

## 현직 수학교사 문항 개발 연수의 평가도구 개발 전문성 향상 효과

최 지 선\*

교사의 평가도구 개발 전문성은 자동적으로 개발되지 않으며 실습을 통해서 개발될 수 있다. 본 연구의 목적은 평가 문항 개발 연수가 평가도구 개발 전문성 향상에 효과에 있는지를 분석하기 위한 것이다. 연수에 참여한 수학 교사들이 문항 개발에 대한 실습을 할 수 있도록 연수 과정을 구성하였고, 연수 전과 후로 평가도구 개발 전문성 평가지, 문항 검토 의견서를 작성하게 하고, 연수 후에 일부 교사들을 대상으로 연수를 평가하도록 하였다. 이 연구의 결과는 첫째, 교사들이 스스로 판단하기에 문항 개발 연수는 평가도구 개발 전문성 향상에 효과가 있었다. 특히 표준화 검사에 대한 이해와 문항의 난이도와 변별도에 대한 이해가 증가하였다. 둘째, 구체적인 근거를 제시하여 검토 의견을 제시하는 비율이 증가하였다. 셋째, 짧은 연수시간은 연수의 효과를 저해하는 요인이었고, 팀내 토론과 팀간 토론 과정, 강사의 피드백, 적극적인 참여와 전문성 부족의 자각은 효과적인 요인이었다.

### 1. 서론

#### 1. 교사의 평가도구 개발 전문성

교사는 학생의 학업성취 정도를 파악하기 위하여 여러 가지 정보를 수집하고 해석하여 활용하는 학생평가를 하게 된다(McMillan, 2004). 교사가 학생평가를 하기 위하여 필요한 전문지식을 개념적으로 구분하면, 평가방법의 선정, 평가도구의 개발, 평가 실시-채점-성적부여, 평가 결과의 분석-해석-활용-의사소통, 평가의 윤리성 인식으로 구분할 수 있다(김수동, 김선희, 2005). 교사는 교수-학습이 이루어지는 과정에서 여러 가지 형태의 평가 목적을 가지고 평가 도구를 결정하게 되는데, 적절한 평가 도구를 결정하지 못

한다면 평가의 목적을 달성할 수 없다. 평가도구를 선택할 때에는 기존에 개발된 검사 도구를 선택하여 사용할 수도 있지만, 평가 목적에 부합하는 적절한 평가도구를 찾지 못했을 때에는 새로운 평가도구를 개발하게 된다. 교사가 의도하는 평가 목적에 맞는 적절한 평가도구를 찾는 것이 항상 가능한 것은 아니므로 교사들은 상황과 목적에 맞는 평가도구를 개발할 수 있어야 한다. 또는 기개발된 평가도구의 질을 스스로 점검하고 부분적으로 혹은 전반적으로 수정할 수 있어야 한다. 이하에서는 김수동, 김선희(2005)의 용어 정의를 사용하여 학생평가 전문성의 일부로 ‘평가의 목적과 내용에 적합한 평가도구를 개발하거나 선택하는 능력과 평가도구의 질을 점검하고 개선하는 능력’을 ‘평가도구 개발 전문성’이라고 하겠다.

\* 한국교육과정평가원, jschoi@kice.re.kr

평가도구 개발 전문성은 평가에 대한 이론적인 지식과 교사의 경험을 통해서 상호보완적으로 향상될 수 있다. 그런데 교사들의 평가도구 개발 전문성은 이론적인 지식과 경험이 조화되지 못하고 분리되어 발달될 가능성이 있다. 예비교사 시절에 이론적인 지식을 습득하고, 학교 현장에서 경험을 통해서 평가 문항을 개발하는 경우가 많기 때문이다. 문항 자체에 결함이 있는 문항을 개발하는 경우(이석현, 한길준, 2012), 주어진 조건을 만족하는 값이 없는 문항, 조건 부족으로 성립되지 않는 문항 등을 개발하는 경우가 있다(신범영, 2010). 또한 이와 같은 어려움이 있기 때문에 교사들은 문항을 제작하기 보다는 기존의 문항들을 변형하여 사용하는 경향이 있다(김선희, 2006).

한편 교사들은 스스로 평가도구 개발 전문성이 향상되어야 할 필요를 느끼고 있다. 장민아(2009)는 서울 소재 중등학교 교사 146명을 대상으로 설문 조사한 결과, 평가 관련 연수 과정에서 중요하다고 생각되는 평가 관련 지식이나 능력 중에서 ‘평가도구의 개발 방법에 관한 지식’을 선택한 교사가 37.7%로 가장 높았다. 또한 강현영 외(2011)는 현직 교사들을 대상으로 교사에게 필요한 역량을 조사 분석한 결과, 현직 교사들이 과제나 평가를 위해 문항을 개발하고 구성하는 능력을 중요하게 판단하고 있음을 밝혔다.

## 2. 평가 문항 개발 연수

수학과에서 평가도구는 주로 문항으로 사용된다. 따라서 평가 문항 개발 연수는 평가도구 개발 전문성 향상을 목적으로 한다고 할 수 있다. 평가도구 개발 전문성은 이론적 지식이 실천적 지식으로 전환되는 과정을 통해서 향상되며, 이러한 전환은 자동적으로 이루어지지 않고 실습이 요구되는 과정이다(최승현, 강대현, 박경순,

장경숙, 2008). 따라서 교사들은 문항 개발이 실천적 지식이 될 수 있도록 실습을 해야 한다. 하지만 일반적으로 실습 위주의 교사 연수 과정이 많지 않은 것처럼, 실습 위주의 평가 문항 개발 연수도 많지 않은 실정이다(박경미 외, 2010; 김선희, 2012a). 이러한 측면에서 특히 평가 문항 개발 연수는 실습 위주로 이루어져야 한다는 주장이 제기되어 왔다(장민아, 2009; 강현영 외, 2011; 김선희, 2012a).

최근 실습 위주의 평가 문항 개발 연수의 효과에 대한 연구 결과가 확인되고 있다. 김선희(2012)는 예비교사 대상 평가 전문성 향상을 위한 프로그램을 구성하고 그 효과를 일부 확인하였다. 이 연구는 실습 위주의 연수 프로그램의 효과를 확인하였는데 의의가 있지만, 예비교사들을 대상으로 했다는 점 그리고 대학교 과목으로써 15주 동안 이루어졌다는 점에서 현직 교사를 위한 연수로 일반화하기 한계가 있다. 최지선(2013)은 현직 고등학교 수학교사들을 대상으로 3박4일(30시간) 간 합숙형태로 이루어진 문항 개발 연수의 효과를 확인하였다. 하지만 이 연구는 연수 구조가 합숙이라는 특수한 형태이므로, 일반화하기에는 한계가 있다.

교사 연수는 15주 동안, 혹은 합숙형태로 이루어지기 어렵고, 방과 후나 방학 중으로 단기적으로 이루어지는 경우가 많다. 특히 교사들이 일반적으로 선택할 수 있는 평가 문항 개발 연수는 6시간 이내로 가정할 수 있다. 예를 들어, 박경미 외(2010)에 따르면 4개의 교육연수원의 1급 정교사 자격 연수 과정은 평가도구 개발과 관련된 시간을 4시간으로 배정하였고, 이환철, 장미숙(2013)이 제시한 모범적인 연수 사례에서도 문항 개발과 관련된 연수 시간은 5시간이었다.

학교 현장에서 교사들이 여러 가지 형태의 평가 문항을 개발하고 있다는 점과 자격 연수와 직무 연수 과정에서 평가 문항 개발 연수가 계

속 개설되고 있다는 점을 고려할 때, 현직 교사 대상 평가 문항 개발 연수에 대한 연구와 분석이 이루어질 필요가 있다.

### 3. 연구 목적 및 연구 질문

이 연구는 중고등학교 수학교사들이 쉽게 선택할 수 있는 6차시 평가 문항 개발 연수를 실습 위주로 구성한 후에, 수학과 평가 문항 개발 연수가 교사의 평가도구 개발 전문성 향상에 효과가 있는지를 분석하기 위한 목적으로 이루어졌다. 이와 같은 연구 목적을 달성하기 위하여 다음과 같이 연구 질문을 설정하였다.

첫째, 연수를 통해서 교사들이 스스로 판단하는 평가도구 개발 전문성은 어떻게 변하는가? 교사의 평가도구 개발 전문성을 검사하기 위한 기개발된 검사지를 활용하여, 연수 전에 교사들의 평가도구 개발 전문성 정도를 파악한다. 그리고 연수 후에 교사들의 평가도구 개발 전문성을 다시 측정하여, 연수 전과의 변화를 비교한다.

둘째, 연수를 통해서 교사들이 제시하는 문항 검토 의견은 어떻게 변하는가? 문항 검토 의견의 수준을 3개의 수준으로 구분한 후에, 교사들의 의견 수준이 어떻게 변화되는지를 분석한다.

셋째, 어떤 연수의 요인이 평가도구 개발 전문성 향상에 도움이 되었는가? 연수 프로그램의 요인별로 연수 효과에 주는 영향을 평가지와 면담을 통해서 분석한다.

수는 P교육청이 주관하는 30차시 수학교사 직무 연수의 일부로, 교사들은 60차시 중에서 자신의 필요에 맞게 30차시를 구성하여 연수를 이수할 수 있었다. 그 중 “문항 개발” 연수는 1일간 6차시로 기획되어 있었다. 따라서 “문항 개발” 연수를 신청한 교사들은 자발적으로 참여한 것으로 볼 수 있다. 총 25명이 참여하였으나, 6차시 모두 참여한 교사는 21명으로, 이하에서는 21명에 대해서만 분석하였다. 연수 참여 교사의 배경 정보는 <표 II-1>과 같다.

성별로 분류하면 남자 10명, 여자 11명이었다. 이 중 평가 관련 연수에 참여했던 경험이 있는 교사는 8명이었다. 교육경력은 짧게는 1년에서부터 길게는 34년에 이르기까지 다양하였고, 평균 교육경력은 13.8년이었다.

<표 II-1> 연수 참여 교사의 배경 정보

구분	인원		
	수(명)	비율(%)	
성별	남자	10	47.6
	여자	11	52.4
교사 경력 (만)	0-5년	5	23.8
	6-10년	5	23.8
	10년-15년	3	14.3
	15-20년	1	4.8
	20-25년	6	28.6
	25-30년	0	0.0
	30-35년	1	4.8
담당 학교	중학교	8	38.1
	고등학교	13	61.9
평가 관련 연수 참여 경험	0회	13	61.9
	1회	6	28.6
	2회	2	9.5

## II. 연구 방법

### 1. 연구 참여자

연구 참여자는 평가 문항 개발 연수에 자발적으로 참여한 중등학교 수학교사들이다. 해당 연

### 2. 문항 개발 연수

문항 개발 연수는 다음과 같은 순서로 구성되었다. 첫째, 가용 가능한 시간인 6시간을 기준으로 연수 프로그램을 설계하였다. 해당 연수의 강사는 본 연구의 연구자이었다. 연구자는 연수에서 이루어지는 6차시를 교사들이 일반적으로 이

수할 수 있는 연수 사례로 가정하였다. 교사 연수의 경우에 특히 교사의 전문성 중의 일부만을 다루는 연수의 경우에는 짧은 시간으로 이루어지는 경우가 많기 때문이다.

둘째, 문항 개발 연수는 실습 위주로 이루어져야 하므로 PBL(Problem based learning) 단계를 따르기로 하였다. PBL과 문항 개발 연수는 문제, 학생, 교사의 측면에서 일관되기 때문이다(최지선, 2013). 교사들이 자발적으로 논의를 이끌어갈 수 있기 위해서는 교사와 학생이 동시에 사용할 수 있는 안내지침으로, ‘문항 개발 원칙’을 활용하였다. ‘문항 개발 원칙’에 평가도구 개발의 일련의 절차와 목적 그리고 구체적인 개발 방법 등을 포함시킴으로써 교사들이 자발적으로 논의를 이끌어갈 수 있다. 연수 참여 교사들은 ‘문항 개발 원칙’을 통해서 문항 개발과 관련된 이론에 기초하여 문항을 개발, 검토, 수정하게 된다.

본 연구에서는 ‘문항 개발 원칙’을 전영주(2012)가 제시한 평가문항 제작 과정에서 유의할 점, 고정화(2010)가 제시한 문항 수정 유형 그리고 김선희(2012b)가 제시한 평가 문항 검토 틀을 참조하여 구성하였다. 전영주(2012)는 문항 제작시의 유의점으로 교육과정을 벗어나는 문항, 수학적으로 오류가 있는 문항, 출제 의도가 명확하지 않은 문항, 반복되는 계산이나 많은 계산을 요구하는 문항, 불분명한 용어를 사용하는 문항 등이 있음을 사례별로 안내하였다. 고정화(2010)는 평가 문항 수정 유형을 성취기준, 행동영역, 문제진술, 문제구성요소, 문항 내용, 문항 형식, 오류, 답지, 문항 수로 구분하였다. 김선희(2012b)는 교육과정 충실도와 문항 완성도로 구분하였다. 이와 같은 사항들을 고려하여, 문항 개발 원칙을 교육과정 수준, 문항 내적 수준, 문항 외적 수준으로 선형적으로 구분한 후에, 하위 항목을 재배열하여, <표 II-2>와 같이 정리하였다. 교육과정 수준에는 교육과정의 내용과 범위

에 부합하는지 그리고 평가 목적에 부합하는지의 여부를 포함하였다. 문항 내적 요소 수준에는 문항의 완성도를 결정짓는 요소들을 포함하였다. 문항 오류, 문항의 타당도, 변별도, 난이도, 용어의 적절성, 정확성, 계산량 등이 해당한다. 문항 외적 요소에는 진술의 일관성, 선택지의 배열, 배점, 편집 등을 포함시켰다.

<표 II-2> 문항 개발 원칙 수준별 하위요소

수준	하위 요소
교육과정 수준	교육과정 및 교과내용의 범위, 수준
	평가 목적 부합 여부
문항 내적 수준	문항 오류
	문항의 타당도, 변별도, 난이도
	용어 수준
	정확성, 단서
문항 외적 수준	문두(발문) 표현
	선택지
	배점
	편집 체제

셋째, 연수 과정은 최지선(2013)의 4단계를 일부 변형하여 구성하였다. ‘문제이해’ 단계에서는 미리 제작된 평가 문항 초안을 검토하는 과정을 통해서, 평가 문항 개발과 검토의 어려움을 느끼도록 하였다. 교사들이 직접 문항 초안을 구안하는 것이 본질적이고 효과적인 방법이지만, 사전에 문항 개발을 과제로 제시할 수가 없었고 연수 시간이 충분하지 못하여 대안적인 방법으로 사전에 미리 제작된 문항 초안을 활용하였다. 교사들은 이 단계에서 문항 초안을 개별적으로 검토한다. ‘미니강의’ 단계에서 평가 이론에 대한 전체 강의를 듣고 수학과 문항 개발에서 필요한 이론을 제시하였다. 문항 개발과 검토의 기준으로 ‘문항 개발 원칙’을 안내하였다. 그리고 평가 목적, 평가도구 개발과 검토, 그리고 결과의 해

<표 II-3> 평가 문항 개발 연수 과정

연수 과정	차시 및 시간	학습 활동	수업 방법
사전 평가	10분	• 평가도구 개발 전문성 평가지 자가 평가	
문제이해	1차시 (20분)	• 개발된 문항(2문항)에 대한 개발 검토 의견 제시	
미니강의	1-3차시 (130분)	• 문항 개발 관련 이론 소개 - 문항 개발 원칙 안내	강의
		• 문항 개발 사례 소개 - 문항 개발 원칙에 따라 사례(3문항) 논의	강의, 토론
평가 문항 개발	4-6차시 (150분)	• 문항 검토 및 수정하기 - 문항 개발 원칙에 의거하여 문항(4문항) 검토하기	팀내 토론 팀간 토론
사후 평가	10분	• 평가도구 전문성 평가지 자가 평가	

석에 대해서 이론을 적용하도록 하였다. ‘평가 문항 개발’ 단계에서는 앞서 개별 검토하였던 2개 문항을 포함하여 모두 4개의 문항에 대해서 3-4인으로 구성된 팀으로 토론을 통해서 공동으로 문항을 수정하였다. 평가 문항 제작은 계획 단계, 초안 작성 단계, 검토·수정 단계, 완성 단계로 이루어지는데(전영주, 2012), 본 연수는 검토·수정 단계와 완성 단계 중심으로 이루어졌다. 수정된 문항은 팀별로 차이가 있었고, 팀별로 논의 결과를 발표함으로써 다른 팀의 의견이 공유되도록 하였다. ‘반성’ 단계에서는 논의 사항들을 반영한 수정 문항을 만들어보고, 검토 사항들이 어떻게 반영되었는지를 팀별로 정리하였다.

넷째, 평가 문항 개발 연수의 효과를 높이기 위한 요인을 반영하였다. 최지선(2013)이 제시한 선택형 평가 문항 개발 연수 프로그램의 효과를 높이기 위한 4가지 요인을 반영하였다. 집중 훈련 가능한 ‘연수 프로그램 구조’를 위해서 소규모 집단으로 구성하여 팀 작업이 가능하도록 하였다. 학교급별로 분리하여, 1팀에 3-4명의 교사를 임의로 배정하였다. ‘문항 개발 습관 표면화와 수정’을 위하여 연수 초반에 문항 검토 의견을 제시하도록 하였다. 셋째, ‘교사간 공감대 형성’을 위하여 팀별로 문항 수정 사항을 논의하

도록 하고, 다른 팀의 발표와 토론을 통해서 팀별간으로 검토 사항이 공유되도록 하였다. 그리고 본인의 ‘전문성 자각’을 위해서 평가도구 개발 전문성 평가지를 활용하여 자가 진단하도록 하였다.

이와 같이 구성된 평가문항 개발 연수 과정의 차시 및 시간, 학습 활동, 수업 방법은 <표 II-3>과 같다.

### 3. 자료 수집과 분석

자료 수집은 2014년 1-2월에 이루어졌고, 자료는 연수 전후에 이루어진 평가도구 개발 전문성 평가지, 연수 초반에 작성한 문항 검토 의견서, 연수 과정에서 제작된 교사들의 노트, 연수 과정을 녹음한 음성파일과 녹취록, 연수 이후에 이루어진 연수 평가지와 개별 면담을 녹음한 음성파일과 녹취록, 연구자의 연구일지 등이다. 개별 면담은 연수의 효과를 요인별로 분석하기 위하여 이루어졌으며, 4명의 교사를 대상으로 하였다. 해당 교사들은 연수 과정에서 자신의 의사를 적극적으로 표현하는 교사들로, 편의 추출되었다. 연구 질문과 연구 질문을 위해 수집된 자료의 관계는 <표 II-4>와 같다.

<표 II-4> 연구 질문과 수집 자료 관계

연구질문	수집 자료
1	사전 평가도구 개발 전문성 평가지 사후 평가도구 개발 전문성 평가지
2	(개별) 문항 검토 의견서 팀내, 팀간 토론 자료(음성파일, 녹취록)
3	연수 평가지 면담 자료(음성파일, 녹취록)
기타	연수 과정에서 제작된 교사들의 노트 연구자의 연구일지

가. 평가도구 개발 전문성 평가지

교사들의 평가도구 개발 전문성을 분석하기 위하여, 송미영, 김경희(2007)의 ‘교사의 학생평가 전문성 진단 도구’를 활용하였다. 이 도구는 교사의 전문성 지식과 기술을 교사 스스로 진단할 수 있는 검사 도구로 타당도와 신뢰도가 확보된 것이다. 이 도구의 총 60문항 중 ‘평가도구의 개발’과 관련된 다음과 같은 13개 항목을 활용하여 교사의 평가도구 개발 전문성을 조사하였다. 교사들은 자신의 전문성을 판단하여 그 정도에 따라 ‘전혀 그렇지 않다(1점)’, ‘별로 그렇지 않다(2점)’, ‘보통이다(3점)’, ‘약간 그렇다(4점)’, ‘매우 그렇다(5점)’으로 응답하였다. 연수의 처음과 마지막에 실시하여 사전, 사후 점수를 비교하였다.

- (1) 평가의 목적과 내용에 적합한 평가도구를 선택할 수 있다.
- (2) 평가의 목적과 내용에 부합하는 평가도구가 없을 경우 평가도구를 직접 개발할 수 있다.
- (3) 검사제작의 일반적 절차에 따라 평가도구를 개발하기 위한 계획을 구체적으로 세울 수 있다.
- (4) 선택형 문항제작원리를 충실히 반영하여 평가도구를 개발할 수 있다.
- (5) 서답형 문항제작원리를 충실히 반영하여 평가도구를 개발할 수 있다.
- (6) 수행평가 과제 및 채점기준 개발 원리를 충실히

반영하여 평가도구를 개발할 수 있다.

- (7) 표준화 검사의 개념과 특징 및 종류를 알고, 목적에 맞게 선택하여 사용할 수 있다.
- (8) 타당도의 개념을 알고, 타당도를 평가할 수 있다.
- (9) 신뢰도의 개념을 알고, 신뢰도를 평가할 수 있다.
- (10) 채점자 신뢰도의 개념을 알고, 채점자 신뢰도를 평가할 수 있다.
- (11) 문항 난이도의 개념을 알고, 문항 난이도를 평가할 수 있다.
- (12) 문항 변별도의 개념을 알고, 문항 변별도를 평가할 수 있다.
- (13) 오답지 매력도의 개념을 알고, 오답지 매력도를 평가할 수 있다.

전문성의 정도는 송미영, 김경희(2007)이 제시한 바와 같이, 척도 평균의 기준선 범위를 정하여 다음과 같이 해석하였다.

- 4.5 이상: 평가 전문성이 높다.
- 3.5 이상 4.5 미만: 평가 전문성이 있다.
- 2.5 이상 3.5 미만: 평가 전문성이 낮다.
- 2.5 미만: 평가 전문성이 매우 낮다.

나. 문항 검토 의견서

연수 참여 교사들의 문항을 검토하고 수정 방향을 제시하는 정도를 파악하기 위하여, 연수 초반에 준비된 2개의 문항 초안을 교사들에게 제시하고 문항 검토 의견을 작성하도록 하였다. 교사들은 주어진 문항 검토 의견서를 개별적으로 작성하였고, 개인의 의견을 충분히 이끌어 낼 수 있도록 충분한 시간을 주었다. 15여분이 지나자 많은 교사들이 문항 검토 의견서 작성을 마쳤고 20분이 경과할 때에 개별 문항 검토 시간을 종료하였다. 검토 문항은 [그림 II-1]과 같다.

연수 중에는 4개의 문항 검토가 이루어졌으나, 문항 검토 의견 수준의 변화를 연수 전과 비교하기 위해서 2개의 문항에 대해서만 팀내, 팀간 검토 과정에서 제시된 의견을 분석하였다. 연수

과정을 녹음한 음성파일과 녹취록을 작성한 후에 교사들의 의견 정도를 코드화하여 분석하였다. 코드는 ‘문항 개발 원칙’의 분류에 따라 이루어졌다.

◆ 내용 영역 : 미적분과 통계 기본, 정적분의 활용  
 ◆ 행동 영역 : 이해  
 ◆ 출제 의도 : 함수의 주기성과 정적분의 성질을 이용하여 정적분을 할 수 있는지를 평가한다.

1. 실수 전체에서 정의된 함수  $f(x)$ 가 다음 조건을 만족시킨다.

$$(가) f(x) = \begin{cases} |x+1|-1 & (-2 \leq x < 0) \\ -|x-1|+1 & (0 \leq x \leq 2) \end{cases}$$

(나) 모든 실수  $x$ 에 대하여  $f(x) = f(x+4)$

이 때,  $\int_{-2014}^{-2016} \{-f(-x)\}dx$ 의 값은?

① -2      ② -1      ③ 0  
 ④ 1      ⑤ 2

---

◆ 내용 영역 : 수학 I, 여러 가지 수열  
 ◆ 행동 영역 : 발견적 추론  
 ◆ 출제 의도 : 주어진 도형에 나타난 수열의 일반항을 구할 수 있는지를 평가한다.

2.  $T_n$ 은 한 변의 길이가 1인 정사각형의 각 변을  $(n+1)$ 등분하여 대각선 방향으로 각 정사각형에 문양 을 넣은 것이다.

...

$T_1$ 
 $T_2$ 
 $T_3$

이때  $T_n$ 에 문양 의 개수를  $a_n$ 이라 하자. 예를 들어  $a_1 = 4, a_2 = 5, a_3 = 8$  이다.  $a_{20} + a_{31}$ 의 값은?

① 90      ② 95      ③ 100  
 ④ 105      ⑤ 110

[그림 II-1] 검토할 문항 초안

#### 다. 연수 평가지

연수가 효과적으로 이루어지는데 영향을 준 요인을 파악하기 위하여 연수 종료 후에 연수 과정에 참여한 4명의 교사를 대상으로 개별 면담을 하였다. 연구자는 사전에 연수를 기획하는 과정에서 고려하였던 연수의 효과 요인을 반영하여 <표 II-5>와 같이 5점 척도 연수 평가지를 구성하였다. 항목별로 연수의 효과 정도를 ‘전혀 그렇지 않다(1점)’, ‘별로 그렇지 않다(2점)’, ‘보통이다(3점)’, ‘약간 그렇다(4점)’, ‘매우 그렇다(5점)’으로 응답하는 방식이다. 개별 면담 과정에서 연수 참여 교사는 연수 평가지를 기록한 후에 개별 면담을 하였다. 면담 시간은 교사 1인당 평균 8분이 소요되었다.

<표 II-5> 연수 요인별 연수 평가지 항목

연수 요인	연수 평가지 항목
연수 프로그램 구조	1. 소그룹으로 문항을 수정하는 실습이 효과적이었다.
	2-1. 연수 시간이 충분하였다. 2-2. (현재 주어진 시간 기준으로) 더 많은 문항을 실습해야 한다.
문항 개발 습관 표면화와 수정	3. 연수 초반에 문항 수정 의견을 혼자서 고민해본 것이 연수를 진행하는데 도움이 되었다.
	4. 강사의 피드백이 도움이 되었다.
교사간 공감대 형성	5. 나는 평가 문항 개발 전문성을 향상시키고 싶었다.
	6. (혼자 문항을 수정하는 것과 비교하여) 소그룹 논의 속에서 문항 수정 방향을 탐색하게 되었다.
전문성 자각	7. 다른 그룹 발표를 통해 문항 수정에 다양한 생각을 할 수 있었다.
	8. 이번 연수를 통해 앞으로 평가 문항 개발과 관련된 전문성 향상을 시킬 필요성을 느꼈다.

### Ⅲ. 연구 결과

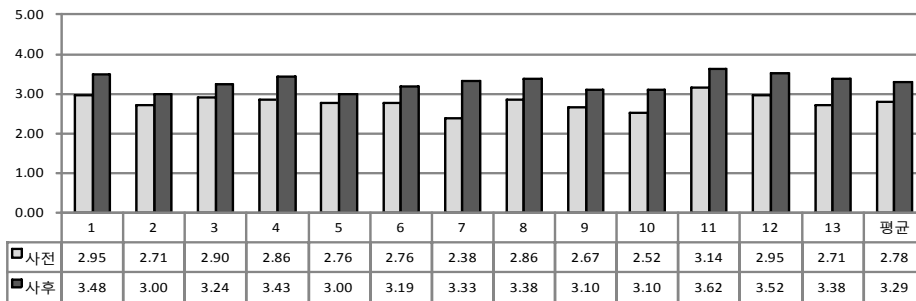
문항 개발 연수의 평가도구 개발 전문성 향상 효과를 파악하기 위하여, 연수 전후 평가도구 개발 전문성 변화, 문항 검토 의견의 변화, 그리고 연수 요인에 대한 교사들의 의견을 분석하였다.

#### 1. 평가도구 개발 전문성의 변화

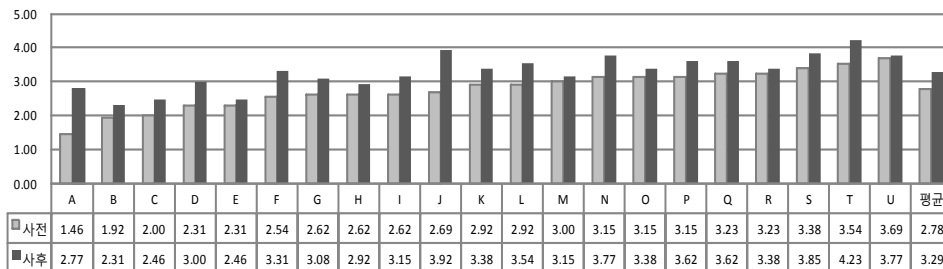
연수 초반에 교사들은 자신의 평가도구 개발 전문성을 보통 이하라고 인식하고 있었다. 전체 평균은 2.78, 분산은 0.20으로 나타났고 항목별로 최소 2.38에서 최대 3.38까지 차이가 약간 있었다. 7번을 제외한 모든 항목이 평가 전문성이 낮은 것으로 나타났다. 표준화 검사와 관련된 7번 항목은 매우 낮은 수준이었다. 참고로 교사별로

는 최소 1.46에서 최대 3.69까지로 큰 차이가 있었다.

연수 후에 교사들은 평가도구 개발 전문성을 스스로 점검하였다. 그 결과 평균 3.29 분산 0.20으로 나타났고 항목별로 살펴보면 [그림 Ⅲ-1]과 같다. 사후 결과는 사전 결과보다 평균 약 0.51 상승한 값으로, t-검정으로 유의미하게 상승하였다(유의수준 0.05). 전문성이 매우 낮다고 보는 항목은 없었고, 11, 12번 항목에서는 전문성이 있는 편으로 인식하였다. 즉, 교사들은 문항 난이도와 변별도 개념을 이용하여 문항을 검토하는 전문성이 발전하였다고 생각하였다. 이를 제외한 나머지 항목은 전문성이 낮은 편이었다. 가장 두드러진 항목은 7번으로 0.95정도가 상승하여 매우 낮은 편에서 낮은 편으로 상승하여, 가장 높은 상승률을 나타내었다. 사후 결과가 여전히



[그림 Ⅲ-1] 평가도구 개발 전문성 항목별 평균 점수의 상승 정도



[그림 Ⅲ-2] 평가도구 개발 전문성 교사별 평균 점수의 상승 정도



히 전문성이 낮은 편이지만, 문항 개발 연수가 표준화 검사를 이해하는데 도움이 되었음을 말해준다.

각 교사들이 제시한 평가도구 개발 전문성 점수의 평균은 사전에 비해 사후에 높아졌으며 [그림 III-2]와 같다.

#### 나. 문항 검토 의견의 변화

연수 초반에 교사들이 개별적으로 작성한 문항 검토 의견서를 분석하여 문항 개발 원칙 요소별 반응을 분석하고, 연수 중에 교사들이 팀별로 문항을 검토하고 그 결과를 팀별로 발표하는 과정에서 나타난 검토 의견을 문항 개발 원칙의 요소별 빈도와 비율을 분석한 결과는 <표 III-1>과 같다. 개별 검토 의견을 분류한 연수 초와 팀내 검토 의견을 분류한 연수 중의 빈도를 직접적으로 비교하기는 어렵기 때문에, 비율을 비교하였다. 그 결과, 교사들이 문항 개발 원칙에 따라서 문항을 검토하고 논의하는 과정을 거치면서 교육과정 수준의 검토 의견의 비율이 증가하였다. 교육과정 수준의 검토 의견이 가장 높은 비율로 나타났으며, 이들 검토 의견은 연수 초반에 개별 검토에서 제시된 의견들 중 문항 내적

수준의 의견들이 정교화된 것이다. 또한 문항의 출제의도가 교육과정의 내용과 수준에 맞는지에 대한 검토 의견이 연수 초반보다 확연히 증가하였다.

문항 별로 살펴보면, 1번 문항에서 교사들은 연수 초에 문항 타당도에 관한 의견을 가장 많이 제시하였다. 문항 타당도 관련 의견들은 근거 없이 제시된 경우가 많아, 교사들이 직관적으로 문항 타당도가 낮다는 것을 인식하고 있으나 논리적으로는 설명하지 않았다. 예를 들어, <표 III-2>에 제시된 바와 같이, 계산량이 많다, 불필요하게 복잡하다. 숫자를 조정해야 한다 등의 의견을 제시하였다. 연수 중에는 교육과정의 범위와 수준에 적합한지 혹은 평가 목적에 부합하는지에 대한 의견의 비율이 크게 증가하였다.

2번 문항에서도 연수 초에는 문항 내적 수준의 검토 의견이 가장 많이 나타났고, 교육과정의 범위와 수준에 적합한지 혹은 평가 목적에 부합하는지에 대한 의견은 소수로 나타났다. <표 III-3>과 같이, 대부분의 의견이 검토자의 관점에서 문항을 이해하는데 초점이 있었으며, 이유나 근거를 논리적으로 제시한 의견은 거의 없었다. 연수 중에 나타난 검토 의견은 1번 문항에 비해 변화가 두드러지지 않았다. 팀내 검토에서 처음으로 나타난 대표적인 검토 의견은 교대수열이

<표 III-1> 연수 초반 문항 개발 원칙 요소별 교사 반응 빈도 및 비율(%)

수준	하위 요소	1번				2번			
		연수 초		연수 중		연수 초		연수 중	
		빈도	비율	빈도	비율	빈도	비율	빈도	비율
교육과정	교육과정 및 교과 내용의 범위, 수준		13.3	1	54.5	1	21.4	1	20.0
	평가 목적 부합 여부	2		5		2		1	
문항 내적	문항 오류		73.3		27.3		71.4		70.0
	문항의 타당도, 변별도, 난이도	11		3		6		2	
	용어 수준								
	정확성, 단서					4		5	
문항 외적	문두(발문) 표현		13.3	1	18.2	1	7.1		10.0
	선택지	2		1				1	
	배점								
	편집 체제								
합		15	100	11	100	14	100	10	100

<표 III-2> 1번 문항에 대한 검토 의견 (예시)

수준	연수 초	연수 중
교육 과정	<ul style="list-style-type: none"> <li>출제 의도와 무관한 개념을 너무 많이 요구하고 있다.</li> <li>치환은 출제 의도와 무관한 것 같다.</li> </ul>	<ul style="list-style-type: none"> <li>출제 의도를 보니까, '정적분을 할 수 있다'가 있는데 이것을 판단하기 위해서 넓이 구할 때 (정적분) 계산을 할 수 있도록 2차 함수나 3차 함수를 고려했다. 이것저것해 보다가 3차 함수로 바꾸기로 하였다.</li> <li>출제 의도에서 정적분의 계산을 할 수 있다고 해서, 정적분 계산을 강조하기 위해서 (함수의) 양수 구간을 2차 함수로 바꾸었다.</li> <li>(발문에서) 마이너스를 다 빼버렸는데, 의미가 없다고 생각했다. 있으면 평가요소가 하나가 더 들어가니까 그럴 필요가 없다고 생각한다.</li> <li>(평가 목표가 포함되어 있는) 정적분의 성질을 (이해하는지를 문항에) 넣기 위해 적분구간은 2015부터 2018까지로 바꾸었다.</li> </ul>
문항 내적	<ul style="list-style-type: none"> <li>주기성을 좀 더 간단하게 할 수 있도록 숫자를 조절할 필요가 있다.</li> <li>구하고자 하는 범위를 수정할 필요가 있다.</li> <li>쓸데없이 어렵게 느껴진다.</li> <li>주어진 함수 <math>f(x)</math>가 원점대칭인 함수이므로 <math>f(x) = -f(-x)</math>가 동일한데 지나치게 도형의 변환이 많다.</li> <li>윗끝과 아래끝이 크기순으로 배열되지 않은 점에서 학생들이 한 번 더 꼬여서 정답을 구하는데 오류를 많이 발생할 것 같다.</li> <li>지나치게 너무 많은 음수가 있다.</li> <li>(가), (나) 조건을 이해하고 함수를 분석하는 시간까지 한다면 시간이 부족하다.</li> </ul>	<ul style="list-style-type: none"> <li>구간이 음수구간이고 너무 커서, 학생들이 실수할 수 있으니까 양수 구간으로 바꾸었다.</li> </ul>
문항 외적	<ul style="list-style-type: none"> <li>적분구간의 길이가 2이므로 선택지의 5가지 중 정답과 전혀 무관하다고 판단되는 것이 있다.</li> <li>주어진 함수 <math>f(x)</math>의 주기 2에서의 넓이가 1이므로 ②, ④번 중 하나로 학생들이 짚을 것 같다.</li> </ul>	<ul style="list-style-type: none"> <li>박스형으로 물어볼 필요가 없다고 생각한다. 학생들은 박스형으로 물어보면 더 어려워한다. 그러니까 박스를 빼고 그냥 기술하는 방식으로 하는 게 더 좋을 것 같다.</li> </ul>

<표 III-3> 2번 문항에 대한 검토 의견 (예시)

구분	연수 초	연수 중
교육 과정	<ul style="list-style-type: none"> <li>고등학교 교육과정의 수열 개념이 필요하지 않다. 수열의 개념이 없어도 답을 찾을 수가 있다.</li> <li>출제자의 평가의도와 다른 방법으로 문항을 해결하게 된다.</li> <li>규칙성을 금방 발견하게 된다. 귀납적 추론을 하지 않아도 문제가 해결된다.</li> </ul>	<ul style="list-style-type: none"> <li>이 문항은 교대수열을 다루고 있기 때문에 고등학교 교육과정을 벗어난 것이 아닌가 고민을 했다. 식이 아니라 그림으로 주어졌으니까 될 것 같기도 하고 안 될 것 같기도 하다.</li> <li>규칙성을 금방 발견하게 된다. 귀납적 추론을 하지 않아도 문제가 해결된다.</li> </ul>
문항 내적	<ul style="list-style-type: none"> <li>일반항을 2번 구하게 되고 불필요한 문구나 숫자가 있다.</li> <li>문제를 명확하게 이해하기가 어렵다.</li> <li>학생들의 수학실력에 대한 평가보다는 느낌만 어렵게 느껴진다.</li> <li>문제파악에 어려움이 있었다. 대각선 방향으로 넣는다는 말이 무슨 말인지 이해가 잘 안돼서 <math>T_1</math>까지 그림을 그려줬으면 좋겠다.</li> <li>일반항을 안 구해도 답이 나온다. 일반항을 구하는데 더 복잡하다.</li> </ul>	<ul style="list-style-type: none"> <li>대각선을 표시해줘야 할 것 같다.</li> <li>귀납적인 추론이 가능하도록 정확하게 설명해줘야 한다.</li> <li>일반항을 구하지 않고도 학생들이 짚을 수 있기 때문에 일반항을 구할 수 있도록 문항을 수정해야 한다.</li> <li>문양이 필요하지 않다.</li> <li>문제를 읽고 이해하기 어려웠다. 특히 3번째 그림까지 보고 예상한 방법으로 하면 5번째 그림이 다른 그림이 나왔다. 그래서 그림이 하나 더 있어야 한다고 생각한다.</li> </ul>
문항 외적	<ul style="list-style-type: none"> <li>굳이 20번째, 31번째 항을 물어볼 이유가 없다.</li> </ul>	<ul style="list-style-type: none"> <li>답이 홀수밖에 되지 않기 때문에 ②, ④번만 답이 되요. 그래서 답지를 수정할 필요가 있다.</li> </ul>

고등학교 교육과정의 내용과 수준인가를 점검하는 논의와 문양이 필요하지 않기 때문에 문양을 제외하는 방향으로 문항을 수정하자는 의견이었다.

### 3. 연수 요인별 효과

연수 3주 후에 4명의 교사(교사 I, M, P, R)를 대상으로 개별 면담을 하였다. 4명의 교사는 연수 평가지를 작성한 후에, 각 항목별로 자신의 의견을 적극적으로 설명하였다. 교사들이 5점 척도로 제시한 항목별 평균은 3점에서 5점까지 다양하게 나타났고 <표 III-4>와 같다.

<표 III-4> 연수 평가지 결과

연수 요인	연수 프로그램 구조			문항 개발 습관 표면 화와 수정		교사간 공감대 형성		전문성 자각	
	1	2(1)	2(2)	3	4	5	6	7	8
평균	4.75	3	4.5	4	5	4.75	4.5	4.5	4.75

전반적으로 높은 점수를 주는 경향이 있고 최하점이 3점임을 고려하여, 3점은 부정적인 답변으로 4점은 보통, 5점은 긍정적인 답변으로 해석할 수 있다. 특히, 3점에 해당하는 2(1)번 항목에 대한 4명 교사의 면담을 통해서 부정적인 답변임을 확인할 수 있었다.

연구자 : (2(1)을 가리키며) 연수 시간이 충분하다에 보통이라고 하셨습니다. 보통이라는 것은 어떤 뜻인가요?

교사 R : 충분하지는 않다는 이야기예요. 충분하면 이런 연수 안 받아도 되죠.

연수 프로그램의 구조적인 측면에서는 부정적인 요인과 긍정적인 요인이 모두 있었다. 부정적인 요인은 연수 시간이었다. 참여 교사들이 연수 시간이 충분하지 않았다고 답하였고 더 많은 문

항에 대한 실습이 필요하다고 생각하였다. 긍정적인 요인은 소그룹 활동(1번, 4.75점)이었다. 소그룹 활동은 팀내 검토와 팀간 검토 과정이 이루어지는데 효과적이었다고 판단된다. 교사들이 개별적으로 검토하는 것보다는 소그룹을 통해서 타교사와의 논의 과정 속에서 그리고 다른 팀의 의견을 통해서 문항 개발에 대한 안목이 향상되었다고 인식하였다.

문항 개발 습관 표면화와 수정의 측면에서는 강사의 피드백(4번, 5점)이 효과적인 요인이었다. 강사는 문항 개발 원칙에 따라서 안내자의 역할을 담당하였음을 고려할 때, 이는 문항 개발 원칙에 근거하여 문항을 검토하고 이것을 논의하는 과정이 전문성 향상에 도움이 되었음을 의미한다.

교사간 공감대 형성 측면에서는 교사의 적극적인 참여가 중요한 요인이었다. 그리고 전문성 자각 측면에서는 교사들이 자신의 전문성 부족을 인식했던 점이 효과적인 요인이었다. 전문성 부족 자각은 교사가 짧은 연수 시간 내에 모든 지식을 습득하기에는 어려웠음을 인식하였음을 의미하며, 동시에 연수 프로그램이 교사에게 향후 스스로 학습할 수 있는 동기를 제공해 준 것으로 볼 수 있다. 연수의 필요성을 느꼈다는 8번 항목의 점수가 매우 높았고 면담 과정에서도 확인되었다.

교사 I : (이 연수는) 도움이 많이 되었어요. 그런데 그렇게 많이 해보았는데 아직까지는 잘 모르겠어요.

연구자 : 연수는 6시간으로 이루어졌습니다. 이 시간을 기준으로 봤을 때는 가장 도움이 된 부분은 무엇인가요?

교사 I : (...) 문제를 이런 식으로 변화시키면 되는구나. 타당도 신뢰도 이런 것은 아직 잘 모르겠고 이론적으로는 알겠는데 이 문제는 이렇게 변형시키면 더 좋겠구나 이런 생각을 할 수 있다는 것이 가장 큰 변화예요.

종합적으로 연수의 효과를 저해하는 요인은 짧은 연수 시간이었고, 효과를 증진하는 요인은 소그룹 활동을 통한 팀내/팀간 토론, 강사의 피드백, 적극적인 참여와 전문성 부족의 자각이었다.

#### IV. 요약 및 시사

본 연구의 목적은 중고등학교 수학교사의 평가도구 개발 전문성 향상을 위한 문항 개발 연수의 효과를 분석하는 것이다. 본 연구에서는 자격 연수나 직무 연수에서 문항 개발 연수가 짧은 시간으로 이루어진다는 점을 고려하여, 6차시 연수를 일반적인 경우로 가정하였다. 연수 과정은 실습 위주로 구성하였으며, 연수의 효과는 평가도구 개발 전문성 평가지, 문항 검토 의견서, 연수 중의 발표, 연수 후의 연수 평가지와 면담 등을 통하여 분석하였다. 연구의 결과를 연구 질문별로 세 가지로 정리할 수 있다.

첫째, 문항 개발 연수는 평가도구 개발 전문성 향상에 효과가 있다고 참여 교사들이 인식하였다. 특히 연수 후에 표준화 검사의 개념, 특징, 종류, 활용법에 대해 자신감이 낮았던 교사들이 연수를 통해서 이 부분에 대한 전문성이 향상되었다고 판단하였다. 또한 문항의 난이도와 변별도를 이용하여 문항을 검토하는 능력이 향상되었다고 생각하였다. 이 결과는 교사들이 인식하는 전문성 정도는 교사의 경력보다는 평가 연수를 이수하였는가에 의해 크게 영향을 받는다는 Zhang & Burry-Stock(2003)의 연구와 일맥상통한다. 문항 개발 연수는 평가도구 개발 전문성에 대한 교사의 자신감을 높이는데 효과가 있다고 볼 수 있다.

둘째, 문항 개발 연수가 진행되면서 근거 없는 검토 의견의 비율이 줄어들고 구체적인 근거를 제시하는 검토 의견의 비율이 증가하였다. 연수

전에 교사들의 문항 검토 의견 수준은 단순히 해당 문항이 학생들의 이해 정도를 측정하기에 타당하지 않을 것 같다는 정도이었다. 너무 어렵다 혹은 문제를 해결하는데 시간이 오래 걸린다 등 단편적인 근거를 대는 경향이 있었다. 분석을 한 2문항 중 1문항에서 연수가 진행됨에 따라서 교육과정 수준의 검토 의견의 비율이 확연하게 증가하였다. 하지만 다른 1문항에서는 큰 차이가 나타나지 않았기 때문에 이 부분에 대해서는 추가적인 검증이 필요하다.

셋째, 평가도구 개발 전문성 향상 효과를 저해하는 연수의 요인은 제한된 연수시간(6시간)이었다. 일반적으로 개설되는 문항 개발 연수가 6시간 이내에 이루어지고 있다고 가정할 때, 이러한 종류의 연수의 효과를 기대하기는 어렵다고 볼 수 있다. 한편 30시간의 문항 개발 연수를 분석한 최지선(2013)에 의하면 최소 16시간 후에 연수 효과가 나타나기 시작한다고 보고한 바 있다. 이들 연구 결과는 연수의 효과는 연수 시간과도 관련되어 있음을 말해준다. 효과적인 연수의 요인은 소그룹 활동을 통한 팀내/팀간 토론, 강사의 피드백, 적극적인 참여와 전문성 부족의 자각이었다. 교사 개인이 문항 의견에 대해 고민한 이후에 미처 생각하지 못했던 또는 관점이 다른 타인의 의견을 통해서 교사들의 메타인지 기능이 활성화된 것으로 해석할 수 있다.

이 연구의 결과를 바탕으로 평가도구 개발 전문성 향상을 위한 교사 연수에 대해 도출할 수 있는 시사점은 문항 개발 연수를 단기적인 연수와 장기적인 연수로 이원화하여 개발할 필요가 있다는 점이다. 우선, 단기적으로 이루어지는 문항 개발 연수를 통해서 문항 개발 지침과 전문성 향상을 위한 방향을 탐색하고, 평가도구 개발 전문성에 대한 자신감을 향상시킬 필요가 있다. 문항 개발 연수 과정에서 교사들은 자신의 의견을 개진하고 동료 교사들과 논의하고 이전에 미

쳐 생각지 못했던 의견을 공유하면서 자신이 인식하는 평가도구 개발 전문성이 증진되었기 때문이다. 적정 연수 시간을 고려한 연수 프로그램을 개발하여 교사들의 적극적인 실습과 토론 참여를 독려하고, 지속적인 전문성 향상의 필요성을 자각하도록 해야 한다.

단기연수를 통해서 평가도구 개발 전문성 향상을 위해 지속적으로 노력할 필요성과 자신감을 얻은 후에, 다른 한편으로는 장기적인 관점에서 교사 공동체 활동을 통한 교사의 자기연수가 이루어져야 한다. 교사의 자기연수는 일종의 학습이며, 학습은 사회적이며 물질적인 맥락에서 분리될 수 없는 상황 의존적 활동이라는 관점에서 볼 때(Brown, 1989; Greeno et al., 1998), 교사의 공동체 활동은 전문성 개발에 영향을 준다(김희경, 2007; 박영희, 2011; 박영순, 2012). 단기적인 연수를 통해서 지속적인 문항 개발 실습을 위한 발돋움을 하고, 장기적으로 교사 공동체 활동을 통해서 지속적인 노력을 할 때, 교사의 평가도구 개발 전문성이 향상될 수 있다.

이 연구는 몇 가지 제한점을 가지고 있다. 첫째, 해당 연수는 지역교육청이 주관하였기 때문에 연수 사전에 강사와 연수 참여자 사이의 의사소통이 원활하게 이루어지지 않았다. 사전에 참여자들에게 과제를 부여하기가 어려웠고 소수의 교사는 연수 과정보다는 출석여부를 중요시하였다. 둘째, 교사들이 직접 문항 초안을 구안하는 대신 미리 구안된 문항 초안을 수정하는 과정을 통해서 평가도구 개발 과정의 일부만을 실습하였다는 한계가 있다. 셋째, 평가도구 개발 전문성 평가지를 통해서 연수의 효과를 간접적으로 측정하였다. 넷째, 연수의 효과를 요인별로 파악하기 위하여 활용한 연수 평가지에 대한 타당도를 검증하지 않고 활용하였다.

본 연구의 제한점에도 불구하고, 교사들을 대상으로 하는 평가 문항 개발 연수들이 주로 6시

간 이내의 짧은 시간동안에 이루어진다는 점을 고려할 때, 본 연구의 결과는 실제 현직 교사들을 대상으로 하는 연수 과정을 개발하고 발전시키는 데 시사점을 제공한다. 이후에 단기적으로 이루어진 연수에 대한 거듭된 연구가 이루어진다면, 교사들이 주로 활용하는 연수의 효과를 높이는 방안을 마련하는데 일조할 것으로 보인다.

## 참고문헌

- 강현영, 고은성, 김태순, 조완영, 이경화, 이동환 (2011). 좋은 수학수업을 위해 수학교사에게 필요한 역량과 교사교육에 대한 현직교사의 인식조사. **학교수학**, 13(4), 633-649.
- 고정화(2010). 평가 문항의 질 향상을 위한 문항 수정 유형 분석. **학교수학**, 12(2), 113-136.
- 박영순(2012). 학습자의 핵심역량 제고를 위한 과학 교사교육 개선 방안. **한국지구과학회지**, 33(2), 162-169.
- 김선희(2006). 학생평가 전문성을 갖춘 수학교사 양성을 위한 "수학학습평가" 강좌의 교육 내용과 방법에 대한 제안. **학교수학**, 8(3), 301-326.
- 김선희(2012a). 실습과 반성을 통한 수학 예비교사의 평가 전문성 신장. **수학교육학연구**, 22(2), 277-292.
- 김선희(2012b). 수학 예비교사가 개발한 평가 문항의 교육과정 충실도와 문항 완성도. **학교수학**, 14(4), 517-529.
- 김수동, 김선희(2005). **수학 수업에서 학생평가를 잘 하려면**, 한국교육과정평가원 연구보고서 ORM 2005-51-4. 한국교육과정평가원.
- 김희경(2007). 과학 교사의 전문성 개발 프로그램의 조건과 모형. **초등과학교육**, 26(3), 295-308.
- 문영수(2008). **총괄평가 문항 개발에 따른 내용**

- 과 편집상의 문제점 및 단원별 평가 내용에 관한 연구 : '수학 10-가'를 중심으로. 한국교육대 학원 석사학위논문.
- 박경미, 정영옥, 김화경, 김동원, 최수일, 최지선 (2010). 우리나라 초·중등학교 수학 교육 발전 방안 기획연구. 한국과학창의재단.
- 박영희(2011). 초등 수학 수업 전문성 신장을 위한 대학과 초등학교의 학습공동체 사례 연구. **수학교육 논문집**, 25(1), 47-61.
- 송미영, 김경희(2007). 교사의 학생평가 전문성 진단도구 개발. **교육과정평가연구**, 10(1), 47-74.
- 신범영(2010). 바람직한 수학 평가 문항 개발에 대한 연구. 서울시립대학교 교육대학원 석사학위논문.
- 윤열현(2012). 중등 수학과 자격 연수의 운영 실태, 연수 효과 및 개선 방안 연구. 한국교원대학교 석사학위논문.
- 이광호, 장은하(2012) 문제중심학습(PBL)에서 초등 예비교사들의 문제해결과정. **초등수학교육**, 15(2), 91-106.
- 이석현, 한길준(2012). 교수에 대한 수학적 지식의 관점에서 본 지필평가문항 분석. **한국수학사학회지**, 25(2), 97-111.
- 이경화, 나귀수, 권나영, 김동원, 이환철, 이동환, 고은성, 박민선, 박미미, 이은정, 조진우, 박진형(2012). 한국형 수학교사 전문성 개발 체제(PDS) 모델 구축을 위한 기초 연구. **수학교육학연구**, 22(4), 581-602.
- 이환철, 장미숙(2012). 수학교육 관련 교사 연수 현황 분석을 통한 정책 방안 모색. **한국학교수학회논문집**, 15(1), 171-182.
- 장민아(2009). 수학교사의 학생평가 실태 조사 및 평가 활용 방안 연구. 이화여자대학교 교육대학원 석사학위논문.
- 전영주(2012). 수학과 평가 문항제작의 실제. **한국학교수학회논문집**, 15(2), 281-297.
- 최승현, 강대현, 곽영순, 장경숙(2008). 교과별 내용교수지식(PCK) 연구(II)-중등 초임교사 수업컨설팅을 중심으로. 한국교육과정평가원 보고서 RRI 2008-3.
- 최지선(2013). 선택형 평가 문항 제작 전문성 향상 연수 프로그램 효과 제고에 대한 사례연구. **수학교육학연구**, 23(2), 193-212.
- McMillan, J. H. (2004). *Classroom assessment principle and practice for effective instruction*. Boston: Allyn & Bacon.
- Brown, J. S. (1989). Situated cognition and the culture of learning. *Educational Researcher*, 18(1), 32-42.
- Greeno, J. G., & The Middle School Mathematics Through Applications Project Group. (1998). The situativity of knowing, learning, and research. *American Psychologist*, 53(1), 5-26.
- Zhang, Z., & Burry-Stock, J. A. (2003). Classroom Assessment Practices and Teachers' Self-Perceived Assessment Skills. *Applied Measurement In Education*, 16(4), 323 - 342

# The Effectiveness of the Training Program to Improve Mathematics Teachers' Professional Competency of Developing Assessment Instrument

Choi, Jiseon (Korea Institute for Curriculum and Evaluation)

This study aims to analyse the effectiveness of the mathematics teacher's training program: whether the training program is effective to improve mathematics teachers' professional competency of the developing assessment instrument (including items) or not.

The teachers who were attendee of the program did pre-test before the program and post-test after the program. They wrote the opinions on the pre-developed items which had many errors in the beginning and discussed their opinions each others in the middle of program. The differences between pre-test and post-test and between opinions in the beginning and in the middle were analysed.

The first result of the study is that the teacher's professional competency which is evaluated as self-perceived points is improved with regard to understanding the standardized test, item difficulty and item discrimination.

Secondly, the proportion of the opinion with concrete reasons is increased as the program progressed.

Thirdly, the effective elements of the program are the discussion in the group, discussion between groups, and feedback from the instructor. But the ineffective element of the program is the insufficient time for practicing.

\* Key Words : mathematics teacher(수학교사), teacher's training program (교사 연수 프로그램), Professional Competency of Developing Assessment Instrument (평가도구 개발 전문성)

논문접수 : 2014. 4. 11

논문수정 : 2014. 5. 5

심사완료 : 2014. 5. 9