

자연과 수업에 증거집(포트폴리오) 평가의 적용이 초등학교 학생들의 과학 지식, 탐구능력 및 태도에 미치는 영향

김혜정 · 김찬종¹

(대전 보덕초등학교) · ¹(충주교육대학교 과학교육과)

The Effects of Portfolio Assessment on Elementary School Students' Science Knowledge, Inquiry Ability and Science Attitudes

Kim, Hyejeong · Kim, Chan-Jong¹

(Boduk Elementary School, Daejeon-si) · ¹(Chongju National University of Education)

ABSTRACT

The major purposes of this study are to examine the effects of portfolio assessment on elementary school student's science knowledge, inquiry ability, science attitudes and to investigate students' perceptions on portfolio assessment.

Control group consists of 45 fourth-graders at M-Elementary school located at Miwon, Chongwon-gun, Chung-buk and experimental group 36 fourth-graders of G-Elementary school located in Daejeon-si.

The inventories of scientific knowledge 1, inquiry ability, and science attitudes were administered to both groups as a pre-test. The experimental group was given portfolio assessment instruction and control group traditional instruction for about six weeks. Inventories about scientific knowledge 2, inquiry ability, and science attitudes were administered to both groups as a post-test. A questionnaire on the perception on portfolio assessment was given to experimental group after the treatment. The results were statistically analyzed with SPSS.

Control group showed higher score on scientific knowledge than that of experimental group ($p < 0.5$). No statistically meaningful difference was identified in inquiry ability and scientific attitude. More in-depth analysis revealed that scientific attitudes were improved statistically meaningfully by portfolio assessment. The students' perceptions on portfolio assessment is very positive. Students have positive responses on interests in portfolio assessment, feelings of involvement in learning, self-regulated learning, higher levels of thinking, intentions of participation in portfolio assessment.

Key words : portfolio assessment, elementary school science instruction and learning, science assessment, performance assessment, science knowledge, inquiry ability, science attitudes

¹1998년 10월 20일 받음

1. 서론

우리 나라에서는 교육개혁의 일환으로 열린교육이 권장되고 있으며, 수행평가와 증거집(포트폴리오) 평가를 새로운 대안으로 장려하고 있다(국립교육평가원, 1996; 박도순, 1995; 백순근, 1995).

새로운 교수학습지도나 평가 방법의 도입을 위해서는 교육부나 교육청의 의지와 권장도 중요하지만, 새로운 방법과 우리 교육 현장의 특성에 관한 체계적이고 실증적인 연구가 기반이 되어야 할 것이다. 그러나 여러 가지 수행 평가가 가진 많은 장점이 강조되고 있고(Barton & Collins, 1997; Collins, 1992; Hart, 1994; Ruiz-Primo & Shavelson, 1996; Wiggins, 1989), 전통적인 평가 대신 수행평가가 널리 권장되고 있으며, 이를 실천한 사례를 소개하는 책자가 많이 있음에도 불구하고, 교실에서 수행평가의 적용과 관련된 체계적인 연구가 많지 않다(Kamen, 1996).

우리 나라의 경우에도 사정은 마찬가지여서, 수행평가나 증거집 평가에 대한 체계적인 연구가 본격적으로 실시되기 시작한 것은 그리 오래되지 않았으며, 연구 성과도 많지 않다(김찬중, 김혜정, 1998). 따라서 현장에서는 외국의 사례를 바탕으로(석문주 외, 1997)으로 경험과 직관에 의존한 채 기존의 평가 방식을 변형하여 실시하고 있는 실정이다. 그러나 이처럼 기초 연구가 결여된 실천과 개혁은 새로운 평가 방법이 우리 실정과 맥락에 잘 접목되지 못하고, 제대로 정착되기 어렵게 한다. 수행 평가의 적용 효과에 대해서는 학습효과를 높이고, 교사가 학생들이 무엇을 알고 있는지 더 잘 이해할 수 있고, 교사의 학습지도 능력을 향상시키며(Dana et al., 1991; Haertel, 1991; Martin-Kniep et al., 1998; Seidel et al., 1997), 교사와 학생들 사이에 보다 긍정적인 관계를 수립할 수 있다(Grace, 1992)는 주장과 수행평가의 개념이 불확실하고, 교사의 교수활동과 학생의 학습 성취에 대한 영향도 분명치 않다(Ruiz-Primo & Shavelson, 1996)는 상반된 주장이 있다. 증거집 평가에 대한 외국의 실증적 연구는 많지 않으며, 있다 해도 주로 대학생을 대상으로 삼았다.

대학 일반물리를 수강하는 학생들에게 증거집 평가를 적용한 결과 학생들은 이 평가 방법에 호의적인 반응을 보였으며, 시험 불안이 감소하고, 수업에 대한 태도가 긍정적으로 변화하였으며, 물리 개념 이해가 증진되었다고 응답하였다(Slater, 1994).

그러나 전통적인 평가를 실시한 집단과 비교하여 성취도는 유의미한 차이를 보이지 않았다(Slater, Ryan & Samson, 1997). 대학의 환경 지질학 수강자를 대상으로 한 연구에서도 유사한 결과가 나왔다. 성취도는 전통적 평가 방법을 도입한 집단과 유의미한 차이가 없었으나, 학생들은 능동적으로 성취 증거를 작성하는 증거집 평가 방법을 긍정적으로 받아들였다(Astwood & Slater, 1997).

우리의 상황과 여건, 또는 실정에 적합한 과학과 증거집 평가 방식의 탐색 연구가 초등학교(김수균, 송명섭, 1997; 김찬중과 김혜정, 1988; 김혜정, 1988)와 중등학교 수준(김영학 외, 1998), 그리고 과학 영재를 대상으로(박윤배와 전명남, 1998) 시작되었으며, 증거집 평가의 적용 효과에 대한 연구가 부분적으로 시행되어 과학적 태도 중 일부 영역에 대한 긍정적 효과가 보고되었다(이수환과 송명섭, 1997). 또한 자연과 수업을 위한 증거집 평가 도구의 개발 연구가 수행되었다(김찬중 외, 1988). 따라서 증거집 평가의 효과에 대한 실증적 연구는 극히 미미한 실정이다.

이 연구의 주요 목적은 자연과 수업에 증거집 평가를 적용하여 그것이 초등학생들의 과학 지식과 탐구 능력 및 과학에 관련된 태도에 미치는 영향을 살펴보고 초등학생들의 증거집 평가에 대한 인식을 알아보는 것이다. 이러한 목적을 달성하기 위한 구체적인 연구 문제는 다음과 같다.

- (1) 증거집 평가의 적용이 과학 지식에 미치는 영향은 어떠한가?
- (2) 증거집 평가의 적용이 과학 탐구능력에 미치는 영향은 어떠한가?
- (3) 증거집 평가의 적용이 과학 태도에 미치는 영향은 어떠한가?
- (4) 증거집 평가에 대한 학생들의 인식은 어떠한가?

II. 연구 방법 및 절차

이 연구는 김찬중과 김혜정(1998)의 연구 결과를 바탕으로 구성된 증거집 평가의 적용 효과를 조사하는 것이다.

1. 연구대상

전통적 수업 및 전통적인 평가 집단(이하 '비교집단')은 충북 청원군 미원면에 소재한 M초등학교 4학년 45명이고, 포트폴리오 수업 및 증거집 평가집단(이하 '실험집단')은 대전광역시 유성구에 위치한 G초등학교 4학년 36명이었다. 비교집단과 실험집단은 모두 농촌 지역에 위치하고 있으며, 학부모의 사회경제적 지위(SES: Socio Economic Status)가 낮으며 방과 후 학생들이 과외 교습을 받는 비율이 낮은 점 등에서 유사한 특성을 가지고 있다.

2. 검사도구

증거집 평가의 적용 효과를 알아보기 위해서 지식, 탐구 능력, 태도, 그리고 증거집 평가에 대한 인식 검사지를 사용하였다. 지식 검사지는 사전 지식 검사지와 사후 지식 검사지의 2가지 종류가 있다. 사전 지식 검사지는 본 연구의 적용 대상 내용인 4학년 '지층과 화석'의 관련된 선수 학습 단원으로 3학년 '돌과 흙' 4학년 1학기 '강과 바다' 단원을 내용으로 한다. 이 내용에 대한 2원분류표를 작성하여 중요한과학 개념을 중심으로 평가 문항을 개발하였다. 문항수는 20문항으로 1차 개발 후 초등과학교육 전공자와 현장 교사들의 검토를 거쳐 수정 보완하였다. 이 검사의 내용 타당도는 1.0, 신뢰도(Cronbach- α)는 0.81로 나타났다.

사후 지식 검사지는 4학년 '지층과 화석' 단원의 주요 과학 개념을 중심으로 개발된 평가 문항으로 구성되어 있다. 평가 문항 수는 20문항으로 초등과학교육 전공자와 현장 교사들의 검토를 바탕으로 수정하였다. 이 검사지의 내용 타당도는 0.97, 신뢰도(Cronbach - α)는 0.73이었다. 사전·사후 지식 검사

지가 비동형 검사이므로 각 검사지의 문항 특성을 비교하기 위하여 문항반응이론에 입각한 문항 분석을 실시하였다.

문항 분석을 위해 사용한 프로그램은 ASC사의 문항 및 검사 분석용 소프트웨어(ASC Item and Test Analysis Package)이다. 그 결과 사전 및 사후 지식 검사지는 문항의 난이도(두 검사지 모두 -2.0 이상: 쉬움에서 매우 어려움)와 변별도(0.39에서 0.99 사이: 적절함에서 높음)에서 매우 유사한 능력을 요구하고 있는 것으로 나타났다 (김혜정, 1988). 탐구 능력 평가에는 권재술, 김범기(1994)가 개발한 과학 탐구능력 검사지를 사용하였다. 이 검사지의 신뢰도 지수가 0.78, 난이도 지수 0.59, 타당도 지수가 0.89 (고혁민, 1994)였다. 이 검사지는 30문항으로 관찰, 분류, 측정, 예상, 자료변환, 자료해석, 추리, 가설설정, 변인통제, 일반화 각각 3문제씩으로 되어 있다.

과학 태도의 평가에는 정완호 등(1994)이 초등학교생의 과학적 태도 측정을 위해 개발한 검사지를 사용했다. 전체 문항수는 37문항으로 검사시간은 30분 이내에 할 수 있는 것으로 긍정적 입장과 부정적 입장이 고르게 분포되어 있으며 전체 문항에 대한 신뢰도(Cronbach α)는 0.91로 나타났다.

증거집 평가에 대한 인식 검사지는 연구자들이 직접 제작한 것으로 증거집에 대한 이해, 태도, 효과, 타인의 도움, 증거 작성 시간과 증거집 평가에 대한 전체적인 생각과 느낌에 대하여 묻는 18개의 문항으로 구성되었다. 각 문항마다 5단계로 구성된 답지에 피험자의 의견을 표시하도록 되어 있다. 이 검사지는 사후 검사로 실험 집단에서만 실시하였다.

3. 연구절차

본 연구는 증거집 평가를 위한 준비, 사전 검사, 증거집 평가의 적용, 사후 검사, 결과 분석 및 처리의 순으로 실시되었다. 사전 검사를 위한 검사지의 투입은 1997년 9월 11일에서 13일 사이에 실험 집단과 비교 집단의 남녀학생 모두에게 실시되었다.

검사에 소요된 시간은 사전 지식 검사지 40분, 탐구 능력 검사지 40분, 과학 태도 검사지 30분이었다. 1997

년 9월 12일부터 1997년 10월 27일 까지 약 6주간 걸쳐서 증거집 평가를 이용하여 수업을 실시하였다.

비교 집단에서는 전통적인 수업과 전통적인 평가가 이루어졌고, 실험 집단에서는 약 2주간 증거집 평가에 대해서 소개하고 증거를 작성하는 기회를 주었다. 그리고 난 뒤 약 4주간 4학년 2학기 '지층과 화석' 단원의 포트폴리오 학습과 평가를 본격적으로 실시하였다. 증거집 평가 목표를 먼저 작성하여 수업을

시작할 때 제시하였다.

학생들과 제시한 목표에 대한 충분한 토의를 통하여 이해를 시킨 뒤 포트폴리오 수업을 하면서 증거를 수집하게 하였다. 학생들은 자기 나름대로 이해한 목표를 진술하였다. 학생들이 증거집에 포함시킨 증거의 유형은 상당히 다양했다. 대표적인 유형은 글, 실험보고서, 학습지, 증명서, 복사물, 사진 등이다. 두 가지 또는 그 이상의 유형이 복합된 경우도 있었다.

Table 1. The results of scoring of students' evidence for portfolio objectives

Portfolio assessment objectives	Grade				Self-reflection
	A	B	C	D	
Explaining strata	21	0	0	16	1
Comparing the shapes, thickness, color, grain size of strata at various places	21	4	0	12	12
Explaining strata forming process	25	2	0	10	18
Making the model of strata, and comparing the model with real strata	16	2	1	18	13
Explaining why the grain size of strata is not varying	15	4	1	17	12
Comparing the precipitation patterns of gravels, sand, and clay in water	25	1	1	10	14
Observing, comparing and explaining rocks which form strata in various ways	20	3	3	11	12
Conducting an experiment to compare water leaking between sand and clay layer and explaining water leakage of a stratum	28	0	2	7	3
Explaining how the groundwater is formed	25	1	0	11	23
Explaining the results of observing fossils	13	2	1	21	0
Comparing fossils with fossil models in various ways	9	3	0	25	1
Explaining the process of fossil formation in various way	18	3	0	16	10
Explaining how we identify the sequence and period of strata by fossils	11	5	2	19	3
Explaining the process of coal and petroleum formation	10	7	2	18	3
Investigating the use of fossils	10	2	2	23	5
Investigating what we can learn about the past with strata and fossils	8	3	2	24	7

A: sound understanding of objectives and competent evidence

B: sound understanding of objectives and poor evidence

C: partial understanding of objectives and poor evidence

D: no evidence

Table 2. The types of students' evidence

Types of evidence	Proportion(%)
Written explanation	29.0
Drawing with written explanation	22.0
Model	0.8
Pictures	21.7
Laboratory report	12.9
Experiments	2.3
Newspaper articles	0.2
Poem	0.2
Quiz	1.8
Attestations	6.6
Others	2.5
Total	100.0

목표에 대한 증거를 수집하면서 또는 수집하고 난 뒤 자기 반성을 적은 증거집도 있었다.

실험 집단 37명의 학생들 중 증거집을 제출한 학생은 31명이었다. Table 1은 작성한 포트폴리오의 채점 결과를 보여준다. 목표를 잘 이해하고 이에 잘 부합되는 증거를 수집한 학생은 A, 목표는 바르게 이해했으나 증거의 내용이 미흡한 것은 B, 목표를 이해하지 못하여 증거의 내용이 많이 불충분한 것은 C로 채점하였다. 또한 증거집에 자기 반성의 내용을 적은 수를 조사한 것이다. 대부분의 학생들이 목표를 바르게 이해한 내용을 적고 목표에 알맞은 증거를 제시하였다. 또한, 목표와 증거 외에 자기 반성에 대한 내용을 적은 학생도 많았다. Table 2는 증거의 유형을 분석하여 백분율(%)로 나타낸 것이다.

대부분의 학생들이 글로 설명하거나 그림을 그리고 설명하기, 사진이나 복사물을 많이 사용하여 증거를 제시하였다. 동시를 쓰거나 퀴즈를 증거로 낸 학생들도 있었다. 자기 반성의 내용은 제시한 증거들이 왜 증거로서 합당한가에 대하여 쓴 것이 많았다. 또한, 새로운 사실을 알게 되어 기쁘다는 내용도 많았다. 어떤 학생들은 수업시 교사를 평가한 내용을 쓰기도 했다.

1997년 10월 28일부터 30일 사이에 실험집단과 비교집단의 남녀학생 모두에게 사후 검사를 실시하였다. 사후 지식 검사 외에 탐구능력 검사와 과학에 관

련된 태도 검사지는 동일하였다. 각 검사지에 소요된 시간도 사전 검사에서와 같았다. 증거집에 대한 인식 검사지는 30분 동안 실시하였다. 본 연구는 사전검사 결과를 공변량으로 하는 공분산분석(ANCOVA)을 실시하였다. 통계분석에는 SPSS통계 패키지를 사용하였다.

III. 연구 결과

증거집 평가의 적용 효과를 과학 지식, 탐구 능력, 태도의 순으로 정리하였다. 과학 지식 검사 분석 결과는 Table 3에 제시되어 있다. 실험 집단과 비교 집단의 사전 검사를 이용하여 공분산 분석을 실시하였다. 두 집단간에 통계적으로 유의미한 차이가 나타났다(p<.05), 비교 집단이 더 높은 성취를 보였다(Table 4).

과학 탐구 능력 검사 결과는 Table 5에 제시되어

Table 3. Mean and standard deviation of science achievement test score

	Group	n	M	SD	Adj.M
Pre-test	Experimental	36	21.00	6.23	
	Control	45	19.56	4.98	
Post-test	Experimental	36	17.56	3.86	17.26
	Control	45	18.53	3.30	18.77

Table 4. ANCOVA results on the science achievement test

	Group	df	SS	MS	F	p
Experimental						
Control		1	44.51	44.51	4.73	0.03

있다. 사전 검사에서 탐구 능력의 평균은 두 집단이 비슷하고, Table 6에서 알 수 있는 것처럼 통계적으로 의미있는 차이가 나타나지 않았다(p>.05). 집단간의 평균의 차의 유의도를 검증하기 위해 사전 검사 결과를 공변량으로 하여 공분산 분석(ANCOVA)을 실시하였다. 두 집단간에 유의미한 차이가 나타나지

Table 5. Mean and standard deviation of the science inquiry test score

	Group	n	Pre-test		Post-test		
			\bar{M}	\bar{SD}	\bar{M}	\bar{SD}	Adj.M
Science inquiry(30)	Experimental	36	13.97	4.03	16.03	4.25	16.12
	Control	45	14.24	4.20	15.11	3.52	15.04
Observation(3)	Experimental	36	1.86	0.68	2.17	0.81	2.17
	Control	45	1.91	0.70	1.98	0.69	1.97
Classification(3)	Experimental	36	2.17	0.65	2.06	0.86	2.00
	Control	45	1.67	0.95	2.07	0.75	2.11
Measurement(3)	Experimental	36	1.50	1.00	1.72	0.89	1.74
	Control	45	1.73	0.89	1.80	0.84	1.78
Prediction(3)	Experimental	36	1.58	0.84	1.89	0.92	1.91
	Control	45	1.69	0.97	1.78	0.99	1.76
Inference(3)	Experimental	36	1.56	0.88	1.78	0.98	1.78
	Control	45	1.56	0.89	1.67	0.88	1.67
Data transformation(3)	Experimental	36	1.06	0.84	1.06	0.77	1.06
	Control	45	1.04	1.04	1.04	0.77	1.05
Data interpretation(3)	Experimental	36	1.00	1.00	1.06	0.91	1.08
	Control	45	1.27	0.69	1.16	0.77	1.14
Controlling variables(3)	Experimental	36	1.43	0.85	1.57	0.97	1.55
	Control	45	1.29	0.82	1.29	0.79	1.30
Hypothesis formation(3)	Experimental	36	1.11	0.76	1.23	0.66	1.22
	Control	45	1.02	0.78	1.16	0.82	1.16
Generalization(3)	Experimental	36	0.86	0.78	1.51	1.25	1.53
	Control	45	1.07	0.81	1.18	0.75	1.17

않았다($p>.05$).

탐구 능력을 다시 하위 영역으로 세분화하여 분석한 결과, 두 집단 사이에 통계적으로 유의미하게 차이가 있는 영역은 없었다.

Table 7은 과학 태도를 과학에 대한 태도, 과학 교육에 대한 태도, 과학적 태도로 세분화 한 것과 전체 과학 태도 검사 결과이다. 과학 태도는 두 집단 사이에 유의미한 차이가 없었다(Table 8).

하위 영역별 분석 결과 과학적 태도에서 실험 집단이 비교 집단보다 통계적으로 더 유의미한 차이를 보였다. 이는 증거를 수집, 작성하고 증거집을 구성하는 과정에서 학생들은 왜 자신의 증거가 목표에 대한

증거가 되는지를 판단하는 등과 같은 과학적 태도를 신장시킬 수 있는 경험을 하기 때문인 것으로 보인다.

전체적으로 볼 때, 과학 태도에 대한 사후 검사에서 실험 집단이 비교집단보다 평균 성적이 약 5점 정도 높아졌지만 통계적으로 유의미하게 차이가 나지 않은 것은 실험기간이 약 6주 정도로 태도에 대하여 변화를 보이기에 짧았기 때문이라고 생각된다. 실험 집단에 한해서 증거집 평가를 적용한 후에 증거집 평가에 대한 인식 검사를 실시하였다. 각 항목에 대하여 학생들은 강한 부정(1), 부정(2), 보통(3), 긍정(4), 강한 긍정(5) 중에서 하나를 선택하도록 되

Table 6. ANCOVA results of science inquiry test score

	DF	SS	MS	F	P
Science inquiry(30)	1	23.03	23.03	2.68	0.11
Observation(3)	1	0.81	0.81	1.53	0.22
Classification(3)	1	0.25	0.25	0.41	0.53
Measurement(3)	1	0.035	0.035	0.048	0.83
Prediction(3)	1	0.40	0.40	0.46	0.50
Inference(3)	1	0.25	0.25	0.37	0.54
Data transformation(3)	1	0.0018	0.0018	0.003	0.95
Data interpretation(3)	1	0.078	0.078	0.11	0.74
Controlling variables(3)	1	1.24	1.24	1.67	0.20
Hypothesis formation(3)	1	5.59	5.59	0.11	0.74
Generalization(3)	1	2.54	2.54	2.49	0.12

Table 7. Mean and standard deviation of the science attitudes test score

	Group	n	Pre-test		Post-test		Adj.M
			\bar{M}	\bar{SD}	\bar{M}	\bar{SD}	
Science attitudes(37)	Experimenta	36	122.22	18.92	128.76	16.35	127.47
	Control	45	121.24	17.77	121.82	18.00	122.04
Attitudes toward science(12)	Experimenta	36	37.58	10.29	39.39	7.44	39.73
	Control	45	38.84	0.05	40.64	7.66	40.37
Attitude toward science education(6)	Experimenta	36	18.03	4.93	18.78	3.84	19.04
	Control	45	19.13	4.58	19.38	4.30	19.17
Scientific attitudes(19)	Experimenta	36	66.61	6.94	69.58	8.47	68.67
	Control	45	63.27	7.77	61.80	9.53	62.53

Table 8. ANCOVA results on science attitudes tests

	df	SS	MS	F	p
Science attitudes(37)	1	558.92	558.92	2.75	0.10
Attitudes toward science(12)	1	8.21	8.21	0.23	0.63
Attitude toward science education(6)	1	0.33	0.33	0.026	0.87
Scientific attitudes(19)	1	716.46	716.46	10.24	0.002

어 있었다. Table 9는 학생들의 의견을 인식 요소별 평균으로 제시한 것이다. 전체적으로 볼 때 학생들은 증거집 평가를 비교적 쉽게 이해하고 받아들이고 있으며, 증거 작성 활동이 흥미롭고, 유익한 것으로 판단하고 있다.

IV. 결론 및 토의

세계화·개방화 시대를 맞이하여 창의성과 자기 주도적 학습 능력을 길러주는 것이 매우 중요한 과제로 대두되었다. 교육 개혁도 이를 반영하여 수요자

Table 9. Students' Perceptions on Portfolio Assessment

Domain	Items	Mean
Understanding	I already know what portfolio assessment is before the lesson.	1.21
	Portfolio assessment is easy to understand from the beginning.	3.60
	Developing evidence is easy from the beginning.	3.14
Attitudes	Science classes with portfolio assessment are more interesting than before.	3.90
	Portfolio assessment requires more time without gain.	1.86
	Making portfolio makes me confident.	3.67
	I hope to develop portfolios in the classes to come.	3.14
Effects	Portfolio assessment makes me think a lot.	3.67
	Portfolio assessment makes me study for myself.	3.56
	Portfolio assessment makes me participate class more actively.	3.43
	Portfolio assessment makes me understand science more easily.	3.83
	I am satisfied with my evidence for portfolio.	3.62
	I had many new ideas when I developed evidence for portfolio.	3.43
	I could understand science concepts more clearly with portfolio assessment.	3.95

strongly disagree = 1, disagree = 2, moderate = 3, agree = 4, strongly agree = 5

중심의 개별화 학습의 확산과 열린교육과 수행평가가 강조되고 있다. 수행평가 중에서 수업과 병행하여 질적평가를 할 수 있는 증거집 평가는 창의성과 자기 주도성 신장에 크게 기여할 수 있는 매우 발전적인 평가 방법이다. 본 연구에서는 증거집 평가 도구를 개발하여 4학년 2학기 '지층과 화석' 단원에서 약 6주 동안 적용하여 보았다. 연구 결과 얻은 결론은 다음과 같다.

첫째, 증거집 평가는 전통적 방식과 비교하여 과학 지식의 성취에 더 효율적이 아닌 것으로 나타났다. 선행 연구에서도 지필검사를 이용한 성취도 검사 결과에서 증거집 평가가 더 좋은 효과를 내지 못하는 것으로 보고되었다. 증거집 평가를 적용하였을 때 학생들은 상당한 학습 향상 효과를 느끼는 설문 결과로 보아 새로운 성취도 측정 방법의 도입이 필요한 것으로 보인다.

둘째, 증거집 평가는 학생들의 과학 탐구 능력 신장에서 전통적인 방법과 거의 같은 효과를 보인다. 또한 탐구 능력에 대한 하위 능력에서도 실험집단과 비교 집단 사이에 통계적으로 의미있는 차이를 발견할 수 없었다.

셋째, 증거집 평가는 과학 태도 전반에서는 전통적

방식과 유사한 효과를 보인다. 그러나 증거집 평가는 과학 태도 중에서 과학적 태도 영역에서 비교집단보다 통계적으로 유의미한 차이를 보였다.

넷째, 증거집 평가에 참여한 학생들은 매우 긍정적인 인식을 가지게 되었다. 증거집 평가에 대한 인식 검사지 분석 결과, 증거집 평가에 대한 흥미, 참여도, 자기 학습, 다양한 생각, 공부에 도움이 되는 정도, 계속 참여 희망 등에서 매우 긍정적인 반응을 보였다.

이 연구에서는 증거집 평가가 과학 학업성취도, 과학 탐구능력, 과학 태도에 미치는 영향을 정량적으로 조사하였다. 대부분의 영역에서 유의미한 차이가 발견되지 않은 것은 증거집 평가의 적용 기간이 비교적 짧아서 증거집 평가를 통한 능력의 변화가 크지 않았고, 자료 수집에 전통적 지필 검사를 사용하였기 때문에 증거집 평가를 통해서 길러진 능력이 잘 측정되지 않았을 가능성이 크다.

그럼에도 증거집 평가에 대한 학생들의 인식 검사 분석 결과는 매우 고무적이다. 약 6주의 짧은 기간에 학생들이 흥미, 참여도, 자기 학습, 다양한 생각 등에서 매우 긍정적인 반응을 보였다. 이는 증거집 평가의 장기적인 활용으로 과학 태도의 긍정적 변화를 가

겨울 수 있고, 나아가서는 성취도와 탐구능력의 신장으로 이어질 수 있을 것으로 보이기 때문이다.

앞으로 우리 교실 현장에서 보다 사용하기 쉬운 증거집 평가를 개발하고, 증거집 평가가 교실 환경 변화와 창의성, 자기 주도성 신장에 미치는 영향을 조사해 나가야 한다. 또한 증거집 평가를 통해서 길러진 능력을 정확하게 측정할 수 있는 검사 도구 및 방법의 개발이 뒤따라야 할 것으로 본다.

적 요

이 연구의 주요 목적은 증거집(포트폴리오) 평가를 적용할 때 초등학교 학생들의 과학 지식, 탐구능력, 태도에 미치는 영향을 알아보고 학생들의 증거집 평가에 대한 인식을 조사하는 것이다. 비교집단은 충북 청원군 미원면의 M초등학교 4학년 45명이며, 실험집단은 대전광역시 G초등학교 4학년 36명이다.

사전 검사로 과학지식 검사1, 과학탐구능력검사, 과학 태도 검사를 두 집단에 실시하였다. 실험집단은 증거집 평가를 적용한 수업을 비교집단은 전통적인 수업을 약 6주 동안 실시하였다. 두 집단 모두 사후 검사로 과학지식 검사2, 과학 탐 구능력검사, 과학 태도 검사를 실시하였다. 증거집 평가에 대한 인식 검사지를 실험집단에 실시하였다. 결과는 SPSS를 이용하여 통계적으로 분석하였다. 비교집단이 실험집단보다 과학 지식 검사에서 높은 성취를 보였다. 과학 탐구능력과 과학 태도에서는 의미있는 차이가 없었다. 하위 영역 분석 결과 과학적 태도에서 실험집단이 비교집단보다 통계적으로 유의미하게 높은 성취를 보였다. 학생들의 증거집 평가에 대한 인식은 매우 긍정적으로 나타났다. 학생들은 증거집 평가에 대한 흥미, 학습 참여도, 자기 주도적 학습, 고차적 사고, 앞으로 참여 의사 등과 같은 대부분의 하위 영역에서 긍정적인 반응을 보였다.

참 고 문 헌

고혁민(1994). 초·중학생들을 위한 과학 탐구능력 측정도구의 개발과정 및 타당성 분석. 한국교원

대학교 대학원, 석사학위 논문.

국립교육평가원(1996). 수행 평가의 이론과 실제. 서울: 국립교육평가원.

권재술, 김범기(1994). 초·중학생들의 과학탐구능력 측정도구의 개발. 한국과학교육학회지, 14(3), pp. 251-264.

김수균, 송명섭(1997). 용해와 용액 단원에 대한 증거집 개발의 효과. 한국초등과학교육학회 1997년 하계 학술발표회, 청주교육대학교.

김영학, 서혜애, 윤세진, 정은영(1998). 과학 수업의 증거집 적용에 대한 연구. 한국과학교육학회 1998년 하계학술논문발표회.

김찬중, 김혜정(1998). 초등학교 자연과 포트폴리오 평가의 구성 요소. 한국과학교육학회지, 18(2), 233-243.

김찬중, 윤선아, 최승희, 홍은석, 김명수, 여원미, 김미숙, 김순영, 이주술, 이지연, 편준호(1998). 초등 과학 포트폴리오 평가 도구 개발 연구. 한국초등과학교육학회지, 17(1), 11-21.

김혜정(1998). 초등학교 과학 평가에서 포트폴리오 평가의 적용이 과학 지식과 탐구능력 및 태도에 미치는 영향. 청주교육대학교 교육대학원 석사학위 논문.

박도순(1995). 수행 평가의 문제. 한국교육학회 소식, 31(4), pp. 9-11.

박윤배, 전명남(1998. 8). 과학 영재아의 교수-학습에서 포트폴리오 평가의 활용 방안. 한국과학교육학회 '98 하계 학술대회, 서울교육대학교.

백순근(1995). 창의성 신장에 적합한 평가 방법. 서울 교육, pp. 36-39.

석문주, 송명섭, 이명숙, 이원희, 이종일, 조용기, 최호성, 홍종관 (1997). 학습을 위한 수행평가. 서울: 교육과학사.

이수환, 송명섭(1997. 8). 포트폴리오 학습이 초등학교 학생들의 과학에 관련된 태도에 미치는 영향. 한국초등과학교육학회 1997년 하계 학술발표회, 청주교육대학교.

정완호, 허명, 윤병호(1994). 국민학생들의 과학적 태도 측정을 위한 도구 개발. 한국과학교육학회

- 지, 14(3), pp. 265-271.
- Astwood, P.M., & Slater, T.F. (1997). Effectiveness and management of portfolio assessment in high-enrollment courses. *Journal of Geoscience Education*, 45, 238-242.
- Barton, J., & Collins, A. (Eds.) (1997). *Portfolio assessment: A handbook for educators*. Addison-Wesley: Menlo Park, CA, USA.
- Collins, A. (1992). Portfolios for science education: Issues in purpose, structure, and authenticity. *Science Education*, 76 (4), 451-463.
- Dana, T.M., Lorsbach, A.W., Hook, K., & Briscoe, C. (1991). Students showing what they know: A look at alternative assessment. In G. Kulm & S.M. Malcom (Eds.), *Science assessment in the service of reform* (pp. 331-337). Washington, D.C.: American Association for the Advancement of Science.
- Grace, C. (1992). *The portfolio and its use: Developmentally appropriate assessment of young children*. ERIC Digest, Urbana, IL: ERIC Clearinghouse on Elementary and Early Childhood Education. (ED No. 351 150)
- Haertel, E.H. (1991). Form and function in assessing science education. In G. Kulm & S.M. Malcom (Eds.), *Science assessment in the service of reform* (pp. 233-245). Washington, D.C.: American Association for the Advancement of Science.
- Hart, D. (1994). *Authentic assessment: A handbook for educators*. Addison-Wesley: Menlo Park, CA, USA.
- Kamen, M. (1996). A teacher's implementation of authentic assessment in an elementary science classroom. *Journal of Research in Science Teaching*, 33(8), 859-877.
- Kranz, R. (1994). *Portfolio assessment across the curriculum*. Troll Associates.
- Martin-Kniep, G.O., Cunningham, D., & Feige, D.M. (1998). *Why am I doing this?: Purposeful teaching through portfolio assessment*. Portsmouth, NH: Heinemann.
- Ruiz-Primo, M.A., & Shavelson, R.J. (1996). Rhetoric and reality in science performance assessments: An update. *Journal of Research in Science Teaching*, 33(10), 1045-1063.
- Seidel, S., Wlaters, J., Kirby, E., Olf, N., Powell, K., Scripp, L., & Veenema, S. (1997). *Portfolio practices: Thinking through the assessment of children's work*. Washington, D.C. : National Education Association.
- Slater, T.F. (1994). Portfolios for learning and assessment in physics. *The Physics Teacher*, 32 (Sept.), 370-373.
- Slater, T.F., Ryan, J.M., & Samson, S.L. (1997). Impact and dynamics of portfolio assessment and traditional assessment in a college physics course. *Journal of Research in Science Teaching*, 34(3), 255-271.
- Wiggins, G. (1989). A true test: Toward more authentic and equitable assessment. *Phi Delta Kappan*, (May), 703-713.