

실험실습법에 의한 수행평가가 중학생의 과학성취도 및 정의적 영역에 미치는 영향

이지현¹ · 남정희 · 문성배

¹(김해 영운중학교) · (부산대학교 사범대학 화학교육과)

The Effects of a Performance Assessment Based on the Experimental Practice on Student's Science Achievement and Affective Domain in the Middle School Science

Lee, Jihyun¹ · Nam, Jeonghee · Moon, Sungbae

¹(Young-Woon Middle School) · (Pusan National University)

ABSTRACT

The purpose of this study was to examine the effects of an application of the experimental practice which is one of the performance assessments on student's science achievement and affective domain in the middle school science. For this study, two classes were chosen from a co-ed middle school in Gimhae, and assigned to the experimental and the control groups. Prior to the instructions, a science achievement test and a affective domain test were administered. After the instructions, post-tests which were similar to the pretests were also administered.

The results showed that the performance assessment based on an experimental practice was more effective of enhancing students' science achievement. However, there was no significant difference in an affective domain test between the two groups. For the control group, there was a significant difference between the pretest and post test in affective domain. It means that there was a negative change in affective domain of the control group.

Key words: performance assessment, an experimental practice, science achievement, affective domain

I. 서 론

이제까지 우리나라는 수 차례 교육과정을 개정해 왔으며, 이러한 교육과정의 개정과 더불어 평가방법

의 변화도 시도해 왔다. 그럼에도 불구하고 우리나라의 학교교육은 지식위주의 주입식 암기 교육, 입시위주의 교육이라는 비판을 받고 있다. 이러한 문제는 근본적으로는 교수·학습과정의 문제에 기인하지만

*2002.9.24(집수) 2002.12.6(최종 통과)

동시에 평가의 문제에 기인한다고도 볼 수 있다.

지금까지 우리나라의 학교교육에서 평가는 학생 선발의 목적과 기능을 수행해 왔으며, 이에 따라 학생들을 서열화하는 상대 평가에만 치중하였다. 이를 위해 주로 선택형 시험문항을 이용하여 결과중심의 성취도 평가를 행하여 왔다. 그러나 이러한 형태의 평가방법에 의해서는 학생의 창의성이나 문제해결력, 과학적 의사소통능력 등 고등사고 기능의 평가가 어렵다고 비판받아왔다(김성훈, 1993; 백순근, 1995; 허경철, 1988; Ebenezer & Haggerty, 1999). 특히 학교교육에서 이러한 선택형 문항의 객관식 평가는 과학 지식을 측정하는데는 용이하지만, 탐구 기능 및 문제해결력이나 태도를 측정하는데는 거의 불가능하기 때문에(Herman et al., 1992) 오래 전부터 탐구능력 평가를 위한 실험 실기 평가 또는 실험 보고서 평가와 같은 이름으로 수행평가를 실시하여 왔다. 그러나 실험평가의 경우에 교사가 제시한 실험방법을 기계적으로 따라하거나 교과서에 제시된 방법을 그대로 따라 하고 결과를 확인하는데 그치고 있으며 학생 스스로 실험을 계획하고 실행하여 결론을 도출해 가는 수업이 되지 못하고 있기 때문에 평가 또한 결과 중심으로 이루어 졌다(윤동영, 1999).

이와 같은 현행의 학교교육을 개선하기 위해서 교수·학습 방법 및 평가가 개선되어야 한다는 주장에 대해서는 많은 사람들이 동의하고 있다. 즉 앞으로의 교수·학습 방법 및 평가는 전인 교육을 신장시킬 수 있고 학생 개개인의 소질과 개성을 중시함으로써 시대가 요구하고 있는 창의성이나 문제해결력 등 고등사고 기능을 신장시킬 수 있는 방향으로 변해가야 한다는 것이다. 따라서 교수·학습 과정의 질적인 측면을 평가하고 그 과정의 개선을 도움으로써 교사 개개인의 자질을 향상시키고 학생 개개인의 교육적 성장을 돕기 위해 전통적인 교수·학습평가의 새로운 대안을 제시하고 있는데 그 중에 하나가 바로 수행평가이다.

수행평가의 필요성은 학습 및 학습자에 관한 구성주의 이론에도 바탕을 둔다(국립교육평가원, 1996). 구성주의에 따르면, 학습은 학습자가 이미 가지고 있는 인지구조의 변화이며, 그 과정은 사회적 합의의

과정이다(조희형, 1998). 따라서 이러한 구성주의 학습이론에 따라 평가도 교수-학습 과정의 통합적 구성요소로 포함되어야 한다(Doran et al., 1998). 이러한 관점에서 볼 때 수행평가는 전통적인 평가보다 더 다양한 평가방법을 적용하기 때문에 고차원적인 사고력과 기능을 평가할 수 있으므로 교수-학습과정의 통합적인 요소가 될 수 있다(Hibbard, 2000).

수행평가의 구체적인 방법으로는 서술형 주관식 검사법, 논술형 검사법, 구술시험법, 실기시험법, 실험실습법, 관찰법, 포트폴리오법 등이 있으며(박승재와 조희영, 1995), 그 중에서도 과학과에서는 실험실습법이 많이 활용되고 있다. 현재 과학 교육에 있어서도 수행평가의 방법이 도입되어 학교현장에서 실시되고 있다. 그러나 많은 교사들이 수행평가의 장점을 인정하고 있음에도 불구하고 일선 학교에서는 아직도 수업의 종결 단계에서만 인지적 목표를 겨냥한 결과 위주의 교사 주도적이고 지필 교사 위주의 단편 평가가 실시되고 있어 수행평가가 제대로 자리잡아가고 있지 못한 실정이다(백순근, 1998).

또한 기존에 이루어졌던 수행평가에 관한 연구들이 대체로 초등학교에 관한 연구들이 대부분이고(윤동영, 1999, 왕경순, 1998, 엄기영, 1999, 김영순, 1992), 중학교 과학에서 수행평가 관련 연구들이 일부 진행되어 왔으나 대부분 수행평가 도구개발에 관한 연구들이다. 따라서 수행평가가 실제 중학교 과학수업에 미치는 영향을 구체적으로 알아 볼 필요가 있다고 본다.

따라서, 이 연구에서는 다양한 수행평가의 방법 중 실험실습법을 이용한 수행평가를 중학교 과학수업에 적용하여 학생들의 과학성취도와 정의적 영역에 미치는 영향을 알아보고자 한다.

II. 연구방법

1. 연구 대상

경남 마산시 소재 중학교 1학년 2개 학급 82명을 연구 대상으로 선정하였으며, 임의로 각 학급을 실험 집단과 비교집단으로 구분하였다. 또한 각 학급을 사

전 과학성취도 검사의 결과에 따라 상·중·하의 세 그룹으로 나누었다.

2. 검사 도구 및 방법

1) 사전 과학성취도 검사지

사전 과학성취도 검사지는 실험집단과 비교집단의 동질집단 여부와 집단 내 이질 소집단을 편성을 위한 자료로 활용하기 위하여 연구자가 근무하고 있는 학교의 과학교사들과의 협의를 거쳐 개발하였다. 검사지는 총 20문항으로, 중학교 2학년 1학기에 학습한 과학내용의 이해정도를 알아보기 위한 문항으로 구성되었다. 실험집단과 비교집단 모두 수업처치에 들어가기 전에 사전 과학성취도 검사를 실시하였다.

2) 사후 과학성취도 검사지

사후 과학성취도 검사지는 실험집단에 실험처치를 한 후 실험·실습법에 의한 수행평가가 학생들의 과학성취도에 미치는 영향을 알아보기 위하여, 사전 과학성취도 검사지의 문제와 비슷한 난이도로 연구자가 근무하고 있는 학교의 과학교사들과의 협의를 거쳐 개발하였다. 검사지는 총 20문항으로, 중학교 2학년 2학기에 학습한 과학내용의 이해정도를 알아보기 위한 문항으로 구성되었다. 실험집단과 비교집단 모두 수업처치 이후에 사후 과학성취도 검사를 실시하였다.

3) 정의적 영역 검사지

정의적 영역에 대한 검사를 위해 한국교육개발원에서 개발한 과학과에 대한 학습태도 검사지를 사용하였다. 검사지의 신뢰도(Cronbach's α)는 .865이었다. 검사지는 자아개념, 학습에 대한 태도, 학습습관의 3개 하위 영역으로 구성되어 있으며, 각각 10문항, 15문항, 15문항의 총 40문항으로 구성되었다. 문항은 5단계 Likert 척도로 되어 있다.

4) 수행평가 도구 개발 및 수업 처치

실험실습법에 의한 수행평가를 위해 중학교 1학년 과학의 화학단원을 중심으로 교과서에 제시된 실험을 탐구적인 형태의 실험으로 재구성하였다. 수업의 처

치는 개발된 실험을 학생들에게 배부한 후 학생들이 실험을 하는 동안 평가를 실시하는 수행평가 형식으로 실시하였다. 수행평가는 개발된 10개 실험에 대하여 총 10회 실시하였으며, 수업은 45분을 1차시로 하여 실험의 내용에 따라 1차시 내에 평가가 이루어지기도 하고 때로는 2차시에 걸쳐 평가가 이루어지기도 하였다.

수행평가를 실시하기 전 미리 수행평가의 척도표를 제시하여 학생들이 평가의 내용과 항목, 평가 시간 등을 미리 알고 평가에 임하도록 하였다. 평가의 방법은 주로 학생들이 개인별로 과학실의 한 테이블에 한 명이 앉아 스스로 실험을 수행하고 연구자가 척도표를 따라 그 과정을 평가하였다. 평가는 실험에 대한 각 평가의 요소를 Likert 척도로 만들어 연구자가 직접 관찰하여 평가하였다. 학생들의 협동심을 향상시키기 위해 때로는 2-3명이 짝이 되어 실험을 수행하거나 조별로 실험을 수행하고 동일한 방법으로 평가하였다.

비교집단도 실험집단과 동일한 탐구적인 형태로 재구성한 실험을 수행하였다. 그러나 비교집단에는 실험을 하는 동안 실시하는 수행평가는 실시하지 않았으며 실험 후에 실험의 결과만을 평가하였다.

III. 연구결과 및 논의

1. 과학성취도에 미치는 영향

실험실습법을 이용한 수행평가가 학생들의 과학성취도의 향상에 미치는 영향을 알아보기 위하여 수업처치 후에 실험집단과 비교집단에 사후 과학성취도 검사를 실시하였다.

1) 실험집단과 비교집단의 사전 과학성취도 비교
사후검사 결과의 비교에 앞서 두 집단이 동질 집단 인지를 알아보기 위해 실시한 사전 과학성취도 검사의 결과를 비교하였다. t-검증 결과 두 집단 사이에는 통계적으로 유의미한 차이가 없는 것으로 나타났다 ($p>.05$). 따라서 두 집단은 동질집단으로 간주하였다 (Table 1).

2) 실험집단과 비교집단의 사후검사 비교

수업처치 후 실시한 사후 과학성취도 검사 결과를 이용한 t-검증 결과 실험집단의 평균이 비교집단의 평균보다 통계적으로 유의미하게 높은 것으로 나타났다(p<.05). 이것은 실험실습법에 의한 수행평가의 실시가 학생들의 과학성취도의 향상에 효과가 있음을 시사한다고 할 수 있다.

3) 과학성적에 따른 실험집단과 비교집단의 사후 과학성취도 비교

실험실습법에 의한 수행평가가 과학성취도 상·중·하위 집단에 따라 사후 과학성취 정도에 차이가 있는지를 검증하기 위하여 실험집단과 비교집단을 사전 과학성취도 검사 결과에 의해 상·중·하의 세 집단으로 나누어 ANOVA 분석을 하였고, 그 결과를 Table 3에 제시하였다.

Table 3에서 보듯이 사전검사에서는 과학성취도에 따른 상·중·하위의 세 집단에서 실험집단과 비교집단 사이에 통계적으로 유의미한 차이가 없음을 보여 주고 있다(p>.05).

Table 1. Results of t-test on the scores of the previous-science achievement test by group

Source	Group	Number	Mean	Std. dev.	t	p
Previous science achievement	Control	41	9.1829	4.5080	.362	.719
	Experimental	41	9.0000	3.9576		

Table 2. Results of t-test on the scores of the post-science achievement test by group

Source	Group	Number	Mean	Std. dev.	t	p
Post science achievement	Control	41	11.9024	4.3520	-2.227,362	.032
	Experimental	41	13.2439	4.1548		

Table 3. ANOVA results on the gain scores of science achievement test by the students' science achievement levels

Source	Level	Source of variation	SS	df	MS	F	p
Previous test	High	Between levels	3,471	1	3,471	.555	.464
		Within levels	150,192	24	6,258		
	Medium	Between levels	3,571	1	3,571	.548	.466
		Within levels	169,393	26	6,515		
	Low	Between levels	5,143	1	5,143	2.813	.105
		Within levels	47,536	26	1,828		
Post test	High	Between levels	.962	1	.962	.212	.650
		Within levels	109,077	24	4,545		
	Medium	Between levels	54,321	1	54,321	11,904	.002
		Within levels	118,643	26	4,563		
	Low	Between levels	4,321	1	4,321	.548	.466
		Within levels	204,929	26	7,882		

실험처치를 한 후 두 집단의 사후 학업 성취도 검사의 점수를 과학성취도 상·중·하의 집단별로 분석한 결과, 두 집단의 중위 그룹에서만 통계적으로 유의미한 차이를 보여주었다($p < .05$).

위의 결과로부터 볼 때 실험실습법에 의한 수행평가가 실험집단 전체 학생에 대해서는 비교집단에 비하여 학생들의 과학성취도 향상의 정도에서 유의미한 차이를 보였으나 과학성취 집단별로 비교했을 때는 중위 집단의 학생들만이 사후 과학성취도에서 유의미한 차이가 나타났고, 상위집단과 하위집단의 학생들의 경우에는 비교집단과 실험집단 사이에 유의미한 차이가 없었다.

2 정의적 영역에 미치는 영향

정의적 영역 검사지는 자아개념, 학습에 대한 태도, 학습습관의 3개 하위 영역으로 구성되어 있으며, 정의적 영역의 점수는 각 문항의 답에 해당하는 점수의 합으로 구하였다. 각 문항의 답은 1번 항상 그렇다, 2번 대체로 그렇다, 3번 보통이다, 4번 대체로 그렇지 않다, 5번 전혀 그렇지 않다고 되어 있으며 각 번호를 점수로 하여 그 합을 구하되 부정적 질문은 역으로 하여 계산하였다. 그러므로 정의적 영역에서는 점수가 낮을수록 바람직하다고 본다.

1) 실험집단과 비교집단의 전체 비교

실험집단과 비교집단 사이에 실험실습법에 의한 수행평가로 인하여 정의적인 영역에서 차이가 있는가를 알아보기 위하여 수업 전, 후에 실시한 정의적 영역 검사에 대한 t-검증 결과를 Table 4에 나타내었다.

사전검사에서는 두 집단 사이에 평균점수의 차이가 거의 없었으나 사후검사에서는 실험집단의 평균이 108.32로 비교집단의 평균 123.29에 비해 낮게 나타났다. 이것은 실험집단의 학생들이 실험실습법에 의한 수행평가 수업을 받은 후 정의적 영역에서 비교집단 보다는 바람직한 태도를 나타낸 것으로 볼 수 있다. 그러나 사후검사에 대한 t-검증 결과 두 집단 간에 통계적으로 유의미한 차이는 없는 것으로 나타났다($p > .05$). 따라서 이 연구에서 적용한 실험실습법을 이용한 수행평가방법이 전통적인 실험에 의한 수업과 비교해 볼 때 정의적인 영역에 영향을 미치지 않았음을 알 수 있다.

2) 비교집단과 실험집단의 각각에 대한 사전·사후 검사 비교

정의적 영역에 대한 항상 정도를 더욱 구체적으로 알아보기 위해 비교집단과 실험집단의 사전·사후 검사를 각각 t-검증으로 분석해 보았다.

Table 5에서 보듯이 비교집단의 정의적 영역의 점

Table 4. Results of t-test on the scores of the previous and the post affective domain test by group

Source	Variable	Frequency	Mean	Std.dev.	t	p
Previous affective domain	Control	41	114.317	21.201	-.044	.965
	Experimental	41	114.488	21.382		
Post affective domain	Control	41	123.293	22.410	3.03	.085
	Experimental	41	108.317	22.238		

Table 5. Results of t-test on the scores of the affective domain tests for the control group

Source	Number	Mean	Std.dev.	t	p
pre-test	41	114.317	21.201	-2.911	.006
post-test	41	123.293	22.410		

수는 평균이 수업처치 전 114.317에서 수업처치 후 123.293으로 점수가 올라가 더 부정적으로 변화했으며 그 차이는 통계적으로 유의미하였다($p < .05$).

실험집단의 경우는 Table 6에서 알 수 있듯이 수업처치 전 평균이 114.49에서 수업처치 후 108.32로 더 긍정적으로 변화하였으나 통계적으로는 유의미한 차이를 보이지 않았다($p > .05$). 그러나 수업처치 후 비교집단의 정의적 영역은 부정적으로 변했음에 반해 실험집단은 비록 유의미하지는 않지만 긍정적으로 변화하였음을 보여주고 있다.

2) 실험집단과 비교집단의 상·중·하 집단별 비교

실험실습법을 이용한 수행평가가 과학성적 상, 중, 하 집단에 따라 정의적 영역에 미치는 영향에 차이가 있는지를 알아보았다. 이를 위하여 정의적 영역의 하위요소를 자아개념, 학습에 대한 태도, 학습습관으로 나누고 과학성적 상·중·하 집단으로 나누어 비교·

분석하였다. 과학성적별 정의적 점수가 실험집단과 비교집단에서 각각 어떤 차이를 보여주는지에 대한 ANOVA 결과를 Table 7에 제시하였다.

사전검사에서는 실험집단과 비교집단 사이에 과학성적 집단에 따라 통계적으로 유의미한 차이를 보이지 않았으나 사후검사의 결과 성적이 중위권 학생과 하위권 학생들에게 통계적으로 유의미한 차이를 보여주고 있다($p < .05$). 따라서 실험실습법에 의한 수행평가의 적용이 과학성적이 상위권인 학생들의 정의적 영역에는 큰 영향을 미치지 못하였지만, 과학성적이 중·하위권인 학생들에게는 영향을 주었다고 볼 수 있다.

3) 실험집단의 사전·사후검사 결과 비교

실험집단에 대해 수업 전 후에 실시한 정의적 영역에 대한 점수 평균 및 표준편차와 t-검증 결과를 정의적 영역의 각 하위요소에 따라 Table 8에 제시하

Table 6. Results of t-test on the scores of the affective domain tests for the experimental group

Source	Number	Mean	Std.dev.	t	p
pre-test	41	114.488	21.382	1.764	.085
post-test	41	108.317	22.238		

Table 7. ANOVA results on the gain scores of the affective domain test by the students' science achievement levels

Source	Level	Source of variation	SS	df	MS	F	p
Pre-test	High	Between levels	113.23	1	113.28	1.28	.261
		Within levels	6724.72	24	88.48		
	Medium	Between levels	2.92	1	2.92	.03	.863
		Within levels	7941.47	26	98.04		
	Low	Between levels	22.85	1	22.85	.245	.619
		Within levels	7425.76	26	91.68		
Post-test	High	Between levels	193.96	1	193.96	1.45	.230
		Within levels	10054.87	24	132.30		
	Medium	Between levels	505.19	1	505.19	4.62	.035
		Within levels	8967.10	26	109.36		
	Low	Between levels	685.71	1	685.71	5.18	.026
		Within levels	10865.43	26	132.51		

Table 8. Results of t-test on the sub-domains of the affective domain tests for the experimental group

Sub-domain	Source	Mean	Std.dev.	t	p
Self-consciousness	pre-test	32.49	8.47	2.77	.01
	post-test	29.10	8.22		
Attitudes towards learning	pre-test	36.10	9.05	.82	.42
	post-test	34.66	9.67		
Learning attitude	pre-test	45.71	9.23	.08	.08
	post-test	45.61	9.76		

Table 9. Gain score and standards deviations of the affective domain test by the students' science achievement levels

Level	Gain score	Std. dev.
high	4.846	15.410
medium	3.714	18.837
low	12.571	27.729
Total	7.098	21.274

였다.

Table 8에서 알 수 있듯이 정의적 영역의 하위요소 중 자아개념에 대한 실험집단의 사전검사 평균이 32.48이고, 사후검사 평균이 29.10으로 수업 후 긍정적으로 변화했으며 t-검증 결과, 두 검사 사이에 통계적으로 유의미한 차이가 있는 것으로 나타났다 ($p < .05$). 그러나 학습에 대한 태도와 학습습관에 대해서는 수업전후 두 검사에서의 평균점수의 차이가 통계적으로 유의미한 차이를 보여주고 있지 않다 ($p > .05$). 따라서 수행평가의 실시가 학생들의 자아개념의 변화에는 긍정적인 영향을 미치지만 학습에 대한 태도와 학습습관에는 유의미한 영향을 미치지 않는다고 말할 수 있다. 이러한 결과는 자아개념은 비록 짧은 기간의 수업처치에 의해서도 긍정적으로 변할 수 있으나 학습습관이나 학습에 대한 태도는 더 장기적인 수업처치가 필요함을 시사해준다.

4) 과학성적에 따른 실험집단의 향상 정도 비교
실험실습법에 의한 수행평가의 적용이 실험집단 학생들의 과학성취도 상·중·하위 집단에 따라 정의적

영역의 향상 정도에 차이가 있는가를 알아보았다. 실험 처치 전·후에 실시한 정의적 영역 검사의 향상점수와 표준편차는 Table 9에 제시하였으며 이에 대한 분산분석의 결과를 Table 10에 제시하였다.

Table 9에서 보듯이 과학성취도 상위집단과 중위집단의 향상점수에 비해 하위 집단의 평균이 12.517로 상위나 중위 집단의 학생들보다 더 큰 향상을 보이는 것으로 나타났다. 이것은 성적이 하위인 학생들이 그들의 지적인 수준에 대한 평가보다 실험의 과정에 대한 평가에 흥미를 가졌으며, 집중력이 약하고 산만하여 강의식 수업을 지루하게 느꼈던 학생들이 구체적인 실험의 상황을 평가받음으로써 자신들도 할 수 있다는 자신감을 주었던 것으로 생각된다.

과학성적 집단별 정의적 영역검사의 향상 점수에 대한 분산분석의 결과 집단별로는 통계적으로 유의미한 차이가 나타나지 않았다($p > .05$). 별도로 제시하지는 않았으나 정의적 영역의 하위요소인 자아개념, 학습에 대한 태도, 학습습관의 세 가지 영역에서 과학성적 상·중·하 집단별로는 유의미한 향상의 차이는 나타나지 않았다($p > .05$).

Table 10. ANOVA result of the gain scores of the affective domain test by the students' science achievement levels

Source of variation	SS	df	MS	F	p
Between levels	645.632	2	322.816	.703	.502
Within levels	17457.978	38	459.420		
Total	18103.610	40			

IV. 결론 및 제언

현재 학교 현장에서는 수행평가를 강조하고 있으며 다양한 방법의 수행평가를 권장하고 있다. 이 연구에서는 수행평가의 한 방법으로 실험실습법에 의한 수행평가가 중학교 학생들의 과학성취도와 정의적 영역에 미치는 영향을 알아보았다.

실험실습에 의한 수행평가를 적용한 실험집단이 전통적인 실험을 실시한 비교집단에 비해 과학성취도에서 통계적으로 유의미한 향상을 나타내었다($p < .05$). 사후 과학성취도를 학생들의 사전 과학성취도에 의해 상·중·하 집단으로 구분하여 분석한 결과, 중위권 학생들이 통계적으로 유의미한 향상을 나타내었다. 이러한 결과는 실험실습법에 의한 수행평가가 과학성취도가 그리 높지 않은 학생들의 과학개념 이해에 도움이 되나 과학성취도가 높은 학생들과 매우 낮은 하위권 학생들에게는 효과적이지 않았음을 보여준다. 이러한 결과가 나온 것은 성적이 중위권인 학생들이 실험·실습법의 수행평가를 통해 수업의 내용을 더욱 잘 이해하고, 실험과정을 평가하는 방법에 대해 흥미를 가지고 잘 참여한 영향도 있으리라고 생각된다. 반면 성적이 상위권인 학생들에게는 실험실습법에 의한 수행평가가 그들의 과학성취도를 크게 향상시킬 만큼의 자극을 주지 못하였으며, 하위권인 학생들은 실험실습법에 의한 평가에는 열심히 참여하고 흥미를 가졌으나 이것이 과학개념이해에 까지는 영향을 미치지 못한 것이라고 판단된다. 따라서 다양한 능력 집단의 학생들에게 적용할 수 있는 다양한 수행평가 방법이 필요함을 시사해준다.

정의적 영역에서는 실험집단과 비교집단의 사전·

사후 검사 비교에서 통계적으로 유의미한 차이가 없었다. 그러나 두 집단 각각의 사전·사후 검사 비교에서는 실험집단은 통계적으로 유의미하지는 않지만 ($p > .05$) 긍정적으로 변화를 보인 반면, 비교집단은 통계적으로 유의미하게 부정적인 방향으로 변화하였다 ($p < .05$ 보). 이는 학생들이 초등학교에서 중학교에 입학하여 과학내용이 점점 어려워짐에 따라 학습을 어려워하고 과학에 대한 부정적인 견해가 많아지는 것으로 생각된다.

실험집단과 비교집단의 과학성취도에 따른 정의적 영역의 비교분석에서는 중·하위권 학생들의 경우 실험집단이 비교집단보다 긍정적인 효과를 나타냈으며 ($p < .05$), 실험집단의 경우 정의적 요소의 하위요소 중 자아개념에서만 수업처치 후 통계적으로 유의미한 변화를 보여주었다($p < .05$). 실험집단에서 과학성취도가 상·중·하인 집단에 대한 정의적 영역 점수의 향상 정도의 분석 결과 집단별로는 향상정도가 유의미한 차이를 나타내지 않았다($p > .05$).

따라서 실험실습법에 의한 수행평가의 실시가 실험집단의 정의적 영역 중 자아개념을 긍정적으로 변화시키는데 영향을 미쳤고, 학습에 대한 태도와 학습습관을 변화시키는데는 영향을 미치지 못했다고 할 수 있다. 이러한 결과는 자아개념은 비록 짧은 기간의 수업처치에 의해서도 긍정적으로 변할 수 있으나 학습습관이나 학습에 대한 태도는 더 장기적인 수업처치가 필요함을 시사해준다.

위 결과에서 보는 바와 같이 실험실습법에 의한 수행평가의 실시가 중학생의 과학성취도와 정의적 영역에 부분적이지만 긍정적인 영향을 미친다고 할 수 있다. 따라서 앞으로 현장에서는 수행평가가 자리잡을

수 있도록 교사들의 노력이 요구되며 이에 대한 연구가 더 많이 이루어져야 한다고 생각된다.

적 요

이 연구는 현재 학교에서 실시되고 있는 수행평가가 중학생들의 과학성취도와 정의적 영역에 어떠한 영향을 미치는지를 알아보는데 목적이 있다. 수행평가의 다양한 방법 중에서도 과학과에서 가장 많이 실시되고 있는 실험실습법에 의한 수행평가의 방법이 중학교 과학학습에 미치는 영향을 알아보았다.

연구 결과로부터 볼 때, 과학성취도의 평균이 동일한 두 집단에서 수행평가를 적용한 실험집단의 과학성취도 점수가 유의미하게 향상되었으며, 특히 사전 과학성취도가 중, 하위권인 학생들의 과학성취도에 통계적으로 유의미하게 영향을 미쳤다. 상위권 학생은 비록 통계적으로 유의미하지는 않지만 성적이 약간 향상되었다. 정의적 영역에서는 실험집단의 사전·사후 검사에서 유의미한 차이는 없었다. 그러나 실험집단은 통계적으로 유의미하지는 않지만 긍정적으로 변화를 보인 반면, 비교집단은 통계적으로 유의미하게 부정적인 방향으로 변화하였다. 정의적 영역에 대한 실험집단의 사전·사후검사 비교에서 정의적 영역의 하위요소 중 자아개념과 학습습관의 변화는 유의미하게 상승하였으며, 학습에 대한 태도의 변화에는 효과가 없었다. 실험집단의 과학성취도가 상·중·하인 집단별 정의적 영역점수의 향상정도에서 모두 유의미한 차이는 없었고, 실험집단과 비교집단의 비교분석에서는 정의적 영역의 하위요소 중 학습에 대한 태도에서 중, 하위권 학생들에게 긍정적인 효과가 나타났으며, 학습습관에 있어서는 하위권 학생들에게 긍정적인 효과를 보였다.

따라서 과학과에서 실험실습법에 의한 수행평가의 실시가 중학생들의 과학성취도와 정의적 영역의 향상에 영향을 미치고 있는 것으로 나타났으며 특히 성적이 중, 하위권인 학생들에게 미치는 영향이 더 크게 나타났다.

참 고 문 헌

국립교육평가원(1996). 수행평가의 이론과 실제. 서

울: 대한교과서주식회사.

김영순(1992). 초등학교 자연과 수행평가에 대한 문
화기술적 연구. 한국교원대학교 교육대학원 석사
학위 논문.

김성훈(1993). 새로운 교육평가를 위한 인지심리학적
교육측정의 활용 가능성 탐색. 한국교육, 20,
131-154.

박승재, 조희영(1995). 과학론과 과학교육. 서울: 교육
과학사.

백순근(1995). 학력평가를 위한 새로운 대안, 교육개
발, 97(9), 44-48.

백순근(1998). 수행평가의 이론과 실제. 서울: 원미사

엄기영(1999). 수행평가가 학업성취와 창의적 사고력
에 미치는 영향. 공주대학교 교육대학원 석사학
위 논문.

윤동영(1999). 수행평가피드백이 아동의 학업성취 및
학습태도에 미치는 효과. 한국교원대학교 교육대
학원 석사학위 논문.

왕경순.(1998). 다중지능에 기초한 프로젝트 활동이
초등학교 아동의 과학 학업 성취도 및 문제 해결
에 미치는 영향. 부산대학교 교육대학원 석사학
위 논문.

조희형(1998). 과학교육의 이론적 배경과 그 시사점.
한국과학교육학회지, 18(2), 183-200.

허경철(1988). 객관식 일변도의 평가방식의 문제점과
개선방향. 교육과정연구, 7, 155-182.

Doran, R., Chan, F., & Tamir, P. (1998).
Science educator's guide to assessment.
National Science Teachers Association:
Arlington, Virginia.

Ebenezer, J. V., & Haggerty, S. M. (1999).
*Becoming a secondary school science
teacher*. Meril: Upper Saddle River, New
Jersey.

Herman, J. L., Aschbacher, P. R., & Winter,
L.(1992). *A practical guide to alternative
assessment*. U.S.A: University of California.

Hibbard, K. M.(2000). *Performance-based
learning and assessment in middle school
science*. Eye On Education: Larchmont,
NY.