

## CRESST 형성평가 프로그램의 이해 및 효과성 - 중학교 2학년 대수 관련 내용을 중심으로 -

최 승 현\* · 황 혜 정\*\* · 류 현 아\*\*\*

미 연방정부 교육부로부터 재정 지원을 받고 있는 평가·기준·검사 전문연구센터인 CRESST (The National Center for Research on Evaluation, Standards, and Student Testing at UCLA)에서는 지속적인 형성평가 실시에 따른 결과 반응을 통하여 교사의 수업 개선과 학생의 내용 숙달을 지원하고, 투입한 형성평가 프로그램의 교육 효과성을 검증하고자 하는 5년 일정(2007년~2011년)의 장기적 연구를 수행 중에 있다. 2007년과 2008년의 사전 연구를 토대로, 2009년도 연구에서는 우리나라 중학교 1학년 학생들에게 PowerSource©를 투입하여 대수 영역 성취도의 변화 추이를 분석하고, 해당 프로그램의 활용에 대한 교사와 학생의 반응을 탐색하고자 하였다. 또한, 우리나라 중학교 2학년 학생들에 PowerSource©를 투입하여 대수 영역 성취도의 변화 추이를 분석하고자 하였다. 본고에서는 CRESST의 PowerSource©의 이해 및 이의 활용을 토대로, 중학교 2학년 학생들을 대상으로 하는 대수 영역 성취도의 변화 추이를 분석하는 데에 초점을 두었다.

### 1. 연구의 필요성 및 목적

최근 한국교육과정평가원에서 실시한 국내·외 교실 학습 연구 결과에 따르면 영국, 프랑스, 일본 학생에 비하여 우리나라 학생은 교사로부터 학습 과정이나 결과에 대하여 적절한 피드백을 받지 못하고 있으며, 학습 과정에서 자신이 공부를 제대로 하고 있는지에 관심을 두기보다는 얼마나 과제를 빨리 해결하느냐에 초점을 두는 경향이 있는 것으로 나타났다(전효선 외, 2007). 이는 학습이 끝난 다음에 학생의 학습 결과에 대해 평점을 매기는 총괄평가가 강조되어 온 것에 기인한 것으로 학기 중에 수시로 형성평가를 실시하고 이에 대한 피드백을 적절히 받을 수 있다면 학생의 입장에서

자신이 공부를 제대로 하고 있는지에 대하여 좀 더 관심을 기울이게 될 것이다. 또한 교사의 입장에서라도 학습자의 성취수준을 파악할 수 있고 이를 반영하여 수업 전략 및 구체적인 실천 방식을 변화시켜나갈 수 있을 것이다. 따라서 형성평가 체제는 학생 중심 교육과정을 표방하는 국가수준 교육과정 지원 체제의 일환이기도 하다.

한편, NAEP(National Assessment of Educational Progress)의 2000년 당시 연구 결과에 따르면, 높은 수준의 학생들의 점수는 시간이 흐르면서 향상되어온 반면 낮은 수준의 학생들의 점수는 계속 하강하고 있다(NAEP, 2007). 또한 미국의 고등학교 학생들은 41개국에 참여한 PISA 평가 결과 31위로 나타났으며, 대학 신입생 중 약 1/3 정도는 정규 대학 강좌를 시작하

\* 한국교육과정평가원 (jhtina@kice.re.kr)

\*\* 조선대학교 (sh0502@chosun.ac.kr, 교신저자)

\*\*\* 한국교육과정평가원 (ryuha29@kice.re.kr)

기 전에 교정 교육을 받아야만 하는 상황에 처해 있다. 이러한 점은 CRESST(The National Center for Research on Evaluation, Standards, and Student Testing at UCLA)<sup>1)</sup>가 교수·학습 방법 및 평가에 대한 지속적인 연구에 대한 당위성을 강조하게 하였다. 이에 대한 한 가지 방법으로 CRESST는 수업 방법의 변화와 함께 형성평가의 활용을 제안하고, 형성평가 체제 도입의 교육 효과를 제고하기 위하여 5년(2007년~2011년) 일정으로 연구를 추진하고 있다. 이는 지속적으로 형성평가를 투입하고, 그 결과를 반영하여 교사의 수업 개선과 학생의 내용 숙달을 지원하고, 투입한 형성평가 프로그램의 교육 효과성을 검증하려는 연구이다. 이를 위하여, CRESST에서는 일차년도인 2007년에 대수 영역의 내용을 중심으로 중학교 1학년에 해당하는 자료를 개발하였으며, 이를 당해년도 7월부터 우리나라와 공유하면서 본격적으로 공동 연구가 착수되었다. 이 연구에서 대수 영역을 선택한 일차적 이유는 수학의 여러 영역 중에서 특별히 미국 학생들의 성취 수준이 낮기 때문이었으며, 또한 대학 입학에 포함된 후속되는 학업에서 대수에서의 성취 수준이 중요하다고 판단하였기 때문이다(Brown & Niemi, 2007). 따라서 학생들이 이후 상급 학교 교육에서 성공적인 교육 기회를 보장받고 자신의 활동 영역에서 능숙하게 대처할 수 있기 위해서는 해당 학교급에서 관련된 대수 영역 성취도를 보장해 줄 필요가 있다는 문제의식에서 출발한 것이라 하겠다.

이차년도인 2008년도에는 차기년도 연구에 앞서 사전적 의미에서의 연구가 실시되었다.

즉, 우리나라 13개 중학교의 1학년 학생 800여 명을 대상으로 PowerSource©를 시행하였으며 그 결과, 형성평가 실시 후 계속적으로 발전하는 긍정적인 효과가 있는 것으로 나타났다. 이러한 긍정적인 효과는 기존 학습 방법이나 1회성 평가 방법과는 달리 형성평가와 이를 기초로 한 피드백을 중시하는 형태로 진행되는 데서 비롯된 것으로 판단된다. 2009년도 연구에서는 우리나라 중학교 1학년 학생들에 PowerSource©를 투입하여 대수 영역 성취도의 변화 추이를 분석하고, 해당 프로그램의 활용에 대한 교사와 학생의 반응을 탐색하고자 하였다. 또한, 우리나라 중학교 2학년 학생들에 PowerSource©를 투입하여 대수 영역 성취도의 변화 추이를 분석하고자 하였다. 연구 계획에 따르면 2008년에 1학년, 2009년에 1, 2학년, 2010년에 1, 2, 3학년을 대상으로 시행하여 연차적으로 학생 개인별 성장 변화를 보고자 하였으며, 이 계획에 따라 우리나라도 2009년에 1학년과 2학년에 동시에 투입하였다. 보고에서는 CRESST의 PowerSource©의 이해 및 이의 활용을 토대로, 중학교 2학년 학생들을 대상으로 하는 대수 영역 성취도의 변화 추이를 분석하는 데에 초점을 두고자 하였다.

## II. PowerSource© 프로그램의 이해

이 장에서는 CRESST에서 개발한 PowerSource© 프로그램의 특징을 살펴보고, 이를 토대로 개발된 대수 영역에서의 형성평가 문항 및 수업 자료에 관하여 간략히 제시하고자 한다.

1) CRESST는 미 연방정부 교육부로부터 재정지원을 받고 있는 평가·기준·검사 전문연구센터로, UCLA의 교육·정보학 대학원 및 평가연구센터(CSE)와 제휴하여 운영되고 있다. 뿐만 아니라 콜로라도 대학, 스텐포드 대학, 랜드 연구소, 피츠버그 대학, 캘리포니아 산타 바바라 대학, 남캘리포니아 대학(USC), 그리고 ETS등과도 연합하여 구성되었다. CRESST는 설립 이래 주로 공교육의 주요 문제를 다루어왔으며, 교육의 질 평가, 평가 시스템의 설계와 활용에 있어서의 문제를 해결하기 위한 다양한 방안 제시 등의 평가 및 책무성 분야의 발전에 도움이 되는 연구를 선도적으로 수행하였음.

## 1. PowerSource©의 구성 및 절차

현재, 미국에는 국가 수준의 교육과정이 존재하지 않고 주 단위에서 교육과정을 마련하여 운영하고 있으나 CRESST에서는 어느 한 주에서만 사용되는 것이 아닌 핵심 개념(Big Idea)을 대수 영역을 중심으로 도출하여 PowerSource©를 개발하였다. 핵심 개념은 수학의 한 영역을 이루는 기본적인 개념을 일컫는 것으로, 이를테면 대수 영역에서는 수, 정수, 유리수, 집합, 수식, 등식, 부등식 이외의 여러 가지 내용을 포함하며, 이 개념들 간의 연계성이나 상호 관계를 대수 지식 내용의 지도로 나타낸 것이다(최승현 외, 2008). 이를 도식화하여 나타내면 [그림 II-1]과 같으며, 음영 처리된 부분이 CRESST에서 PowerSource©로 개발된 것이다. 실험 연구에 앞서, PowerSource©에서 제공하는 교수 학습 목표, 내용, 방법 등이 우리나라 교육 현장에서도 활용 가능한지 살펴보기 위하여, 우리나라 2007 개정 교육과정 내용을 바탕으로 하는 핵심 개념도를 마련하고 CRESST의 것과 비교하였다. 그 결과 CRESST에서 마련한 대수에 관한 핵심 개념도와 한국교육과정평가원의 것이 유사함을 인식하고, 이에 따라 미국 측에서 제안한 핵심 개념도를 그대로 적용하여 실험 연구를 실시하였다.<sup>2)</sup>

한편, 대수 영역 PowerSource©의 구성을 살펴보면 다음과 같다. CRESST에서 제공하는 중학교 2학년 PowerSource© 자료는 ‘유리수 동치’, ‘분배법칙’, ‘방정식 풀이’, ‘복습과 활용’

영역으로 구분되어 있으며, [그림 II-2]에서와 같이 각각 2차시 분량의 수업 자료와 3회의 형성평가로 구성되어 있다.<sup>3)</sup> 부연 설명하면, 각 주제별로 해당 주제에 대한 초기 평가(평가1), 1차시 학습지도안(수업1), 1차시 학습지도에 대한 형성평가(평가2), 평가2의 결과에 대한 피드백 위주의 2차시 학습지도안(수업2), 2차시 학습지도안에 대한 형성평가(평가3)로 구성되어 있다. 여기서, 평가1은 초기평가로서 학습시키고자 하는 내용에 기초한 문항들로 이루어져 있고, 이 문항들과 유사한 형태의 문항을 평가2와 평가3에서 다시 접하도록 하였다. 수업1은 각 주제별로 다뤄야 할 개념 및 성질에 대하여, 교사 발문, 설명 내용, 구체적인 예 등을 상세히 수록한 것이다. 이 지도안에 나와 있는 표현이나 구조를 정확하게 따르지 않고 교사 자신의 아이디어를 융통성 있게 활용하여 수업하도록 허용하되, 수업 자료에 제시된 개념들은 반드시 모두 다루도록 하였다. 평가2는 수업1에서 다룬 내용에 기초한 것으로 평가1에 포함된 문항 중 일부와 유사한 형태를 취하고 있다. 수업2는 개별 학습이나 소집단 학습 그리고 전체 학습하는데 모두 이용할 수 있도록 구성되어 있으며, 평가1의 결과에 따라 지도안의 활용 방법을 교사가 결정할 수 있도록 하였다.<sup>4)</sup> 평가3은 수업2에서 다룬 내용을 기초한 것으로 평가1에 포함된 문항 중 일부와 유사한 형태를 취하고 있다.

2) CRESST의 PowerSource© 프로그램이 개발된 미국의 상황과 우리나라의 상황, 교육과정이나 교과서 등이 완전히 일치하는 것이 아니므로 프로그램을 수정, 변형 없이 적용했다는 점이 본 연구의 제한점이 될 수 있겠으나 학생들이 알아야 할 기본적인 대수 내용은 같으며, 문항 유형에서 차이가 있다 하더라도 개념을 정확히 이해한다면 해결 가능하기 때문에 프로그램의 수정, 변형 없이 그대로 적용하였음.

3) 실험 수행 당시, ‘복습과 활용’ 관련 자료가 CRESST 측으로부터 전달되지 않아서 이 부분을 포함시키지 못하였음.

4) 본 연구에서는 1차 형성평가 실시 후 문제별 오답 수와 응답 유형을 정리하여 상대적으로 오답 수가 많은 문항과 상당수의 학생들이 오류를 범하고 있는 문항을 파악하여 2차시 수업 자료에서 해당 부분을 선택하여 전체 학생 대상으로 피드백 수업하거나 개인별 또는 그룹별 활동지로 활용할 수 있도록 하였다. 교사의 지도 방법이 학생들의 평가 결과에 영향을 미칠 수 있음을 고려할 때 교사별 자료 활용 및 수업 방법에 관한 질적 분석 또한 의미 있는 연구가 될 수 있음.



요소의 이해와 설명 등을 사용하여 각 핵심 개념 영역에 필요한 지식을 평가하기 위한 문항들을 개발하였다. 수학 교육자들과 평가 전문가들로 이루어진 평가 문항 개발 팀이 기본 문항들을 개발한 후 이를 수학 교육자들과 연구자들이 각 문항들의 일관성과 편향성을 검토하였다. 평가 문항은 7가지 유형으로, 평가에 사용하기 쉽고 문항 간 신뢰도를 높일 수 있는 형태로 개발하였다. 문항 유형은 다음과 같다.

가. 기초 계산과 기호 표현 과제(basic computation and symbolic representation tasks)

이 문항은 학생들이 문제를 인식하고 성공적으로 해결할 수 있는지를 보기 위한 유형으로, 간단하면서 대표적인 예는 ‘ $6(3+1) = 6 \times \square + 6 \times 1$ ’과 같다.

단계	문제 풀이 과정
1	$x + 5 = 10$
2	$x + 5 - 5 = 10 - \square$
3	$x = \square$

[그림 II-3a] 문항의 예 1

나. 풀이 과정을 빈 칸에 채우는 문제(partially worked problems without explanation)

이 문항은 주어진 문제 풀이 과정의 내용을 읽고 이해한 다음, [그림 II-3a]와 같이 빈 칸에 숫자나 문자를 채워 넣거나 [그림 II-3b]와 같이 풀이 단계를 채워 넣는 유형이다.

다. 풀이 과정에 대한 이유를 빈 칸에 채우는 문제(partially worked problems with explanations)

이는 문제에서 주어진 풀이 과정의 이해를 토대로 이 과정에 요구되는 원리나 법칙 등을 설명하도록 하는 유형이며, 이의 예는 다음 [그림 II-4]와 같다.

단계	단순화 과정
1	$2(7+4)$
2	$2 \cdot 7 + 2 \cdot 4$
3	
4	22


[그림 II-3b] 문항의 예 2

④ 한 학생이  $\frac{7}{12} + \frac{1}{4}$  을 다음과 같은 방법으로 풀었다. 이 학생이 두 번째 단계에서  $\frac{1}{4}$  에  $\frac{3}{3}$  을 곱한 이유를 설명하여라. 이 때 사용한 수학적 규칙이나 원리에 대해 설명하여라.

문제 해결 단계	설명
1	$\frac{7}{12} + \frac{1}{4} =$ 이 분수들은 분모가 다르다. 이 분수들을 더하려면 공통분모를 찾아야 한다.
2	$\frac{7}{12} + (\frac{1}{4} \times \frac{3}{3})$ a) 이 학생은 왜 $\frac{1}{4}$ 에 $\frac{3}{3}$ 을 곱하였는가? <input type="text"/> b) $\frac{1}{4}$ 에 $\frac{3}{3}$ 을 곱하는 것이 수학적으로 옳은가? 이유를 설명하여라. <input type="text"/>
3	$\frac{7}{12} + (\frac{3}{12}) =$ $(\frac{1}{4} \times \frac{3}{3})$ 의 괄호 안의 분수를 곱하였다.
4	$\frac{10}{12}$

[그림 II-4] 문항의 예 3



구체물	시각자료 1
발문	두 개의 액수를 더하면 얼마일까요? (학생들은 \$7이라고 말할 것이다). (\$5짜리 지폐를 흔들면서 “하나”라고 하고, \$2짜리 지폐를 흔들면서 “둘”이라고 하면서) 만약에 어떤 초등학교 1학년 학생이 둘을 더하면 둘이라고 말하면 어떻게 하겠어요?
화면	화면2 
구체물	2달러 지폐를 1달러짜리 2장으로, 5달러 지폐를 1달러짜리 다섯 장으로 교체하는 것을 보여준다.
	(중략)
판서	$\frac{1}{4}, \frac{1}{2}, \frac{1}{10}, \frac{1}{100}$
	(이하 생략)

[그림 II-8] 중2 유리수 동치 수업 자료1의 일부

수업 자료2	유리수 동치
<p>분수를 더하기 위해 활용한 똑같은 규칙을 분수들의 뺄셈에서도 적용할 수 있다. 다음 두 분수들을 더해 보자.</p> $\frac{4}{5} + \frac{-3}{b} =$	
<p>a) 이 문제의 분모들은 무엇인가? <i>분모들은 5와 b 이다.</i></p>	
<p>b) 왜 이 분수들을 당장 더할 수 없는지 그 이유를 설명하시오. <i>두 개의 분모가 다르다. 우리는 공통분모를 구해서 같은 것끼리 더해야 한다.</i></p>	
<p>c) 공통분모는 무엇인가? <i>공통분모는 5b 이다.</i></p>	
<p>d) 첫 번째 계산 단계는 무엇인가? 공통분모를 가진 값이 같은 분수들을 찾기 위해 다음 박스들을 활용하시오.</p>	
<p>Step 1: <math>\frac{4}{5} \cdot \begin{array}{ c } \hline b \\ \hline b \\ \hline \end{array} + \frac{-3}{b} \cdot \begin{array}{ c } \hline 5 \\ \hline 5 \\ \hline \end{array} = ?</math></p>	
<p>e) 다음 빈칸에 값이 같은 분수를 쓰시오.</p>	
<p>Step 2: <math>\frac{4b}{5b} + \frac{-15}{5b} = ?</math></p>	
<p>f) 1단계에서 그렇게 계산한 이유를 설명하시오. <i>양쪽 분수값을 바꾸지 않기 위해서 양쪽 분수에 1을 곱했다. 0이 아닌 숫자인 경우 그 숫자 분의 그 숫자는 1이 된다.</i></p>	
<p>g) 2단계에서 그렇게 계산한 이유를 설명하시오. <i>분수들을 곱하기 위해 분자들끼리 곱하고, 그리고 나서 분모들끼리 곱한다.</i></p>	
(이하 생략)	

[그림 II-9] 중2 유리수 동치 수업 자료2의 일부

### 3. 대수 영역의 수업 자료

PowerSource<sup>®</sup>는 교사가 진행할 수업을 지원하기 위해 판서와 도식 자료, 설명 방법 및 내용을 포함하여 수업 방법을 상세하게 제시하고 있다. [그림 II-8]은 유리수 동치의 기본 개념 및 원리를 학습시키기 위한 수업 자료로써 1차시에 해당하며, [그림 II-9]는 전 평가에 대한 피드백을 제공하는 수업 자료로써 2차시에 해당한다. 이전 평가의 결과에 따라 교사가 개별 학습이나 조별 활동 등 선택적으로 사용할 수 있다.

## III. 연구 방법 및 절차

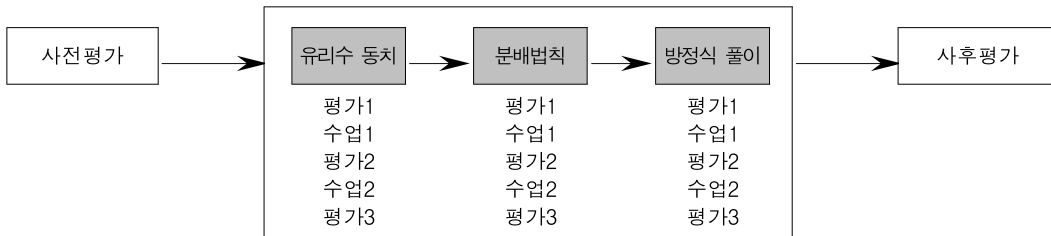
### 1. 표집 선정

미국의 경우 최소한 2개 주 이상을 선정하여 연구를 실시하였으며, 우리나라 연구에서는 학교 및 지역 분포를 고려하여 총 10개 학교(서울 5개, 경기 3개, 인천 1개, 광주 1개)를 선정하고, 각 학교에서 무작위로 2개 반을 선정하여 한 반은 실험반, 다른 한 반은 통제군으로 배정하였다. 총 실험반 학생의 수는 385명이고 통제반 학생의 수는 239명이다.<sup>6)</sup> 이 실험 연구에 참여할 교사들에게 두 차례 연수를 실시

하였다. 첫 번째는 PowerSource<sup>®</sup> 자료의 개발 배경, CRESST의 연구 내용 및 결과에 대한 연수이며, 두 번째는 PowerSource<sup>®</sup> 자료의 내용과 활용 방법, 형성평가 실시 방법 및 사후 조치에 대한 연수이다. 이때 제시된 자료들은 PowerSource<sup>®</sup> 평가 문항 및 교수 학습 지침서에 대한 활용 방법뿐만 아니라 CRESST가 교사 연수 시 사용하였던 연수 자료 및 관련 연구물도 포함한다.

### 2. 연구 실행 절차

전체 PowerSource<sup>®</sup> 프로그램을 실행하기 전과 후에, 학생들의 성취도 향상 정도를 파악하기 위해 사전평가와 사후평가를 실시하였다. 이 평가는 27개의 객관식 문제로 유리수 동치, 분배법칙, 방정식 풀이 세 주제에 관한 문항을 모두 포함하고 있다. 통제집단의 수업은 PowerSource<sup>®</sup> 프로그램 절차를 따르지 않고 사전평가 후에 각 주제별로 학습 요소를 추출하여 일상적인 방법으로 수업을 진행하고 사후평가를 실시하였으며, 이때 특별한 용어나 별도의 설명이 필요한 경우 교육과정에서 벗어나지 않는 수준에서 간결하게 설명하는 것을 허용하였다. 실험반은 PowerSource<sup>®</sup> 프로그램 절차에 따라 진행하였다. 프로그램 절차는 [그림 III-1]과 같이 각 주제별로 평가1→수업1→평가2→수업2→



[그림 III-1] 프로그램 절차

6) 실험반과 통제반 모두 평가 시간의 학생 출결 상황에 따라 각 평가에 참여하는 전체 학생 수에 다소 차이가 있음.



평가3의 순서로 진행된다. 학생들에게 평가1은 20~25분, 평가2와 평가3은 15분을 주었다. 각각의 평가를 실시한 후 학생들의 부족한 부분을 보충하기 위해 문제별 오답 수와 응답 유형을 정리함으로써 복습하거나 다시 가르쳐야 할 개념을 파악하여 다음 수업에 반영하도록 하였다.

#### IV. 대수 영역의 PowerSource© 실행 결과 분석

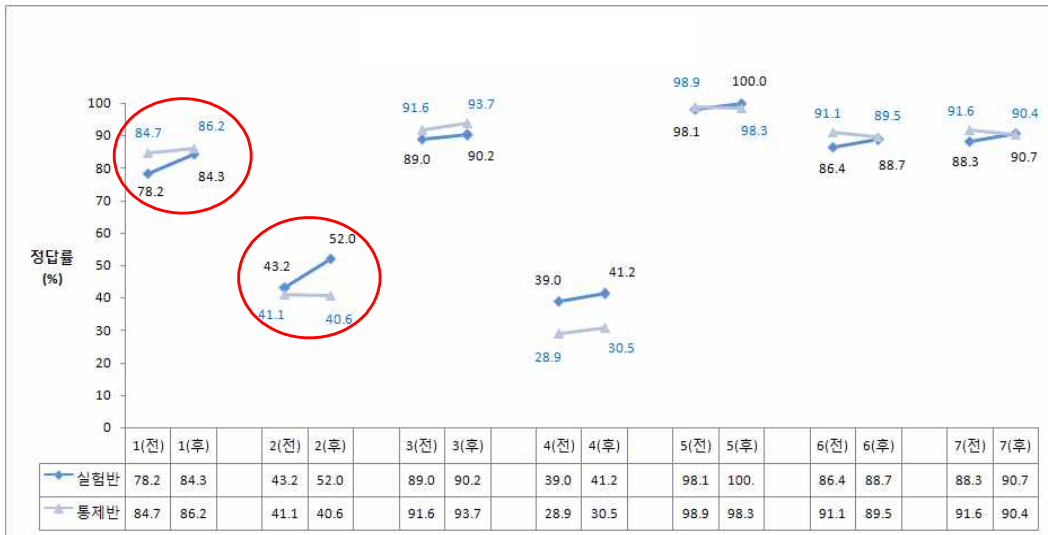
PowerSource© 투입 전과 후에 실시한 진단 평가의 정답률을 비교하고, 형성평가에서 동형 문항에 대해 응답의 변화를 분석하였다.

##### 1. 사전·사후평가의 정답률 비교)

##### 가. 사전·사후평가의 정답률 변화

실험반과 통제반의 사전·사후평가 결과를 문항별 정답률의 변화로 비교·분석하였다. 다음 [그림 IV-1], [그림 IV-2], [그림 IV-3], [그림 IV-4]는 각각 문항별 정답률을 나타낸 그래프이다. 실험반과 통제반 모두 전 문항에서 전반적으로 정답률이 상승하였음을 알 수 있으며, 전체적으로 실험반 결과의 상승률이 더 높음을 알 수 있다. 특히 그래프에서 실선으로 표시된 진단평가의 1번, 2번, 9번, 10번, 11번, 18번, 21번, 24번, 25번 문항의 뚜렷한 차이는 PowerSource©의 확연한 학습 효과가 돋보인다고 판단할 수 있다. 한편, 학습한 후에 오히려 정답률이 떨어지거나, 변화가 거의 없는 문항들은 점선으로 표시하였다.

그래프 분석에서도 전체적으로 실험반의 결과가 상승하였음을 알 수 있지만 실제 평균

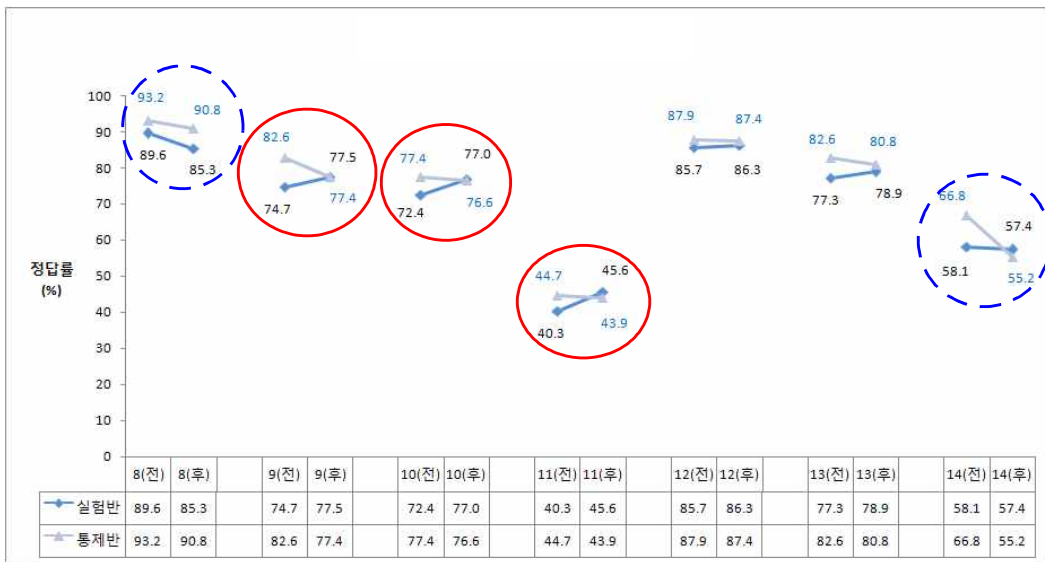


[그림 IV-1] 실험반과 통제반의 사전·사후평가 결과 비교 (1번~7번 문제)

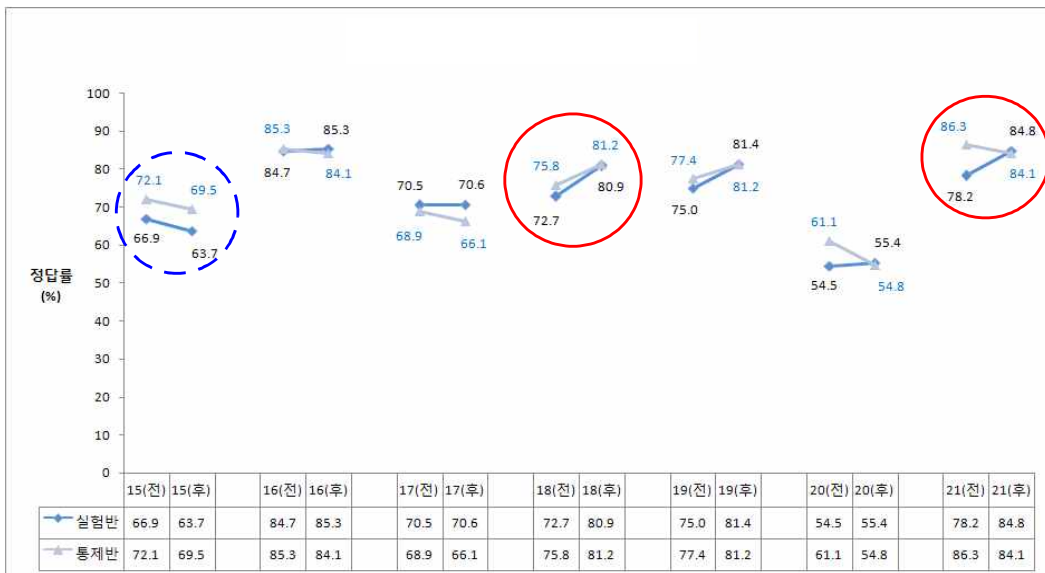
7) 실험반과 통제반 각각의 사전·사후평가에 대한 변화를 분석하기 위해 t-검증을 시도하였으나 참여한 교사들의 데이터 관리가 제대로 이루어지지 않아 학생 개인별 자료 획득에 어려움이 있었다. 일부 보존된 데이터로부터 t-검증을 실시하였고, 실험반과 통제반의 차이는 있으나 수치상 유의미한 차이가 없는 것으로 나타났다. 그러나 이는 전체 데이터가 아니라 일부 학교 학생들의 데이터로서 전체에 대한 설명을 대표할 수 없으므로, 본 연구에서는 실험반과 통제반의 평균 정답률 변화로 향상 정도를 파악하였다. 본 연구 결과는 사전·사후 평가의 정답률 변화뿐만 아니라 세 차례의 형성평가 결과를 통해 응답 유형에서 질적으로 유의미한 차이를 볼 수 있었음.

정답률에서도 나타났다. 실험반 평균 정답률은 70.9%에서 73.7%로 상승하였으나, 통제반 평균 정답률은 74.8%에서 73.7%로 하락하였다. 실험반의 사전평가 결과를 볼 때 PowerSource© 투입 전에는 통제반보다 학습성취도가 낮았던 것으로 나타났다. 이는 연구에 참여한 교사들이

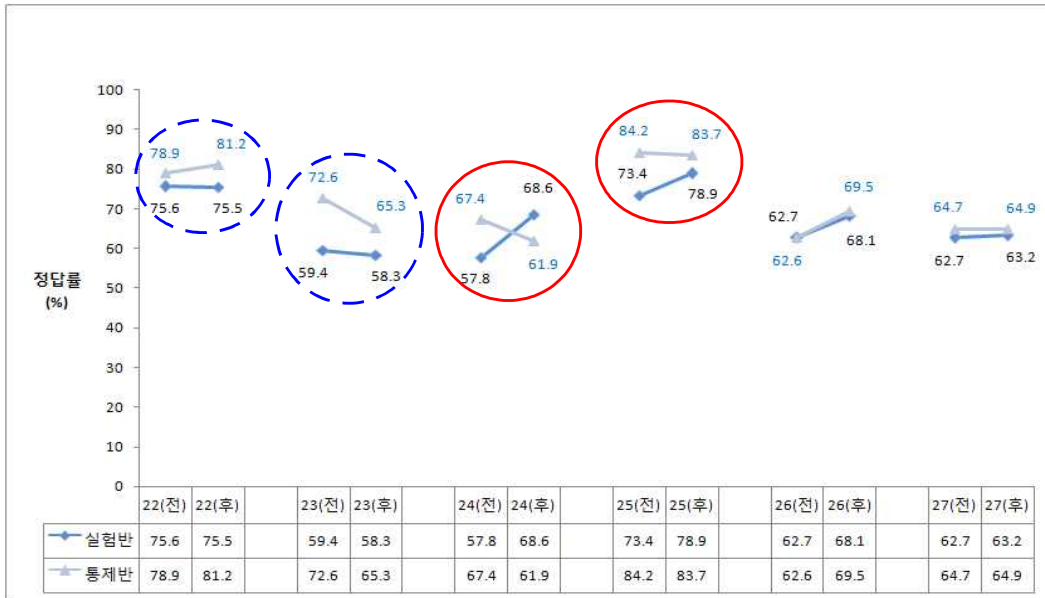
실험반을 통제반에 비해 학습 성취 수준이 더 낮은 반을 선택하였기 때문이다. 그러나 PowerSource© 사용 후 실험반의 사후평가 결과를 통제반과 비교할 때 실험반이 더 향상되었음을 알 수 있다. 한편 각 문항별로 정답률의 변화를 보면 실험반의 정답률이 통제반에 비해



[그림 IV-2] 실험반과 통제반의 사전·사후평가 결과 비교 (8번~14번 문제)



[그림 IV-3] 실험반과 통제반의 사전·사후평가 결과 비교 (15번~21번 문제)



[그림 IV-4] 실험반과 통제반의 사전·사후평가 결과 비교 (22번~27번 문제)

많이 상승한 경우가 있는 반면, 변화가 없거나 다소 떨어진 경우도 발견할 수 있었다. 후자의 문항에 대해서는 통제반 학생들도 어려움이 있는 것으로 보이며, 관련 개념 및 원리에 대한 수업과 평가 문항의 보강이 필요하겠다.

#### 나. 문항별 학생 답안의 변화

사후평가에서 통제반과 비교할 때 실험반의 정답률이 큰 폭으로 상승한 문항에 대해서 학생들의 답안의 변화를 분석하였다.<sup>8)</sup> 각 문항의 물음 내용 및 사전 사후 평가 결과는 <부록 1>에 일괄 제시되어 있다.

진단평가 1번 문항의 경우, 정답인 ①번 답안을 선택한 학생이 사전평가에서 78.2%, 사후평가에서 84.3%로 증가하였다. 반면 ②번과 ④번 답안을 선택한 학생의 비율이 감소하였다. 이는 PowerSource© 프로그램에서 문자가 있는 식에서 분배법칙을 사용할 수 있음을 학습할

때 다루는 기초적인 내용으로 수업 전반에 걸쳐서 다룸으로써 학습되었을 것으로 여겨진다. 진단평가 2번 문항의 경우에는 정답인 ③번 답안을 선택한 학생이 사전평가에서 43.2%, 사후평가에서 52.2%로 증가하였다. 사전평가에서는 ④번 답안을 선택한 학생이 51.3%로 가장 많았으나 사후평가에서는 ④번 답안이 줄고 정답인 ③번 답안을 선택한 것으로 보인다. 통제반에서는 정답률이 41.1%에서 40%로 줄고, ④번 답안을 선택한 학생이 52%에서 55.6%로 늘어난 것으로 볼 때, 실험반 학생들이 Power- Source© 자료로 학습한 후 괄호 묶는 방법에 대한 학습이 원만히 이루어진 것으로 판단된다. 진단평가 9번 문항의 경우에는, 실험반에서 정답인 ②번 답안을 선택한 학생이 사전평가에서 74.7%, 사후평가에서 77.5%로 증가하였다. 상대적으로 ③번 답안을 선택한 학생의 비율이 16.2%에서 14.5%로 감소하였다. 비교하여 통제반에서는 정

8) 사후평가의 정답률이 하락하거나 변화가 없는 문항에 관해서는 정도가 1~2% 내외의 매우 근소한 차이며 내용에서도 단순 계산의 오류나 부주위로 인한 오류이므로 분석 결과를 기술하지 않음.

답률이 82.6%에서 77.4%로 상당히 감소하면서 다른 오답들이 약간씩 증가하였다. 이는 PowerSource©의 유리수 동치 수업에서 종이테이프를 이용하여 4등분한 점을  $\frac{1}{4}$ 로 표시하고 다시 5등분한 점을  $\frac{1}{20}$ 으로 표시하는 등 1보다 작은 분수에 대한 이해를 돕기 위한 활동 수업의 효과라 볼 수 있다.

또, 진단평가 10번 문항의 경우, 실험반에서 정답인 ①번 답안을 선택한 학생이 사전평가에서 72.4%, 사후평가에서 77.0%로 증가하였다. 상대적으로 무응답의 비율과 ②번 답안을 선택한 학생의 비율이 감소하였다. 비교하여 통제반의 정답률은 77.4%에서 76.6%로 하락하면서 무응답이 증가하였다. 이는 PowerSource©의 유리수 동치 수업에서 분수의 값이 1이 되도록 하는 반복 학습의 효과라 볼 수 있다. 진단평가 11번 문항의 경우에는 실험반에서 정답인 ②번 답안을 선택한 학생이 사전평가에서 40.3%, 사후평가에서 45.6%로 증가하였다. 비교하여 통제반에서는 정답률이 44.7%에서 43.9%로 하락하면서 ①번 답안을 선택한 학생의 비율이 44.7%에서 49.0%로 증가하였다. 이것은 PowerSource©의 유리수 동치 수업에서 분모와 분자가 같으면 1이 되고, 이러한 1을 어떤 분수에 곱하더라도 원래 분수의 크기는 변하지 않음을 여러 가지 예를 통해 반복 학습한 효과라 볼 수 있다. 다만, ①번 답안을 선택한 경우도 46.1%로 여전히 정답보다 많은 것은 학생들이 단순히 한번 약분한 경우의 답인 ①번 답안  $\frac{26}{39}$ 을 선택한 것으로 보인다. 그러나 만약 문제에서 답안의 순서가 바뀌었다면 그 결과는 달라졌으리라 본다. 즉, 학생들은 일차적으로 약분한 답이라고 생각한 것이 앞쪽에 있으므로 이를 선택한 것으로 여겨진다. 이는 학생들이 간단히 하는 방법을 알고 있으나 매력적인 오

답으로 인하여 오류를 범한 것으로 판단된다.

진단평가 18번 문항의 경우, 정답인 ③번 답안을 선택한 학생이 사전평가에서 72.7%, 사후평가에서 80.9%로 증가하였다. 상대적으로 ①번 답안을 선택한 학생이 15.6%에서 8.8%로 감소하였다. 이는 통제반의 정답률이 86.3%에서 84.1%로 감소한 것으로 볼 때 PowerSource© 프로그램에서 분배법칙에 대한 학습이 원만히 이루어진 결과라 할 수 있다. 진단평가 21번 문항의 경우에는 정답인 ①번 답안을 선택한 학생이 사전평가에서 78.2%, 사후평가에서 84.8%로 증가하였으며 상대적으로 무응답의 비율이 11.0%에서 4.9%로 감소하였다. 비교하여 통제반에서는 정답률이 86.3%에서 84.1%로 줄고, 무응답이 많아졌다. 이는 PowerSource© 프로그램에서 일차방정식의 풀이 과정의 모든 단계를 순차적으로 제시하면서 설명하고, 각 단계에서 식이 전개되는 이유를 밝히고 학생들이 이해 시키는 수업 내용의 영향으로 보인다. 진단평가 24번 문항의 경우에는 정답인 ③번 답안을 선택한 학생이 사전평가에서 57.8%, 사후평가에서 68.6%로 증가하였고 무응답이 상당 수 줄었다. 통제반의 경우 정답률이 67.4%에서 61.7%로 감소한 것으로 볼 때, PowerSource©에서는 방정식의 풀이에서 첫 번째 단계를 강조하여 반복적으로 학습하도록 한 것에 영향을 받은 것으로 판단된다. 끝으로, 진단평가 25번 문항의 경우에는 정답인 ②번 답안을 선택한 학생이 사전평가에서 73.4%, 사후평가에서 78.9%로 증가하였다. 통제반에서는 정답률이 67.4%에서 61.9%로 감소하면서 무응답이 많아졌다. 이 문제의 경우 문장제 유형으로 선택형으로 제시되어 학생들이 어떤 과정으로 이러한 답을 구했는지 제대로 파악하기 어렵다. 그러나 통제반에서는 무응답이 5%이상 증가한 것에 반해 실험반에서는 6% 이상 감소하였음을

볼 때 실험반 학생들이 문제를 해결하고자 시도하고 있음을 알 수 있다.

## 2. 평가 간 동형 문항별 응답 유형 비교 · 분석

각 주제에서 초기 평가인 평가1의 각 문항과 동형은 평가2, 평가3에 포함되어 있으며, 이에 관한 ‘유리수 동치’ 부분의 예가 <부록 2>에 제시되어 있다. 이 절에서는 동형 문항의 응답 유형을 분석함으로써 학생들의 이해도와 학습의 효과를 검증하고자 하였다.

### 가. 유리수 동치

유리수 동치에서 초기 평가인 평가1의 각 문항에 대한 평가2와 평가3에서의 동형 문항 중에서 응답 유형에 의미가 있거나 변화가 보이는 몇몇 문항들을 분석하면 다음과 같다.

평가1의 3번 문항은 분모에 문자가 포함된 두 분수의 덧셈이다. <표 IV-1>에서 학생들의 응답 유형을 볼 때 정답률뿐만 아니라 학생들의 생각을 읽을 수 있는 오답 유형까지도 알 수 있다. 수업 전에 이루어진 평가1에서는 학생들이 문자와 숫자를 통분하여 덧셈을 하는 것을 생각조차 못하는 학생들이 50% 이상이었으나, 수업1을 진행한 후 평가2에서는 정답률이 70.4%로 상당히 향상된 것으로 나타났다. 뿐만 아니라 평가1에서는 무응답이 28.1%였던 것에 비해 평가2, 평가3에서는 10% 정도로 나타난 것은 학생들이 분모에 문자를 포함하는 다소 복잡한 형태의 분수의 덧셈 문제에 접근할 수 있게 된 것으로 판단된다. 이는 평가 결과를 다음 수업에 반영하는 형태의 PowerSource© 프로그램이 학생들의 이해를 높이고 수학에 대한 의욕을 불러일으키는데 영향을 미쳤음을 보여주는 결과이다.

평가1의 6번과 7번 문항은 분수의 덧셈에서

제시한 첫 번째 단계의 풀이가 옳은지 판단하고, 그 이유를 설명하는 문제이다. 6번의 식은 문자를 포함하지 않은 반면 7번의 식에서는 분모가 문자인 분수라는 차이를 두고 있다. <표 IV-2>에서 평가1의 정답률은 58.4%, 평가2의 정답률은 60.6%로 큰 변화가 없는 듯 보이나 그 응답 유형을 보면 큰 차이가 있음을 알 수 있다. 평가1에서는 학생들이 문제의 답을 알기는 하지만 제대로 용어를 사용하지 못하는 답이 많았던 반면, 평가2에서는 정확한 설명과 용어를 사용하여 답하였다. 이는 PowerSource© 프로그램의 수업 방법에 있어서 여러 가지 예를 최대한 활용하여 반복적으로 설명하거나 일부 내용을 집중적으로 설명하는 등의 형태로 진행됨에 따라 학생들로 하여금 개념을 명확히 하고 명료한 용어를 사용하게 하였음을 보여준다.

또, <표 IV-3>에서와 같이 평가1의 정답률은 39.5%, 평가3의 정답률은 60.9%로 큰 변화가 나타났으며 그 응답 유형 또한 큰 차이가 있음을 알 수 있다. 평가1에서는 학생들이 개념을 알기는 하지만 제대로 용어를 사용하지 못한 정답이 많았던 반면, 평가3에서는 다양한 설명 방법을 동원하여 답하였다는 점이다. 이 문제의 경우, 앞서 제시된 평가1의 6번 문항, 평가2의 5번 문항과 비슷한 형태인 계산 과정을 쓰고 그 이유가 정당함을 밝히는 문제이나 이 문제의 형태는 문자도 포함되었다는 점에서 차이가 있다. 이러한 문항은 이유를 설명하는 경우로 학생들이 많이 어려워하는 유형 중 하나이다. 그러나 PowerSource© 수업에서 반복적으로 교사의 발문을 통해 학생들의 반응을 이끌어내면서 정당한 이유를 설명하도록 하고 있음을 볼 때, <표 IV-3>에서와 같이 학생들이 여러 가지 방법으로 이유를 설명할 수 있었다는 것은 수업의 큰 효과라 볼 수 있다.

<표 IV-1> 중2 유리수 동치 평가1의 3번, 평가2의 3번, 평가3의 2번 응답 유형 비교<sup>9)</sup>

평가1		평가2		평가3	
문3) $\frac{3}{7} + \frac{2}{m}$		문3) $\frac{5}{7} + \frac{2}{p}$		문2) $\frac{2}{n} + \frac{4}{9}$	
응답 유형	응답률(%)	응답 유형	응답률(%)	응답 유형	응답률(%)
$\frac{3m+14}{7m},$ $\frac{3m}{7m} + \frac{14}{7m}$	42.6	$\frac{5p+14}{7p}$	68.8	$\frac{18+4n}{9n},$ $\frac{18}{9n} + \frac{4n}{9n}$	71.5
$\frac{17m}{7m}$	4.4	$\frac{5}{7} \times \frac{p}{p} + \frac{2}{p} \times \frac{7}{7}$	1.6	$\frac{22n}{9n}$	4.9
$\frac{17}{7}, 2\frac{3}{7}$	4.4	$\frac{19p}{7p}, \frac{19}{7}$	6.9	$\frac{2}{9n} + \frac{4}{9n}$	1.3
$\frac{5}{7}$	3.4	기타	11.1	$\frac{22}{9n}, \frac{22n}{9}$	0.5
$\frac{5}{7m}$	3.9	무응답	11.6	기타	10.9
기타	13.2	계	100.0	무응답	10.9
무응답	28.1			계	100.0
계	100.0				

<표 IV-2> 중2 유리수 동치 평가1의 6b)번, 평가2의 5b)번 응답 유형 비교

평가1		평가2	
문6b) $\frac{4}{5} + \frac{3}{7} = \frac{4}{5} \times \frac{7}{7} + \frac{3}{7} \times \frac{5}{5}$ 옳은 이유?		문5b) $\frac{3}{4} + \frac{2}{5} = \frac{3}{4} \times \frac{5}{5} + \frac{2}{5} \times \frac{4}{4}$ 옳은 이유?	
응답 유형	응답률(%)	응답 유형	응답률(%)
분수의 값은 똑같이 유지하면서 공통분모를 찾기 위해	23.6	분수의 값은 똑같이 유지하면서 공통분모를 찾기 위해	23.3
통분해야 하므로	15.6	통분해야 하므로	15.6
통분이 필요하기 때문에 분모의 최소공배수를 구해야하고, 분자와 분모에 같은 수를 곱해야하기 때문	4.2	분모를 같게 만들어야 하므로	12.7
분수의 값이 변하지 않으므로	3.6	곱셈의 항등원 1을 곱해도 그 값이 변하지 않으므로	4.5
분모를 35로 (같게) 만들어야 하므로	3.1	분자와 분모에 같은 수를 곱해야 하므로	2.4
최소공배수가 나오게 통분하여 계산해야 하므로	3.1	공통분모를 구하기 위해 분모의 공배수를 구해야 하고 분모에 곱한 수를 분자에도 곱해야 하므로	2.1
5와 7의 최소공배수는 35로 $\frac{4}{5}$ 에 7을 분자와 분모에 곱하고 $\frac{3}{7}$ 에 5를 분자와 분모에 곱해야 하므로	2.3	통분해야 하므로 최소공배수를 만들기 위해 4에는 5를 5에는 4를 곱하고 분자에도 같은 수를 곱했으므로	2.9
분모를 통분할 때 분자도 같은 수로 곱했으므로	1.6	최소공배수가 되도록 분자와 분모에 곱한다	0.5
1을 곱하는 것이므로	1.0	기타 및 옳지 않다	15.6
항등원 1을 곱해도 같은 값이다	0.3	무응답	20.4
최소공배수를 곱해야 한다	0.3	계	100.0
계산이 옳다	2.3		
기타 및 옳지 않다	13.2		
무응답	25.7		
계	100.0		

9) <표 IV-1>에서 음영 처리된 부분은 정답에 관한 부분을 나타낸 것이며, 이후 문항의 경우에도 마찬가지임.

<표 IV-3> 중2 유리수 동치 평가1의 7b)번, 평가3의 4b)번 응답 유형 비교

평가1		평가 3	
문7b) $\frac{2}{a} + \frac{5}{b} = \frac{2}{a} \times \frac{b}{b} + \frac{5}{b} \times \frac{a}{a}$ 옳은 이유?		문4b) $\frac{4}{7a} + \frac{3}{5b} = \frac{4}{7a} \times \frac{5b}{5b} + \frac{3}{5b} \times \frac{7a}{7a}$ 옳은 이유?	
응답 유형	응답률(%)	응답 유형	응답률(%)
분수의 값은 똑같이 유지하면서 공통분모를 찾기 위해	20.8	분수의 값은 똑같이 유지하면서 공통분모를 찾기 위해	16.1
분모를 같게 만들어야 하므로	2.6	분수의 값은 변하지 않으므로	5.7
항등원 1을 곱해도 같은 값이다	1.3	(공통분모를 만들기 위해, 분모가 같아지게, 통분하기위해)1을 곱했으므로	3.4
(최소공배수를 이용하여) 통분해야 하므로	14.0	공통분모를 만들기 위한 과정이므로	2.6
분수의 값이 변하지 않으므로	0.8	항등원 성질 활용해서 통분하였으므로	7.0
서로의 분모를 곱했기 때문	3.1	분모를 같게 만들어야 하므로 분모를 같게 만들면서 분자에도 같은 수를 곱했으므로	8.3
$\frac{2}{a} + \frac{5}{b} = \frac{2}{a} \times \frac{1}{b} + \frac{5}{b} \times \frac{1}{a}$	0.3	분모를 통분해야 하므로	8.3
설명 없이 계산과정 씌	0.5	통분으로 단위를 맞춰줘야 하므로	8.0
최소공배수를 곱해야 한다	3.9	통분해서 분자 분모가 같아지게 하기 위해 최소공배수를 구하여 통분하고 같은 수를 분자에도 곱했으므로	1.6
기타 및 옳지 않다	14.3	최소공배수가 되도록 같은 수를 곱했으므로	3.1
무응답	38.3	각 분수에 서로의 분모를 곱했으므로	0.8
계	100.0	설명 없이 식을 폼	0.5
		기타 및 옳지 않다	8.3
		무응답	26.4
		계	100.0

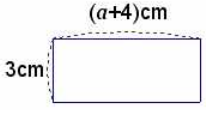
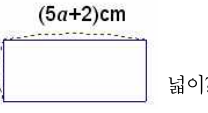
나. 분배법칙

분배법칙에서 초기평가인 평가1의 응답 유형과 비교할 때, 평가2와 평가3의 응답 유형에

의미가 있거나 변화가 보이는 몇몇 동형 문제들을 분석하면 다음과 같다.

평가1의 8번 문항은 대수의 성질에서 분배법

<표 IV-4> 중2 분배법칙 평가1의 8번, 평가2의 5번 응답 유형 비교

평가1		평가2	
문8)  둘레?		문5)  넓이?	
응답 유형	응답률(%)	응답 유형	응답률(%)
$2(a+4) + 2 \times 3$ , $2(a+4+3)$ $2(a+7)$ , $2(a+4)+6$	54.9	$7(5a+2)$ $7 \times 5a + 7 \times 2$	73.8
$3+3+a+4+a+4$	4.1	$35a+14$	6.7
$3(a+4)$ , $3a+12$ (넓이를 구한 경우)	5.4	$2(5a+2+7)$ 등 (둘레를 구한 경우)	5.4
기타	11.1	기타	6.7
무응답	24.4	무응답	7.5
계	100.0	계	100.0

칙을 이용하여 직사각형 둘레의 길이를 구하는 문제이고, 평가2의 5번 문항은 직사각형의 넓이를 구하는 문제이다. <표 IV-4>에서와 같이 문제의 내용상 응답 유형이 다양하게 나타날 수 있는 문항은 아니므로 정답률의 변화를 보면, 평가1의 정답률은 59.0%, 평가2의 정답률은 80.4%로 상승하였다. 특히 평가1에서는 답을 하지 못하는 무응답이 24.4%로 나타난 반면, 평가2에서는 7.5%로 상당히 낮아졌다. 이를 통해 미지수가 포함된 식에서 분배법칙을 이용할 수 있음을 학습한 효과임을 알 수 있다. 평가1의 10번은 <표 IV-5>에서와 같이 분배법칙을 사용하는 이유를 묻는 문제로 평가3의 5번 문항이 동형이다. 평가1의 정답률은 48.3%, 평가3의 정답률은 58.0%로 큰 변화가 나타났으며 그 응답 유형을 보면 큰 차이가 있음을 알 수 있다. 평가1에서는 제대로 설명을 못하여 무응답이 38.1%로 높았던 반면, 평가3에서는 무응답이 29.8%로 낮아졌다는 점이다. 이는 PowerSource© 수업에서 반복적으로 교사의 발문을 통해 학생들의 반응을 이끌어내면서 정당한 이유를 설명하도록 하는 방법의 효과라 할 수 있다.

다. 방정식의 풀이

방정식의 풀이에서 초기평가인 평가1의 응답 유형과 비교할 때, 평가2와 평가3의 응답 유형에 의미가 있거나 변화가 보이는 몇몇 동형 문제들을 분석하면 다음과 같다.

평가1의 6번은 일차방정식의 풀이 과정에서 제시한 첫 번째 단계의 옳고 그름을 판단하고 이유를 설명하는 문제이다. 평가1의 정답률은 34.7%, 평가2의 동형문항의 정답률은 46.8%로 증가하였으며, <표 IV-6>의 응답 유형을 보면 큰 차이가 있음을 알 수 있다. 특히 평가2에서는 풀이 과정 설명을 다양한 방법으로 서술하였음을 알 수 있다. 또, 평가1에서는 무응답이 43.5%인 반면, 평가2에서는 무응답이 32.8%로 줄어든 것으로 보아 수업 후 학생들이 답안에 접근할 수 있게 된 것으로 사료된다. 또, 평가1의 10번 문항은 문장제이며 동형인 문항은 평가3의 3번이다. 이 문제는 다른 문장제와 비교할 때 비교적 높은 정답률을 보였는데, 이는 문제에 주어진 요소에 대해 변수를 지정해 주었기 때문이라 사료된다. <표 IV-7>에서 정답률과 응답 유형을 보면, 평가1의 10번 문항의

<표 IV-5> 중2 대수의 성질 평가1의 10번, 평가3의 5번 응답 유형 비교

평가1		평가3	
문10) $5 \times 3\frac{1}{4}$ 을 $(5 \times 3) + (5 \times \frac{1}{4})$ 와 같이		문5) $4 \times 7\frac{1}{2}$ 을 $(4 \times 7) + (4 \times \frac{1}{2})$ 와 같이	
계산하는 이유는?		계산하는 이유는?	
응답 유형	응답률(%)	응답 유형	응답률(%)
분배해서 곱해야 하므로	7.3	분배해서 곱해야 하므로	13.9
두 수에 공평하게 곱해주어야 하므로	3.8	두 수에 공평하게 곱해 주어야 하므로	5.9
쉽게 풀기 위해	21.9	쉽게 풀기 위해	27.7
간단히 하기 위해	12.3	간단히 하기 위해	10.5
따로따로 계산해서 더해야 답이 나오므로	1.6	괄호를 풀어야 하므로	1.3
$3\frac{1}{4}$ 은 $3 + \frac{1}{4}$ 이므로	0.3	자연수와 분수를 각각 계산해야 하기 때문	1.3
기타	13.7	규칙이다, 그래야 풀 수 있다	2.1
무응답	38.1	기타	7.6
계	100.0	무응답	29.8
		계	100.0



<표 IV-6> 중2 일차방정식 평가1의 6번, 평가2의 3번 응답 유형 비교

평가1		평가2	
문6) $6a + 12 = 48$ 의 풀이 과정의 첫 번째 단계 $\frac{6p}{p} + \frac{12}{6} = \frac{48}{6}$		문3) $48 = 8p + 16$ 의 풀이 과정의 첫 번째 단계 $\frac{48}{8} = \frac{8p}{8} + \frac{16}{8}$	
응답 유형	응답률(%)	응답 유형	응답률(%)
옳지 않다. 같은 수로 나누어야 한다.	34.7	옳다. $p$ 를 구하기 위해 $p$ 의 계수를 1로 하기 위해	6.4
옳지 않다. 12를 먼저 이항시켜야 한다.	13.4	옳다. 문자( $p$ )만 남기기 위해	11.1
기타	8.4	옳다. 같은 수로 나누어도 등호가 성립하므로	17.0
무응답	43.5	옳다. 양변을 모두 8로 나누었으므로	6.8
		옳다. 분모를 없애기 위해	2.1
		옳다. 간단히 하기 위해	3.4
		이항하여 푼다	6.8
		기타	12.6
		무응답	32.8
계	100.0	계	100.0

<표 IV-7> 중2 일차방정식 평가1의 10번, 평가3의 3번 응답 유형 비교

평가1		평가3	
문10) 통밀 과자의 3배에 해당하는 머쉬멜로우가 있다. 머쉬멜로우는 $m$ , 통밀 과자는 $g$		문3) 기린 수의 4배에 해당하는 뿔권이 있다. 뿔권의 수는 $p$ , 기린의 수는 $q$	
응답 유형	응답률(%)	응답 유형	응답률(%)
$m = 3g$	45.6	$p = 4q$	70.5
$3m = g$	8.4	$q = 4p$	7.7
$3m + g$	3.3	$x = 4x, y = 4x$ (다른 문자 사용)	0.4
기타	4.2	기타	6.8
무응답	38.5	무응답	14.5
계	100.0	계	100.0

정답률은 45.6%, 평가3의 3번 문항의 정답률은 70.5%로 급격히 증가하였으며, 평가1에서는 무응답이 38.5%인 반면, 평가3에서는 무응답이 14.5%로 줄어, 학생들이 문제를 해결하는데 필요한 학습이 이루어진 것으로 판단할 수 있다.

## V. 결론 및 제언

CRESST에서 개발한 PowerSource©로 우리나라 중학교 2학년 학생들을 대상으로 유리수 동치,

분배법칙, 일차방정식에 관한 수업 및 형성평가를 실시하였다. 사전·사후평가의 진단평가 결과를 분석해 보면, 실험반과 통제반 모두 모든 문항에서 전반적으로 정답률이 상승하였음을 알 수 있었다. 통제반과 실험반의 결과를 비교할 때 전반적으로 실험반 결과의 상승률이 더 높았으며, 특히 일부 문항의 경우 Power- Source©의 확연한 학습 효과가 나타나는 것으로 판단할 수 있었다. 또한 실험반의 세 차례의 형성평가에서 동형 문항에 대한 응답의 변화를 살펴보면 정답률은 차츰 높아졌고, 무응답의 수는 확연히 줄어들었

며, 애매모호한 풀이 과정이나 답을 제시하였던 학생들이 수학적 용어를 사용하여 좀 더 구체적으로 답하는 것을 확인할 수 있었다. 이는 여러 가지 예를 통한 반복적인 설명과 발문을 통해 학생들의 생각을 계속적으로 이끌어내어 설명하도록 하는 수업과 수 년 간의 평가 모델로부터 효과적이었던 문항들을 모아 각 핵심 개념 영역에 필요한 지식을 평가하기 위해 개발한 문항을 포함하는 PowerSource© 프로그램의 효과라 할 수 있다. 이러한 결과에 더하여, 본 연구로부터 도출될 수 있는 결론 및 제언은 다음과 같다.

현재 우리나라의 경우 학생에게 시험을 치르게 하여 성적 또는 등급을 부여하거나 선발을 위해 평가를 실시하는 선발적 교육관에 치중하고 있다. 많은 교사들은 학생을 평가하기 위한 출제, 채점, 우열 판정을 위해 많은 시간을 보내지만, 평가 결과를 통해 해당 학생이 가지고 있는 오개념이나 오류 유형에 대한 처치는 소홀히 하고 있다는 지적과 우려가 그다지 새삼스럽지 않다. Black과 Wiliam(1998a, 1998b)에 따르면, 교실에서의 주된 평가는 ‘학습을 위한 평가(assessment for learning)’이어야 하며, 이는 기본적으로 학생들의 학습을 위해 고안되고 실행되어야 한다. 여기서 학습을 위한 평가는 과거의 전통적인 형성평가에 비해서 실제 교육 현장에서의 활용 방법과 철학에 대하여 보다 더 깊이 있게 하였다는 점에서 기존의 것과 차별화될 수 있다(김명숙, 2008). 이는 곧 학습을 위한 평가에서의 시험은 교수·학습과 괴리되어 일련의 수업이 실시된 후에 추가되는 것이 아니라 교수·학습을 위한 것이며 교수·학습과 통합되어 실시하는 것이어야 함을 의미한다고 하겠다.

또한, 학습을 위한 평가에서 피드백은 학습자가 현재 자기 자신이 도달(성취)하고 있는 성취 수준과 미래에 자신이 성취하고자 하는 학

습 수준 내지 목표와의 차이를 알아, 그 간격을 좁히기 위해 필요한 행동을 안내한다(Ramaprasad, 1983; Sadler, 1989). 이때 제안되는 피드백의 형태는 단순히 그 해결 방법이나 답을 알려주는 것에서 벗어나 학습자가 가지고 있는 오개념이나 오류에 대한 적절하고 구체적인 처치 방안을 제시하여야 하며, 뿐만 아니라 제시한 처치 방안에 따라 학습자가 제대로 이행하는지를 집중적으로 관리, 감독하여야 할 것이다. 이러한 점에서, 학생들의 형성평가 결과를 반영하여 수업 운영 방법을 조정하고 피드백을 주도록 하는 PowerSource© 프로그램은 학생들의 오류 유형이나 오개념을 파악하여 보충 학습은 물론 완전 학습으로의 길로 나아가게 함으로써 학습 부진을 최소화 하는데 기여할 수 있을 것으로 여겨진다.

한편, 우리나라 국가 수준 학업성취도 평가에서 수학의 경우 가장 큰 비율을 차지하는 학생들의 성취 수준은 학교급이 올라감에 따라 점차 낮아지는 것으로 나타났다(고정화 외, 2007). 이러한 현상은 ‘계통성’이라는 기초적인 내용을 기반으로 하여 새로운 내용을 첨가함으로써 일관성 있게 순차적으로 이루어지는 수학의 특성에서 기인하는 것으로 해석되고 있다. 즉, 학생들이 수학 학습을 할 때 선행 지식을 충분히 숙지하고 있어야 하는데 학년이 올라갈수록 선행 지식의 양은 점차 많아져 학습에 대한 부담이 점점 커지게 되므로 이전 단계의 학습 결손은 다음 단계의 학습 결손을 초래하는 중요한 원인이 되어 결국 학습 부진으로 이르게 될 수 있다. 따라서 계통성을 갖는 수학은 형성평가를 실시하여 적절하게 피드백을 받을 수 있으면 학생의 입장에서는 보다 충실한 학습을 할 수 있고 나아가 수학의 학업 성취 수준을 높이는데 매우 중요한 역할을 할 것으로 판단된다. 이러한 학습 방법이 바로 CRESST에

서 개발한 PowerSource©에서 추구하는 것이며, 이러한 수업 방법에 따라 수업 자료들을 개발하여 재조직하는 노력들이 지속된다면 학생들의 수학 성취 수준은 점차 높아질 것으로 사료된다.

본 연구의 PowerSource© 형성평가 결과에서도 현재 우리나라 학생들이 특정의 수학적 개념이나 원리에 대한 정확한 이해와 적용 능력이 다소 미흡함을 알 수 있었다. 학생들이 수학 공식을 암기하여 문제를 해결하는 방식으로 수학을 학습하고 있으나 이러한 학습 방법으로는 의미 있는 수학적 능력이 신장되기는 어려울 것이다. 학생들의 수학 성취 수준과 수학적 소양이 신장되기 위해서는 수학적 개념에 대한 정확한 이해와 그 개념이 적용되는 다양한 맥락에서 활용해 보는 경험을 충분히 쌓는 것이 필요하다. 이를 위해 수학 수업의 개선뿐만 아니라 수학 학습 평가에서도 수학적 개념이나 원리와 관련된 이해 및 적용 능력의 평가를 위한 형태로 문제해결 과정을 설명하는 문항이 보다 적극적으로 활용되어야 할 것이다. 이러한 관점에서 볼 때, 답에 대한 근거나 풀이 과정을 제시하는 문항으로 구성된 PowerSource© 형성평가는 학생들로 하여금 자신의 수학적 개념에 대한 이해를 보다 확고히 할 수 있도록 하는데 도움이 된다고 하겠다.

또한, 수학적 사고를 글이나 수식으로 표현하는 것은 수학적 의사소통 능력과 관계되며 2007 개정 교육과정에서도 강조하고 있는 요소이다. 그런데, 2006년의 PISA 연구에서도 이미 나타난 바와 같이(이미경 외, 2007), 본 연구의 PowerSource© 평가 결과에서도 우리나라 학생들의 수학적 의사소통 능력이 여전히 미흡한 것으로 나타났다. 따라서, 수학적 의사소통 활동이 충분히 반영된 교수 학습 자료가 개발되어 보다 적극적으로 수학적 의사소통이 활발하

게 일어날 수 있는 수업 분위기가 조성되어야 함은 중요한 수학교육 과제일 것이다.

끝으로, TIMSS에서와 같이 있는 그대로의 교실을 두고 교실 학습이나 교육 상황에 대하여 국가 간 비교 연구를 수행한 적은 있지만, 특정 목적으로 개발된 교수·학습 프로그램을 공동으로 투입하고 그 성과를 연구한 사례는 드물다. 이번 PowerSource© 형성평가 프로그램 효과성 평가 연구의 경우, 미국과 한국이라는 국가 단위에서 공통의 학습 내용에 대한 학습 성과를 개선하기 위해 프로그램 개발에서부터 연구에 참여할 교사 연수, 투입 후 성과 측정에 이르기까지 일련의 프로그램 실행과 평가를 시도하였다는 것 자체로 의미를 찾아볼 수 있을 것이다. 한 마디로, 본 연구는 CRESST의 PowerSource© 프로그램의 투입과 결과를 통해 형성평가의 효과성을 재고할 수 있는 좋은 계기가 되었음은 물론, 한국과 미국의 국간 간 공동 연구로서의 의미를 담고 있다. 그러나 이렇게 공동 프로그램을 투입하는 경우 어려움이 적지 않게 발생하게 되는데, 이는 표준화된 평가도구를 활용한 국제 비교 연구의 문제점은 왜 그러한 결과가 나타났는지에 대한 진단이 쉽지 않다는 점이다. PowerSource©와 같은 공동 프로그램을 투입할 때 교육 상황의 차이, 투입의 세부 과정에서의 변이, 행위자나 참여자 측면의 차이 등으로 인해 명확한 비교가 어렵다. 즉, 한국의 교사와 미국의 교사들이 생각하는 수학 교과에 대한 인식이 다르고, 한국과 미국 학생들이 수학을 배우는 목적도 각각 다른 상황에서 프로그램의 효과를 정확히 검증하기란 어렵다.

일련의 프로그램을 투입하거나 참여자들의 반응을 설문 조사하기 이전에 비교 당사자들의 인식론적 특징과 문화적 맥락에 대한 이해가 병행되어야 제대로 된 결과 해석과 성과 분석

이 가능할 것이다. 예컨대 ‘유리수 동치’를 주제로 한 동일한 PowerSource© 프로그램이 한국적 상황 속에서 투입 학년이 어떻게 조정되는지, 같은 프로그램을 가지고 한국의 교실 맥락에서 어떻게 활용 방법이 달라지는지, 프로그램에 대한 학생과 교사의 반응은 어떻게 달라지는지, 연구에 참여하는 관계 당사자들의 관점은 어떻게 달라지는지 등을 들여다볼 필요가 있다. 한마디로, PowerSource©와 같은 프로그램을 투입하고 그 성과를 비교하는 연구를 수행하더라도 획일적인 처치와 연구 방법을 고집하기보다는 동일한 프로그램이 각국이 처한 고유의 교육상황 속에서 어떤 변이와 조정 과정을 거치는지를 문화적 맥락에서 보다 신중히 탐구할 필요가 있다.

## 참고문헌

고정화 · 도종훈 · 조지민 · 김명화 · 최인봉 · 송미영 · 김수진(2007). **2006년 국가수준 학업성취도 평가 연구 -수학-**. 한국교육과정평가원 연구보고 RRE-2007-3-4.

김명숙(2008). 교수 · 학습 혁신 세미나: 교수 · 학습과 평가의 연계. 한국교육과정평가원 연구보고 ORM 2008-18.

이미경 · 손원숙 · 노연경(2007). **PISA 2006 결과 분석 연구 - 과학적 소양, 읽기 소양, 수학적 소양 수준 및 배경 변인 분석**. 한국교육과정평가원 연구보고 RRE 2007-1.

전효선 · 오상철 · 박진용 · 권순달 · 정미경 · 김이성(2007). **국내외 교실 학습 연구(I) - 한국, 영국, 프랑스, 일본의 초등학교를 중심으로**. 한국교육과정평가원 연구보고 RRI-2007-1.

최승현 · 광영순 · 김주훈(2008). **2008년 CRESST 형성평가 프로그램 (PowerSource©)과 한국교**

**육과정평가원 공동 연구**. 한국교육과정평가원 연구보고 RRI-2008-6.

Baker, E. L.(1997). Model based performance assessment. *Theory into Practice*, 36(4), 247-254.

Baker, E. L., Freeman, M., & Clayton, S. (1991). Cognitive assessment of history for largescale testing. In M. C. Wittrock & E. L. Baker (Eds.), *Testing and cognition* (pp. 131-153). Englewood Cliffs, NJ: Prentice-Hall.

Black, P. J., & Wiliam, D. (1998a). *Assessment and classroom learning. Assessment in Education: Principles, Policy, and Practice*, 5(1), 7-74.

Black, P. J., & Wiliam, D. (1998b). *Inside the black box: Raising standards through classroom assessment*. London: School of Education, King's College. (See also article with the same title, 1998, in Phi Delta Kappan, 80(2), 139-148.)

Brown, R. S. & Niemi, D. N. (2007). *Investigating alignment of high school and community college assessments in California*. San Jose, CA: National Center for Public Policy in Higher Education

Chi, M. T. H., Bassok, M., Lewis, M., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science*, 13, 145-182.

NAEP(2007) [Http://www.ed.gov/print/nclb/overview/intro](http://www.ed.gov/print/nclb/overview/intro)(검색일: 2009.5.1)

Niemi, D. N. (1996). *Instructional influences on content area explanations and representational knowledge: Evidence for the construct validity of measures of principled understanding*. (CRESST Tech. Rep. No. 403). Los Angeles: University of California, National Center for

- Research on Evaluation, Standards and Student Testing (CRESST).
- Niemi, D. N., Baker, E. L., & Sylvester, R. (2007). Scaling up, scaling down: seven years of performance assessment development in the nation's second largest school district. *Educational Assessment, 12*, 195 - 214.
- Ramaprasad, A. (1983). On the definition of feedback. *Behavioral Science, 28*(1), 4-13.
- Renkl, A., Atkinson, R. K., & Maier, U. H. (2000). From studying examples to solving problems: Fading worked-out solution steps helps learning. In L. Gleitman & A. K. Joshi (Eds.), *Proceeding of the 22nd Annual Conference of the Cognitive Science Society* (pp. 393-398). Mahwah, NJ: Lawrence Erlbaum Associates.
- Sadler, R. (1989). Formative assessment and the design of instructional systems. *Instructional Science, 18*, 1 - 25.
- Sweller, J., van Merriënboer, J., & Paas, F. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*, 251 - 296.
- VanLehn, K. (1996). Cognitive skill acquisition. In J. Spence, J. Darly & D. J. Foss (Eds.), *Annual review of psychology* (Vol. 42, pp. 513 - 539). Palo Alto, CA: Annual Reviews.

# Understanding and Effectiveness of Formative Assessment Program in CRESST Focused on the Algebra Domain in the 8th Grade

Choe, Seung Hyun (Korea Institute of Curriculum and Evaluation)

Hwang, Hye Jeang (Chosun University)

Ryu, Hyun Ah (Korea Institute of Curriculum and Evaluation)

CRESST(the National Center for Research on Evaluation, Standards, and Student Testing at UCLA) is now carrying out the research, which was scheduled for a five year period from 2007 to 2011. This research aimed at testing the effectiveness of the formative assessment program by continuously conducting the program on the target group and steadily applying the recurring feedback, in order to reform the teachers' teaching and to facilitate students' learning. To do this, CRESST has set out to develop the material for 7th graders since

January 2007, and KICE(Korea Institute of Curriculum and Evaluation) have been running a collaborated research since July 2007, while sharing the instructional materials developed by CRESST.

In 2008, the pre-test was conducted prior to this study in 2009. Especially, this paper deals with the Korean 8th graders' scholastic achievements in algebra domain measured by PowerSource©. In addition, this study would examine the responses of teachers and students on its application.

\* key words : CRESST, Formative Assessment(형성평가), Big ideas(핵심개념), Algebra domain (대수영역)

논문접수 : 2010. 4. 23

논문수정 : 2010. 6. 1

논문완료 : 2010. 6. 11

<부록 1> 각 문항의 사전사후 응답 변화

1. 다음 중  $3x$  와 같은 것을 고르면?

- ①  $x+x+x$       ②  $x^3$       ③  $x \times x \times x$       ④ 정답 없음

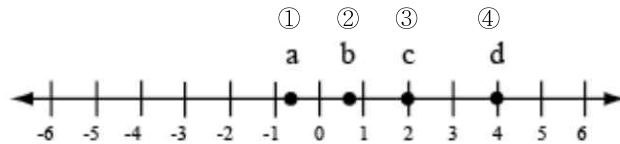
	①	②	③	④	무응답
사전평가	78.2	3.9	10.7	6.5	0.6
사후평가	84.3	1.5	11.3	2.9	0.0

2. 다음 중  $a \times (2 \times 3)$  과 같은 것은?

- ①  $2a+3$       ②  $a+2+3$       ③  $(a \times 2) \times 3$       ④  $a \times 2 \times a \times 3$

	①	②	③	④	무응답
사전평가	1.9	2.6	43.2	51.3	1.0
사후평가	6.9	2.0	52.0	38.2	1.0

9. 다음 수직선에서  $\frac{2}{3}$  와 근사한 위치를 나타내는 점에 표시하여라.



	①	②	③	④	무응답
사전평가	1.9	74.7	16.2	3.6	3.6
사후평가	3.4	77.5	14.5	2.9	2.0

10. 다음 방정식이 참이 되게 하는  $p$  의 값은?

$$\left(\frac{1}{p}\right)(4) = 1$$

- ①  $p = 4$       ②  $p = \frac{1}{4}$       ③  $p = 16$       ④  $p = -4$

	①	②	③	④	무응답
사전평가	72.4	18.2	1.9	2.3	5.2
사후평가	77.0	16.7	2.5	2.5	1.5

11. 분수  $\frac{52}{78}$  를 최대한 간단히 하면?

- ①  $\frac{26}{39}$       ②  $\frac{2}{3}$       ③  $\frac{4}{6}$       ④ 분수  $\frac{52}{78}$  는 더 이상 간단히 할 수 없다.

	①	②	③	④	무응답
사전평가	52.6	40.3	1.9	2.9	2.3
사후평가	46.1	45.6	3.9	2.9	1.5

18.  $3+2(x-2)$  와 같은 것을 고르면?

- ①  $5(x-2)$       ②  $5(x+4)$       ③  $3+2x-4$       ④  $3+2+x-2$

	①	②	③	④	무응답
사전평가	15.6	4.2	72.7	3.2	4.2
사후평가	8.8	2.5	80.9	3.9	3.9

21.  $6x+0=4x-3$  과 같은 것은?

- ①  $6x=4x-3$       ②  $6x+0=4x$       ③  $6x+4x=1x$       ④  $6x-3=0$

	①	②	③	④	무응답
사전평가	78.2	3.2	4.2	3.2	11.0
사후평가	84.8	5.4	2.1	2.9	4.9

24. 한 학생이 방정식  $3x+6=-9$  를 풀려고 하고 있다. 다음 중 첫 번째 단계로 옳은 것은?

- ①  $\frac{3x}{3}+6=\frac{-9}{3}$       ②  $3x+6+(-9)=(-9)-(-9)$       ③  $\frac{3x}{3}+\frac{6}{3}=\frac{-9}{3}$       ④  $9x=-9$

	①	②	③	④	무응답
사전평가	3.9	15.3	57.8	6.5	16.6
사후평가	5.9	13.7	68.6	4.4	7.4

25. 재경의 나이는 선우의 나이의 두 배보다 6살이 적다. 선우가 6살일 때, 재경은 몇 살인가?

- ① 0      ② 6      ③ 12      ④ 18

	①	②	③	④	무응답
사전평가	2.3	73.4	7.8	4.9	11.7
사후평가	3.4	78.9	9.3	2.9	5.4



<부록 2> 평가 간 동형 문항의 예 (유리수 동치)

평가1	평가2와 평가3
1. $\frac{3}{7} + \frac{1}{7} =$	[평가2] 1. $\frac{2}{9} + \frac{5}{9} =$ [평가3] 1. $\frac{5}{11} + \frac{3}{11} =$
2. $\frac{2}{3} + \frac{3}{5} =$	[평가2] 2. $\frac{1}{3} + \frac{2}{7} =$
3. $\frac{3}{7} + \frac{2}{m} =$ ( $m \neq 0$ )	[평가2] 3. $\frac{5}{7} + \frac{2}{p} =$ ( $p \neq 0$ ) [평가3] 2. $\frac{2}{n} + \frac{4}{9} =$ ( $n \neq 0$ )
4. 다음 분수를 더하기 위한 첫 번째 풀이 과정을 쓰시오. $\frac{1}{7} + \frac{2}{5}$	[평가2] 4. 다음 분수를 더하기 위한 첫 번째 풀이 과정을 쓰시오. $\frac{3}{5} + \frac{1}{8} =$
5. $\frac{3}{7} + \frac{2}{7}$ 을 계산할 때, 왜 답이 $\frac{5}{14}$ 가 아닌지 설명하시오.	[평가3] 3. $\frac{1}{9} + \frac{5}{9}$ 를 계산할 때, 왜 답이 $\frac{5}{18}$ 가 아닌지 설명하시오.
6. 한 학생이 분수의 덧셈 $\frac{4}{5} + \frac{3}{7}$ 을 하고 있다. 풀이의 첫 번째 단계는 다음과 같다. $\frac{4}{5} + \frac{3}{7} = \frac{4 \times 7}{5 \times 7} + \frac{3 \times 5}{7 \times 5}$ a) 과정이 옳은가? b) 이유를 설명하시오.	[평가2] 5. 한 학생이 분수의 덧셈 $\frac{3}{4} + \frac{2}{5}$ 를 하고 있다. 풀이의 첫 번째 단계는 다음과 같다. $\frac{3}{4} + \frac{2}{5} = \frac{3 \times 5}{4 \times 5} + \frac{2 \times 4}{5 \times 4}$ a) 과정이 옳은가? b) 이유를 설명하시오.
7. 한 학생이 $a \neq 0, b \neq 0$ 일 때, 분수의 덧셈 $\frac{2}{a} + \frac{5}{b}$ 를 하고 있다. 풀이의 첫 번째 단계는 다음과 같다. $\frac{2}{a} + \frac{5}{b} = \frac{2 \times b}{a \times b} + \frac{5 \times a}{b \times a}$ a) 과정이 옳은가? b) 이유를 설명하시오.	[평가3] 4. 한 학생이 $a \neq 0, b \neq 0$ 일 때, 분수의 덧셈 $\frac{4}{7a} + \frac{3}{5b}$ 을 풀고 있다. 풀이 과정의 첫 단계는 다음과 같다. $\frac{4}{7a} + \frac{3}{5b} = \frac{4 \times 5b}{7a \times 5b} + \frac{3 \times 7a}{5b \times 7a}$ a) 과정이 옳은가? b) 이유를 설명하시오.
8. $\frac{1}{4}$ 마일 걷고 난 후 다시 $\frac{2}{5}$ 마일을 걸었다. a) 총 거리는 얼마인가? 풀이 과정을 모두 쓰시오. b) 풀이 과정의 각 단계를 설명하시오. 사용한 수학적 성질을 쓰고, 그 중요성을 설명하시오.	[평가3] 5. 케익을 만들면서 흑설탕 $\frac{2}{3}$ 컵과 백설탕 $\frac{1}{4}$ 컵을 사용하였다. a) 사용한 설탕의 전체 양은? b) 풀이 과정의 각 단계를 설명하시오. 사용한 수학적 성질을 쓰고, 그 중요성을 설명하시오.