

메타분석을 통한 STEAM 교육의 효과검증

김지원 · 원효현*
(부경대학교)

Effectiveness of STEAM Education applying a meta-analysis

Ji-Won KIM · Hyo-Heon WON*
(Pukyong National University)

Abstract

The purpose of this study was to investigate the effectiveness of integrated education research conducted in Korea and to propose a meaningful discussion for further research.

Among the studies conducted for last five years, the relevant 69 research articles were selected, and 211 effect sizes were calculated. Effect sizes were analyzed with different dependent variables including creativity, problem solving ability, inquiry skills, scientific recognition, attitude, interest, motivation and self-esteem. In addition, effect sizes with different moderating variable, such as samples sizes, characteristics of subjects, class types, core disciplines were compared.

The conclusions of this study was that STEAM education had a relatively middle effect size. Finally, researchers discussed the results related with previous results, and provided the implications and suggestion for future studies.

Key words : STEAM Education, Meta-analysis, Educational effectiveness

I. 서론

교육과학기술부는 범교과적인 학습을 위해 초·중등학교 수준에서부터 다양한 학문을 연결하여 통합적으로 사고하고 문제를 해결할 수 있는 능력을 기르기 위한 STEAM(Science, Technology, Engineering, Arts, & Mathematics) 교육을 제시하였다(MEST, 2010). 2011년에 본격적으로 STEAM 교육이 실시된 이후 다양한 분야에서 STEAM 교육 관련 연구들이 활발하게 진행되고 있다(SHIN Jin-kyung et al.,2013). STEAM 교육 연구의 유형은 STEAM 교육 프로그램 개발, 분석, 효과, 사례, 인식, 동향으로 분류할 수 있으며 그 중 특히 개발과 효과를 연구한 논문이 가장 많다(Oh,

hyun-sook, 2012; Lee, ji-hae, 2014; An, hae-ran, 2015). 이와 관련하여 2011년부터 2014년 8월말까지 발표한 STEAM 관련 약 380편의 논문을 분석한 결과에 따르면, STEAM 교육의 효과성 검증 논문은 크게 창의성 영역, 정의적 영역, 인지적 영역에 관한 연구로 나타나고 있다(An, hae-ran, 2015). 한편, 최근에는 교육학 전반에 걸쳐 증거에 기반한 교육 효과의 검증이 보편화되면서, 효과성에 대한 관심이 높아지고 있다. 바로 이러한 점에서 STEAM 교육의 효과를 종합적으로 검증할 필요성이 제기된다. 또한 최근 몇 년간 폭발적으로 증가한 많은 양의 STEAM 교육 연구가 진행되어 다양한 연구결과가 축적되어 온 현 시

* Corresponding author : 051-629-5972, wonhyo@pknu.ac.kr

점에서 개별적 연구 단위로 보고된 STEAM 교육의 효과에 대해 포괄적으로 분석하는 작업이 필요할 것으로 생각된다. 이에 본 연구는 이러한 의문과 필요성에 대한 해답을 통계적 기법으로 제공하려는 시도로 메타분석의 방법을 활용하고자 하는 것이다. 향후 보다 효과적인 STEAM 교육 프로그램을 개발하기 위해서도 선행 연구들의 총합을 통해 다양한 변인의 영향을 분석할 필요가 있으며 기존 연구들을 종합하고, 분류하여 함축성 있고 의미 있는 일반적인 결론을 추출하고자 하였다. 이와 같은 본 연구의 목적을 위해 설정한 구체적 연구 문제를 제시하면 다음과 같다.

첫째, STEAM 교육의 전체적인 평균 효과 크기는 어느 정도인가?

둘째, STEAM 교육이 중속변인(창의성 영역, 인지적 영역, 정의적 영역)에 미치는 효과 크기는 어느 정도인가?

셋째, 조절변인(실험집단 인원수, 실험처치 횟수, 수업유형, 학급유형, 학교유형, 중심교과)에 따른 STEAM 교육의 평균 효과 크기는 어느 정도인가?

II. 연구 방법

1. 분석대상

본 연구는 STEAM 교육의 효과를 규명하고자 2010년부터 2015년 12월까지 국내에서 발표된 관련 분야 석·박사학위논문, 학술지 게재논문, 간행물 및 개인 연구물 등을 대상으로 수집하였다. 자료를 수집하기 위하여, 2015년 12월 현재, 국회도서관(<http://www.nanet.go.kr>), 한국교육학술정보원 학술연구정보서비스(KERIS, <http://www.riss.kr>), 국가전자도서관(<http://www.dlibrary.go.kr>), 국립중앙도서관(<http://www.nl.go.kr>), 한국학술정보원(<http://Kiss.kstudy.com>), 누리미디어 DB-PIA(<http://www.dbpia.co.kr>) 등의 온라인 검색 DB(DataBase)를 활용하여 분석대상 논문을 검색

하였다. 논문 목록을 추출하기 위해 사용한 주요 검색어는 ‘STEAM’과 ‘융합인재교육’이었으며, 가급적 다양한 키워드를 검색하였으며 관련 문헌을 최대한 수집하고자 연구물의 참고문헌까지 탐색하였다. 또한 한국과학창의재단(www.kofac.or.kr)의 DB에서 2012년, 2013년, 2014년 학술대회 자료집을 함께 수집하여 대상 논문을 하나도 빠짐없이 수집하려고 하였다. 대상 논문의 탐색하는 기간은 일정한 기간이 없이 연구 초기단계에서 코딩 단계, 분석 단계에서도 수시로 문헌탐색을 실시하여 최대한 많은 논문을 수집하려고 하였다. 이러한 과정을 통해 검색된 논문 중에서 아래의 절차에 의해 자료 선정 기준을 만족하는 논문을 최종 분석 대상으로 선정하였다.

첫째, STEAM 교육 프로그램을 활용하여 STEAM 교육의 효과를 연구한 논문이어야 한다.

둘째, 실험연구나 준실험연구 설계를 이용했으며, 실험집단과 통제집단을 설정한 연구결과를 다룬 연구물이어야 한다. 셋째, 실험집단과 통제집단의 평균치, 표준편차 및 사례수 등의 기초통계 자료를 가지고 있는 논문으로 제한한다. 그 외의 경우에는 효과 크기를 산출하는 데 필요한 통계치(t-test, F-test, 상관관계수 r 등)가 제시되어 있는 논문을 분석 대상으로 선정하였다. 넷째, 연구의 대상은 유치원, 초등학교, 중학교, 고등학교에 해당하는 논문을 분석대상으로 선정하였다. 다섯째, 본 연구에서는 학업성취도 변인에 관한 연구물은 제외하였다. 그 이유는 평가 전문가의 의견을 반영하여 각 학교급, 학생 및 학교 특성이 반영된 성취도 점수가 그 타당도가 부족하다고 판단하였기 때문이다. 최초로 검색되어진 STEAM 관련 논문은 총 8,119편으로 학위논문 3,229편, 학술지논문 4,890편, 학술대회연구물 171편이었으며, 자료 선정 기준을 통해 본격적 논문 분석을 실시하였다. 1차 논문분석은 논문의 제목과 국문초록을 토대로 이루어졌으며, 1차 논문분석을 통해 총 8,119편 중 주제와 관련 없는 연구, 연구대상이 적절하지 않은 논문 7,708편을 제외

하여 총 582편의 논문을 추출하였다. 1차 논문분석을 통해 선정된 582편의 논문들 중 자료선정 기준을 만족시키는 논문을 추출하기 위해 2차 논문분석을 실시하였다. 이때 1차 논문분석을 통해 선정된 582편의 논문은 전문에 대해 세밀히 읽고 검토되었다. 이 중 실험연구가 아닌 논문 443편, 비교집단이 없는 논문 34편, 다른 연구와 중복되는 연구 13편, 자료가 부적절한 논문 16편을 제외하여 총 76편의 논문을 추출하였다. 정상범위를 상당히 벗어난 극단치(extreme value)인 이상치(outlier)는 전체 평균 효과 크기를 구하는데 방해가 될 수 있어 제외하여 최종 선정 한 논문은 총 69편으로, 학위논문 51편, 학술지 논문 15편, 학술대회 3편이었다.

2. 연구 절차

체계적이고 종합적인 분석을 위해 위의 기준으로 코딩 매뉴얼과 코딩표를 개발하여 자료코딩을 실시하였다. 코딩 매뉴얼과 코딩표 및 자료코딩 과정의 모든 과정은 연구자를 포함하여 교육학 박사 1명(상담전공)과 STEAM 교육 현장 연구자 3명(교육학 석사), 통계 및 평가 전문가 1명(교육학 박사)의 의견을 반영하여 실시하였으며, 코딩을 실시하는 과정에서 코딩에 참여한 연구자간에 충분한 코딩 전후에 범주의 분류 및 입력 방식에 대한 사전·사후 협의를 통하여 코딩의 불일치 및 신뢰성 문제를 해결하고자 노력하였다. 전문가들의 코딩 범주 간에 불일치가 나타난 경우 충분한 논의를 통해 해결하였다. 코딩 입력 등 협의과정 중 의견이 불일치한 논문은 없었으므로 코딩자간 신뢰도나 일치도는 별도로 계산하지 않았다. 분석 범주를 설정하는 방법으로는 연구자의 직관에 따라 설정하는 방법, 이론적 배경 또는 선행연구에 따라 설정하는 방법, 연구결과들에 제시된 공통 변수들을 추출하여 범주화하는 방법 등이 있다(Guzzo, Jackson, & Katzell, 1987). 본 연구에서는 선행연구를 포함한 이론적 배경과 연구결과들

에 제시된 공통 변수 추출방법을 함께 사용하였다.

먼저, 종속변인은 An(2015)의 STEAM 교육의 효과성 검증과 선행연구를 통해서 확인된 창의성 영역, 인지적 영역, 정의적 영역으로 분류하였다. STEAM 교육과 창의성에 관한 연구에서는 창의성, 창의적 인성, 창의·인성, 창의적 사고활동 등 다양한 표현이 사용되고 있고 창의성의 정의 및 창의성 구성 요인은 학자들마다 의견이 다양하여 본 연구에서는 Torrance(1962)와 Lee et al(2010)의 구성 요인을 참고로 창의성의 지적요인(유창성, 융통성, 제목의 추상성, 독창성, 정교성, 성급한 종결에 대한 저항)을 사고 기능으로 정의적 요인(호기심, 탐구심, 자신감, 자발성, 정직성, 개방성, 독자성, 집중성 등)을 사고 성향으로 재분류하였다. 그리고, Park et al(2010)를 참고하여 인지적 요인(지식, 학업성취, 이해정도, 인지적 변화, 융합적사고력, 문제해결력, 탐구력 등) 중 문제해결력과 탐구능력을 인지적 영역으로 구분하였고, 정의적 요인(정의적 특성, 프로그램만족도, 참여도, 선호도, 흥미도, 태도, 동기, 인식, 실생활 적용력, 자아효능감, 자신감 등) 중 인식, 흥미, 태도, 학습동기, 자아효능감을 정의적 영역으로 분류하였다.

조절변인은 Oh(2012)의 연구에서 제시된 분석의 대상과 선행 논문의 특징을 범주화 하는 과정을 반영하여 실험집단 인원수, 실험처치 횟수, 수업유형, 학급유형, 학교유형, 중심교과로 분류하였다. 수업유형은 Kofac(2012)에서 제시한 일반적인 수업 유형을 바탕으로 교과 내, 교과 연계, 창의 활용형, 학교밖으로 분류하였고, 학급유형은 일반학급, 영재학급, 특수학급으로, 학교유형은 유치원, 초등학교, 중학교, 고등학교로 분류하였다. 중심교과는 STEAM 교육 프로그램을 설계할 때 내용이나 운영의 중심이 되는 교과가 무엇인지에 따라 분석하였다. 과학, 수학, 기술, 예술 교과 중심으로 분류하였으며, 예술 교과에는 미술, 음악 외에 사회, 국어 교과 등을 포함하였다.

3. 자료 분석

가. 효과크기의 산출

효과크기란 실험집단의 평균값에서 비교집단의 평균값을 뺀 값을 두 집단의 통합 표준편차 값으로 나눈 값을 의미한다(Borenstein M. et al., 2009). 본 연구에서는 STEAM교육 프로그램의 효과를 검증하기 위해 (식 1)과 같이 Cohen's *d*의 효과크기 계산법을 사용하였다.

$$d = \frac{\bar{X}_1 - \bar{X}_2}{S_p} \quad (\text{Formula 1})$$

$$S_p = \sqrt{\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{(n_1 + n_2 - 2)}}$$

d=Effect size, \bar{X}_1 =The average of the experimental group, \bar{X}_2 =The average of the comparison group, S_p =Integrated SD

그러나 앞서 제시한 Cohen's *d*는 사례수가 작은 경우 과대평가되는 경향이 있어 정확한 모집단의 표준편차를 추정하기 어렵다. 이를 교정하고자 효과크기 값을 (식 2)에 제시한 Hedges' *g*로 교정하였다.

$$g = J \times d$$

$$J = \left[1 - \frac{3}{4(n_1 + n_2) - 9} \right] \text{ or } \left(1 - \frac{3}{4df - 1} \right) \quad (\text{Formula 2})$$

J=Correction factor, *d*=Effect size

나. 동질성 검증

동질성 검증은 모든 효과크기들이 동일한 모집단의 효과크기를 추정하고 있는가의 여부를 검증하는 것을 의미한다(Oh, Sung. Sam. 2009). 이를 위해 동질성 검증의 통계량 *Q*값을 산출하였다. *Q*값의 산출 공식은 다음의 (식 3)과 같다.

$$Q = \sum W_i (g - \bar{g})^2 = \sum \left(\frac{g - \bar{g}}{S} \right)^2 \quad (\text{Formula 3})$$

Q=The observed degree of distribution, *W*=Weight, *g*=Effect size(Hedges' *g*)

동질성 검증 통계량 *Q*값에 대한 해석은 카이 제곱 분포에 근거한다. 그 이유는 *Q*값이 카이 제곱 분포와 동일하기 때문이다. 하지만 *Q*값은 귀무가설의 유무만을 검증한다. 즉, 각 연구의 모집단의 효과크기들이 모두 같다는 귀무가설만을 검증할 뿐이다(Borenstein M. et al., 2009). 따라서 효과크기의 동질성을 넘어선 초과분산을 산출하기 위해 *T*²과 *I*²값을 산출한다. *T*²과 *I*²값의 산출 공식은 다음의 (식 4), (식 5)와 같다.

$$T^2 = \frac{Q - df}{C} \quad (\text{Formula 4})$$

$$C = \sum W - \frac{\sum W^2}{\sum W}$$

*T*²=Distributed between research, *Q*=The observed variance(Total variance), *C*=Scaling factor, *df*=Degrees of Freedom, *W*=Weight

$$I^2 = \frac{Q - df}{Q} \times 100\% \quad (\text{Formula 5})$$

*I*²=The proportion of variance between studies *Q*=The observed variance(Total variance), *df*=Degrees of Freedom

*T*²값이 0이면 랜덤효과 모형은 고정모형이 되므로, 랜덤효과 모형을 채택하더라도 문제가 되지 않으며, *p*값이 .10보다 작고, *I*²값이 50% 이상이면 상당한 정도의 이질성이 있다고 판단한다(Borenstein M. et al., 2009).

다. 이질성에 대한 분석

만약 효과크기들의 분포가 동질적이지 않은 경우 산출된 평균 효과크기의 값은 집단분포를 대표하는 것으로서 적절한 값이 되지 못한다(Oh, Sung. Sam. 2009). 즉, 앞서 살펴본 동질성 검증

(Q, T^2, I^2)에서 이질성이 나타난다면 이질성이 왜 생겼는지를 살펴볼 필요가 있다. 본 연구 효과크기의 이질성을 설명하기 위해 효과크기를 설명하는 조절변인을 활용하여 메타 회귀분석과 Meta-ANOVA를 실시하였다.

라. 출간 오류(publication bias)검증

메타분석에 실제 포함된 연구들로 얻은 결과가 분석에 포함되어야 할 연구들이 모두 포함되었을 때의 결과와 다르다면 출간 오류가 존재한다고 본다(Borenstein M. et al., 2009). 이는 그림(Funnel Plot)을 통해 이를 확인 할 수 있으며 수직된 평균 효과크기를 중심으로 좌우대칭이 이루어지면 출간오류가 없다고 결론을 내릴 수 있다. 출간오류가 존재할 경우에는 메타분석을 통해 산출된 효과크기를 보정할 필요가 있다(Oh, soon-moon, 2011). 본 연구에서는 출간오류가 발생할 경우 'Trim and Fill' 방법과 '누적메타분석(cumulative meta-analysis)'를 사용하였다.

마. 민감성 분석(sensitivity analysis)

출간 오류 분석과 함께 메타분석에서 데이터 오류에 대한 분석 방법으로 민감성 분석이 있다. 어떤 특정 연구를 제외 또는 포함시켰을 때 전체 평균 효과 크기가 어떻게 달라져서 연구 결과에 얼마나 변화가 있는지를 검증하는 것이 민감성 분석의 목적이다. 여기에는 누적메타분석(cumulative meta-analysis)도 함께 포함된다(Borenstein M. et al., 2009).

바. 데이터 분석방법

본 연구에서는 STEAM 프로그램의 효과를 분

석하기 위해 CMA(Comprehensive Meta Analysis) 프로그램을 사용하였다.

III. 분석 결과

STEAM 교육의 효과를 규명하고자 2010년부터 2015년 12월까지 국내에서 발표된 학위논문 및 학술지 논문, 기타 연구물에서 관련된 연구물 69편을 분석대상으로 효과 크기 사례수 211개를 산출하여 분석하였다. 이는 STEAM 교육 프로그램 전체에 대한 결과를 제시함과 동시에 효과에 해당하는 하위 영역을 구분하여 각각의 결과를 제시한 경우에는 각각의 효과 크기를 사례로 산출하였기 때문이다.

1. STEAM 교육의 평균 효과 크기

STEAM 교육의 전체 평균 효과 크기를 산출한 결과는 다음과 같다. <Table 1>에서 보는 바와 같이 효과 크기의 분포를 측정하는 관찰된 분포 Q 값은 164.313이고, 모든 연구들이 동일한 효과 크기의 추정치를 갖고 있다고 가정할 때 기대되는 기대 분포 값은 $df=68$ 로 나타났다. 그리고 연구 간 분산, 즉 실제 분산은 $T^2=0.097$ 로 나타났다. 또, I^2 의 값은 58.616으로 추출된 효과 크기는 서로 이질적인 것으로 나타났다($p=0.000$). 즉, 연구간 효과 크기가 일관되지 않고 효과 크기 간 차이가 발생하였다. 따라서 본 연구에서는 무선 효과 모형을 이용해서 전체적인 효과 크기를 측정하였다. 본 연구에서 설정한 STEAM 교육의

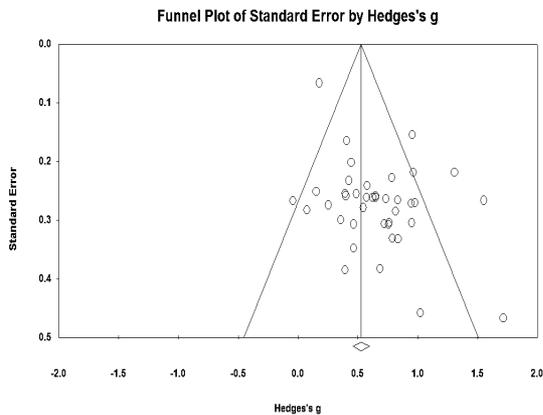
<Table 1> Mean effect sizes of STEAM education

Variables	Model	k	ES(g)	U ₃	95% CI		Q	df	p	I ²
					Lower	Upper				
STEAM education	Fixed	69	0.684	75.17	0.623	0.746	164.313	68	0.000	58.616
	Random	69	0.719	76.11	0.621	0.817				

k=number of studies; ES(g)=Effect Size(Hedges'g); U₃=Non-redundant percentile index; 95% CI=95% confidence interval; Q=The observed variance; df=Degrees of Freedom; I²=The proportion of variance between studies

전체 평균 효과 크기는 0.719(95% CI 0.621-0.817)인 것으로 나타났으며, 95% 신뢰구간에서 0을 포함하고 있지 않기 때문에 효과 크기는 통계적으로 유의한 것으로 나타났다. 효과 크기의 누적분포적 해석방법인 U_3 는 76.11의 백분위로 나타났다.

이어서 최종 선정된 69편의 논문(효과 크기 사례수= 211)에 대한 출간 오류분석을 실시하였으며, 그 결과는 [Fig. 1]과 같다. 출간 오류를 분석한 결과, 좌우 대칭형을 완전히 이루지 못하고 있어 출간 오류가 다소 존재하는 것으로 나타났다. Egger의 회귀분석을 통해서도 회귀식 초기값(Intercept)의 유의확률이 통계적으로 유의함($p = 0.0213$)이 나타났으며, 오류가 있다고 말할 수 있다.



[Fig. 1] Funnel plot

이러한 출간 오류를 보정하는 Duval과 Tweedie(2000)의 ‘Trim and Fill’을 이용하여 비대칭을 대칭으로 교정한 결과 출간 오류로 인해 메타분석에 포함하지 못한 효과 크기 사례들(16개)이 포함된 경우 보정된 평균 효과 크기(adjusted value)가 0.563(95% CI 0.452-0.674)인 것으로 나타났으며, 기존분석 결과인 관찰된 평균 효과 크기(observed value)가 0.719(95% CI 0.621-0.817)인 것으로 나타났다. 이러한 결과는 observed value와

adjusted value를 비교했을 때 원래 값보다 감소했지만 여전히 통계적으로 의미 있는 결과라 할 수 있다.

또한 ‘cumulative meta-analysis’를 활용하여 누적분석을 실시하였다. 연구의 크기순으로 62개(누적 가중치가 90%에 해당)의 연구를 투입했을 때 효과 크기는 0.684로 나타났고, 전체 효과 크기인 0.719의 전체 결과에 큰 영향을 미치지 못하는 것으로 나타났다. 따라서 효과의 크기가 큰 것으로 나타난 작은 표본의 연구들이 전체 결과에 영향을 주는, 즉 오류를 일으키는 원인이라고 말할 수 없다. 그리고, ‘one study removed’를 이용한 민감성 분석의 결과, 특정연구에서 전체효과($g = 0.719$)에 영향을 줄 수 있는 효과의 차이는 발견하지 못하였다.

2. STEAM 교육이 증속변인에 미치는 평균 효과 크기

가. 창의성 영역

STEAM 교육이 창의성 영역의 변인에 미치는 평균 효과 크기를 산출한 결과는 <Table 2>와 같다. STEAM 교육이 창의성 영역의 창의적 사고기능의 효과 크기는 0.706(95% CI 0.544-0.867), 창의적 사고성향의 효과 크기는 0.836(95% CI 0.600-1.072)인 것으로 나타났다. 두 영역 모두 95% 신뢰구간에서 0을 포함하고 있지 않기 때문에 STEAM 교육이 창의성 영역인 창의적 사고기능, 창의적 사고성향에 미치는 효과 크기는 모두 통계적으로 유의한 것으로 나타났다. 효과 크기의 누적분포적 해석방법인 U_3 는 창의적 사고기능 76.11의 백분위, 창의적 성향 79.95의 백분위로 나타났다. STEAM 교육에 대한 창의성 영역들의 동질성을 검증한 결과 창의적 사고기능($Q = 36.911, df=19, p < .05, I^2=48.525$), 창의적 사고성향($Q=72.577, df=21, p < .001, I^2=71.065$)으로 나타나 창의적 사고기능과 창의적 사고성향은 효과 크기가 모두 이질적인 것으로 나타났다. 즉, 연구

<Table 2> Comparison of effect sizes with different dependant variables

Variables	Model	k	ES(g)	U ₃	95% CI		Q	df	p	I ²
					Lower	Upper				
creativity (Thinking skills)	Fixed	20	0.700	75.80	0.588	0.811	36.911	19	0.008	48.525
	Random	20	0.706	76.11	0.544	0.867				
creativity (Thinking tendency)	Fixed	22	0.784	78.23	0.662	0.907	72.577	21	0.000	71.065
	Random	22	0.836	79.95	0.600	1.072				
problem solving ability	Fixed	9	0.584	71.90	0.431	0.737	8.4501	8	0.391	5.326
	Random	9	0.592	72.24	0.432	0.751				
inquiry skill	Fixed	9	0.656	74.53	0.482	0.829	37.751	8	0.000	78.808
	Random	9	0.813	79.10	0.426	1.201				
scientific recognition	Fixed	9	0.525	70.19	0.346	0.704	56.725	8	0.000	85.897
	Random	9	0.565	71.56	0.083	1.046				
interest	Fixed	25	0.635	73.89	0.537	0.734	49.986	24	0.001	51.987
	Random	25	0.641	73.89	0.487	0.785				
attitude	Fixed	35	0.683	75.17	0.597	0.769	76.547	34	0.000	55.583
	Random	35	0.696	75.80	0.565	0.827				
motivation	Fixed	4	0.632	73.56	0.397	0.868	0.453	3	0.929	0.000
	Random	4	0.632	73.56	0.397	0.868				
self-esteem	Fixed	4	0.628	73.56	0.353	0.903	2.319	3	0.509	0.000
	Random	4	0.628	73.56	0.353	0.903				

간 효과크기가 일관되지 않고 효과 크기 간 차이가 발생하였다.

나. 인지적 영역

STEAM 교육이 인지적 영역의 변인에 미치는 평균 효과 크기를 산출한 결과는 <Table 2>와 같다. STEAM 교육이 인지적 영역의 변인인 문제해결력의 효과 크기는 0.592(95% CI 0.432-0.751), 탐구능력의 효과 크기는 0.813(95% CI 0.426-1.201)인 것으로 나타났다. 두 영역 모두 95% 신뢰구간에서 0을 포함하고 있지 않기 때문에 STEAM 교육이 인지적 영역인 문제해결력, 탐구능력에 미치는 효과 크기는 모두 통계적으로 유의한 것으로 나타났다. 효과 크기의 누적분포적 해석방법인 U₃는 문제해결력 72.24의 백분위, 탐구능력 79.10의 백분위로 나타났다. STEAM 교육에 대한 인지적 영역들의 동질성을 검증한 결과 문제해결력(Q=8.4501, df=8, p>.05, I²=5.326), 탐구능력(Q=37.751, df=8, p<.001, I²=78.808)으로 나타나 문제해결력은 효과 크기가 비교적 동질한 것으로 나타난 반면 탐구능력은 효과 크기가

이질한 것으로 나타났다. 즉, 탐구능력은 연구 간 효과 크기가 일관되지 않고 효과 크기 간 차이가 발생하였다.

다. 정의적 영역

STEAM 교육이 정의적 영역의 변인에 미치는 평균 효과 크기를 산출한 결과는 <Table 2>와 같다. STEAM 교육이 정의적 영역의 변인인 인식의 효과 크기는 0.565(95% CI 0.083-1.046), 흥미의 효과 크기는 0.641(95% CI 0.487-0.785), 태도의 효과 크기는 0.696(95% CI 0.565-0.827), 학습동기의 효과 크기는 0.632(95% CI 0.397-0.868), 자아효능감의 효과 크기는 0.628(95% CI 0.353-0.903)인 것으로 나타났다. 영역 모두 95% 신뢰구간에서 0을 포함하고 있지 않기 때문에 STEAM 교육이 정의적 영역인 인식, 흥미, 태도, 학습동기, 자아효능감에 미치는 효과 크기는 모두 통계적으로 유의한 것으로 나타났다. 효과 크기의 누적분포적 해석방법인 U₃는 인식 71.56의 백분위, 흥미 73.89의 백분위, 태도 75.80의 백분위, 학습동기 73.56의 백분위, 자아효능감 73.56의 백분위로 나

타났다. STEAM 교육에 대한 인지적 영역들의 동질성을 검증한 결과 인식($Q=56.725$, $df=8$, $p<.001$, $I^2=85.897$), 흥미($Q=49.986$, $df=24$, $p<.05$, $I^2=51.987$), 태도($Q=76.547$, $df=34$, $p<.001$, $I^2=55.583$), 학습동기($Q=0.453$, $df=3$, $p>.05$, $I^2=0.000$), 자아효능감($Q=2.319$, $df=3$, $p>.05$, $I^2=0.000$)으로 나타나 학습동기, 자아효능감은 효과 크기가 비교적 동질한 것으로 나타난 반면 인식, 태도, 흥미는 효과 크기가 이질한 것으로 나타났다. 즉, 인식, 태도, 흥미는 연구간 효과 크기가 일관되지 않고 효과 크기 간 차이가 발생하였다.

3. 조절변인에 따른 STEAM 교육의 평균 효과 크기

앞서 각 요인에 미치는 효과를 얻은 평균 효과 크기가 메타분석에 포함된 연구의 성질에 따라 다른 효과를 보이는 것으로 확인되었다. 따라서 이처럼 효과 크기가 달리 나타나게 되는 요인, 즉 조절변인에 대한 분석을 실시하였다.

가. 실험집단 인원수

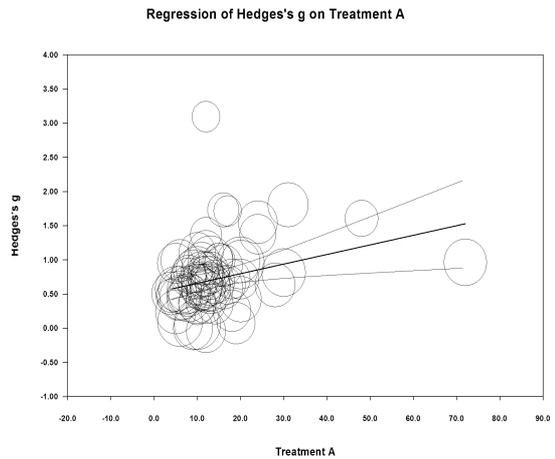
실험집단 인원수에 따라 평균 효과 크기에 유의한 차이가 있는지를 알아보기 위해 메타회귀분석을 실시하였다. 실험집단 인원수는 최소 10명~최대 40명까지이다. 메타회귀분석을 실시한 결과, 회귀모형의 적합성을 나타내는 Model $Q=1.60(df=1, p>.05)$ 이므로 회귀식 자체가 유의한 모형이 아님을 알 수 있었다. 따라서 통계적으로 유의미한 조절변수가 아닌 것으로 나타났다.

나. 실험처치 횟수(수업시수)

실험처치 횟수(수업시수)에 따라 평균 효과 크기에 유의한 차이가 있는지를 알아보기 위해 메타회귀분석을 실시하였다. 프로그램의 실험처치 횟수는 최소 10차시~최대 50차시까지이다.

메타회귀분석을 실시한 결과, 회귀모형의 적합성을 나타내는 Model $Q=9.75(df=1, p<.01)$ 이므로

로 통계적으로 유의미한 모형으로 나타났다. 그리고 조절변수인 횟수의 회귀계수 $b=0.0141$ ($z=3.12, p<.01$)로서 유의한 것으로 나타났다. 메타회귀분석의 결과로 도출한 [Fig. 2]의 regression plot의 결과로 확인 할 수 있었다.



[Fig. 2] Regression plot

조절변수에 의해 설명되는 연구 간 실제 분산의 비율이 $R^2=20\%$ 로 나타났다. 다음으로 조절변인 실험집단의 인원수별, 실험처치 횟수(수업시수) 모두 투입하여 다중 회귀분석을 수행하였다. 다중 회귀분석을 실시한 결과는 회귀모형의 적합성을 나타내는 Model $Q=11.25(df=2, p<.01)$ 이므로 통계적으로 유의미한 모형으로 나타났다. 그러나 조절변인 중 실험처치 횟수(수업시수)만 유의한 것($b=0.0140(z=3.09, p<.01)$)으로 나타나, 결과적으로 수업시수만 효과가 있는 것으로 나타났다.

다. 수업유형

수업유형에 따라 평균 효과 크기에 유의한 차이가 있는지를 알아보기 위해 집단의 크기를 네 집단(교과내, 교과연계, 창체활용, 학교밖)으로 구분하여 메타 ANOVA를 실시하였다. 그 결과는 교과내 수업의 효과 크기는 $0.613(95\% CI 0.462-0.765)$, 교과연계 수업의 효과 크기는

0.597(95% CI 0.399-0.796), 창의활용 수업의 효과 크기는 0.935(95% CI 0.651-1.218), 학교밖 수업의 효과 크기는 0.918(95% CI 0.717-1.119)로 나타났으며, Q 값은 9.312($df=3$, $p<.05$)로 나타나 집단 간의 효과 크기가 동일하다는 귀무가설을 기각할 수 있다. 즉, 수업유형에 따른 평균 효과 크기의 유의한 차이가 확인되었다. 그리고 창의활용, 학교밖 수업이 효과가 높은 수업유형으로 확인되었다.

라. 학급유형

학급유형에 따라 평균 효과 크기에 유의한 차이가 있는지를 알아보기 위해 집단의 크기를 세 집단(영재학급, 일반학급, 특수학급)으로 구분하여 메타 ANOVA를 실시하였다. 그 결과는 영재학급의 효과 크기는 0.722(95% CI 0.364-1.079), 일반학급의 효과 크기는 0.705(95% CI 0.602-0.807), 특수학급의 효과 크기는 1.305(95% CI 0.637-1.973)으로 나타났으며, Q 값은 3.028($df=2$, $p>.05$)로 나타나 집단 간의 효과 크기가 동일하다는 귀무가설을 기각할 수 없다. 즉, 학급유형에 따른 평균 효과 크기의 유의한 차이가 확인되지 않았다.

마. 학교유형

학교유형에 따라 평균 효과 크기에 유의한 차이가 있는지를 알아보기 위해 집단의 크기를 네 집단(유치원, 초등학교, 중학교, 고등학교)으로 구분하여 메타 ANOVA를 실시하였다.

그 결과는 유치원의 효과 크기는 1.140(95% CI 0.734-1.546), 초등학교의 효과 크기는 0.682(95% CI 0.564-0.799), 중학교의 효과 크기는 0.730(95% CI 0.462-0.997), 고등학교의 효과 크기는 0.716(95% CI 0.435-0.996)으로 나타났으며, Q 값은 4.532($df=3$, $p>.05$)로 나타나 집단 간의 효과 크기가 동일하다는 귀무가설을 기각할 수 없다. 즉, 학교유형에 따른 평균 효과 크기의 유의한 차이가 확인되지 않았다.

바. 중심교과

중심교과에 따라 평균 효과크기에 유의한 차이가 있는지를 알아보기 위해 집단의 크기를 네 집단(과학, 수학, 기술, 예술)으로 구분하여 메타 ANOVA를 실시하였다. 그 결과는 과학중심교과의 효과 크기는 0.646(95% CI 0.517-0.776), 수학중심교과의 효과 크기는 0.615(95% CI 0.308-0.921), 기술중심교과의 효과 크기는 0.773(95% CI 0.480-1.065), 예술중심교과의 효과 크기는 0.839(95% CI 0.573-1.106)으로 나타났으며, Q 값은 6.356($df=4$, $p>.05$)로 나타나 집단 간의 효과 크기가 동일하다는 귀무가설을 기각할 수 없다. 즉, 중심교과에 따른 평균 효과 크기의 유의한 차이가 확인되지 않았다.

IV. 결론 및 제언

우리나라에서는 STEAM 교육을 통해 세계적 과학기술 인재 양성을 위한 목적으로 모든 학교급을 포함하는 교육 정책으로 제시되고 있다. 본 연구는 STEAM 교육과 관련된 선행연구와 STEAM 교육의 효과를 분석한 연구를 통해 보다 객관적이고 종합적인 의미 있는 일반적 결과를 도출하고자 메타분석을 실시하였다. 2010년 이후의 연구물을 분석 대상으로 선정한 이유는 우리나라에 STEAM 교육이 2010년부터 본격적으로 도입되었고, 또한 관련 논문이 양적으로 증가한 시점이라고 판단하였기 때문이다. STEAM 교육의 메타 분석 결과와 논의를 통해 얻어진 결론은 다음과 같다.

첫째, 본 연구의 핵심인 메타 분석을 통한 STEAM 교육의 평균 효과 크기 결과에 따르면 평균 효과($g=0.719$)는 그 실제적 유용성이 입증되었으며, 중간 정도의 효과 크기를 나타내었다.

구체적으로 창의적 사고기능과 창의적 사고성향, 문제해결력과 탐구력, 인식, 흥미, 태도, 학습동기, 자아효능감 등에서 통계적으로 유의한 것으로 확인되었다. 출간 오류를 검토한 결과 또한

산출된 효과 크기가 안정적이었음을 고려하면, 현재까지 이루어지고 있는 STEAM 교육은 긍정적인 효과가 있다는 시사점을 얻을 수 있다.

둘째, STEAM 교육이 종속 변인에 미치는 평균 효과 크기는 창의적 사고기능과 창의적 사고성향, 문제해결력, 탐구능력, 인식, 흥미, 태도, 학습동기, 자아효능감 등 전 영역에서 모두 효과적임을 확인할 수 있었다. 창의성 영역에서는 창의적 사고기능 보다는 창의적 사고성향의 효과가 큰 것으로 나타났고, 인지적 영역에서는 문제해결력 보다 탐구능력이 큰 것으로 나타났다. 그리고 정의적 영역에서는 태도가 가장 큰 효과 크기를 나타냈다. 다만, 정의적 영역에서 태도에 관한 연구가 과학과 중심으로 이루어져 STEAM 교육에서 요구하는 태도나 역량을 잘 반영하지 못하고 있었다. 또, STEAM 교육은 동질성 검증에 의해 효과 크기의 차이 및 그 이질성이 발생하였고, STEAM 교육의 효과 크기의 차이 및 이질성은 구체적으로 창의성 영역의 창의적 사고기능과 창의적 사고성향, 정의적 영역의 태도, 흥미 등에서 발생한다고 판단 할 수 있었다. 따라서 앞으로 STEAM 교육 프로그램의 개발을 위해 이러한 요인을 고려할 필요가 있다.

셋째, STEAM 교육의 효과 크기의 차이 및 그 이질성에 대한 원인을 알아보기 위하여 조절변인에 따른 STEAM 교육의 평균 효과를 탐색한 결과 실험처지 횡수와 수업유형에 대한 변인이 연구 간 효과 크기의 차이를 설명하는 중요한 조절변인으로 나타났다. 실험처지 횡수 즉, 수업 시수가 증가할수록 STEAM 교육의 효과가 분명하게 나타나는 결과는 교육의 대상이 교육내용을 받아들이고 변화하는 데는 충분한 시간이 필요함을 의미한다. 또, 수업유형에서 창체활용형 수업이 가장 좋은 수업 유형으로 나타나고 있으나 특정 유형을 제시하기 보다는 학교급별로 지역적 특성, 기반 조성의 양상, 연계 기관의 유무와 형태, 학교 규모와 재정상태, 행정적 지원 등 실정에 맞도록 다양한 유형의 활동으로 시행하되, 교과

서와 교육과정 중심의 활동 보다는 보다 자율적이며 창의적인 형태의 교육활동으로 전환하여야 할 것이다. STEAM 교육의 효과에 대한 메타분석 결과는 이처럼 비교적 명확한 결론을 제시할 수 있음에도 불구하고, 본 연구의 제한점 및 논의 결과를 중심으로 후속연구를 위한 제언을 제시하면 다음과 같다.

첫째, 본 연구에서 인지적 영역인 학업성취도를 학교급, 학생 및 학교 특성이 반영되어 있어 그 타당도가 부족하다고 판단되어 연구대상에서 제외하였다. 후속연구에서는 이 부분을 보완하여 제시해야 할 필요가 있으며 STEAM 교육의 적용 후 남학생이 여학생 보다 과학에 대한 태도와 흥미에서 향상도가 높았다(Park jee-hyeon, 2014)는 연구도 있어 다른 조절변인을 통해 이를 더 심도 있게 탐색해 볼 필요가 있다. 둘째, 메타분석을 위해 본 연구에서 선정된 논문 중 학술논문의 대부분이 학위논문을 게재한 것은 아쉬운 부분이다. 뿐만 아니라 메타분석에서 이상치나 출간 오류의 분석에서는 실험처지의 효과가 긍정적인 연구가 출판된 가능성이 많은 반면 실험처지의 효과를 밝히지 못했거나 부정적인 효과를 얻은 연구물의 출판 가능성이 낮음을 경계해야 한다(Higgins, J. P. et al, 2011). 따라서 연구자들은 연구 설계와 진행과정을 명확히 진술하고 논문을 이해하는 데 필요한 자료가 누락되지 않도록 최대한 상세하게 기술하는 등의 노력이 필요하다. 셋째, 본 연구는 STEAM 교육의 효과성을 증명하는데 초점을 맞추고 있어 난점이 있다. 교육은 인간 내부에서 일어나는 복합적인 현상으로 가설과 검증이라는 단순한 흐름으로 증명하는데 한계가 있으며, 교육 대상이 교육내용을 받아들이고 충분히 무르익을 시간이 필요하다. 또한, 교육은 교사와 학생들의 공식적인 교육과정의 목적 이외에도 수업을 둘러싼 교육적, 사회적 맥락을 통해 수많은 변수가 존재하기 때문에 이를 종합적으로 관찰하기 위해서는 직접 그 안으로 들어가 현상을 심층적으로 연구해 볼 필요가 있다.

References

- An, hae-ran(2015). Analysis of Research Trends in STEAM education for the Gifted.. ajou university.
- Borenstein, M. · Hedges, L. V. · Higgins, J. P. & Rothstein, H. R.(2009). Introduction to Meta-Analysis. Chichester, UK: Wiley.
- Duval, S. & Tweedie, R.(2000). A nonparametric 'trim and fill' method of accounting for publication bias in meta-analysis, Journal of the American Statistical Associaton, 95, 89-98.
- Guzzo, R. A. · Jackson, S. E. & Katzell, R. A. (1987). Meta analysis. In L. L. Cummings, & B. M. Staw (Eds.). Research in organizational behavior. Stanford: Elevier Ltd.
- Higgins, J. P. & Green, S.(Eds.)(2011). Cochrane Handbook for Systematic Review of Interventions Version 5.1.0, The Cochrane Collaboration. Retrieved from www. cochrane-handbook.org.
- Korea Foundation for the Advancement of Science&Creativity(2012). STEAM education basic research for establishing the direction of running.
- Lee, gyeong-eon · Lee, gwang-woo · Kim, hyeon-mi · Lim, soe-ha(2010). Reorganization Study courses for enhancing creativity. (research report RRC). Korea Institute of Curriculum and Evaluation.
- Lee, ji-hae(2014). The analysis on domestic research trends of STEAM education of Mathematics. ajou university.
- Ministry of Education, Science and Technology. (2010). Republic of Korea driven by creative person and advanced technology. 2011 business report.
- Oh, hyun-sook(2012). STEAM education program development strategies for secondary schools based on the study of convergence and integrated education. dankook university.
- Oh, soon-moon(2011). A meta-analysis on the relationship between the principal's leadership and school organizational effectiveness. department of education graduate school korea university.
- Oh, Sung-Sam(2009). Theory and practice of meta-analysis. Konkuk University Press.
- Park, do-soon(2010). Educational Evaluation- Understand and apply. Education and Science Press.
- Park, jee-hyeon(2014). The effects of STEAM education on the attitude, interest level, and academic achievement of elementary 6th graders with regards to science.
- SHIN, jin-kung · CHOI, Dong-Kyu · KIM Ji-Won · HEO, Gyun · PARK, Jong-Un · JU, Dong-Beom · WON, Hyo-Heon(2013). An analysis of core competence and core element on the STEAM program in the research report of school, JFMSE, 25(4), 898~914.
- Torrance, P.(1992). The Torrance tests of creative thinking. Bensenville, IL: Scholastic Testing Services.

-
- Received : 09 September, 2016
 - Revised : 04 October, 2016
 - Accepted : 10 October, 2016