 사전검사로는 Torrance의 창의적 사고력 도형검사 A형을 실시. 사후검사로는 Torrance의 창의적 사고력 도형검사 B형을 실시하였다. 이 검사는 세가지의 하위 활동으로 이루어져 있다. : 활동 1은 그림 구성하기, 활동 2는 그림 완성하기, 활동 3은 선 더하기로 각각 10분씩 총 30분의 시간이 소요된다. 이 검사는 창의적 사고의 영역인 유창성, 독창성, 제목의 추상성, 정교성, 성급한 종결에 대한 저항 등을 측정한다.

TTCT 도형검사는 2003년 3월 5일과 6월 11일에 각각 사전검사와 사후 검사가 집단으로 실시되었고, 검사 요강에 제시된 대로 엄격하게 실시되었다. 채점은 대학원에서 창의성을 전공한 박사 2인이 채점하여 그 평균을 최종 점수로 삼았다. 채점자간 신뢰도는 유창성 .89, 독창성 .93, 제목의 추상성 .88, 정교성 .91, 성급한 종결에 대한 저항 .80 등으로 높게 나타났다.

### 3. 연구 절차

본 연구를 진행하기 위해 정은이(2003)의 연구에서 사용한 대학생용 창의성 프로그램을 일부 수정하여 16주 동안 적용하였다. 프로그램의 내용은 표 1과 같다.

연구자는 각 주에 해당되는 수행과제와 방법을 설명한 활동지를 그 전 주에 미리 제시하고 학생들은 준비물과 활동 내용을 주지하고 창의성 활동에 참여하게 된다. 그 이후 수행 결과물과 체크리스트 및 평가용 질문지로 구성된 평가지에 평가자 수준에 따른 평가를 받게 된다. 첫 주와 마지막 주에는 Torrance의 창의성 도형검사 A, B형을 각각 실시한다. 학생들은 기말에 이 모든 자료를 포함한 포트폴리오 형태의 파일을 각자 제출하게 된다.

집단별 평가방식을 자세히 살펴보면 다음과 같다.

실험집단은 공통적으로 1시간에 걸친 창의성 이론 수업을 받고 자기평가, 동료평가, 전문가 평가 집단으로 분류되어 각각 1시간의 창의성 활동을 한 후 각기 다른 평가자에 의해 평가를 받게 된다. 자기 평가 집단은 자신이 스스로 창의성 결과 및 과정에 대해 연구자가 각 주마다 제작해 온 체크리스트와 자기 평가 질문지를 이용하여 평가를 하게 된다. 동료 평가 집단은 역시 연구자가 제작한 자료를 가지고 같은 조의 동료 2명이 공동으로 평가를 하게 된다. 전문가 집단은 창의성 활동 후 결과물과 과정을 본 연구자와 창의성을 전공한 박사 과정 대학원생 1명이 평가를 해 그 다음 주에 평가물을 제시받게 된다. 평가자 수준에 따른 이 세 집단은 공통적으로 연구자가 제작한 체크리스트와 평가용 질문지를 통해 평가를 받게 된다. 학생들은 해당 주에 한 결과물에 대한 평가를 그 다음 주에 제시받게 되며 평가에 제시된 대로 자신의 결과물을 다시 한번 점검하고 보완점을 수정하는 시간을 갖게 된다.

&lt;표 1&gt; 창의성 프로그램의 내용

주	이론 수업 (60분)	워밍업 (10분)	창의성 훈련 (50분)	활동목표
1주	강의에 대한 개요, 창의성 프로그램 소개		TTCT 창의성 도형 검사 A 형	
2주	창의성 연구의 역사	유머 1	브레인스토밍 : 효과적인 다이어트 방법	유창성, 독창성
3주	창의성 연구의 필요성 (창의성의 신화와 현실)	고전명화를 보고 들리는 소리 찾기	창의적인 마인드맵 만들기 : 사랑, 봄, 자유 등을 주제로	융통성, 독창성
4주	창의성의 본질 (정의, 개념)	보여주는 색의 이름 붙이기	만화 말주머니 채우고 미완성 만화 그리기	독창성
5주	창의성 이론1 (인지적, 사회심리적 접근)	끝말잇기, 중간말 잇기	이야기 완성하기 (주어진 문장 이용 하여 이야기 완성하기)	독창성, 논리적 사고
6주	창의성 이론2 (다원적 접근)	시장에 가면-놀이	창의적 퍼즐 문제 만들기 : 바둑판에 단어 조합을 한 후 가로, 세로 문제 만들기	독창성, 유창성
7주	창의적 성격과 능력	유머 2	그림 문장 만들기(기호 조합하여 의미있는 문장 만들기)	독창성, 논리적 사고
8주	창의성 기법 1 (브레인스토밍, 색다른 용도법, 강제결합법)	그림에 제목 붙이기	색다른 용도법 : 칫솔, 우산	독창성, 유창성
9주	창의성 기법 2 (괴상한 해결법, 스킴 퍼, 희망-결점 열거법)	상형문자보고 연상하기	스캠퍼(대치/순응/확대/제거/결합/수정/다른 용도/순서바꾸기) : 핸드폰	독창성, 유창성
10주	창의성 기법 3 (특성 목록법, 형태학적 분석법, 시네틱스)	영화장면 보고 대사쓰기	특성 목록법 : 우리집 화장실을 개선하는 방법	독창성, 유창성
11주	창의적인 문제 해결력	음악 듣고 제목 쓰기	형태학적 분석법 : 재활용품 이용 강의실 개선법	독창성, 유창성
12주	영재 교육과 창의성	유머 3	시네틱스(환상유추, 직접유추, 개인유추) : 효과적인 대학 홍보 전략	독창성, 상상력
13주	창의성 교육과 프로그램	수수께끼	미래의 뉴스 특종을 정하고 미래의 모습 그리기	독창성, 상상력
14주	창의성 측정	서로 다른 그림 찾기	S 대학교 상징물 만들기(로고, 상징탑 및 상징물)	독창성, 상상력
15주	창의성 발달과 장애요인	음악 듣고 그림 그리기	불완전 도형 이용해 그림 그리기	독창성, 융통성, 유창성
16주	기말고사		TTCT 창의성 도형검사 B 형	



## 5. 자료 처리

본 연구에서는 SPSS/PC+과 SAS 프로그램을 사용하여, 공분산분석(Analysis of Covariance) 및 Schéffe 검증으로 수집된 자료를 분석하였다.

## IV. 연구결과

본 연구의 창의성 프로그램 효과를 알아보기 위해 공분산 분석(Analysis of Covariance)을 사용하여 창의성 요인별 점수 차이가 있는지 알아보았다. 이는 사전 검사 점수가 종속변수에 영향을 미치는 효과를 제거하고 순수한 효과만을 분석하기 위해 실시한 것이다. 사전 검사 점수를 살펴본 결과 집단간 차이가 있어 사전 검사(Torrance 도형검사 A형)의 영향력을 통계적으로 사후 검사에서 배제하고 교정평균으로 분산분석을 실시하였다.

창의성 총점의 평균과 각 하위 요인별 평균 차이의 유의성을 검증하기 위하여 공분산 분석한 결과는 다음과 같다. <표 2-1>에서 <표 7-2>까지에 제시된 바와 같이 '제목의 추상성'을 제외한 '유창성', '독창성', '정교성', '종결에 대한 저항' 하위 요인과 '창의성 총점의 평균'에서 세 집단별로 통계적으로 유의한 차이가 나타났다.

<표 2-1> 자기평가, 동료평가, 전문가평가 집단의 '유창성' 점수의 서술통계

		자기평가	동료평가	전문가평가
사전검사	Mean	100.31	101.28	96.10
	SD	16.65	25.11	22.62
사후검사	Mean	115.03	98.78	98.37
	SD	19.68	23.33	19.64
교정평균	Mean	114.52	97.78	99.97
	SD	3.18	3.19	3.30
사례수		32	32	30

<표 2-2> 자기평가, 동료평가, 전문가평가 집단의 '유창성' 점수에 대한 공분산 분석

분산원	SS	df	MS	F
공분산	10818.17457	1	10818.17459	33.28***
집단간	5259.82127	2	2629.91089	8.09***
오차	29253.22958	90	325.03588	
총	45785.92553	93		

\*P<.05, \*\*P<.01, \*\*\*P<.001

구체적으로 각 집단별 '유창성'의 교정평균 점수를 살펴보면 동료평가집단(M=97.78)에 비해 전문가평가 집단(M=99.97)과 자기평가 집단(M=114.52)이 교정 평균 점수가 높았다. 세 집단 중에서는 자기 평가 집단이 가장 높은 점수를 얻었다. 집단별 차이를 알아보기 위해 교정평균 점수를 Schéffe 검증한 결과 자기 평가 집단이 동료평가 집단(p<.001)과 전문가 평가 집단(p<.01)에 비해 유의한 차이를 나타냈다. 하지만 동료평가 집단과 전문가 평가 집단 사이에는 유의한 차이가 없었다.

<표 3-1> 자기평가, 동료평가, 전문가평가 집단의 '독창성' 점수의 서술통계

		자기평가	동료평가	전문가평가
사전검사	Mean	123.44	115.84	116.33
	SD	19.45	26.78	28.38
사후검사	Mean	140.44	121.56	122.10
	SD	12.06	24.31	26.32
교정평균	Mean	137.73	123.09	123.35
	SD	2.97	2.95	3.05
사례수		32	32	30

<표 3-2> 자기평가, 동료평가, 전문가평가 집단의 '독창성' 점수에 대한 공분산 분석

분산원	SS	df	MS	F
공분산	17811.59620	1	17811.59620	63.86***
집단간	4360.21190	2	2180.10595	7.82***
오차	25100.85380	90	278.89838	
총	50230.60638	93		

\*P<.05, \*\*P<.01, \*\*\*P<.001

그 다음 각 집단별로 '독창성'의 교정평균 점수를 살펴보면 동료평가집단(M=123.09)과 전문가평가 집단(M=123.35)에 비해 자기평가 집단(M=137.73)이 교정 평균 점수가 높았다. 세 집단 중에서는 자기 평가 집단이 가장 높은 점수를 얻었다. 집단별 차이를 알아보기 위해 교정평균 점수를 Schéffe 검증한 결과 자기 평가 집단이 동료평가 집단(p<.01)과 전문가 평가 집단(p<.01)에 비해 유의한 차이를 나타냈다. 하지만 동료평가 집단과 전문가 평가 집단 사이에는 유의한 차이가 없었다.

&lt;표 4-1&gt; 자기평가, 동료평가, 전문가평가 집단의 '제목의 추상성' 점수의 서술통계

		자기평가	동료평가	전문가평가
사전검사	Mean	81.19	68.53	75.33
	SD	35.10	36.10	30.25
사후검사	Mean	103.84	95.81	89.83
	SD	33.83	28.02	35.04
교정평균	Mean	102.39	97.33	89.95
	SD	5.61	5.61	5.76
사례수		32	32	30

&lt;표 4-2&gt; 자기평가, 동료평가, 전문가평가 집단의 '제목의 추상성' 점수에 대한 공분산 분석

분산원	SS	df	MS	F
공분산	5815.293124	1	5815.293124	5.84
집단간	2486.985689	2	1243.492845	1.25
오차	89615.96729	90	995.73297	
총	98503.70213	93		

각 집단별로 '제목의 추상성'의 교정 평균 점수는 집단별로 유의한 차이가 없는 것으로 나타났다.

&lt;표 5-1&gt; 자기평가, 동료평가, 전문가평가 집단의 '정교성' 점수의 서술통계

		자기평가	동료평가	전문가평가
사전검사	Mean	95.34	92.72	89.43
	SD	18.18	23.87	20.40
사후검사	Mean	115.84	97.03	99.40
	SD	17.04	26.87	25.48
교정평균	Mean	114.15	96.93	101.30
	SD	3.51	3.50	3.63
사례수		32	32	30

&lt;표 5-2&gt; 자기평가, 동료평가, 전문가평가 집단의 '정교성' 점수에 대한 공분산 분석

분산원	SS	df	MS	F
공분산	14805.75600	1	14805.75600	37.64***
집단간	5077.76957	2	2538.88479	6.45**
오차	35402.63150	90	393.36257	
총	56882.55319	93		

\*P<.05, \*\*P<.01, \*\*\*P<.001

그 다음 각 집단별 '정교성'의 교정평균 점수를 살펴보면 동료평가집단(M=96.93)에 비해 전문가평가 집단(M=101.30)과 자기평가 집단(M=114.15)이 교정 평균 점수가 높았다. 세 집단 중에서는

자기 평가 집단이 가장 높은 점수를 얻었다. 집단별 차이를 알아보기 위해 교정평균 점수를 Schéffe 검증한 결과 자기 평가 집단이 동료평가 집단( $p<.01$ )과 전문가 평가 집단( $p<.05$ )에 비해 유의한 차이를 나타냈다. 하지만 동료평가 집단과 전문가 평가 집단 사이에는 유의한 차이가 없었다.

<표 6-1> 자기평가, 동료평가, 전문가평가 집단의 '종결에 대한 저항' 점수의 서술통계

		자기평가	동료평가	전문가평가
사전검사	Mean	77.22	74.97	77.43
	SD	17.02	17.41	20.68
사후검사	Mean	91.13	77.78	80.40
	SD	17.95	17.05	20.78
교정평균	Mean	90.83	78.43	80.01
	SD	3.00	3.01	3.10
사례수		32	32	30

<표 6-2> 자기평가, 동료평가, 전문가평가 집단의 '종결에 대한 저항' 점수에 대한 공분산 분석

분산원	SS	df	MS	F
공분산	5458.48062	1	5458.048062	18.85***
집단간	2890.780199	2	1445.390099	4.99**
오차	26066.12069	90	289.62356	
총	34708.60638	93		

\* $P<.05$ , \*\* $P<.01$ , \*\*\* $P<.001$

그 다음 각 집단별 '종결에 대한 저항'의 교정평균 점수를 살펴보면 동료평가집단( $M=78.43$ )에 비해 전문가평가 집단( $M=80.01$ )과 자기평가 집단( $M=90.83$ )이 교정 평균 점수가 높았다. 세 집단 중에서는 자기 평가 집단이 가장 높은 점수를 얻었다. 집단별 차이를 알아보기 위해 교정 평균 점수를 Schéffe 검증한 결과 자기 평가 집단이 동료평가 집단( $p<.01$ )과 전문가 평가 집단( $p<.05$ )에 비해 유의한 차이를 나타냈다. 하지만 동료평가 집단과 전문가 평가 집단 사이에는 유의한 차이가 없었다.

<표 7-1> 자기평가, 동료평가, 전문가평가 집단의 '창의성 평균' 점수의 서술통계

		자기평가	동료평가	전문가평가
사전검사	Mean	95.50	90.67	90.93
	SD	13.03	20.12	20.06
사후검사	Mean	113.26	98.19	98.02
	SD	15.13	19.75	19.08
교정평균	Mean	111.40	99.22	98.89
	SD	2.65	2.64	2.72
사례수		32	32	30

마지막으로 각 집단별 '창의성 총점의 평균'에 대한 교정평균 점수를 살펴보면 전문가평가 집단(M=98.89)에 비해 동료평가 집단(M=99.22)과 자기평가 집단(M=111.40)이 교정 평균 점수가 높았다. 세 집단 중에서는 자기 평가 집단이 가장 높은 점수를 얻었다.

집단별 차이를 알아보기 위해 교정평균 점수를 Schéffe 검증한 결과 자기 평가 집단이 동료평가 집단(p<.01)과 전문가 평가 집단(p<.01)에 비해 유의한 차이를 나타냈다. 하지만 동료평가 집단과 전문가 평가 집단 사이에는 유의한 차이가 없었다.

<표 7-2> 자기평가, 동료평가, 전문가평가 집단의 '창의성 평균' 점수에 대한 공분산 분석

분산원	SS	df	MS	F
공분산	10522.61614	1	10522.61614	47.26***
집단간	3163.10065	2	1581.55033	7.01***
오차	20039.54936	90	222.66166	
총	35404.83106	93		

\*P<.05, \*\*P<.01, \*\*\*P<.001

<표 8> 집단별 Scheffe 검증 결과

창의성 하위 요인	집단별	교정평균	자기평가	동료평가	전문가평가
유창성	자기평가	114.52		***	**
	동료평가	97.78			
	전문가평가	99.97			
독창성	자기평가	137.73		**	**
	동료평가	123.09			
	전문가평가	123.35			
제목의 추상성	자기평가	102.39			
	동료평가	97.33			
	전문가평가	89.75			
정교성	자기평가	114.15		**	*
	동료평가	96.93			
	전문가평가	101.30			
종결에 대한 저항	자기평가	90.83		**	*
	동료평가	78.43			
	전문가평가	80.01			
창의성 평균	자기평가	111.40		**	**
	동료평가	99.22			
	전문가평가	98.75			

\*P<.05, \*\*P<.01, \*\*\*P<.001

이러한 결과를 종합해 보면 자기평가 집단과 나머지 두 집단(동료평가 집단과 전문가평가 집단)간에 창의성 평균과 유창성, 독창성, 정교성 및 종결에 대한 저항에서 통계적으로 유의한 차이를 나타냈다.

## V. 결론 및 논의

본 연구에서는 창의성 프로그램을 적용하여 대학생들에게 창의성 훈련을 받게 한 후 자기평가, 동료평가, 전문가 평가 등 각기 다른 평가자에 의한 창의성 평가를 실시하였다. 이후 창의성 증진 효과가 어떻게 다른지 알아보았다.

첫째, 자기평가 집단이 나머지 두 집단(동료평가, 전문가평가)보다 창의성 평균과 유창성, 독창성, 정교성 및 종결에 대한 저항에서 높은 점수를 나타냈고 이는 통계적으로 유의한 결과였다. Schéffe 검증 결과 집단간 차이도 유의했다.

이는 Amabile(1998)의 실험 결과에서도 언급했듯이 타인의 시선과 평가에 대한 지나친 부담으로 인해 동료평가와 전문가 평가를 한 집단에서는 창의성 증진 효과가 나타나지 않은 것이다.

또한 이러한 결과는 자기평가를 할 때 스스로 자신의 수행과정 및 결과를 돌아보고 성찰해보는 시간 자체가 창의성을 발현하는 하나의 과정이 되기 때문인 것으로 보인다. 이 과정에서 개인은 자신이 스스로 문제를 해결하고 개선하려는 자발성을 키우며 자신의 평가에 대한 자신감과 만족감을 느끼게 되는 것이다. Guilford(1960)는 창의성과 정적 상관성이 있는 요인 중의 하나로 자발적 융통성을 들었고 Barron(1953)은 창의적 성격 특성으로 독자적인 비판을 내리는 특성을 언급했다. 이러한 기존의 연구들을 볼 때 자기평가를 통해서 자발성과 독자적인 판단력 및 평가에 대한 만족 등을 얻게 되고 이것은 다시 창의적 능력에 영향을 미치게 되어 창의성 증진의 효과를 가져오게 되는 것이다. 또한 자기 평가를 하게 됨으로써 스스로 내적인 동기 유발을 일으켜 호기심과 만족감을 충족시키는 결과를 가져오는 것이다. 그러나, 동료평가와 전문가 평가에서는 개인의 이러한 성찰 및 자발적으로 사고하는 과정이 없기 때문에 창의성 증진 효과가 미미한 것으로 보인다. 이는 평가가 타인에 의해서 주어지기 때문에 내적인 동기유발보다는 외적인 동기 유발에 의존해 자발적인 동기의식이 떨어지게 되는 것이다.

둘째, TTCT 검사의 창의성 하위 요인 중에서 제목의 추상성 하위 요인에서는 유의한 차이가 나타나지 않았는데 이는 제목의 추상성 요인 자체가 언어적인 측면을 많이 내포하고 있기 때문으로 보인다. 본 연구에서 사용한 창의성 프로그램 자체가 언어적인 내용이 부족하기 때문에 언어 창의성 증진에는 효과적이지 못하며 그렇기 때문에 제목의 추상성 요인에서는 유의미한 차이가 나타나지 못한 것이다.

셋째, 동료평가 집단과 전문가 평가 집단간에는 통계적으로 유의한 차이가 나타나지 않았다. 이러한 결과는 창의성 증진의 차원에서 보면 같은 또래가 평가를 하든 전문가가 평가를 하든 별 차이가 없다는 것이다.

이러한 결과를 볼 때, 평가자 유형 측면에서 보면 가장 창의성을 증진시키는 방법은 자기 평가이므로 학교에서는 창의성 프로그램을 운영하거나 교과와 관련된 특별한 과제를 수행할 경우 자기 평가를 할 수 있는 시간을 따로 마련하는 것이 효과적일 것이다. 실제 학교 현장에서 다수의 학생들을 한, 두 명의 교사나 특별히 초빙된 전문가가 평가를 하는 것은 무리인 경우가 많고 시간의 제약으로 동료 평가를 하는 것도 어렵다. 따라서 논술형 자기 평가 등을 수업 중에 실시하거나 과제로 제시하여 피드백을 하는 방법이 효과적일 것이다. 또한 교과 영역에서 교과서 내용 자체에 창의성 증진을 위한 소단원이 개설될 경우 자기 평가 체크리스트와 간단한 설문 등을 삽입하는 것이 학생들의 창의성 증진에 큰 도움이 될 것이다.

연구의 제한점 및 향후 연구를 위한 제언을 하면 다음과 같다.

첫째, 본 연구에서 창의성 측정 도구로 Torrance의 TTCT 도형검사를 활용하였는데, 이 외에 다른 검사 도구를 사용하여 연구하였을 때 결과에 어떤 차이가 있는지 알아볼 필요가 있다. TTCT 언어 검사나 다른 유형의 창의성 검사를 활용했을 때에도 같은 결과를 나타내는지 추후 연구가 필요하다.

둘째, 창의성 평균 점수를 기준으로 할 때 창의성이 높은 집단과 낮은 집단에서도 동일한 결과가 나오는지 살펴볼 필요가 있다. 창의성 고저집단별로 창의성 증진을 위한 다른 평가 방식이 요구될 가능성이 있기 때문이다.

셋째, 다양한 창의성 평가 방법 모색 및 이의 효과에 대한 추후 연구가 필요할 것이다. 본 연구에서는 평가자 수준에 따른 차이만을 살펴보았으나 주·객관식 지필평가, 논술형 평가, 포트폴리오 평가 등 다양한 평가 방법에 따라 창의성 증진에 어떠한 차이를 가져오는지 살펴보아야 할 것이다.

넷째, 본 연구에서는 대학생만을 대상으로 하였기 때문에 연구 결과를 대학생에게만 한정시켜서 적용해야 한다. 따라서, 초중고 등 다른 연령대의 학생들도 동일한 결과가 나오는지 연구할 필요가 있다. 연령대별 학습자 특성으로 인해 다른 평가 방식이 더 창의성 증진에 효과가 있을 수 있기 때문이다.

## 참고문헌

## 참 고 문 헌

- 박동준·허경조(1988). 대학생을 위한 창의력 개발 프로그램. *충북대학생활연구*, 제12집, 15-48.
- 박동준·임성문·정영수(1989). 창의력 개발 프로그램의 효과 검증. *충북대학생활연구*, 제 2 집, 1-20.
- 박병기(2004). 교양강좌를 이용한 대학생 창의성 교육의 효과 분석. *교육심리연구*. 제 18 권 제 2 호., 69-81.
- 백순근(1998). 수행평가에 대한 이론적 기초. *중학교 각 교과별 수행평가의 이론과 실제*. 서울: 원미사.
- 부지영(2000). 창의성 훈련이 초등학교 아동의 창의적 사고력 및 정의적 특성에 미치는 효과. *전남대학교 대학원 석사학위논문*.
- 장좌욱(2002). 유아를 위한 창의성 프로그램이 유아의 창의성 증진에 미치는 효과. *성균관대학교 대학원 석사학위 논문*.
- 전경원(1997). 창의성 프로그램이 대학생의 창의성 계발에 미치는 효과. *교육심리연구*. 제 11 권, 제 2 호. pp. 223-253.
- 전경원(2000). 창의성을 저해하는 요인 조사 연구 : 유아교육과 여학생을 중심으로. *열린 유아 교육*, 5(1), 163-180.
- 전평국·이재학·백석윤·박성선·감성만(2001). 열린교육에서의 평가와 창의성 계발. *청람수 학교육*. 제 9 권. pp. 1-25.
- 정은이(2003). 개별·협동학습에 따른 창의성 프로그램이 대학생의 창의성 계발에 미치는 효과. *교육심리연구*. 제 17 권. 제 1호. pp. 281-297.
- 정황순(2001). 창의성 계발 프로그램의 적용이 창의성, 정서 지능 및 다중 지능에 미치는 효과. *원광대학교 대학원 박사학위논문*.
- 조성수(1996). 초등학교 아동의 창의성 신장을 위한 프로그램의 효과 연구. *고려대학교 대학원 석사학위 논문*.
- 최미정(2000). 창의성 교육 프로그램이 아동의 창의성 신장에 미치는 효과. *광주교육대학교 교육대학원 석사학위논문*.
- 호사라(2001). 창의성 교육 프로그램의 유형이 초등학생의 창의성 신장에 미치는 효과. *서울대학교 대학원 석사학위 논문*.
- Amabile, T. M.(1998). *Creativity in context*(3rd. ed.) Boulder, Colorado: Westview Press.
- Barron, F. (1953). Complexity-simplicity as a personality dimension. *Journal of Abnormal, Social, Psychology*, 48. 163.



- Burns, M. B. (1983). *A comparison of three creative problem-solving methodologies*. Doctoral Dissertation, University of Denver.
- deBono, E. (1971). *Lateral thinking*. London: Ward Lock: Education.
- deBono, E. (1975). *Cort thinking*. Dorset, England: Direct Educational Services.
- deBono, E. (1978). *Teaching thinking*. London: Pelican Books.
- Feldhusen, J. F., Treffinger, D. J., & Bahlke, S. J. (1970). Developing creative thinking: The Purdue creativity program. *Journal of Creative Behavior*, 4(2), 85-90.
- Getzel, J., & Jackson, P., (1962). *Creativity and intelligence: Explorations with gifted students*. New York: Wiley.
- Guilford, J. P. (1960). *Research conference on the identification of creative, scientific talent*, In M. I. Stein, & S. J. Heinge (ed.), *Creativity and the individual, III: The free Press of Glencoe*.
- Khatena, J. (1971). A second study training college adults to think creatively with words. *Psychological Report*, 28, 385-386.
- Korth, W. L. (1973). Training in creative thinking: The effect on the individual of training in the "Synectics" method of group problem solving. *Dissertation Abstract Internation*, 33, 3947B.
- Noller, R. B., Parnes, S. J., & Biondi, A. M. (1976). *Creative actionbook*. New York: Scribners.
- Osborn, A. (1963). *Applied Imagination*. New York: Scribner's.
- Torrance, E. P. (1974). *Torrance tests of creative thinking: Directions manual and scoring guide* (Verbal test booklet A, B). Scholastic Testing Service, Inc.
- Treadwell, Y. (1970). Humor and creativity. *Psychological reports*, 26, 55-58.
- Wey, B. L. (1983). *The effect of a guidebook to improve creative thinking abilities*, Doctoral dissertation, University of Northern, Colorado.
- Ziv, A. (1976). Facilitating effects of humor on creativity. *Journal of Educational Psychology*, 68, 318-322.
- Ziv, A. (1988). Using humor to develop creative thinking. *Journal of Children in Contemporary Society*, 2, 99-116.

\* 2004년 6월 16일 / 1차 심사 2004년 9월 7일 / 2차 심사 2004년 11월 9일

\* 정은이: 고려대학교 대학원 교육학과에서 교육심리 전공으로 박사학위를 취득함. 현재 청운대학교 교수로 재직 중. 관심 연구 분야는 창의성, 상담심리, 정신건강 등이다.

\* e-mail: jejei@dreamwiz.com

Abstract

## Effects of Creativity Training Programs On College Students' Creativity According to Evaluator Types

Jeong Eun I\*

In this study of creativity program, college students participated in creativity enhancement training and three groups of them were evaluated by different ways such as self-evaluation, peer-evaluation, and expert-evaluation. Then this study tried to find out how different ways of evaluation influence the enhancement of creativity, using the TTCT: Figural. The followings are the results of this study.

First, the self-evaluation group received significantly higher scores on fluency, originality, elaboration, resistance to closure and the average score than the other two groups. This result implies that the self-evaluation may facilitate participants' reflection of their own performance by which creativity emerges. In this process of self-evaluation, individuals develop the initiative and satisfaction about their evaluation which influence creative abilities.

Second, there were no significant difference on creativity scores between the peer-evaluation group and the expert-evaluation group. This implies that the evaluations by peers and experts have similar effect on creativity enhancement.

Key words: Evaluator Types, Creativity Training Programs

---

\* Chungwoon University