



저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

교육학박사학위논문

예비 화학교사의 수행평가 전문성 향상을
위한 프로그램의 개발 및 적용

Development and Application of the Program
to Improve Pre-Service Chemistry Teachers'
Performance Assessment Professionalism

2019년 2월

서울대학교 대학원
과학교육과 화학전공
이 재 원

예비 화학교사의 수행평가 전문성 향상을
위한 프로그램의 개발 및 적용

Development and Application of the Program
to Improve Pre-Service Chemistry Teachers'
Performance Assessment Professionalism

指導教授 盧 泰 熙

이 論文을 教育學博士學位 論文으로 提出함.

2018年 11月

서울大學校 大學院

科學教育科 化學專攻

李 在 原

李在原의 博士學位論文을 認准함.

2018年 12月

委 員 長 _____

副委員長 _____

委 員 _____

委 員 _____

委 員 _____

국 문 초 록

학생 중심의 수준별·맞춤형 수업을 강조하는 국가 수준의 교육과정이 학교 현장에 올바르게 정착하기 위해서는 평가도 이에 부합하도록 변화해야 한다. 수행평가는 학생들의 상대적인 석차를 산출하는 지필평가에 비해 개별 학생의 성취수준을 타당하게 판단할 수 있는 평가 방법이며, 현재 과학교육의 패러다임으로 자리잡은 구성주의적 평가를 구현할 수 있는 이상적인 평가 체계이다. 최근 성취평가제가 도입됨에 따라, 교사의 수행평가 전문성 향상은 더욱 중요해지고 있다. 그러나 많은 과학교사는 수행평가를 구성주의적으로 개발하고 실행하는 데 어려움을 겪고 있다. 따라서 과학교사의 수행평가 전문성 향상을 위한 전략의 모색이 필요하다. 수행평가의 개발 과정에는 과학 평가에 관한 지식뿐 아니라 과학교육과정, 과학 교수 전략, 학생, 과학 내용에 관한 지식과 과학 평가에 관한 신념 등 다양한 교과교육학 지식(pedagogical content knowledge; PCK) 구성요소가 활용된다. 따라서 PCK를 효과적으로 활용할 수 있는 평가 관련 활동을 제공하는 프로그램을 통하여 수행평가 전문성 향상을 도모하는 접근 방법은 유용할 수 있다.

과학교사의 수행평가 전문성 향상 프로그램을 개발하기 위한 기초 연구로서 중등 과학교사의 평가 전문성에 관한 수준과 문제점을 조사하였고, 예비교사 단계에서의 교육 필요성을 도출하였다(연구 I). 이에 중등 예비 화학교사를 대상으로 수행평가 개발 과정에서 나타나는 특징을 PCK 구성요소와 구성요소 사이 연계의 측면에서 조사하였다(연구 II). 이후 연구 I과 II의 결과를 분석하여 전략과 활동을 구성하고, 교사 전문성 개발 원리에 기초하여 예비 화학교사의 수행평가 전문성 향상을 위한 프로그램을 개발하였다. 개발한 프로그램을 예비 화학교사에게 적용한 후, 수행평가 과제 개발 과정에서 나타나는 특징을 PCK 구성요소와 구성요소 사이 연계의 측면에서 조사하고 연구 II의 결과와 비교하였다(연구 III).

연구 I에서는 92명의 중등 과학교사를 대상으로 그들이 실시하는 과학 평가와 지향하는 평가 사례를 평가 목적과 방법의 측면에서 조사하였다. 교사의 배경 변인에 따른 형성평가의 활용 정도도 조사하였다. 교사들이 실시하는 과학 평가 사례를 분석한 결과, 대부분의 교사들은 전통적인 총괄평가 목적의 평가를 실시하는 경향이 있었고, 선다형 평가, 실험, 보고서, 논술형 평가가 가장 많이 사용하는 평가 방

법이였다. 전형적인 수행평가 상황에서 과학교사들은 다양한 유형의 구성주의적 평가 방법을 지향했지만, 여전히 전통적 평가 목적에서 벗어나지는 못하였다. 실험 및 자유탐구 상황에서 구성주의적 과학 평가관 수준이 높은 교사가 형성평가를 지향하는 비율은 구성주의적 과학 평가관 수준이 낮은 교사들에 비해 유의미하게 높았다($p=.018$, $p=.023$). 연구 I을 통하여 교사의 구성주의적 평가에 대한 인식 함양의 중요성을 알 수 있었다. 하지만 이는 단기간에 이루어지기 어려우므로 구성주의적 수행평가의 확산을 위해서는 체계적이고 장기적인 교육이 가능한 예비교사 단계에서부터 교육이 이루어져야 한다. 이에 이 연구의 목표를 예비 화학교사의 수행평가 전문성 향상으로 구체화하였다.

연구 II에서는 예비 화학교사 8명을 대상으로 수행평가 과제 개발 과정에서 나타나는 PCK 구성요소와 구성요소 사이 연계의 특징을 분석하였다. 이들은 간단한 수행평가 워크숍 후 구상 단계, 구체화 단계로 이루어진 개별 과제 개발 과정과 조별 논의 과정을 거쳐 수행평가 과제를 개발하였다. 활동 과정에 대한 녹음 및 녹화와 반구조화된 면담을 실시하였다. 분석 결과, 개별 과제 개발 과정에서 나타난 PCK 구성요소는 과학 평가에 관한 지식에 편중되어 있었다. 과학교육과정, 과학 교수 전략, 학생에 관한 지식의 비율은 낮았으며 학생에게 제공되는 피드백도 고려되지 않았다. 구상 단계보다 구체화 단계에서 평가의 공정성과 객관성을 중시하는 전통적 평가 지향이 강하게 나타났다. PCK 구성요소 사이의 연계는 구상 단계에서 과학 교수 전략에 관한 지식을 중심으로 이루어졌다. 그러나 구체화 단계 및 조별 논의 과정에서 PCK 구성요소 사이의 연계는 거의 나타나지 않았다. 조별 논의 과정에서 나타난 PCK 구성요소의 특징은 구체화 단계와 유사하였다. 조별 논의를 통하여 예비교사들이 고려하지 못했던 일부 PCK 구성요소가 보완되었으나 논의가 주로 전통적 관점에서 이루어졌다. 이는 단순히 수행평가 과제 개발 기회를 제공하는 것만으로는 예비교사의 수행평가 전문성 향상에 큰 한계가 있으며, PCK 구성요소의 활용 측면에서 더욱 실질적인 도움이 제공되어야 함을 시사한다.

연구 III에서는 연구 I과 II의 결과를 종합하여 교사 전문성 개발 원리와 전문성 개발 설계 프레임워크를 바탕으로 예비 화학교사의 수행평가 전문성 향상 프로그램을 개발하였다. 연구 II의 결과를 고려하여 예비교사들이 PCK를 활용하고 개발할 기회를 제공하였으며, 점검표와 학생 오개념 자료를 제공하였다. 이와 동시에 연구 I의

결과를 토대로 구성주의적 수행평가에 대한 이론과 실행을 통합할 수 있도록 연구자-예비교사 사이의 협력적 논의 및 교사와 학생의 이중적 입장에서의 논의 활동을 실시하였다. 예비 화학교사 8명이 3주 동안 프로그램에 참여한 후 연구 II와 같은 방법으로 수행평가 과제를 개발하였다. 자료 수집 및 분석 방법도 연구 II와 같게 하여 두 연구의 결과를 비교하였고, 추가로 발성 사고법을 활용하여 수행평가 과제의 개발 과정을 분석하였다. 개별 과제 개발 과정의 분석 결과, 구상 단계에서는 과학 평가, 과학 교수 전략, 학생, 과학 내용에 관한 지식이 거의 동일한 비율로 나타났고, 구체화 단계에서는 과학 평가에 관한 지식이 주로 나타났다. 조별 논의 과정에서는 과학 평가, 과학 교수 전략, 학생에 관한 지식의 비율이 거의 같았다. PCK 구성요소 사이의 연계는 개별 과제 개발 과정의 구상 단계에서 과학 교수 전략에 관한 지식을 중심으로 나타났고, 구체화 단계에서는 학생에 관한 지식을 중심으로 나타났다. 조별 논의 과정에서는 학생에 관한 지식이 적극적으로 고려되면서 다른 PCK 구성요소와의 연계가 활발하게 이루어지는 계기가 되었으며, 구성주의적 수행평가에 대한 이론과 실행의 간극을 구성주의적인 방향으로 극복하려는 노력도 이루어졌다. 하지만 과학교육과정에 관한 지식과 하위 요소 중 과학 과정 지식을 고려한 경우는 드물었고 과학교육과정과 과학 내용에 관한 지식이 다른 PCK 구성요소와 연계되는 경우는 적은 것으로 나타났다.

이상의 연구 결과를 바탕으로 이 연구에서 개발한 수행평가 전문성 향상 프로그램을 보완을 위해 협력적 논의에 참여하는 연구자의 역할 및 조별 논의가 효과적으로 이루어질 수 있도록 하는 방안에 대하여 논의하였다. 또한, 프로그램의 효과적인 활용을 위하여 사범대학에 개설된 강좌에 이 프로그램을 적용하는 구체적인 방안 및 연구자와 예비교사의 연결을 위한 지원 체계의 필요성을 제안하였다.

주요어: 수행평가, 구성주의적 평가, 교과교육학 지식, 교사 전문성 개발 원리, 프로그램 개발, 예비 화학교사

학 번: 2013-21428

목 차

국 문 초 록	i
목 차	iv
표 목 차	viii
그 립 목 차	x
부 록 목 차	xi

제 1 장. 서 론

1. 연구 배경과 목적	1
2. 연구 내용	9
3. 연구의 제한점	11
4. 용어의 정의	13

제 2 장. 이론적 배경 및 선행연구

1. 과학교사의 PCK와 평가 전문성	
1.1 PCK	15
(1) PCK의 개념	15
(2) PCK의 구성요소	18
1.2 평가 전문성	21
(1) 학생평가의 개념과 목적	21
(2) 수행평가의 정의와 특징	24
(3) 평가 전문성의 개념과 구인	33
1.3 PCK와 평가 전문성의 관계	38
2. 과학교사의 평가 전문성 향상	
2.1 과학교사의 전문성 향상	42

(1) 교사 전문성 개발 원리	43
(2) 전문성 개발 설계 프레임	49
2.2 과학교사의 수행평가 전문성 관련 선행연구	55
(1) 평가 실태와 인식 조사	55
(2) 평가 모형, 도구, 과제 개발	58
(3) 평가 채점 및 측정	61
(4) 평가 전문성 향상	63

제 3 장. 중등 과학교사의 과학 평가 실태와 지향 (연구 I)

1. 서 론	67
2. 연구 방법	69
2.1 연구 참여자	69
2.2 연구 절차	70
2.3 분석 방법	71
3. 연구 결과 및 논의	73
3.1 중등 과학교사의 과학 평가 실태	73
(1) 평가 목적	73
(2) 평가 방법	76
3.2 중등 과학교사의 과학 평가 지향	78
(1) 평가 목적	78
(2) 평가 방법	84
4. 결 론	88

제 4 장. 중등 예비 화학교사의 수행평가 과제 개발 과정에서 나타나는 PCK 구성 요소의 특징 (연구 II)

1. 서 론	91
2. 연구 방법	92

2.1 연구 참여자	92
2.2 연구 절차	93
2.3 분석 방법	97
3. 연구 결과 및 논의	99
3.1 개별 과제 개발 과정의 PCK 구성요소	99
3.2 개별 과제 개발 과정의 PCK 구성요소 사이의 연계	107
3.3 조별 논의 과정의 PCK 구성요소 및 구성요소 사이의 연계	110
4. 결 론	119

제 5 장. 수행평가 전문성 향상 프로그램에 참여한 예비 화학교사의 수행평가 과제 개발 과정에서 고려된 PCK 구성요소 및 연계의 특징 (연구 III)

1. 서 론	122
2. 연구 방법	125
2.1 연구 참여자	125
2.2 수행평가 전문성 향상 프로그램	125
2.3 연구 절차	132
2.4 분석 방법	134
3. 연구 결과 및 논의	135
3.1 개별 과제 개발 과정의 PCK 구성요소	135
3.2 개별 과제 개발 과정의 PCK 구성요소 사이의 연계	143
3.3 조별 논의 과정의 PCK 구성요소	154
3.4 조별 논의 과정의 PCK 구성요소 사이의 연계	161
4. 결 론	166

제 6 장. 결론 및 제언

1. 연구의 요약	169
2. 결론 및 제언	171

3. 추후 연구 과제	174
참고 문헌	177
부 록	200
ABSTRACT	214

표 목 차

<표 1> 주요 연구자들의 PCK 구성요소	20
<표 2> 수행평가와 유사한 용어의 정의와 수행평가와의 관련성	27
<표 3> 전통적 평가와 수행평가의 특징 비교	30
<표 I-1> 연구에 참여한 교사의 배경	70
<표 I-2> 교사가 실시하는 평가 목적	73
<표 I-3> 교사의 배경 변인에 따른 형성평가의 사용 정도에 대한 독립표본 t 검정 결과	76
<표 I-4> 교사가 실시하는 평가 방법	77
<표 I-5> 전형적인 과학 수행평가 상황에서 교사가 지향하는 평가 목적 ..	79
<표 I-6> 전형적인 수행평가 상황에서 교사의 배경 변인에 따른 형성평가의 지향 정도에 대한 독립표본 t 검정 결과	82
<표 I-7> 전형적인 과학 수행평가 상황에서 교사가 지향하는 평가 방법 ..	85
<표 II-1> 연구에 참여한 예비교사의 배경	93
<표 II-2> 반구조화된 면담의 개요	94
<표 II-3> 과학 수행평가에 관한 PCK 구성요소별 하위 요소와 정의	98
<표 II-4> 개별 과제 개발 과정의 PCK 구성요소	100
<표 II-5> 개별 과제 개발 과정의 PCK 구성요소 사이의 연계 (통합 유형별)	107
<표 II-6> 개별 과제 개발 과정의 PCK 구성요소 사이의 연계 (PCK 구성요소별)	108
<표 II-7> 조별 논의 과정의 PCK 구성요소	110
<표 II-8> 조별 논의 과정의 PCK 구성요소 사이의 연계 (통합 유형별)	116
<표 II-9> 조별 논의 과정의 PCK 구성요소 사이의 연계 (PCK 구성요소별)	116
<표 III-1> 연구에 참여한 예비교사의 배경	125
<표 III-2> 교사 전문성 개발 원리와 관련 프로그램 활동 및 도입 배경	130
<표 III-3> 프로그램 관련 반구조화된 면담의 개요	133

<표 III-4> 개별 과제 개발 과정의 PCK 구성요소	136
<표 III-5> 개별 과제 개발 과정의 PCK 구성요소 사이의 연계 (통합 유형별)	144
<표 III-6> 개별 과제 개발 과정의 PCK 구성요소 사이의 연계 (PCK 구성요소별)	145
<표 III-7> 조별 논의 과정의 PCK 구성요소	155
<표 III-8> 조별 논의 과정의 PCK 구성요소 사이의 연계 (통합 유형별)	162
<표 III-9> 조별 논의 과정의 PCK 구성요소 사이의 연계 (PCK 구성요소별)	163

그 립 목 차

〈그림 1〉 과학교사의 수행평가 전문성 향상과 PCK 개발의 관계	5
〈그림 2〉 연구 개요	8
〈그림 3〉 Magnusson, Krajcik, & Borko(1999)의 PCK 구성요소	19
〈그림 4〉 예비교사의 수행평가 전문성 구인	38
〈그림 5〉 수행평가 과제 개발 과정에서 PCK 구성요소의 활용	40
〈그림 6〉 과학 수행평가에 관한 PCK 구성요소	41
〈그림 7〉 예비교사를 위한 교사 전문성 개발 원리	49
〈그림 8〉 전문성 개발 설계 프레임	50
〈그림 II-1〉 예비교사 A가 개발한 수행평가 과제의 평가 준거	101
〈그림 II-2〉 예비교사 D가 개발한 수행평가 과제	106
〈그림 II-3〉 예비교사 E가 개발한 수행평가 과제	112
〈그림 II-4〉 예비교사 H가 개발한 수행평가 과제	115
〈그림 II-5〉 예비교사 C가 개발한 수행평가 과제	118
〈그림 III-1〉 수행평가 전문성 향상 프로그램의 주차별 활동	131
〈그림 III-2〉 예비교사 D가 개발한 수행평가 과제	140
〈그림 III-3〉 예비교사 H가 개발한 수행평가 과제	142
〈그림 III-4〉 예비교사 A가 개발한 수행평가 과제	148
〈그림 III-5〉 예비교사 G가 개발한 수행평가 과제	150
〈그림 III-6〉 예비교사 C가 개발한 수행평가 과제	152
〈그림 III-7〉 예비교사 B가 개발한 수행평가 과제	156
〈그림 III-8〉 예비교사 F가 개발한 수행평가 과제	158
〈그림 III-9〉 예비교사 G가 작성한 수행평가 아이디어	159

부 록 목 차

<부록 I-1> 과학 평가 실태와 인식 조사 설문지	200
<부록 II-1> 개별 과제 개발 단계의 활동지	206
<부록 III-1> 수행평가 전문성 향상 프로그램의 PPT 자료	208
<부록 III-2> 논의 방법 안내 자료	211
<부록 III-3> 학생 오개념 자료	212
<부록 III-4> 구성주의적 수행평가 과제 개발을 위한 점검표	213

제 1 장. 서 론

1. 연구 배경과 목적

융합적이고 창의적인 사고력이 중요한 현대 사회에서 과학과 기술에 대한 통섭의 능력과 과학적 소양을 지닌 인재양성은 국가경쟁력 확보를 위한 핵심과제라 할 수 있다(김진수, 2007; Katchi, Pearson, & Feder, 2009). 이에 우리나라에서도 2009 개정 교육과정에서부터 과학 교과에 융합인재교육(STEAM)을 비롯한 창의·인성 교육을 강조하고, 2018년도부터 적용되는 2015 개정 교육과정에서도 학생들의 과학적 소양과 핵심역량의 함양을 통한 창의·융합 인재양성을 강조하고 있다(교육과학기술부, 2011; 교육부, 2015). 또한, 이를 구현하기 위한 구체적인 방안으로 집중이수제 또는 블록타임제 등과 연계한 교과교실제를 확대 도입하여 경직된 교육과정 운영을 탈피하고 각급 학교의 상황에 따라 학생 중심의 수준별·맞춤형 수업이 이루어질 수 있는 여건 조성을 위해 노력하고 있다. 하지만 제도적으로 창의적이고 학생 중심적인 수업을 강조하더라도, 이에 대한 평가가 기존의 상대평가 방식에 머물러 있으면 수업의 변화에는 한계가 있다. 상대평가는 학생의 등급과 석차를 산출하는 평가 방식이므로 학생들 사이에 배타적 경쟁심과 이로 인한 스트레스를 유발하여 협력 학습을 어렵게 하고, 평가의 객관성이 지나치게 강조되어 교사가 시험을 대비한 수업(teaching to test; 남명호, 1995)과 줄 세우기 목적의 평가를 하도록 왜곡시킨다. 교육과정 운영의 측면에서도 상대평가는 학생의 수와 분포에 따라 등급과 석차가 달라지기 때문에 수준별·맞춤형 수업의 실행을 어렵게 한다(지은림, 2011). 따라서 교육과정의 변화가 초·중등학교 현장에 올바르게 정착하기 위해서는 수업뿐 아니라 평가도 이에 부합하는 방향으로 개선되어야 한다(Jonassen, Peck, & Wilson, 1999).

이러한 맥락에서 개별 학생이 무엇을 얼마나 성취하였는지 평가하는 성취평가에 관한 관심이 높아지고 있다. 성취평가는 준거지향평가의 일종으로 교육과정에 근거하여 설정된 성취수준을 준거로 학생들의 성취도를 평가하기 때문에 다른 학생과의 비교가 필요하지 않아 학생들 사이의 협력 학습 분위기를 조성할 수 있고, 교사의 수업 활동을 왜곡하지 않으며 학교별 상황에 맞춘 유연한 교육과정의 운영

이 가능하다는 장점이 있다(원효현, 2014). 이에 우리나라에서도 교육과정의 변화에 맞게 2012년부터 순차적으로 중등학교에 기존의 상대평가를 대체하는 성취평가제를 도입하고(교육과학기술부, 2011), 2015 개정 교육과정에서도 지속해서 성취평가제를 강조함으로써 성취평가가 학교 현장에 정착할 수 있도록 노력하고 있다(교육부, 2015).

이때 선택형, 단답형 평가 중심의 전통적 평가는 학습 목표나 내용의 지식적 측면에 대한 단편적인 이해 여부만을 평가하므로 개별 학생의 성취수준을 타당하게 파악하는 데 한계가 있다. 학생의 성취수준은 실제와 가까운 상황이 제시될 때 타당한 평가가 가능하며, 이를 위해서는 실험, 토의, 프로젝트형 평가 등 다양한 평가방법을 활용할 필요가 있다. 성취평가의 이러한 특징은 이미 1998년부터 정책적으로 도입되어 학교 현장에 시행 중인 수행평가와 맥락을 같이 한다(원효현, 2014). 수행평가는 학생의 고등 사고능력이나 지식의 실생활 적용 능력을 평가하기 어려운 전통적 평가의 단점을 극복하기 위해 도입된 평가 체계이다(교육부, 1998). 객관적 인식론에 기초한 전통적 평가와 달리, 수행평가는 현재 과학교육의 패러다임으로 자리 잡은 구성주의적 인식론에 이론적 배경을 두고 있다. 구성주의적 학습관에서 학생은 외부의 지식이나 정보를 수동적으로 받아들이는 것이 아니라 자신의 선지식과 경험을 바탕으로 능동적으로 재구성하고, 교사는 이를 돕는 조력자의 역할을 한다. 또한, 구성주의적 관점에서의 평가는 지식의 구성 과정과 고등 사고능력에 대한 평가, 다양한 관점과 양식으로 이루어지는 평가를 의미한다(박선미, 1998; 백순근, 1998; 성태제, 1999). 이에 수행평가는 학생들에게 실제적이고 유의미한 평가 과제를 제시하고, 지식의 정답 여부나 수행 결과뿐 아니라 과정과 질을 종합적으로 판단하여 평가할 것을 강조한다. 평가 결과의 활용 측면에서도 성적 산출과 같은 전통적인 목적뿐 아니라 평가를 통한 개별적 피드백 제공과 학생들의 발달과정 추적을 중시한다(백순근, 2000; Torrance, 1995). 따라서 수행평가는 성취평가가 학교 현장에 안정적으로 정착할 수 있는 기반을 제공할 뿐 아니라 구성주의적 교수학습 및 평가를 구현할 수 있는 이상적인 평가 체계로서도 교육적 의의가 크다고 할 수 있다(성태제, 2012; 원효현, 2014).

수행평가가 도입된 이래로 지금까지 수행평가에 대한 꾸준한 정책적 강조가 이어져 왔으며, 이로 인해 현재 수행평가는 학교에서 이루어지는 평가 중 낮지 않은

비율을 차지하고 있다(정은영, 2018). 특히 최근에는 성취평가제가 도입되면서 교육부에서도 수행평가의 내실화를 강조하며 교사에게 더욱 큰 수행평가 전문성을 요구하고 있다. 예를 들어, 교사가 성취평가제를 고려하여 수행평가를 계획하고 실행하기 위해서는 직접 학생의 수준을 고려한 성취기준을 설정하고, 학생들이 성취해야 할 지식, 기능, 태도의 특성을 설명하는 학기 단위의 성취수준을 작성해야 한다. 또한, 성취기준의 도달 정도를 측정할 수 있는 평가도구를 개발하고, 평가 결과를 종합하여 개별 학생의 성취수준을 최종 판단하는 등의 평가 활동을 해야 한다(김경희 등, 2012). 따라서 과학교사의 수행평가 전문성은 최근 그 의미와 중요성이 매우 커지고 있다.

그러나 많은 과학교사는 여전히 수행평가를 시행하는 데 어려움을 겪고 있으며, 특히 구성주의적 관점에서 의미 있는 수행평가도 제대로 이루어지지 않고 있다(김석우, 2007; 김수동, 김경희, 김수진, 2005; 김정환, 2005; 민희정, 2012; 손정숙, 2011; 신혜영, 2001; 심화선, 2006; 윤민화, 2010; 정은영, 최원호, 2014; 최은주, 2011; Dorman & Knightley, 2006). 예를 들어, 과학 교과의 경우 인지적 능력뿐 아니라 탐구능력과 태도 영역에 대한 평가가 중요하게 다루어져야 하지만 각 영역을 포괄하는 균형 잡힌 평가는 보기 어려우며, 특히 정의적 영역의 평가는 이루어지지 않는 경우가 많다(신정윤, 양일호, 2017). 또한, 수행평가의 타당도와 신뢰도, 평가관점의 구체성에 한계가 있었고(김석우, 2007), 평가 결과를 적절히 활용하지 못하거나(정시온, 2007), 평가 방법으로는 평가 영역과 관계없이 여전히 전통적인 지필평가를 선호하는 것으로 보고된다(고민석 등, 2013).

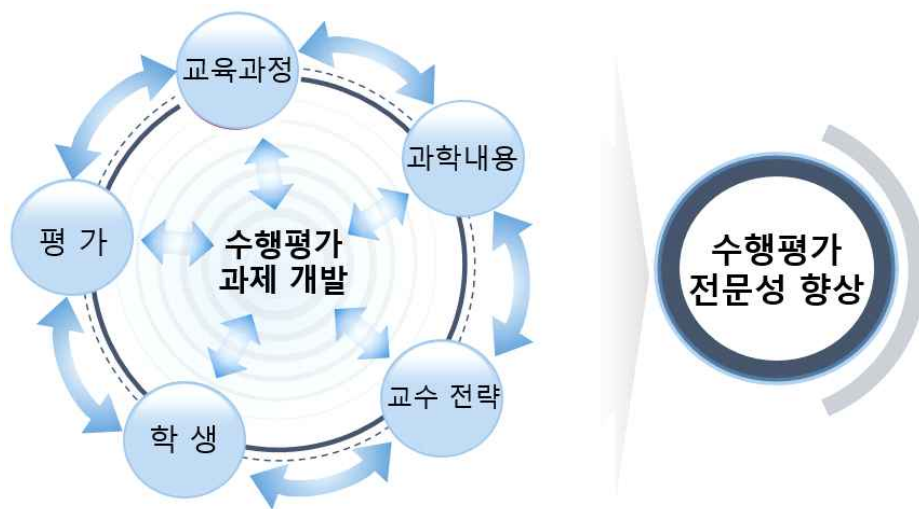
그럼에도 지금까지 이루어진 과학교사의 수행평가 전문성에 관한 연구는 전반적으로 매우 부족하다. 일부 진행된 연구들도 주로 특정 평가 영역이나 방법에 대한 교사의 인식(김경미, 김성원, 2002; 김성원, 현미숙, 2005; 김지영, 2005; 손정숙, 2011; 심애순, 2008; 심화선, 2006; 신혜영, 2001; 최은주, 2011; 홍혜정, 2012), 평가관(노태희, 윤지현, 강석진, 2009), 평가 문항 제작 수준(윤민화, 2010; 최은주, 2011; 최현숙, 김중복, 2013), 평가 실태(김호진, 성민웅, 박대오, 2000; 민희정, 2012; 정은영, 최원호, 2014), 특정 평가 방법의 효과(최경희, 이현주, 남정희, 2000) 등을 조사하는 측면에 치중되어 있다. 즉, 과학교사의 수행평가 전문성 향상에 관한 연구(김동영, 2006; 민희정, 2012; Falk, 2012)는 매우 부족한 실정이다. 이로 인해 교사양성

과정이나 각종 현직교사 연수 과정에서 교사들에게 수행평가 전문성 향상을 위한 교육 기회가 충분히 제공되지 못할 뿐만 아니라, 심지어는 사범대학에서 예비교사를 위한 평가와 관련된 강좌 자체도 매우 부족한 것으로 보고되고 있다(김선희, 2006). 따라서 과학교사의 수행평가 전문성 향상 방안을 모색하는 연구는 매우 중요하고 필요하다.

과학교사의 수행평가 전문성 향상을 위해서는 먼저 교사들이 수행평가를 계획하고 실행하는 과정에 주목할 필요가 있다. 과학 평가의 계획, 평가 과제의 개발 및 실행 과정에서는 과학 평가에 관한 지식뿐만 아니라 다양한 교과교육학 지식(pedagogical content knowledge; PCK) 구성요소의 종합적인 활용이 요구되는 것으로 보고되고 있다(민희정, 2012; 서미영, 2007; 윤민화, 2010; Alonzo & Aschbacher, 2005; Bell & Cowie, 2001; Black *et al.*, 2004; Falk, 2012). 교사의 전문성을 나타내는 대표적인 지표로 활용되는 PCK는 일반 교육학 지식(pedagogical knowledge; PK)과 교과 내용 지식(content knowledge; CK)의 결합체로서, PCK의 구성요소에는 대체로 과학 평가, 과학교육과정, 과학 교수 전략, 과학 내용, 학생에 관한 지식 및 과학 평가에 관한 교사의 신념 등이 포함된다(Loughran, Berry, & Mulhall, 2006; Magnusson, Krajcik, & Borko, 1999; Park & Oliver, 2008).

특히 실험, 토의 등 다양한 방법으로 이루어지는 수행평가는 수업과 함께 이루어지는 경우가 많고 학생의 능동적인 참여가 필수적이므로 과제 개발 과정에서 다양한 PCK 구성요소의 활용이 더욱 중요하다. 예를 들어, 과학 수행평가 과제 개발의 일반적인 절차(Doran, Chan, & Tamir, 1998)를 살펴보면, 성취기준 설정 단계에서는 과학교육과정, 교육목표 및 내용, 학생 등에 대한 이해가 수반되어야 하므로 과학교육과정, 과학 내용, 학생에 관한 지식 등이 요구된다. 평가 영역과 유형, 방법 선정 및 평가도구 제작 단계에서는 학생들의 과학 내용 지식과 과학적 탐구력 등의 인지적 특성이나 학습 동기, 호기심, 수업 참여도 등의 정의적 특성을 고려하여 평가 영역과 유형을 정한 후, 이를 토대로 평가 방법을 결정하고 적절한 평가도구를 개발해야 한다. 따라서 이 단계에서는 학생, 과학 교수 전략, 과학 내용에 관한 지식이 요구된다. 그리고 평가도구 제작의 전반적인 과정에서는 과학 평가에 관한 지식과 과학 평가에 관한 신념이 많은 영향을 미친다. 평가를 실행한 후 그 결과를 분석하고 피드백하는 단계에도 교사의 PCK가 큰 영향을 미칠 수 있다. 그러

므로 PCK는 교사의 수행평가에 관련된 문제점을 진단하고 개선하기 위한 관점을 제공할 수 있다. 실례로 교사의 PCK와 형성평가 실행의 관계를 조사한 연구(Falk, 2012)에 따르면, 교사들은 형성평가 과정에서 과학교육과정, 과학 교수 전략, 학생에 관한 지식을 활용하였으며, 동시에 형성평가의 경험이 교사들의 PCK를 계발할 기회를 제공하는 것으로 나타났다. 이는 교사들의 과학 수행평가 계획 및 실행 과정에 PCK에 관한 적절한 도움과 훈련을 제공하면 그들의 수행평가 전문성 향상에 기여할 수 있음을 의미한다(그림 1).



PCK 구성요소들의 종합적인 활용

<그림 1> 과학교사의 수행평가 전문성 향상과 PCK 계발의 관계

지금까지 개발된 과학교사의 PCK 계발을 위한 전략들을 살펴보면, 대체로 PCK에 관한 충분한 지식을 갖춘 경력교사나 교사 교육자, 전문가 등의 외부의 전문가가 코칭이나 멘토링, 컨설팅 등을 통하여 예비 및 현직교사들에게 PCK의 모든 요소에 대한 포괄적인 도움을 직접 제공하는 방식의 접근이 주를 이루고 있다(고문숙 등, 2009; 남정희 등, 2010; 노태희, 강석주, 강훈식, 2012; 노태희 등, 2012; 양찬호, 2015; 정금순, 강훈식, 2011; 한재영, 윤지현, 2009; Marable & Raimondi, 2007; Roth & Tobin, 2005; Stanulis, Little, & Wibbens, 2012; Tobin, Roth, & Zimmermann, 2001). 이러한 전략은 과학교사의 PCK 계발에 도움이 되지만 시간과

인력 등의 여건 문제로 우리나라의 교사 교육에 일반적으로 적용되기에는 현실적인 한계가 있다. 따라서 이런 전략들과 함께 실천적 맥락에서의 다양한 접근이 요구된다.

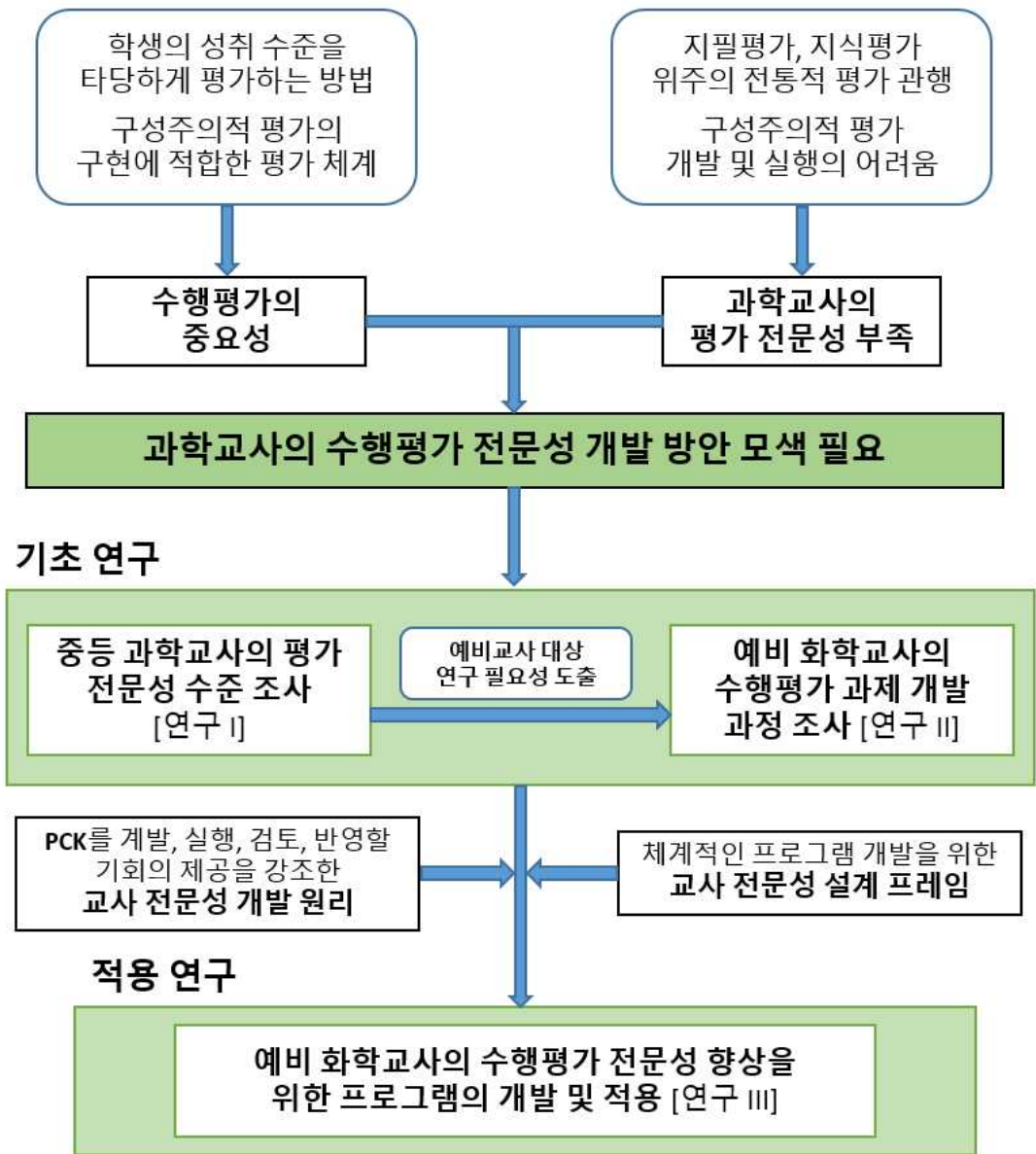
이를 위한 방법의 하나로, 교사들이 스스로 자신의 PCK를 효과적으로 활용하고 계발할 수 있는 평가 관련 활동의 기회와 도움을 제공함으로써 평가 전문성 향상을 도모하는 접근 방법이 유용할 수 있다. 예를 들어, 교육과정 자료의 분석과 수업 구성 과정에는 교사가 PCK 구성요소를 종합적으로 활용하는 것이 중요하다는 관점을 바탕으로 교사들의 PCK와 교수 설계 능력 사이의 관계를 규명하려는 연구가 있었으며(Brown, 2009; Forbes & Davis, 2008; Schwarz *et al.*, 2008), 교수 설계 관련 활동을 통한 PCK 계발의 가능성도 일부 확인하였다(Beyer & Davis, 2012). 즉, PCK 구성요소들이 서로 밀접한 관련이 있다는 점에 기초하여, PCK의 구성요소 중 하나인 교수 전략에 관한 지식 측면에서의 활동 경험과 자극을 제공함으로써 교사들이 자신의 PCK를 계발하는 방안의 유용성을 확인할 수 있었다. 하지만 과학 교사의 수행평가 전문성 향상 전략을 PCK 계발 관점에서 체계적으로 접근한 연구는 매우 부족하며, 특히 예비교사를 대상으로 한 연구는 거의 없다.

이에 이 연구에서는 예비교사에게 PCK를 활용하고 계발 및 검토할 수 있는 기회를 제공하는 예비교사의 수행평가 전문성 향상을 위한 프로그램을 개발하여 적용하였다. 이때, Loucks-Horsley 등(2009)은 과학교사의 전문성을 효과적으로 향상할 수 있는 프로그램을 개발하기 위해서는 교사가 PCK 및 기술을 계발하고 비판적으로 검토 및 반영할 기회를 제공해야 한다는 교사 전문성 개발 원리를 제안하였다. 이 원리는 구성주의적 관점에서 효과적인 교수학습에 대한 정립된 이미지 제공, 학생의 학습 목표 및 요구 충족, 동료 및 다른 전문가들과 함께 하는 기회 제공, 성인 학습자로서 교사의 능동적 참여 등의 요소를 포함하고 있다. 또한, 그들은 이러한 원리가 프로그램에 체계적으로 반영될 수 있도록 비전 및 기준 설정, 학생 학습 및 기타 자료 분석, 목표 설정, 계획, 실행, 결과 평가의 여섯 단계로 구성된 전문성 개발 설계 프레임틀을 제안하였다. 이 프레임에서 연구자는 자신의 지식과 신념을 바탕으로 프로그램의 비전 및 기준을 설정하며, 맥락을 고려하여 학습 및 기타 자료를 분석한다. 또한, 자료 분석을 통해 도출된 핵심적 문제를 바탕으로 전문성 향상 프로그램의 전체 목표가 설정되며, 프로그램의 계획 및 실행 단계에서는

목표를 달성하기 위한 구체적인 전략이 투입된다. 마지막으로 프로그램 적용 후 분석 결과를 바탕으로 반성 및 수정 과정을 거친다. 따라서 이와 같은 교사 전문성 개발 원리와 전문성 개발 설계 프레임에 기초한다면, 양질의 수행평가 전문성 향상 프로그램을 체계적으로 개발할 수 있을 것으로 기대된다.

이상의 논의를 종합하여 이 연구에서는 먼저 과학교사의 수행평가 전문성 향상 프로그램을 개발하기 위한 기초 연구로서, 현직 중등 과학교사들의 평가 전문성 수준을 조사하였다. 이를 통해 과학교사의 평가 전문성의 특징과 부족한 점을 구체화하였고, 예비교사 단계에서의 교육의 필요성을 도출하였다(연구 I). 이에 중등 예비 화학교사를 대상으로 수행평가 개발 과정에서 나타나는 특징을 PCK 구성요소 및 구성요소 사이 연계의 측면에서 조사하였다(연구 II). 이러한 결과를 토대로 구체적인 전략과 활동을 선정하고, Loucks-Horsley 등(2009)의 교사 전문성 개발 원리와 전문성 개발 설계 프레임을 바탕으로 예비교사의 평가 전문성 향상 프로그램을 개발하였다. 그리고 이를 예비 화학교사에게 적용하여 프로그램이 예비교사의 수행평가 전문성에 미치는 영향을 PCK 구성요소 및 구성요소 사이 연계의 측면에서 조사하였다(연구 III).

연구 배경과 목적 및 내용에 대한 개요는 <그림 2>와 같다.



<그림 2> 연구 개요

2. 연구 내용

각 연구의 구체적인 내용과 연구 문제는 다음과 같다.

연구 I(노태희 등, 2015)에서는 과학교사의 평가 전문성 수준과 문제점을 탐색하기 위한 기초 연구로 중등 과학교사들이 실시하는 과학 평가 사례와 그들이 지향하는 평가 사례를 각각 평가 목적과 방법의 측면에서 조사하였다. 연구 대상은 중등 과학교사 92명으로, 그들이 학교에서 실시했던 평가 사례와 구성주의적 평가를 실시할 수 있는 가상의 전형적인 수행평가 상황에 대하여 지향하는 평가 사례를 조사하기 위한 설문지를 개발하였다. 교사의 성별, 교직 경력, 학교급, 최종 학위, 평가 관련 교육 경험의 유무, 구성주의적 과학 평가관 수준 등 교사의 배경 변인도 조사하였다. 연구 I의 구체적인 목표는 다음과 같다.

- 1) 중등 과학교사가 실시하는 평가 사례를 평가 방법과 목적의 측면에서 조사한다.
- 2) 전형적인 수행평가 상황에서 교사가 지향하는 평가를 평가 방법과 목적의 측면에서 조사하고, 실태 조사 결과와 비교한다.
- 3) 교사의 배경 변인에 따른 형성평가의 활용 정도를 조사한다.

연구 II(노태희 등, 2017)에서는 예비 화학교사 8명을 대상으로 간단한 수행평가 워크숍을 실시한 후, 구상 단계와 구체화 단계로 이루어진 개별 과제 개발 과정과 조별 논의 과정을 거쳐 수행평가 과제를 개발하도록 하였다. 활동 과정에 대한 녹음 및 녹화와 반구조화된 면담을 실시하였고, PCK의 관점에서 분석하였다. 연구 II의 구체적인 목표는 다음과 같다.

- 1) 중등 예비 화학교사에게 수행평가 과제를 개발하도록 한 후, 개별 과제 개발 과정에서 나타나는 특징을 PCK 구성요소 및 연계의 측면에서 심층적으로 조사한다.
- 2) 조별 논의 과정에서 나타나는 특징을 PCK 구성요소 및 연계의 측면에서 심층적으로 조사한다.

연구 III(이재원 등, 2018)에서는 연구 I과 II의 결과를 종합하여 Loucks-Horsley 등(2009)의 교사 전문성 개발 원리와 전문성 개발 설계 프레임틀을 바탕으로 예비교사의 수행평가 전문성 향상을 위한 프로그램을 개발하였다. 8명의 예비교사가 3주 동안 프로그램에 참여한 후 연구 II와 같은 방법으로 수행평가 과제를 개발하였다. 예비교사의 수행평가 과제 개발 과정을 PCK 구성요소와 구성요소 사이의 연계 측면에서 조사하기 위해 발성 사고법을 사용하였고, 활동을 녹음 및 녹화하였으며 반구조화된 면담을 통하여 자료를 수집하였다. 연구 III의 구체적인 목표는 다음과 같다.

- 1) 연구 I과 II의 결과를 종합하여 교사 전문성 개발 원리와 전문성 개발 설계 프레임틀을 바탕으로 예비 화학교사의 수행평가 전문성 향상을 위한 프로그램을 개발한다.
- 2) 개발한 프로그램에 참여한 중등 예비 화학교사에게 수행평가 과제를 개발하도록 한 후, 개별 과제 개발 과정에서 나타나는 특징을 PCK 구성요소 및 연계의 측면에서 심층적으로 조사하고, 프로그램이 미친 영향을 분석한다.
- 3) 조별 논의 과정에서 나타나는 특징을 PCK 구성요소 및 연계의 측면에서 심층적으로 조사하고, 프로그램이 미친 영향을 분석한다.

3. 연구의 제한점

이 연구는 다음과 같은 제한점을 지닌다.

첫째, 연구 I에서는 연구 참여자로 다수의 중등 과학교사를 표집하였으나, 주로 수도권 지역에 근무하는 교사가 연구에 참여하였다. 또한, 교사의 교직 경력, 최종 학위, 평가 관련 교육 경험 등 연구 참여자의 배경이 모집단의 비율과 차이가 있다. 따라서 이 연구 결과를 우리나라 전체의 중등 과학교사에 대한 평가 전문성의 수준으로 일반화하는 데 한계가 있다.

둘째, 연구 II와 III은 사례 연구로, 각각 중등 예비 화학교사 8명을 의도적 표집에 의해 선정하였다. 이들은 충북 지역의 한 사범대학교에 개설된 평가 관련 강좌의 수강생이었으므로, 성별과 나이 등 연구 참여자의 특성과 개인적 성향을 통제하는 데 한계가 있었다. 따라서 이 연구의 결과를 중등 예비 화학교사 전체로 일반화하기 어렵다.

셋째, 연구 II와 III은 평가 관련 강좌의 학사일정에 따른 교수학습의 일부로 이루어졌으므로 사례 연구의 특성상 이러한 맥락을 벗어나면 연구 결과가 적용되지 않을 수 있다. 다만, 연구 참여자의 개인적 특성 및 맥락에 따른 영향을 최대한 통제할 수 있도록 사범대학교 한 곳에서 3학년에 재학 중인 화학교육과 학생 중, 같은 교수가 개설한 평가 관련 강좌를 수강중인 예비교사를 선정하여 1년 간격으로 연구 II와 III을 실시하였다.

넷째, 연구 II와 III에서 개발한 수행평가 과제는 모두 2009 개정 교육과정에 따른 중학교 1학년 과학 교과서의 ‘분자 운동과 상태 변화’ 단원을 대상으로 하였다. 따라서 예비교사가 개발한 수행평가 과제에 학습 내용이나 단원 특성이 반영되었을 수 있다. 그러므로 이 연구의 결과를 다른 과학 교과나 단원에 일반화하여 해석하는 데 한계가 있다. 다만, 연구 II와 III의 결과를 타당하게 비교할 수 있도록 같은 단원을 대상으로 동일한 조건에서 수행평가 과제를 개발하도록 하였다.

다섯째, 연구 III은 수행평가 전문성 향상을 위한 프로그램 3주와 자료 수집을 위한 수행평가 과제 개발 1주를 포함, 총 4주라는 비교적 짧은 시간 동안 적용되었다. 따라서 이 기간에 예비교사들이 평가 전문성 향상을 위한 프로그램의 내용을 완전히 이해하고 내면화하는 데 한계가 있을 수 있다. 또한, 프로그램의 효과는 8

명을 대상으로 나타나는 PCK 구성요소의 특징을 양적, 질적으로 분석하여 기술하였기 때문에 연구 III에서 나타난 특징을 프로그램의 효과로 해석함에 주의가 필요하다.

4. 용어의 정의

이 연구에서 사용되는 주요 용어들의 정의는 다음과 같다.

(1) 교과교육학 지식 (pedagogical content knowledge; PCK)

PCK는 일반 교육학 지식(pedagogical knowledge; PK)과 교과 내용 지식(content knowledge; CK)의 결합체로서, 교사의 전문성을 나타내는 대표적인 지표이다. PCK의 주제 특이성(topic specificity; van Driel, Verloop, De Vos, 1998)을 고려하여, 이 연구에서의 PCK는 구성주의적 수행평가 과제 개발의 맥락에서 과학교사가 갖추어야 할 전문적 지식을 의미한다. 구체적인 PCK의 구성요소는 Magnusson과 Krajcik, Borko(1999)의 PCK 구성요소에 교과 내용 지식을 추가하여 과학 평가에 관한 지식, 과학교육과정에 관한 지식, 과학 교수 전략에 관한 지식, 과학 내용에 관한 지식, 학생에 관한 지식, 과학 평가에 관한 신념의 6가지로 정의하였다.

(2) 구성주의적 수행평가 (constructivist performance assessment)

수행평가는 학생 스스로가 자신의 지식이나 기능이나 태도를 나타낼 수 있는 답을 구성하거나, 산출물을 만들거나, 행동을 나타내도록 요구하는 평가 방식을 의미한다(백순근, 1998). 이때 평가의 전통적 측면을 고려하되 구성주의적 평가의 특징인 지식 구성의 과정에 대한 평가, 단순 암기를 지양하는 고등 사고능력에 대한 평가, 다양한 관점과 양식이 존중되는 평가 등의 특징(박선미, 1998)이 구현되었거나 이를 지향하는 수행평가를 구성주의적 수행평가로 정의한다.

(3) 수행평가 전문성 (performance assessment professionalism)

이 연구에서는 과학교사의 수행평가 전문성을 구성주의적 수행평가에 관한 전문성으로 정의하였다. 즉, ‘수행평가의 구성주의적 특징과 지향점에 대한 이해를 바탕으로 이를 고려한 수행평가 과제를 개발하고, 학생의 학습과 성취에 관한 정보

를 체계적으로 수집, 해석, 활용할 수 있는 능력' 을 의미한다. 수행평가 전문성의 구인으로는 예비교사의 맥락을 고려하여 평가 방법의 선정, 평가도구의 개발, 평가 실시, 평가 결과의 분석 및 해석, 평가 결과의 활용을 선정하였다(김경희 등, 2006).

(4) 교사 전문성 개발 원리 (principles of effective professional development)

Loucks-Horsley 등(2009)이 제안한 과학교사의 전문성 개발 원리 8가지 중에서 예비교사의 맥락에 적용할 수 있는 5가지 원리를 의미한다. 즉, 효과적인 교수학습에 대한 정립된 이미지의 제공, 학생의 학습 목표 및 요구의 충족, 교사가 PCK 및 기술을 개발하고 비판적으로 검토 및 반영할 기회의 제공, 동료 및 다른 전문가들과 함께하는 기회의 제공, 성인 학습자로서 예비교사의 능동적 참여 촉진을 의미한다.

제 2 장. 이론적 배경 및 선행연구

1. 과학교사의 PCK와 평가 전문성

1.1 PCK

특정 직업군의 전문성은 해당 직업 종사자만이 가지고 있는 전문적 지식에 기반을 둔다. 예를 들어, 의사는 의학적 지식을 통해 의료계 전문가로, 판사는 법학적 지식을 통해 법조계 전문가로 인정을 받는다. 그렇다면 교사를 교육계의 전문가라고 할 때, 교사의 전문성은 어떤 전문적 지식에 근거한다고 할 수 있는가? 이러한 질문과 함께 과학교육계에서도 과학교사의 전문성에 관한 연구가 활발히 진행되었다.

(1) PCK의 개념

1980년대 이전에는 과학교사의 전문성에 대하여 교과 내용 지식과 교육학적 지식으로 나누어 접근하였다. 하지만 Shulman(1986, 1987)은 교사의 전문적 지식에 관한 논의가 교과 내용 지식 또는 교육학적 지식의 극단에 치우쳐 있음을 지적하며, 두 지식의 교집합에 교사의 전문적 지식이 존재한다고 주장하였다. 이를 개념화한 PCK는 일반 교육학 지식(pedagogical knowledge; PK)과 교과 내용 지식(content knowledge; CK)의 결합체로 정의된다. 과학자와 교육학자, 교사를 서로 비교하면 PCK의 의미를 쉽게 이해할 수 있는데, 과학자는 과학 지식(CK)의 전문가이지만 교육적 상황에서 학생들에게 과학 지식을 효과적으로 이해시킬 수 있는 방법에 관한 지식은 부족하다고 할 수 있다. 반면 교육학자는 교육학적 지식(PK)의 전문가이지만 구체적인 과학 지식에 대한 이해가 부족할 수 있다. 하지만 교사는 자신의 전공 교과에 대한 구체적인 과학 내용 지식과 교육학적 지식을 결합하여 학생들이 과학 내용을 효과적으로 이해할 수 있도록 하는 자료나 교수 전략을 개발하고 수업에서 이를 구현할 수 있다. 그러므로 교사는 과학자 또는 교육학자와 차별화된 전문성을

지닌다고 할 수 있으며, 이러한 점에서 Carter(1990)와 Gedds 등(1993)은 교사의 PCK를 ‘구체적인 과학 내용을 학생들이 이해할 수 있는 형태로 변형하는 데 필요한 지식’으로 정의한 바 있다. 또한, van Driel과 Verloop, De Vos(1998)는 PCK를 ‘특정 주제에 관하여 교사가 발달시킨 지식’으로 정의하고, 한 교과 영역의 PCK는 다른 교과로 쉽게 전이되지 않는다고 하였다. 같은 맥락에서 Loughran 등(2001)도 PCK를 ‘특정 과학 내용에 관한 학생의 이해를 돕기 위해 사용되는 지식’으로 정의하였다. 이는 PCK에 주제 특이성(topic specificity) 즉, 교사의 PCK가 특정 교과 또는 주제에만 한정되는 특징이 있음을 나타낸다. PCK의 주제 특이성은 교사가 자신의 전공 교과가 아닌 다른 교과를 가르쳐야 하는 상황에서 겪을 수 있는 어려움을 의미한다고 할 수 있다.

또한, 연구자들은 PCK의 본성이나 특징과 관련하여 자신들이 강조하고자 하는 점을 반영한 다양한 용어들을 제안하였다. 예를 들어, Cochran과 DeRuiter, King(1993)은 구성주의적 입장에서 PCK가 특정한 상황과 맥락 속에서 구성원들 간의 상호작용에 따라 계발되고 변형될 수 있으므로 교사가 PCK를 알고(knowing) 이해하는 것이 중요하다고 강조하며 PCKg(pedagogical content knowing)라는 용어를 사용하였다. Barnett과 Hodson(2001)은 PCK 형성에 사회적 맥락(context)의 기여를 강조하며 PCxK(pedagogical context knowledge)라는 용어를 제안하였다. 앞서 논의한 PCK의 주제 특이성을 강조한 용어도 제안되었는데, Hashweh(2005)가 제안한 TPCs(teacher pedagogical constructions)는 구체적인 교수학습 주제에 관한 PCK를 의미하며, 이는 해당 주제에 대한 지속적인 교수학습 및 반성의 경험으로부터 발달될 수 있다고 하였다. Mishra와 Koehler(2006)는 교육현장에 교과서나 칠판과 같은 전통적인 도구인 뿐 아니라 컴퓨터와 인터넷 등 새로운 도구의 활용이 빠르게 확산됨에 따라 이에 대한 교사의 이해와 교육적 활용 능력이 필요하다고 강조하며 기존의 PCK와 기술적 지식(technological knowledge; TK)을 결합한 TPCCK(technological pedagogical content knowledge) 개념을 제안하였다. 이와 유사한 방식으로 Ab Kadir(2017)는 비판적 사고의 개념, 기술 및 과정에 대한 이해 및 학생들의 비판적 사고능력을 함양하는 교육방법 등에 관한 교사의 비판적 사고 지식(critical thinking knowledge; CTK)과 PCK를 결합하여 CTPCK(critical thinking pedagogical content knowledge)를 제안하였다. 이와 같이 많은 연구자들은 PCK로

부터 파생된 다양한 용어를 꾸준히 제안하면서 PCK의 개념을 재정립하고 있다. 하지만 본질적인 의미는 Shulman(1986)이 처음 제안한 일반 교육학 지식과 교과 내용 지식의 결합체라는 큰 틀에서는 벗어나지는 않는다고 볼 수 있다.

이와 같은 노력으로 PCK는 현재까지 국내외 과학교육계에서 교사의 전문성을 나타내는 대표적인 지표로써 널리 활용되고 있다(Alonzo & Aschbacher, 2005; Bell & Cowie, 2001; Black *et al.*, 2004). 예를 들어, 맹희주와 손연아(2011)는 과학 수업에서 통합적 적용 경험에 따른 초등 교사들의 PCK 차이를 분석하였으며, 곽영순(2008)은 과학교사의 PCK 유형별 교사 전문성의 특징을 조사하였다. 또한, 고미례와 남정희, 임재항(2009), 곽영순(2009)은 각각 초임 과학교사의 PCK를 분석하였다. 이 외에도 다양한 상황에서 교사의 PCK를 분석하기 위한 연구들이 꾸준히 이루어지고 있다(박성혜, 2003a, 2003b, 2006; 임청환, 2003; 조희형, 고영자, 2008; Keller, Neumann, & Fischer, 2017; Loughran, Mulhall, & Berry, 2004; Segall, 2004).

또한, 과학교사의 PCK를 계발함으로써 그들의 수업 전문성을 향상하는 연구도 활발하게 이루어지고 있다. 예를 들어, 고문숙 등(2009), 남정희 등(2010), 노태희와 강석주, 강훈식(2012), Marable과 Raimondi(2007)는 각각 멘토가 직접 교사들에게 PCK 측면에서의 도움을 제공하는 멘토링 전략을 활용하였다. 노태희 등(2012), 정금순과 강훈식(2011), 한재영과 윤지현(2009), Roth와 Tobin(2005)은 코칭 전략을 활용하였다. 양찬호(2015)는 멘토링과 코칭을 상호보완적으로 결합한 전략을 활용함으로써 초임 과학영재 교사의 수업 전문성 향상을 도모하기도 하였다. 이처럼 과학교사의 PCK 계발은 외부 전문가가 교사에게 PCK와 관련한 도움을 포괄적으로 제공하는 활동을 중심으로 하는 경우가 많다.

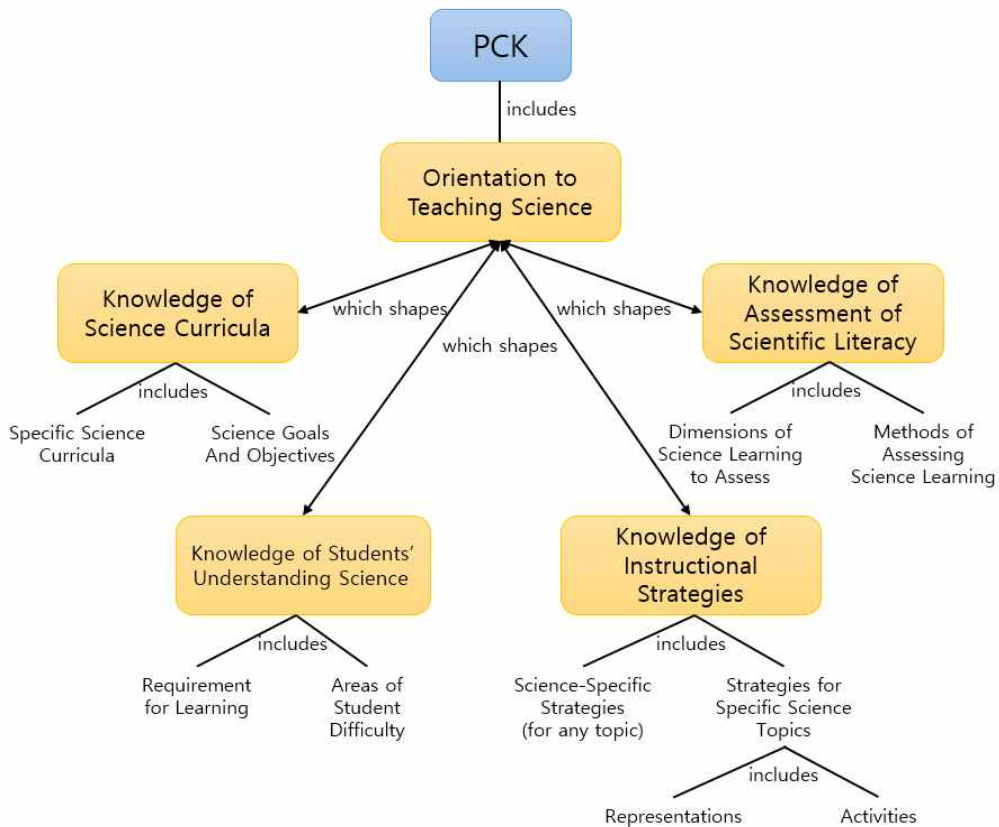
교사가 반성적 활동을 하면서 스스로 PCK를 계발할 기회를 제공하는 방식도 일부 이루어지고 있다. 예를 들어, 장효순과 최병순(2010)은 교사가 특정 영역이나 주제에 대해 서술적으로 기술하는 프레임인 CoRe(content representation) 개발 전략을 활용하였고, 박재성과 강훈식, 한재영(2017)은 해결자·청취자 활동 전략을 활용한 지필평가 문항 제작 활동을 통하여 예비 화학교사의 PCK를 계발하고자 하였다. 이 전략은 발성 사고법에 기반을 두고 있으며 해결자는 과제를 해결하는 과정에서 청취자의 점검과 도움을 받으면서, 동시에 청취자는 해결자의 과제 해결 과정을 점검해주면서 반성적 사고를 할 수 있다.

(2) PCK의 구성요소

Shulman(1987)은 PCK의 구성요소로 학생에 관한 지식과 교수 전략에 관한 지식을 제안하였다. 그리고 PCK를 교과 내용 지식, 일반 교육학 지식, 교육과정에 관한 지식, 상황 지식, 가치에 관한 지식, 교수 목적에 관한 지식과 함께 교사의 지식을 구성하는 요소로 보았다. 이후 연구자들은 Shulman(1987)이 제시한 PCK의 구성요소를 분석하고, PCK의 본성과 영역, 구성요소에 관한 연구를 활발히 수행하였다. 예를 들어, Marks(1990)는 PCK의 구성요소에 교과 내용 지식도 포함되어야 한다고 주장하였고, Grossman(1990)은 PCK의 구성요소에 교육과정에 관한 지식과 교수 목적에 관한 지식을 포함하였다. Hashweh(2005)는 교과 내용에 관한 지식, 교수 전략에 관한 지식, 학생에 관한 지식, 교육과정에 관한 지식, 상황에 관한 지식, 평가에 관한 지식을 모두 PCK의 구성요소로 보았다.

과학 교과에서 이루어진 PCK 구성요소에 관한 연구는 Magnusson과 Krajcik, Borko(1999)가 대표적이다. 그들은 PCK의 구성요소에 먼저 교사의 과학 교수학습에 관한 지향을 포함하였다. 과학교사가 지향하는 교수학습의 유형에는 강의식, 개념 변화, 과정 중심, 학문 중심, 활동 중심, 탐구 중심, 구조화된 탐구 중심, 발견학습, 과제 기반 수업 등이 있으며 대부분의 교사들은 서로 모순되더라도 동시에 여러 가지의 지향을 한다고 보았다. 그리고 과학 교수학습에 관한 지향을 구체화하는 요소로 과학교육과정에 관한 지식, 학생에 관한 지식, 과학 교수 전략에 관한 지식, 과학 평가에 관한 지식을 제안하였다. 이때 각 PCK 구성요소별 하위 요소도 제시되었는데, 과학 평가에 관한 지식에서는 평가 방법과 평가 영역을 하위 요소로 하였다. 과학교육과정에 관한 지식의 하위 요소에는 특정 과학교육과정에 관한 지식과 교수학습의 목적과 목표에 관한 지식이 포함된다. 학생에 관한 지식의 하위 구성요소에는 학생의 교육 요구와 어려움이 있다. 과학 교수 전략에 관한 지식의 하위 구성요소는 주제와 무관하게 과학 교과 전반에서 활용할 수 있는 교과 전략(science specific strategies)과 특정 주제에서만 활용할 수 있는 주제 전략(strategies for specific science topics)으로 나누었다. 이 중 주제 전략은 다시 개념 또는 원리를 표현하는 예시, 모형 등의 활용 전략과 관련된 표상(representations)과

실험, 문제, 시뮬레이션 등의 활동(activities)으로 나눌 수 있다(그림 3).



<그림 3> Magnusson, Krajcik, & Borko(1999)의 PCK 구성요소

PCK의 개념을 처음 제안한 Shulman(1987)의 PCK 구성요소와 <그림 3>을 비교하면, Shulman(1987)은 과학교육과정에 관한 지식이 PCK와 함께 교사의 지식 체계를 이루는 것으로 보았지만, Magnusson과 Krajcik, Borko(1999)는 이를 PCK의 구성요소로 편입시켰다는 점과 평가에 관한 지식이 새롭게 고려되었다는 점에서 주요한 차이가 있다.

<표 1>에 앞서 논의한 바를 포함하여 주요 연구자들이 주장한 PCK 구성요소를 연도순으로 정리하였다(조희형 등, 2006; Park & Oliver, 2008).

표 1. 주요 연구자들의 PCK 구성요소

연구자	교수 목적	학생	교육 과정	교수 전략	매체, 자원	평가	교과 내용	맥락	교육 학
Shulman (1987)	D	O	D	O			D	D	D
Tamir (1988)		O	O	O		O	D		D
Smith & Neale(1989)	O	O		O			D		
Grossman (1990)	O	O	O	O			D		
Marks (1990)		O		O	O		O		
Cochran, DeRuiter, & King(1993)		O		N			O	O	O
Geddis <i>et al.</i> (1993)		O	O	O					
Fernandez- Balboa & Stiehl(1995)	O	O		O			O	O	
Magnusson, Krajcik, & Borko(1999)	O	O	O	O		O			
Carlsen (1999)	O	O	O	O					
Gess-Newso me(1999)							O	O	O
Morine-Der shimer & Kent(1999)		O	O				O	O	O
Hashweh (2005)	O	O	O	O		O	O	O	O
Lee(2005)	O	O	O	O	O	O			
Loughran, Berry, & Mulhall(2006)	O	O		O			O	O	O
Park & Oliver(2008)	O	O	O	O		O			

* O: PCK 구성요소, D: PCK와 함께 교사의 지식을 구성하는 요소, N: 명시적으로 언급하지 않음.

<표 1>을 살펴보면, PCK의 구성요소는 연구자의 관점과 강조하고자 하는 바에 따라 다양하지만, 대체로 학생에 관한 지식, 교수 전략에 관한 지식, 교육과정에 관한 지식 등을 포함하고 있다는 것을 알 수 있다. 또한, 대부분의 연구에서 PCK 구성요소는 각각 하위 구성요소를 포함하고 있는 경우가 많다. 따라서 하위 구성요소를 포함하여 PCK의 구성요소에 관한 정의는 통일되기 어렵다(이연숙, 2006).

PCK가 교사의 전문적 지식 체계를 분석하는 데 유용한 개념적 도구를 제공한다는 점은 부인하기 어렵지만, 일각에서는 교사의 지식 체계를 PCK 구성요소에 따라 분절적으로 나누어 분석하게 되면 오히려 교사가 가지는 지식 체계의 전체적이고 실제적인 모습을 파악하기는 어려워질 수 있다는 지적도 제기되었다(Avraamidou & Zembal-Saul, 2005; Borgia, 2001). 이에 많은 PCK 관련 연구에서는 이러한 단점을 보완하기 위하여 PCK 구성요소 사이의 상호작용을 고려하고, PCK 구성요소 사이의 연계 내지는 통합을 강조하는 추세이다(Park & Oliver, 2008).

1.2 평가 전문성

교육과정이 내실화를 갖추기 위해서는 교수학습과 평가가 모두 변화해야 한다. 새 교육과정에서 창의·인성 교육을 강조하고 이를 함양하기 위한 수업이 이루어지더라도 평가가 전통적인 관점에 머문다면 교육현장에서의 변화는 피상적인 수준에 그칠 수밖에 없기 때문이다. 따라서 교사의 평가 전문성은 곧 교수학습 활동의 결과 직접적으로 관련되어 있으며, 나아가 국가 수준의 교육과정의 성패로도 연결된다. 그러므로 교사의 평가 전문성을 개념화하고 이를 함양하기 위한 노력은 매우 중요하다.

(1) 학생평가의 개념과 목적

학생평가는 ‘교사가 학생의 학습과 성취, 자신의 수업에 관한 의사 결정을 하는 데 필요한 정보를 수집, 해석, 활용하는 활동’으로 정의된다(이인제 등, 2004; McMillan, 2007). 이때 교사가 필요로 하는 정보는 수업의 계획 단계에서부터 실행과 반성에 이르는 수업의 전 과정에서 수집, 해석, 활용될 수 있으며 교사의 의사

결정 또한 수업의 전 과정에서 이루어진다. 따라서 학생평가는 수업과 통합된 과정이며 분리할 수 없는 관계에 있다(김신영, 2004).

먼저 학생평가의 의미와 기능을 살펴보면, 성태제(2002)는 학생평가의 의미를 세 가지로 구분하였다. 첫째는 교육과정에서 설정한 목표가 얼마나 달성되었는지 판단하는 것으로, 이는 교육과정의 목표에 대한 학생의 성취수준을 판단하는 전통적인 관점이다. 학교에서 점수화를 목적으로 시행되는 지필평가가 일반적으로 이러한 기능을 수행한다고 할 수 있다. 둘째는 학생, 교사, 학부모 등 교수학습과 관련된 이들에게 의사 결정에 필요한 정보를 제공하는 것이다. 김신영(2002, 2007)도 교사가 수집한 학생의 학습과 성취에 관한 정보를 관련 대상자에게 제공하여 그들의 의사 결정과 교육적 노력을 돕는 데 학생평가의 의미가 있다고 보았다. 셋째는 교육의 목적에 대한 가치 판단에 필요한 정보를 제공하는 것이다. 이 관점에서는 학생평가의 기능이 성취수준 판단, 의사 결정에 필요한 정보 제공과 같이 도구적, 기능적 측면에 머물지 않고 교육 자체의 목적에 대한 가치 판단에도 평가의 기능이 있음을 강조한다.

Popham(2009)은 교사의 학생평가를 교실 평가와 책무성 평가로 구분하였다. 교실 평가는 교사가 학생의 학습과 성취에 관한 정보를 얻기 위한 평가와 학생, 교사, 학부모 등 관련 대상자들의 의사 결정에 필요한 정보를 수집하고 제공하기 위한 평가를 의미한다. 책무성 평가는 전국 또는 시·도 단위의 학업 성취도 평가와 같이 지역 또는 국가 단위에서 각 학교의 교육적 노력 및 교육과정 목표의 달성 정도 파악을 위한 정보를 수집하는 평가이다.

Davis와 Neitzel(2011)은 수집한 평가 정보를 활용하는 주체에 따라 학생평가의 기능을 나누었다. 먼저 교사를 위한 평가의 기능에는 학생들의 교육적 활동에 관한 정보를 얻고 측정하는 것, 추가 지원이 필요한 학생을 식별하는 것, 학업 성적을 산출하는 것, 학생을 선발하는 것, 학생의 개념 이해의 깊이와 폭을 측정하는 것, 교실 활동의 참여도를 측정하는 것이 있다고 보았다. 다음으로 학생을 위한 평가의 기능에는 학생의 학업 노력이나 시간 투자에 대한 보상이나 처벌, 학생의 학습 과정이나 결과에 관하여 성적표 등을 통해 규범적 정보를 제공하는 평가 피드백(evaluation feedback) 또는 학습을 향상할 수 있는 전략이나 조언을 제공하는 기술 피드백(descriptive feedback)이 있다고 하였다. 마지막으로 외부인(external

audience)을 위한 평가가 있다. 외부인에는 학부모, 교장 또는 지역 또는 국가 단위 상급 교육기관의 인사가 해당하며, 이들에게 학생들의 학습 수준에 관한 정보나 교육과정 또는 정책의 효과에 관한 정보를 제공하는 것이다. 또한, 교사가 학생들에게 수능 기출 문제를 활용한 평가를 하는 것과 같이 특정 시험 형식과 내용에 친숙해지기 위한 평가를 실시하는 것도 외부인을 위한 평가 범주에 포함된다.

하지만 일반적으로 교사가 수집한 학생평가 정보는 앞서 논의한 다양한 기능을 동시에 수행할 수 있다. 그러므로 학생평가에 있어 평가의 기능보다 중요하게 여겨지는 것은 평가 정보를 수집, 해석, 활용하는 교사의 평가 목적이다. 교사의 평가 목적은 크게 교수학습이나 교육과정 목표의 달성 및 성취 여부를 판정하고, 수업의 교육적 효과를 알아보는 총괄평가(summative assessment)와 학생에게 피드백을 제공하고 교수학습이나 교육과정을 개선하기 위한 목적으로 실시되는 형성평가(formative assessment)로 구분된다(Hanna & Dettmer, 2004). 교사가 학생들로부터 같은 정보를 수집한 경우에도 상급 학교로의 진학과 성적 산출 등에 필요한 정보를 제공하는 의미가 강조되면 총괄평가, 학생의 적성과 진로선택 등에 필요한 정보를 제공하는 의미가 강조되면 형성평가로 볼 수 있다.

Lee와 Lin, Tsai(2013)는 고등학생이 인식하는 평가의 목적을 조사한 후, 평가의 목적을 총괄평가, 형성평가와 더불어 피상적 평가(surface assessment)의 세 가지로 나누었다. 총괄평가는 책무성 평가를 포함하여 학습 후 등급을 결정하기 위한 평가이며, 형성평가는 학습에 도움이 되는 비판적 판단, 문제 해결 등의 역할을 하는 평가이다. 피상적 평가는 파편화된 지식이나 용어들을 기계적인 반복을 통해 암기하는 능력에 관한 매우 낮은 수준의 평가를 의미한다.

James(2006)는 평가 목적을 행동주의적 평가, 구성주의적 평가, 사회문화적 평가로 분류하였다. 행동주의적(behaviorist) 평가는 학습 결과의 측정 및 교육과정의 차별화에 초점을 두며, 역사적으로는 직업 선발, 공공질서의 통제 및 유지에 기원을 두고 있는 평가 목적으로 총괄평가에 대응하는 개념이다. 행동주의적 평가 과제의 조건은 평가의 타당도와 신뢰도를 중시하며 환경을 통제하고, 교수학습이 끝난 후 평가를 시행하는 것이다. 구성주의적(constructivist) 평가 목적은 학생의 학습에 도움을 주기 위한 평가, 결과뿐 아니라 학습 과정에도 초점을 두는 평가, 학생이 갖는 오개념과 이해의 질에 대한 평가로서 형성평가에 대응하는 개념이다. 구성주의

적 평가 과제의 조건은 실제적 상황에서의 문제 해결, 개념을 중심으로 한 평가, 학습 과정에서 언제든 일어날 수 있는 평가이다. 사회문화적(sociocultural) 평가 목적은 구성주의적 평가와 마찬가지로 형성평가에 대응하지만, 구성주의적 평가가 학생 개인만을 고려한다면 사회문화적 평가는 개인이 속한 사회적, 문화적 맥락을 함께 고려한다는 점에서 구별된다. 사회문화적 평가 목적의 특징은 학생 스스로 능동적이고 독립적인 학습자임을 보여줄 기회를 제공하는 평가, 학생의 학습 전략과 과정을 살피고 강화하기 위한 최선의 도구와 도움을 제공하는 평가, 근접 발달 영역의 관점에서 학생에게 도움이 필요한 정도를 측정하는 평가이다. 평가 과제의 조건으로는 실제적 상호작용이 이루어지는 맥락 속에서의 평가, 역동적인 평가, 전문성 향상의 지표로서 언어적, 행동적 변화가 이루어지는 평가를 제시하였다.

이와 같이 평가 목적에 관한 용어와 유형은 연구자에 따라 조금씩 다르지만, 큰 틀에서는 전통적 평가 관점에 기반한 총괄평가와 구성주의적 평가 관점에 기반한 형성평가로 구분된다는 것을 알 수 있다. 이때 Earl(2012)은 총괄평가가 학습의 과정보다는 결과를 강조한다는 측면에서 학습에 관한 평가(assessment of learning)라 할 수 있고, 형성평가는 학습 과정을 중시하는 학습을 위한 평가(assessment for learning)라 할 수 있으며, 최근에는 학습 활동이 곧 평가 활동이 되는 학습으로서의 평가(assessment as learning)의 관점도 나타나고 있다고 하였다.

형성평가의 특징은 구성주의적 평가의 특징인 학생의 개인차를 존중하는 평가, 고등 사고능력에 관한 평가, 학생들이 개념을 더 잘 이해할 수 있도록 돕는 평가와 일맥상통한다(James, 2006). 최근의 교육 패러다임으로 자리 잡은 구성주의 학습관은 학습이 교사와 학생, 학생과 다른 학생들과의 상호작용을 통해 구성된다고 보며, 교사의 역할도 지식의 전달자에서 학생의 지식 구성을 돕는 조력자로 변해왔다. 이에 평가의 패러다임도 양적평가에서 질적평가로, 총괄평가에서 형성평가로 변화하고 있다(민희정, 2012). 따라서 이와 같은 평가 패러다임 변화에 부합하는 교사의 평가 전문성 향상은 필수적이라 할 수 있다.

(2) 수행평가의 정의와 특징

먼저 수행평가의 정의를 살펴보면, 교육부(1998)에서는 수행평가를 ‘교사가 학생

들의 학습 과제의 수행 과정 및 결과를 직접 관찰하고, 그 관찰 결과를 전문적으로 판단하는 평가 방식'으로 정의하였다. 여기서 학습 과제란 성취기준이나 교육목표를 통해 구체화할 수 있는, 가능한 실제 생활과 관련하여 유의미하고 중요하며 유용한 과제를 의미한다. 수행이란 학생이 정해진 답을 선택하는 것이 아니라 스스로 답을 구성하는 것을 의미하며, 때로는 산출물을 만들어 내거나 태도나 가치관을 행동으로 드러내는 것을 모두 포함한다. 마지막으로 관찰이란 학생의 수행 과정과 결과에 관하여 교사가 읽고, 보고, 듣고, 느끼는 모든 활동을 포함한다. 따라서 넓은 의미에서의 수행평가는 선택형 또는 단답형 지필평가 이외의 다른 모든 방법으로 평가하는 방식으로 정의될 수 있다.

연구자들이 정의하는 수행평가는 그들이 강조하는 바에 따라 조금씩 차이가 있다. 예를 들어, 남명호(1995)는 수행평가를 '학습 과정, 결과물, 수행을 모두 포함하는 넓은 평가 개념'으로 정의하였고, 백순근(1998)은 '학생 스스로 자신의 지식이나 기능이나 태도를 나타낼 수 있는 답을 구성하거나, 산출물을 만들거나, 행동을 나타내도록 요구하는 평가 방식'으로 정의하였다. Brown과 Shavelson(1996)은 실제적인 상황과 자료가 제시된 '수행 과제', 구체적인 활동을 이끌어낼 수 있는 '학생용 활동지', 과정별로 구체적인 '채점 체계'의 세 요소를 갖춘 평가를 수행평가로 정의하였다. 이러한 정의는 수행평가의 특징 중에서 수행 측면을 특히 강조한 것으로 볼 수 있다.

김은진(2000)은 수행평가를 '과학적 지식의 적용, 과학적 탐구 과정, 창의력, 과학적 태도, 의사소통능력 등 학생의 다양한 능력을 총체적으로 평가하는 방법'으로 정의하고, 수행평가 과제를 '과학적 지식을 활용하여 실생활에서 부딪힐 수 있는 실제적인 문제를 과학적 사고력과 탐구 과정을 거쳐 해결하도록 하는 평가 과제'로 정의하였다. Oberg(2010)는 수행평가를 수업과 평가를 연결하는 궁극적인 평가 형태라고 주장하면서 '학생의 발달과정, 기술, 성취를 측정하기 위한 한 가지 이상의 접근 방식'으로 폭넓게 정의하였다. 또한, 좋은 수행평가 과제는 교실 환경에서 설정된 인위적인 문제가 아니라 실제 상황에서의 이해와 기술을 보여줄 수 있는 과제가 되어야 한다고 하였다. Koh, Tan과 Ng(2011)는 수행평가와 전통적 평가를 비교하며 '전통적 평가보다 더 깊은 수준에서 학생들의 지식과 기술을 측정하는 평가'로 정의하였다. 이러한 정의는 수행평가에 활용할 수 있는 평가 방법

또는 평가 영역의 다양성을 강조한 정의라 할 수 있다.

한편, Mehrens(1992)는 수행평가를 ‘학생들의 반응을 평가하는 과정에서 교사의 관찰과 전문적인 판단을 특히 중시하는 평가 방식’으로 정의하였고, 박도순(1995)은 ‘교사가 학생 개개인이 처한 상황을 고려하여 과제 수행 과정과 결과에 대한 포괄적인 정보를 수집하는 평가 방식’으로 정의하였다. 이러한 정의에서는 수행평가를 실시하는 교사의 관찰과 판단을 중요시한다.

또한, 연구자들은 자신들이 강조하고자 하는 점을 반영하여 수행평가와 유사하거나 같은 의미를 지닌 여러 용어를 제안하거나 수행평가와의 관련성을 논의하였다. 예를 들어, Kan과 Bulut(2014)은 수행평가가 학생들이 무엇을 알고 할 수 있는지 측정하는 전통적 평가의 대안적 평가(alternative assessment)라 하였다. Adeyemi(2015)는 포트폴리오 평가에 관한 실증적 연구를 통해 진정한 수행평가를 구현하는 한 가지 형태로써 수행평가와 포트폴리오 평가의 관계를 논의하였다. 이처럼 수행평가와 유사하거나 같은 의미를 지닌 용어의 정의 및 수행평가와의 관련성을 <표 2>에 정리하였다(교육부, 1998).

<표 2> 수행평가와 유사한 용어의 정의와 수행평가와의 관련성

용어	정의	수행평가와의 관련성
대안적 평가 alternative assessment	<ul style="list-style-type: none"> - 한 시대의 주류를 이루는 평가체제와 성질을 달리하는 평가체제 - 선택형 문항을 사용하는 표준화된 검사의 대안적인 평가 	수행평가는 대안적 평가의 한 사례
포트폴리오 평가 portfolio	<ul style="list-style-type: none"> - 시험이 아니라 학생이 쓰거나 만든 작품집이나 서류철 등을 이용한 평가 - 결과가 나오게 된 과정 및 변화에 대한 평가를 중시함 	수행평가의 대표적인 방식
진정한 평가 authentic assessment	<ul style="list-style-type: none"> - 평가 상황이 실제 상황과 동일한 상태에서 이루어지는 평가 	수행평가의 특수한 사례
직접평가 direct assessment	<ul style="list-style-type: none"> - 간접적인 방법보다는 직접적인 방법을 중시하는 평가 - 예를 들어, 도덕성을 지필평가나 구두시험이 아닌 관찰을 통해 평가함 	수행평가는 가능한 직접평가의 성격을 포함하려 함
실기평가 practical skill assessment	<ul style="list-style-type: none"> - 아는 것(지식)이 아니라 할 수 있는 것(기능)을 재는 평가 	수행평가의 유형 중 하나
과정 (중심) 평가 process (centered) assessment	<ul style="list-style-type: none"> - 학습의 결과가 아니라 학습의 과정을 평가의 주요 대상으로 설정하는 평가 	수행평가의 중요한 측면 중 하나.

우리나라에 수행평가가 정책적으로 도입된 시기는 1998년으로, ‘교육비전 2002: 새 학교 문화 창조 방안’ 공고를 통하여 국가 수준의 교육정책으로 채택되었다(교육부, 1998). 일선 학교에는 1996년 서울시교육청 소속 초등학교에 처음으로 수행평가가 시범 적용된 이후, 1997년부터 1999년까지 전국의 초, 중, 고등학교에 확대 적용되었다. 교육부(1998)가 밝히는 수행평가의 도입 배경을 정리하면 다음과 같다.

첫째, 미래 정보화 사회에 대비하기 위함이다. 정보화 사회에서는 단편적·사실적 지식을 암기하고 이해하는 능력보다 정보의 탐색, 수집, 분석, 비판, 종합, 창출 등의 고등 사고능력, 자기 주도적인 평생 학습능력, 효율적인 의사소통 능력, 협동

적 문제 해결 능력이 절실히 요구된다. 수행평가는 이런 능력을 타당하게 평가할 수 있는 방법이다.

둘째, 학교 교육의 정상화를 도모하기 위함이다. 우리나라는 그동안 교육목표로 고등 사고능력, 바람직한 태도 및 정서 함양 등 전인적 인간 양성을 제시해 왔다. 그러나 교육방법과 평가는 이러한 능력이나 품성 함양을 위한 방향에 부합하지 못했다. 우리나라 학교 교육의 비정상성은 이러한 ‘교육목표-교육방법-교육평가’ 사이의 연계가 부족한 데서 기인한 바가 크다. 수행평가는 전인적 인간 양성이란 교육목표의 달성 여부를 평가하는 데 적합한 방법이다.

셋째, 수업과 평가의 긴밀한 연계를 도모하기 위함이다. 학교의 교육목표는 교실 수업을 통해서 실현되며, 교실 수업이 평가의 영향을 받게 되는 것은 불가피하다. 따라서 교실 수업과 평가가 긴밀하게 연계되지 않으면 교육목표의 효과적인 달성은 불가능하다. 전통적인 지필평가는 수업의 결과를 확인하는 방식으로 수업과 간접적으로 연계된다. 그러나 수행평가는 수업의 과정에서 교수학습 활동과 평가 활동을 접목하는데 적절한 평가 방법이다.

넷째, 학생들에게 더욱 의미 있는 평가 결과를 제공하기 위함이다. 평가 결과는 학생의 능력에 대해 적절하고 정확한 정보를 제공해주고 교사에게는 수업개선에 필요한 정보를 제공할 수 있어야 한다. 그러기 위해서 평가는 학생들이 수업을 통해 기르기를 바라는 능력 중 가장 중요한 부분을 포함해야 한다. 그러나 기존의 지필평가는 교육목표나 내용의 중요한 부분 중에서 지필평가가 가능한 지식적 측면만을 간접적으로 평가하는 데 그치는 한계를 지닌다. 따라서 지필평가의 결과만으로 학생의 능력을 종합적으로 이해하거나 수업개선에 필요한 정보를 얻기에는 부족함이 많다. 반면, 수행평가는 학생의 종합적인 능력을 정확하고 타당하게 파악하고, 수업개선에 도움이 되는 의미 있는 평가 결과를 제공할 수 있다.

다섯째, 구성주의적 지식관과 교수학습관에 부응하기 위함이다. 기존의 객관적·절대적 지식관은 주관적·상대적 지식관을 강조하는 방향으로 변화하고 있다. 즉, 지식을 인간의 외부에 독립적으로 존재하는 것이 아니라, 인간 내부에서 개별적인 경험에 의해 주관적으로 구성되어 가는 것으로 보게 된 것이다. 이에 따라 교사의 학습관도 학습자 중심 교육을 지향하게 되었다. 학습은 더 이상 학습자 외부에 객관적·독립적으로 존재하는 지식 체계를 수동적으로 받아들이는 과정으로 간

주되지 않는다. 학습이란 지식의 요소들을 받아들여 학습자 개개인의 경험에 의해 능동적·주관적으로 구성해 나가는 과정으로 해석하게 되었다. 이처럼 변화하는 지식관·학습관에 가장 부합하는 평가체제로써 수행평가가 도입되었다.

이상의 도입 배경에는 전통적 지필평가와 구분되는 수행평가의 특징과 지향점이 잘 드러나 있다. 즉, 우리나라에 수행평가가 도입된 것은 단지 평가 방법만의 변화가 아닌, 선택형 또는 단답형 위주의 지필평가 방식과 수업의 결과만을 확인하는 총괄평가 목적 위주로 이루어졌던 전통적 평가 패러다임의 한계를 극복하기 위함이라 할 수 있다. <표 3>은 여러 연구자가 논의한 수행평가의 특징을 전통적 평가와 비교하여 정리한 것이다(교육부, 1998; 박도순, 1999; 백순근, 1998; 심주옥, 2005; 이효정, 2008).

<표 3> 전통적 평가와 수행평가의 특징 비교

구분	전통적 평가	수행평가
이론적 배경	절대주의, 합리주의 객관적·절대적 지식관	구성주의, 인지심리학, 동기심리학 주관적·상대적 지식관
수업 특성	교사 중심 추상적이고 객관적인 수업 상황 직선적, 위계적 과정으로 진행	학생 중심 구체적이고 실제적인 수업 상황 불규칙적, 지속적 과정으로 진행
교사와 학생의 역할	교사: 지식의 전달자 학생: 수동적인 학습자	교사: 학습의 안내자·촉진자 학생: 능동적인 학습자
평가 시기	학습이 종료된 후 평가가 이루어짐	모든 학습 과정에서 평가가 이루어짐
평가 목적	총괄평가(선발, 분류, 배치) 중심	총괄평가 목적뿐 아니라 형성평가(지도, 조언, 개선) 지향
평가 내용	학문적, 선언적 지식(내용적 지식) 암기력과 이해력 등의 기본 학습능력 학습의 결과 중시 객관성·일관성·공정성 강조	실천적, 절차적 지식(방법적 지식) 비판적 사고, 창의력 등의 고등 사고능력 학습의 결과뿐 아니라 과정도 중시 전문성·타당성·적합성 강조
평가 방법	표준화된 선택형 지필평가 위주 대규모 평가(양적 평가) 중시 일회적·부분적 평가 학생 개개인을 평가 인위적인 시험상황	개별 교사에 의한 관찰과 판단 위주 소규모 평가(질적 평가) 중시 지속적·종합적 평가 집단 단위의 평가도 중시 실제 상황이나 유사 모의상황

<표 3>에서 주의해야 할 점은 수행평가와 전통적 평가의 특징이 상대적인 것이며, 전통적 평가 또는 수행평가만의 절대적 특징은 아니라는 것이다. 즉, 전통적 총괄평가도 학생들의 학습 개선에 도움을 주는 형성평가의 기능을 수행할 수 있고(박정, 2013a), 수행평가도 성적 산출을 통한 총괄평가의 기능을 수행할 수 있다. 하지만 상대적으로 전통적 평가는 수행평가와 비교하면 총괄평가 목적에 집중하는 경향이 있다고 할 수 있다. 마찬가지로 평가의 객관성과 공정성, 타당성과 적합성은

전통적 평가와 수행평가에서 모두 중요하다. 하지만 전통적 평가는 수행평가에 비해 상대적으로 공정성과 객관성을 강조하는 평가 방법이며, 수행평가는 더 실제적이고 구체적인 상황을 제시하며 다양한 평가 방법을 활용하므로 상대적으로 평가의 타당성과 적합성을 더 강조한다고 할 수 있다. 따라서 <표 3>을 해석할 때 이러한 점을 고려할 필요가 있다.

수행평가의 이론적, 철학적 배경을 보다 구체적으로 살펴보면, 수행평가는 구성주의적 평가관과 인지심리학, 동기심리학에 기반을 두고 있다. 먼저 구성주의적 평가관은 학습을 학생이 능동적으로 지식을 구성하는 과정으로 간주하며, 따라서 수업 활동이 학생들에게 유의미 학습이 될 수 있도록 해야 하는 배경을 제공한다(Dana & Nichols, 1992; Driver & Bell, 1986; Kulm & Malcolm, 1991; Roth, 1994). 이러한 관점에서 수행평가는 학생의 개념 변화 과정이나 반성적 사고 과정, 과학 탐구 과정 등을 관찰할 방법을 제공하므로 구성주의적 평가를 구현하는 이상적인 체계로 볼 수 있다(Mintze, Wandersee, & Novak, 1998). 또한, 인지심리학은 수행평가에 적합한 수업 환경 조성에 필요한 이론적 기반을 제공한다. 즉, 인지심리학의 관점에서 바람직한 수행평가는 학생들이 능동적으로 수업 활동에 참여하고, 선개념과 새로운 개념을 연계하며, 다른 학생들과 상호 협력하고, 자신의 학습 과정을 관찰할 수 있는 환경을 조성할 필요가 있다(Glaser, 1984; Glaser & Pellegrino, 1987; Hibbard & Baron, 1990; Johnson & Johnson, 1990). 마지막으로 동기심리학에서는 학생들이 최상의 학습 효과를 내기 위해서는 학습 동기가 유발되어야 한다고 주장한다(Baron, 1991). 학습 동기는 학생의 학습 과정에 선택권이 있을 때, 과제를 수행하는 자신의 능력에 자신감과 책임감을 가질 때, 학습에 대한 자신의 공헌과 그 가치를 인식할 때, 유의미하고 도전적인 과제가 주어질 때 나타난다. 이 또한 이상적인 수행평가 과제가 가져야 하는 조건과 깊은 관련이 있다.

수업과의 연계 측면에서 보면, 전통적 평가에서는 교사 중심의 수업이 이루어진 후 수업의 결과를 확인하는 방식으로 평가가 이루어졌으며, 주로 정해진 답을 선택하도록 하였다. 하지만 수행평가는 수업 과정에서 교수학습 활동과 접목되어 이루어지는 평가로서 더욱 긴밀하게 수업과 연계되며, 학생이 정해진 답을 선택하는 것이 아니라 스스로 계획하고 판단하며 답을 작성하거나 행동으로 나타낼 기회를 제공한다. 이를 위해 수행평가에서는 학생의 환경을 고려하여 가능한 매력적이고, 추

상적인 상황보다는 구체적이고 실제적인 수업 상황에서 평가하고자 한다. 따라서 수행평가는 문제 해결 능력과 확산적, 창의적 사고 중심의 교육이 이루어지도록 하는 역할을 할 수도 있다.

평가 내용 및 방법 측면에서 보면, 과학적 지식의 습득이 평가 목표인 전통적 평가와 달리 수행평가에서 과학적 지식은 과정을 파악하기 위한 수단적 가치를 지닌다. 이에 수행평가에서는 과정과 결과가 가능한 통합된 상태에서 평가가 이루어지도록 노력한다. 또한 기억이나 이해와 같은 단순 사고능력보다 창의, 비판과 같은 고등 사고능력의 측정을 중시하며, 사고능력과 같은 인지적 영역뿐 아니라 행동 발달, 흥미와 태도 등 정의적 영역을 포함하여 종합적이고 전인적인 평가를 중시한다. 그리고 일회적 평가에 그치지보다 학생의 성취나 변화를 누적적, 종합적으로 평가 방법을 중시한다. 이에 관찰 평가나 면담 평가 등 평가 결과를 누적적으로 기록할 수 있는 평가 방법이 권장된다. 마지막으로 수행평가는 학생 개개인을 평가하기도 하지만, 집단 단위의 평가도 중시한다. 학생들은 집단 단위의 수행평가를 통하여 다른 학생 또는 교사와 상호작용하는 기회를 얻고, 과학 지식의 특성인 사회적 합의를 통한 잠정성의 의미를 이해할 수 있다. 이러한 집단적 활동은 사회적 구성주의 관점에서 협동심과 자기표현 능력 및 의사소통 능력의 발달을 촉진하므로 중요하다.

따라서 수행평가가 이러한 본질적 특성과 도입 의도를 살려서 시행된다면, 수행평가는 개별 학생의 성취수준을 타당하게 평가할 수 있는 평가 체계이자 구성주의적 평가를 구현할 수 있는 이상적인 평가 체계로서 궁극적으로 교수학습을 개선하고 학교 교육의 질을 향상하는데 기여할 수 있을 것이다. 그러나 교사와 학생들이 수행평가를 성적 산출을 위한 새로운 평가 방법의 하나로만 인식하고 전통적 총괄평가 목적에서 수행평가를 실시한다면 수행평가가 아무리 제도적으로 확대 시행되더라도 그 의미는 퇴색될 수밖에 없다. 이상의 논의를 종합하여 이 연구에서는 백순근(1998)의 정의를 참고하여 수행평가를 ‘학생 스스로 자신의 지식이나 기능이나 태도를 나타낼 수 있는 답을 구성하거나, 산출물을 만들거나, 행동을 나타내도록 요구하는 평가 방식’으로 정의하고자 한다. 이때 수행평가 과제 측면에서 구성주의적 평가 특성을 고려한 수행평가 즉, 구성주의적 수행평가(constructivist performance assessment)를 바람직한 수행평가 과제의 지향점으로 삼고자 한다.

(3) 평가 전문성의 개념과 구인

일반적으로 교사의 평가 전문성은 바람직한 학생평가를 위하여 교사가 반드시 알고 있어야 하는 지식 및 이를 실천하고 활용하는 실제적인 능력을 의미한다. 이러한 의미에서 교사의 평가 전문성은 곧 학생평가 전문성이라 할 수 있다. 교사가 평가에 관한 의사 결정을 내릴 때는 관련 수업과의 연계를 고려할 필요가 있는데, 수업과 연계된 평가는 학생의 학습과 성취에 관한 정보를 제공하여 수업 활동에서 학생의 수준을 파악하고 이를 개선하기 위한 형성평가 목적으로 활용될 수 있기 때문이다(김수동, 김경희, 김수진, 2005). 이러한 능력은 평가뿐 아니라 교사의 전반적인 업무 수행에 있어 기초적이고 필수적인 능력으로도 볼 수 있으므로, 이러한 관점에서 교사의 평가 전문성은 평가 소양(assessment literacy)이라고도 한다(김경희 등, 2006; Plake, Impara & Frager, 1993; Schafer, 1993). 교사의 평가 전문성은 어느 순간 완성될 수 있는 성질의 것이 아니라, 끊임없는 자기 개발 과정에서 지속해서 향상된다(김신영, 2004).

Popham(2009)은 교사의 평가 소양이 교실 평가뿐 아니라 일상적인 교육 활동에 많은 영향을 미치고 중요한 역할을 한다고 하였다. 예를 들어, 교사가 수업에서 형성평가를 자주 사용한다면 이는 수업의 중요한 일부가 되고 교사의 평가 전문성은 학생의 학습을 돕는 방향으로 발달하게 된다. 반면, 교사가 지역 또는 국가 수준에서 시행되는 표준화된 평가가 학생평가에 더 적합하다고 생각한다면 교사의 평가 전문성은 교실에서 이루어지는 수업과 관련이 적어지게 된다. 그러므로 교사는 어떤 평가가 적절한지 판단하고 최선의 의사 결정을 할 수 있어야 하며, 평가 결과를 해석할 때 학생들의 개인적 특성을 고려하고 반영하는 능력도 필요하다. 이러한 능력을 통틀어 교사의 평가 전문성이라 할 수 있다.

박도순(2007)은 교사의 평가 전문성에 전제될 필요가 있는 요소를 다음과 같이 제시하였다.

첫째, 인본주의에 기반을 둔 평가관의 확립이다. 교사의 평가 활동이 대학 등 상급 교육기관으로의 진학을 위한 수단일 뿐이라면, 이에 따르는 평가 활동은 평가의 신뢰도와 타당도 등과는 별개로 오히려 교육적 의미와 학생의 발달을 저해

할 수 있다. 그러므로 교사의 평가 전문성이 궁극적으로 학교 교육의 정상화에 이바지하기 위해서는 교사의 평가관이 인본주의에 기반을 두고 확립되어야 한다. 이러한 관점에서 학생 간의 상대적인 위치만을 파악하여 학생을 서열화하는 상대평가는 학생 중심의 평가가 되기 어렵다.

둘째, 교육 개선과 학생 중심의 평가이다. 교사가 평가 과정에서 수집하는 정보는 수업과 학습의 개선뿐 아니라 전반적인 학교 교육과 교육과정의 개선을 위해서도 필수적인 활동이다. 그러므로 교사의 평가 전문성을 개발하는 본질적인 이유는 교육 상황을 개선하고 학생들의 학습을 지원하기 위함이 되어야 한다.

셋째, 교사에게 평가의 자율성이 확보되어야 한다. 평가의 자율성이란 교사가 외부의 지시나 간섭 없이 독자적인 기준과 자신의 평가 전문성에 근거하여 평가를 계획, 실행하고 이에 따른 책임도 지는 것을 의미한다. 평가의 자율성이 없으면 평가의 다양성이 확보되기 어렵고, 평가의 다양성이 없으면 교사의 평가 전문성이 발휘될 여지가 거의 없다. 그러므로 교사의 평가 전문성 향상은 교사의 평가권 확보가 전제되어야만 한다. 물론 이를 위해서는 자율성에 따른 책임과 평가 과정 및 결과에 관한 투명성도 부여되어야 할 것이다.

넷째, 평가 과정과 결과 활용은 교육 혁신을 현장에 정착시키는 기능을 수행해야 한다. 이는 교육 혁신이 학교 현장에 구현되고 정착되기 위한 모든 요소를 평가가 규정하고 있기 때문이다.

다섯째, 교사의 평가 전문성은 윤리적, 도덕적 문제를 포함해야 한다. 교사가 지켜야 할 윤리적, 도덕적 문제에는 교사의 비도덕적 행위나 왜곡, 사생활 침해 등이 있을 수 있다. 평가의 윤리와 도덕성을 확보하고 유지하기 위해서는 학교 구성원들의 협동적인 노력이 필요하다.

다음으로 평가 전문성의 구인(construct)에 대하여 알아보고자 한다. 일반적으로 평가 전문성의 구인에는 평가하려는 특성과 영역, 평가의 목적과 방법, 적절한 평가도구를 개발하는 방법과 절차, 원하는 정보를 수집, 해석하는 능력, 결과를 활용하는 방법에 대한 이해 등을 들 수 있지만, 평가의 어떤 측면을 강조하는가에 따라 강조점이 다를 수 있다. 예를 들어, 평가의 전통적 측면을 강조한다면 평가의 타당도와 신뢰도, 평가의 객관성과 공정성 등을 고려한 평가도구나 문항의 개발 능력이 평가 전문성의 구인으로 강조된다. 하지만 평가의 구성주의적 측면을 강조한다면

평가를 교수학습의 맥락에서 활용할 수 있는 태도와 능력이 평가 전문성의 구인으로 강조될 것이다. 선행연구에서 제시한 과학교사의 학생평가 전문성의 구인과 기준은 다음과 같다.

김경희 등(2006)은 교실에서 이루어지는 학생평가를 기준으로 교사의 평가 활동에서 드러나는 평가 전문성의 구인으로 다섯 가지를 제안하였다.

첫째, 평가 방법의 선정 능력이다. 이는 교사가 교수학습 목표를 이해하고 평가 목적을 구체화하며, 수업의 진행 및 교수학습 활동을 고려하여 평가 방법을 선정할 수 있는 능력이다.

둘째, 평가도구의 개발 능력이다. 이는 타당한 평가도구의 개발 계획을 세울 수 있는 능력, 계획에 따라 실제로 평가도구를 개발할 수 있는 능력, 평가도구의 질을 관리할 수 있는 능력을 의미한다.

셋째, 평가의 실시, 채점 및 성적 부여 능력이다. 이는 평가의 타당도와 공정성과 관련되며 교사가 적절한 시기와 상황에 맞게 절차에 맞는 평가를 실시하는 능력이다. 또한, 평가가 이루어지기 위한 환경을 조성하고 관리하는 능력, 정확한 채점 및 성적 부여 방법에 대한 이해 및 적용 능력이 포함된다.

넷째, 평가 결과의 분석, 해석, 활용 및 의사소통 능력이다. 이는 교사가 평가 결과를 타당하게 분석하고 해석하여 교육적 의사 결정에 활용할 수 있는 능력과 교사가 학생, 학부모 등 관련자들에게 평가 결과를 설명하는 등 의사소통을 할 수 있는 능력을 의미한다.

다섯째, 평가의 윤리성이다. 이는 학생에 대한 인격 존중, 법적 윤리, 자신의 평가 활동에 대한 자기반성의 내면화 등 평가와 관련한 지식 이외에 교사가 기본적으로 갖추어야 할 윤리성을 의미한다.

또한, 김신영(2002)은 평가의 질이 곧 교수학습의 질을 좌우한다고 강조하며 교사의 학생평가 전문성의 기준 여섯 가지를 다음과 같이 제시하였다.

첫째, 교실 내에서의 평가 활용 능력이다. 이는 수업과 평가를 불가분의 관계로 보는 관점으로, 교사가 평가 실행이나 과정을 수업 전략으로 활용할 수 있어야 함을 의미한다. 교사가 수집한 평가 정보는 교실 환경에서 교육적 의사 결정의 수단으로 활용되는데, 교사는 학생들의 성적을 부여하고 학습 부진아를 판별하는 전통적 측면뿐 아니라, 수업의 성취 목표에 관하여 학생들과 의사소통하고 학생에게 더

나은 학습의 기회를 제공하기 위해서도 평가 정보를 활용할 수 있어야 한다.

둘째, 평가하고자 하는 학생의 특성에 대한 이해 능력이다. 교사는 개념 이해도나 학업 성취도와 같은 인지적 능력 이 외에도 사고력, 문제해결력 등의 고등 사고 능력이나 정의적 영역에 대한 평가를 실시할 수 있다. 그러므로 교사가 평가할 수 있는 다양한 학생의 특성 및 평가 영역에 대한 이해, 그리고 이를 타당하게 평가할 수 있는 평가 방법에 대한 이해가 요구된다.

셋째, 평가의 질에 대한 이해 능력이다. 교사의 평가는 수업 상황 등 평가 맥락에 따라 많은 영향을 받는다. 따라서 교사는 맥락에 따라 평가의 질이 불안정해지지 않도록 유의하며 상황을 관리하는 능력을 가져야 한다.

넷째, 다양한 평가 방법과 도구에 관한 이해 능력이다. 교사는 같은 평가 특성과 영역에 대해서도 다양한 평가 방법과 도구를 활용할 수 있다. 이 능력에는 평가 결과를 종합하고 토론할 수 있는 능력도 포함된다.

다섯째, 평가에서 교사와 학생의 상호 관계에 대한 이해 능력이다. 평가는 교사가 학생에게 일방적으로 실시하는 것이 아니라, 학생들과 상호작용함으로써 이루어지는 것이다. 따라서 교사들은 평가에서 학생과의 관계 및 학생의 심리적인 요소를 이해하고 활용할 필요가 있다.

여섯째, 평가 결과의 활용 방법에 관한 이해 능력이다. 교사는 평가 결과를 수업과 교수학습 개선에 활용하고, 평가 결과에 대하여 학생과 학부모 등 교육 관계자들과 의사소통할 수 있어야 한다. 그리고 평가 결과의 적절하지 못한 활용이 미칠 수 있는 부정적 영향에 대해서도 이해할 필요가 있다.

다음으로 외국에서 제시하는 평가 전문성의 구인을 보면, 미국의 교육 관련 기관에서 공동으로 제안한 교사의 평가 전문성 기준을 들 수 있다(AFT, NCME, & NEA, 1990). 이 기준은 교사의 평가 전문성의 구인을 7가지 영역으로 제시하고 있다.

첫째, 평가 방법의 선정이다. 교사가 수업에서 의사 결정을 내리기 위해서는 적절한 평가 방법을 선정해야 한다. 이때 측정 오차와 타당도, 학생의 배경 등을 고려할 수 있어야 한다. 이는 평가의 타당도와 관련이 있다.

둘째, 평가 방법의 개발이다. 평가 방법을 선정한 후, 교사는 의사 결정에 필요하거나 수업 의도에 부합하는 정보를 수집하기 위한 구체적인 평가 방법과 도구를

개발하거나 타인이 개발한 것을 선택할 수 있어야 한다. 이는 교사가 단순히 학생 평가 방법을 선정하는 것이 아니라 구체적인 평가도구의 질을 점검할 필요성이 있음을 강조한다.

셋째, 평가 실행이다. 이 영역은 교사가 평가를 실시한 후 채점하고 결과를 해석하는 것까지 포함한다. 교사는 채점을 위한 기준을 마련하고 활용하며, 평균과 분산 등의 정량적 평가 결과를 분석하고 이해할 수 있어야 한다.

넷째, 평가 결과의 활용이다. 교사는 학생 개인에 관한 판단, 수업 계획, 교육과정의 개발, 학교의 향상 등 다양한 분야에 평가 결과를 활용할 수 있고, 평가 정보를 축적하여 활용할 수도 있다. 정량적 평가 분석 결과는 학생의 특성을 구체적으로 파악하는데 활용될 수 있다. 또한, 자신이 실시한 평가뿐 아니라 지방 또는 국가 수준의 평가 결과도 교육의 질을 개선하는 데 활용할 수 있어야 한다.

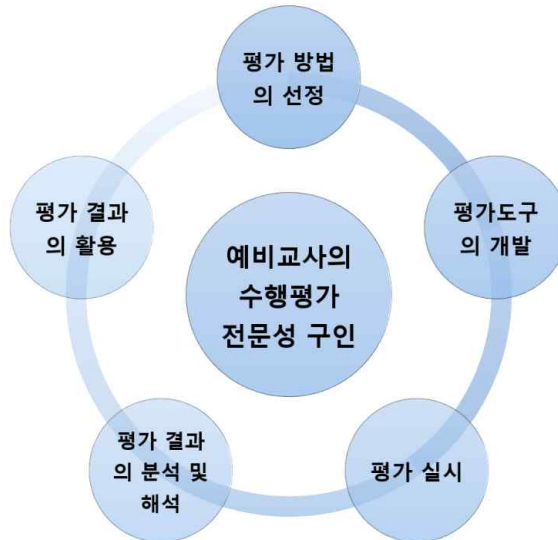
다섯째, 성적 부여이다. 교사는 성적 부여 절차를 고안하고 실행하기 위한 절차와 방법을 알아야 하며, 성적 부여 방식에 따른 결과를 산출하고, 반성하고 수정할 수 있어야 한다.

여섯째, 평가 결과에 관한 의사소통이다. 교사는 학생, 학부모, 동료 교사 등 관계자들에게 평가 결과에 대한 적절한 설명을 제공하고 학생들의 교육적 성취에 관한 의견을 교환할 수 있어야 한다. 평가 결과를 해석할 때 교사는 학생의 배경 변인을 고려할 줄 알아야 하고, 측정의 오차와 평가 방법의 제한점에 대해 폭넓은 인식을 하고 이를 염두에 두어야 한다.

일곱째, 평가에 관한 윤리적 의식이다. 교사는 평가와 관련한 학교와 국가 수준의 법조항에 대해 이해하고 있어야 한다. 공정한 절차를 준수해야 하며, 학생의 권리를 존중하고 사생활을 침해하거나 윤리적으로 문제 있는 평가를 미리 방지해야 한다. 그리고 지역 또는 국가 단위로 실시되는 표준화된 학력 검사가 교수 효과 측정에 부적절할 수도 있음을 알고 있어야 한다.

평가 전문성에 관한 이상의 논의를 바탕으로, 예비교사의 수행평가 전문성 향상을 목적으로 하는 이 연구에서는 수행평가 전문성을 ‘평가의 전통적 측면을 고려하되 구성주의적 평가를 지향하는 구성주의적 수행평가에 관한 전문성’으로 정의하고, 구체적으로 ‘수행평가의 구성주의적 특징과 지향점에 대한 이해를 바탕으로 이를 고려하여 학생의 학습과 성취에 관한 정보를 체계적으로 수집, 해석, 활용할

수 있는 능력' 으로 구체화하였다. 또한, 수행평가 전문성의 구인으로는 예비교사의 맥락을 고려하여 학부모와의 의사소통과 같은 현직교사 맥락의 구인을 제외하고 평가 방법의 선정, 평가도구의 개발, 평가 실시, 평가 결과의 분석 및 해석, 평가 결과의 활용을 선정하였다(그림 4).



<그림 4> 예비교사의 수행평가 전문성 구인

1.3 PCK와 평가 전문성의 관계

앞서 논의한 바와 같이, PCK의 구성요소에는 평가에 관한 지식이 포함된다. 하지만 이것이 교사가 평가를 계획하고 실행할 때 평가 방법이나 영역과 같은 평가에 관한 지식만 고려하면 된다는 의미를 지니는 것은 아니다. 특히 구성주의적 평가 관점에서 평가는 수업의 계획, 실행, 반성 단계에서 모두 일어나며 수업과 평가를 별개의 활동으로 보지 않으므로 평가를 계획하고 실행할 때 모든 PCK 구성요소를 고려할 필요가 있다. 이러한 점에서 초등 과학교사 11명을 대상으로 형성평가와 PCK 사이의 관계를 조사한 Falk(2012)는 형성평가와 PCK가 상호보완적 관계에 있다고 하였다. 선행 연구 결과, 교사의 PCK는 형성평가 분석을 위한 지식의 원천이 되었으며 그중에서도 과학 교수 전략에 관한 지식과 과학교육과정에 관한 지식이

가장 자주 사용되었다. 또한, 과학교사들은 형성평가의 분석 과정에서 교육과정의 목표와 학생에 관한 지식을 구성하고 정교화하는 등 PCK가 계발되기도 하였다. 따라서 형성평가는 교사의 PCK를 활용하고, 통합하고, 계발할 기회를 제공하지만, 과학 교수 전략에 관한 지식을 계발하기 위해서는 추가적인 자료가 필요함을 알 수 있었다. 한편, 민희정(2012)은 현직 과학교사의 학생평가 실태에 대한 조사를 바탕으로 PCK의 관점에서 과학 평가에 관한 지식, 과학교육과정에 관한 지식, 학생에 관한 지식, 과학 교수 전략에 관한 지식의 네 가지로 구성된 과학교사의 학생평가 요소 모델을 도출하였다.

평가 전문성의 구인으로 PCK를 제안한 연구도 이루어졌는데, 한국교육과정평가원에서는 이인제 등(2004)의 연구를 통하여 과학교사의 평가 전문성 요소로 PCK의 구성요소인 과학 평가에 관한 지식, 과학교육과정에 관한 지식, 과학 교수학습에 관한 지식, 과학 교과 내용에 관한 기준을 제안하였다. 이때 과학 평가에 관한 지식에는 평가 방법과 유형, 유사 개념 등에 관한 지식이 포함된다. 그리고 과학교육과정에 관한 지식은 우리나라 과학교육과정의 목표와 내용 체계, 과학교육과정이 추구하는 인간상과 철학에 관한 이해 및 학년별 내용 구성에 관한 이해를 포함한다. 과학 교수학습에 관한 지식에는 학생의 학습 상황, 교육목표 달성에 적합한 교수학습 전략과 교수 자료에 관한 지식이 포함된다. 과학 교과 내용에 관한 기준에는 과학 및 과학 교과의 학문적 성격과 구조, 과학 교과에서 다루는 핵심적 내용 지식, 과학적 탐구 방법과 태도 등이 포함된다.

교사의 평가 전문성과 PCK 사이의 관계는 평가 과제의 개발 과정에 따라서도 살펴볼 수 있다. 예를 들어, 이인제 등(2004)은 학생평가 과정을 네 단계로 나누었는데, 각 단계에서 이루어져야 할 작업과 이때 주로 활용되는 PCK 구성요소를 연결하면 <그림 5>와 같이 기술할 수 있다.

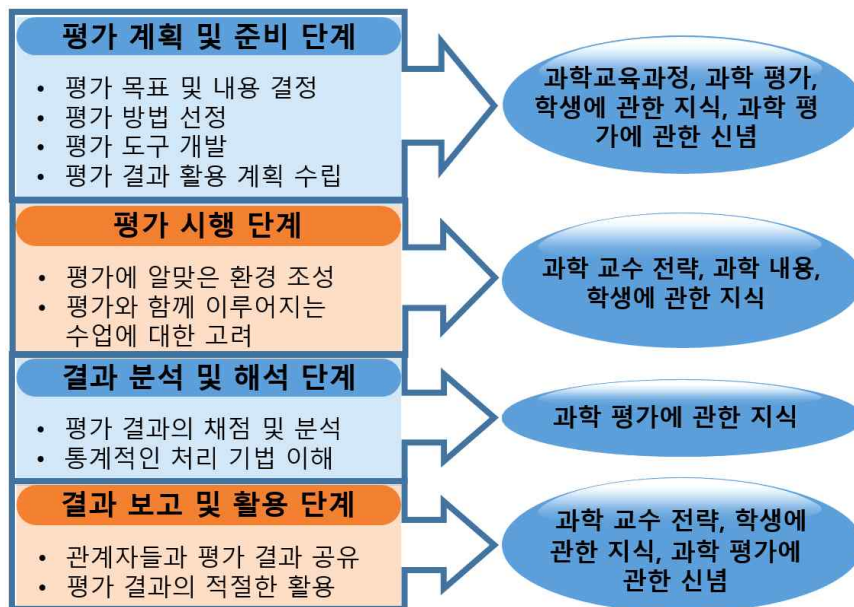
첫 번째는 평가 계획 및 준비 단계이다. 이 단계에서 구체적인 평가 목표와 내용을 결정할 때에는 과학교육과정에 관한 지식이 요구된다. 평가 방법을 선정하고 도구를 개발하는 과정에서는 학생들의 인지적, 정의적 특성을 고려하여 평가 영역과 방법을 결정해야 하므로 평가에 관한 지식과 학생에 관한 지식이 요구된다. 또한, 교사는 평가의 실행 및 결과 분석, 활용 등에 관한 전반적인 계획을 세워야 하는데 이는 모든 PCK 구성요소의 영향을 받지만, 특히 교사의 과학 평가에 관한 신

념이 가장 큰 영향을 미친다고 할 수 있다.

두 번째는 평가 시행 단계이다. 이때 교사는 평가에 알맞은 환경 조성을 해야 한다. 특히 형성평가의 경우 대개 교수학습의 일환으로 이루어지므로 과학 교수학습에 관한 지식은 필수적으로 활용되며, 과학 내용, 학생에 관한 지식도 요구된다.

세 번째는 평가 결과의 분석 및 해석 단계이다. 이 단계에서는 평가 결과의 통계적인 처리 기법을 이해하고 채점 및 평가 결과를 분석하고 해석해야 하므로 과학 평가에 관한 지식이 많은 영향을 미친다.

네 번째는 평가 결과 보고 및 활용 단계이다. 이 단계에서는 평가 결과를 학생, 학부모 등 관계자들과 공유하고 의사소통하는 것이 중요하며, 평가 결과를 적절하게 활용하여 학생의 학습에 도움을 줄 수 있는 전략을 세울 수 있어야 한다. 따라서 이 단계에서는 과학 교수 전략에 관한 지식과 학생에 관한 지식, 과학 평가에 관한 신념이 많은 영향을 미칠 수 있다.

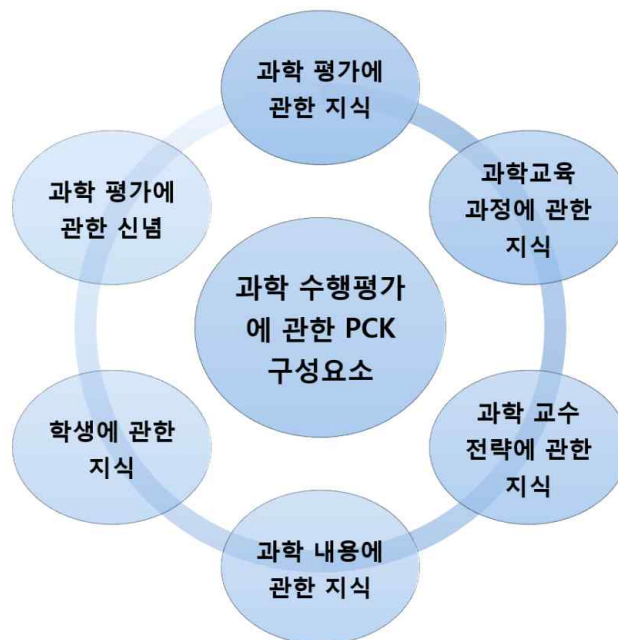


<그림 5> 수행평가 과제 개발 과정에서 PCK 구성요소의 활용

한편, 최근에는 수행평가뿐 아니라 전통적 평가 형태로 여겨지는 지필평가에서도 PCK와의 관계를 분석한 연구가 이루어지고 있다. Five와 Barnes(2017)는 예비교

사들이 총괄평가 문항의 선정 과정에서 학생의 인지적, 정의적 특성과 같은 학생에 관한 지식, 학습 목표 등의 과학교육과정에 관한 지식, 학습 계획과 같은 과학 교수학습 전략에 관한 지식을 고려한다고 하였다. 그리고 노태희와 박재성, 강훈식(2016)은 예비교사의 지필평가 문항 개발 과정을 PCK의 관점에서 분석하였으며, 그 결과 과학 평가에 관한 지식이 가장 많이 고려되었으나 과학교육과정, 과학 내용, 학생에 관한 지식도 비교적 많이 고려된 것으로 나타났다. 또한, PCK 구성요소의 통합 측면에서는 두 가지 혹은 세 가지 측면의 통합이 나타나 지필평가 문항의 개발 활동도 예비교사의 PCK 개발 기회를 제공할 수 있다고 주장하였다.

평가 전문성과 PCK의 관계에 관한 이상의 논의를 바탕으로 과학 수행평가의 개발에 필요한 PCK 구성요소를 정의하였다. 즉, Magnusson과 Krajcik, Borko(1999)의 PCK 구성요소에 과학 내용에 관한 지식을 더하여 과학 평가에 관한 지식, 과학교육과정에 관한 지식, 과학 교수 전략에 관한 지식, 과학 내용에 관한 지식, 학생에 관한 지식, 과학 평가에 관한 신념의 6가지를 과학 수행평가에 관한 PCK 구성요소로 정의하였다(그림 6).



<그림 6> 과학 수행평가에 관한 PCK 구성요소

2. 과학교사의 평가 전문성 향상

2.1 과학교사의 전문성 향상

1980년대까지 사범대학의 예비교사 양성과정이나 교사 전문성 향상 또는 연수 프로그램에서는 과학교육의 맥락을 고려한 내용보다 일반적인 교육학 내용에 더 비중을 두었다. 하지만 Shulman(1986, 1987)에 의해 PCK 개념이 도입된 이후로는 과학 내용 지식과 일반 교육학 지식을 모두 강조하고, 과학교사의 전문성 향상 프로그램도 교사의 PCK 개발을 중심으로 해야 한다는 논의가 전개되었다(Cochran, DeRuiter, & King, 1993; Veal & MaKinster, 1999).

대표적으로 Magnusson과 Krajcik, Borko(1999)는 과학교사의 전문성 향상 프로그램에 앞서 논의한 PCK의 구성요소를 포함해야 하며, 구체적으로 과학교육의 목적, 과학 교수학습 목표와 과학교육의 목적 사이의 관계, 특정 주제에 관한 교수학습 전략, 수업의 계획, 실행 및 반성, 평가의 계획과 실행이 고려될 필요가 있음을 주장하였다. Veal과 MaKinster(1999), De Jong과 van Driel, Verloop(2005), Cochran과 DeRuiter, King(1993)은 과학교사의 양성을 위한 교육과정의 목표가 예비교사의 PCK 개발이 되어야 한다고 주장하였으며, 이를 위해서는 프로그램에 과학 교수학습 전략에 관한 지식과 오개념 등 학생에 관한 지식 요소를 포함해야 한다고 보았다. 같은 맥락에서 Short(2006)는 교사의 전문성 향상 프로그램이 교과 내용, 학생, 교수학습 방법을 잘 이해할 수 있도록 개발되어야 한다고 강조하였다. Rhoton과 Wojnowski(2006)도 교과 내용 지식과 실제 교수학습에 대한 이해를 함양해야 한다고 강조하였다. Beck과 Kosnik(2006)은 PCK에 구성주의적 관점을 더하여 이론과 실제, 학습과 생활을 연결하는 통합적 방법, 탐구 지향 방법, 사회기반 방법을 통하여 과학교사를 양성해야 한다고 하였다.

Loucks-Horsley와 Stiles(2001)는 그동안 과학 및 수학교사의 전문성 향상은 학생의 학습을 촉진하는 교육 체계의 여러 가지 구성요소를 분리해 왔기 때문에 어려움을 겪어 왔다고 지적하며, 교사의 신념을 변화시키는 변혁적 학습의 관점에서 선행연구 분석을 통해 과학교사의 전문성 향상 프로그램의 특성과 원리를 종합하고, 이를 바탕으로 교사 전문성 개발 프레임 개발하였다. 이 프레임은 개발자의 지식

과 신념, 비판적 논쟁거리, 전략 요소를 바탕으로 한 프로그램의 목적과 계획의 수립, 프로그램을 실시하고 보완하는 실행과 반성의 네 단계로 구성되어 있다. 이후에도 그들은 광범위한 선행연구 조사 및 연구를 지속해서 수행하여 전문성 향상 프로그램과 프레임의 발전을 꾀했다. 그 결과 Loucks-Horsley 등(2003)을 거쳐 Loucks-Horsley 등(2009)에서는 수업, 평가, 교육과정 등 교육 시스템의 전반적 요소에 대한 교사의 전문성을 모두 포괄하는 교사 전문성 개발 원리(principles of effective professional development)를 제시하고, 교사 전문성 향상 프로그램에 교사 전문성 개발 원리가 체계적으로 반영될 수 있도록 전문성 개발 설계 프레임(professional development design framework)을 함께 제시하였다. 이 연구에서 개발한 수행평가 전문성 향상 프로그램은 이를 기반으로 하였으므로 Loucks-Horsley 등(2009)이 논의한 교사 전문성 개발 원리와 전문성 개발 설계 프레임에 관하여 자세히 다루고자 한다.

(1) 교사 전문성 개발 원리

역사적으로 교사의 전문성 향상은 교사들이 기존의 생각과 신념을 버리거나 바꾸도록 돕는 것보다 교사가 알아야 하는 새로운 기술과 지식을 가르치는 데에만 집중하였다. 이에 교사들은 자신들이 알아야 하는 지식의 양에 압도되어 기본 의도나 배경 등을 충분히 이해하지 못하는 경향이 있었다. 이 경우 교사는 학생들의 학습을 오히려 저해하는 부주의한 결정을 내릴 수도 있다(Appleton, 2003). 이에 교사의 전문성 향상을 위해 진정으로 필요한 것이 무엇인지에 대한 인식과 논의가 이루어지게 되었으며, 과학교사에게 필요한 내용 지식을 체계적으로 전달하는 것보다 교사가 프로그램의 내용을 역동적인 지식 체계로 이해하고, 스스로 자신이 더 많은 내용을 가르칠 준비가 되었다고 느낄 수 있도록 하는 것이 더 중요하다는 주장이 제기되었다(Banilower *et al.*, 2006; Weiss & Pasley, 2009). 실제로 과학교사의 전문성 향상을 위한 프로그램은 교사들에게 새로운 지식과 기술을 더해줄 수 있을 뿐 아니라 교수학습에 대한 생각과 신념을 변화시키는 계기를 제공할 수 있다. 이러한 관점에서 전자가 ‘추가적 학습(additive learning)’의 기회를 제공한다면, 후자는 ‘변혁적 학습(transformative learning)’의 기회를 제공한다고 할 수 있을 것이다.

변혁적 학습은 교사의 신념, 지식, 습관을 변화시키는 전략에 교사를 적극적으로 참여시키는 것으로, 변혁적 학습을 위한 조건에는 다음과 같은 요구사항이 있다 (Thompson & Zeuli, 1999).

첫째, 학생, 교과 내용, 학습에 관한 교사의 신념과 관행의 균형을 깰 수 있는 높은 수준의 인지적 부조화를 창출한다. 이로써 교사가 교과 내용, 교수 전략 또는 수업 방법 측면에서 더 좋은 방법이 있다는 것을 이해하고 더 많이 배우도록 할 수 있다.

둘째, 교사가 부조화를 경험하고 생각할 기회를 충분히 제공한다. 이를 위해 교사들이 읽고, 토론하고, 경험한 것을 이해할 충분한 시간과 구조, 지원이 제공되어야 한다.

셋째, 교사 자신의 상황과 교수학습 관행에 관한 부조화를 생성하고 해결하기 위해 학생의 과제, 수업 녹화 자료를 활용하고 분석하는 활동을 포함한다.

넷째, 교사가 다음 질문에 답할 기회를 제공함으로써 교사가 새로운 이해에 맞는 실천을 할 수 있도록 한다. ‘새로운 이해를 얻었는가.’, ‘이제 교실에서 무엇을 변화시킬 것인가.’, ‘학생들이 새로운 이해를 하도록 도울 방법은 무엇인가.’

다섯째, 교사들이 교수학습과 관련한 새로운 문제를 파악하고, 문제에 대해 새로운 이해를 하며, 실행을 변화시키는 과정을 반복하는 지속적인 향상 과정에 참여할 수 있도록 한다.

변혁적 학습은 교사가 단순히 새로운 기술을 익히거나, 현재 지식에 새로운 지식을 통합하는 추가적 학습과 다르다고 할 수 있다(Thompson & Zeuli, 1999). 교사의 전문성 향상을 위해 변혁적 학습과 추가적 학습 방법을 모두 활용할 수 있지만, 프로그램에서 다루는 내용 및 이유에 대하여 교사가 의식적으로 선택할 기회를 제공하는 것은 중요하며, 이러한 점에서 변혁적 학습이 바람직하다고 할 수 있다. 이러한 배경을 바탕으로 Loucks-Horsley 등(2009)이 제안한 교사 전문성 개발 원리의 구성요소는 다음과 같다. 아울러 그들은 이 원리가 현직교사를 중심으로 기술되었으나, 예비교사 교육 프로그램에도 적용할 수 있다고 하였다.

첫째, 효과적인 교사 전문성 개발은 학생의 학습 목표 및 요구를 충족하도록 설계된다. 즉, 학생들의 성장에 대한 증거 자료에 근거하여 교사가 지속적이고 중요한 내용에 집중할 수 있도록 하고, 학생의 학습 향상에 필요한 지식과 기술을 개발

할 수 있도록 도와준다.

둘째, 효과적인 전문성 개발 경험은 효과적인 교실 교수학습에 관한 잘 정의된 이미지에 의해 좌우된다. 예를 들어, 이 이미지에는 과학을 학습하는 모든 학생에 대한 헌신, 탐구 기반 학습에 대한 강조, 조사, 문제 해결 및 지식 적용, 핵심 개념에 대한 심층적인 이해를 강조하는 접근법, 학생들이 새로운 이해를 구축하도록 유도하는 것, 의미 있는 성취를 측정하기 위한 명확한 수단 등이 포함된다.

셋째, 효과적인 전문성 개발 경험은 교사가 내용 지식과 PCK 및 기술을 개발하고 실행을 비판적으로 검토 및 반영할 기회를 제공한다. 즉, 교사가 과학에 대한 심층적인 지식과 PCK를 개발하고, 교육과정 및 학습 경험을 선택하고 통합하는 데 도움을 준다.

넷째, 효과적인 전문성 개발 경험은 연구를 기반으로 하며 교사들이 성인 학습자로서 참여한다. 효과적인 전문성 개발은 교사들이 있는 곳에서부터 시작하며, 심층 조사, 공동 작업 및 반성을 위한 충분한 시간을 제공한다. 그리고 교사들의 다른 전문성 개발 경험 및 관련 활동과 명시적으로 연결될 수 있도록 한다.

다섯째, 효과적인 전문성 개발은 교사가 자신의 실행을 지속해서 향상하기 위해 동료 및 학습 공동체의 다른 전문가들과 함께 일할 기회를 제공한다. 지속적인 학습이 학교의 규범과 문화의 일부가 되고, 교사는 함께 배우고 모범 사례를 공유할 수 있도록 한다.

여섯째, 효과적인 전문성 개발 경험은 교사가 직업 경력 전반에 걸쳐 전문성을 심화시키고 리더십을 익히도록 지원한다. 예를 들어, 교사는 다른 교사의 지지자, 변화의 대리인, 개혁의 촉진자 역할을 한다.

일곱째, 효과적인 전문성 개발 경험은 교육 시스템의 다른 부분에 대한 연결을 제공한다. 예를 들어, 교사의 전문성 개발은 다른 지역 또는 학교의 계획, 지역 또는 주 단위의 교육과정 체계 및 평가와 통합되며 지역 사회 내에서 적극적인 지원을 받을 때 효과적으로 이루어질 수 있다.

여덟째, 효과적인 전문성 개발 경험은 교사의 효율성, 학생 학습, 리더십 및 학교 공동체에 긍정적인 영향을 주기 위해 지속적으로 평가된다.

이 원리는 지난 약 40년간 교사 전문성 개발에 관한 신념이 어떻게 변화했는지 보여준다. 1970년대 초반의 교사 전문성 향상 프로그램은 주로 교사 훈련(training)

이라 불렀다. 또한, 교사 전문성 개발자는 흔히 강사(trainer)라고 불렸으며, 주로 강의와 시연을 학습 도구로 사용하였다. 이때 교사 교육의 목표는 주로 새로운 프로그램이나 접근법에 관한 전문적 지식을 제공함으로써 교사들의 지식을 향상하는 것이었다. 이러한 프로그램에서 제공된 교육 자료는 학습에 관한 단순한 ‘요리법’에 가까웠고, 엄격히 말해 추가적 학습이었기 때문에 교사의 생각이나 근본적인 가정을 변화시키거나 전문성 문화를 형성하는 데에는 미흡했다. 오늘날 이것의 대부분은 바뀌었지만, 오래된 패러다임의 흔적은 여전히 남아있다. 따라서 지엽적인 부분에 초점을 맞춘 교사 전문성 개발 패러다임에서 벗어나기 위한 지속적인 노력이 필요하다.

특히 학교는 교사들의 지속적인 학습 공동체에 관심을 기울이고 이를 조직해야 할 많은 이유가 있는데, Mundry와 Stiles(2009)는 이러한 방식이 학습의 일관성을 높이고, 교사들의 고립감을 줄이며, 교사의 지식을 효과적으로 개발하고, 학생들이 조기에 참여할 수 있는 체계를 마련할 수 있는 장점이 있다고 하였다. Hord와 Sommers(2008)는 광범위한 문헌 조사를 통하여 학교의 교사 학습 공동체에 관하여 다음과 같은 다섯 가지 특성을 제안하였다.

첫째, 교사들은 학생의 성공적인 학습에 관하여 끊임없이 신념, 가치관, 비전을 공유한다. 그들은 학생 학습뿐 아니라 성인 학습에도 중점을 둔다. 또한 필요시 어떤 조치를 취하기 위한 리더십 및 의사 결정 권한을 공유하고, 학교는 그들의 의사 결정 과정을 명확히 한다.

둘째, 교사들은 학교에서 함께 학습하고, 학습한 것을 학생들에게 함께 적용할 수 있도록 조직된다. 이때 학교는 교사가 학습 공동체를 형성하기 위한 시간, 장소, 자원, 리더십 및 지원 체계를 가졌는지 확인하고, 교사들이 개방적이고 정직하며 배려하는 기여자가 되는데 필요한 대면 기술을 가졌는지 확인한다.

셋째, 교사들의 학습 공동체가 조직된 학교의 특징은 교사들의 교육 실천을 다른 사람들에게 소개한다는 것이다. 교사는 서로 관찰하고 피드백을 제공하면서 교육에 대한 접근법을 공유하고, 이를 통해 교사의 전문성은 지속해서 향상된다.

교사 전문성 향상을 위한 프로그램의 개발자에게 교사의 학습을 최적화하는 조력자의 역할은 매우 중요하다(Remillard & Geist, 2002). 이때 조력자의 지위(과학자, 교수, 교사 교육자 등)보다 조력자의 지식, 기술 및 준비가 프로그램의 질에 더 큰

영향을 미칠 수 있다(Banilower *et al.*, 2006). 또한, 프로그램에서 조력자와 교사가 하는 활동은 교실에서 교사와 학생들이 하는 활동과 달라야 하며, 성인 학습의 조건을 충족시키기 위한 그 이상의 무언가가 필요하다는 인식이 커지고 있다. 예를 들어, 성인 학습자는 새로운 지식이 실제 상황에서 제시될 때 가장 효과적으로 배우며, 새로운 아이디어와 기술을 자신이 가진 다양한 배경에 연결할 시간이 필요하다(Knowles, Holton, & Swanson, 2000). 교사 전문성 향상을 위한 프로그램의 조력자는 내용 지식 및 교사가 내용에 대해 어떻게 생각하는지에 관한 지식, 교사와 내용 사이의 연결을 촉진시키기 위한 효과적인 전략과 기술을 가지고 있어야 한다(Carroll & Mumme, 2007). 조력자는 성인 학습자인 교사가 스스로 학습 목표를 설정하고 자신의 사전 지식을 활성화하며, 새로운 정보, 자원, 아이디어, 현상과 상호작용할 수 있는 경험을 제공해야 한다. 학생들에게 필요한 학습주기와 마찬가지로, 성인 학습에서도 경험, 반영 및 평가에 이르는 학습의 뚜렷한 단계를 포함한 프로그램의 설계는 여러 가지 이점이 있다. 따라서 효과적인 조력자는 교사를 성인과 전문가로서 존중하는 규범과 기본 원칙을 수립하고, 여러 학습 양식을 사용하며, 교사가 다양한 방식으로 정보를 처리하고 생각하며 상호작용하기 위한 충분한 시간을 제공하는 등 교사에게 필요한 학습 환경을 만드는데 주의를 기울여야 한다(Garmston & Wellman, 2009).

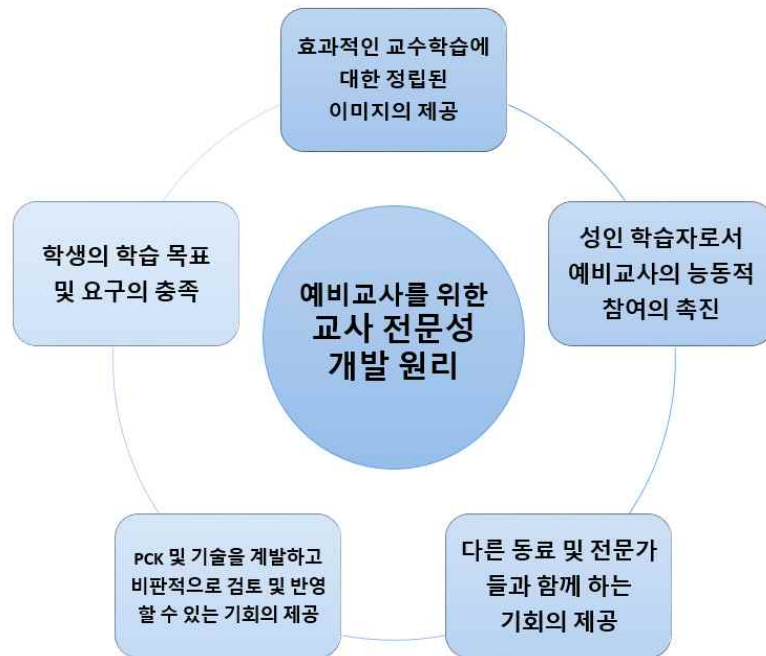
또한, 현직교사를 대상으로 하는 프로그램 개발자는 프로그램 설계에 교사의 경력을 고려해야 하는 경우가 많다. 예를 들어, 초임교사인지 경력교사인지 여부에 따라 전문성 향상 전략과 목표가 달라질 수 있다. 경력 3년 미만의 초임교사는 소개 또는 멘토링 프로그램을 통한 전략이 유효할 수 있고, 경력 4-10년 사이의 교사는 다른 교사를 지도하는 멘토 또는 코치 활동을 통하여 전문성을 향상시킬 수 있다. 이처럼 교사의 경험을 인정하는 것은 교사의 참여를 촉진하거나 이끌기 위한 특정 전략에 영향을 미칠 수 있다(Johnson & The Project on the Next Generation of Teachers, 2007). 또한, 프로그램 개발자가 초임 및 경력교사가 어떻게 학습하는지 고려하는 것도 프로그램 개발에 도움이 된다. 예를 들어, 경력교사들은 초임교사보다 주어진 정보의 기능과 의미 있는 패턴에 주목하고, 주제에 대한 깊은 이해를 반영하는 체계화된 많은 내용 지식을 가지고 있으며, 추가적인 노력을 거의 들이지 않고도 지식의 중요한 측면을 검색할 수 있다(Bransford, Brown, & Cocking,

1999). Stiles와 Mundry(2002)는 교사의 전문성 개발을 위한 전략을 선정할 때 다음과 같은 점을 고려하는 것이 도움이 될 수 있다고 하였다.

- 과학 내용 지식에 관한 기본 개념과 원리에 대해 깊은 이해를 할 수 있도록 명시적으로 설계된 활동.
- 학생들이 과학을 어떻게 생각하고 배우는지 이해할 기회의 제공.
- 사전 지식과 학습에 기초한 학습 기회의 제공.
- 동료와의 상호작용 및 협력을 위한 시간과 구조의 마련.
- 학생 학습을 분석하고 반영할 기회의 제공.
- 새로운 학습을 교수 전략으로 전환할 충분한 기회의 제공.

특히, 경력교사와 초임교사는 동일한 교사 전문성 개발 프로그램에 참여하더라도 학습의 양상이 다르게 나타나는 것으로 보고되고 있다(Stiles & Mundry, 2002). 예를 들어, 교실 수업 비디오 분석에서 초임교사는 수업이 조직되는 방식과 교사의 특정 동작에 더 집중하는 반면, 경력교사는 학생들이 말하고 행동하는 것과 교사가 개별 학생들의 말과 행동에 어떻게 반응하는지에 더 초점을 두었다. 따라서 프로그램 개발자는 초임교사와 경력교사가 다양한 방식으로 참여할 수 있다는 것을 알고 프로그램을 설계해야 하며, 따라서 비슷한 배경을 가진 교사들끼리 그룹화하여 학습 방법을 결정하는 것은 유용할 수 있다.

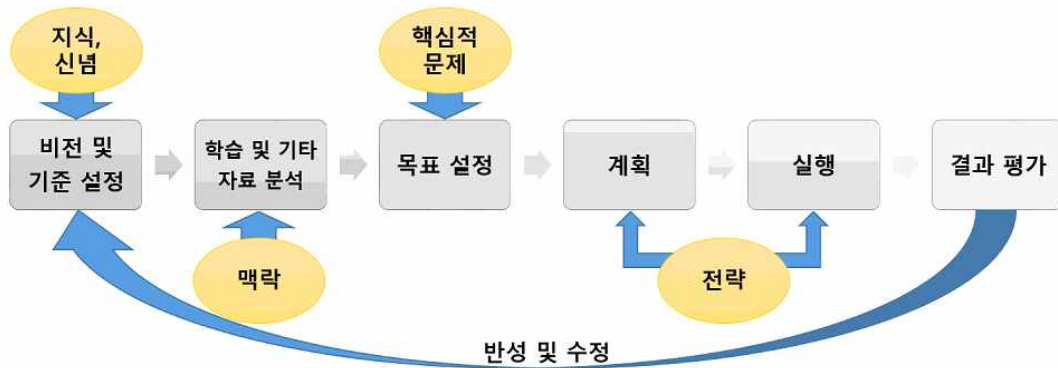
Loucks-Horsley 등(2009)은 교사 전문성 개발 원리가 교사 전문성 향상 프로그램의 개발 과정뿐 아니라 사범대학의 예비교사 양성과정에서도 중요하게 고려되어야 한다고 주장하였다. 교사 전문성 개발 원리 중에서 직업 경력 전반에 걸친 전문성 심화, 교육 시스템의 다른 부분에 대한 연결 제공, 학교 공동체에 긍정적인 영향의 요소는 현직교사에 적용되는 맥락이므로, 이 연구에서는 예비교사 맥락에 적용이 가능한 교수학습에 대한 정립된 이미지의 제공, 학생의 학습 목표 및 요구의 충족, PCK 및 기술을 계발하고 비판적으로 검토 및 반영할 기회의 제공, 동료 및 다른 전문가들과 함께하는 기회의 제공, 성인 학습자로서 예비교사의 능동적 참여 촉진의 다섯 가지 요소를 바탕으로 평가 전문성 향상 프로그램을 개발하였다(그림 7).



<그림 7> 예비교사를 위한 교사 전문성 개발 원리

(2) 전문성 개발 설계 프레임

Loucks-Horsley 등(2009)은 교사 전문성 향상을 위한 프로그램에 교사 전문성 개발 원리가 체계적으로 반영될 수 있도록 전문성 개발 설계 프레임(그림 8)을 함께 제시하였다.



<그림 8> 전문성 개발 설계 프레임

전문성 개발 설계 프레임의 첫 번째 단계는 비전 및 기준 설정(commit to vision & standards)이다. 과학교육의 궁극적인 비전은 학생이 학습을 통해 과학적 소양을 함양하고 관련 지식과 기술, 추론 과정을 깊이 이해하도록 하는 것이라 할 수 있다. 이를 성취하기 위해 교사는 확고한 교과 내용 지식과 기술, 행동, 과학적 소양을 갖추어야 한다. 또한, 교사는 높은 수준의 비전에 대한 주인의식을 가지고 적절한 교육 환경을 조성하고 교육의 질을 유지할 수 있다는 자신감을 가져야 한다. 이는 교사가 교과 내용 지식을 충분히 숙지하고 있고, 성공적인 학습을 위한 다양한 범주의 교수 전략을 숙지하고 있다고 스스로 느끼는 감각을 포함한다.

교사 전문성 향상 프로그램의 목표를 달성하기 위해서는 프로그램에 참여하는 교사들도 그들이 성취해야 하는 비전에 대해 명확히 알 필요가 있다. 특히 학교는 교사의 협력과 효과적인 교수 실행의 공유를 저해하는 장애물을 없애야 하고, 학교 자체가 교사들의 학습 공동체가 될 수 있어야 한다. 그러나 학교에서 이러한 기회를 충분히 가지는 것이 어렵다면 교사 전문성 향상 프로그램에서 이를 제공할 수 있어야 한다. 이것이 전문성 개발 설계 프레임의 첫 단계가 비전 및 기준 설정이 되는 이유이다. 이때 고려해야 할 구체적인 질문에는 다음과 같은 것이 있다.

- 과학 교수학습에서 비전은 무엇인가?
- 학생들이 과학을 잘 할 수 있으려면 무엇을 알아야 하는가?
- 학생이 지식을 이해했다면 우리가 그것을 어떻게 알 수 있는가?

- 학생이 지식을 얻지 못했을 경우 우리는 무엇을 해야 하는가?
- 새로운 비전이 펼쳐지는 교실의 모습은 어떠해야 하는가?
- 학생이 이러한 기준을 달성하기 위해서 교사는 무엇을 알아야 하며, 무엇을 할 수 있어야 하는가?
- 교사 학습의 비전은 무엇인가?
- 새로운 비전이 펼쳐지는 교사 전문성 향상 프로그램의 모습은 어떠해야 하는가?
- 과학 교수학습과 전문성 향상 프로그램의 비전을 지원하기 위해 어떤 조직이 필요한가?

이때, 지식과 신념(knowledge & beliefs)은 교사 전문성 향상 프로그램 개발자의 전문적 식견을 구성한다. 지식이란 선행연구로 뒷받침되는 사실을 의미하며, 신념은 개인의 경험, 관찰, 확신에 기초한 믿음을 의미한다. 이는 프로그램에서 무엇을 하고 하지 말아야 하는지에 대한 모든 전문적인 판단을 내리고 구체화하는 근거가 된다. 지식과 신념은 비전 및 기준 설정에 투입되는 시간과 노력을 경감시키고, 확고한 이론적 기반을 제공하며, 불필요한 시행착오를 피할 수 있도록 도와주므로 전문성 향상 프로그램의 개발 과정에서 반드시 고려되어야 한다. 고려되어야 하는 지식과 신념의 영역에는 학생과 학습자, 교사와 교수학습, 과학의 본성, 성인 학습과 전문성 향상, 지식과 생각에 관한 변화 과정 등이 있다.

두 번째 단계는 학습 및 기타 자료 분석(analyze student learning & other data)이다. 이 단계에서 개발자들은 비전과 기준에 근거하여 현실을 돌아보며 현재 상태와 원하는 상태 사이의 격차를 탐색하고, 개선을 위한 목표를 설정한다. 만약 학교 공동체가 모든 학생들에게 높은 기준을 설정한다면 학생들의 학습 자료를 더 엄격하게 분석할 수 있으므로 기준에 미치지 못하는 결과에 대하여 격차를 경험할 가능성이 더 클 것이다(Love *et al.*, 2008). 학생의 학습 및 기타 자료를 분석하는 것은 학생의 학습을 향상하기 위한 구체적인 목표를 확인하고, 학생의 결과에 분명히 집중하기 위한 교사 전문성 향상 프로그램의 기초를 형성하기 위함이다. 만약 교사 전문성 향상 프로그램이 학교나 지역 단위에서 개발된다면, 해당 학교나 지역의 목표와 프로그램의 계획을 연계하는 것도 중요하다.

이 단계에서 고려해야 하는 요소는 맥락(context)이다. 어떤 상황에서도 항상 완벽하게 적용되는 교사 전문성 향상 프로그램은 없다. 따라서 교사 전문성 향상 프로그램 개발자는 이론(지식과 신념)과 맥락을 모두 고려할 필요가 있다. 즉, 전문성 향상 프로그램은 과학의 교수학습과 전문성 향상 프로그램이 어떠한 형태를 갖춰야 하는지에 관한 비전에 영향을 받으며, 동시에 프로그램이 교사의 맥락과 얼마나 밀접하게 관련이 있는지 고려하고 신중하게 분석해야 한다. 개발자가 고려해야 할 맥락적 요소에는 학생과 그들의 교육적 요구, 교사와 그들의 교육적 요구, 교육과정·교수·평가 실행·학습에 관한 환경, 조직 문화와 교사 공동체, 리더십, 지역 및 국가 수준의 정책, 가용한 자원, 가족과 공동체 등이 있다.

세 번째 단계는 목표 설정(set goals)이다. 비전 및 기준 설정을 통하여 개발자가 바라는 미래를 묘사하고, 학생의 학습 및 기타 자료에 대한 철저한 분석을 통하여 현실을 기술한다면, 목표는 비전을 향하여 전진할 기준점 혹은 주춧돌이 된다. 적은 수의 분명하고, 구체적이며, 이를 수 있는 목표는 교사 전문성 향상에 동기를 부여하고 집중할 수 있게 한다. 목표 설정에는 학생의 학습을 위한 목표, 교사의 학습을 위한 목표, 교사의 교수 실재를 위한 목표, 기관을 위한 목표의 네 가지 영역을 고려할 필요가 있다.

한편, 지금까지 개발된 교사 전문성 향상 프로그램을 살펴보면 맥락과 관계없이 개발자가 직면한 공통적인 문제를 발견할 수 있는데, 이를 핵심적 문제(critical issue)라고 한다. 그러므로 개발자들은 목표 설정 단계에서 이 문제를 어떻게 다루어야 할지 미리 고민해야 하며, 그렇지 않으면 이 문제는 나중에 문제가 될 가능성이 있다. 개발자가 고려해야 할 핵심적 문제에는 지속 가능성을 위한 역량 구축, 전문성 향상을 위한 충분한 시간 확보, 리더십 개발, 형평성 보장, 전문성 학습 문화의 구축, 공공 지원 확보, 규모 키우기 방법 등이 있다.

네 번째 단계는 계획(plan)이다. 구체적인 목표가 설정되면, 개발자는 프로그램의 성과를 어떻게 측정할지 생각하고 프로그램의 설계를 구상해야 한다. 예를 들면, ‘목표 달성을 위해 어떤 전략의 조합이 최선인가.’, ‘만약 우리의 목표가 내용 지식의 향상에 초점을 둔다면, 어떤 전략을 사용해야 하는가.’, ‘교사들이 새로운 지식을 실천으로 옮길 수 있게 하려면 어떤 전략을 사용해야 하는가.’, ‘우리는 프로그램의 목표 달성에 필요한 리더십을 가지고 있는가, 그렇지 않다면 어떤

전략이 필요한가.’와 같은 질문을 할 수 있다. 이 과정에서 개발자들은 교사의 맥락을 재검토하고 중요한 요소를 다시 발굴하며, 학습 및 기타 자료 분석 결과를 다시 살펴보고 프로그램을 환경에 맞추는 것을 고려한다.

이때 교수학습, 과학 내용, 전문성 향상 프로그램, 변화 과정 등에 관한 정보가 더 필요할 수도 있다. 이럴 때는 연구에 기반을 두어 전문성 향상 프로그램의 비전과 기준, 핵심적 문제를 다시 살펴보는 것이 계획을 위한 새로운 아이디어를 생성하는 데 도움이 되며, 다른 프로그램의 계획이나 교육 문헌 등을 폭넓게 검토하는 것도 도움이 될 수 있다. 계획은 프로그램의 기저를 이루는 신념을 재검토하며 정교화하는 작업이기 때문이다.

다섯 번째 단계는 실행(do)이다. 계획 단계가 스케치라면, 실행 단계는 이를 실제로 구현하는 페인팅과 같다. 이 단계에서 개발자들은 내용 지식의 전문가이자 촉진자로서, 프로그램의 구현 및 변화 과정에 관한 자신들의 지식과 기술을 모두 활용한다. 근본적인 변화가 일어나기 위해서 교사는 교실에서 구현이 가능한 방식으로 학습을 경험해야 하고, 새로운 행동을 실험하고 이를 충분히 이해할 시간적 여유가 필요하다. 교사들이 어떤 것을 느끼고 새로운 접근법을 활용할 때, 대체로 개발자가 예측 가능한 단계를 따라 움직일 것이다. 때로는 교사가 새로운 교수 전략을 시도하며 어려움을 겪는 과정에서 종종 더 나빠지는 현상이 나타나기도 한다. 이때 아무런 도움이 제공되지 않고 같은 어려움을 겪는 동료 교사와의 대화 기회가 주어지지 않는다면 빠르게 프로그램 참여 이전의 상태로 되돌아갈 위험도 있다. 그러므로 지속적인 도움의 제공, 교실 방문, 교사들을 돕기 위한 다른 지원 체계를 제공하는 것이 필수적이다. 그러므로 실행 단계에서 개발자들이 교사들의 활동과 프로그램의 진행 과정을 주의 깊게 살펴보고 이를 평가하는 것은 필수적이다.

계획과 실행 단계에서 프로그램에 투입되어야 하는 것이 전략(strategies)이다. 이 시점에서 개발자들은 교사들의 전문성 향상을 위해 어떤 전략을 도입할 것인가에 대하여 고민하게 된다. 교사와 마찬가지로, 효과적인 개발자들은 학습자들의 전문성을 끌어낼 다양한 전략을 이해하고 있어야 하며, 능숙하게 전략을 선정하고 종합하여 전문성 향상 프로그램의 목표를 달성할 수 있어야 한다. 또한, 예비 또는 현직교사, 초임 또는 경력교사와 같은 참여 교사의 특성이나 요구, 학습 유형에 따라 서로 여러 집단의 교사들에게 동시에 다른 전략을 제공할 수 있다.

여섯 번째 단계는 결과 평가(evaluate results)이다. 교사 전문성 향상 프로그램에서 필수적이지만 자주 간과되는 단계가 바로 이 단계이다. 교사 전문성 향상 프로그램은 구체적이고 특정한 목표 또는 넓은 범위의 다양한 목표 달성을 위해 설계될 수 있다. 이때 목표를 성공적으로 달성하는 방법이 무엇인지 결정하는 것이 평가의 역할이다. 그러나 이러한 과정은 다음의 몇 가지 이유로 쉽지만은 않다.

첫째, 개발자는 프로그램의 목적과 무관하게 가장 측정하기 쉬운 참여자의 만족도를 통하여 프로그램을 평가하려는 경향이 있다. 이는 프로그램의 결과 해석에 관하여 심도 있게 생각하는 것을 어렵게 한다. 둘째, 참여 교사들이 가르치는 학생의 성취를 통해 전문성 향상 프로그램의 가치를 평가하려는 요구가 커지고 있다. 이는 많은 자원이 투입되는 교사 전문성 향상 프로그램에서 학생의 학습 향상을 이뤄내는 것이 필수적이라는 점을 근거로 한다. 하지만 학생의 학습 성과를 성급하게 기대하는 요구는 교사 전문성 향상 프로그램의 효과가 교사에게 완전히 받아들여져 교사의 학습과 실행의 변화를 끌어내는 데 시간이 필요하다는 점에서 문제가 있다. 셋째, 학습 경험이 프로그램의 개발자, 참여자 및 다른 사람들에게 가치 있는 학습 경험으로 작용할 수 있지만, 이러한 점이 충분히 고려되지 않는다는 점도 있다. 이러한 점을 고려하여 결과 평가 단계에서 개발자가 할 수 있는 질문은 다음과 같다.

- 프로그램이 원하는 결과나 목표는 무엇이었는가?
- 프로그램의 결과와 성취를 보여주는 증거는 무엇인가?
- 프로그램의 결과와 실행에서의 변화 측정에 관한 자료를 어떻게 얻었는가?
- 학습 경험으로 평가를 어떻게 활용하는가?

마지막 일곱 번째 단계는 반성 및 수정(reflect & revise)이다. 이 단계는 전문성 개발 설계 프레임의 모든 단계에 내포되어 있으며, 연속적이고 항상 진행 중인 단계이다. 예를 들어, 프로그램에 투입된 요소 중 고정된 것은 아무것도 없다. 교수학습에 기초한 지식, 수학과 과학의 본성, 전문성 향상 프로그램, 변화 과정 등은 모두 지속해서 변화하고 성장하는 것이다. 따라서 개발자들은 끊임없이 ‘교사 전문성 향상 프로그램에 새로운 정보를 어떻게 반영할 것인가’를 자문하고 반영할 필요가 있다. 개발자의 신념 또한 변화할 수 있는데, 개발자들은 프로그램의 효과를

분석하며 학생, 교사, 규율, 전문성 향상 프로그램, 변화에 대하여 이전과는 다르게 생각할 수 있고, 이는 계획과 실행 단계에 다시 영향을 미칠 수 있다. 핵심적 문제 또한 역동적이다. 평가 전문성 향상 프로그램의 실행 경험은 개발자들이 이 전까지 알지 못했던 새로운 문제를 고려하게 하고 더 깊은 수준의 이해를 얻을 수 있게 한다.

2.2 과학교사의 수행평가 전문성 관련 선행연구

우리나라에 수행평가가 도입된 1998년 이후 이루어진 과학교사의 수행평가 전문성과 관련하여 이루어진 선행연구를 분석하였다. 선행연구를 유형에 따라 평가의 실태와 인식 조사 연구, 평가 모형, 도구, 과제 개발 연구, 평가 채점 및 측정 연구, 과학교사의 평가 전문성 향상을 위한 전략 또는 프로그램 개발 연구로 범주화하여 정리하였다.

(1) 평가 실태와 인식 조사

홍정림과 최은경(2002)은 449명의 중등학생을 대상으로 수행평가에 관한 인식과 자아효능감을 분석하였다. 학생들은 수행평가의 목적, 전통적 지필평가와의 차별성 등 기본적인 수행평가의 특성에 대해 이해하고 있었다. 그러나 수행평가가 자신의 학습이나 잠재적인 능력 개발에 도움이 되지 않는 못하며, 평가의 신뢰도나 타당도에 대해서도 만족하지 못한다고 응답하였다. 이때 고등학생이 중학생보다, 남학생이 여학생보다 더 부정적으로 응답한 비율이 높았다. 학생들의 자아효능감 수준은 실험 절차와 실험 도구의 활용 등 과제 수행 과정에 대해서는 높았으나, 실험 목표나 토론 주제에 대한 이해 등 과제 수행 결과의 정리에 대해서는 상대적으로 낮았다. 이때 여학생은 남학생보다 자아효능감 수준과 수행평가 점수가 유의미하게 높았고, 자아효능감과 수행평가 점수 사이에도 유의미한 상관성이 있는 것으로 나타났다.

김성원과 현미숙(2005)은 중학교 과학교사 80명을 대상으로 수행평가의 일반적 사항과 방법에 관한 인식을 조사하였다. 연구 결과, 교사들은 수행평가에서 교과 내용과 과정, 과학적 원리와 개념의 통합이 중요하며, 학생들이 배워야 할 지식과

기능을 고려하여 과제를 선정해야 한다고 인식하고 있었다. 평가 방법 측면에서는 물리, 화학 단원의 경우 서술형 평가, 실험 실습평가가 적절하다고 여겼으며, 지구 과학 및 환경 단원에서는 서술형 평가, 연구 보고서 평가, 포트폴리오 평가가 적절하다고 인식하고 있었다.

김석우(2007)는 52명의 과학교사를 대상으로 설문 조사와 면담, 수행평가 관련 자료 분석을 통하여 고등학교 과학과 수행평가의 실태를 분석하였다. 연구 결과, 수행평가의 타당도와 신뢰도, 평가 관점의 구체성에 있어 한계가 있었고, 평가 시기는 학기 말에 편중되어 있었다. 수행평가에 대한 교사들의 실천 의지도 부족하였다. 이에 수행평가의 공정성과 신뢰도를 높이기 위한 교내 연수, 성적관리위원회의 내실 있는 운영 등 제도적 개선과 교사, 학생, 학부모 등 관계자들의 의견수렴 및 반영의 필요성을 주장하였다.

정시온(2007)은 중학교 과학교사를 대상으로 수행평가에 대한 실태, 인식, 이해 정도를 조사하였다. 연구 결과, 수행평가의 성적 반영 비율은 30-40%였으며, 기본 점수는 10-70%까지 매우 다양하였다. 평가 방법으로는 실험 실습평가와 보고서 평가를 가장 많이 사용하였고, 수업은 주로 전통적인 강의 형태 위주로 이루어졌다. 대부분의 교사들은 수행평가를 수업 중 수시로 실시하고 있었고, 평가 결과는 성적 산출 및 학습 결과물로 활용하였다. 수행평가 과제 개발에서는 학습 내용과의 관련성을 가장 많이 고려하였다. 교사들은 대체로 수행평가에 대해 긍정적으로 인식하고 있었지만, 소요 시간, 업무 과중, 과제 개발의 어려움, 평가의 신뢰도와 객관성, 학생 과다 등으로 인한 어려움을 겪는 것으로 나타났다.

서정권(2008)은 실험 실습법을 중심으로 고등학교 과학과에서 실시되고 있는 수행평가의 실태와 문제점을 조사하였다. 연구 결과, 한 학기당 수행평가 실시 횟수는 1시간씩 1-2회였으며, 주로 사용하는 평가 방법은 실험 실습평가, 서술형 및 논술형 평가, 보고서 평가의 순서였다. 실험 실습평가의 기본 점수는 10-70%로 다양했으며, 점수 차이는 결과 및 자료 해석 부분에서 주로 나타났다. 태도 점수의 반영 범위는 5-20%였다. 수행평가 과제 개발은 학습 내용과의 연관성과 타당성을 중점으로 하였고, 실험 실습법에 대한 교사의 인식은 대체로 긍정적이었다. 교사가 겪는 어려움의 원인은 학생 수의 과다, 평가 과제의 개발과 준비, 평가의 신뢰도와 객관성 순서로 나타났다.

고민석 등(2013)은 설문지와 면담을 통하여 초등 교사의 탐구에 대한 신념과 과학 수행평가의 실제 사이의 관계를 조사하였다. 연구 결과, 대부분의 초등 교사들은 탐구 실험에서 학생의 사고 과정과 행동을 관찰하고, 선지식 구조의 습득보다 스스로 해석을 만드는 사고 과정을 포함한 구성주의적 탐구 관점을 가지고 있었다. 또한, 교사 대부분은 수행평가의 목적이 과학 지식의 적용을 평가하는 데 있다고 보았다. 그러나 실제 평가에서는 대부분 평가의 용이성, 객관성 확보와 수행평가 과정과 방법에 관한 이해 부족 등으로 지필평가를 사용하는 것으로 나타났다.

정은영과 최원호(2014)는 518명의 교사를 대상으로 설문을 통하여 우리나라 초, 중, 고등학교 과학과의 평가 실태를 조사하였다. 설문 조사 결과, 학교급이 올라갈수록 수행평가의 비율이 줄어드는 것으로 나타났다. 서술형 평가의 경우 채점 시간과 채점 결과에 대한 신뢰도를 고려하여 제한된 범위에서 실시하는 경향이 중, 고등학교에서 더 많이 나타났다. 고등학교에서는 수행평가를 도입 목적에 부합하게 실시하는 비율이 가장 낮았다. 수행평가에서 다양한 능력을 평가하는 경우도 있었지만, 주로 수행과 관련한 평가만 평가하고 수행평가에 소요되는 시간과 노력 문제로 인해 수행평가가 형식적으로 실시되는 경우도 있었다.

신정윤과 유선아, 양일호(2016)는 초등 교사 203명을 대상으로 설문을 통하여 과학 수행평가의 실시 현황과 자료 개발에 대한 교사들의 요구를 분석하였다. 연구 결과, 교사들은 수행평가 과제를 개발할 때 기존에 제작된 문항에 의존하는 경향이 높았고, 다양성 부족과 체계적인 평가 방법의 부재를 느끼고 있었다. 특히 주제별 평가 문항과 탐구 과정을 평가할 수 있는 문항, 평가 후 피드백과 관련한 자료에 대한 요구가 높았다. 따라서 수행평가 자료의 현장 활용성을 높이려면 학습 과정을 평가할 수 있는 탐구 중심의 주제별 수행평가 과제와 평가 기준에 대한 명확한 기준, 예시자료, 피드백 자료가 개발되어야 한다고 주장하였다.

신정윤과 양일호(2017)는 과학과 수행평가에 관한 기존 문헌을 검토하여 과학과 수행평가의 문제점과 개선 방향을 분석하였다. 연구 결과, 교사들은 과정을 경시하고 결과물과 성취도, 과학적 지식만을 측정하고 있으며, 수업과 수행평가가 분리되어 실시되고 있었다. 따라서 수행평가의 목적과 방법에 대한 인식 개선과 평가 소양의 강조가 필요하고, 교사의 평가 전문성을 향상할 기회를 제공할 필요가 있다고 주장하였다.

과학 교과와 수행평가 인식과 실태에 관한 연구는 수행평가 제도의 도입 이후 꾸준히 이루어져 왔으며, 연구 대상도 초·중등 교사로 다양하였다. 그러나 대부분의 연구는 수행평가가 본래의 도입 취지에 맞게 구성주의적 평가로 활용되지 못하고 형식적으로 시행되거나 전통적 평가에 가깝게 시행되고 있으며, 교사의 인식 또한 부정적인 비율이 높음을 보고하고 있다. 따라서 학교 현장에서 수행평가에 대한 교사들의 인식과 실태가 만족할만한 수준에 도달하지 못하는 이유를 밝히고, 교사들의 수행평가 전문성 향상을 위한 연구가 이루어질 필요가 있다.

(2) 평가 모형, 도구, 과제 개발

박종원 등(2000)은 현직 중등 교사와의 모임을 분석하여 수행평가의 성공적인 현장 적용을 위해 고려할 점을 수업과의 연관성, 수행평가 과제 개발 과정, 수행평가 적용 시 어려운 점, 채점 기준의 개발의 네 가지로 나누었다. 이를 통해 바람직한 수행평가 개발 방향으로 첫째, 평가가 수업의 일환이 되도록 개발할 것. 둘째, 다원적 평가가 이루어지도록 가능한 여러 가지 유형의 평가를 개발할 것. 셋째, 교사가 여러 측면을 함께 고려하여 선택하여 활용할 수 있도록 데이터베이스 식으로 개발할 것. 넷째, 인지적 영역뿐 아니라 정의적, 사회적 영역도 포함되도록 개발할 것. 다섯째, 객관성이 확보되도록 점수 부여 방식을 구체적으로 제시할 것을 제안하였다.

김정숙과 여철현(2002)은 화학교육 분야에서 이루어지는 수행평가를 개선하고자 5개 화학실험 주제에 대하여 시범 실험 관찰, 문제 인식과 가설 설정, 탐구 설계, 자료 기록, 자료 변환과 해석, 결론 도출과 평가로 이루어진 탐구 과정의 기능 요소별 수행평가 모형을 개발하였다. 그리고 과학적 탐구능력을 신장시킬 수 있는 구체적인 학생용 탐구 활동 수행평가 문항지와 교사용 탐구 활동 수행평가 안내서를 개발하였다.

최문수와 정인기(2002)는 학생들의 탐구능력을 측정하고, 주제별 토론 과제를 제시하고, 다양한 문제 유형을 지원하는 웹 기반 평가도구를 개발하였다. 이를 통해 교사의 평가 소요 시간 및 관리의 제약성을 해소하고, 학생들에게 즉각적인 피드백을 제공하여 학습 향상을 도모하였다.

이택수(2003)는 초등과학에서 창의적 문제해결력 신장을 위한 수행평가 자료를 개발하였다. 전문가 집단의 협의를 통하여 학생용 활동지와 교사용 지침서로 이루어진 수행평가의 구성 체계를 개발하였고, 교수학습을 통하여 평가가 자연스럽게 이루어지도록 하였다. 영역 체계는 창의성을 포함하여 개발하였다. 개발한 수행평가 과제를 적용한 후 평가 결과를 분석하였고, 창의성과 관련한 측면에서 질적인 향상이 있었음을 보고하였다.

최희정과 임채성, 김은진(2003)은 초등과학의 생물 단원에서 과학 문제 해결 과정을 적용한 수행평가 도구를 개발하였다. 먼저 학생들이 실제적인 활동으로 느낄 수 있도록 동물원으로 구체적인 상황을 설정하고, 그 동물에 대한 자료를 분석하는 과제를 개발하였다. 과학 지식을 활용하여 과학 문제를 해결하는 활동을 중심으로 학생용 활동지와 채점 체계를 개발한 후, 예비 연구를 통해 수정 및 보완하여 수행평가 도구를 개발하였다.

이종기와 임낙룡, 황성수(2004)는 고등학교 생물 단원을 중심으로 과학영재의 탐구능력 신장을 위한 서술형 수행평가 문항을 개발하였다. 이를 위해 교과서 탐구 활동 분석, 탐구 활동 주제선정, 성취기준 개발, 평가 기준 개발 및 평가 영역 선정, 등급화 방안 개발 과정을 거쳤고, 최종적으로 서술형 평가 문항과 모범 답안, 채점 기준표를 개발하였다.

정영란(2004)은 고등학교 생물 단원을 중심으로 준거지향적 수행평가 문항을 개발하였다. 이를 위해 교과서를 분석하여 필수 학습 요소를 추출하고 성취기준 및 평가 기준을 개발하였다. 이 기준에 따라 서술형, 논술형, 포트폴리오 문항을 개발하고 문항별 구체적인 채점 기준을 제시하였으며, 고등학교 1학년 학생에게 적용하고 양호도 검증을 실시하였다.

심주옥(2005)은 초등과학에서 STS 주제에 대한 수행평가 도구를 개발하고 적용하여 효과를 분석하였다. 이를 위해 교육과정 단원을 재구성하고 STS 주제의 단계별로 수행 과제, 반응 양식(활동지), 채점 체계를 포함한 활동지를 구성하였다. 효과는 과학 태도, 과학 탐구능력, 과학 지식의 측면에서 조사하였다. 그 결과, 과학 태도에서는 유의미한 차이가 나타나지 않았으나 과학 지식과 과학 탐구능력에서는 유의미한 향상이 있었다.

홍수연(2006)은 구성주의적 관점에서 학생들의 자기 주도적 학습능력을 신장시

키고자 탐구학습 모형을 활용한 수행평가 과제를 개발하였다. 즉, 학생들이 직접 탐구 과정에 참여하여 자신의 학습 욕구와 필요에 부합하는 유의미한 지식을 구성하며 문제를 해결하도록 하였다. 10개의 과학과 수행평가 과제를 개발한 후 적용하였다. 분석 결과, 개발한 수행평가 과제는 과학적 탐구능력, 과학 학습에 대한 동기 및 태도에 긍정적인 효과를 보였고, 고차원적 사고능력 함양에도 효과가 있었다.

김상달과 최성봉, 한상아(2008)은 지구과학교육 분야에서 수행평가 모형을 개발하고 활용방안을 제시하였다. 중학교의 기상, 천문, 지질 영역에서 한 가지씩 주제를 선정한 다음 핵심 구성 질문, 수행 능력, 주요 개념 및 내용, 수행 과제, 질적 준거가 포함된 수행 척도 모형을 개발하였다. 그리고 구체적인 단원과 성취기준, 수행 능력, 준비물, 평가 기준과 방법 등이 제시된 수행 활동지와 채점 기준, 활용 시 고려사항 등이 담긴 교사용 자료를 개발하였다.

최혁준(2013)은 전류계와 전압계 회로 연결 및 구성 실험을 주제로 실제 실험을 대체할 수 있는 컴퓨터 기반 수행평가 문항을 개발하였다. 개발된 문항은 실제 실험 도구를 활용한 평가와 유사하게 구성하였고, 수행에 따른 피드백과 결과 채점이 자동으로 진행될 수 있는 알고리즘을 개발하였다.

Gale 등(2016)은 알짜 힘, 가속도, 속도 등의 물리 개념과 관련한 시뮬레이션 기반 수행평가 도구를 개발하여 중학생들에게 적용하였다. 이를 바탕으로 시뮬레이션 기반 수행평가 과제를 개발할 때 자료의 개발, 관리 및 분석 과정에서 고려해야 할 다양한 요소에 대하여 논의하였다.

전혜진(2017)은 핵심역량에 기반을 둔 과정 중심 수행평가를 개발하고, 이를 적용한 수업이 학생들의 과학적 개념과 과학 관련 태도에 미치는 영향을 조사하였다. 수행평가 과제 개발을 위하여 교육과정 자료를 분석하고 핵심 성취기준을 중심으로 과학적 핵심역량을 기를 수 있는 14 차시의 수행평가 자료를 개발하였다. 수행평가 자료의 적용 결과, 통제 집단과 비교하여 과학 관련 태도에 긍정적인 영향이 있었으나 과학 개념의 이해에는 유의미한 차이가 없었다.

수행평가의 모형, 도구 또는 과제를 개발한 연구는 국내 수행평가 관련 연구에서 특히 높은 비율을 차지하였다. 탐구 과정, 창의적 문제해결력, 문제 해결 과정, 자기 주도적 학습능력 등 다양한 측면을 강조한 수행평가 모형과 예시 과제를 개발하고 적용하여 효과를 조사한 연구가 대부분이었고, 컴퓨터 또는 웹 기반 수행평

가 과제를 개발한 예도 있었으며, 단순히 특정 단원에 대한 수행평가 과제와 활동지 등을 개발한 연구도 있었다. 이러한 연구는 수행평가 과제 개발에 어려움을 겪는 과학교사가 참고할 수 있는 데이터베이스를 제공한다는 측면에서 실용적인 의미가 있다. 하지만 수행평가는 표준화된 책무성 평가나 선택형 지필평가와 달리 수업과 연계되는 경우가 많으므로 개별 학교의 상황과 학생의 특성 등 맥락에 많은 영향을 받아 해당 도구를 적용하기 어려울 수 있다. 또한, 교육과정이 개정되면 해당 수행평가 자료를 활용하기 어렵게 될 가능성이 크다. 이러한 점을 고려할 때 수행평가 과제, 모형, 도구 등을 개발하는 연구보다는 교사들이 양질의 수행평가 과제를 직접 개발하고 활용할 수 있도록 과학교사의 수행평가 전문성을 함양하는 연구가 이루어져야 할 것이다.

(3) 평가 채점 및 측정

서미령(2004)은 동일한 평가지를 활용하여 다른 시기에 평가를 했을 때 학업 성취도에 차이가 있는지를 실험 실습평가와 서술형 평가에 대하여 각각 조사하였다. 이를 위해 고등학교 6개 학급을 3개씩 나눈 후 시일 차를 두고 수행평가를 실시한 다음 결과를 비교하였다. 연구 결과, 실험 실습평가 또는 서술형 평가를 활용한 수행평가 모두 평가 시기에 따른 차이는 나타나지 않았다. 이는 실험 실습평가의 경우 탐구 과정이나 실험 기능에 관한 평가가 이루어졌기 때문에 평가지가 미리 공개되더라도 영향이 없는 것으로 분석되었으며, 서술형 평가의 경우 몇 가지의 동형평가지 중 하나를 선택하여 수행평가를 실시하였기 때문으로 분석되었다.

이기영과 안희수(2005)는 7개의 중등학교에서 이루어지는 과학 교과의 수행평가가 주로 어떤 평가 유형과 채점 방식을 사용하는지 분석하였다. 또한, 일반화가능도 이론을 이용하여 과학 수행평가 점수의 신뢰도를 분석하였다. 연구 결과, 평가 유형은 크게 응답 제한형 서술형 검사 위주의 지필형 평가, 과정 평가와 결과물 평가로 나누어지는 과제형 평가로 나누어졌다. 채점은 1명의 교사가 모든 과제를 채점하거나 2명의 교사가 과제를 나누어 채점하는 방식이었고, 2명 이상의 교사가 서술형 문항을 교차하여 채점하는 경우는 없었다. 신뢰도 분석 결과, 수행 과제의 특성보다 서술형 문항의 특성이 수행평가 점수에 더 큰 영향을 미쳤으며, 학교에 따

라 채점자가 학생을 다르게 채점하는 정도는 상당한 차이가 있었다. 신뢰도는 대부분 적정 수준에 미치지 못했으며, 신뢰도의 확보를 위해 더 많은 수의 문항과 수행 과제, 채점자가 필요한 것으로 나타났다.

이효정(2009)은 7개 중학교에서 실시한 과학과 수행평가의 원점수 자료를 통하여 일반화가능도(신뢰도) 계수와 다양한 오차 요인에 의한 영향력을 분석하였다. 분석 결과, 각 학교의 수행평가 과제의 난이도는 상당히 달랐으며, 이는 학생의 수행평가 점수에 큰 영향을 미치는 것으로 나타났다. 수행평가가 적정 수준의 일반화가능도에 도달하기 위해서는 교사 1명이 채점할 때를 기준으로 적어도 12개의 수행과제 수가 필요한 것으로 분석되었다.

김형준과 유준희(2010, 2012)는 중학교의 과학 탐구 활동에 관한 수행평가에서 총체적 채점의 채점자간 불일치 정도와 유형을 분석하고, 총체적 채점과 분석적 채점의 신뢰도를 비교하였다. 연구 결과, 총체적 채점 방식은 내적 일치도가 높았으며, 분석적 채점 방식은 채점자간 신뢰도가 상대적으로 높았다. 분석적 채점 방식에서 채점척도는 3수준으로도 충분히 신뢰도를 확보할 수 있는 것으로 나타났다. 전체 240개의 총체적 채점 사례 중 4명의 채점자간 불일치가 나타나는 사례는 189개(78.8%)였다. 불일치가 나타나는 원인의 유형은 주로 채점자가 중요하게 여기는 부분의 차이, 채점자의 관대함과 엄격함에 따른 차이, 채점자가 중요하게 여기지 않고 간과한 부분에서 발생한 실수 등으로 나타났다. 이러한 불일치를 줄이려면 채점자가 사전에 중요한 과제 요소와 평가 요소에 대한 협의를 진행할 필요가 있으며, 학생의 응답에 대한 주의 깊은 판단과 채점자 간의 교차 채점이 필요하다고 주장하였다.

Kan과 Bulut(2014)는 수행평가 과제 채점에서 채점용 루브릭의 활용 여부 및 교사의 경험이 미치는 영향을 조사하였다. 연구 결과, 교사의 경험은 수행평가 과제의 채점 결과에 영향을 미치며 경험이 풍부한 교사가 더 융통성 있고 신뢰도가 높은 채점을 하는 것으로 나타났다. 이때, 채점용 루브릭의 활용은 평가의 신뢰도를 상당히 증가시켰을 뿐 아니라 교사의 경험으로 인한 채점 결과의 차이를 무시할 수 있을 정도로 줄이는 것으로 나타났다.

수행평가의 채점 및 측정에 관한 선행연구는 주로 채점의 신뢰도 향상을 중심으로 이루어졌다. 하지만 수행평가의 신뢰도를 지나치게 고려하여 수행평가 과제를

전통적 평가에 가깝게 구성하는 방식으로 개발하게 된다면 수행평가의 도입 취지에도 어긋날 뿐 아니라 평가의 타당도가 낮아지므로 그 의미에 한계가 있다. 따라서 교사가 수행평가의 구성주의적 평가 목적과 타당도를 고려하여 수행평가 과제를 개발한 다음, 평가의 신뢰도를 고려할 수 있도록 할 필요가 있다.

(4) 평가 전문성 향상

김신영과 송미영(2008)은 예비교사의 학생평가 전문성 수준을 확인한 후, 예비교사 교육과정에서 제공되는 교육평가 관련 과목과 교육실습 과정이 예비교사의 평가 전문성 신장에 실질적인 도움이 되는지 분석하였다. 연구 결과, 예비교사들의 초기 평가 전문성은 매우 부족한 수준이었다. 교육실습 경험은 그들의 평가 전문성을 유의미하게 신장시키는 것으로 나타났으나 객관적으로 기대되는 수준에 도달하지는 못하였다. 이에 연구자는 예비교사의 평가 전문성 향상을 위한 더욱 전문적이고 체계적인 교육 프로그램의 필요성을 주장하였다.

김수동과 전영석(2005)은 초보자(apprentice), 동기화(motivated), 인식(awareness), 개선(improving)의 네 단계 상태로 이루어진 AMAI 평가 전문성 모형을 제안하였다. 초보자 상태(apprentice state)에서 교사들은 일상적인 평가를 시행한다. 이때 글이나 강연, 평가 연수 등의 기회를 통하여 교사들의 평가 실태와 문제점, 다른 교사의 평가 사례 등을 접하면 해당 교사는 동기화된 상태(motivated state)가 된다. 이때 교사는 자신의 평가 실태를 파악하고 개선해야겠다는 내적 요구가 발생하며, 자신의 평가를 다른 교사의 평가와 비교하고 평가 진단 도구 등을 활용하는 등의 자기 개발에 자발적으로 나선다. 이때 동료 또는 교사 교육자 등으로부터 외부의 지원이 함께 이루어진다면 교사는 자신의 단점을 파악한 상태인 인식 상태(awareness state)가 된다. 이 상태에서 교사가 내부적으로는 지속적인 독서와 성찰, 다른 교사의 평가 사례 수집 및 참관, 토의 과정을 거치고, 외부에서 지속해서 강의 및 워크숍과 침식 지도를 제공하면 해당 교사는 개선 상태(improving state)가 된다. 이후 교사의 평가 전문성은 개선 상태에 머물지 않고 끊임없이 자신의 단점을 파악하여 동기화된 상태와 인식 상태로 돌아가 지속적으로 발전하게 된다. 그러므로 이 연구에서는 교사의 평가 전문성 향상이 이루어지려면 교사 개인의 내적 요구와 외부의

자극 또는 도움이 함께 이루어져야 함을 강조하였다.

Sato와 Coffey, Moorthy(2005)는 ‘교수학습의 개선을 위한 교실 평가 프로젝트(The Classroom Assessment Project to Improve Teaching and Learning; CAPITAL)’에 관한 연구를 수행하였다. CAPITAL 프로젝트에서는 평가를 사용하여 학습을 개선하는 방법을 두 명의 교사에게 소개하되, 실제 수업에서의 평가는 교사 자율에 맡겨두고 그들의 수업과 평가 실행이 어떻게 달라지는지 조사하였다. 그 결과, 두 교사에게서 나타난 변화는 교사 개인의 본성과 상황에 많은 영향을 받았다. 한 교사는 반성적 사고를 통해 학생들이 정답을 맞히기를 기대하는 모습에서 학생들의 이해를 향상하기 위한 모습으로 변화하였으며, 다른 교사는 동료 평가를 도입하고 형성평가가 다른 교사들에게도 확장되기를 원하였다. 즉, CAPITAL 프로젝트는 교사들에게 단순히 평가 목적과 방법을 알려주는 수준을 넘어, 그들의 신념을 변화시키는 계기가 되었다.

김동영(2006)은 과학교사가 자신의 평가 전문성을 진단하고 부족한 부분에 대해 스스로 수정 및 보완할 수 있는 평가 전문성에 대한 자기 진단 도구를 개발하여 과학교사의 평가 전문성을 향상시키고자 하였다. 이를 위해 과학교사의 평가 전문성에 관한 문헌 연구와 전문가들과의 검토 및 협의 과정을 바탕으로 과학교사의 평가에 관한 지식, 실무 능력, 평가의 실천 내용을 포함한 자기 진단 도구를 개발하였다. 또한, 교사들이 스스로 자신의 평가 전문성 수준을 판단할 수 있도록 평가 전문성 등급 척도와 안내 자료를 함께 개발하였다.

민희정(2012)은 과학교사의 평가 전문성 향상을 위한 연수 프로그램을 개발하였다. 이 프로그램은 서답형 평가와 수행평가를 두 축으로 하며 3개의 국면(phase)과 4개의 단계(step)로 이루어져 있다. 모든 국면과 단계에는 자기반성 및 에세이 작성을 통한 자기성찰 및 반성 과정을 포함시켰다. 국면 1은 서답형 평가에 관한 것으로, 교사의 평가 전문성과 서답형 평가에 관한 이론적 고찰(단계 1), 모의 서답형 평가 협의(단계 2), 개인별 서답형 문항 발표 및 토론(단계 3)으로 이루어져 있다. 국면 2는 수행평가에 관한 것으로, 수행평가의 의미 고찰(단계 1), 수행평가에 관한 토론(단계 2), 개인별 수행평가 사례 발표 및 토론(단계 3)으로 이루어져 있다. 국면 3은 국면 1, 2의 전 과정을 종합 토론을 통하여 내면화시키는 과정으로 구성하였다. 또한, 연수에 참여한 현직교사들을 대상으로 설문 조사를 실시하여 교사의 평

가 전문성의 발달과정을 7단계로 제시하였다. 각 단계의 순서는 낮은 위계부터 순서대로 교사가 평가에 관한 지식을 새롭게 습득하는 획득하기(acquiring) 단계, 변화의 필요성을 인식하는 인식하기(recognizing) 단계, 기존의 신념을 포기하고 새로운 신념을 받아들이는 깨뜨리기(breaking) 단계, 기존의 신념을 새로운 신념으로 대체하는 변화시키기(changing) 단계, 구체적인 계획을 세우지 못하고 계속된 숙고를 하는 고민하기(considering) 단계, 숙고 과정을 마치고 구체적으로 평가를 계획하기 시작하는 계획하기(planning) 단계, 계획한 내용을 실천으로 옮기는 실행하기(practicing) 단계이다.

Furtak 등(2012)은 학습 발달과정(learning progression) 이론을 평가에 적용하여 교사의 평가 전문성 향상을 위한 순환적 사이클을 제안하였다. 이 사이클은 반성 및 수정(reflect & revise), 학생의 사고 탐구(explore student thinking), 새로운 도구를 사용한 실행(practice using new tools), 제정(enact)의 네 단계로 이루어져 있다. 반성 및 수정 단계에서 교사는 자신의 평가 실재를 반성하고, 장단점을 토론한다. 학생의 사고 탐구 단계에서는 학습 발달과정의 관점에서 학생의 특정 개념에 대한 이해 수준과 정교화 과정, 각 단계에서 학생들이 일반적으로 가지는 오개념이나 대안적 개념 등에 대한 이해를 심화시킨다. 새로운 도구를 사용한 실행 단계에서 교사들은 새로운 평가도구를 사용하여 형성평가를 실행하고, 학생의 개념을 확인하고 학습 발달과정에 맞는 피드백을 논의한다. 제정 단계에서는 학습 발달과정에 따라 개발된 평가도구를 사용하여 평가를 수행하고, 평가 결과를 반성 및 수정하는 첫 단계로 돌아가는 순환적 사이클을 시행한다.

Falk(2012)는 초등학교 과학교사의 전기와 전기 회로 주제에 관한 평가 전문성 향상 프로그램을 개발하고 적용하였다. 이 프로그램은 3시간씩 8회에 걸쳐 이루어졌으며, 각 회차별 활동은 교사가 학생들의 평가 자료를 분석하는 ‘학생 자료 분석하기(Looking at Student Work; LASW)’ 활동과 교사들이 전기 회로에 관한 핵심 개념을 탐구하고 논의하는 ‘과학 조사(science investigation)’ 활동으로 이루어졌다. 한 차시의 과학 조사 활동은 다음 차시의 LASW 활동을 위한 내용으로 구성되었으며, 이는 교사들이 분석하고자 하는 학생들의 평가 과제에 대한 최소한의 초기 이해를 하도록 하기 위함이다. 학생들의 평가 결과물은 전기 회로의 특정 주제와 관련한 학생들의 생각과 설명을 기록한 문제 은행이다. 교사는 각 회차의 주제

별로 선정된 학생들의 평가 과제를 분석하였다. 이때 무조건 많은 평가 과제를 분석하기보다 이전 회차의 경험을 바탕으로 새로운 면을 분석할 수 있도록 초기에는 3-4개의 과제를 분석하였고, 후반에는 9-12개의 수업 세트를 조사하였다. 그 결과, 형성평가 과제의 분석 활동이 교사의 PCK를 활용하고 계발할 기회를 제공함으로써 평가 전문성을 향상시킬 수 있다고 주장하였다.

한편, 김선희(2012)는 수학 교과에서 예비교사의 평가 전문성 신장을 위해 평가 과정 전반을 경험할 수 있는 실습과 학습일지를 통한 반성을 두 축으로 하는 평가 전문성 향상 전략을 개발하였다. 학습일지는 예비교사의 평가에 관한 지식, 학생과 교사 사이에 있는 이중적인 예비교사의 입장, 평가 실행의 세 가지 영역으로 분석되었다. 그 결과, 지식, 입장, 실행의 세 가지 영역은 서로 상호작용하며 예비교사의 평가 전문성 신장에 기여하며 평가 전문성을 구성하는 요소가 되는 것으로 나타났다.

제 3 장. 중등 과학교사의 과학 평가 실태와 지향 (연구 D)¹⁾

1. 서 론

과학교육의 패러다임으로 자리 잡은 구성주의 학습관에 따르면, 학생들은 객관적인 과학 지식을 단순히 수용하는 것이 아니라 능동적으로 지식을 구성하고 교사나 다른 학생과의 상호작용을 통해 의미를 조정하는 순환적이고 누적적인 학습 과정을 거친다(조희형, 최경희, 2002). 교육현장에서 구성주의 학습이 구현되기 위해서는 이에 부합하는 평가 활동이 수반되어야 한다(Jonassen, Peck, & Wilson, 1999). 평가는 학생의 지식이나 성취를 판단하기 위하여 증거를 수집하고 해석하는 일련의 과정인데(Atjonen, 2014), 목표 지향적이고 정답을 요구하며 학습 후에 총괄적으로 이루어지는 전통적인 방식의 평가는 구성주의 학습을 구현하는 데 한계가 있다. 구성주의 관점에서는 주어진 문제에 관한 결과보다는 학생들이 문제를 해결하는 과정과 방법에 대한 평가를 중요시한다(Winterbottom *et al.*, 2008). 또한, 정해진 정답을 요구하는 전통적 평가와 달리 개방적 응답을 중요시하는 등 보다 다양한 방법으로 평가가 이루어진다. 즉, 구성주의적 평가는 지식 구성의 과정에 대한 평가, 단순 암기를 지양하는 고등 사고능력에 대한 평가, 다양한 관점과 양식이 존중되는 평가라고 할 수 있다(박선미, 1998).

교사는 학생들의 성취도 확인, 포상, 동기 부여 등 다양한 목적으로 평가를 실시하며, 이는 크게 성취에 대한 정보를 수집하는 총괄평가와 학습 보조를 위한 정보를 수집하는 형성평가로 구분할 수 있다(Hanna & Dettmer, 2004). 총괄평가도 교수학습 개선에 도움을 주는 구성주의적 평가로 기능할 수 있지만(박정, 2013a), 주로 평가 대상인 학생들의 성취 여부에 집중한다. 반면, 형성평가는 교사-학생 간의 상호작용과 피드백에 주목하여 학생들이 학습 과정에서 겪는 어려움을 개선하고 학습을 촉진시키므로 구성주의적 평가에 보다 가깝다고 볼 수 있다(Black & William, 1998). 형성평가는 각 학생의 상황과 수준에 맞게 개별적으로 진단과 조언

1) 연구 I은 한국과학교육학회지 35호 4권의 725-733쪽에 게재된 ‘중등 과학교사의 과학 평가 실태와 지향’의 내용을 보완하고 재구성한 것임.

을 제공하므로 학생이 스스로 학습을 조정하는 데 도움이 된다. 선행연구에서도 형성평가는 학습 목표 달성과 과학 개념 이해뿐 아니라 학습 동기나 자기 주도적 학습능력의 향상에도 효과적인 것으로 보고되었다(남정희 등, 2005; Shavelson *et al.*, 2008).

이러한 맥락에서 우리나라도 제7차 교육과정 이후 2015 개정 교육과정에 이르기까지 단순 암기식이나 선택형 평가와 지식의 양을 측정하기 위한 평가를 지양하고 다양한 평가도구와 방법을 활용한 평가와 학습과 수업의 질 개선을 위한 평가 실시를 지속적으로 강조하고 있다(교육과학기술부, 2012; 교육부, 1997). 그런데 구성주의적 평가의 대표적인 방법 중 하나인 형성평가는 학교 현장에서 구조화된 상황에서 닫힌 질문으로 실시되며 개별화된 피드백을 제공하지 않는 등 여전히 학습 결과를 확인하는 전통적인 방식으로 사용된다고 보고되었다(박영석, 2009; 조성수, 2003). 따라서 학교 현장에서 구성주의적 평가 환경을 조성하기 위하여 평가를 실시하는 교사들이 구성주의적 과학 평가관과 인식을 가질 수 있도록 하고(남정희 등, 1999; 성을선, 남정희, 최병순, 2000), 바람직한 구성주의적 평가 방법을 확산시키기 위한 방안을 모색해야 한다(김성원, 현미숙, 2005; 정은영, 최원호, 2014). 하지만 교사가 구성주의적 평가에 대한 지식을 가지고 있는 것과 실제로 구성주의적 평가를 실시하는 것은 별개의 문제이다. 즉, 구성주의적 평가에 대하여 알고 있는 교사들도 실제 수업에서는 결과에 초점을 두는 전통적 평가를 실시할 수 있으므로(Thomas *et al.*, 2011), 이상과 실제 사이의 간극이 존재할 가능성이 있다. 그러므로 교사들은 구성주의적 평가 목적에 대한 올바른 인식을 바탕으로 다양한 평가 방법들을 적절하게 활용할 능력을 갖추어야 할 것이다.

그러나 지금까지 국내에서 이루어진 교사의 과학 평가 실태에 대한 연구는 교사가 제작한 평가 문항의 특성과 문제점을 분석하거나(김석우, 2007; 김윤희, 윤기순, 권덕기, 2010; 양일호 등, 2008; 오현석, 이기영, 2006), 교사의 평가 목적과 방법, 사용 시기 등을 단편적으로 분석하는 경우(김동진, 2005; 남정희 등, 1999; 성을선, 남정희, 최병순, 2000; 정은영, 최원호, 2014)에 제한되어 있었다. 이로 인해 다양한 평가 상황에서 평가 목적과 방법의 부합성 및 평가 지향과 실제의 일치성 등에 대한 총체적인 정보는 아직 부족한 실정이다. 즉, 실험, 토의, 자유탐구 등과 같은 여러 평가 상황에서 평가 목적에 맞게 평가 방법을 적절하게 사용하는지, 자신

의 평가 지향에 맞게 실제로 평가 목적을 설정하고 방법을 계획 및 실행하는지에 대한 구체적인 정보가 부족하다고 할 수 있다.

이에 연구 I에서는 중등 과학교사들의 과학 평가 실태와 지향의 특징을 구성주의적 평가 방법과 목적 측면에서 조사하였다. 즉, 중등 교사의 과학 평가 실태를 조사하기 위하여 교사들이 실시한 대표적인 과학 평가 사례를 평가 목적과 방법 측면으로 나누어 분석한 후 비교하였다. 다음으로, 구성주의적 평가에 대한 교사들의 지향을 조사하기 위하여 학교에서 발생할 가능성이 높은 전형적인 수행평가 상황을 제시하고 이 상황에 대하여 교사가 지향하는 평가 사례를 평가 목적과 방법 측면으로 나누어 분석하고, 분석 결과를 과학 평가의 실태에서의 결과와 비교하였다.

한편, 선행연구에서는 교사의 교직 경력, 평가 연수 경험, 성별 등이 교사의 평가 실행에 영향을 미칠 수 있는 변인으로 제안되었다(강대중, 염시창, 2014; 김동진, 2005; 김신영, 2002; 박정, 2013b; 이정애, 2003; Thomas *et al.*, 2011). 또한, 평가에 대한 교사의 인식도 구성주의적 평가 실행에 직접적인 영향을 미치는 것으로 보고되었다(Carless, 2005; Sato, Coffey, & Moorthy, 2005; Tierney, 2006). 따라서 연구 I에서는 교사의 배경 변인 중에서 교사의 성별, 학교급, 교직 경력, 평가 관련 교육 경험, 구성주의적 과학 평가관을 선택하여 과학 평가의 실태와 지향에서의 차이를 평가 목적 측면에서 비교하였다.

2. 연구 방법

2.1 연구 참여자

서울과 경기 지역의 중·고등학교 과학교사 92명이 연구에 참여하였다. 교직 경력 5년 이하의 초임교사가 34명(37.0%)이었고, 교원 임용 후 각종 연수, 컨설팅, 대학원 수업 등에서 평가 관련 교육을 받은 경험이 있는 교사는 51명(55.4%)이었다. 연구 참여자의 구체적인 배경 변인은 <표 I-1>과 같다.

<표 I-1> 연구에 참여한 교사의 배경

	배경	빈도 (%)
성별	남성	20 (21.7)
	여성	72 (78.3)
학교급	중학교	43 (46.7)
	고등학교	49 (53.3)
교직 경력	5년 이하	34 (37.0)
	5년 초과	58 (63.0)
최종 학위	학사	47 (51.1)
	석사 이상	45 (48.9)
평가 관련 교육 경험	있음	51 (55.4)
	없음	41 (44.6)

2.2 연구 절차

중등 과학교사의 과학 평가 실태와 지향을 조사하기 위하여 설문지를 개발하였다. 선택형 문항을 사용할 경우 선택지가 교사들의 응답에 영향을 미칠 수 있으므로, 자신의 생각을 자유롭게 기술하는 서술형 문항을 사용하였다. 먼저, 학교에서 현재 실시하는 과학 평가의 실태를 조사하기 위하여 교사가 지금까지 실시했던 가능한 모든 평가 사례를 목적과 방법에 초점을 두어서 자세히 기술하도록 하였다. 예비 검사 결과, 교사들이 평가의 목적과 방법을 기술하는 방법에 어려움을 겪는 것으로 나타났다. 이에 평가의 목적과 방법을 기술하는 방식에 대한 이해를 돕기 위해 ‘영재교육원에 추천할 학생을 선발하기 위한 목적으로 지필평가를 활용하여 학생들의 과학 탐구력을 평가하였다.’ 라는 문구를 예로 제시하였다. 다음으로, 과학 평가에 대한 교사들의 지향을 조사하기 위하여 결정적 사건 기법(critical incident technique; Shapira-Lishchinsky, 2011)을 활용하여 가상의 전형적인 수행평가 상황에 대하여 교사들이 활용할 것이라고 응답한 평가 사례를 대상으로 평가의 목적과 방법을 분석하였다. 결정적 사건 기법은 참여자의 관점에서 활동을 기술하므로 연구 대상의 생각을 구체적으로 밝힐 수 있는 질적 연구의 장점과 많은 대상

을 연구할 수 있는 양적 연구의 장점을 동시에 지니고 있기 때문에(Gremler, 2004), 구성주의적 평가에 대한 교사들의 평소 선호나 지향을 조사하려는 이 연구에 적합하다. 이 연구에서는 구성주의적 평가의 실시가 가능한 전형적인 수행평가 상황 세 가지를 결정적 사건으로 제시하였는데, 과학실에서의 실험 상황, 사회적 이슈에 대한 토의 상황, 자유탐구 상황이었다. 교사들에게 각 상황을 묘사하는 간단한 제시문을 읽은 뒤, 그 상황에서 자신이 실시할 평가의 목적과 방법을 자세히 기술하도록 하였다.

교사들의 구성주의적 과학 평가관을 조사하기 위하여 James와 Pedder(2006)의 과학 평가관 검사 30문항을 번역한 후 요인 분석을 통하여 우리나라의 상황에 맞도록 타당화 과정을 거친 노태희와 윤지현, 강석진(2009)의 과학 평가관 검사 문항을 사용하였다. 구성주의적 과학 평가관 검사 문항이 과학 평가의 실태와 지향에 대한 교사들의 응답에 영향을 미칠 가능성이 있으므로, 다른 검사를 실시한 후 마지막으로 이 검사를 실시하였다.

연구 대상이 아닌 과학교사 3인을 대상으로 예비 검사를 실시하고, 문항에 대한 이해도를 점검한 후 설문지를 수정하였다. 또한, 과학교육 전문가 3인으로부터 문항에 대한 안면타당도를 점검받았다. 완성된 설문지(부록 I-1)는 온라인 설문지 형태로 제작한 후, 연구 대상 교사들에게 전자 우편과 전화를 통하여 설문 참여를 요청하고 온라인으로 응답을 수집하였다.

2.3 분석 방법

교사들의 응답 중 불성실하거나 부적절한 응답을 제외한 후, 과학 평가의 실태에 대해서는 245개(1인당 평균 2.7개), 과학 평가의 지향에 대해서는 세 가지 상황별로 최대 126개(1인당 평균 1.4개)의 사례를 분석 대상으로 하였다. 교사들이 응답한 과학 평가 사례에 나타난 목적은 크게 총괄평가와 형성평가로 구분한 후, Davis와 Neitzel(2011)의 연구를 바탕으로 세분하였다. 교사들이 응답한 과학 평가 사례의 방법은 크게 측정 방식(measurement mode), 수행 방식(performance mode), 비형식 방식(informal mode)의 3가지 유형으로 구분하였다(Wang, Kao, & Lin, 2010). 측정은 전통적인 지필평가 방법이고, 수행은 관측 가능한 행동이나 사고 과정을 활

용한 평가 방법이며, 비형식은 관찰, 면담 등을 사용하는 평가 방법이다. 평가 방법은 백순근(2000) 및 Hanna와 Dettmer(2004)의 연구를 바탕으로 세분하였는데, 이 연구에서는 동료 평가나 자기 평가와 같이 교사가 직접 정보를 수집하지 않는 평가 방법도 비형식 유형으로 분류하였다.

교사들의 응답을 분석하여 귀납적으로 분류 기준을 설정하였고, 2인의 분석자가 교사들의 응답 중 일부를 이 기준에 따라 각자 분석한 뒤 일치도를 확인하였다. 분석자 간의 의견이 일치하지 않는 응답은 논의를 통하여 분류 기준을 수정하였다. 수정한 기준을 바탕으로 다시 분석자 간 일치도를 확인하여 일치도가 .96에 도달한 뒤, 연구자 중 1인이 모든 응답을 다시 분류하였다.

교사의 배경 변인에 따른 형성평가의 사용 정도의 비교에서는 중복 응답에 의한 효과를 통제하기 위하여 총괄평가는 0점, 형성평가는 1점을 부여하였고, 여러 가지 사례를 중복으로 응답한 경우에는 평균 점수를 부여하였다. 예를 들어, 한 명의 교사가 총괄평가 사례 두 가지와 형성평가 사례 한 가지를 기록한 경우, 각각 0점, 0점, 1점을 부여하여 해당 교사에게 0.33점을 부여하였다. 이후 성별(남성/여성), 학교급(중학교/고등학교), 교직 경력(초임교사/경력교사), 최종 학위(학사/석사 이상), 평가 관련 교육 경험(유/무), 구성주의적 과학 평가관 수준(상위/하위)을 독립변인으로 사용하여 독립표본 t 검정을 실시하였다. 교직 경력은 5년을 기준으로 초임교사와 경력교사로 구분하였고(Stansbury & Zimmerman, 2000), 구성주의적 과학 평가관 수준은 과학 평가관 검사 점수의 중앙값(3.36)을 기준으로 구분하였다(상위 48명, 하위 44명). 모수 통계를 위한 정규성 가정의 경우 집단의 크기가 30 이상인 경우에는 중심극한정리에 의하여 정규성 가정을 만족하며, 12 이상인 경우에는 유의수준을 .01로 더 엄격하게 적용함으로써 이에 대한 보정이 가능하다(Clinch & Keselman, 1982; Tan, 1982). 이 연구에서는 모든 경우에 대하여 집단의 크기가 12 이상이었으므로 t 검정 결과를 먼저 제시하고, Levene의 등분산 검정을 통하여 등분산성 가정을 만족하지 못하는 경우에는 Welch-Aspin 검정을 실시하였다. 이때 집단의 크기가 30 미만인 경우에는 Kolmogorov-Smirnov 검정을 통해 정규성 검정을 실시한 뒤 이를 만족하지 못한 경우에는 Mann-Whitney U 검정을 실시하여 결과를 본문에 제시하였다.

3. 연구 결과 및 논의

3.1 중등 과학교사의 과학 평가 실태

(1) 평가 목적

중등 교사들이 제시한 과학 평가 사례의 목적을 분석한 결과를 <표 I-2>에 제시하였다. 학습 목표 달성에 대한 정보를 수집하는 총괄평가가 전체의 88.2%를 차지하였는데, 학습 수준 측정, 학생 선발, 학습 부진아 확인, 학생의 수업 참여도 측정, 학습 노력에 대한 보상 등 다섯 가지 하위 유형으로 세분할 수 있었다.

<표 I-2> 교사가 실시하는 평가 목적

평가 목적		빈도 ^a (%)
총괄평가	학습 수준 측정	134 (54.7)
	학생 선발	74 (30.2)
	학습 부진아 확인	3 (1.2)
	수업 참여도 측정	3 (1.2)
	학습 노력에 대한 보상	2 (0.8)
소계		216 (88.2)
형성평가	학습 기회 제공	23 (9.4)
	피드백 제공	-
소계		23 (9.4)
기타	시험 대비	6 (2.4)
소계		6 (2.4)
계		245 (100.0)

^a 중복 응답.

가장 많았던 유형은 전체의 54.7%를 차지한 학습 수준 측정이었으며, 중간고사나 기말고사와 같은 성취도 평가와 논술 문항으로 과학적 탐구력을 평가하는 경우

가 대표적인 예이다. 학생 선발 목적도 전체의 30.2%를 차지하였는데, 하위 유형으로 과학 관련 대회의 수상자 선정, 과학 관련 대회에 참가할 대표학생 선발, 영재 학급이나 과학 동아리 추천 학생 선발 등이 있었다. 한편, 학습 부진아 확인, 학생의 수업 참여도 측정, 학습 노력에 대한 보상을 목적으로 하는 사례는 소수였다.

과학 평가의 목적이 대부분 학습 수준 측정이나 학생 선발 등의 총괄평가에 치우친 결과는 교사의 평가가 대부분 전통적인 방식으로 이루어진다는 선행연구 결과(Brown, 2004; Wang, Kao, & Lin, 2010)와 유사하다. 특히 우리나라에서는 평가의 가장 중요한 목적이 상급 학교 진학에 필요한 정보 제공이라는 인식이 지배적이므로, 이러한 현상이 심화되었을 수 있다. 즉, 학생들은 과학에 대한 흥미와 같은 내적 동기보다는 시험 대비나 상급 학교 진학과 같은 외적 동기에 의해 과학을 공부하는 경우가 많으므로(곽영순 등, 2006), 교사들도 자연히 평가의 공정성과 객관성을 중시하고 정확하고 엄밀한 성취도 측정을 평가의 초점으로 삼았을 가능성이 높다.

교사와 학생의 상호작용과 피드백을 강조하는 형성평가는 전체의 9.4%로 적었다. 형성평가에 해당하는 평가 사례는 아래와 같이 주로 학생의 과학 개념 이해나 과학적 사고를 내면화하기 위한 기회 제공을 목적으로 하였다. 즉, 평가 활동을 교수-학습의 도구로 활용함으로써 과학 개념의 이해, 흥미 유발, 협동성과 발표 능력의 향상 등의 목표를 달성하려는 구성주의적 평가를 시도하였다.

원자 모형의 변천을 역사적 관점에서 이해할 수 있도록 하기 위해 원자 모형 변천에 기여한 과학자들의 삶을 조사하고 조사한 내용을 바탕으로 다양한 형태의 발표 자료를 만들어 발표하게 하였다. 친구들의 발표를 듣고 발표 내용과 전달력 등을 기준으로 동료 평가를 하고, 조사 자료의 충실성 및 발표의 전달력을 토대로 학생 과제를 평가하였다.

실제 형성평가에 해당하는 사례는 적었지만 교사들의 응답에서 형성평가라는 용어는 적지 않게 등장한다. 그러나 아래의 예에서 볼 수 있듯이, 교사들이 언급한 형성평가는 대부분 구성주의적 평가가 아닌 중간고사나 기말고사 이외의 단원평가나 수업의 정리 단계에서 학생의 목표 도달 정도를 파악하는 평가와 같은 실질적인 의미에서의 작은 총괄평가(Popham, 2009)에 해당한다. 선행연구에서도 교사들이

수업의 정리 단계에서 실시하는 평가를 형성평가라고 생각한다는 결과가 보고된 바 있다(박영석, 2009).

진행되는 과학 수업의 내용에 대한 학습 정도를 파악하기 위해 객관식+서술형 형성 평가 문제를 출제하고 평가를 실시하였다.

총괄평거나 형성평가 이외에 시험을 대비하기 위한 목적의 평가(2.4%)가 있었다. 수학능력시험, 학업 성취도 평가, 경시대회 등을 대비하기 위하여 기출 문제나 예상 문제를 제시하여 학생들에게 문제 유형과 형식에 관한 정보를 제공하는 목적의 평가가 이 유형에 해당한다.

3학년의 경우, 수능을 준비하고자 수업 시간에 복습한 내용에 해당되는 기출문제를 뽑아 매 시간 20문제씩 풀게 했으며 그 결과를 직접 평가할 수 있도록 하였다.

평가 사례에 나타난 형성평가의 사용 정도를 교사의 배경 변인에 따라 분석하기 위하여 독립표본 t 검정을 실시하였다(표 I-3). 평가 목적 분류에서 기타 유형에 해당하는 사례는 소수였으므로 분석에서 제외하였다.

<표 I-3> 교사의 배경 변인에 따른 형성평가의 사용 정도에 대한 독립표본 *t* 검정 결과

배경		<i>n</i>	<i>M</i> (<i>SD</i>)	<i>t</i>	<i>p</i>
성별	남성	19	.132 (.23)	1.050	.305
	여성	71	.074 (.14)		
학교급	중학교	43	.100 (.18)	1.297	.199
	고등학교	47	.074 (.12)		
교직 경력	5년 이하	33	.051 (.13)	-1.60	.113
	5년 초과	57	.107 (.20)		
최종 학위	학사	46	.099 (.19)	.725	.470
	석사 이상	44	.074 (.14)		
평가 관련 교육 경험	있음	50	.119 (.22)	2.07	.042*
	없음	40	.046 (.11)		
구성주의적 과학 평가관	상위	46	.110 (.21)	1.28	.204
	하위	44	.062 (.15)		

* $p < .05$, ^a 평균 점수(총괄평가 0점, 형성평가 1점).

분석 결과, 성별, 학교급, 교직 경력, 최종 학위와 구성주의적 과학 평가관 수준에 따라서는 유의미한 차이가 없었지만, 평가 관련 교육 경험이 있는 교사들은 그렇지 않은 교사들에 비해 형성평가 목적이 통계적으로 유의미하게 많았다. 집단의 크기가 30 미만인 성별 변인에 대해서는 정규성 검정 및 Mann-Whitney U 검정을 별도로 실시하였으며, 유의미한 차이는 없었다($U=663.5$, $p=.559$). 이러한 결과는 평가에 대한 강의나 연수가 구성주의적 과학 평가의 실행에 기여할 가능성을 의미한다고 볼 수 있다(박가나, 2012). 그러나 형성평가 목적을 지닌 평가 사례가 전체의 10%에도 미치지 못하는 점을 고려할 때, 단순히 평가 관련 교육 경험을 늘리는 것 보다는 평가 관련 강의나 연수에서 구성주의적 평가를 확산시킬 수 있는 방안에 대한 고려가 더 중요할 것이다.

(2) 평가 방법

중등 교사들이 제시한 과학 평가 사례의 방법을 분석한 결과를 <표 I-4>에 제시하였다. 분석 결과, 측정(54.3%)과 수행(40.7%) 방식의 평가가 대부분을 차지하였으며, 비형식 방식의 평가(4.9%)의 비율은 낮았다.

<표 I-4> 교사가 실시하는 평가 방법

평가 방법		빈도 ^a (%)
측정 방식	선택형 평가	123 (50.6)
	서답형 평가	9 (3.7)
소계		132 (54.3)
수행 방식	실험 평가	37 (15.2)
	보고서 평가	23 (9.5)
	논술형 평가	19 (7.8)
	발표 평가	13 (5.3)
	포트폴리오 평가	4 (1.6)
	토의 평가	3 (1.2)
	소계	99 (40.7)
비형식 방식	관찰 평가	6 (2.5)
	동료 평가	4 (1.6)
	면담 평가	2 (0.8)
	자기 평가	-
소계		12 (4.9)
계		243 (100.0)

^a 중복 응답.

측정 방식의 평가는 제시된 선택지에서 정답을 고르는 선택형 평가와 단어나 문장으로 답을 쓰는 서답형 평가로 나눌 수 있다. 교사들은 선택형 평가(50.6%)를 주로 사용하였고, 서답형 평가(3.7%)는 아래의 예와 같이 학생들의 이해도를 심층적으로 평가할 때 사용하였다.

교내 화학 경시 대회 수상자를 선발하기 위해 중, 고등학교 교육과정 범위 내에서 암기가 필요한 내용보다 전체적인 화학 교과 이해를 필요로 하는 내용 중심으로 서술형 문제를 출제하여 전반적인 교과 이해도를 평가하였다.

수행 방식에서는 실험(15.2%)과 보고서 평가(9.5%)의 빈도가 높았다. 선행연구(김성원, 현미숙, 2005)에서도 과학 교과에서 지필평가를 제외할 경우 실험과 보고서 평가가 가장 흔하게 사용된다고 하였다. 과학 독후감이나 글짓기를 활용한 논술형 평가(7.8%), 문제의 답을 설명하는 구술시험이나 탐구 활동의 결과를 조별로 발표하는 발표 평가(5.3%)도 적지 않게 나타났다. 기록을 체계적으로 누적한 자료를 이용한 포트폴리오 평가는 결과물에 이르기까지의 전 과정을 누적하여 평가하므로 학생의 변화 과정을 종합적으로 파악할 수 있는 구성주의적 평가 방법이다(김찬중, 2012). 특정 주제에 대한 학생들의 토의 과정을 보고 평가하며 필요한 경우 피드백을 제시해줄 수 있는 토의 평가 또한 학습과 평가 활동이 통합적이고 반복적으로 이루어지는 구성주의적 평가 방법이라 할 수 있다(백순근, 2000). 그러나 이 연구에서는 포트폴리오 평가와 토의 평가의 빈도가 매우 낮게 나타났다.

비형식 방식의 평가에는 관찰 평가, 동료 평가, 면담 평가가 있었다. 학생이 평가 대상이 아니라 평가 과정의 주체로 참여하는 동료 평가나 자기 평가는 대표적인 구성주의적 평가이지만(초한무, 2001), 이 연구에서는 발생 비율이 매우 낮았다.

3.2 중등 과학교사의 과학 평가 지향

(1) 평가 목적

실험, 토의, 자유탐구의 세 가지 전형적인 과학 수행평가 상황에 대한 교사들의 응답에서 평가의 목적을 분석한 결과는 <표 I-5>와 같다.

<표 I-5> 전형적인 과학 수행평가 상황에서 교사가 지향하는 평가 목적

평가 목적	평가 상황 ^a (%)			
	실험	토의	자유탐구	
총괄평가	학습 수준 측정	84 (77.8)	53 (45.7)	45 (38.1)
	학생 선발	-	-	1 (0.8)
	학습 부진아 확인	-	-	-
	수업 참여도 측정	9 (8.3)	23 (19.8)	21 (17.8)
	학습 노력에 대한 보상	-	1 (0.9)	2 (1.7)
소계		93 (86.1)	77 (66.4)	69 (58.5)
형성평가	학습 기회 제공	12 (11.1)	39 (33.6)	34 (28.8)
	피드백 제공	3 (2.8)	-	15 (12.7)
	소계	15 (13.9)	39 (33.6)	49 (41.5)
계		108 (100.0)	116 (100.0)	118 (100.0)

^a 중복 응답.

실험 상황에서는 총괄평가가 86.1%로 대부분을 차지하였고, 실태에서의 결과(88.2%)와 비율이 비슷하였다. 교사들의 응답은 대부분 실험 과정이나 결과를 바탕으로 과학적 탐구력이나 실험 설계 능력의 정확도 측정을 지향하였다(77.8%).

혼합물 분리를 수업시간에 배운 내용을 적용하여 실제로 수행할 수 있는지 평가한다. 학생들이 수업시간에 배운 사물들을 이용하여 혼합물을 만들어 섞은 후 단계를 거쳐 분리하게 한다. 모두 분리하면 만점, 분리를 못한 경우 단계별로 점수를 차등하여 감점한다. 실험 보고서에 그 과정을 쓰고 원리를 설명하도록 하여 이를 함께 평가한다.

반면, 학생의 실험 과정을 확인하여 적절한 기술 피드백을 제공(2.8%)하거나 평가 결과를 학습 기회 제공 목적(11.1%)으로 활용하는 형성평가는 13.9%에 불과하였다.

평가는 혼합물 분리라는 결과뿐 아니라, 실험 과정 동안 어떻게 생각(가설)이 변하

였는지 그리고 피드백을 통해 어떻게 실험 방법 등을 변화시켰는지 등을 스토리텔링 형식으로 발표하도록 하여 평가할 것이다.

토의 상황에서는 총괄평가의 비율이 다소 감소하였지만, 여전히 높은 비율(66.4%)을 차지하고 있었으며, 그중에서도 학습 수준 측정(45.7%) 목적의 비율이 가장 높았다. 그런데 총괄평가 중 학생의 수업 참여도를 측정하기 위한 목적(19.8%)이 증가하였다. 즉, 토의 상황에서 교사들은 토의의 과정과 결과뿐 아니라 학생들의 토의 참여도나 기여도도 평가의 중요한 목적으로 생각하였다.

평가 목적은 주어진 주제를 정확히 파악한 뒤 올바른 해결 방안을 제시할 수 있는 것이다. 위의 과정에서 평가를 할 때는 교사의 주관에 따라 좋은 의견으로 평가받기도 아니기도 하다고 생각한다. 그래서 조별 토의 결과를 교사의 주관으로 판단하는 것이 아니라 좋은 의견을 수렴하는 과정도 학생들이 고민해보고 결정하도록 하는 게 좋다고 생각한다. 그리고 평가 시에는 토의에 적극적으로 참여하는 참여도도 평가에 포함되어야 한다고 생각한다.

토의 상황에서 형성평가의 비율은 33.6%로 증가하였다. 토의 상황에서 교사들은 학생에게 발언권을 주거나 역할을 부여함으로써 평가와 학습을 연계하려고 시도하였다.

학생들의 토론을 관찰하면서 적극적으로 발표에 임하는 학생들을 체크리스트를 통해 체크하면서 평가한다. 잘 참여하지 않는 학생들에게 발언권을 주면서 토의에 참여하도록 독려하며, 그렇게 해서 잘 참여하는 학생도 점수를 줄 것이다.

자유탐구 상황에서도 총괄평가의 비율은 58.5%로 가장 높았는데, 자유탐구는 비교적 오랜 기간 평가가 지속되어야 하므로 무임승차를 방지하기 위해 학생의 기여도를 측정하려는 목적이 상대적으로 많았다. 한편, 실험이나 토의 상황과 비교하면 형성평가의 비율이 41.5%로 높았으며, 학습 기회 제공 목적(28.8%) 뿐 아니라 기술 피드백 제공 목적(12.7%)도 나타났다. 이는 자유탐구가 주제 선정 단계에서부터 탐구의 진행 상황을 주기적으로 평가하면서 적절한 피드백을 제공해야 하는 성격을

지니고 있기 때문으로 보인다.

스스로 탐구할 수 있는 능력과 협동 능력, 흥미 고취 등을 위해 소집단별로 한 탐구 활동의 과정들을 시기별로 4단계 정도로 분할하여 단계별로 어디까지 와 있는지 교사와 함께 검토하고, 필요하면 주제도 조금씩 바꿀 수 있는 여지를 준다.

전체적으로, 교사들이 구성주의적 평가를 활용할 것이라고 예상하여 제시한 전형적인 수행평가 상황에서도 약 60%의 평가 사례가 전통적인 목적의 총괄평가로 나타난 것으로 보아, 중등 과학교사들이 과학 평가에서 전통적인 목적을 지향하는 경향이 있음을 알 수 있다. 즉, 평가 실태에서 구성주의적 평가가 거의 나타나지 않았던 것은 여러 가지 제한점으로 인하여 우리의 교육에서 구성주의적 평가 환경이 조성되지 못했기 때문일 수도 있지만, 교사들의 과학 평가에 대한 인식이 전통적인 관점에서 벗어나지 못하기 때문일 수도 있음을 시사한다.

한편, 교사들이 지향하는 총괄평가와 형성평가의 비율은 상황에 따라 다르게 나타났다. 즉, 실험 상황에 비해 토론과 자유탐구 상황에서 형성평가 비율이 증가하였다. 실험은 과학에서 흔히 사용되는 평가 상황 중 하나이지만(김성원, 현미숙, 2005), 학교 실험은 대부분 과학 이론을 확인하는 목적으로 활용되므로 실험의 방법과 결과가 정해진 경우가 많다(Chinn & Malhotra, 2002). 이러한 상황에서 교사는 학생들에게 유의미한 피드백을 제공하기보다는 실험 과정이나 결과의 정확도를 측정하는 평가를 실시하게 된다. 따라서 실험 평가를 학습의 일부로 활용하는 구성주의적 평가의 비율이 다른 상황에 비해 낮았던 것으로 보인다. 반면, 토의나 자유탐구는 형성평가를 수행하기에 적절한 상황임을 교사들이 비교적 잘 인지했기 때문에 그 상황에서 형성평가의 비율이 높게 나타난 것으로 보인다. 토의나 자유탐구는 학교 현장에서 자주 실시하지 않는 평가 상황이므로, 교사들이 지향하는 이상적인 관점이 상대적으로 높게 나타났다고 볼 수 있다.

각 평가 상황에서 교사가 지향하는 평가에 나타난 형성평가의 사용 정도를 교사의 배경 변인에 따라 비교하기 위하여 독립표본 t 검정을 실시하였다(표 I-6).

<표 I-6> 전형적인 수행평가 상황에서 교사의 배경 변인에 따른 형성평가의 지향 정도에 대한 독립표본 *t* 검정 결과

상황	배경	<i>n</i>	<i>M</i> (<i>SD</i>)	<i>t</i>	<i>p</i>	
실험	성별	남성	19	.184 (.38)	1.072	.295
		여성	70	.085 (.24)		
	학교급	중학교	42	.036 (.17)	-2.43	.018*
		고등학교	47	.170 (.33)		
	교직 경력	5년 이하	33	.091 (.23)	-.41	.327
		5년 초과	56	.116 (.30)		
	최종 학위	학사	45	.067 (.20)	-1.382	.171
		석사 이상	44	.148 (.33)		
	평가 관련 교육 경험	있음	51	.088 (.26)	-.73	.186
		없음	38	.131 (.30)		
	구성주의적 과학 평가관	상위	47	.170 (.33)	2.43	.018*
		하위	42	.036 (.17)		
토의	성별	남성	20	.300 (.44)	-.697	.488
		여성	70	.379 (.45)		
	학교급	중학교	42	.369 (.46)	.16	.875
		고등학교	48	.354 (.44)		
	교직 경력	5년 이하	34	.368 (.43)	.11	.384
		5년 초과	56	.357 (.45)		
	최종 학위	학사	46	.391 (.12)	.658	.512
		석사 이상	44	.330 (.44)		
	평가 관련 교육 경험	있음	51	.294 (.43)	-1.65	.387
		없음	39	.449 (.46)		
	구성주의적 과학 평가관	상위	46	.380 (.45)	.42	.675
		하위	44	.341 (.44)		

* $p < .05$, ^a 평균 점수(총괄평가 0점, 형성평가 1점).

<표 I-6> 전형적인 수행평가 상황에서 교사의 배경 변인에 따른 형성평가의 지향 정도에 대한 독립표본 *t* 검정 결과 (계속)

상황	배경	<i>n</i>	<i>M</i> (<i>SD</i>)	<i>t</i>	<i>p</i>	
자유탐구	성별	남성	18	.361 (.48)	-.647	.520
		여성	67	.440 (.46)		
	학교급	중학교	39	.410 (.47)	-.244	.808
		고등학교	46	.435 (.45)		
	교직 경력	5년 이하	33	.515 (.46)	1.47	.791
		5년 초과	52	.365 (.46)		
	최종 학위	학사	42	.464 (.47)	.806	.422
		석사 이상	43	.384 (.45)		
	평가 관련 교육 경험	있음	48	.438 (.47)	.55	.548
		없음	37	.405 (.45)		
구성주의적 과학 평가관	상위	43	.535 (.47)	2.32	.023*	
	하위	42	.310 (.43)			

* $p < .05$, ^a 평균 점수(총괄평가 0점, 형성평가 1점).

분석 결과, 성별, 최종 학위, 교직 경력과 평가 관련 교육 경험은 교사들의 형성평가에 대한 지향에 영향을 미치지 못하는 것으로 나타났다. 집단의 크기가 30 미만인 성별 변인에 대해서는 정규성 검정 및 Mann-Whitney U 검정을 별도로 실시하였으며, 모든 상황에서 유의미한 차이는 없었다($U=603.0$, $p=.312$; $U=768.5$, $p=.455$; $U=662.5$, $p=.481$). 그러나 학교급과 교사들의 구성주의적 과학 평가관 수준에 따라서는 구성주의적 평가에 대한 지향에서 유의미한 차이가 나타났는데, 실험 상황에서만 고등학교 교사들이 중학교 교사보다 형성평가를 지향하는 비율이 더 높았고, 실험과 자유탐구 상황에서 구성주의적 과학 평가관 수준이 높은 교사가 형성평가를 지향하는 비율이 더 높았다. 학교급에 따른 차이는 고등학교 교사가 중학교 교사보다 실험 상황의 수행평가를 어려운 개념 이해의 방법으로 활용하려는 지향이 더 높았기 때문에 나타났다. 한편, 구성주의적 과학 평가관에서 나타난 결과는 구성주의적 평가에 대한 교사의 이해와 신념이 구성주의적 평가 실행에 중요하다는

선행연구(Sato, Coffey, & Moorthy, 2005)의 주장을 뒷받침한다. 즉, 구성주의적 평가관을 지닌 교사들일수록 구성주의적 평가를 시행하려는 지향이 높으므로, 교사들이 구성주의적 평가관을 지닐 수 있도록 지속해서 노력해야 할 것이다. 예를 들어 교직 경력과 평가 관련 교육 경험이 구성주의적 평가 지향에 별 영향을 미치지 못하는 것으로 나타났으므로, 학교 현장이나 예비 및 현직교사의 교육과정에서의 적극적인 개선 노력이 요구된다. 또한 토의 상황에서는 구성주의적 과학 평가관 수준의 영향이 적었으므로, 교사들에게 토의 상황에서의 형성평가 실행 방안에 대한 실천적인 교육 기회를 제공하기 위한 노력도 필요하다.

(2) 평가 방법

실험, 토의, 자유탐구의 전형적인 과학 수행평가 상황에서 평가의 방법을 분석한 결과를 <표 I-7>에 제시하였다.

<표 I-7> 전형적인 과학 수행평가 상황에서 교사가 지향하는 평가 방법

평가 방법		평가 상황 ^a (%)		
		실험	토의	자유탐구
측정 방식	선택형 평가	5 (4.1)	-	-
	서답형 평가	3 (2.5)	-	-
	소계	8 (6.6)	-	-
수행 방식	실험 평가	25 (20.7)	-	-
	보고서 평가	43 (35.5)	22 (17.5)	51 (40.5)
	논술형 평가	-	-	-
	발표 평가	5 (4.1)	9 (7.1)	25 (19.8)
	포트폴리오 평가	-	5 (4.0)	17 (13.5)
	토의 평가	1 (0.8)	19 (15.1)	-
	소계	74 (61.2)	55 (43.7)	93 (73.8)
비형식 방식	관찰 평가	36 (29.8)	44 (34.9)	8 (6.3)
	동료 평가	3 (2.5)	24 (19.0)	14 (11.1)
	면담 평가	-	-	6 (4.8)
	자기 평가	-	3 (2.4)	5 (4.0)
소계	39 (32.2)	71 (56.3)	33 (26.2)	
계	121 (100.0)	126 (100.0)	126 (100.0)	

^a 중복 응답.

실험 상황에서는 측정 방식의 평가 방법이 6.6%로 매우 적었고, 토의와 자유탐구 상황에서는 전혀 나타나지 않았다. 측정 방식의 평가 방법이 50% 이상이었던 평가 실태의 결과와 달리, 전형적인 수행평가 상황에서는 거의 모든 교사가 수행(61.2%)이나 비형식(32.2%) 방식의 평가를 지향하였다. 또한, 실태 분석에서는 거의 나타나지 않았던 논술형 평가, 발표, 포트폴리오, 토론과 같은 수행 방식의 평가와 관찰, 자기 평가, 동료 평가, 면담 등의 비형식 평가도 낮지 않은 비율을 차지하였다. 평가 방법은 제시 상황에 따라 지향하는 비율의 차이가 컸는데, 실험 상황에서는 보고서(35.5%), 관찰 평가(29.8%), 실험(20.7%)의 비율이 높았고, 토의 상황에서는

관찰(34.9%), 동료 평가(19.0%), 보고서(17.5%), 토의(15.1%)가 높은 비율을 차지했으며, 자유탐구 상황에서는 보고서(40.5%), 발표(19.8%), 포트폴리오(13.5%), 동료 평가(11.1%)에 대한 교사들의 지향이 높았다.

세 가지 상황에서 교사들이 지향한 평가 방법 중 공통적으로 많았던 것은 보고서 평가였다. 교사들은 자료 수집과 조사 활동 혹은 직접적인 실험·실습이 어려운 주제에 대하여 보고서 평가를 선호하는데(김성원, 현미숙, 2005), 보고서 평가는 지필평가로 측정하기 어려운 과학 탐구력이나 과학적 창의성을 효과적으로 측정할 수 있고, 형식을 구조화하여 채점할 수 있는 문항을 포함시키면 객관적인 점수를 산출할 수도 있기 때문이다(손정숙, 2011). 따라서 과정에 대한 점검보다 신뢰도 높은 성적 산출이 목적인 경우, 보고서 평가를 많이 실시하는 경향이 있다(김석우, 2007; 서정권, 2008). 대부분의 보고서 평가가 결과 보고서의 형태였던 이 연구의 결과도 이 맥락의 연장선상에서 해석할 수 있다.

한편, 보고서 평가 중에는 토의 주제에 대해 학생들이 미리 조사해오도록 하는 예비 보고서 평가도 일부 있었다(7회). 예비 보고서 평가는 토의 상황에서만 나타났는데, 토의 주제에 대한 기본적인 지식과 이해를 바탕으로 학생들을 토의에 적극적으로 참여시키는 것을 목적으로 하였다. 이러한 예비 보고서 평가는 원활한 토의 진행을 유도하면서 동시에 토의에 참여하는 학생의 논리적 사고력과 의사소통 능력의 향상도 추구하는 구성주의적 평가로 볼 수 있다.

평가 목적은 토론의 규칙을 준수하며 과학적인 근거를 바탕으로 논리적으로 말하는 능력의 향상과 타인을 배려하며 의사소통하는 능력의 향상입니다. 평가 방법은 1) 예비 보고서(토론 전 미리 제출) 평가, 2) 토론의 규칙 준수 여부, 3) 과학적인 근거를 바탕으로 한 논리적 말하기 여부 등으로 평가하겠습니다.

발표 평가는 주제에 대한 학생의 준비도와 이해도 외에도 표현력과 의사소통능력 등을 함께 평가할 수 있는 장점이 있으며, 발표 후 교사나 동료와의 상호작용을 기대할 수 있는 평가 방법이다. 발표 평가는 주로 자유탐구 상황에서 많이 나타났는데(19.8%), 발표가 주제 선정 단계나 계획 단계, 수행 단계 등 탐구 활동 진행 과정에서 이루어지는 경우에는 평가와 함께 피드백을 제공하는 구성주의적 평가로

나타났다. 하지만 교사들은 탐구 결과 발표 단계에서 보고서 평가를 발표 평가로 대체하거나 함께 활용하는 방법으로 발표 평가를 활용하려는 경우가 많았다.

자유탐구를 성실하게 수행했는지에 대해 평가하기 위해 주제 선정 기간과 탐구 기간을 각각 정해 그 기간 내에 과제를 수행했는지 여부를 판단하여 평가를 하고, 유의미한 탐구 활동이 되었는지 확인하기 위해 탐구 활동 수행 이후 결과를 발표하게 하여 발표에 대한 평가를 실시할 것이다.

포트폴리오는 학습의 산출물을 체계적이고 조직적으로 배열한 것으로서, 학습 과정과 성과를 누적하여 나타낸다(김찬중, 2012). 따라서 포트폴리오 평가는 학생의 성취와 능력을 입체적으로 파악할 수 있기 때문에 대표적인 구성주의적 평가 방법으로 알려져 있다(von Glasersfeld, 1993). 면담은 흔히 선발을 목적으로 하는 평가 방법이지만 대화를 통하여 정확한 의사 전달과 즉각적인 피드백이 가능하다는 장점이 있다. 따라서 학생들이 학습 과정에서 느끼는 어려움을 효과적으로 개선할 수 있는 구성주의적 평가로 활용할 수 있다(Thorndike, 1997). 이 연구에서도 포트폴리오 평가와 면담은 교사가 과정 평가를 지향하는 경우에 구성주의적으로 활용할 예정이다라고 했으나, 그 빈도는 상대적으로 낮았다.

평가의 대상은 결과뿐 아니라 과정을 중요시할 것입니다. 중간 과정에서 가설을 어떻게 수정해갔으며 피드백을 통한 구체적 실천의 변화가 어떻게 진행되었는지를 수시로 조별 상담 등을 통해 확인해 가며 평가할 것입니다. 또한, 결과물이 나왔을 때 다른 사람들이 잘 이해할 수 있도록 표현할 수 있는지 커뮤니케이션 능력을 평가할 것입니다.

관찰 평가는 과학 수업에서의 활동을 평가하는 유용한 구성주의적 평가 방법인데(손정숙, 2011), 이 연구에서는 실험과 자유탐구 상황에서 많이 나타났다. 하지만 적절한 피드백을 제공하고 소극적인 학생들을 토의에 참여시키는 구성주의적 평가를 지향하는 경우보다 성적 산출을 위한 자료를 얻는 한 가지 방법만을 지향하는 경우가 더 많았다. 즉, 교사들이 객관적인 점수를 얻을 수 있는 경우에만 관찰 평가를 사용하려는 경향이 있었는데, 이는 교사들이 평가의 목적에 대해 전통적인 관

점을 지니고 있기 때문으로 해석된다.

조별로 혼합물 한 가지를 선택하여 이를 분리할 수 있는 방법에 대해 토의를 통해 가설을 세우게 하고, 실험 과정에 대해 설계한 후 실제 실험하여 결과를 얻도록 한다. (중략) 또한, 수업 전체를 교사가 지속적으로 순회 관찰하면서 특별히 좋지 않은 태도를 보이는 학생은 기록하여 평가 결과에 반영한다.

자기 평가는 학습 후 자신의 이해 수준과 개선할 점에 대하여 스스로 평가하는 방법으로서, 반성적 자기 평가는 타인에 의한 평가보다 자기 발전에 더 큰 영향을 미친다(Wilson, 1992). 동료 평가는 사회적 구성주의 관점에서 학생 간의 상호작용을 강조한 것으로서, 상호작용을 통해 제공되는 피드백이 모든 학생들에게 도움이 되는 평가이다. 따라서 자기 평가와 동료 평가는 성적 산출보다 학습 지원을 목적으로 하는 구성주의적 평가라고 할 수 있다(초한무, 2001). 하지만 이 연구에서는 교사들이 무임승차자 방지 등 수업 참여도를 평가하기 위한 목적으로만 자기 평가와 동료 평가를 활용하려는 경우가 많았다.

소집단에서 무임승차자가 발생하지 않는지 파악하기 위해 학생들로 하여금 자기 평가 및 동료 평가를 각자 실시하도록 함. 탐구 수행결과는 시간이 허락하는 한 조별로 발표하도록 하며 이때 주제선정, 탐구 과정, 결과정리 등의 체크리스트를 만들어 평가함.

4. 결론

연구 I에서는 중등 과학교사들이 학교에서 실시하는 과학 평가의 실태와 전형적인 수행평가 상황에서 나타난 과학 평가의 지향을 평가의 목적과 방법 측면에서 분석하고, 교사의 배경 변인에 따른 차이도 조사하였다.

교사들이 실시하는 과학 평가 사례를 분석한 결과, 교사 대부분은 주로 학생에게 피드백을 제공하거나 학습에 도움을 주는 목적의 형성평가보다는 성적 산출, 대표 선발, 수업 참여도 측정을 위해 교사가 필요한 정보를 얻는 목적의 총괄평가를 먼저 떠올리며, 이를 자신의 주요 평가 활동으로 인식하는 것으로 나타났다. 성적

을 통해 학생에게 도움이 되는 피드백을 제공하기 위한 구성주의적인 총괄평가는 나타나지 않았으며, 교사에게 필요한 정보를 얻기 위한 전통적인 목적의 총괄평가가 주로 나타났다. 평가 방법 측면에서는 단답형, 서술형 등 측정을 활용하는 평가와 실험, 보고서, 논술, 발표 등 학생 수행을 활용하는 평가가 대부분이었으며, 자기 평가, 동료 평가 등의 비형식 평가 방법은 거의 사용되지 않았다. 즉, 교사들은 과학 평가 방법에 비하여 구성주의적 평가 목적에 대한 이해가 부족하다고 볼 수 있다. 교사들이 형성평가라는 용어를 흔히 사용하지만, 형성평가가 구성주의적 평가보다는 작은 총괄평가의 의미를 지니는 것도 이러한 해석을 뒷받침한다. 따라서 앞으로의 평가 전문성 향상 프로그램에서는 구성주의적 평가 방법을 단순히 소개하는 것보다 교사들의 구성주의적 평가의 중요성에 대한 인식과 실천 능력을 증진시키기 위한 방안을 모색해야 할 것이다.

다음으로 전형적인 수행평가 상황에서 나타난 교사들의 과학 평가에 대한 지향을 분석해보면, 실태 분석에서 나타났던 평가 목적과 방법의 불일치가 더욱 크게 나타났다. 구성주의적 평가를 실시할 수 있는 가상의 전형적인 수행평가 상황에서 학생들에게 형성평가를 시행하겠다고 응답한 교사의 비율은 실태 조사 결과보다는 다소 높았으나, 여전히 모든 평가 상황에서 총괄평가 목적의 평가 비율이 더 높았다. 하지만 평가 방법 측면에서 대부분의 교사들은 수행 방식과 비형식 방식의 평가를 지향하였다. 즉, 대부분의 교사들이 실태 조사에서는 거의 활용하지 않았던 대표적인 구성주의적 평가 방법을 지향하는 것으로 나타나 평가 방법 측면에서는 교사들의 평가 실태와 지향 사이에 큰 차이가 있었음을 알 수 있었다. 하지만 교사들이 지향하는 평가 목적은 여전히 총괄평가가 우세하게 나타나 상대적으로 평가 실태와 지향 사이에 큰 차이가 없었다고 할 수 있었다. 요약하면, 교사들이 실제로 사용하는 평가와 비교할 때 교사들이 지향하는 평가는 방법 측면에서는 보다 구성주의적 평가인 것처럼 보이지만, 목적 측면에서는 여전히 전통적 평가에서 벗어나지 못한다고 볼 수 있다. 이러한 결과는 교사들이 구성주의적 평가 방법에 대한 지식은 있으나 구성주의적 평가가 무엇인지에 대해서는 피상적으로 이해하고 있을 가능성을 의미한다. 따라서 구성주의적 평가에 대한 교사들의 인식과 이해를 증진시키기 위한 노력을 지속적으로 기울여야 할 필요성이 있다. 구성주의적 과학 평가관을 지닌 교사가 형성평가를 지향하는 비율이 높은 것으로 나타난 결과도 교사의

구성주의적 평가에 대한 인식 함양을 위한 지속적인 노력의 필요성을 뒷받침한다. 하지만 구성주의적 평가에 대한 교사들의 인식 전환은 단기간에 이루어지기 어려운데, 현직교사를 대상으로 실시되는 연수는 기간이 비교적 짧으므로 한계가 있다(최명숙, 2001). 따라서 집중적이고 지속적인 교육이 가능한 예비교사 교육과정에 관심을 가져야 할 것이다. 선행연구에서도 평가 강좌가 예비교사의 평가 인식 개선에 도움이 되는 것으로 보고되었으므로(박정, 2013b), 예비교사를 대상으로 구성주의적 평가에 대한 지속적인 교육이 필요할 것이다.

한편, 교사들이 지향하는 평가 목적과 방법은 평가 상황에 따라 다양하였으며, 몇 가지 목적과 방법에 편중되어 있었다. 이때, 교사들은 구성주의적 평가 방법으로 알려진 다양한 평가 방법들을 전통적 평가 목적으로만 활용하고자 하는 경우가 많았다. 이러한 결과는 제시된 전형적인 수행평가 상황에서 활용할 수 있는 다양한 구성주의적 평가 방법에 대한 교사들의 이해도나 친숙도가 부족했기 때문일 가능성이 의미한다. 따라서 평가 전문성 향상 프로그램에서 구성주의적 평가를 다룰 때는 원론적이고 이론적인 교육을 탈피하여 구체적이고 전형적인 수행평가 상황을 바탕으로 다양한 평가 방법에 대한 실천적인 교육을 진행할 필요가 있다.

제 4 장. 중등 예비 화학교사의 수행평가 과제 개발 과정에서 나타나는 PCK 구성요소의 특징 (연구 II)²⁾

1. 서 론

지식 기반 사회에서는 단순한 지식의 습득보다 지식의 재구성을 위한 통찰력과 문제 해결 능력의 함양이 필요하다. 이러한 사회적 요구에 부응하고자 우리나라에서도 기존의 단답형, 선택형 문항 중심의 지필평가를 지양하고 학생들의 고등 사고 능력을 기르고 평가하기 위한 수행평가를 도입하여 시행하고 있다. 학습자의 능동적인 지식 구성을 강조하는 구성주의 학습관에 기초한 수행평가는 학습의 결과보다 과정과 질을 측정하는 평가로서 학생들에게 실제적이고 유의미한 평가 과제를 제시한다(박선미, 1998). 또한 성적 산출과 같은 전통적인 목적뿐 아니라 평가를 통한 개별적 피드백 제공과 학생들의 발달과정 추적을 중시한다(백순근, 2000; Torrance, 1995). 그러므로 수행평가는 이상적인 구성주의적 평가를 구현하는 평가 체계이다. 그동안 수행평가에 대한 정책적 강조로 인하여 교육현장에서 실시되는 수행평가가 양적으로는 늘어났다. 그러나 연구 I에 따르면, 수행평가가 질적인 측면에서 축소된 지필평가나 단편적인 지식 습득 여부에 관한 평가에 머무르는 문제점이 지적되었다.

수행평가가 구성주의적으로 이루어지기 위해서는 개발 과정에서부터 다양한 PCK 구성요소가 종합적으로 활용되어야 한다(Falk, 2012; 민희정, 2012). 그러므로 수행평가 개발 과정에서 활용되는 PCK 구성요소의 활용 수준을 분석함으로써 수행평가의 질을 평가하고 개선하는 방안을 도출할 수 있다. 지금까지 우리나라의 수행평가 관련 연구는 수행평가 실태 조사(고민석 등, 2013; 김석우, 2007), 수행평가에 대한 교사의 인식이나 효과 분석(김성원, 현미숙, 2005; 이현주, 최경희, 남정희, 2000; 홍정림, 최은경, 2002), 수행평가 채점 방식에 따른 신뢰도 분석(김형준, 유준희, 2010; 이규민, 황경현, 2007), 수행평가 자료 개발(최혁준, 2013; 최희정, 임채성,

2) 연구 II는 한국과학교육학회지 37호 2권의 291-299쪽에 게재된 ‘중등 예비 화학교사의 수행평가 개발 과정에서 나타나는 교과교육학 지식(PCK) 요소의 특징’의 내용을 보완하고 재구성한 것임.

김은진, 2003) 등이 대부분이었다. 최근 수행평가 과정에서 나타나는 교사의 PCK를 분석한 연구가 일부 이루어졌지만, 교사들이 협력하여 수행평가를 계획하거나(Falk, 2012) 교육현장에서의 수행평가 실행 상황을 다루었다(민희정, 2012). 따라서 교사들이 수행평가의 개발 과정에서 중점적으로 고려하는 PCK 구성요소가 무엇인지는 밝혀내지 못하였다. 연구 I에서 현직교사를 대상으로 하는 평가 연수는 교사의 평가 전문성을 향상시키고 구성주의적 평가의 실행에 기여하는 것으로 나타났지만, 교사의 평가 전문성은 단기간에 향상시키기 어려우므로 지속적인 노력이 요구된다(강대중, 2016). 즉, 교사의 구성주의적 수행평가 능력을 향상시키기 위해서는 현직 교사에게 평가 연수 프로그램을 제공해야 하겠지만, 동시에 예비교사들에게도 수행평가에 대한 전문성을 갖출 수 있도록 교육해야 할 필요성이 있다.

이에 연구 II에서는 예비교사를 대상으로 한 수행평가 전문성 향상을 위한 기초 자료를 얻기 위하여 중등 예비 화학교사들의 수행평가 개발 과정에서 나타나는 PCK 구성요소와 PCK 구성요소 사이 연계의 특징을 분석하였다. 일반적으로 수행평가 개발 과정은 평가 소재 및 아이디어를 떠올리는 구상 단계와 평가 준거와 내용을 상세하게 결정하는 구체화 단계로 나눌 수 있다(장수웅, 2002; Doran, Chan, & Tamir, 1998). 따라서 수행평가의 각 단계에 필요한 PCK 구성요소와 요소별 특징이 달라질 수 있으므로 연구 II에서는 예비교사의 수행평가 개발 과정을 세분하여 구상, 구체화 단계 및 이후 조별 논의 과정에서 나타나는 PCK 구성요소의 특징을 분석하였다.

2. 연구 방법

2.1 연구 참여자

충청북도 소재 사범대학의 화학교육과 3학년에 재학 중인 학부생 중, 졸업 후 교직 진출을 희망하는 예비교사 8명(남학생 7명, 여학생 1명)이 연구에 참여하였다. 연구 참여자로 3학년을 선정한 이유는 이들이 사범대학에 재학 중인 일반적인 예비교사의 수준과 큰 차이가 없다고 볼 수 있으며, 2년 동안 다양한 교육학 강의를 수강하여 기초적인 교육학적 지식을 갖추고 있는 상태이므로 예비교사의 PCK 개발

및 수행평가 전문성 향상을 위한 이 연구의 목적에 적합하였기 때문이다. 연구에 참여한 예비교사의 구체적인 배경은 <표 II-1>과 같다. 예비교사들은 22-25세였으며, 모두 화학교재연구 및 지도법과 교육평가 등의 평가 관련 강좌를 수강하여 수행평가의 개념, 도입 배경, 필요성, 수행평가 방법의 종류와 특징, 장단점 등 수행평가에 관련된 기초적인 내용을 학습한 상태였다.

<표 II-1> 연구에 참여한 예비교사의 배경

예비교사	1조				2조			
	A	B	C	D	E	F	G	H
성별	남성	여성	남성	남성	남성	남성	남성	남성
나이	24	22	22	24	25	24	25	24
평가 관련 강좌* 경험	있음	있음	있음	있음	있음	있음	있음	있음

* 과학교육론, 교육평가.

2.2 연구 절차

자료 수집은 매주 3시간씩 3주에 걸쳐 이루어졌다. 첫째 주에는 연구 안내 후 수행평가 개발에 관한 워크숍을 진행하였다. 워크숍에서는 과학 평가의 목적과 방법, 평가의 타당도와 신뢰도, 일반적인 수행평가의 개발 절차(평가 목표 진술, 과제 형식 결정, 과제작성, 채점 기준 개발 등)를 설명하고, 한국교육과정평가원의 수행평가 지원 포털에 탑재된 과학 교과목의 수행평가 예시 중 연구 II의 과제와 무관한 내용을 선정하여 여러 가지 수행평가 방법의 종류와 특징을 구체적인 사례와 함께 소개하였다.

둘째 주에는 2009 개정 과학과 교육과정의 중학교 1학년 과학의 ‘분자 운동과 상태 변화’ 단원 내에서 한 차시 분량의 수행평가를 개별적으로 개발하도록 하였다. 모든 예비교사에게 2종류의 중학교 1학년 교과서와 지도서를 제공하고 필요시에는 스마트폰을 활용하여 인터넷 검색을 할 수 있도록 하였다. 수행평가 개발은 구상 단계와 구체화 단계로 나누어 실시하였는데(장수웅, 2002; Doran, Chan, &

Tamir, 1998), 구상 단계에서는 여러 가지 수행평가 아이디어를 떠올리도록 하고 구체화 단계에서는 그 중 한 가지를 선택하여 자세히 수행평가 과제를 개발하도록 하였다. 즉, 수행평가의 구조와 흐름을 파악하기 위하여 교사와 학생이 수행하는 구체적인 활동이 포함된 수행평가 계획안, 학생들에게 제공할 활동지, 평가 방법 및 준거 등을 자세히 작성하도록 하였다(부록 II-1). 4명의 면담자가 각 1명의 예비교사를 전담 관찰하여 특징적인 사항을 필드 노트에 기록하고 이후 중간 면담의 기초 자료로 활용하였다. 중간 면담은 수행평가 개발이 끝난 후 면담자가 관찰했던 예비교사를 대상으로 필드 노트와 수행평가 계획안을 바탕으로 30-50분 동안 실시하였다. 예비교사의 대답이 불충분한 경우 후속질문을 하였으나 3주차 활동에 영향을 미치는 질문은 하지 않았다(표 II-2). 개별 과제 개발 과정 및 중간 면담은 녹음 및 녹화하였다.

셋째 주에는 수행평가 계획안에 대한 조별 논의와 수정을 실시하였다. 예비교사들을 4명씩 2개 조로 편성하여 작성한 수행평가 계획안을 조별로 검토하도록 하였다. 조별 논의는 한 명의 예비교사가 자신의 평가 계획안을 설명한 후 다른 조원들이 검토 및 수정 의견을 제시하고 논의하는 방식으로 이루어졌으며, 이 과정을 4인의 면담자가 관찰하며 필드 노트를 작성하여 면담의 기초 자료로 활용하였다. 조별 논의 후, 개별로 수행평가 계획안을 완성하도록 하였다. 모든 활동이 끝난 후 반구조화된 최종 면담을 40-60분 동안 실시하였다. 최종면담에서는 조별 논의 과정에서 자신이 지적했던 내용과 의도, 다른 조원에게 지적받았던 내용과 이에 대한 본인의 생각, 최종 계획안 반영 여부와 그 이유에 대한 질문을 하였다. 이후 예비교사가 고려한 PCK 구성요소 중 면담이나 계획안 분석을 통해 잘 드러나지 않는 요소를 구체적으로 파악하기 위한 질문을 하였다. 또한 전체 연구 과정에 대한 예비교사의 생각과 어려웠던 점에 대해서도 질문하였다(표 II-2). 조별 논의 과정 및 최종 면담도 녹음 및 녹화하였다.

<표 II-2> 반구조화된 면담의 개요

<2주차 중간 면담>	
영역	면담질문
배경	• 대학에서 평가 관련 강좌를 수강한 적이 있나요? (과학교육론, 교육평가 등)

	Y: 어떤 과목인가요? 내용은?
변인	<ul style="list-style-type: none"> 수행평가 과제 개발 활동을 해본적이 있나요?
	Y: 언제, 어떤 내용에 대해 만들어보셨나요?
	<ul style="list-style-type: none"> (구상 단계 활동지를 같이 보면서, 각 활동 별로) 이걸 어떤 의도로 작성한 것인지 설명해주세요. (활동지에 기록된 내용을 바탕으로 하나씩 상황, 방법, 도입 의도, 이유, 고려한 점 등을 구체적으로 질문)
개별	<ul style="list-style-type: none"> 수행평가 아이디어를 떠올릴 때, 어떤 점을 주로 고려하셨나요?
개발	<ul style="list-style-type: none"> 어떤 자료를 살펴보았나요? 자료의 어떤 측면을 중점적으로 보았나요?
과정	<ul style="list-style-type: none"> 참고했던 자료의 내용을 반영한 부분이 있나요?
(구상 단계)	<p>N: 반영하지 않은 이유는 무엇인가요?</p> <p>Y: 어떻게 반영하였으며, 의도는 무엇인가요?</p> <ul style="list-style-type: none"> 수행평가 아이디어를 검토하고 수정하신 내용이 있나요? <p>N: 수정할지 고민했던 내용이 있나요? 그렇다면 수정하지 않은 이유는 무엇인가요?</p> <p>Y: 원래 어떤 내용이었는데 어떻게 수정을 했나요? 수정한 이유는 무엇인가요?</p> <ul style="list-style-type: none"> 아이디어 중에서 이것을 선택한 이유는 무엇인가요? (이 아이디어가 다른 아이디어보다 낫다고 판단한 이유는?) 아이디어 구상 단계에서 어려웠던 점, 가장 고민했던 점이나 시간이 오래 걸렸던 점은?
	<ul style="list-style-type: none"> (구체화 단계 활동지를 같이 보면서, 각 활동 별로) 이걸 어떤 의도로 작성한 것인지 설명해주세요. (활동지에 기록된 내용을 바탕으로 하나씩 도입 의도, 상황, 방법, 이유 등을 구체적으로 질문)
	<ul style="list-style-type: none"> 이 수행평가 과제에서 평가하고자 했던 가장 핵심적인 영역은 무엇인가요? 그것을 선택한 이유는?
개별	<ul style="list-style-type: none"> 어떤 자료를 살펴보았나요? 자료의 어떤 측면을 중점적으로 보았나요?
개발	<ul style="list-style-type: none"> 참고했던 자료의 내용을 반영한 부분이 있나요?
과정	<p>N: 반영하지 않은 이유는 무엇인가요?</p> <p>Y: 어떻게 반영하였으며, 의도는 무엇인가요?</p>
(구체화 단계)	<ul style="list-style-type: none"> 수행평가 과제를 검토하고 수정하신 내용이 있나요? <p>N: 수정할지 고민했던 내용이 있나요? 그렇다면 수정하지 않은 이유는 무엇인가요?</p> <p>Y: 원래 어떤 내용이었는데 어떻게 수정을 했나요? 수정한 이유는 무엇인가요?</p> <ul style="list-style-type: none"> 구체화 단계에서 어려웠던 점, 가장 고민했던 점이나 시간이 오래 걸렸던

점은?

- 구체화한 수행평가 과제가 아이디어 구상 단계 때의 계획대로 잘 되었다고 생각하나요?

N: 구체적으로 어떤 점에서 잘 안 되었다고 생각하나요?

- 구체화 단계에서 어려웠던 점, 가장 고민했던 점이나 시간이 오래 걸렸던 점은?

<3주차 최종 면담>

영역	면담질문
조별 논의 과정	<ul style="list-style-type: none">• 본인의 수행평가 과제에 대하여 친구들이 지적해 준 점에는 어떤 것들이 있었나요? (각각에 대해서) 친구들이 지적해주었던 점을 반영하였나요? N: 반영하지 않은 이유는 무엇인가요? Y: 반영한 이유는 무엇인가요? 어떻게 반영하였나요?• 다른 친구의 수행평가 계획안을 검토할 때 어떤 점을 주로 검토하였나요? (어떤 점을 주로 지적하셨나요?) 그 이유는 무엇인가요?• 친구들의 지적을 받았던 부분 외에 본인이 추가하거나 수정한 내용이 있나요? Y: 추가하거나 수정한 이유는 무엇인가요?• 조별 논의 과정에서 어려웠던 점, 가장 고민했던 점이나 시간이 오래 걸렸던 점은?
PCK 관련	<p>(본인의 생각을 묻는 것으로, 정답이 있는 것이 아님을 강조)</p> <p>1) 과학교육과정에 관한 지식</p> <ul style="list-style-type: none">• 교과서나 지도서에 제시된 평가 내용이나 목표를 참고하셨나요? N: 참고하지 않은 이유는 무엇인가요? Y: 참고한 이유는 무엇인가요? 참고하신 내용을 평가에 (어떻게) 반영하셨나요?• 수행평가 전후 수업 차시에 어떤 내용을 다룰지 생각을 해보셨나요? Y: 구체적으로 어떤 내용을 생각해보셨나요? <p>2) 과학 교수 전략에 관한 지식</p> <ul style="list-style-type: none">• 수행평가 과제에 학생이 참여할 수 있도록 고려한 부분이 있나요? (그러한 과정이 들어갈 필요가 있다고 생각하시나요?) N: 그 이유는 무엇인가요? Y: 구체적으로 어떤 부분에서 반영이 되었나요?• (실험인 경우) 과정이나 재료를 다양해서 학생들이 여러 가지 결과를 낼 수 있도록 하는 것에 대해서는 어떻게 생각하시나요?

3) 학생에 관한 지식

- 학생의 특성에 대해서 수준이나 오개념 등을 고려해보셨나요?
Y: 구체적으로 어떤 과정에서 이런 점을 고려하셨나요?

4) 과학 평가에 관한 지식

- 채점 기준(루브릭, 배점표 등)을 설정한 과정에 대해 설명해주세요. 무엇을 주로 고려하였나요?
- 평가 기준(체크리스트 등)을 학생에게 알려줄 생각인가요?
Y, N: 그 이유는 무엇인가요?

5) 과학 내용에 관한 지식

- 평가 과제와 관련된 과학 내용에 대해서는 어떤 점을 주로 생각하셨나요?
- 지식적인 측면에서 부족함을 느낀 점이 있나요?

전체
과정

- 수행평가 과제 개발 및 조별 논의 과정을 다른 예비교사에게 적용한다고 했을 때, 개선할 점이나 보완할 점, 추가로 제공해주었으면 하는 점이 있다면 무엇인가요?
-

2.3 분석 방법

과학 교과 및 평가에서의 PCK를 조사한 선행연구(민희정, 2012; Falk, 2012; Kind, 2009; Magnusson, Krajcik, & Borko, 1999; Park & Chen, 2012; Park & Oliver, 2008)를 바탕으로 수행평가 개발에서 고려해야 할 PCK 구성요소를 선정하여 예비 분석틀을 구성하였다. 이를 이용하여 면담 및 조별 논의 전사본을 분석한 후 예비 분석틀을 수정, 보완하여 PCK 구성요소의 정의와 하위 요소를 확정하였다. <표 II-3>에 제시한 최종 분석틀에서는 PCK 구성요소로 과학 평가에 관한 지식, 과학교육과정에 관한 지식, 과학 교수 전략에 관한 지식, 과학 내용에 관한 지식, 학생에 관한 지식, 과학 평가에 대한 신념 등을 설정하였고 각 요소별로 2-3가지의 하위 요소를 정의하였다.

<표 II-3> 과학 수행평가에 관한 PCK 구성요소별 하위 요소와 정의

구성요소	하위 요소	정의
과학 평가에 관한 지식(A)	평가 방법(A1)	여러 가지 평가 방법 및 평가 방법의 신뢰도와 타당도에 대한 지식
	평가 영역(A2)	개념 이해도, 참여도 등 교사가 평가하는 영역에 대한 지식
과학교육과정에 관한 지식(C)	연계(C1)	수행평가 전후 차시와의 연계, 개념의 교육과정상 위치 등 교육과정의 연계에 대한 지식
	평가 목표(C2)	교육과정에 제시된 평가 목표에 대한 지식
과학 교수 전략에 관한 지식(I)	주제 전략(I1)	특정 주제에 대한 표현(그림, 예시, 모형, 비유 등)과 활동(문제풀이, 시연, 탐구, 실험 등)에 대한 지식
	교과 전략(I2)	수업 모형 등 과학 교과의 일반적 전략에 대한 지식
과학 내용에 관한 지식(K)	개념 및 이론(K1)	과학 개념 및 이론에 관한 지식
	과학 과정(K2)	실험이나 탐구 과정에 관한 지식
학생에 관한 지식(L)	오개념(L1)	개념, 이론, 실험 등과 관련하여 학생들이 흔히 갖는 오개념에 관한 지식
	발달 수준(L2)	문항, 지문, 실험 등을 이해할 수 있는 학생들의 발달 수준에 관한 지식
	동기 및 흥미(L3)	학생들이 흥미를 느끼고 동기를 부여할 수 있는 활동이나 소재에 관한 지식
과학 평가에 관한 신념(O)	전통적 평가 지향(O1)	선발, 성적 산출 등 전통적 평가 목적에 대한 지향
	구성주의적 평가 지향(O2)	개념이나 과학의 본성 등에 관한 학습 기회 제공, 피드백 등 구성주의적 평가 목적에 대한 지향

수행평가 개발에 사용된 PCK 구성요소의 분석에서는 예비교사의 중간 면담 자료와 조별 논의 전사본을 중심으로 하고 예비교사들이 작성한 활동지, 연구자의 필드 노트, 최종 면담 자료 등은 보조적인 자료로 사용하였다. PCK 구성요소의 분석은 예비교사가 수행평가 개발 과정에서 제시한 세부 활동 단위로 분석하였다. 중간 면담의 경우, 한 가지 주제에 대한 면담자의 질문과 이에 대한 예비교사의 응답을 기본 단위로 분석하였다. 예비교사의 의도를 파악하기 위한 후속 질문의 경우에는 같은 분석 단위로 처리하였으며, 질의-응답 과정에서 다른 활동으로 초점이 이동한

경우에는 분석 단위를 구분하였다. 조별 논의에서는 예비교사들의 논의 대상이 이동하는 경우 분석 단위를 구분하였다. 하나의 분석 단위에 두 가지 이상의 PCK 구성요소가 동시에 나타나는 등 동일한 활동에 대하여 여러 PCK 구성요소들을 동시에 고려하는 경우 또는 한 PCK 구성요소가 다른 PCK 구성요소의 이유나 근거로 활용된 경우에는 두 PCK 구성요소가 연계된 것으로 분석하였으며(노태희, 박재성, 강훈식, 2016), 동일한 활동에 대하여 같은 PCK 구성요소가 여러 자료에 중복하여 드러나는 경우에는 한 번만 코딩하였다.

분석의 신뢰도를 높이기 위하여 2인의 분석자가 조별로 한 명씩 무작위로 선정된 예비교사의 면담 자료와 1개 조의 논의 자료를 각자 분석하고 비교하였으며, 분석자 사이에 의견이 일치하지 않는 경우에는 논의를 거쳐 합의를 도출하였다. 이 과정을 반복하여 분석자 간 일치도가 .93에 도달한 이후 1인의 연구자가 모든 자료의 PCK 구성요소를 분석하였고, 분석 결과는 수차례의 세미나를 통하여 수정 및 보완하였다.

3. 연구 결과 및 논의

3.1 개별 과제 개발 과정의 PCK 구성요소

개별 과제 개발 과정에서 나타난 예비교사들의 PCK 구성요소는 <표 II-4>와 같다. 전체 PCK 구성요소는 189개였으며, 이 중 58.2%는 구상 단계에서, 41.8%는 구체화 단계에서 나타났다. 구체화 단계에서 구상 단계보다 빈도가 늘어난 PCK 구성요소는 없었다. 구상 단계에서 가장 많았던 PCK 구성요소는 과학 평가에 관한 지식(23.3%)이었고, 과학 교수 전략에 관한 지식(12.2%)과 과학 내용에 관한 지식(10.6%)의 비율도 높았다. 과학교육과정에 관한 지식(4.2%), 과학 평가에 관한 신념(4.2%), 학생에 관한 지식(3.7%)의 비율은 상대적으로 낮았다. 구체화 단계에서 가장 많이 나타난 PCK 구성요소는 과학 평가에 관한 지식(20.6%)으로 구상 단계와 빈도는 비슷하였지만, 구체화 단계에서 나타난 PCK 구성요소의 절반(49.4%)을 차지하였다.

<표 II-4> 개별 과제 개발 과정의 PCK 구성요소

PCK 구성요소		구상 단계 (%)		구체화 단계 (%)		계 (%)	
A	A1	44 (23.3)	23 (12.2)	39 (20.6)	23 (12.2)	83 (43.9)	46 (24.3)
	A2		21 (11.1)		16 (8.5)		37 (19.6)
C	C1	8 (4.2)	8 (4.2)	3 (1.6)	3 (1.6)	11 (5.8)	11 (5.8)
	C2		-		-		-
I	I1	23 (12.2)	22 (11.6)	16 (8.5)	15 (7.9)	39 (20.6)	37 (19.6)
	I2		1 (0.5)		1 (0.5)		2 (1.1)
K	K1	20 (10.6)	14 (7.4)	6 (3.2)	1 (0.5)	26 (13.8)	15 (7.9)
	K2		6 (3.2)		5 (2.6)		11 (5.8)
	L1		-		1 (0.5)		1 (0.5)
L	L2	7 (3.7)	3 (1.6)	7 (3.7)	2 (1.1)	14 (7.4)	5 (2.6)
	L3		4 (2.1)		4 (2.1)		8 (4.2)
O	O1	8 (4.2)	5 (2.6)	8 (4.2)	7 (3.7)	16 (8.5)	12 (6.3)
	O2		3 (1.6)		1 (0.5)		4 (2.1)
합계		110 (58.2)		79 (41.8)		189 (100.0)	

* A: 과학 평가에 관한 지식(A1: 평가 방법, A2: 평가 영역), C: 과학교육과정에 관한 지식(C1: 연계, C2: 평가 목표), I: 과학 교수 전략에 관한 지식(I1: 주제 전략, I2: 교과 전략), K: 과학 내용에 관한 지식(K1: 개념 및 이론, K2: 과학 탐구 과정), L: 학생에 관한 지식(L1: 오개념, L2: 발달 수준, L3: 동기 및 흥미), O: 과학 평가에 관한 신념(O1: 전통적 평가 지향, O2: 구성주의적 평가 지향).

과학 평가에 관한 지식 중 평가 방법(A1)의 경우, 구상 단계에서는 어떤 평가 방법을 활용할지에 대한 평가 방법 선정이 8.5%(16회), 평가 방법의 적용 방안을 검토하는 평가 방법 검토가 2.1%(4회), 평가 준거 및 배점 설정이 1.6%(3회) 나타났다. 구체화 단계에서는 평가 준거 및 배점 설정이 6.3%(12회), 평가 방법 검토가 4.2%(8회), 문항 형태 검토가 1.6%(3회)로 나타났다. 구상 단계에서 높은 비율을 차지했던 평가 방법의 선정은 이미 결정되었기 때문에 구체화 단계에서는 나타나지 않았다. 이 단계에서 예비교사들은 구상 단계에서 도입했던 여러 가지 평가 방법을 구체화하여 평가 문항을 제작하고, 평가 유형에 따라 평가 준거표, 배점표, 평가 루브릭 등을 만들었다.

다. 즉, 관찰 평가나 동료 평가와 같이 주관적 요소가 개입될 수 있는 평가 방법을 사용할 때는 기본 점수를 부여하거나 배점을 줄이는 등 교사의 주관을 가능한 한 배제하려고 노력하였다. 예를 들어, 예비교사 C는 샤를 법칙과 관련하여 계란 삶기를 주제로 한 수행평가 과제를 개발하고, 결과 기술 20%, 결과 해석 60%, 참여도 20%의 비율로 점수를 부여하였다. 이때 참여도는 교사의 관찰 평가와 동료 평가 점수를 50%씩 합산하였으나, 20점 만점에 기본 점수 16점을 부여하여 참여도가 수행평가 점수에 미치는 영향을 최소화하였다.

면담자 : 참여도가 1점부터 5점. 교사가 매기는 거예요?

예비교사 : (중략) 그래도 교사의 눈이 정확하지 않을까 해서 교사를 기본으로 하고. 이제 거기에 (동료 평가는) 추가적으로 학생들끼리 서로 눈빛 주고 '알지?, 알지?' 이러는 문제가 있어서.

면담자 : 이런 영향을 줄이려고. 아 그래서 이것도 이게 편차가 5점밖에 안 되는 거구나, 기본 점수가 16점이 되는 거고.

(예비교사 C의 최종 면담 중)

또한 8명 중 7명의 예비교사는 평가 기준 공개에 부정적이었다. 하지만 평가의 신뢰도와 타당도를 높이기 위해서는 평가 기준을 학생에게 공개하는 것이 바람직하며(김찬중, 김진규, 임형, 2001), 이 외에도 학생의 입장에서 문항을 풀어보거나 각 수준별 예시 답안을 만들어보는 등 평가 방법과 준거의 신뢰도와 타당도 확보를 위한 노력이 중요하다(Doran, Chan, & Tamir, 1998).

면담자 : 채점 기준들을 학생에게 미리 알려줄 생각인가요? 수행평가를 하기에 앞서서.

예비교사 : 아니요. 저는 그거를 따로 얘기할 필요는 없다고 생각해요. 학생들이 그걸 알면서 '아, 이게 성적에 들어가는구나, 안 들어가는구나.' 그런 거를 알게 되면 참여도에 변화가 생길 수가 있기 때문에.

(예비교사 A의 최종 면담 중)

과학교육과정에 관한 지식의 비율은 구상 단계에서 4.2%, 구체화 단계에서 1.6%

에 불과하였다. 연계(C1)를 고려한 예비교사들은 이상기체 상태방정식이나 질량 보존의 법칙 등 단원 외의 개념을 수행평가에 도입할 수 있는지 확인하였으며 교육과정의 범위를 벗어난 개념은 도입하지 않았다. 그러나 구상 및 구체화 단계에서 모두 교육과정의 연계를 확인하지 않은 예비교사도 2명 있었고, 교육과정을 벗어나는 개념을 심화 개념이라고 생각하는 예비교사도 있었다. 하지만 이는 심화 학습이 아니라 속진 학습의 성격이 강하므로(노태희 등, 2013), 교육과정 범위 내에서 고차원적 사고력을 계발할 수 있는 심화 학습 평가가 이루어지도록 해야 한다.

평가 목표(C2)의 경우, 교과서나 지도서에 제시된 평가 목표를 확인하거나 참조하는 예비교사가 일부 있었으나 적극적으로 평가에 반영하는 예비교사는 없었다. 대부분의 예비교사들은 학생이 습득해야 할 개념이나 기능 등을 스스로 설정하였으며 교과서나 지도서에 제시된 평가 목표를 따르는 것보다 평가 의도에 맞게 스스로 평가 목표를 설정하는 것이 더 바람직하다고 인식하고 있었다. 교사의 교육과정 운영은 교육과정의 의도를 교육현장의 사정에 맞게 수정하여 실천하는 것을 포함하고(김두정, 2006), 예비교사가 설정한 평가 목표가 교과서나 지도서에 제시된 핵심을 대부분 포괄하고 있었으므로, 평가 목표에 해당하는 PCK 구성요소가 구상 과정에 고려된 것으로 해석할 수 있다.

면담자 : 교과서나 지도서에 제시된 어떤 평가 목표 아니면 수업목표 그런 걸 혹시 찾아봤나요?

예비교사 : 맨 처음 계획 짤 때는 한번 읽어보고 반영을 해야겠구나 생각하고 했는데요. 근데 반영을 하면 만들기가 어려웠을 거 같아요. (중략) 오히려 그러려면 평가 기준 같은 것도 거기에 맞춰야 되고 그러면 또 이 방법적인 것도 제한이 될 거 같아서.

(예비교사 F의 최종 면담 중)

과학 교수 전략에 관한 지식 중 주제 전략(I1)과 교과 전략(I2)의 비율은 각각 19.6%와 1.1%였고 구상 및 구체화 단계별 큰 차이는 없었다. 예비교사들은 수업 모형 등 과학 교과에서 일반적으로 쓰이는 교과 전략보다 특정 단원과 주제에 특화된 전략을 주로 고려하였다. 구체적으로 구상 단계에서는 실험 등 수업 활동을 설계하는 수업 기법이 9.5%(18회), 상호작용 촉진 전략, 동기 유발 전략 등의 수업 전

략이 2.1%(4회)로 나타났고, 구체화 단계에서는 수업 기법이 5.3%(10회), 수업 전략이 0.5%(1회), 시간 배분 또는 실험 안전 교육이 2.1%(4회)로 나타났다.

과학 내용에 관한 지식 중 개념 및 이론(K1)의 비율은 구상 단계에서 7.4%였으며, 확산, 증발, 보일의 법칙, 샤를의 법칙 등 과학 법칙이나 이론에 관한 지식을 떠올리는 과정에서 주로 나타났다. 그러나 구체화 단계에서 개념 및 이론은 거의 나타나지 않았는데(0.5%), 이는 수행평가 과제 개발에 필요한 과학적 지식은 구상 단계에서 대부분 고려되었기 때문으로 볼 수 있다. 과학 과정 지식(K2)은 수업 기법으로 실험을 선택한 예비교사에게서 낮은 비율로 나타났다(5.8%). 과학 과정 지식은 예비교사들이 개념이나 이론에 대한 이해 여부를 평가하기 위한 실험을 설계할 때 주로 나타났지만, 학생들이 실험을 수행하고 실험 결과의 의미를 이해하는지 평가하는 데 목적이 있었기 때문에 과학 과정 지식에 대한 평가 계획은 이루어지지 않았다. 과학적 개념과 이론 평가는 지필평가로도 쉽게 할 수 있으므로 과정에 관한 평가를 강조하는 수행평가에서는 학생들의 과학 과정 지식이나 기능을 중점적으로 평가할 필요가 있다(교육부, 1998).

학생에 관한 지식은 평가 소재나 교과 전략을 도입할 때 학생의 수준, 동기, 흥미를 고려하는 형태로 나타났으나 비율은 구상 및 구체화 단계에서 각각 3.7%로 높지 않았다. 예비교사들은 대부분 학생들이 일반적으로 가지는 오개념, 선지식, 어려움, 해당 개념의 발달과정 등 학생에 관한 구체적인 지식(Falk, 2012)은 중요하게 고려하지 않았다.

면담자 : 평가를 계획할 때 학생의 특성을 고려한 점이 있다면? 그러니까 예를 들어 학생들의 인지적 수준이라든지, 친숙도, 이해도, 오개념, 선지식 이런 부분.

예비교사 : 학생들이 그니까 이해를 얼마나 하고 있는지를 사실 보고 싶은 거라서, 선개념이나 선지식 같은 거는... (중략) 수업을 하고 내용을 가지고 연결 지을 거만 생각하고 그런 건 딱히 고려를 안 해본 거 같아요.

(예비교사 G의 최종 면담 중)

현직교사와 달리 예비교사들은 학생을 가르친 경험이 부족하므로 수행평가 구상 단계에서 학생의 특성을 반영하는 것이 어려울 수 있다. 그러므로 평가 계획의

구상 단계에서 학생에 관한 지식을 구체적으로 제공하여 예비교사들이 학생의 흥미, 오개념, 발달 수준 등을 활용할 수 있도록 도울 필요가 있다.

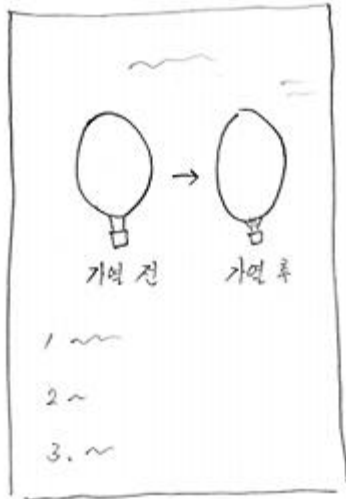
과학 평가에 관한 신념의 경우, 전통적 평가 지향(O1)은 6.3%였고 기술 피드백의 제공 등 구성주의적 평가 지향(O2)은 2.1%로 매우 적게 나타났다. 관찰 평가나 동료 평가는 피드백을 제공하여 구성주의적 평가를 구현할 수 있는 대표적인 평가 방법이지만(김송자, 최창우, 2006), 예비교사들은 무임승차 방지 등 학생의 수업 참여도 평가를 목적으로 활용하는 경향이 있었다. 또한, 예비교사들은 학생 통제를 위해 교사 주도의 구조화된 평가를 구상하는 경우가 많았다. 다음 사례에서 예비교사 D는 ‘온도가 증가하면 기체 분자는 어떻게 될까?’ 를 주제로 <그림 II-2>와 같이 유인물을 활용한 지필평가 형식의 수행평가 과제를 개발하였는데, 이러한 평가 방법을 선택한 이유로 소요 시간, 효율성, 교사의 학생 통제 등을 언급하였다.

제가 예전에 학교에서 그런 종류의 수업 시연을 많이 했거든요. 좀 열려있는 상태로 해서 아이들의 생각을 끌어내는 그런 걸 많이 했는데. 그런 게 좋다고 생각을 해요. 수업을 해보니까 좋은 건 좋은 건데 힘든 건 또 힘든 거더라고요. 하나하나의 생각도 굉장히 타당할 수도 있고 거기에 많은 과학적 내용이 있을 수도 있는데, 아무리 그래도 제가 가르치려는 집중적인 내용이 있잖아요. (중략) 아이들이 많이 못 좇아 오는 경우도 있고 아무래도 시간도 많이 걸리고, 왜냐면 생각이 이것저것 많이 나오고 중구난방 하다보니까 시간도 오래 걸리고 해서 좋다는 건 알겠지만 효율에 관해선 잘 모르겠어요.

(예비교사 D의 최종 면담 중)

이러한 결과는 과학 평가에 관한 지식에서 예비교사들이 평가의 공정성과 객관성 유지를 중시한 것과 맥락을 같이한다. 예를 들어, 구상 단계에서는 평가에 관한 교사의 편의성을 고려하는 경우가 있었고, 구체화 단계에서는 특히 평가 문항의 객관성에 대해 고려할 때 전통적 평가 지향이 강하게 나타났다.

[세부 계획안]



· 과학 교사 같은 유인물을 내어준 후 각각에 대한 설명을 진행

· 학생은 교사의 설명을 들은 후 가기에 맞게 자신의 생각등을 작성

준강

1. 위 가에서 각 속의 공기 분자는 어떻게 되어있는지 그려보시오.

· 과학 원리를 이용해 분자를 그려보도록 설명하고 질도에 따라 화살표 등을 자유롭게 사용하도록 한다.

2. 거품 전과 후의 공기 분자가 어떻게 되는지 설명해보시오

· 리선이 안고 있는 책을 자유롭게 글로 소개 한다.

3. 위 그림과 같은 현상의 예를 두 가지 이상 서술하시오.

[평가 준거]

총 5점 만점

· 1번 문항

- 그림을 이용하여 거품의 부피 증가(분자 사이 간격이 멀어짐) 그림의 유무만을 평가 (유: 1점, 무: 0점)

· 2번 문항

- 사물의 변화를 이용한 설명, '원리에 의한 분자 속의 증감'을 이용한 설명, '생분현' 등 3가지의 경우를 각각 2, 1, 0점으로 평가

· 3번 문항

- 2가지, 1가지, 사실 인화 3가지 경우를 각각 2, 1, 0점으로 평가

<그림 II-2> 예비교사 D가 개발한 수행평가 과제

3.2 개별 과제 개발 과정의 PCK 구성요소 사이의 연계

개별 과제 개발 과정에서 나타난 PCK 구성요소 사이 연계의 분석 결과는 <표 II-5>, <표 II-6>과 같다. 구상 단계에서는 주제 전략을 중심으로 PCK 구성요소 사이의 연계가 많았지만(20회, 87.0%), 구체화 단계에서는 PCK 구성요소 사이의 연계가 거의 나타나지 않았다(3회, 13.0%). 이는 예비교사들이 구상 단계에서는 수업 기법을 선택하고 도입할 때 다양한 요소를 고려하였지만, 구체화 단계에서는 선택한 수업 기법을 바탕으로 평가를 구체화하는 데 집중하였고, 이때 문항의 난이도와 변별도 등 평가 관련 요소만 검토하였기 때문으로 보인다.

<표 II-5> 개별 과제 개발 과정의 PCK 구성요소 사이의 연계 (통합 유형별)

통합 유형	구상 단계 (%)	구체화 단계 (%)	계 (%)	
A - I	4 (17.4)	-	4 (17.4)	
A - L	-	1 (4.3)	1 (4.3)	
A - O	-	1 (4.3)	1 (4.3)	
2가지 PCK 구성요소의 통합	I - C	4 (17.4)	-	4 (17.4)
	I - K	3 (13.0)	-	3 (13.0)
	I - L	5 (21.7)	1 (4.3)	6 (26.1)
	I - O	2 (8.7)	-	2 (8.7)
	K - L	2 (8.7)	-	2 (8.7)
계	20 (87.0)	3 (13.0)	23 (100.0)	

* A: 과학 평가에 관한 지식, C: 과학교육과정에 관한 지식, I: 과학 교수 전략에 관한 지식, K: 과학 내용에 관한 지식, L: 학생에 관한 지식, O: 과학 평가에 관한 신념.

<표 II-6> 개별 과제 개발 과정의 PCK 구성요소 사이의 연계 (PCK 구성요소별)

PCK 구성요소	구상 단계 (%)	구체화 단계 (%)	계 (%)
A와 다른 PCK 구성요소의 연계	4 (8.7)	2 (4.3)	6 (13.0)
C와 다른 PCK 구성요소의 연계	4 (8.7)	-	4 (8.7)
I와 다른 PCK 구성요소의 연계	18 (39.1)	1 (2.2)	19 (41.3)
K와 다른 PCK 구성요소의 연계	5 (10.9)	-	5 (10.9)
L과 다른 PCK 구성요소의 연계	7 (15.2)	2 (4.3)	9 (19.6)
O와 다른 PCK 구성요소의 연계	2 (4.3)	1 (2.2)	3 (6.5)
계	40 (87.0)	6 (13.0)	46 (100.0)

* A: 과학 평가에 관한 지식, C: 과학교육과정에 관한 지식, I: 과학 교수 전략에 관한 지식, K: 과학 내용에 관한 지식, L: 학생에 관한 지식, O: 과학 평가에 관한 신념.

구상 단계에서 나타났던 모든 PCK 구성요소 사이의 연계(20회) 중 90.0%(18회)는 과학 교수 전략에 관한 지식과 연계되었다. 따라서 구상 단계에서 PCK 구성요소 사이 연계는 주로 과학 교수 전략에 관한 지식을 중심으로 이루어졌다고 볼 수 있다. 이때 학생에 관한 지식과의 연계는 주로 예비교사가 수업 기법이나 전략을 구상하고 선택할 때 학생의 흥미와 동기, 발달 수준을 고려하는 과정에서 나타났다. 교수 전략과 학생에 관한 지식은 교사 PCK의 핵심 요소이므로 두 지식의 연계가 많이 일어나는 것이 바람직하다(Park & Oliver, 2008). 그러나 이 연구에서는 교수 전략과 학생에 관한 지식의 연계가 나타나지 않았던 예비교사도 있었으며, 질적인 측면에서도 특정한 수업 기법이나 전략을 도입하는 과정에서 일회적으로 다른 PCK 구성요소를 고려한 경우가 많았다. 예를 들어, 흥미와 동기는 수업의 도입이나 평가 소재 선택 과정에서 일부 나타났을 뿐, 학생 간 상호작용 촉진 전략과 같이 학생의 흥미와 동기를 유발할 수 있는 수업 기법이나 전략은 거의 나타나지 않았다. 과학 교수 전략에 관한 지식은 수행평가의 기본 틀을 결정하는 중요한 PCK 구

성요소이며 다양한 PCK 구성요소와 연계될 수 있는 잠재력이 높다. 따라서 예비교사가 교수 전략을 구상할 때 다양한 PCK 구성요소를 고려하고 연계할 수 있도록 도울 필요가 있다.

학생에 관한 지식은 과학 교수 전략에 관한 지식을 제외하면 다른 PCK 구성요소와 연계되는 비율이 매우 낮았다. 수행평가에서 학생들에게 제시하는 지문과 실험 내용은 모든 수준의 학생들이 이해할 수 있도록 작성해야 한다(Doran, Chan, & Tamir, 1998). 그리고 교수 전략이나 평가 전략을 구체화하는 단계에서도 학생들의 오개념이나 발달 수준 등을 고려해야 하므로 학생에 관한 지식은 다른 PCK 구성요소와 폭넓게 연계될 필요가 있다. 따라서 예비교사가 교수 전략이나 평가 내용을 구체화할 때 학생에 관한 지식을 적극적으로 고려할 수 있도록 지침을 제공해야 할 것이다.

과학 평가에 관한 지식은 개별 과제 개발 과정에서 나타난 전체 PCK 구성요소의 43.9%를 차지하였음에도 불구하고 다른 PCK 구성요소와의 연계 비율은 매우 낮았다. 현직교사의 경우, 과학 평가에 관한 지식은 주로 과학 교수 전략에 관한 지식이나 학생에 관한 지식과 연계되는 것으로 나타났으므로(Park & Chen, 2012), 예비교사들이 수행평가 구상 과정에서 평가 계획을 수립할 때 학생의 특성과 교수 전략을 보다 충분히 고려하도록 해야 할 것이다.

과학 평가에 관한 신념을 중심으로 한 연계는 PCK 구성요소 중에서 가장 낮았다. 특히 수행평가가 전반적으로 전통적인 관점에서 개발되었음에도 전통적 평가 지향이 연계된 경우는 없었고, 드물게 나타난 연계(3회)는 모두 구성주의적 평가 지향이 다른 PCK 구성요소가 연계된 경우였다. 이는 예비교사들이 수행평가에 어떤 활동을 도입할 때 대체로 전통적인 관점을 견지하였으나 이는 명시적이고 의도적이었다기보다 암묵적이고 당위적인 수준이었음을 의미한다. 구상 단계에서는 먼저 포괄적인 수행평가 목표를 설정해야 한다(Doran, Chan, & Tamir, 1998). 이때 예비교사의 과학 평가에 관한 신념은 목표 설정에 큰 영향을 미친다. 구성주의적인 수행평가를 위해서는 계획 단계에서부터 피드백 등 구성주의적 평가 목적에 대한 올바른 이해가 요구되므로 구상 단계에서는 예비교사들이 구성주의적 평가 목적을 인식하고 이를 다른 PCK 구성요소와 연계할 수 있도록 안내해야 할 것이다.

3.3 조별 논의 과정의 PCK 구성요소 및 구성요소 사이의 연계

예비교사의 조별 논의 과정에서 나타난 PCK 구성요소는 <표 II-7>과 같다. 가장 많이 나타난 PCK 구성요소는 과학 평가에 관한 지식이었다(48.3%). 과학 교수 전략에 관한 지식(9.0%)은 개별 과제 개발 과정의 구상 및 구체화 단계와 비교하여 비율이 크게 감소하였다. 다른 PCK 구성요소는 구상 및 구체화 단계와 비슷하거나 조금 높았다.

<표 II-7> 조별 논의 과정의 PCK 구성요소

PCK 구성요소		1조 (%)		2조 (%)		계 (%)	
A	A1		13 (14.6)		13 (14.6)		26 (29.2)
	A2	21 (23.6)	8 (9.0)	22 (24.7)	9 (10.1)	43 (48.3)	17 (19.1)
C	C1		2 (2.2)		2 (2.2)		4 (4.5)
	C2	2 (2.2)	-	2 (2.2)	-	4 (4.5)	-
I	I1		2 (2.2)		5 (5.6)		7 (7.9)
	I2	2 (2.2)	-	6 (6.7)	1 (1.1)	8 (9.0)	1 (1.1)
K	K1		3 (3.4)		1 (1.1)		4 (4.5)
	K2	5 (5.6)	2 (2.2)	6 (6.7)	5 (5.6)	11 (12.4)	7 (7.9)
L	L1		3 (3.4)		1 (1.1)		4 (4.5)
	L2	5 (5.6)	2 (2.2)	5 (5.6)	3 (3.4)	10 (11.2)	5 (5.6)
	L3		-		1 (1.1)		1 (1.1)
O	O1		9 (10.1)		4 (4.5)		13 (14.6)
	O2	9 (10.1)	-	4 (4.5)	-	13 (14.6)	-
합계		44 (49.4)		45 (50.6)		89 (100.0)	

* A: 과학 평가에 관한 지식(A1: 평가 방법, A2: 평가 영역), C: 과학교육과정에 관한 지식(C1: 연계, C2: 평가 목표), I: 과학 교수 전략에 관한 지식(I1: 주제 전략, I2: 교과 전략), K: 과학 내용에 관한 지식(K1: 개념 및 이론, K2: 과학 탐구 과정), L: 학생에 관한 지식(L1: 오개념, L2: 발달 수준, L3: 동기 및 흥미), O: 과학 평가에 관한 신념(O1: 전통적 평가 지향, O2: 구성주의적 평가 지향).

과학 평가에 관한 지식 중 평가 방법(A1)은 평가 준거 및 배점 설정 13.5%(12

회), 문항 구성 및 유형 검토 7.9%(7회), 평가 방법 선정 및 검토 5.6%(5회)로 나타났다. 개별 과제 개발 과정의 구체화 단계에서와 마찬가지로 조별 논의 과정에서도 예비교사들은 주로 평가의 객관성이나 변별력에 대해 언급하는 경우가 많아서 전반적으로 평가 논의가 전통적 평가 관점에서 이루어졌음을 알 수 있다. 즉, 예비교사들이 대부분 전통적 평가 관점을 유지하고 있을 경우에는 조별 논의를 거치더라도 저절로 구성주의적인 수행평가가 계획되지 않음을 의미한다. 이와 관련된 사례로, 1조의 예비교사들은 관찰 평가를 활용한 예비교사 A, C의 수행평가 과제에 대하여 5단계 리커트 척도를 3단계로 줄이거나 문항 수를 줄이는 등 관찰 평가의 영향을 더욱 줄이는 것을 제안하였다. 또한, 2조의 예비교사들은 <그림 II-3>과 같이 실험을 활용한 예비교사 E의 수행평가 과제에 대하여 주로 변별력 측면에서의 논의만을 진행하였다.

예비교사 D : 난 애매하게 생각하는 게 뭐냐면, 리커트 척도를 사용하는 건 좋은데
1 2 3 4 5 단계가 있잖아. 과연 어디까지가 1단계고 어디까지가 2단계고 3단계고 이게 가장 문제인거지. 차라리 문항을 줄이던가. 1 2 3 4 5 하지 말고.

예비교사 A : 그냥 1 3 5 이런 식으로?

예비교사 D : 아니 그냥 척도를 체크리스트 식으로 해서 뭐 보통 참여, 미참여 이런 식으로. '적극적인 참여' 아니면 '적당한 참여' 아니면 '참여', '미참여' 두 개로 하면 나쁘지 않을 것 같아.

(1조의 조별 논의 중)

예비교사 H : 평가지 양식을 보면 압력과 부피 변화를 그리는 거 아냐. 그렇지? 그러면 그냥 관찰한대로만 그리면, 앞면은 그냥 다 맞을 수 있는 거 아닌가? 그래서 변별력은 뒤쪽에서 두는가? 수행평가의 변별력은.

예비교사 G : 근데 뒤쪽도 문항 그냥 예시로 생각했던 건데. 너무 쉬워가지고 또.

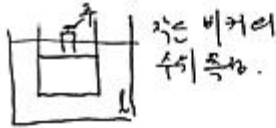
예비교사 F : 이것을 그냥 서술형처럼 바꿔. 그렇게 하면 변별력이 생길 거 같은데.

(2조의 조별 논의 중)

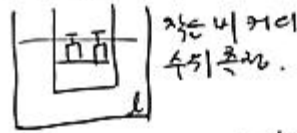
[세부 계획안]

각종 네 번의 실험을 수행하게 될 것이다. 학생들이 모두 관찰할 수 있도록 자리배치를 하거나, 각급내 TV 등을 이용 ~ 등기.

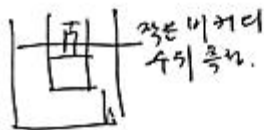
1. 얼음물 속에서.



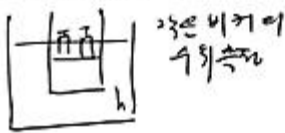
2. 얼음물속에서



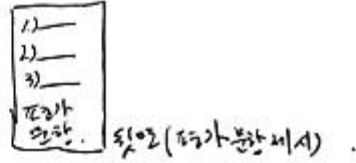
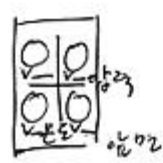
3. 뜨거운 물 속에서



4. 뜨거운 물속에서



나. 각각의 실험은 동시에 진행할 수도 있고, 나누어 진행할 수도 있다. 시간이 준비한 평가표에 각각의 실험을 배치할 수 있다. (학생들 스스로).



[평가 준거]

가. 평가표의 앞면을 올바르게 배치. ... 5점. (실험이 대강 이미 될 때)

뒷면 분할

- 1) 압력이 높을수록 부피가 커진다 (O, X) ... 2점 (내용이 대강 이해)
- 2) 온도가 높을수록 부피가 커진다 (O, X) ... 2점.

다. 앞면을 통해 실험을 제대로 관찰했는지. 확인할 수 있고,

평가표 뒷면을 통해 온도나 부피, 압력나 부피의 상관관계를 올바르게 이해하고 있는지 확인할 수 있다.

<그림 II-3> 예비교사 E가 개발한 수행평가 과제

과학교육과정에 관한 지식에서는 각 조에서 수행평가를 교육과정의 연계(C1) 측면에서 검토하는 예비교사가 있었다. 따라서 과학교육과정에 관한 지식이 나타난 빈도는 매우 낮았지만, 개별 과제 개발 과정에서 교육과정의 연계를 고려하지 못했던 예비교사들의 계획안이 조별 논의를 통해 보완될 수 있었다.

예비교사 D : 내가 알기론 이 단원에 상태변화가 안 나온 거로 알고 있거든. 상태변화가 이 단원에 이어서 나오는 거로 알고 있는데, 벌써 상태변화에 대한 개념을 물어보는 건 틀린 거 같은데. 이거 상태변화는 아직 안 배웠어.

예비교사 B : 상태변화가 뭐?

예비교사 D : 안 나와. (중략) 이 바로 다음 단원 이어서 나오는 게 상태변화야. 상태변화라는 말을 지우고 오로지 분자 운동, 분자 간의 거리 이런 거로만 얘기하니깐. 고체가 기체가 되고 기체가 액체가 되는 이런 거. 아무튼, 아직 안 나오거든. 그래서 아마 상태변화라는 말을 빼야 될 거야. 빼면 될 것 같고.

(1조의 조별 논의 중)

과학 교수 전략에 관한 지식에서는 주제 전략(I1)의 비율이 크게 감소하였는데, 이는 조별 논의가 기존 평가 계획을 수정, 보완하는 방식으로만 진행되고 새로운 수업 기법이나 전략에 관한 논의는 거의 없었기 때문이다. 예비교사들의 주된 검토 대상은 오개념을 유발할 요소, 이론적 측면에서 실험 설계의 타당성, 실험의 안전이나 시간 배분 등이었다. 하지만 이러한 검토도 대부분 전통적 평가 관점에서 이루어졌다.

학생에 관한 지식은 논의 과정에서 전반적으로 낮은 빈도를 보였으나 개별 과제 개발 과정과는 다른 특징이 나타났다. 학생의 오개념(L1)은 개별 과제 개발 과정에서는 거의 고려되지 않았던 PCK 구성요소였으나 조별 논의 과정에서는 학생의 오개념에 관한 지식이 상대적으로 많이 나타났다(40.0%, 4회). 구성주의적 평가 관점에서 학생에 관한 지식은 개별 학생의 특성을 고려하고 학생이 일반적으로 가지는 흥미와 동기, 오개념, 발달 수준 등을 종합적으로 고려하는 것을 의미한다(박선미, 1998). 그러나 논의 과정에서 이러한 고려는 거의 없었으므로 조별 논의는 수행

평가의 과학적 오류 제거나 타당도와 신뢰도 개선 측면에서는 어느 정도 기여하였으나 구성주의적 관점에서는 충분하지 못했다고 볼 수 있다.

과학 평가에 관한 신념은 두 조 모두 전통적 평가 지향(O1)이 강하게 나타났다. 다음 사례는 <그림 II-4>와 같이 직소 모형을 활용한 예비교사 H의 ‘일일 교사되기’ 수행평가 과제에 대한 2조의 조별 논의 내용이다. 예비교사 H는 수행평가 과제에서 학생 중심의 활동과 이를 보조하는 교사의 역할을 계획하고, 개별 과제 개발 과정에서도 학습 기회 제공 등의 구성주의적 평가 지향을 나타냈으나, 이에 관하여 동료 예비교사들은 다음과 같이 전통적인 관점에서 부정적인 의견을 주로 제시하였다.

예비교사 E : 나 이거 근데 개인적으로 생각되는 건, 조별 간의 수준을 과연 일정하게 유지할 수 있을까라는 생각이.

예비교사 G : 어느 팀은 수준이 높고 낮고 서로 가르쳐 주기에는 부족하고.

예비교사 E : 어떤 조는 너무 의욕이 없어서 아예 압력에 대해서 그냥 ‘압력이 뭐다 모르겠어요.’ 하면, 개네는 그냥 평가가 엄청 낮은 거잖아. 그니까 이런 거는 교사가 처음부터 고려할 수가 있을까? (중략) 시도는 좋은 거 같아. 애들이 굉장히 자율적이고 지적이고 탐구심이 강하고 이런 존재라고 가정 하에, 굉장히 이상적인 학생들이란 가정 하에 할 수 있는 거 같은데.

예비교사 F : 참여도 외에 이게 과학 주제를 서로 가르쳐 주고 배우고 하는 거니까 그 주제에 대한 이해도를 이 체크리스트를 통해서 평가를 하겠다는 거 같은데 (중략) 교사평가를 하면 이제 그 한 클래스에서 대단히 진행하기 좀.

예비교사 E : 개인적으로 드는 생각이 우리는 참 조별과제에 대해서 다 부정적인 거 같은 말이야. 거의 다 그렇지, 다 부정적이잖아. 정말 조별과제에서 생길 수 있는 모든 폐해를 다 가질 수 있는 수행평가인 거 같아, 잘못된다면.

(2조의 조별 논의 중)

[세부 계획안]

시작할 때, 학생들에게 숙형평가가 미리 설명한다. 교사는 학생들이 수업을 들으면서 한번에 이해하기 힘든, 예로 들면 압력과 같은 개념 등을 제시 한다. 그다음 조를 나누어서 (3~4명) 조별로 주제를 선택하게 한다.

주제를 모두 나누어 주고 학생들에게 조별로 주제에 관해 의견등을 느꼈는지, 그래서 자신들은 어떻게 이해를 했는지, 어떻게 하면 더 쉽게 이해할 수 있는지, 디테일하게 설명해 주는 것을 학생들에게 말해주고 조별회의 시간을 준다.

이때 교사는 학생들이 토의가 원활하게 이루어지는지, 또 학생들이 도어를 하면서 올바른 방향으로 접근하는지, 또 주제에 어려움은 느끼는 학생들에게 도움을 주며, 모든 학생들이 참여가 이루어지도록 한다.

모두 토의를 마치면 학생들은 다른 조에게 처음에 소개해줬던 것 그대로 자신들의 주제에 대한 토의 내용을 설명해준다. 다른 학생들은 발표를 들으면서 질문하거나 발표한 학생들이 내용을 정리한다.

교사는 학생들이 발표할 때에 틀린 점이나 보강할 것을 적어 두며 모든 조의 발표가 끝나면 함께 적어 놓았던 모든 학생들에게 알려준다.

<그림 II-4> 예비교사 H가 개발한 수행평가 과제

2조의 예비교사 H는 각 조에서 객관성과 공정성을 강조하는 다른 조원들보다 좀 더 구성주의적 시각을 가지고 있었지만, 조별 논의 과정에서는 자신의 생각을 적극적으로 다른 조원들에게 표현하지 못하였다. 1조에서도 예비교사 C는 다른 조원들로부터 객관성 측면에서 문제가 있다는 지적을 받았지만, 자신의 과제가 학생의 흥미와 참여를 이끌어낼 수 있는 좋은 방법이라고 생각하여 조별 논의 후에도 실험을 대체하지 않고 보완하고자 하였다. 따라서 조별 논의에서 구성원이 적극적으로 구성주의적 측면의 의견을 개진할 수 있는 분위기를 조성할 필요가 있다.

한편, 조별 논의 과정에서 나타난 PCK 구성요소 사이의 연계 분석 결과는 <표 II-8>, <표 II-9>와 같다. PCK 구성요소 사이의 연계 빈도는 총합 7회로 연계가 거

의 나타나지 않았다고 볼 수 있다.

<표 II-8> 조별 논의 과정의 PCK 구성요소 사이의 연계 (통합 유형별)

통합 유형	1조 (%)	2조 (%)	계 (%)	
2가지 PCK 구성요소의 통합	A - C	1 (14.3)	-	1 (14.3)
	A - I	1 (14.3)	-	1 (14.3)
	C - I	1 (14.3)	-	1 (14.3)
	I - L	2 (28.6)	2 (28.6)	4 (57.1)
계	5 (71.4)	2 (28.6)	7 (100.0)	

* A: 과학 평가에 관한 지식, C: 과학교육과정에 관한 지식, I: 과학 교수 전략에 관한 지식, K: 과학 내용에 관한 지식, L: 학생에 관한 지식, O: 과학 평가에 관한 신념.

<표 II-9> 조별 논의 과정의 PCK 구성요소 사이의 연계 (PCK 구성요소별)

PCK 구성요소	1조 (%)	2조 (%)	계 (%)
A와 다른 PCK 구성요소의 연계	2 (14.3)	-	2 (14.3)
C와 다른 PCK 구성요소의 연계	2 (14.3)	-	2 (14.3)
I와 다른 PCK 구성요소의 연계	4 (28.6)	2 (14.3)	6 (42.9)
K와 다른 PCK 구성요소의 연계	-	-	-
L과 다른 PCK 구성요소의 연계	2 (14.3)	2 (14.3)	4 (28.6)
O와 다른 PCK 구성요소의 연계	-	-	-
계	10 (71.4)	4 (28.6)	14 (100.0)

* A: 과학 평가에 관한 지식, C: 과학교육과정에 관한 지식, I: 과학 교수 전략에 관한 지식, K: 과학 내용에 관한 지식, L: 학생에 관한 지식, O: 과학 평가에 관한 신념.

두 개 조에서 유일하게 공통으로 나타났던 PCK 구성요소 사이의 연계는 과학 교수 전략에 관한 지식과 학생에 관한 지식 사이의 연계였으며, 주로 평가 문항이

나 실험 설계가 학생의 오개념을 유발할 가능성이 있는지에 대한 논의가 이루어졌다. 다음은 <그림 II-5>와 같이 달걀을 삶는 활동을 통해 샤를 범칙에 관한 학생들의 이해를 평가하려는 예비교사 C의 수행평가 계획안에 대한 1조의 논의 중 일부이다.

예비교사 D : 근데 이것만 하면 되지. '온도가 올라가서 부피가 커져서 터졌다.'라는 사실만 알면 돼.

예비교사 A : 아니 오개념을 유발하잖아, 실험이.

예비교사 B : 근데 애가 분자운동을 빨리 할 거 아냐. 그럼 껍질들이 분자운동해서 터진다고 생각하면.

예비교사 D : 오 그것도 있네. 그 정도까지 생각할까요? 애들이.

예비교사 C : 교과서에 보니까 식초나 소금을 넣으면 껍질이 단단해져서 쉽게 깨지지 않는다는데 그런 것까지 넣으면 복잡해지니까.

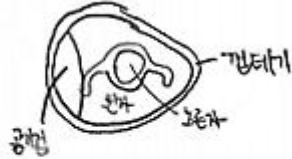
예비교사 D : 그리고 분자 운동 개념은 제자리에서 진동만 할 뿐 이동하진 않는다는 얘기가 나오니까, 그거 때문에 깨진다는 생각하지 않지 않을까?

예비교사 A : 재밌는데 오개념이 생기기 쉬운 실험이네.

(1조의 조별 논의 중)

[세부 계획안]

실험 전에 계란의 구조와 사육의 방식에 대해 설명을 한다.



"계란의 부패는 온도에 관계있다"

4번 노가 되어
병실에 넣어 온 계란과 실외 온도를 비교
실험 마친 뒤 결과를 나타낸다.



[평가 준거]

<p>설명 (30%)</p> <p>1. 학생의 강요도: 과학 탐구의 과정의 순서까지 상세 설명을 하였다. 과정에 대한 설명은 생략된 부분이 나오지 않았고, 실험의 관련 과정에 대해서도 잘 설명하였다.</p> <p>2. 결과: 계란의 상태 (계란의 크기)에 대해 잘 설명하였다.</p>	<p>평가 (70%)</p> <p>1. 물리 평가: 과학 실험에 관한 기본 사항들이 보인 (10%) (실험 방법, 안전 등)</p> <p>2. 결과 평가: 실험의 결과가 온도에 따라 어떻게 변하는지 잘 설명하였다. (60%)</p> <p>(논의는) "물의 온도, 실험을 할 때 온도를 어떻게 조절하는지" 등 관련 사항의 설명 - 30. 결과 설명 경우 - 0</p>
--	--

<그림 II-5> 예비교사 C가 개발한 수행평가 과제

하지만 조별 논의 과정에서 이러한 PCK 구성요소 사이의 연계는 거의 나타나지 않았으며, 대부분 특정 PCK 구성요소에 관한 논의가 주로 이루어졌다. 교사의 PCK 수준은 PCK 구성요소의 연계 능력에 좌우되므로 (Friedrichsen *et al.*, 2009; Krauss *et al.*, 2008) 교사는 다양한 PCK 구성요소를 활용하는 능력 뿐 아니라 PCK를 통합적으로 연계하여 활용하는 능력을 갖추어야 한다(Loughran, Berry, & Mulhall, 2006). 따라서 조별 논의 과정에서도 여러 가지 PCK 구성요소를 통합적으로 검토할 수 있는 방안을 마련할 필요가 있다.

4. 결론

연구 II에서는 예비교사의 수행평가 전문성 향상 방안을 마련하기 위한 기초 연구의 일환으로 중등 예비 화학교사들의 수행평가 개별 과제 개발 및 조별 논의 과정에서 나타나는 PCK 구성요소와 PCK 구성요소 사이의 연계를 분석하였다.

우선 예비교사들은 평가의 공정성과 객관성을 중시하는 전통적 평가관을 바탕으로 수행평가를 개발하는 것으로 나타났다. 예비교사들이 주로 고려한 PCK 구성요소는 과학 평가에 관한 지식에 편중되어 있었으며, 구상 단계보다 구체화 단계에서 전통적 평가 지향이 더 강하게 나타났다. 또한 구상 단계에서는 여러 가지 평가 영역과 다른 PCK 구성요소들을 고려하였으나 구체화 단계에서는 교사의 주관을 가능한 배제하는 방향으로 평가 준거 및 배점을 설정하였다. 평가 영역에서 과학 과정 지식이 배제된 것도 이를 객관적으로 평가하는 데 예비교사들이 어려움을 느꼈기 때문으로 볼 수 있다. 교수 전략이나 평가 결과의 활용 측면에서도 학생에 대한 피드백은 거의 고려되지 않았으며, 대부분의 평가는 학생 활동이 끝난 후 이루어지는 전통적인 총괄평가와 유사한 형태로 개발되었다. 그러므로 예비교사의 수행평가에 대한 전문성 향상을 위해서는 구성주의적 평가를 구현하는 평가 체계로서 수행평가의 가치와 구성주의적 평가의 중요성에 대한 인식이 선행되어야 한다. 이를 바탕으로 수행평가에서 구체적인 세부 활동의 도입 의도와 효과, 평가 결과의 활용 방법 등을 구성주의적 평가 관점에서 미리 생각해보는 사례 분석 활동을 실시하는 것도 필요할 것이다.

PCK 구성요소 사이의 연계는 새로운 아이디어의 도입이나 완성된 수행평가의 검토 과정 등 특정한 상황에서 주로 이루어졌는데, 대부분의 연계는 구상 단계에서 과학 교수 전략에 관한 지식을 중심으로 나타났다. 나머지 PCK 구성요소 사이의 연계는 양과 질 측면에서 모두 부족하였다. 구체화 단계 및 조별 논의 과정에서는 PCK 구성요소 사이의 연계가 거의 나타나지 않았다. 수행평가 개발 과정에서 PCK 구성요소 사이의 연계가 충분히 이루어지기 위해서는 구체화 단계에서도 학생이나 교육과정에 대한 고려가 이루어져야 한다. 완성된 수행평가 문항에는 개발자가 고려했던 PCK 구성요소가 생략되거나 구체적으로 드러나지 않을 수 있으므로, 예비교사들에게 실질적인 도움이 되기 위해서는 수행평가의 개발 과정을 구체적으로

제시하고 다양한 PCK 구성요소가 어떻게 적용되어 있는지 분석하는 활동을 실시할 필요가 있다. 한편, 예비교사들은 학생에 대한 경험적 지식이 부족하므로 일반적인 학생의 오개념과 발달 수준 등에 대해서는 구체적인 관련 자료를 제공하여 이를 반영하도록 할 필요가 있다. 또한, 구상 및 구체화 단계에서 예비교사들이 잘 고려하지 않는 것으로 나타난 PCK 구성요소를 점검표 형태로 제공한다면 예비교사들의 수행평가 개발에 도움을 줄 수 있을 것이다.

조별 논의 과정에서 나타난 PCK 구성요소의 특징은 전반적으로 구체화 단계와 유사하였다. 논의 중 고려된 PCK 구성요소는 과학 평가에 관한 지식에 편중되어 있었고 대부분의 논의는 목표 개념을 정확하게 전달하고 공정하게 평가하는 것을 목표로 하는 전통적 평가 관점에서 이루어졌다. 조별 논의를 통하여 교육과정의 연계, 오개념 등 예비교사들이 고려하지 못했던 일부 PCK 구성요소가 보완된 측면이 있었지만, 평가의 타당도와 신뢰도를 높여려는 목적으로 이루어졌으므로 수행평가가 구성주의적으로 개선되지는 못하였다. 즉, 대부분 전통적 평가관을 가지고 있는 예비교사들 사이에 논의 기회를 제공하는 것만으로는 수행평가가 저절로 구성주의적으로 개선되기는 어렵다고 볼 수 있다. 따라서 조별 논의에서는 수행평가 경험이 풍부한 전문가가 멘토링을 제공하거나 조별 논의 시 고려해야 할 사항들을 점검표 형태로 제공하여 예비교사의 반성적 활동을 촉진하고 구성주의적 측면에서의 논의가 이루어질 수 있도록 도울 필요가 있다.

이때, 점검표는 예비교사가 반드시 따라야 할 상명하달식 지침을 제공하는 것이 아니라, 자신과 동료의 수행평가 과제를 분석하고 유의미한 피드백을 제공하는 능력이 충분치 못한 예비교사의 반성적 활동을 효과적으로 지원하기 위해 제공되는 도구이다(이선, 2017). 그러므로 점검표를 활용할 때에는 예비교사가 각 PCK 구성요소의 의미를 이해할 수 있는 형태로 명료하게 제시하고, 점검표에 대한 예비교사의 이해도를 높이고 효과적으로 활용할 수 있도록 점검표의 각 항목에 대하여 바람직한 예시를 제공하거나 점검표를 활용한 연습 활동을 함께 할 필요가 있다(김덕영, 박종원, 2015). 또한, 점검표에서 예비교사들이 여러 가지 PCK 구성요소를 함께 검토하고, 수행평가 과제에 전통적, 구성주의적 평가 특성이 어떻게 반영되어 있는지 검토할 수 있도록 명시적인 지침을 제공함으로써 예비교사가 구성주의적 수행평가를 지향하면서도 평가의 전통적 측면을 함께 고려하도록 할 수 있다. 이와 같

은 기준에 따라 제작된 점검표를 예비교사의 개별 과제 개발 과정 전반에 걸쳐 참고하도록 하고 조별 논의 과정에서도 활용하도록 한다면, 예비교사가 수행평가 과제 개발 과정에서 다양한 PCK 구성요소를 통합적으로 활용하고, 조별 논의 과정에서도 이에 관한 논의가 일관성 있게 이루어지도록 하는 데 도움이 될 수 있을 것이다.

제 5 장. 수행평가 전문성 향상 프로그램에 참여한 예비 화학교사의 수행평가 과제 개발 과정에서 고려된 PCK 구성요소 및 연계의 특징 (연구 III)³⁾

1. 서 론

과학교사의 전문성 향상은 지식 기반 사회의 도래에 따른 과학교육의 변화와 혁신에서 핵심 요소이다(임청환, 2003; van Driel, Beijard, & Verloop, 2001). 교사의 전문적 지식 체계를 나타내는 PCK는 교사의 전문성 수준을 대표하는 지표이므로(김경순 등, 2011; 노태희 등, 2010), 과학교사의 PCK를 개발하기 위한 전략의 모색이 이루어지고 있다. 교사의 PCK는 그 구성요소를 활용하거나 구성요소 사이의 상호작용을 촉진하는 활동으로 개발할 수 있다(Aydin & Boz, 2013; Park & Oliver, 2008). 예를 들어, 다양한 PCK 구성요소를 활용하는 교수 설계 활동은 초등 예비교사의 PCK 사용을 촉진함으로써 PCK를 향상시키는 것으로 나타났다(Beyer & Davis, 2012).

과학 평가에서도 PCK 구성요소의 활용은 중요한데(노태희, 박재성, 강훈식, 2016; 조희형, 고영자, 2008; Falk, 2012), 특히 학생의 능동적인 참여가 필수적이고 실험, 토의 등 다양한 방법으로 이루어지는 수행평가 과정에서는 다양한 PCK 구성요소의 활용이 더욱 중요하다. 이러한 맥락에서 현직교사들을 대상으로 협력적 수행평가 워크숍 활동의 효과(Falk, 2012)나 학교 현장의 수행평가 상황에서 나타나는 교사의 PCK에 대한 연구(민희정, 2012)가 이루어졌다. 또한 교사의 평가 전문성은 단기간에 향상되기 어려우므로 예비교사 단계에서부터 지속적인 교육이 이루어져야 한다는 관점에서 예비 화학교사의 수행평가 과제 개발 과정에서 나타난 PCK 구성요소의 특징을 분석한 연구 II도 진행되었다. 그 결과, 예비교사들이 수행평가 과제 개발 과정에서 고려하는 PCK 구성요소는 과학 평가에 관한 지식에 편중되어 있

3) 연구 III은 한국과학교육학회지 38호 4권의 505-518쪽에 게재된 ‘구성주의적 수행평가 워크숍에 참여한 예비 화학교사의 수행평가 과제 개발 과정에서 고려된 교과교육학 지식(PCK) 구성 요소 및 연계의 특징’의 내용을 보완하고 재구성한 것임.

있고 PCK 구성요소 사이의 연계도 거의 이루어지지 않은 것으로 나타났다. 또한 수행평가 과제가 평가의 공정성과 객관성에 중점을 두는 총괄평가의 형태로 개발되어 구성주의적 평가와도 거리가 멀었다.

수행평가는 성적 산출과 선발 위주의 전통적 평가 패러다임을 극복하고 구성주의적 평가를 구현하기 위해 도입된 평가 체계이다(교육부, 1998). 따라서 교사의 수행평가 전문성의 핵심은 구성주의적 수행평가 과제를 개발 및 실행하는 능력이라고 볼 수 있다. 구성주의적 수행평가란 구성주의적 평가관에 기초하여 이루어지는 수행평가로서, 학습 결과뿐 아니라 지식의 구성 과정과 고등 사고능력에 대한 수행평가, 다양한 관점과 양식으로 이루어지는 수행평가를 의미한다(박선미, 1998). 또한 PCK 구성요소들이 골고루 활용되고 PCK 구성요소들 사이에 충분한 상호작용이 이루어질 때 교사의 PCK는 효과적으로 발달된다(Friedrichsen *et al.*, 2009; Park & Chen, 2012). 그러므로 예비교사에게 단순히 수행평가 과제 개발 기회를 제공하는 것만으로는 예비교사의 PCK 계발과 수행평가 과제 개발 능력 향상에 한계가 있다.

따라서 예비교사의 PCK를 계발하고 수행평가 전문성을 함양하기 위해서는 구성주의적 평가 목적에 대한 이해를 바탕으로 수행평가 과제 개발 과정에서 고려해야 하는 다양한 PCK 구성요소를 접하고 활용하는 기회를 제공할 필요성이 있다. 이에 선행연구에서는 PCK를 계발하는 방안으로 숙련된 교사나 전문가의 컨설팅, 멘토링, 코칭 등을 통하여 적절한 도움을 제공하는 방법을 많이 활용하였다(고문숙, 남정희, 2013; 양찬호, 2015; 정금순, 강훈식, 2011; Stanulis, Little, & Wibbens, 2012). 또한 PCK가 경험을 통해 발달하는 실천적 지식임을 고려할 때(임청환, 2003; Bond-Robinson, 2005), 예비교사가 교육현장에서 학생들을 대상으로 수업을 실시한 후 이를 평가하기 위한 수행평가 과제를 개발하도록 하는 것도 이상적이다. 하지만 이러한 방법들을 예비교사 교육현장에 폭넓게 적용하기에는 현실적으로 한계가 있으므로 보다 실천적인 접근법을 검토할 필요성이 있다(노태희, 박재성, 강훈식, 2016; 박재성, 강훈식, 한재영, 2017). 이에 연구 III에서는 예비교사가 PCK를 효과적으로 계발하고 수행평가 전문성을 향상시킬 수 있는 프로그램을 고안하여 적용하였다.

프로그램이 결과론적 지식을 전달하는 전통적인 강의 형태로 이루어지면 해당 내용에 대한 교사의 이해는 피상적 수준에 머물기 때문에 교사에게서 의미 있는

변화를 끌어내기 어렵다(최진영, 송경오, 2005; van Driel, Beijard, & Verloop, 2001). 따라서 프로그램 또한 구성주의적 관점에서 개발되어야 한다. 이러한 맥락에서 Loucks-Horsley 등(2009)은 프로그램에서의 교사 전문성 개발 원리와 전문성 개발 설계 프레임틀을 제안하였다. 연구 III에서 고안한 프로그램은 이에 기반하여 구성주의적 수행평가에 대한 이론과 실행의 통합을 도모하는 방향으로 구성하였다. 먼저 구성주의적 수행평가의 이론적 근거를 소개하여 예비교사가 스스로 구성주의적 수행평가의 타당성에 공감하고 적용 방법을 판단할 수 있도록 하였다(Beyer & Davis, 2009). 또한 다양한 수행평가 사례와 예시 분석 및 과제 개발 실습 과정에서는 연구자와 예비교사들이 서로의 관점을 공유하며 공동으로 이론과 실행의 통합을 추구하는 협력적 논의 활동을 실시하였다(김구연, 2010; 이은진 등, 2007). 이러한 협력적 논의를 통해 예비교사는 연구자나 다른 예비교사의 견해를 접하며 자신이 지닌 평가관의 특징과 한계를 발견하고 구성주의적 평가에 대한 견해를 정교화할 수 있다. 특히 예비교사는 실행 중심의 관점을 지니는 것에 비해 연구자는 상대적으로 이론 중심의 관점을 지니므로, 연구자가 유기적으로 참여하는 협력적 논의 활동은 구성주의적 평가의 이론과 실행 사이의 간극을 좁히는 데 기여할 수 있다(신옥순, 1997; Herrenkhol, Kawasaki, & Dewater, 2010). 또한 중등학교를 졸업하고 현직교사가 되기 위한 중간 과정에 있는 예비교사의 특수한 입장을 고려하여(김선희, 2012), 구성주의적 수행평가를 현장에 적용할 때의 어려움과 극복 방안을 교사와 중등학생의 입장에서 논의하도록 하였다. 예비교사들은 구성주의적 평가 및 교수학습관 측면에는 공감하면서도 이를 우리나라의 교육현장에 실제로 적용하기는 어렵다고 생각하는 경향이 있으므로(김선영, 2014), 이러한 활동도 구성주의적 평가의 이론과 실행의 통합을 촉진하는 데 도움을 줄 것으로 기대된다.

이에 연구 III에서는 수행평가 전문성 향상 프로그램에 참여한 예비 화학교사들이 개별로 수행평가 과제를 개발하는 과정과 이에 대하여 조별로 논의하는 과정에서 고려하는 PCK 구성요소와 요소 사이의 연계의 특징을 분석하였다. 이를 통하여 예비교사에게 제공한 프로그램이 미친 영향과 예비교사가 인식하는 구성주의적 수행평가의 특징을 파악하였고, 예비교사의 PCK를 효과적으로 개발하고 수행평가 전문성을 향상시키기 위한 시사점을 탐색하였다.

2. 연구 방법

2.1 연구 참여자

연구 II와 동일한 조건으로, 충청북도 소재의 1개 대학교에서 과학과 교재연구 및 지도법 강좌를 수강 중인 화학교육과 3학년 학부생 중 졸업 후 교직 진출을 희망하는 예비교사 8명(남학생 3명, 여학생 5명)이 연구에 참여하였다. 연구 참여자 중 7명은 지난 학기 다른 강좌에서 수행평가의 개념과 방법 등 기초적인 수행평가 이론을 학습한 상태였고, 다른 1명의 연구 참여자는 이번 학기에 과학과 교재연구 및 지도법과 교육평가 강좌를 동시에 수강중인 상태였다. 또한, 연구 참여자 중 1명은 다른 강좌에서 수행평가 과제를 개발한 경험이 있었다. 연구 참여자의 구체적인 정보는 <표 III-1>과 같다.

<표 III-1> 연구에 참여한 예비교사의 배경

예비교사	1조				2조			
	A	B	C	D	E	F	G	H
성별	여성	여성	남성	남성	여성	여성	여성	남성
나이	23	23	25	25	23	23	23	25
평가 관련 강좌* 경험	있음	있음	있음	수강중	있음	있음	있음	있음

* 과학교육론, 교육평가.

2.2 수행평가 전문성 향상 프로그램

자료 수집은 주당 3시간씩 4주 동안 이루어졌으며, 첫째 주부터 셋째 주는 수행평가 전문성 향상 프로그램을 진행하였다. 프로그램의 설계는 이론적 배경에서 논의한 <그림 3>의 전문성 개발 설계 프레임을 이용하여 다음의 과정을 거쳤다.

첫째 단계는 연구자의 지식과 신념을 반영한 비전 및 기준 설정이다. 이론적 배

경에서 논의한 바와 같이 수행평가의 이론적, 철학적 배경과 특징은 구성주의적 평가와 일맥상통하며, 우리나라에 수행평가가 도입된 이유도 학생의 학습 결과만을 측정하고 평가와 수업이 분리된 전통적 지필평가의 한계를 극복하기 위함이다. 따라서 이 연구에서 개발하고자 하는 수행평가 전문성 향상 프로그램의 비전 및 목표는 구성주의적 평가를 구현하는 평가 체계로서의 수행평가 즉, 구성주의적 수행평가가 개발될 수 있도록 하는 것이다. 이에 수행평가 과제에서 구성주의적 평가의 특징인 지식 구성의 과정에 대한 평가, 단순 암기를 지양하는 고등 사고능력에 대한 평가, 다양한 관점과 양식이 존중되는 평가(박선미, 1998)가 구현되도록 하는 것을 비전 및 기준으로 설정하였다.

둘째 단계는 학습 및 기타 자료 분석이다. 이는 프로그램의 구체적인 목표 설정을 위한 자료 수집의 단계로, 이 연구에서는 연구 I과 II의 결과를 분석하였다. 먼저 연구 I에서는 교사들이 구성주의적 평가의 목적과 방법을 이해하는 것과 실제로 수행평가 과제를 구성주의적으로 개발하는 것은 별개의 문제로, 평가 실태와 지향 및 평가 방법과 목적 사이의 간극이 존재할 수 있음을 밝혔다. 연구 II에서는 예비교사들이 전통적 평가관을 바탕으로 수행평가 과제를 개발하여 프로그램의 비전 및 기준인 구성주의적 수행평가와는 거리가 먼 것으로 나타났다. 또한, PCK 측면에서는 대부분의 PCK 구성요소가 과학 평가에 관한 지식에 편중되어 있었고, PCK 구성요소 사이의 연계도 구상 단계를 제외하면 거의 나타나지 않았다. 한편, 이 단계에서는 프로그램을 적용할 구체적인 맥락을 고려해야 한다. 예비교사를 대상으로 한 평가 전문성 향상 프로그램의 개발 필요성이 도출된 연구 I의 결과에 따라 사범대학의 예비교사 교육 환경을 고려하였고, 담당 교수의 협조를 요청하여 과학 평가 및 교수학습에 관한 이론과 실습을 활동이 이루어지는 과학과 교재연구 및 지도법 강좌에 평가 전문성 향상 프로그램을 도입하고, 해당 강좌의 수강생 중 일부를 연구 참여자로 선정하였다.

셋째 단계는 목표 설정이다. 이 단계에서는 연구 I과 II의 결과를 토대로 두 가지의 목표를 설정하였다. 먼저 연구 I의 결과를 바탕으로 설정한 목표는 예비교사들의 구성주의적 수행평가 과제 개발 과정에서 나타날 수 있는 간극을 좁혀 구성주의적 수행평가의 이론과 실행을 통합시키는 것이다. 연구 II의 결과를 바탕으로 설정한 목표는 PCK 측면에서 예비교사들이 다양한 PCK 구성요소를 활용하면서도

PCK 구성요소 사이의 연계가 활발히 일어나도록 하는 것이다. 이 단계에서는 고려해야 할 핵심적 문제 중 충분한 시간 및 공공 지원 확보 등의 문제는 과학과 교재 연구 및 지도법 강좌 담당 교수의 협조를 통하여 해결하였고, 리더십 개발, 역량 구축 등의 문제는 다음 단계의 전략에 최대한 반영하였다.

넷째와 다섯 번째 단계는 계획 및 실행 단계이다. 이 단계에서는 목표 달성을 위한 전략이 선정되고 투입되어야 한다. 이를 위해 Loucks-Horsley 등(2009)의 교사 전문성 개발 원리 중에서 선정한 예비교사 맥락에 적용할 수 있는 5가지 요소를 고려하여 프로그램에 투입할 전략을 선정하고 활동을 구성하였다. 먼저 ‘효과적인 교수학습에 대한 정립된 이미지의 제공’을 위하여 구성주의적 수행평가에 대한 이론적 배경과 특징 및 PCK와 구성주의적 수행평가의 관계를 소개하였다. 다음으로 ‘학생의 학습 목표 및 요구의 충족’을 위하여 구성주의적 수행평가를 현장에 적용할 때의 어려움과 극복 방안을 교사와 중등학생의 입장에서 분석하고 논의하였으며 수행평가 과제 개발 시 학생 오개념 자료를 제공하였다. ‘PCK 및 기술을 개발하고 비판적으로 검토 및 반영할 기회의 제공’을 위해서는 연구자가 선정한 구성주의적 수행평가 예시를 PCK 및 구성주의적 관점에서 분석하고 논의하는 활동을 실시하였고, 직접 구성주의적 수행평가 과제를 개발하는 실습 활동에서는 PCK 측면에서의 점검표를 제공하였다. ‘다른 동료 및 전문가들과 함께 하는 기회의 제공’ 측면에서는 다양한 평가 방법을 구성주의적 목적으로 활용하는 방법에 관한 논의 활동과 예비교사들이 자신이 학생일 때 경험했던 수행평가 사례를 공유하고 논의하는 활동을 도입하였고, 자신이 개발한 수행평가 과제의 개선 방향에 대해서도 조별 논의 활동을 실시하였다. 이때 ‘성인 학습자로서 예비교사의 능동적 참여 촉진’을 위하여 연구자와 예비교사 사이의 협력적 논의 활동을 실시하였다. 연구자는 논의 주제를 제공한 후 예비교사들과 동등한 위치에서 서로에게 ‘비판적 친구(critical friend; Fenstermacher & Richardson, 1993)’의 역할을 하며 논의에 참여하였다. 즉, 논의 과정에서 연구자는 이론 중심의 관점에서 예비교사의 의견에 대한 평가와 대안을 제시하였으며, 예비교사는 실행 중심의 관점에서 연구자의 이론적 의견에 대한 현장 적용 가능성을 검토하고 지적하면서 구성주의적 수행평가의 이론과 실행의 통합을 위해 유기적으로 협력하였다(이은진 등, 2007; 신옥순, 1997). 프로그램의 구체적인 활동은 연구 I, 연구 II의 결과 및 워크숍 또는 교사연

수 관련 선행연구를 분석하여 구성하였다.

첫째 주에는 구성주의적 수행평가에 대한 이론적 근거를 소개하였다. 연구 I에서 교사들은 일반적으로 과학 평가 방법에 비해 구성주의적 평가 목적에 대한 이해가 부족한 것으로 보고되었으므로, 예비교사들이 구성주의적 평가의 의미와 중요성을 이해하고 내면화하는 데 중점을 두었다. 선행연구(백순근, 2000; 박선미, 1998)를 바탕으로 구성주의적 평가의 개념과 주요 특징을 설명하고 수행평가의 국내 도입 배경에 근거하여 구성주의적 평가를 구현할 수 있는 것이 수행평가임을 강조하였다. 또한, 수행평가와 PCK의 관계를 다룬 연구 II와 선행연구(민희정, 2012; Falk, 2012; Kind, 2009; Park & Chen, 2012; Park & Oliver, 2008)를 바탕으로 구성주의적 수행평가 과제 개발에서 고려할 요소를 평가, 교육과정, 교수 전략, 과학 내용, 학생의 측면에서 정리하여 예시와 함께 소개하였다. 이후 연구자는 평가 방법의 다양한 예를 제시하였고, 예비교사들은 연구자와 함께 다양한 평가 방법을 구성주의적 목적으로 활용하는 방법에 대해 논의하였다. <부록 III-1>에 이때 활용한 PPT 자료를 제시하였다.

둘째 주에는 구성주의적 수행평가의 실천 능력을 증진시키기 위하여 예비교사들을 4명씩 2개 조로 편성한 후 다양한 수행평가의 예시 분석 및 논의 활동을 하였다. 이때 선행연구(장수웅, 2002; 홍수연, 2006)와 한국교육과정평가원의 수행평가 지원 포털⁴⁾에서 이 연구의 과제와 무관하고 개발 과정이 포함된 구성주의적 수행평가 과제 예시 5가지를 소개하여, 예비교사들이 수행평가 과제 개발 과정에서 고려된 PCK 구성요소도 분석할 수 있도록 하였다. 예비교사들은 첫째 주의 프로그램 내용을 바탕으로 각 수행평가 예시에 구성주의적 평가의 특징과 요소가 어떻게 고려되어 있는지 조별로 분석한 후, 연구자와 함께 분석 결과 및 구성주의적 관점에서의 개선 방향에 대하여 논의하였다. 이후 예비교사들은 자신이 중등학생일 때 경험했던 수행평가 사례를 조별로 발표하고, 같은 관점에서 문제점과 개선 방향 등을 논의하였다. 이때 연구자는 연구 I과 선행연구(민희정, 2012)에 보고된 과학 수행평가 사례를 제시하며 협력적 논의에 참여하였다. 또한, 논의 전반에 걸쳐 구성주의

4) 연구 당시에는 '수행평가 지원 포털'을 이용하였으나, 현재 한국교육과정평가원은 이용자가 학생 평가 관련 모든 자료를 한 번에 이용할 수 있도록 기존에 운영하던 '수행평가 지원 포털', '성취평가제 지원 포털', '서술형 평가 지원 포털'을 하나로 통합한 '학생평가 지원 포털' (<https://stassess.kice.re.kr/>)을 운영하고 있음. (2018년 12월 기준)

적 수행평가 과제를 개발하고 실행할 때 직면할 수 있는 어려움과 이에 대한 극복 방안 등을 교사와 학생 입장에서 논의하였다. <부록 III-2>에 연구자-예비교사의 협력적 논의, 교사-학생의 이중적 입장에서의 논의 방법에 관한 안내 자료를 제시하였다. 한편, 예비교사의 개별 과제 개발 과정을 심층적으로 이해하기 위하여 추가로 약 30분간 발생 사고법에 대한 오리엔테이션을 실시하였다. 즉 발생 사고의 개념과 방법을 설명한 후 발생 사고를 통한 문제 해결 과정이 담긴 영상을 보여주었다. 예비교사들은 영상을 시청한 후 연구와 무관한 문제를 해결하며 발생 사고법을 연습하였다.

셋째 주에는 구성주의적 수행평가에 대한 내면화를 목표로, 수행평가 과제의 개발 실습 및 개선 방향에 대한 논의 활동을 하였다. 예비교사들에게 2009 개정 과학 교육과정에 따른 중학교 2학년 과학 교과서와 교사용 지도서 각각 2종을 제공한 후 각자 ‘물질의 특성’ 단원에서 수행평가 과제를 개발하도록 하였다. 예비교사들이 발생 사고에 익숙해질 수 있도록 수행평가 과제 개발 중 발생 사고법을 사용하도록 하였다. 연구 II에 따르면 학생에 대한 경험적 지식이 부족한 예비교사들은 학생들의 오개념이나 발달 수준을 고려하지 못하는 것으로 알려져 있으므로 물질의 특성에 대한 오개념 자료(부록 III-3)도 예비교사들에게 제공하였다. 오개념 자료에는 “용해와 화학 변화를 혼동한다.”와 같이 해당 단원의 대표적인 학생들의 오개념 유형 10여 가지를 제시하였고, 각 유형마다 구체적인 오개념의 예시와 학생이 그렇게 생각하는 이유 등을 간단히 소개하였다. 이후 각자 개발한 수행평가 과제를 실제 수업에 적용할 때 발생할 수 있는 어려움과 이에 대한 극복 방안 등을 교사와 학생 입장에서 논의하였다. 이때 예비교사들의 반성적 활동을 촉진하고 조별 논의가 구성주의적 관점에서 이루어질 수 있도록 첫째 주에 소개했던 수행평가 과제 개발 과정에서 고려할 요소를 점검표 형태로 정리하여 제공하였다(부록 III-4).

각 프로그램 활동과 관련 교사 전문성 개발 원리와 전략 및 관련 자료로써 활용 또는 참고한 선행연구를 <표 III-2>에 정리하였다. 또한, 예비 화학교사의 수행평가 전문성 향상을 위한 프로그램의 주차별 내용을 <그림 III-1>에 제시하였다.

<표 III-2> 교사 전문성 개발 원리와 관련 프로그램 활동 및 도입 배경

교사 전문성 개발 원리 (Loucks-Horsley <i>et al.</i> , 2009)	관련 프로그램 활동	도입 배경	관련 선행연구
효과적인 교수학습에 대한 정립된 이미지의 제공	구성주의적 수행평가의 이론적 배경과 특징 PCK와 구성주의적 수행평가의 관계	연구 I, II	민희정, 2012; 이은진 등, 2007; Beyer & Davis, 2009; Falk, 2012
학생의 학습 목표 및 요구의 충족	구성주의적 수행평가를 현장에 적용할 때의 어려움과 극복 방 안에 대한 교사와 학생의 이중 적 입장에서의 논의	연구 I	김선희, 2012
	학생 오개념 자료의 제공	연구 II	노태희, 김창민, 1999; 노태희 등 1995
PCK 및 기술을 개발하 고 비판적으로 검토 및 반영할 기회의 제공	다양한 수행평가 예시 과제에 대한 PCK 및 구성주의적 관점 에서의 분석 및 논의	연구 II	장수용, 2002; 홍수연, 2006; 한국교육과정평 가원 수행평가 포털
	구성주의적 수행평가의 과제 개 발 실습	연구 II	
	PCK 측면에서의 점검표 제공	연구 II	
다른 동료 및 전문가들 과 함께 하는 기회의 제공	다양한 평가 방법의 구성주의적 목적으로의 활용 방법에 관한 논의	연구 I	
	자신이 중등학생일 때 경험했던 수행평가 사례를 공유하고 논의	연구 I	노태희 등, 2015; 민희정, 2012
	자신이 개발한 수행평가 과제의 개선 방향에 대한 조별 논의	연구 II	
성인 학습자로서 예비 교사의 능동적 참여의 촉진	연구자와 예비교사 사이의 협력 적 논의	연구 I	김구연, 2010; 이은진 등, 2007

	목표	프로그램 활동	비고
1주	이론적 근거의 이해	<ul style="list-style-type: none"> - 구성주의적 평가의 이론적 배경과 특징^a - PCK와 구성주의적 수행평가의 관계^a - 다양한 평가 방법의 구성주의적 목적으로의 활용 방법에 관한 논의^d 	<p>PPT [부록 III-1]</p> <p>연구자와 예비교사 사이의 협력적 논의^e</p>
2주	실천 능력의 증진	<ul style="list-style-type: none"> - 다양한 수행평가 예시 과제에 대한 PCK 및 구성주의적 관점에서의 분석 및 논의^c - 자신이 중등학생일 때 경험했던 수행평가 사례를 공유하고 논의^d 	<p>교사와 학생의 이중적 입장에서 논의^b [부록 III-2]</p>
3주	실습 및 논의를 통한 내면화	<ul style="list-style-type: none"> - 구성주의적 수행평가 과제의 개발 실습^c - 개발한 수행평가 과제의 개선 방향에 대한 조별 논의^d 	<p>점검표 제공^c [부록 III-3]</p> <p>학생 오개념 자료 제공^b [부록 III-4]</p>

^a 효과적인 교수학습에 대한 정립된 이미지의 제공, ^b 학생의 학습 목표 및 요구의 충족, ^c PCK 및 기술을 계발하고 비판적으로 검토 및 반영할 기회의 제공, ^d 다른 동료 및 전문가들과 함께 하는 기회의 제공, ^e 성인 학습자로서 예비 교사의 능동적 참여의 촉진.

<그림 III-1> 수행평가 전문성 향상 프로그램의 주차별 활동

2.3 연구 절차

3주에 걸쳐 평가 전문성 향상 프로그램을 진행한 후, 넷째 주는 수행평가 전문성 향상 프로그램에 참여한 예비 화학교사의 수행평가 과제 개발 과정을 알아보기 위한 자료 수집 단계로 진행하였다. 예비교사들은 프로그램에 참여한 경험을 토대로 연구 II와 동일한 중학교 1학년 과학의 ‘분자 운동과 상태 변화’ 단원을 대상으로 수행평가 과제를 개발하였다. 수행평가 과제 개발 과정 또한 연구 II와 동일하게 구상 단계와 구체화 단계로 이루어진 개별 과제 개발 과정과 조별 논의 과정으로 나누어 실시하였다. 즉, 구상 단계에서는 활동지를 제공하여 수행평가 아이디어를 자유롭게 기록하도록 하였다. 구체화 단계에서는 아이디어 중 한 가지를 선택한 후 별도의 활동지에 수행평가의 전체적인 흐름과 교사와 학생이 수행하는 활동을 구체적으로 묘사하고 학생들에게 제공할 활동지와 평가 방법 및 준거 등 필요한 모든 자료를 개발하도록 하였다. 또한 수행평가는 실험 수업과 같이 교수학습의 일부로 이루어지는 경우가 많으므로 예비교사가 수업의 일환으로 수행평가 과제를 개발하고자 하는 경우, 수업 지도안 등 해당 수업에 대한 자료도 포함하여 수행평가 과제를 개발하도록 하였다. 예비교사들에게 선행연구(노태희, 김창민, 1999; 노태희, 임희준, 우규환, 1995)를 바탕으로 제작한 분자 운동과 상태 변화에 대한 오개념 자료(부록 III-3)와 수행평가 과제 개발을 위한 점검표(부록 III-4)를 제공하였으며 발성 사고법을 사용하도록 하였다. 즉, 4인의 연구자가 각각 예비교사 두 명을 관찰하며 관찰 노트를 작성하고, 예비교사가 발성 없이 글만 쓰거나 생각만 할 때는 발성 사고를 하도록 격려했다. 조별 논의 과정에서는 개별로 개발한 수행평가 과제의 개선 방향을 논의하였으며, 이때 연구자는 학생들의 조별 논의를 관찰하며 관찰 노트를 작성하였다. 조별 논의 과정에서 예비교사들은 수행평가 계획안 위에 수정할 내용을 기록하였고 이를 바탕으로 수행평가를 수정 및 보완하여 과제 개발을 완료하였다. 이후 예비교사를 관찰했던 연구자가 예비교사와 함께 활동지를 보면서 연구 II와 동일한 면담 시나리오를 활용하여 반구조화된 면담을 실시하였다. 즉, 면담자는 관찰 노트를 참고하면서 각 단계별 활동에서 예비교사의 의도를 구체적으로 조사하였으며, 필요시 추가 질문을 하였다. 또한 평가 전문성 향상 프로그램의 영향 및 이에 관한 예비교사의 인식을 조사하기 위하여 추가 질문을 하였다. 예를

들어, 구상, 구체화 단계로 구성된 개별 과제 개발 과정 및 조별 논의 과정에 프로그램 활동이 미친 영향을 구체적으로 조사하였다. 이후 예비교사의 과학 평가에 관한 신념 등 수행평가 과제 개발 과정에서 잘 드러나지 않는 PCK 구성요소를 구체적으로 파악하기 위해 자신이 생각하는 구성주의적 평가의 가장 큰 특징이나 요소, PCK 구성요소 중 중점적으로 반영하고자 하였던 요소나 구현하기 어려웠던 요소 등을 조사하였다(표 III-3). 각 면담에는 약 50여분이 소요되었으며, 모든 발성 사고, 조별 논의 및 면담 과정은 녹음 및 녹화한 후 전사본을 작성하였다.

<표 III-3> 프로그램 관련 반구조화된 면담 개요

면담질문

(구상 단계 질문 후)

- 지난 3주간의 워크숍 활동이 아이디어 떠올리기(구상 단계) 활동에 도움이 된 부분이 있나요?

Y: 구체적으로 어떤 활동으로부터 도움을 받았으며, 이유는 무엇이라고 생각하나요?

(구체화 단계 질문 후)

- 지난 3주간의 워크숍 활동이 수행평가 과제 개발(구체화 단계) 활동에 도움이 된 부분이 있나요?

Y: 구체적으로 어떤 활동으로부터 도움을 받았으며, 이유는 무엇이라고 생각하나요?

(조별 논의 과정 질문 후)

- 지난 3주간의 워크숍 활동이 조별 논의 활동에 도움이 된 부분이 있나요?

Y: 구체적으로 어떤 활동으로부터 도움을 받았으며, 이유는 무엇이라고 생각하나요?

-
- 구성주의를 강조한 수행평가를 개발해 본 경험에 대한 전체적인 소감을 말해주세요.
 - 좋았던 점, 힘들거나 어려웠던 점이 있다면 구체적으로 질문.
 - 이번 프로그램에서는 좋은 수행평가를 개발하기 위해 고려해야 할 요소로 크게 다섯 가지를 강조하였습니다. (평가, 교육과정, 교수 전략, 과학 내용, 학생) 각 요소를 어떤 식으로 반영하고자 하였나요?
 - 이 항목 중에서 수행평가에 반영(구현)하기 어려웠거나, 이해가 잘 가지 않았던 내용이 있었다면?
 - 이를 위해 어떤 도움이 필요하다고 생각하나요?
 - ‘구성주의적 평가’의 가장 큰 특징이나 요소는 무엇이라고 생각하나요?
 - 향후 교직에 진출하였을 때, 구성주의적 평가를 적용하거나 시도해볼 수 있겠다고 생각하시나요?
-

N: 어떤 어려운 점이 예상되나요? 구체적으로 설명해주세요.

Y: 이번 활동이 구성주의적 평가관을 갖는데 도움이 되었다고 생각하시나요? 구체적으로 설명해주세요.

2.3 분석 방법

중등 예비교사의 수행평가 과제 개발 실태를 PCK 측면에서 분석한 연구 II의 분석틀을 수정 없이 사용하여 중등 예비교사의 구성주의적 수행평가 과제 개발 과정에서 나타나는 PCK 구성요소와 요소 사이의 연계를 분석하였다. 이 분석틀은 예비교사의 PCK를 과학 평가에 관한 지식, 과학교육과정에 관한 지식, 과학 교수 전략에 관한 지식, 과학 내용에 관한 지식, 학생에 관한 지식, 과학 평가에 관한 신념으로 구분하고 각 PCK 구성요소를 2-3가지의 하위 요소로 세분하였다. 수집한 자료 중 면담 전사본과 조별 논의 전사본을 주 분석 자료로 하였으며, 발성 사고 전사본과 예비교사들의 활동지 및 연구자의 관찰 노트 등은 보조 자료로 활용하였다. 분석의 기본 단위도 연구 II와 동일한 기준으로, 면담의 경우 면담자의 질문과 예비교사의 응답으로 하였고, 조별 논의는 한 가지 주제에 대한 예비교사들 사이의 논의로 하였다. 한 가지 주제에 대한 후속 질문과 응답이 이어지는 경우에는 같은 분석 단위로 간주하였고, 질의응답 과정에서 주제가 전환되는 경우에는 분석 단위를 구분하였다. 특정 PCK 구성요소가 다른 PCK 구성요소의 이유나 근거로서 활용되었거나 동일한 활동에서 여러 PCK 구성요소들을 동시에 고려하는 경우와 같이 PCK 구성요소 사이의 연관성이 드러날 때는 연계로 해석하였다(노태희, 박재성, 강훈식, 2016). 예를 들어, 예비교사가 학생의 오개념을 드러내기 위하여 특정한 실험 활동을 구성한 경우에는 학생에 관한 지식 중 오개념과 과학 교수 전략에 관한 지식 중 주제 전략이 연계된 것으로 해석하였다.

결과 해석의 신뢰도와 타당도를 높이기 위하여 2인의 분석자가 분석 기준에 따라 예비교사 2명의 면담 전사본과 한 조의 조별 논의 전사본을 각자 분석하고 비교하는 과정을 반복하였다. 분석자 간 일치도가 .90에 도달한 후 1인의 연구자가 모든 예비교사의 수행평가 과제 개발 과정에서 나타나는 PCK 구성요소와 요소 사이의 연계 빈도 및 특징을 분석하였다. 또한 분석 결과가 발성 사고 전사본과 예비

교사가 작성한 계획안에 나타났는지 여부를 확인하는 삼각측정(triangulation) 과정을 거쳤다. 연구자 간 내부 논의와 과학교육 전문가 3인, 중등 화학교사와 대학원생이 참여한 수차례의 세미나를 통하여 연구 결과를 점검하고 지속적으로 보완하였다.

3. 연구 결과 및 논의

3.1 개별 과제 개발 과정의 PCK 구성요소

개별 과제 개발 과정에서 나타난 PCK 구성요소의 분석 결과는 <표 III-4>와 같다. 전체 206개의 PCK 구성요소 중 44.7%(92회)는 구상 단계에서, 55.3%(114회)는 구체화 단계에서 나타났다. 구상 단계의 PCK 구성요소 중에서는 과학 평가(10.2%), 과학 교수 전략(10.2%), 학생(8.3%), 과학 내용에 관한 지식(7.8%)의 비율이 상대적으로 높았다. 구체화 단계에서는 과학 평가에 관한 지식(22.3%)의 비율이 가장 높았고, 그다음으로는 과학 평가에 관한 신념(12.1%), 과학 교수 전략(8.7%), 학생에 관한 지식(6.8%)의 순서로 나타났다. 과학교육과정에 관한 지식의 비율은 두 단계에서 모두 비교적 낮았다.

<표 III-4> 개별 과제 개발 과정의 PCK 구성요소

PCK 구성요소		구상 단계 (%)		구체화 단계 (%)		계 (%)	
A	A1	21 (10.2)	13 (6.3)	46 (22.3)	25 (12.1)	67 (32.5)	38 (18.4)
	A2		8 (3.9)		21 (10.2)		29 (14.1)
C	C1	4 (1.9)	4 (1.9)	5 (2.4)	3 (1.5)	9 (4.4)	7 (3.4)
	C2		-		2 (1.0)		2 (1.0)
I	I1	21 (10.2)	19 (9.2)	18 (8.7)	17 (8.3)	39 (18.9)	36 (17.5)
	I2		2 (1.0)		1 (0.5)		3 (1.5)
K	K1	16 (7.8)	9 (4.4)	6 (2.9)	4 (1.9)	22 (10.7)	13 (6.3)
	K2		7 (3.4)		2 (1.0)		9 (4.4)
L	L1	17 (8.3)	7 (3.4)	14 (6.8)	4 (1.9)	31 (15.0)	11 (5.3)
	L2		5 (5.4)		8 (3.9)		13 (6.3)
	L3		5 (5.4)		2 (1.0)		7 (3.4)
O	O1	13 (6.3)	3 (1.5)	25 (12.1)	12 (5.8)	38 (18.4)	15 (7.3)
	O2		10 (4.9)		13 (6.3)		23 (11.2)
합계		92 (44.7)		114 (55.3)		206 (100.0)	

* A: 과학 평가에 관한 지식(A1: 평가 방법, A2: 평가 영역), C: 과학교육과정에 관한 지식(C1: 연계, C2: 평가 목표), I: 과학 교수 전략에 관한 지식(I1: 주제 전략, I2: 교과 전략), K: 과학 내용에 관한 지식(K1: 개념 및 이론, K2: 과학 탐구 과정), L: 학생에 관한 지식(L1: 오개념, L2: 발달 수준, L3: 동기 및 흥미), O: 과학 평가에 관한 신념(O1: 전통적 평가 지향, O2: 구성주의적 평가 지향).

PCK 구성요소별로 살펴보면, 과학 평가에 관한 지식(A)은 구상 단계(10.2%)보다 구체화 단계(22.3%)에서 더 많이 나타났다. 평가 방법(A1)의 경우, 구상 단계에서는 평가 방법 선정(13회, 6.3%)만 나타났지만, 구체화 단계에서는 평가 방법 선정(5회, 2.4%), 평가 방법 검토(8회, 3.9%), 채점 준거 및 배점 설정(12회, 5.8%)이 모두 나타났다기 때문에 비율이 높아졌다. 평가 영역(A2)의 경우, 구상 단계에서는 평가 영역을 구체적으로 정하지 않았던 예비교사들도 구체화 단계에서는 새로운 평가 영역을 설정하고 구상 단계에서 정한 평가 영역을 세분화하였기 때문에 비율이 증가하였다. 이때 예비교사들은 프로그램의 수행평가 사례 분석에서 다루었던 평가 영역들을 참고하며 개념 이해도와 참여도뿐만 아니라 의사소통능력, 관찰력, 창의력 등

다양한 인지적, 정의적 측면의 평가 영역을 고려하였다. 한편 예비교사의 수행평가 과제 개발 실태를 조사한 연구 II에서는 예비교사들이 과학 평가에 관한 지식 중 채점 준거 및 배점 설정에 집중하였고, 그 결과 과학 평가에 관한 지식의 비율이 절반에 가까웠다. 이는 예비교사들이 점검표를 활용하여 특정한 과학 평가에 관한 지식 요소에 지나치게 집중하지 않고 다른 요소들을 골고루 고려하였음을 의미하며, 다음은 이러한 주장을 뒷받침하는 면담의 일부이다.

사실 전에는 이렇게 구체적으로 수행평가를 계획한 적이 없었으니까 그냥 막연하게 교과서 보고 비슷하게 문제 만들고 채점표 만들고 그냥 간단 간단하게 했었는데, 이 구성주의적 수행평가라는 걸 만드는 활동을 하면서 더 체계적으로 만들 수 있는 것 같아서 좋았어요. 이 점검표도요. (중략) 오개념을 생각해야 된다는 것도 사실 이렇게 제대로 알고 있지는 않았는데 지금 이렇게 점검표도 보면서 이런 활동을 하니까 오개념도 알아봐야 되고, 또 계속 빠뜨렸던 연계도 생각해 봐야하는 거도 이렇게 알 수 있었고.

(예비교사 B의 추가 면담 중)

과학교육과정에 관한 지식(C)의 전체 비율은 PCK 구성요소 중에서 가장 낮았다. 연계(C1)는 수행평가를 도입하거나 정리 상황을 구상할 때, 그리고 수행평가에서 다루고자 하는 개념의 교육과정 포함 여부를 확인할 때 1-2회씩 나타났다. 그러나 다른 교과와 연계하는 수평 연계를 고려하거나 연계되는 내용을 적극적으로 수행평가에 반영하는 등의 활용은 여전히 나타나지 않았다. 이는 예비교사들이 수행평가 과제 개발 과정에서 과학교육과정에 관한 지식을 활용할 필요성이나 효과적인 활용 방법에 대한 이해가 부족했기 때문일 수 있다. 과학 교과의 다른 단위이나 다른 교과와의 연계는 학생의 선지식에 대한 고려와 함께 지식의 유기적인 연결을 촉진하므로(강신포, 김호선, 2006), 수행평가를 계획할 때 이러한 연계를 고려하도록 유도할 수 있는 방안이 필요하다. 예를 들어 예비교사들에게 교과서와 교사용 지도서를 제공할 때, 학교 현장에서와 같이 수행평가를 실시하고자 하는 단위의 전체적인 교수학습 계획을 체계적으로 검토할 수 있도록 명시적인 지침을 함께 제공한다면 연계의 활용을 보다 촉진할 수 있을 것이다.

과학 교수 전략에 관한 지식(I)의 대부분은 특정 단위에서 활용이 가능한 주제

전략(I1)이었다. 구상 단계에서는 활동지 작성과 발표 등의 구체적인 수업 기법이 13회(6.3%), 동기 유발 전략 등의 수업 전략이 5회(2.4%) 나타났고, 구체화 단계에서는 수업 기법 13회(6.3%), 수업 전략 3회(1.5%), 시간 배분 1회(0.5%)가 나타났다. 즉, 예비교사들은 프로그램에서 POE나 협동학습 등과 같이 교과 내용과 무관한 교과 전략(I2)을 활용한 다수의 수행평가 예시를 보며 관련 내용을 논의하였음에도, 구체적인 수업 기법과 전략을 검토하고 정교화하는 데 집중하였으며 교과 전략은 거의 활용하지 않았다. 면담 결과, 예비교사들은 수업 모형과 같은 교과 전략을 수행평가에 도입할 수 있고 이는 구성주의적 수행평가를 구현하는 데 유용하다는 것을 인식하였지만 다양한 수업 모형에 대한 지식이 부족하거나 특정 교과 전략을 활용할 경우 형식이 제한되어 다양한 아이디어를 생각하는 데 어려움을 겪을 수 있다고 생각하는 것으로 나타났다.

면담자 : 점검표에서 제일 반영하기 어려운 요소는 뭔 거 같아요?

예비교사 : 어려운거요? 이거요. 교수 전략.

면담자 : 이유를 말한다면?

예비교사 : 이것 때문에 생각의 한계가 생기는 거 같아요. 왜냐면 만약 제가 예시 이런 거 하나도 안보고, 무슨 기법이 있는지 하나도 모르고, 어떤 제한이 있는지도 모르고 그러면, 물론 너무 기준이 없이 제가 중구난방으로 뭔가를 작성할 수도 있는데, 어떻게 생각하면 진짜 새로운 걸 생각했을 수도 있잖아요. 그런데 (교과) 전략을 쓰면 그냥 이 안에서 생각하게 되요. 결국에는 이 안에서 생각하게 될 수밖에 없는데, 도움은 되는데 한계가 생기는 거 같아요. 말하다 보니까 장단점인 거 같네요.

(예비교사 G의 구체화 단계 면담 중)

예비교사들이 다양한 주제 전략을 고려한 것은 긍정적으로 볼 수 있지만 교과 전략을 거의 활용하지 않은 점에 대해서는 개선이 필요하다. 이를 위해서는 예비교사들이 수행평가 과제 개발 과정에서 교과 전략을 활용할 때의 장점에 대하여 보다 긍정적으로 인식함은 물론 교과 전략의 효과적인 활용 방법을 체득할 수 있는 기회를 제공할 필요가 있다.

과학 내용에 관한 지식(K) 측면의 경우에도 비교적 적게 나타났는데, 그중에서

도 구상 단계(7.8%)보다 구체화 단계(2.9%)에서 약간 적게 나타났다. 특히 예비교사들은 실험이 학생들의 개념 이해를 돕고 흥미를 촉진할 수 있으며, 타 교과와 차별화되는 과학 교과만의 특징적인 활동임과 동시에 중요한 평가 방법이라고 생각하고 있었다. 이에 대부분 실험을 활용한 수행평가를 계획하였음에도 불구하고 과학 개념 및 이론(K1) 및 과학 탐구 과정 지식(K2)은 각각 6.3%, 4.4%에 불과하였다. 과학 탐구 과정에 관한 평가는 구성주의적 평가의 대표적인 특징 중 하나이다(박선미, 1998). 그러나 대부분의 예비교사들은 학생의 참여도를 평가하는 관찰 평가만을 도입하거나 소요 시간과 교사의 노력 등 현실적인 측면을 고려하여 이를 구현하는데 여전히 어려움을 겪었다. 반면 예비교사 D는 면담 평가를 통해 학생들의 과학 탐구 과정을 평가하고자 하였는데, 실험 재료만을 제공하고 과학 탐구 과정과 조건을 학생 스스로 설계하도록 하는 수행평가를 계획한 특징이 있었다(그림 III-2). 결과 확인 위주의 구조화된 실험에서는 과학 탐구 과정에 대한 학습과 평가가 어렵다는 점을 고려할 때(이근준, 정진우, 2004), 실험을 보다 개방된 형태로 계획하도록 하는 것은 예비교사가 과학 탐구 과정 지식을 활용하도록 하는 방안이 될 수 있을 것으로 보인다.

면담자 : 이번 프로그램에서 다룬 좋은 수행평가 계획을 위한 요소들을 어떤 식으로 반영을 하려고 했는지 얘기해주세요.

예비교사 : 단순하게 학생들이 답만 적은걸 평가하는 거 보다는 (교사가) 직접 그 학생이 생각하는 과정이나 직접 실험하는 과정을 보면서, 즉석에서 평가하는 데 구성주의적인 요소가 들어있다고 생각해요. (중략) 그래서 면담 평가를 한 번 해보자 라고 생각을 해봤어요.

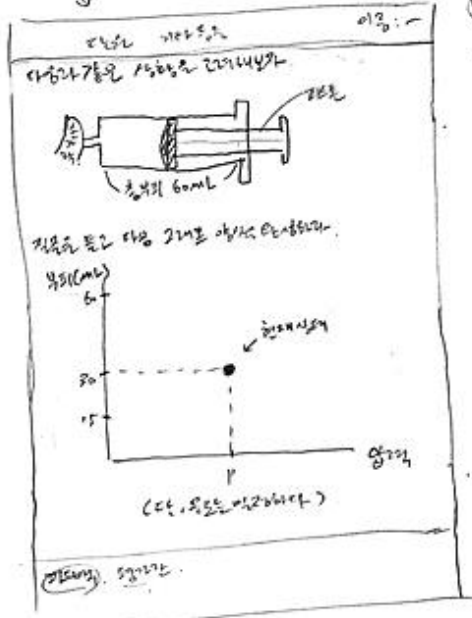
면담자 : 면담 평가를 하면 뭐가 어떤 점이 좋을까요? 그런 점을 생각해 봤어요?

예비교사 : 학생이 눈앞에서 주어진 시간 동안에 자기가 생각하는 것을 말하니까 그 과정을 직접 볼 수 있고. 그러니까 이 학생이 어느 정도의 그 수준을 갖고 있구나, 이것을 쉽게 알 수 있는 것 같아요.

(예비교사 D의 추가 면담 중)

[세부 계획안]

면담을 통하여 학생들에게 직접 질문해서 사고 과정과 결과물 정립할수 있는. 면담 평가 시에 필요한 항목들.



면담 질문

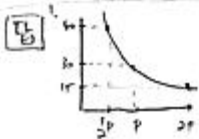
- 총 부피가 60ml 짜리 주사기가 있다고 생각해보자
- 60ml 주사기의 끝은 손가락으로 막아줍니다.
- 처음 주사기의 상단 피스톤이 주사기 부피의 절반에 해당하는 지점에 놓여 있고 이때 주사기 내부의 압력을 P 라고 해봅시다.
- 1) 주사기를 끝까지 당겼을 때다
- 2) 주사기 내부 부피가 15ml 가 될 때 까지 끝까지 당겼을 때의 압력을 생각해봅시다. 그래프에 표시하고 그래프의 대략적인 모양도 그려주세요.
- 3) 그리고 그래프를 그릴 때 단위를 적어주세요.

평가시간 : 25분.

[평가 준거]

면담 평가 시 결과가 그래프와 함께 학생의 답변에 따라 측량에서 정한다.
 학생의 항목들은 오차가 나도록 되도록 해준다.

1. 그래프를 올바르게 완성하였는가? (5점)
 - 그리지 못하였음 ... 0
 - 점만 찍었음 } 하지만 올바르게 찍었음 ... 1
 - 틀리거나 올바르게 찍었음 ... 3
 - 그래프 개형만 고쳤음 ... 3
 - " " 올바르게 고쳤음 ... 5
2. 그래프 완성의 근거를 타당하게 설명하였는가? (10점)
 - 설명하지 못하였음 ... 0
 - 설명하게는 그래프 보거나 정서 못해 그리 못함 ... 2
 - 비정수점에서 압력까지 보거나 부피가 같아도 같은 것은 설명함 ... 5
 - 보일의 법칙이라는 용어거나 잘 이해하고 있지 못함 ... 5
 - 압력과 부피의 반비례 관계라는 설명함 ... 8
 - 압력과 부피의 관계가 보일 법칙이라는 반비례 관계라는 설명함 ... 10



2. 압력과 부피는 보일의 법칙에 따라 서로 반비례 관계에 있으므로 부피가 2배가 되면 압력은 1/2배가 되고, 부피가 3배가 되면 압력은 1/3배가 되기 때문입니다.

<그림 III-2> 예비교사 D가 개발한 수행평가 과제

학생에 관한 지식(L) 측면의 경우 구상 단계에서는 오개념(L1), 발달 수준(L2), 동기 및 흥미(L3)에 대한 고려가 비교적 고르게 나타난 반면, 구체화 단계에서는 주로 발달 수준에 대한 고려가 나타났다. 이러한 결과는 예비교사들이 구상 단계에서는 학생의 동기 및 흥미를 고려하여 소재나 전략을 구상하고 오개념 자료를 활용하여 학생들의 오개념과 발달 수준을 고려하였지만, 구체화 단계에서는 학생의 수준을 고려하여 문항이나 활동의 수준을 의도적으로 조절하는 과정이 주를 이루었기 때문으로 생각할 수 있다. 예를 들어, 숨에 에탄올을 묻히고 증발시키는 실험을 주제로 수행평가 과제를 개발한 예비교사 C는 실험 과정에 POE 모형을 도입한 뒤, 자신이 심화한 실험 활동과 활동지의 지문이 중학교 1학년 수준의 사고력으로 이해할 수 있는 수준인지 검토하였고, ‘어느 것의 확산이 빠를까?’를 주제로 확산 실험에 관한 수행평가 과제를 개발한 예비교사 H는 실험 과정에 대한 학생들의 이해를 돕기 위하여 활동지에 구체적인 그림을 그려주었다(그림 III-3).

좀 더 심화하면서도 너무 난이도 높지 않게, 딱 초등학교 때 배우고 올라온 중학교 1학년한테 이 정도의 난이도면 되겠다는 생각을 했어요. (중략) 중학교 1학년한테 교육과정이 원하는 사고력, 그 정도를 평가하기는 무난하지 않을까.

(예비교사 C의 구체화 단계 면담 중)

면담자 : 활동지를 작성하면서 본인이 특별히 고려한 점이나 특별히 좀 신경을 쓴 부분이 있나요?

예비교사 : 그림을 그렸던 거요. 애들이 실험을 할 때 말만 들어서 ‘이걸 이렇게 해 보세요.’ 하면 어떻게 감이 안 잡히잖아요. 근데 그림 같은 거를 하나 주면 그 말 전체를 한 번에 이해할 수 있는 발판이 되니까 그래서 그림을 그렸어요.

(예비교사 H의 구체화 단계 면담 중)

또한 학생에 관한 지식은 예비교사의 수행평가 과제 개발 실태를 조사한 연구 II와 비교하였을 때 비율이 가장 크게 증가한 PCK 구성요소였다. 이는 면담 사례에서 볼 수 있듯이 프로그램 참여 후 대부분의 예비교사들이 자신이 생각하는 구성주의적 평가의 가장 큰 특징으로 학생 중심의 평가를 생각했기 때문으로 해석할 수 있다.

면담자 : 구성주의적 평가의 가장 큰 특징이 뭐라고 생각하세요?

예비교사 : 학생 위주로 만들어진 평가. 선생님이 편하지 않고 학생들이 이해하기 쉽게, 더 잘 알 수 있게 해주는 평가 같아요. 문제를 만들고 그냥 시험 보고 끝이 아니라 애들이 기억에 남을만한 문제를 만들거나 (중략) 재미있게 풀 수 있도록 하려면 어떻게 해야 할지 생각하거나 잘 기억에 남도록 풀이해주고 이런 것도 중요한 거 같아요.

면담자 : 이제까지의 (프로그램) 활동이 이런 본인의 생각이나 구성주의적 평가에 선 이런 걸 고려해야겠다고 인식을 갖는데 도움이 된 게 있나요?

예비교사 : (3주차에서) 이걸 직접 한번 만들어 봄으로써 생각을 많이 하게 되는 거 같아요. 그 오개념 같은 부분도 그렇고 이게 학생 입장에서 좋은 문제인지 아닌지 생각해 보게 되요. (중략) 논의해본 것도 좋다고 생각해요. 오개념 같은 것도 그렇고 다른 단원과의 연계도 그렇고 그걸 안 했으면 원래는 생각을 안 했을 거 같은데.

(예비교사 F의 추가 면담 중)

3.2 개별 과제 개발 과정의 PCK 구성요소 사이의 연계

개별 과제 개발 과정에서 나타난 PCK 구성요소 사이의 연계 분석 결과는 <표 III-5>, <표 III-6>과 같다. 총 46회의 연계가 나타났으며 이중 47.8%(22회)는 구상 단계에서, 52.2%(24회)는 구체화 단계에서 나타났다. 두 가지 PCK 구성요소의 연계는 73.9%(34회)였고 세 가지 PCK 구성요소 사이의 연계는 26.1%(12회)였다. 각 PCK 구성요소를 중심으로 보면, 과학 교수 전략에 관한 지식을 중심으로 한 연계가 25.7%로 가장 많았으며, 과학 평가에 관한 신념(22.9%), 학생에 관한 지식(21.4%), 과학 평가에 관한 지식(17.9%) 등을 중심으로 한 연계는 유사한 비율로 나타났다. 그러

나 과학교육과정에 관한 지식이나 과학 내용에 관한 지식을 중심으로 한 연계는 적었다.

<표 III-5> 개별 과제 개발 과정의 PCK 구성요소 사이의 연계 (통합 유형별)

통합 유형	구상 단계 (%)	구체화 단계 (%)	계 (%)
A - I	2 (4.3)	1 (2.2)	3 (6.5)
A - L	-	2 (4.3)	2 (4.3)
A - O	1 (2.2)	7 (15.2)	8 (17.4)
C - L	-	1 (2.2)	1 (2.2)
C - O	1 (2.2)	-	1 (2.2)
I - K	3 (6.5)	1 (2.2)	4 (8.7)
I - L	5 (10.9)	3 (6.5)	8 (17.4)
I - O	2 (4.3)	1 (2.2)	3 (6.5)
K - O	1 (2.2)	-	1 (2.2)
L - O	-	3 (6.5)	3 (6.5)
A - C - L	-	1 (2.2)	1 (2.2)
A - I - L	-	1 (2.2)	1 (2.2)
A - I - O	1 (2.2)	1 (2.2)	2 (4.3)
A - K - L	1 (2.2)	-	1 (2.2)
A - L - O	1 (2.2)	-	1 (2.2)
C - I - L	1 (2.2)	-	1 (2.2)
C - I - O	-	1 (2.2)	1 (2.2)
I - K - O	1 (2.2)	-	1 (2.2)
I - L - O	2 (4.3)	1 (2.2)	3 (6.5)
계	22 (47.8)	24 (52.2)	46 (100.0)

* A: 과학 평가에 관한 지식, C: 과학교육과정에 관한 지식, I: 과학 교수 전략에 관한 지식, K: 과학 내용에 관한 지식, L: 학생에 관한 지식, O: 과학 평가에 관한 신념.

<표 III-6> 개별 과제 개발 과정의 PCK 구성요소 사이의 연계 (PCK 구성요소별)

PCK 구성요소	구상 단계 (%)	구체화 단계 (%)	계 (%)
A와 다른 PCK 구성요소의 연계	9 (6.4)	13 (9.3)	25 (17.9)
C와 다른 PCK 구성요소의 연계	3 (2.1)	5 (3.6)	8 (5.7)
I와 다른 PCK 구성요소의 연계	22 (15.7)	14 (10.0)	36 (25.7)
K와 다른 PCK 구성요소의 연계	8 (5.7)	1 (0.7)	9 (6.4)
L과 다른 PCK 구성요소의 연계	15 (10.7)	15 (10.7)	30 (21.4)
O와 다른 PCK 구성요소의 연계	15 (10.7)	17 (12.1)	32 (22.9)
계 (%)	72 (51.4)	68 (48.6)	140 (100.0)

* A: 과학 평가에 관한 지식, C: 과학교육과정에 관한 지식, I: 과학 교수 전략에 관한 지식, K: 과학 내용에 관한 지식, L: 학생에 관한 지식, O: 과학 평가에 관한 신념.

단계별로 살펴보면, 구상 단계에서는 과학 교수 전략에 관한 지식을 중심으로 한 연계가 가장 많이 나타났다(22회, 15.7%). 과학 교수 전략에 관한 지식은 과학 평가, 과학교육과정, 과학 내용에 관한 지식 등 다양한 PCK 구성요소와 연계되었는데, 이중 학생에 관한 지식과의 연계가 가장 활발히 이루어졌다. 즉 예비교사들은 수행평가의 주제를 결정하거나 실험을 설계할 때 학생의 오개념, 발달 수준, 동기 및 흥미 등을 고르게 고려하였으며, 다음이 그 사례들이다. 이 사례에서 예비교사 H는 증발과 확산을 수행평가 주제로 선정한 후 학생들의 미시적 입자 개념에 대한 이해 수준과 흥미 등을 고려하여 교수 전략과 실험을 선정하고 있으며, 예비교사 G는 ‘암모니아의 확산’을 주제로 학생들이 흔히 갖는 오개념을 고려하여 실험을 설계하고 있다.

면담자 : 왜 증발하고 확산 부분이 좋을 거 같다는 생각을 했어요?

예비교사 : 이 부분을 학생들이 헛갈려 하는 경우가 많더라고요. 뭐가 증발이고 뭐가 확산인지를 좀 많이 헛갈려 하는 게 있어서. (중략) 이게 (안 보이니

까) 말로 힘들더라고요. 애들이 이해도 못하고. 그래가지고 그래서 좀 어려울 때는 제가 시청각 동영상을 준비해가지고 이해를 시키려고 노력했던 거 같아요.

면 담 자 : 그럼 본인이 확산을 선택한 이유가 있어요? 이 아이디어를 구체화 시키는 걸 선택한 이유?

예비교사 : 이거는 리트머스 종지로 '아 이게 가고 있구나.' 확인할 수 있잖아요. 관찰이라는 자체가 사람이 확인을 할 때 뭔가 결과물이 보이면 흥미도가 올라가잖아요. 솔직히 둘 다 흥미도는 있지만 마지막 부분에 흥미도가 몰려있는 그런 결과물은 이게 더 많이 있을 것 같아서 그래서 이걸 더 쓴 것 같아요.

(예비교사 H의 구상 단계 면담 중)

이게 아까 오개념으로 나왔던 것 같은데. 그러면 온도에 따른 게 맞아. 이렇게 생각하기 쉬운 것 같아. 그러면 처음에 학생들한테 실험을 하게 하고. 이거 오개념을 잡으려면 어떻게 하는 게 나올까. 기체, 온도, 기체가 달라지는 거니까. 압력은 똑같은 온도 변화하고 변하는 거는.... 아니야. 이거는 압력 아니고 온도가 높아지면 부피가 달라지는 건데. 이 실험이 제일 나올 것 같은데.... 아, 그러면 확산이랑 섞으면 되겠다.

(예비교사 G의 발성 사고 중)

세 가지 PCK 구성요소가 함께 연계된 경우도 나타났다. 예를 들어, ‘보일/샤를 법칙 발표하기’를 주제로 한 예비교사 A는 프로그램 2주차 활동 때 발표했었던 중등학생 시절의 수행평가 경험을 바탕으로 학생의 수행 수준 및 전후 차시 수업 내용을 고려하여 과학 교수 전략을 수립하였는데, 이때 과학교육과정에 관한 지식과 학생에 관한 지식이 모두 연계되었다. 예비교사 A가 개발한 수행평가 과제를 <그림 III-4>에 제시하였다.

저번에도 얘기했었는데 제가 고등학생 때 발표 수행평가를 한 적이 있거든요. 처음에는 되게 '아 왜 나한테 수업을 시키지? 너무 싫다.'이랬는데 나중에는 기억이 되게 오래 남더라고요. (중략) 그래서 첫 번째 시간에 뒷단원의 예고편 같이 수업을 조금 한 다음에 하는 거예요. 학생들한테 이 단원 수업을 안 나가고 '너희들이 공부 해

와서 수업 형식으로 발표해.’ 이러면 너무 어려울 것 같아서. 그 다음 4단원이 보일
이고 5단원이 샤를인데 조를 딱 반으로 나눠서 한 시간은 보일하고 한 시간은 샤를
을 해서 발표를 한 다음에 마지막 한 시간에는 다시 선생님이 수업하면서 학생들
개념을 정리해주는 수업이구요.

(예비교사 A의 구상 단계 면담 중)

[세부 계획안]

- i) '3단원 기체의 압력' 진도를 나갈 때 기체의 압력 & 온도에 관한 내용 예고편처럼 학습
(1차시)
- ii) 4. 기체의 압력과 부피의 관계 (보일 법칙) (2차시)
 - 1조
 - 2조
 - 3조

※ 수업형식의 발표
ppt 활용 가능
모든 조원이 발표
발표 시간 1조당 20분 내다
- 5. 기체의 온도와 부피의 관계 (샤를 법칙) (3차시)
 - 4조
 - 5조
 - 6조
- iii) 발표가 끝난 뒤 1시간 동안 4.5 단원 내용 정리 (4차시)

[평가 준거]

- | | |
|--|--|
| <p>1) 관찰평가</p> <p>상 발표 준비도가 높고 태도가 좋고 발표내용 청신</p> <p>중 위 내용중 2개 해당</p> <p>하 " 1~0개 해당.</p> | <p>3) 동료평가</p> <p>① 조원 내 평가 (자기 평가와 동일)</p> <p>② 조원 외 평가 (9명씩)</p> <p>상 발표 내용이 잘 이해되고 창의적, 흥미력 있음</p> <p>중 위 내용중 2개 해당</p> <p>하 " 1~0개 해당.</p> |
| <p>2) 자기평가</p> <p>상 발표참여도 높고 조원들과 의사소통 잘함</p> <p>중 발표 참여 열심히 했지만 의사소통 X</p> <p>하 방담에 참여하지 않음</p> | |

[평가 결과 활용 방법]

발표 내용을 보고 여러가지 평가후 학생들의 오개념과 어려워하는 부분 확인후 피드백
평가목적은 완벽한 발표가 아니라 보일/샤를의 법칙의 공학력에 초점을 둔 것

<그림 III-4> 예비교사 A가 개발한 수행평가 과제

한편, 학생에 관한 지식을 중심으로 한 연계는 두 번째로 비율이 높았지만 대부분 과학 교수 전략에 관한 지식과 연계되었다. 즉 예비교사들이 과학 교수 전략을 구상할 때에는 다양한 PCK 구성요소를 고려하였으나, 학생에 관한 지식은 주로 교수 전략을 구상할 때만 활용되었다고 볼 수 있다. 따라서 구상 단계에서 PCK 구성요소 사이 연계의 중심은 연계의 빈도도 가장 높고, 다른 PCK 구성요소와 가장 많이 연계되어 있는 과학 교수 전략에 관한 지식이라고 볼 수 있다.

구체화 단계에서는 과학 평가에 관한 신념을 중심으로 한 연계의 비율이 가장 높았는데(17회, 12.1%), 이는 예비교사들이 뚜렷한 평가 지향이나 목적을 기반으로 PCK 구성요소를 활용하여 수행평가 과제 개발을 구체화하는 경우가 많았음을 의미한다. 과학 평가에 관한 신념은 주로 과학 평가에 관한 지식과 연계되었다. 그런데 구상 단계에서 과학 평가에 관한 신념을 중심으로 한 연계는 대부분 구성주의적 평가 지향과의 연계였던 것에 비해, 구체화 단계에서는 전통적 평가 지향과 구성주의적 평가 지향의 연계가 비슷하게 나타났다. 즉, 예비교사들은 구상 단계에서 평가 영역이나 방법에 대해 주로 구성주의적 평가 측면을 고려한 아이디어를 떠올렸지만, 이를 구체화하는 과정에서는 구성주의적 평가 목적과 평가의 전통적 측면을 동시에 고려하는 특징을 보였다고 할 수 있다. 다음 사례에서는 예비교사 G가 학생 중심의 평가 문제 만들기 활동을 설계하며 학습 동기 유발 및 피드백 제공 등 구성주의적 평가 목적을 구현하기 위해 프로그램에서의 협력적 논의 경험을 바탕으로 수행평가에 협동학습을 도입하였으나, 학생이 문제를 만들지 못할 경우 시간 관계상 점수를 깎고 교사가 만든 문제를 제공하기로 한 것과 같이 현실적인 이유로 인해 전통적인 평가 관점도 고려하고 있음을 보여주고 있다(그림 III-5).

(프로그램 중에) 협동학습을 함으로써 분위기를 바꿔서 아이들이 좀 더 이렇게 집중 있게 된다는 얘기가 있어서 협동학습도 꼭 넣게 되었던 것 같고, 또 피드백하는 것도 있었던 거 같은데.... 그러니까 교사가 개입하는 정도를 생각하게 되는 거 같아요. 원래 그거는 생각한 적 없었는데 교사가 어디까지 개입해야지 괜찮을까 생각하게 돼요. (중략) 사실 하고 싶은 건 많은데 너무 어렵지 않나요? 우리나라에서 하기에 는 시간도 너무 안 되고, 사실 모든 게 시간문제죠. (중략) 그래서 학생이 이 부분 (문제 만들기)을 못 하겠다 하면 그건 제가 만들어 놓은 문제로 대체하고 점수는 깎이는 거로 하고.

[세부 계획안]

1. 설명한 것은 분자 운동에 대해 알기까지 예상은이, (0, X 퀴즈)

예) 온도가 높아질수록 분자의 운동이 커진다는 사후 평가에 의한 것이다. (0)

타) 입자의 크기에 따라 속도가 다르다는 사전 평가이다. (X)

회향 원리는 충돌이다. (X)

2. 조건형식에 맞추어, 4인 12조 학습. 총 48명 4인 12조, 사후 평가 각 한 문항씩 기록
 <각과 인터넷 사용 가능>

3. 30분 19문항 x 4 = 76문항

3. 20분 입자 각각의 파동 (비행기) ① 운동은 한 방향으로만 진행되고 수직 방향으로 진행,
1. 조건형식 맞추기, 2. 원리
리 설명
 3) 불안정한 경우,
교과서 기입 → 교과

4. 숙업시간에 개인 이 퀴즈.

5. 해설 각 문항 원리 정리 같이 답을 같이 구분 시 있 는다,
 (수업 평가의 원리 이 유 선 명)

[평가 준거]

1. 만든 문항 (20점) 4문항 의 기박한 정도

각각	5	4	3	2	1
	<u>원리</u> <u>정리</u> <u>같이</u> <u>답을</u> <u>같이</u> <u>구분</u> <u>시</u> <u>있</u> <u>는다</u>	<u>원리</u> <u>정리</u> <u>같이</u> <u>답을</u> <u>같이</u> <u>구분</u> <u>시</u> <u>있</u> <u>는다</u>	<u>원리</u> <u>정리</u> <u>같이</u> <u>답을</u> <u>같이</u> <u>구분</u> <u>시</u> <u>있</u> <u>는다</u>	<u>원리</u> <u>정리</u> <u>같이</u> <u>답을</u> <u>같이</u> <u>구분</u> <u>시</u> <u>있</u> <u>는다</u>	<u>원리</u> <u>정리</u> <u>같이</u> <u>답을</u> <u>같이</u> <u>구분</u> <u>시</u> <u>있</u> <u>는다</u>

2. 개인 퀴즈 (20점) 각 문항 1.5점

3. 제출 오답률 문항 인 문 조 수 평 은 4 문 항 중 다, → 비 었 유 지 이 는 만 이.

<그림 III-5> 예비교사 G가 개발한 수행평가 과제

다음으로는 학생에 관한 지식을 중심으로 한 연계의 비율이 높았으며(15회, 10.7%), 주로 과학 교수 전략 및 과학 평가에 관한 지식과 연계되었다. 과학 교수 전략과의 연계는 실험 등을 구체화하는 과정에서 학생에 관한 지식을 고려한 것으로 구상 단계와 유사하게 나타났고, 과학 평가에 관한 지식과의 연계는 주로 학생의 특성과 수준, 흥미 등을 고려하여 평가 방법을 구체화하거나 점수 배점을 결정하는 형태로 나타났다. 예를 들어, 다음 사례에서 예비교사 C는 중학생의 일반적인 특성을 고려하여 보고서를 활용한 평가 방법을 구체화하고 있다(그림 III-6).

면담자 : 개인 활동지랑 모둠 활동지가 나눠져 있네요. 일부러 나눈 이유가 있나요?

예비교사 : 중학교 1학년에 대한 생각을 좀 많이 했어요. 간단한 거 좋아하고 활발하고 그런 애들이니까 관찰할 때 개인 생각을 쓴 다음 토론 토의를 시키고 발표, 조별 토의 후 모둠 활동지를 쓰는 거죠. 그래야지 참여도 막하니까. (중략) 그러니까 개인 보고서를 먼저 쓰는데, 조별로 토의하면 그 영향을 받기 때문에, 그 이후로는 개인 보고서는 수정을 금지시켜요. 그 전에 이제 걷어간다는 얘기인 거죠.

(예비교사 C의 구체화 단계 면담 중)

[세부 계획안]

POE작성.

학생들의 주어진 질문 및 문제를 풀기 위해 빈칸을 채운다. ⇒ 9등 정도

모든 빈칸은 너의 집 생화 등을 나누어 준다. 물을 적시고 관찰 (목계, 점)

물이 나누어 줄 경우 1개 ⇒ 개인 생각, 아 1개 ⇒ 정리 (모든 것)

목계변화 왜그럴까?

어떤 변화가 있었는가?

자신의 생각 정리.

메달을의 경우 어떤 차이?

생애 ⇒ 확산 개념 경험
대응단위.

학생들끼리 토론, 토의 ⇒ 8분 토론 30초 토론 즉 개인생각보면서 수정
이후에가 에원들 적신후 관찰 ⇒ 개인보면서 (생애 확인 ⇒ 확인력)

<그림 III-6> 예비교사 C가 개발한 수행평가 과제

구체화 단계에서 나타난 세 가지 PCK 구성요소 사이의 연계도 대부분 과학 평가에 관한 신념이나 학생에 관한 지식을 중심으로 이루어졌다. 예를 들어, 다음과 같이 예비교사 A는 학생의 특성을 고려하여 발표 시간과 분량, 도구 등의 교수 전략과 학생의 오개념을 알기 위한 평가 전략을 세부적으로 조정하였고, 이 과정에서 학생, 과학 평가, 과학 교수 전략에 관한 지식이 연계되었다. 따라서 구체화 단계에서 PCK 구성요소 사이 연계의 중심은 학생에 관한 지식이라고 볼 수 있다.

조를 나는 다음에 발표를 하는 건데요. 각각 1시간이고 3조로 나눠서 한 20분 내외, 중학생이면 더 짧게 해야 될 것 같아요. 아무래도 중학생은 발표하기 부담스러워 하니까 최대한 작은 영역으로. 5~6명이 한 조면 한 단원 발표하려면 되게 조금씩 나눠야 되잖아요. 부담감 줄이려고 모든 조원이 발표하고. PPT 활용가능하다고 쓰여

있는데 제가 생각하기엔 중학생은 못 쓸 것 같고, 그렇게 해서 한 시간 한 시간 하고 끝난 뒤에 학생들이 발표하는 것 보면 전체적인 아이들의 오개념이나 그런 걸 볼 수 있잖아요.

(예비교사 A의 구체화 단계 면담 중)

구성주의적 평가 지향의 비율 증가와 구상 및 구체화 단계에서 세 가지 PCK 구성요소 사이의 연계가 나타난 결과는 연구 II에서 나타나지 않았던 특징이다. PCK 구성요소 사이의 연계된 활용은 중요하며(Abell, 2008; Magnusson, Krajcik, & Borko, 1999), 교사가 PCK 구성요소를 통합적으로 고려하면 양질의 수행평가를 계획 및 실행할 수 있다(이인제 등, 2004; 민희정, 2012). 따라서 이러한 결과는 교사 전문성 개발 원리에 기초한 수행평가 전문성 향상 프로그램이 예비교사의 PCK 개발과 수행평가 전문성 향상에 도움이 될 가능성을 보여준다고 할 수 있다.

한편 전통적 신념과 구성주의적 신념이 모두 나타난 결과에서 알 수 있듯이, 예비교사들은 개별 과제 개발 과정에서 구성주의적 평가의 이론과 실행의 통합에 대해 많은 고민을 하는 것으로 나타났다. 그 결과 예비교사 C, E와 같이 학습 부담이 높은 고등학교에서는 구성주의적 수행평가가 여전히 어려울 것이라고 생각하는 경우도 있었지만, 다른 예비교사들은 프로그램에서의 협력적 논의를 통하여 중등학생 때 경험했던 수행평가의 문제점을 분석하고 구성주의적 평가의 이론과 실행의 통합을 고민해보면서 구성주의적 수행평가의 중요성에 대해 공감하고 긍정적으로 생각하게 되었다고 응답하였다. 즉, 개별 과제 개발 과정에서 PCK 구성요소 사이의 연계가 증가한 것은 예비교사들이 다음과 같이 연구 III의 프로그램을 통하여 구성주의적 평가의 중요성을 내면화하고 실행하는 과정에서 다양한 PCK 구성요소를 적극적으로 고려하였기 때문으로 해석할 수 있다.

학교에서 하는 수행평가도 보면 사실 다 지필평가잖아요. 뭐 수행하는 거 거의 없었어요. 그래서 당연히 그냥 그런가 보다 했는데. (중략) 수행평가를 구성주의적으로 하는 것에 대해서 워크숍에서 좀 중요성을 느끼고 나니까, 또 그렇게 만들려고 하다 보니까 이거 진짜 학생들이 하게 할 수도 있겠다, 이런 생각이 들면서 좀 와 닿는 거예요.

(예비교사 D의 추가 면담 중)

구성주의가 좋다는 얘기는 많은데, 그리고 생각해보면 '진짜 교육'이라고 하는 학생 중심 이런 건 제 가치관이었지만 그거에 대한 한계점도 많이 알고 있어서 항상 생각만 해왔거든요. (중략) 토의할 때도 사람들이 대부분 그 애매모호함을 고치려고 하는 거 같았어요. 혼자서 생각하는 거 보다는 다 같이 토론 토의해서 생각이 정리 되는 것이 훨씬 그게 구성주의적으로 된 것 같아요. 이 워크숍을 통해서 더 구성주의적으로 바뀌었다고 말할 수 있을 거 같네요.

(예비교사 G의 추가 면담 중)

3.3 조별 논의 과정의 PCK 구성요소

조별 논의 과정에서 나타난 PCK 구성요소의 분석 결과를 <표 III-7>에 제시하였다. 학생에 관한 지식(16회, 24.6%)의 비율이 가장 높았고 과학 평가(14회, 21.5%)와 과학 교수 전략에 관한 지식(13회, 20.0%)의 비율이 다음으로 높았다.

<표 III-7> 조별 논의 과정의 PCK 구성요소

PCK 구성요소		1조 (%)		2조 (%)		계 (%)	
A	A1	9 (13.8)	6 (9.2)	5 (7.7)	3 (4.6)	14 (21.5)	9 (13.8)
	A2		3 (4.6)		2 (3.1)		5 (7.7)
C	C1	3 (4.6)	2 (3.1)	4 (6.2)	4 (6.2)	7 (10.8)	6 (9.2)
	C2		1 (1.5)		-		1 (1.5)
I	I1	8 (12.3)	8 (12.3)	5 (7.7)	5 (7.7)	13 (20.0)	13 (20.0)
	I2		-		-		-
K	K1	3 (4.6)	2 (3.1)	4 (6.2)	1 (1.5)	7 (10.8)	3 (4.6)
	K2		1 (1.5)		3 (4.6)		4 (6.2)
L	L1	9 (13.8)	4 (6.2)	7 (10.8)	3 (4.6)	16 (24.6)	7 (10.8)
	L2		5 (7.7)		3 (4.6)		8 (12.3)
	L3		-		1 (1.5)		1 (1.5)
O	O1	5 (7.7)	2 (3.1)	3 (4.6)	1 (1.5)	8 (12.3)	3 (4.6)
	O2		3 (4.6)		2 (3.1)		5 (7.7)
합계			37 (56.9)		28 (43.1)		65 (100.0)

* A: 과학 평가에 관한 지식(A1: 평가 방법, A2: 평가 영역), C: 과학교육과정에 관한 지식 (C1: 연계, C2: 평가 목표), I: 과학 교수 전략에 관한 지식(I1: 주제 전략, I2: 교과 전략), K: 과학 내용에 관한 지식(K1: 개념 및 이론, K2: 과학 탐구 과정), L: 학생에 관한 지식 (L1: 오개념, L2: 발달 수준, L3: 동기 및 흥미), O: 과학 평가에 관한 신념(O1: 전통적 평가 지향, O2: 구성주의적 평가 지향).

과학 평가에 관한 지식(21.5%)은 개별 과제 개발 과정(32.5%)에 비해 차지하는 비율이 크게 감소하였다. 그러나 연구 II에서는 오히려 개별 과제 개발 과정보다 조별 논의 과정에서 과학 평가에 관한 지식의 비율이 증가하였다. 이는 예비교사들이 전통적 관점에서 리커트 척도나 채점 준거 및 배점 설정 등의 세부적인 평가 요소를 검토하는 대신, 전반적인 평가 방법과 평가 영역의 적절성에 대한 논의를 주로 하였기 때문에 나타난 결과로, 이는 수행평가의 타당도와 신뢰도를 높이면서도 수행평가의 구성주의적 측면의 개선에도 기여하였다. 예를 들어, 1조의 예시에서는 예비교사 D가 예비교사 B의 압력과 부피의 관계에 관한 수행평가 과제(그림 III-7)

를 검토하며 활동지의 문제가 학생의 사고 과정을 엿볼 수 없어 지필평가 성격이 강함을 지적한 후, 이와 관련한 조별 논의가 이루어지고 있음을 보여주고 있다.

예비교사 D : 여기 이 문제에서는 학생이 뭔가 생각하는 과정이라든가 그런 거는 답을 수가 없지 않아? 답만 있으니까? 오히려 좀 지필평가랑 가까운 것 같아.

예비교사 A : 3번, 4번 문제가 생각하는 거 아니야?

예비교사 C : 그래, 맞아. 근데 생각해 봤는데, 답의 이유를 안 쓰잖아. 서술형이 아니니까.

예비교사 B : 음... 어머, 그러네.

예비교사 A : (문항이) '연관성이 있는가?'만 물으니까 '없다' 하면 끝나는 거네.

(1조의 조별 논의 중)

[세부 계획안]

보일의 법칙 $P_1V_1 = P_2V_2$ $P_1V_1 = P_2V_2$ 잘 알고 있는지 확인

<활동지>

<1기압>

(가)

<3기압>

(나)

<8기압>

(다)

예시로 나타있는 (나)의 그림을 잘 보고 (가), (다)를 직접 그려보아라. (가), (나), (다)는 각각 1기압, 3기압, 8기압에서의 용기의 부피변화를 보일 것이다. 이 때, 기체에 작용하는 압력은 기체의 압력을 압력의 세기와 원시상태의 기체가 비교하여 그려라.

문제

- (가)~(다) 중 내부 압력이 가장 큰 것은 무엇인가? (다)
- (가)~(다) 중 가장 분자가 용기 벽에 충돌하는 횟수가 가장 적은 것은 어느 것인가? (가)
- (가)~(다) 중 어느 것의 부피가 가장 큰가? 또 부피가 큰 것은 그 속에 들어있는 분자의 크기와 어떤 연관성이 있는가? 부피가 클수록 분자 수가 많을 것이다. ~, (가) 분자의 크기는 부피가 크면 ~ (가)~(다) 중 용기 속 분자의 속도가 가장 빠르지는 않은가? 용기 속 분자들의 속도는 모두 ~

<그림 III-7> 예비교사 B가 개발한 수행평가 과제

과학교육과정에 관한 지식(10.8%)은 개별 과제 개발 과정(4.4%)에 비해 약간 많이 나타났는데, 주로 리트머스 종이의 색 변화를 이미 배웠는지, 압력의 단위로 파스칼을 사용해도 무방한지 등과 같이 수행평가 과제가 교육과정에서 벗어나지 않는지 검토하거나 앞뒤 수업과의 연계를 확인하는 과정에서 나타났다(9.2%). 이는 예비교사들이 조별 논의 과정을 통해 개별 과제 개발 과정에서 미처 고려하지 못한 과학교육과정 측면을 보다 고려하게 되었음을 의미한다.

과학 교수 전략에 관한 지식 측면(20.0%)은 개별 과제 개발 과정(18.9%)과 유사하게 나타났다. 즉 조별 논의 과정에서도 개별 과제 개발 과정과 유사하게 교과 전략(I2)에 관한 논의는 거의 나타나지 않았으며, 과학 교수 전략에 관한 지식 측면의 논의는 주제 전략(I1)에 대해서만 이루어졌다. 주로 수업 기법과 관련하여 특정 과학 교수 전략을 도입한 이유나 타당성에 대한 검토가 이루어졌는데, 이때 다른 예비교사가 기존의 계획안이나 문항에 수정을 제안하는 경우도 있었다. 예를 들어, 다음 사례에서는 예비교사 G가 예비교사 F의 ‘압력에 따른 초코파이의 부피 변화’를 주제로 한 수행평가 과제(그림 III-8)의 채점 기준에 대한 문제점을 지적하고, 그 대안으로 자신이 구상 단계에서 떠올렸으나 활용하지는 않았던 수행평가 아이디어(그림 III-9)를 보여주면서 실험 결과에 대한 입자 그림 그리기 전략을 도입할 필요성과 이유를 설명하고 있다.

예비교사 G : 근데 이거 조금 헷갈릴 거 같아. 너가 채점할 때 압력과 부피 반비례 관계에 대한 식을 포함하면 12점이라 했잖아. 식이 ‘PV는 일정’ 그걸 얘기하는 거겠지. 왜냐면은 이거 압력-부피 반비례 관계에 대한 설명이 이거잖아. 그런데 압력이 작아지면 초코파이가 팽창하고 압력 높아지면 초코파이가 작아지잖아. 이게 위에 설명이랑 똑같은 거잖아. 그래서 너가 헷갈릴 거 같아.

예비교사 F : 맞아 그럴 거 같긴 하다.

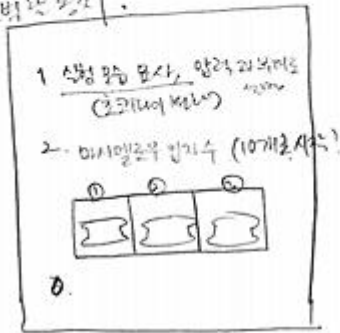
예비교사 G : 그래서 혹시 이거 가져갈 생각 있니? 왜냐면 봐봐. 초코파이 이거에 3인 1조인데 이 실험 모습이 똑같아. 이것 봐, 압력 부피 설명하는 거. 근데 두 번째 초코파이 세 개 그려 넣은 거에서 입자 수 그리기를 했거든. 처음에 마시멜로 입자 수를 10개라고 두고 애들이 압력을 높였을 때와 낮췄을 때 그거 한번 그려보라고 해서. 이거는 입자 수 때문

1. 3인 1조 실험

간이 공기펌프, 환원피이, 활동지 나누어줌

2. 실험 (간이 공기펌프에 환원피이 넣고 공기 빼줌) ← 교사가 잘 실험하는 지 확인하여 줌.

3) 개별 과제



<그림 III-9> 예비교사 G가 작성한 수행평가 아이디어

과학 내용에 관한 지식(10.8%)은 개별 과제 개발 과정(10.7%)과 유사하게 과학 개념 및 이론(4.6%), 과학 탐구 과정(6.2%)에서 모두 적게 나타났다. 이러한 결과는 조별 논의를 통해서도 예비교사들이 과학 내용에 관한 지식을 충분히 고려하여 수행평가 과제를 개발하지 못했음을 의미한다.

학생에 관한 지식(24.6%)은 개별 과제 개발 과정(15.0%)에 비해 다소 증가하였다. 학생의 오개념(10.8%)과 발달 수준(12.3%)에 대한 고려가 주를 이루었고 동기 및 흥미에 대한 고려(1.0%)는 상대적으로 적었다. 예를 들어, 다음 사례에서 예비교사 H는 고체 수산화나트륨과 리트머스 종이를 활용한 확산 실험을 계획하였는데, 이에 대해 예비교사 E는 확산 속도가 느린 고체 수산화나트륨의 활용이 학생의 오개념을 유발할 수 있다고 지적하고 있다. 이러한 결과는 조별 논의 과정이 수행평가 과제 개발 시 학생의 오개념과 발달 수준을 고려하는 데 긍정적인 영향을 미쳤음을 의미한다. 이는 구성주의적 수행평가가 이루어지기 위해서는 학생에 관한 지식의 고려가 매우 중요함을 프로그램에서 강조함은 물론 관련 오개념 자료를 제공하였기 때문일 수 있다. 하지만 학생의 동기 및 흥미 등이 충분히 고려되지 않았던 점에 대해서는 개선이 필요하다.

예비교사 E : 확산 속도는 무거운 분자일수록 더 빠르다고 생각한다고 오개념 증이
에 나와 있는데 (학생들이) 좀 헛갈릴까 봐. 왜냐면 NaOH 고체가 무
겁다고 생각하기 쉬운데 이게 확산이 느리다고 하면 '아 얘는 무거워
서 좀 느리게 확산되는구나.'로 오해하기 쉬울까 봐 이 부분은 걱정돼
요.

(2조의 조별 논의 중)

과학 평가에 관한 신념의 경우(12.3%) 개별 과제 개발 과정(18.4%)보다 약간 적
게 나타났지만, 개별 과제 개발 과정과 유사하게 전통적 평가 지향과 구성주의적
평가 지향이 전반적으로 고르게 나타났다. 즉 전통적 평가 지향은 수행평가의 객관
성이나 공정성 등을 고려할 때 나타났고, 구성주의적 평가 지향은 수행평가를 구성
주의적으로 개선하기 위한 논의를 할 때 나타났다. 이때, 예비교사들은 다음 사례
와 같이 두 평가 지향을 각각 고려하기보다는 예비교사가 도입하였던 수행평가의
구성주의적 평가 특성을 줄이지 않으면서도 평가의 전통적 측면인 객관성과 공정
성을 보장하기 위한 논의를 이어가는 경향이 있었다.

예비교사 G : 좀 더 기발한 아이디어는 맞지만 그런 게 사실 안 나올 수도 있어. 이
문항에서 내가 바라는 건 학습 기회 제공. 애들이 '아 이거는 이게 답
이구나.' 라고 생각하는 과정을 중요시 여겼으면 좋겠다는 거.

...[중략]...

예비교사 H : 근데 만약에 학생이 오줌싸개 인형 문제를 냈어. 그런데 평가하는 교
사 입장에서는 그걸 알고 있으니까 기발하다고 할 수 없잖아. 안 그
래? 이런 식으로 생각이 주관적으로 치우칠 수 있다는 거.

예비교사 G : 근데 그 이면은 추론이 불가능하잖아. 원리를 설명하면 그래도 괜찮은
데. 그럼 기준을 어떻게 만드는 게 좋다고 생각해요?

예비교사 F : 그냥 원리가 정확하게 들어있으면, 개념이 제대로 들어가 있나 이런
식으로 하면 어때?

(2조의 조별 논의 중)

이와 같이 예비교사들은 교사와 학생의 입장을 모두 고려하거나 구성주의적 수
행평가의 이론적 측면과 실행 측면을 모두 고려하는 등 프로그램에서 이루어졌던

협력적 논의와 유사한 과정을 거쳤다. 즉 수행평가를 구성주의적으로 개선시키지 못했던 연구 II의 조별 논의 결과와 달리, 연구 III에서는 프로그램을 통해 예비교사들의 조별 논의가 전통적 관점과 구성주의적 관점에서 균형 있고 체계적으로 이루어졌다. 예비교사들은 프로그램을 통해 구성주의적 평가 관점에서 개발된 점검표를 활용한 조별 논의를 수차례 거쳤는데, 이를 통해 예비교사들이 점검표를 좀 더 숙지하게 된 것이 조별 논의가 더욱 의미 있게 진행될 수 있었던 요인으로 보인다.

지난주(3주차) 토의 활동은 처음이니까 점검표 보면서 하나하나 다 따져보면서 했는데 이번에는 저번 주보다 훨씬 많았는데도 엄청 빨리 끝났더라고요. 그러면서도 질도 더 좋아진 것 같은 그런 느낌. 애들이 예전에는 그냥 두서없이 까는 것 같았는데 이번에는 '뭔가 이거는 좀, 어떤 관점에서 좀 잘못됐다.' 이런 식으로 이렇게 말하는 것 자체가 좀 체계화되었다는 게.

(예비교사 D의 추가 면담 중)

하지만 조별 논의 과정에서도 여전히 과학 교수 전략에 관한 지식 중 교과 전략, 과학교육과정에 관한 지식, 과학 내용에 관한 지식 등과 같이 일부 PCK 구성요소에 대한 고려가 적었으므로, 프로그램 및 조별 논의 과정에서 이를 강화하는 방안을 모색할 필요가 있다. 예를 들어, 예비교사들에게 개발된 수행평가 과제에 적용할 수 있는 교과 전략이 있는지, 수행평가 전후 차시의 수업은 어떻게 구상해야 할지, 평가 방법으로 실험을 선택한 경우에는 실험을 보다 개방적으로 설계하거나 과정을 평가할 방안은 없는지 등 부족한 PCK 구성요소에 관한 논의를 촉진할 수 있는 명시적인 지침을 제공하면 유용할 것이다.

3.4 조별 논의 과정의 PCK 구성요소 사이의 연계

조별 논의 과정에서 나타난 PCK 구성요소 사이의 연계 분석 결과는 <표 III-8>, <표 III-9>와 같다. 총 16회의 연계가 나타났으며, 두 가지 PCK 구성요소 사이의 연계는 87.5%(14회)였고, 세 가지 PCK 구성요소 사이의 연계는 12.5%(2회)였다. 즉 개별 과제 개발 과정에 비하여 두 가지 측면의 연계는 약간 늘었으나, 세 가지 측면

의 연계는 약간 줄었으며 연계의 유형수도 줄었다.

<표 III-8> 조별 논의 과정의 PCK 구성요소 사이의 연계 (통합 유형별)

통합 유형		1조 (%)	2조 (%)	계 (%)
2가지 PCK 구성요소의 통합	A - L	-	2 (12.5)	2 (12.5)
	A - O	2 (12.5)	1 (6.3)	3 (18.8)
	C - L	2 (12.5)	-	2 (12.5)
	C - I	-	1 (6.3)	1 (6.3)
	I - L	-	1 (6.3)	1 (6.3)
	I - O	1 (6.3)	-	1 (6.3)
	K - L	3 (18.8)	1 (6.3)	4 (25.0)
3가지 PCK 구성요소의 통합	A - I - L	-	1 (6.3)	1 (6.3)
	C - I - L	1 (6.3)	-	1 (6.3)
계		9 (56.3)	7 (43.8)	16 (100.0)

* A: 과학 평가에 관한 지식, C: 과학교육과정에 관한 지식, I: 과학 교수 전략에 관한 지식, K: 과학 내용에 관한 지식, L: 학생에 관한 지식, O: 과학 평가에 관한 신념.

<표 III-9> 조별 논의 과정의 PCK 구성요소 사이의 연계 (PCK 구성요소별)

PCK 구성요소	1조 (%)	2조 (%)	계 (%)
A와 다른 PCK 구성요소의 연계	2 (5.0)	5 (12.5)	7 (17.5)
C와 다른 PCK 구성요소의 연계	4 (10.0)	1 (2.5)	5 (12.5)
I와 다른 PCK 구성요소의 연계	3 (7.5)	4 (10.0)	7 (17.5)
K와 다른 PCK 구성요소의 연계	3 (7.5)	1 (2.5)	4 (10.0)
L과 다른 PCK 구성요소의 연계	7 (17.5)	6 (15.0)	13 (32.5)
O와 다른 PCK 구성요소의 연계	3 (7.5)	1 (2.5)	4 (10.0)
계	22 (55.0)	18 (45.0)	40 (100.0)

* A: 과학 평가에 관한 지식, C: 과학교육과정에 관한 지식, I: 과학 교수 전략에 관한 지식, K: 과학 내용에 관한 지식, L: 학생에 관한 지식, O: 과학 평가에 관한 신념.

각 PCK 구성요소를 중심으로 보면 학생에 관한 지식을 중심으로 한 연계의 비율이 32.5%(13회)로 가장 높았는데, 이 비율은 개별 과제 개발 과정에 비해서도 높다. 이때 수행평가 과제가 학생들의 수준에 적절한지 혹은 동기나 흥미를 유발할 수 있는지에 대한 논의가 먼저 이루어진 후, 과학교육과정이나 과학 평가에 관한 지식 등과 같은 다른 PCK 구성요소의 관점에서 추가적인 논의가 이어지는 특징이 있었다. 예를 들어, 1조의 사례에서는 수행평가 과제의 수준이 학생의 수준에 적합한지 논의하는 과정에서 중학교 과학과 및 수학과 교육과정에 대한 논의가 이루어졌다. 또한, 2조에서도 수행평가 과제에 대한 학생의 수준과 흥미, 세부적인 평가 준거에 대한 논의가 이루어진 후 이를 보완하기 위한 교사의 역할에 대한 논의가 이루어졌다.

예비교사 D : 샤를 법칙 배울 때 이 활동지를 다시 나눠줘서 온도-부피 관계 그래프랑 비교하고, 심화 활동으로 두 그래프를 보고 압력-온도도 어떤

관계가 있는지 알아볼 수 있으니 그렇게 한번 해 보는 활동을 할 수 있어.

예비교사 A : 근데 좀 어렵다. 중 1한테 어려워, 이거.

예비교사 D : 그니까 중 1이면 좀 어려울 것 같기도 하고.

예비교사 C : 이게 고등학생 수준인 것 같아.

예비교사 A : 그래, 그래프랑 이거 너무 어려워.

예비교사 D : 근데 중학교 중 1 교과서에도 실험하고 똑같이 그래프 그리는 문제가 있어, 그래서 한 거야.

...[중략]...

예비교사 A : 그래프를 중 1 수학 때 처음 배우지 않나? 일차방정식 할 때.

예비교사 D : 요즘은 다 내려가서.

예비교사 C : 아니 중 3 때, 2차방정식, 그 때 그래프 나와, 중 3 때.

예비교사 A : 그래프 자체가.... 그럼 과학에서는.

예비교사 C : 근데 이게 나왔으면 교육과정, 교과서에 있는 거니까.

(1조의 조별 논의 중)

예비교사 H : 근데 이거 좀 문제 만드는 활동 자체가 중학교 1학년 학생들한테 어렵지 않나? 그러니까 만들 수는 있는데... 그리고 문항에서 기발함을 평가한다고 했잖아 이것도 약간 좀 너무 주관적인.

예비교사 G : 맞아요. 기발하다는 기준이 교과서나 내가 언급한 적이 없는데 학생들이 생각해 낸 문제를 기발하다고 한 거였고.

예비교사 F : 근데 이게 학생들이 만들기 어려울 거 같아요.

...[중략]...

예비교사 G : 이걸 제가 좀 더 생각해 볼게요. 오개념은 진짜 내 바람이야. 이거는 진짜 학생들의 이해 수준에 달려있는 거 같아.

예비교사 E : 그래도 만들면서 재미는 있을 걸?

예비교사 F : 맞아, 맞아.

예비교사 G : 이것도 개념이 들어있는 거에서는 내 역할이 중요한 거 같아. 교사 역할이.

(2조의 조별 논의 중)

이러한 결과는 점검표를 활용한 조별 논의 및 프로그램을 통하여 학생에 관한

지식을 중심으로 한 연계가 보다 활성화되었음을 의미하며, 예비교사와의 면담에서도 이를 확인할 수 있었다.

토론을 하면서 점검표가 가장 많이 도움이 된 것 같기는 해요. 점검표에 있는 내용을 최대한 담으려고 노력을 했지요. (중략) 평가 같은 경우에는 단순하게 학생들이 답만 적은 걸 평가하기보다는 그 학생이 직접 설명이나 생각하는 과정을 직접 보면서 평가하는데 뭔가 좀 구성주의적 요소가 들어있다고 생각하고, 실험 같은 것도 학생들이 보고 더 편하라고 그림 같은 것도 넣어주고. (중략) 근데 수행평가에 흥미 부여를 한다는 게 제일 어려운 것 같네요.

(예비교사 D의 추가 면담 중)

과학 교수 전략에 관한 지식을 중심으로 한 연계(7회, 17.5%)도 자주 나타났는데, 실험 설계가 학생들의 수준에 적합한지 또는 오개념을 유발할 가능성이 있는지에 관한 논의로부터 시작하여 교수 전략 및 평가 문항을 수정, 보완하기 위한 논의가 이루어질 때 주로 나타났다. 예를 들어, 다음 사례에서는 학생들의 오개념을 고려하여 실험 과정 및 문제 상황을 수정하는 방법에 관한 논의가 이루어지고 있다. 과학 수업의 수준을 결정하는 데 중요한 영향을 미치는 과학 교수 전략은 수업과 밀접한 관련이 있는 수행평가의 수준과도 직결되므로, 과학 교수 전략에 관한 지식을 중심으로 한 연계는 매우 중요하다. 이런 점에서 볼 때 조별 논의 과정에서 과학 교수 전략에 관한 지식을 중심으로 한 연계가 자주 나타난 점은 바람직하다고 할 수 있다.

예비교사 E : 보일 법칙에서 펌프를 이용해 (기체의) 반을 빼내었다. 그러면 웬지 여기 기체가 떠 있다고 생각하는 사람 많을 거 같아. 나도 웬지 그렇게 생각했을 거 같아.

예비교사 G : 가열해서?

예비교사 E : 응. 이거를 이렇게 생각하기 너무 쉬운 거야. 근데 이거 온도 문제를 추가적으로 더 하면 그것도 되게 괜찮을 거 같아. 여기는 온도가 일정한데 만약에 똑같은 상태에서 온도를 올리는 거야.

(2조의 조별 논의 중)

과학 평가에 관한 지식을 중심으로 한 연계(7회, 17.5%)는 개별 과제 개발 과정과 유사한 수준에서 나타났는데, 주로 학생에 관한 지식 또는 과학 평가에 관한 신념과 연계되었다. 과학 평가에 관한 신념을 중심으로 한 연계(4회, 10.0%)는 개별 과제 개발 과정에 비하여 다소 낮은 비율로 나타났으며, 주로 과학 평가에 관한 지식과 연계되었다. 한편, 과학교육과정에 관한 지식을 중심으로 한 연계(5회, 12.5%)와 과학 내용에 관한 지식을 중심으로 한 연계(4회, 10.0%)는 개별 과제 개발 과정과 비교하면 소폭 증가하였으나 여전히 다른 PCK 구성요소를 중심으로 한 연계보다는 적은 편이었다.

즉 PCK 구성요소 사이의 연계가 거의 나타나지 않았던 연구 II의 조별 논의 결과보다 연계의 빈도는 증가하였으나, 연구 III의 개별 과제 개발 과정에 비해서는 충분히 향상되지 못했음을 알 수 있다. PCK 구성요소들은 서로 밀접한 관련이 있으므로 각 요소 간의 연계가 충분히 이루어져야 수행평가의 실행 수준도 높아질 가능성이 높다. 이런 측면에서 볼 때 현재의 프로그램은 조별 논의 과정에서 PCK 구성요소 사이의 연계를 강화하는 데에는 일부 제한이 있었던 것으로 보인다. 따라서 이런 점을 개선할 수 있는 조별 논의 및 프로그램 방법을 마련해야 할 것이다. 예를 들어, 프로그램의 협력적 논의에 참여하는 연구자가 예비교사들이 잘 고려하지 못하는 것으로 나타난 과학교육과정 또는 과학 내용에 관한 지식을 포함한 연계를 촉진할 수 있는 의견을 적극적으로 제시하는 방안을 고려할 수 있다. 이때 구상 및 구체화 단계에서 PCK 구성요소 사이 연계의 중심이 되었던 학생과 과학 교수 전략에 관한 지식에 기초하면 효과적일 것이다.

4. 결론

연구 III에서는 예비교사의 수행평가 전문성 향상을 위한 시사점을 얻기 위하여 예비교사의 수행평가 전문성 향상을 위한 프로그램을 개발하여 적용한 후, 예비교사의 개별 과제 개발 과정과 조별 논의 과정에서 고려된 PCK 구성요소와 요소 사이의 연계를 분석하였다.

교사의 PCK와 수행평가 전문성은 경험을 통해 개발되는 실천적 지식이므로(임

청환, 2003; Bond-Robinson, 2005), 예비교사에게 수행평가 과제 개발 경험을 제공할 필요가 있다. 그러나 예비교사의 수행평가 과제 개발 실태를 조사한 연구 II에서는 예비교사들이 구상 단계에서 채점 준거 및 배점 설정에 치중하는 경향이 있어 과학 평가에 관한 지식의 비율이 매우 높았고, 연계된 PCK의 유형과 맥락에 한계가 있었으며, 구체화 단계에서는 PCK 구성요소 사이의 연계가 거의 나타나지 않는 것으로 보고되었다. 그러나 연구 III에서 수행평가 전문성 향상 프로그램에 참여한 예비교사들은 과학 평가에 관한 지식뿐 아니라 교수 전략이나 학생에 관한 지식 등 다른 PCK 구성요소도 적극적으로 고려하기 시작하였다. 또한, PCK 구성요소 사이의 연계도 구상 및 구체화 단계에서 활발히 이루어졌고 세 가지 PCK 구성요소 사이의 연계도 나타났다.

조별 논의 과정에서도 연구 II에서는 주로 전통적 평가 관점에서만 이루어졌고, 논의 과정에서 학생의 응답과 사고의 폭을 제한하는 등 오히려 평가의 구성주의적 특성이 약해지는 문제점이 보고되었다. 그러나 연구 III에서 프로그램과 점검표를 제공한 후 실시한 조별 논의에서는 학생에 관한 지식이 적극적으로 고려되면서 다른 PCK 구성요소와의 연계가 활발하게 이루어지는 계기가 되었으며, 구성주의적 평가에 대한 이론과 실행의 간극을 구성주의적인 방향으로 극복하려는 노력도 이루어지는 것으로 나타났다. 수행평가가 구성주의적 평가를 구현하기 위해 도입된 체계이며, 교사의 PCK 수준은 PCK 구성요소 사이의 연계 능력에 따라 좌우된다는 주장을 고려할 때(Friedrichsen *et al.*, 2009; Park & Chen, 2012). 교사 전문성 개발 원리에 기초한 수행평가 전문성 향상 프로그램은 예비교사의 PCK 개발과 수행평가 전문성 향상에 도움이 될 가능성이 있다고 할 수 있다.

연구 III에서 나타난 구성주의적 수행평가의 핵심적인 특징은 학생 중심의 평가라 할 수 있다. 모든 과정에서 학생에 관한 지식의 비율이 높았으며, 개별 과제 개발 과정의 구체화 단계와 조별 논의 과정에서 나타난 PCK 구성요소 사이의 연계는 학생에 관한 지식을 중심으로 이루어졌다. 구성주의적 평가에 대한 신념이 학생의 다양한 측면을 고려하는 것과 관련하여 주로 나타난 것도 이러한 해석을 뒷받침한다. 특히 예비교사들은 제공받은 오개념 자료를 단순히 평가 문항에 반영하는 수준을 넘어 학생의 선개념이 드러나도록 하는 전략을 구상하거나 학생의 발달 수준, 동기 및 흥미 등을 고려하는 방안을 모색하는 데 다양하게 활용하였다. 구성주의적

평가는 학습자의 능동적인 지식 구성과 참여를 강조하므로(박선미, 1998), 이러한 결과는 구성주의적 평가 관점에서 바람직하다고 볼 수 있다.

하지만 예비교사들은 프로그램 및 점검표를 통해서도 교과 전략이나 과학 탐구 과정 지식 등과 같은 특정 PCK 구성요소를 활용하는 데는 여전히 어려움을 겪었다. 또한, 구성주의적 평가 특성 중 지식의 구성 과정에 관한 평가, 고등 사고능력에 관한 평가 등은 일부 예비교사에게서 제한적으로 나타났다. 연계 측면에서도 과학교육과정에 관한 지식 및 과학 내용에 관한 지식 등과 같이 일부 PCK 구성요소를 중심으로 한 연계가 충분히 이루어지지 않았고, 조별 논의 과정에서 3가지 PCK 구성요소 사이의 연계가 일부 감소하는 등 부족한 점이 있었다. 이러한 원인 중 한 가지는 예비교사들이 중등학생을 대상으로 한 실제 수업 과정을 거치지 않는 상황에서 수행평가 과제를 개발하였기 때문일 수 있다. 즉, 예비교사들이 학교 현장에서와 같이 수업을 실시하는 상황에서 수행평가 과제를 개발하였다면 과학교육과정에 관한 지식 등 미흡했던 PCK 구성요소가 자연스럽게 고려되었을 가능성이 있다. 그러나 예비교사에게 이러한 경험을 제공하는 것은 현실적으로 쉽지 않으므로 향후 수행평가 전문성 향상 프로그램이나 평가 전문성 관련 강좌에서는 이 연구에서 드러난 제한점을 보완하는 방안을 적극적으로 모색할 필요가 있을 것이다.

제 6 장. 결론 및 제언

1. 연구의 요약

전통적 지필평가의 한계를 극복하고자 도입된 수행평가는 학생들의 상대적인 석차를 산출하는 지필평가에 비해 개별 학생의 성취수준을 타당하게 판단할 수 있는 평가 방법이며, 현재 과학교육의 패러다임으로 자리 잡은 구성주의적 평가를 구현할 수 있는 이상적인 평가 체계이다. 특히 최근 성취평가제가 도입됨에 따라, 교사의 수행평가 전문성 향상의 중요성은 더욱 커지고 있다. 그러나 많은 과학교사는 수행평가를 구성주의적으로 개발하고 실행하는 데 어려움을 겪고 있다. 따라서 과학교사의 수행평가 전문성 향상을 위한 전략의 모색이 필요하다. 이에 이 연구에서는 수행평가의 개발 과정에 다양한 PCK 구성요소가 활용된다는 점을 바탕으로, 교사가 PCK를 효과적으로 활용하고 개발할 수 있는 활동 기회를 제공함으로써 교사들의 수행평가 전문성을 향상하고자 하였다.

이를 위한 기초 연구로 연구 I에서 현직 중등 과학교사의 과학 평가 실태와 지향을 평가 방법과 목적 측면에서 분석한 결과, 교사들이 실행하는 과학 평가는 방법적인 측면에서는 비교적 다양한 평가 방법을 활용하였으나, 평가의 목적은 대부분 전통적인 총괄평가였다. 즉, 교사들은 과학 평가 방법에 비하여 구성주의적 평가 목적에 대한 이해가 부족한 것으로 나타났다. 다음으로 전형적인 수행평가 상황에서 교사들이 지향하는 평가 사례를 분석한 결과, 과학교사들은 다양한 유형의 구성주의적 평가 방법을 지향한 반면, 평가 목적 측면에서는 여전히 전통적 총괄평가의 비율이 높았다. 즉, 평가 방법과 목적 측면에서 나타난 평가 실태와 지향 사이의 간극이 과학교사의 평가 전문성에서 나타난 주요 특징이자 문제점으로 도출되었다. 이때 교사들이 형성평가를 지향하는 비율은 구성주의적 과학 평가관 수준이 높은 교사가 구성주의적 과학 평가관 수준이 낮은 교사들에 비해 유의미하게 높았다. 이는 구성주의적 평가에 관한 인식 함양이 매우 중요함을 의미한다. 이에 구성주의적 평가에 관한 인식 함양을 위해서는 체계적이고 장기적인 교육이 필요하다는 관점에서 예비교사 단계에서의 교육 필요성이 도출되었다.

연구 II에서는 중등 예비 화학교사 8명의 수행평가 개발 과정과 조별 논의 과정

에서 나타나는 특징을 PCK 구성요소 및 구성요소 사이 연계의 측면에서 심층적으로 조사하였다. 연구 결과, 개별 과제 개발 과정에서 나타난 PCK 구성요소는 과학 평가에 관한 지식에 편중되어 있었고, 구성주의적 평가 측면도 거의 고려되지 않았다. PCK 구성요소 사이의 연계는 구상 단계에서 과학 교수 전략에 관한 지식을 중심으로 이루어졌으나, 구체화 단계 및 조별 논의 과정에서는 거의 나타나지 않았다. 조별 논의 과정에서는 예비교사들이 고려하지 못했던 일부 PCK 구성요소가 보완되었으나, PCK 구성요소의 특징은 구체화 단계와 유사하였고, 주로 전통적 관점에서 논의가 이루어진 한계가 있었다. 연구 II에서 나타난 예비교사의 수행평가 개발 과정의 특징과 문제점은 수행평가 과제를 전통적 관점에서 개발하고 논의한다는 점, PCK 구성요소의 연계가 거의 이루어지지 않고 과학 평가에 관한 지식에 지나치게 편중되어 있다는 점, 특히 학생에 관한 지식에 대한 고려가 매우 부족하다는 점으로 요약할 수 있다. 따라서 예비교사에게 단순히 수행평가 과제 개발 기회를 제공하는 것만으로는 예비교사의 수행평가 전문성 향상에 큰 한계가 있음을 알 수 있었고, 수행평가 전문성 향상 프로그램 개발에 필요한 구체적인 시사점을 얻을 수 있었다.

연구 III에서는 Loucks-Horsley 등(2009)의 교사 전문성 개발 원리와 전문성 개발 설계 프레임워크를 바탕으로 예비 화학교사의 수행평가 전문성 향상을 위한 프로그램을 개발하였다. 먼저 연구 II의 결과를 고려하여 구성주의적 평가에 대한 이론적 이해를 바탕으로 다양한 수행평가 사례 분석 및 개발 실습 활동을 하도록 하였고, 점검표와 오개념 자료를 제공하였다. 이때 연구 I에서 나타난 평가 실태와 지향 사이의 간극을 줄이기 위해 연구자와 예비교사 사이의 협력적 논의와 교사와 학생의 이중적 입장에서의 논의 활동을 실시하였다. 개발한 프로그램이 예비 화학교사 8명의 수행평가 전문성에 미치는 영향을 연구 II와 같은 관점에서 심층적으로 분석한 결과, 예비교사들은 과학 평가에 관한 지식뿐 아니라 교수 전략, 학생, 내용에 관한 지식 등 다른 PCK 구성요소도 적극적으로 고려하였다. 개별 과제 개발 과정의 구상 단계에서는 과학 평가, 교수 전략, 학생, 내용에 관한 지식이 거의 동일한 비율로 나타났고, 구체화 단계에서는 과학 평가에 관한 지식이 주로 나타났다. 조별 논의 과정에서는 과학 평가, 교수 전략, 학생에 관한 지식의 비율이 거의 같았다. PCK 구성요소 사이의 연계는 개별 과제 개발 과정의 구상 단계에서 과학 교수 전

략에 관한 지식을 중심으로 나타났고, 구체화 단계와 조별 논의 과정에서는 학생에 관한 지식을 중심으로 나타났다. 특히 조별 논의 과정에서는 학생에 관한 지식이 적극적으로 고려되면서 다른 PCK 구성요소와의 연계가 활발하게 이루어지는 계기가 되었으며, 구성주의적 평가에 대한 이론과 실행의 간극을 구성주의적인 방향으로 극복하려는 노력도 이루어졌다. 하지만 예비교사들은 과학교육과정에 관한 지식, 교과 전략, 탐구 과정 지식 등의 일부 PCK 구성요소를 활용하거나 다른 PCK 구성요소와 연계하는 데 여전히 어려움을 겪었으며, 이에 대한 보완 방안의 모색이 필요한 것으로 나타났다.

2. 결론 및 제언

지금까지 과학교사의 수행평가 전문성 향상을 위해 PCK 개발 관점에서 체계적이고 실증적으로 접근한 연구는 부족했으며, 특히 예비교사의 수행평가 전문성 향상을 중심으로 한 연구는 거의 없었다. 현재 사범대학에서 예비 과학교사들에게 수행평가를 개발하는 기회를 일부 제공하고 있지만, 연구 결과를 종합하면 예비교사들은 수행평가 과제 개발 과정에서도 전통적 관점에서 평가의 신뢰도와 타당도만을 주로 고려하는 것으로 드러나, 단지 수행평가 개발 기회를 제공하는 것만으로는 구성주의적 관점에서의 바람직한 수행평가 전문성을 향상시키는 데 한계가 있음을 알 수 있었다. 이러한 점에서 교사 전문성 개발 원리에 기초하여 예비 화학교사의 평가 전문성 향상을 위한 프로그램을 체계적으로 개발 및 적용한 이 연구의 의미가 있다. 이 연구에서 개발한 수행평가 전문성 향상 프로그램에 참여한 예비교사들은 수행평가 과제 개발 과정에서 본래 잘 고려하지 못하였던 PCK 구성요소를 고려할 수 있게 되었으며, 이 과정에서 PCK 구성요소 사이의 연계가 특히 활발하게 나타났다. 또한, 연구자와 예비교사 사이의 협력적 논의 활동을 통해 평가의 전통적 측면과 구성주의적 측면의 접점을 찾으려 노력하였고, 이 과정에서 구성주의적 수행평가가 개발될 수 있었던 것으로 나타났다. 결론적으로, 교사 전문성 개발 원리에 기초한 수행평가 전문성 향상 프로그램은 예비 화학교사의 수행평가 전문성 향상에 효과적이라고 할 수 있다. 그러므로 연구 III에서도 나타난 일부 부족한 측면을 보완하여 예비교사 교육과정의 평가 관련 강좌에 프로그램을 도입하고 활용한

다면, 예비 화학교사의 수행평가 전문성을 향상하여 구성주의적 수행평가의 확산 및 성취평가제의 내실화에 기여할 수 있을 것이다. 이상의 결론으로부터 예비 화학교사의 평가 전문성 향상을 위한 프로그램의 개선 및 활용 방향에 대한 구체적인 제언은 다음과 같다.

첫째, 프로그램에서 협력적 논의에 참여하는 연구자는 예비교사들이 어려워하거나 덜 고려하는 것으로 나타난 PCK 구성요소 및 구성요소 사이의 연계에 대해 부족한 부분을 명시적으로 보완할 수 있도록 도움을 제공할 필요가 있다. 예를 들어, 연구 III에서 과학교육과정에 관한 지식과 과학 내용에 관한 지식의 하위 요소 중 과학 과정 지식을 고려한 경우는 여전히 드물었고, 과학교육과정과 과학 내용에 관한 지식이 다른 PCK 구성요소와 연계되는 경우는 적었다. 협력적 논의에 참여하는 연구자는 예비교사들이 구성주의적 평가의 이론적 측면을 고려할 수 있도록 자연스럽게 촉진하는 역할을 담당한다고 볼 수 있으므로, 해당 PCK 구성요소에 관한 의견을 적극적으로 제시하면 예비교사들이 이에 대한 실천적인 측면을 고려한 의견을 제시하는 과정에서 자연스럽게 해당 PCK 구성요소에 관한 논의와 반성적 사고가 촉진될 것이다. 예를 들어, 실험 및 탐구 과정에 관한 평가에 관한 논의는 과학 과정 지식의 활용을 촉진할 수 있고, 융합인재교육(STEAM) 측면의 논의는 다른 교과와의 연계를 고려해야 하므로 과학교육과정에 관한 지식의 활용을 촉진할 수 있다.

둘째, 조별 논의 과정에서는 예비교사들 사이의 논의가 더욱 체계적으로 이루어질 수 있도록 관련 도움과 지침을 명시적으로 제공하거나 역할을 분담하는 방안이 효과적일 수 있다. 예를 들어, 프로그램에 참여한 예비교사들은 연구자가 조별 논의 과정에 참여하지 않았음에도 한 예비교사가 자신의 수행평가 과제를 발표하면 다른 예비교사들이 자연스럽게 이에 관한 구성주의적 관점에서의 의견을 제시하는 방식으로 연구자의 역할을 대체하는 것으로 나타났다. 그러므로 조별 논의 과정에서 이러한 역할 분담이 보다 분명하게 이루어질 수 있도록 평가의 구성주의적 측면에서의 검토 담당과 전통적 측면에서의 검토 담당을 지정하고, 다음 예비교사가 수행평가 과제를 발표할 때에는 역할을 바꾸어 논의를 진행하는 등의 조별 논의 전략을 적용할 수 있을 것이다. 또한, 이 연구에서는 조별 논의가 특정 PCK 구성요소에 치우치지 않도록 모든 PCK 구성요소를 고르게 담은 점검표를 제공하였으나,

연구 III에서도 부족한 것으로 나타난 과학교육과정 또는 내용에 관한 지식 등의 일부 PCK 구성요소에 대해서는 해당 PCK 구성요소의 논의가 이루어질 수 있는 지침을 구체적으로 제공할 수 있다. 특히 만약 외부 전문가의 멘토링이나 컨설팅이 가능한 상황이라면, 외부 전문가는 예비교사들이 프로그램과 점검표 제공을 통해서도 여전히 어려움을 겪는 해당 부분에 집중적으로 도움을 제공하는 것이 효과적일 것이다.

셋째, 예비교사들은 POE, 순환학습 등의 교과 전략을 수행평가 과제에 도입하는 경우가 거의 없었다. 그러나 이러한 수업 모형이나 교과 전략은 구성주의적 관점에 기반한 경우가 많으므로 이를 활용하면 구성주의적 수행평가 과제를 쉽게 개발할 수 있다는 장점이 있다. 따라서 이 프로그램을 평가 관련 강좌에만 적용하는 것이 아닌, 수업 모형을 다루는 강좌에도 적용하여 교과 전략을 활용한 수행평가 과제 개발 활동 등을 실시한다면 예비교사의 수행평가 전문성 향상에 도움을 줄 수 있을 것이다. 특히 이러한 적용은 수업과 평가가 함께 이루어지는 형성평가의 특성을 강조함으로써 예비교사의 구성주의적 평가의 인식을 높이는데 기여할 수 있으므로 특히 권장된다.

넷째, PCK는 실제 수업이나 평가 활동을 통해 개발되는 실천적 특징이 있고, 예비교사의 수행평가 전문성을 평가하기 위해서는 개발된 수행평가 과제의 수준뿐 아니라, 이를 학생들에게 적용하는 과정과 이후의 결과까지 종합적으로 고려할 필요가 있다. 이러한 점에서 연구 III의 수행평가 전문성 향상 프로그램은 실제 수업 과정을 거치지 않고 가상의 학생을 대상으로 수행평가 과제를 개발하도록 하였다는 점에서 한계가 있다. 이러한 점에서 교육실습의 기회를 이용하는 것이 이상적이지만, 현실적으로 예비교사들이 교육실습 중 실제 학생을 대상으로 수행평가를 개발하고 실행하도록 하는 것은 어려움이 매우 크다. 따라서 평가 관련 강의에 본 프로그램을 적용할 경우, 가능한 실제와 유사한 상황에서 수행평가가 이루어질 수 있도록 다른 예비교사들이 학생의 역할을 수행하도록 하고 시연을 통해 개발한 수행평가 과제를 적용해보는 등 하는 방안 등을 고려할 필요가 있다.

다섯째, 교사 전문성 개발 원리에 기초한 평가 전문성 향상 프로그램의 현장 확산을 위한 방안도 모색할 필요가 있다. 이 연구에서 개발한 프로그램은 연구자-예비교사의 협력적 활동을 포함하므로 연구자의 역할이 매우 중요하다. 즉, 협력적

논의에 참여하는 연구자는 예비교사를 가르치기보다는 능동적인 성인 학습자로 대우하며, 구성주의적 수행평가의 현실적 어려움 속에서도 함께 접점을 찾아가는 자신의 역할을 충분히 이해할 필요가 있다. 만약 이 프로그램을 사범대학의 예비교사 교육현장에 적용할 경우, 강의를 담당하는 교수가 연구자의 역할을 수행할 수 있지만 이외에도 사범대학에 재학 중인 파견 교사, 과학교육 전공 대학원생 등도 연구자의 역할을 할 수 있을 것으로 예상된다. 특히 현직교사 신분이거나 교사 경험이 있는 연구자가 예비교사와의 협력적 논의에 참여한다면, 그들은 이론 및 현실적인 측면을 폭넓게 고려한 의견을 제시할 수 있으므로 구성주의적 수행평가의 이론과 실제의 통합에 많은 도움을 줄 수 있다. 따라서 평가 관련 강좌와 이들을 효과적으로 연결할 수 있는 네트워크 및 행정적, 재정적 지원 체계를 마련하는 것도 필요할 것이다.

3. 추후 연구 과제

이 연구를 바탕으로 다음과 같은 추후 연구가 이루어질 수 있다.

첫째, 연구 I에서 실험 상황보다 토의나 자유탐구 상황에서 형성평가의 비율이 높게 나타난 것은 이 상황들이 더 개방적이고 학생 중심으로 이루어지는 활동과 관련 있기 때문일 가능성이 있다. 그러므로 구성주의적 평가 환경을 조성하기 위해서는 평가 상황의 특성이 평가의 목적과 방법에 미치는 영향에 대해서도 연구가 이루어질 필요성이 있다. 예를 들어, 전통적인 실험과 개방적인 실험 상황에서 교사들의 평가 목적과 방법이 어떻게 달라지는지 연구할 수 있을 것이다. 또한, 평가 계획 및 실행에는 평가 목적에 대한 인식이나 방법에 대한 지식 이외에도 교육과정과 학생 등 여러 측면에 대한 고려가 종합적으로 이루어진다. 따라서 교사들의 과학 평가 과정을 보다 심층적으로 분석하여 교사들이 대표적인 구성주의적 평가 방법을 사용하면서도 평가의 목적은 전통적인 관점에 머무르는 이유를 밝힐 필요성이 있다.

둘째, 연구 I에서 과학교사들은 구성주의적 평가 목적에 대한 인식이 부족한 것으로 나타났지만, 연구자가 설계한 설문지를 바탕으로 정량적인 조사만을 수행하였기 때문에 과학교사의 평가 실태와 지향이 충분히 드러나지 않았을 수 있다. 교사

의 평가 실태에는 교사의 개인적 요인뿐 아니라 우리나라의 과도한 입시 경쟁 분위기나 학급당 학생 수와 같은 사회적, 환경적 측면의 외부 요인도 영향을 미칠 수 있으므로, 교사의 평가 실태를 보다 구체적으로 파악하기 위해서는 이러한 외부 요인까지 고려할 필요가 있다. 이를 위한 분석 방법으로 교사 개인과 그 주변의 대상 사이의 상호작용을 고려하는 문화역사적활동이론(cultural historical activity theory; CHAT; Roth & Tobin, 2002)의 활용을 고려할 수 있다. CHAT와 같은 정성적 분석 방법을 평가 실태 분석에 활용하면 연구 I과 같은 정량적 연구 결과와 차별화된 시사점을 밝힐 수 있을 것이다.

셋째, 연구 II와 연구 III에서 예비교사들은 실험을 화학 교과의 특징적이고 중요한 평가 방법으로 생각하여 대부분 실험을 활용한 수행평가 과제를 개발한 것으로 나타났다. 이는 이 연구에서 수행평가를 개발하도록 한 단위 특성의 영향이 나타난 것일 수 있다. 하지만 수행평가는 토의, 자유탐구, 포트폴리오 등 다양한 평가 방법으로 이루어질 수 있고 때로는 실험을 활용하기 곤란한 단위도 있으므로, 예비교사의 수행평가 전문성 향상을 위해서는 다양한 평가 방법을 활용한 수행평가 과제 개발 과정에서 나타나는 PCK 구성요소와 구성요소 사이의 연계에 관한 연구가 지속해서 이루어질 필요성이 있다.

넷째, 이 연구에서는 연구자 1명이 예비교사 8명과 함께 협력적 논의를 하였으나, 4명으로 편성된 조마다 연구자가 한 명씩 참여하는 등 논의를 진행하는 연구자와 예비교사의 비율을 다르게 하면 협력적 논의의 특징이 달라질 수 있다. 또한, 연구자가 교사로서의 경력이 있는지도 협력적 논의에 영향을 미칠 수 있다. 한편, 교사와 학생의 이중적 입장을 고려한 논의는 예비교사의 학년에 따라 논의 양상이 달라질 수 있다. 이와 같은 다양한 조건에서 나타나는 특징을 조사함으로써 연구 III에서 활용되었던 연구자와 예비교사 사이의 협력적 논의, 교사와 학생의 이중적 입장에서의 논의가 효과적으로 이루어질 수 있는 방법이나 조건을 찾는 연구도 필요하다.

다섯째, 이 연구에서는 소수 인원을 대상으로 양적 및 질적 분석을 하였으므로, 해당 프로그램의 효과를 체계적으로 밝히는 데에는 한계가 있다. 따라서 추후 프로그램의 효과를 검증하기 위한 양적 연구가 다양한 수행평가 상황에서 체계적으로 이루어질 필요가 있다. 또한, 교사의 구성주의적 평가에 관한 인식이나 PCK는 오랜

시간에 걸쳐 형성된다. 따라서 단기간의 프로그램이나 몇 차례의 수행평가 과제 개발 활동만으로는 예비교사의 수행평가 전문성을 충분히 향상시킬 수 없으므로, 예비교사들이 구성주의적 평가관과 PCK를 내면화할 수 있도록 구성주의적 평가나 교수학습에 관한 장기간의 교육이 이루어져야 한다. 이와 더불어 예비교사가 수행평가뿐 아니라 지필평가 등 학교 현장에서 이루어지는 다양한 평가에 대한 전문성을 종합적으로 갖출 수 있도록 이에 관한 교육과 연구가 지속해서 이루어져야 할 것이다.

참 고 문 헌

- 강대중 (2016). 초등학교 초임교사의 학생평가 전문성 수준 분석. 교육과정평가연구, 19(2), 155-170.
- 강대중, 염시창 (2014). 초등학교 교사들의 형성평가에 대한 인식 및 실태 분석. 학습자중심교과교육연구, 14(2), 27-43.
- 강신포, 김호선 (2006). 수학과와 타 교과 내용의 연계성 분석. 과학교육연구, 31, 1-20.
- 고문숙, 남정희 (2013). 협력적 멘토링을 통한 초임 중등과학교사의 교수실행에서 나타나는 반성적 실천의 변화. 한국과학교육학회지, 33(1), 94-113.
- 고문숙, 이순덕, 최정희, 남정희 (2009). 초임 과학교사의 반성적 실천을 위한 협력적 멘토링의 효과. 한국과학교육학회지, 29(5), 564-579.
- 고미례, 남정희, 임재항 (2009). 신입 과학교사의 교과교육학 지식(PCK)의 발달에 관한 사례 연구. 한국과학교육학회지, 29(1), 54-67.
- 고민석, 김은애, 허진미, 양일호 (2013). 초등학교 교사의 탐구에 대한 신념과 과학과 수행평가의 실제. 대한지구과학교육학회지, 6(2), 124-135.
- 곽영순 (2008). 과학과 교과교육학 지식 유형별 교사 전문성의 특징 연구. 한국과학교육학회지, 28(6), 592-602.
- 곽영순 (2009). 중등 초임 과학교사의 수업 전문성 개발 실태 분석. 한국지구과학회지, 30(3), 354-365.
- 곽영순, 김찬중, 이양락, 정득실 (2006). 초, 중등 학생들의 과학 흥미도 조사. 한국지구과학회지, 27(3), 260-268.
- 교육과학기술부 (2011). 과학과 교육과정 총론. 고시 2011-361호. 서울: 교육과학기술부.
- 교육과학기술부 (2012). 2009 개정 교육과정. 고시 2012-14호. 서울: 교육과학기술부.
- 교육부 (1997). 제7차 교육과정. 고시 1997-15호. 서울: 교육부.
- 교육부 (1998). 수행평가의 이해. 서울: 교육부.
- 교육부 (2015). 초·중등학교 교육과정 총론. 고시 2015-74호. 세종: 교육부.

- 김경미, 김성원 (2002). 과학과 평가 영역간의 중요도에 관한 교사들의 인식에 관한 연구. 한국과학교육학회지, 22(3), 540-549.
- 김경순, 윤지현, 박지애, 노태희 (2011). 중등 과학 예비교사들의 수업시연 계획 및 실행에서 나타난 교과교육학지식의 요소. 한국과학교육학회지, 31(1), 99-114.
- 김경희, 김신영, 김성숙, 지은림, 반재천, 김수동 (2006). 교사의 학생평가 전문성 기준 개발. 교육평가연구, 19(2), 89-112.
- 김경희, 박은아, 송미영, 상경아, 김수진, 김희경, 신진아, 서지영, 이채희, 김지영, 김기연, 최숙기 (2012). 중등학교의 성취평가제 운영 방안 연구. 연구보고 CRE 2012-8. 서울: 한국교육과정평가원.
- 김구연 (2010). 수학교사의 수업전문성 신장을 위한 교사 연수 프로그램 개발의 기본 관점. 한국학교수학회논문집, 13(4), 619-633.
- 김덕영, 박종원 (2015). 학생의 열린 과학 탐구 보고서 작성을 돕기 위한 점검표 개발. 한국과학교육학회지, 35(6), 1075-1083.
- 김동영 (2006). 과학교사의 학생 평가 전문성 신장을 위한 자기 진단 프로그램 개발 연구: 지구과학 내용을 중심으로. 서울대학교 대학원 박사학위논문.
- 김동진 (2005). 수행평가에 관한 교사들의 인식 및 실천 분석. 경남대학교 대학원 박사학위논문.
- 김두정 (2006). 교실 수준 교육과정의 운영에 대한 학생의 이론: 포스트모더니즘 시각에서의 분석. 교육과정연구, 24(1), 1-40.
- 김상달, 최성봉, 한상아 (2008). 지구과학 수행평가 모형의 개발 및 활용방안. 대한 지구과학교육학회지, 1(1), 1-13.
- 김석우 (2007). 고등학교 과학과 수행평가 실태분석 및 개선방안. 교육평가연구, 20(4), 53-73.
- 김선영 (2014). 구성주의를 강조한 교사 양성 프로그램을 통한 예비 생물교사의 인식론적 신념 및 세계관 변화 탐색. 교원교육, 30(4), 235-254.
- 김선희 (2006). 학생평가 전문성을 갖춘 수학교사 양성을 위한 ‘수학학습평가’ 강좌의 교육 내용과 방법에 대한 제안. 대한수학교육학회지, 8(3), 301-326.
- 김선희 (2012). 실습과 반성을 통한 수학 예비교사의 평가 전문성 신장. 수학교육학 연구, 22(2), 277-292.

- 김성원, 현미숙 (2005). 중학교 과학교육과정에서 수행평가의 일반적 사항과 적절한 수행평가 방법에 관한 과학 교사들의 인식 조사. 교과교육학연구, 9(2), 213-232.
- 김송자, 최창우 (2006). 초등학교 수학과 수행평가 실천에 관한 연구: 자가평가·동료평가·관찰평가를 중심으로. 한국초등수학교육학회지, 10(1), 67-87.
- 김수동, 김정희, 김수진 (2005). 교사의 학생평가 전문성 기준. 연구보고 ORM 2005-51-1. 서울: 한국교육과정평가원.
- 김수동, 전영석 (2005). 과학수업에서 학생평가를 잘 하려면. 연구보고 ORM 2005-51-5. 서울: 한국교육과정평가원.
- 김신영 (2002). 현장교사의 평가전문성 연구. 교육평가연구, 15(1), 67-85.
- 김신영 (2004). 학생평가 전문성의 개념과 국가수준의 기준개발. 교원의 학생평가 전문성의 개념과 국가기준 개발방향 탐색 워크숍 자료집. 서울: 한국교육과정평가원.
- 김신영 (2007). 교사의 평가 전문성과 중등교사 양성과정. 교육평가연구, 20(1), 1-16.
- 김신영, 송미영 (2008). 예비교사의 학생평가 전문성과 교육실습의 효과. 교육평가연구, 21(1), 79-97.
- 김윤희, 윤기순, 권덕기 (2010). Bloom의 신 교육목표분류에 기초한 중학교 생물 영역 총괄 평가 문항의 목표 분석. 과학교육연구지, 34(1), 164-174.
- 김은진 (2000). 과학 교과 수행평가들의 개발. 서울대학교 대학원 박사학위 논문.
- 김정숙, 여철현 (2002). 탐구지향적 과학의 구체적 수행평가 모형 개발. 연세교육과학, 50, 93-103.
- 김정환 (2005). 초·중등교사의 학습자 평가 능력 수준 분석. 학습자중심교과교육연구, 5(2), 41-60.
- 김지영 (2005). 과학과 수행평가에 대한 초등교사들의 인식과 실행 실태. 대구교육대학교 대학원 석사학위논문.
- 김진수 (2007). 기술교육의 새로운 통합교육 방법인 STEM교육의 탐색. 한국기술교육학회지, 7(3), 1-29.
- 김찬중 (2012). 포트폴리오 교수 학습 및 평가. 서울: 교육과학사.

- 김찬중, 김진규, 임형 (2001). 과학과 수행평가의 이해와 활용. 서울: 경문사.
- 김형준, 유준희 (2010). 중학생 과학탐구활동 수행평가 시 채점 방식 및 척도의 수에 따른 신뢰도 분석. 한국과학교육학회지, 30(2), 275-290.
- 김형준, 유준희 (2012). 중학생 과학탐구활동 수행평가 시 총체적 채점에서 나타나는 채점자간 불일치 유형 분석. 한국과학교육학회지, 32(1), 160-181.
- 김호진, 성민웅, 광대오 (2000). 중등학교 과학교사들의 학습 평가에 관한 실태조사. 한국과학교육학회지, 20(1), 101-111.
- 남명호 (1995). 수행평가의 타당성 연구. 고려대학교 대학원 박사학위논문.
- 남정희, 김현옥, 고문숙, 고미례 (2010). 멘토링을 통한 초임중등과학교사의 탐구지향적 교수실행 변화. 한국과학교육학회지, 30(5), 544-556.
- 남정희, 성을선, 음재호, 김경희, 최병순 (1999). 형성평가에 대한 과학교사들의 인식 및 실태. 대한화학회지, 43(6), 720-727.
- 남정희, 최준환, 고문숙, 김재홍, 강순민, 임재향, 공영대 (2005). 형성평가를 이용한 교수-학습 전략이 중학교 학생들의 과학개념 이해, 학습동기, 메타인지 능력에 미치는 영향. 대한화학회지, 49(3), 311-320.
- 노태희, 강석주, 강훈식 (2012). 중등 과학영재교육에서 초임 교사의 수업 전문성 제고 전략으로써의 멘토링 적용 사례연구. 한국과학교육학회지, 32(2), 331-341.
- 노태희, 김창민 (1999). 협동적인 컴퓨터 보조 수업이 중학생들의 과학 학습에 미치는 효과. 한국과학교육학회지, 19(2), 266-274.
- 노태희, 박재성, 강훈식 (2016). 중등 예비 화학교사의 지필평가 문항 제작 과정에서 고려된 교과교육학 지식(PCK) 구성 요소 사이의 상호작용. 한국과학교육학회지, 36(5), 769-781.
- 노태희, 양찬호, 김영훈, 강훈식 (2012). 코칭을 통한 초임 과학영재교육 담당교사의 수업 전문성 변화에 관한 사례연구. 한국과학교육학회지, 32(4), 655-670.
- 노태희, 양찬호, 이재원, 유지연, 강훈식 (2013). 멘토링을 통한 코칭이 초임 과학영재교육 담당교사들의 교과교육학지식에 미치는 영향. 한국과학교육학회지, 33(5), 1021-1040.
- 노태희, 윤지현, 강석진 (2009). 초등교사의 구성주의적 과학 평가관 및 관련 변인

- 탐색. 초등과학교육, 28(3), 352-360.
- 노태희, 윤지현, 김지영, 임희준 (2010). 초등 예비 교사들이 과학 수업 시연 계획 및 실행에서 고려하는 교과교육학지식 요소. 초등과학교육, 29(3), 350-363.
- 노태희, 이재원, 강석진, 강훈식 (2015). 중등 과학교사의 과학 평가 실태와 지향. 한국과학교육학회지, 35(4), 725-733.
- 노태희, 이재원, 강석진, 한재영, 강훈식 (2017). 중등 예비 화학교사의 수행평가 개발 과정에서 나타나는 교과교육학 지식(PCK) 요소의 특징. 한국과학교육학회지, 37(2), 291-299.
- 노태희, 임희준, 우규환 (1995). 화학양론과 기체 상태에 대한 중·고등학생의 개념 이해도 비교. 한국과학교육학회지, 15(4), 437-451.
- 맹희주, 손연아 (2011). 과학 수업에서 통합적 적용 경험에 따른 초등학교 교사들의 통합과학 교육에 대한 인식 및 교과교육학 지식(PCK)의 차이 분석. 초등과학교육, 30(4), 601-614.
- 민희정 (2012). 과학교사의 학생평가 실태분석 및 교사연수를 통한 평가전문성 모델 개발. 한국교원대학교 대학원 박사학위논문.
- 박가나 (2012). 학생평가에 대한 사회과 교사의 인식 및 평가 전문성 조사. 사회과교육연구, 19(1), 29-46.
- 박도순 (1995). 연구의 최근 동향과 수행평가의 문제. 교육진흥, 27, 120-128.
- 박도순 (1999). 초·중등학교 교과별 수행평가의 실제. 서울: 한국교육과정평가원.
- 박도순 (2007). 학교의 교육문제, 어떻게 생각하여야 하는가?. 서울: 원미사.
- 박선미 (1998). 구성주의와 수행평가: 이론적, 실제적 접점. 사회과교육, 31, 339-356.
- 박성혜 (2003a). 교사들의 과학 교과교육학 지식 측정도구 개발. 한국교원교육연구, 20(1), 105-134.
- 박성혜 (2003b). 교사들의 과학 교과교육학지식과 예측변인. 한국과학교육학회지, 23(6), 671-683.
- 박성혜 (2006). 중등과학교사들의 교수법 및 자기효능감과 태도에 따른 교과교육학 지식. 한국과학교육학회지, 26(1), 122-131.
- 박영석 (2009). 초등학교 사회 수업에서 교사의 형성평가 실행에 관한 연구. 시민교

- 육연구, 41(4), 51-79.
- 박재성, 강훈식, 한재영 (2017). 해결자·청취자 활동이 예비 화학교사의 지필평가 문항 제작 과정에서 고려된 교과교육학 지식(PCK) 구성 요소 사이의 상호작용에 미치는 영향. 한국과학교육학회지, 37(3), 429-440.
- 박정 (2013a). 형성평가의 재등장과 교육 평가적 시사. 교육평가연구, 26(4), 719-738.
- 박정 (2013b). 초등교사의 학생평가전문성에 대한 자기평가 분석. 한국교원교육연구, 30(3), 357-376.
- 박종원, 이수진, 오희균, 조중현 (2000). 중학교 과학 1을 위한 수행평가 도구 개발: 힘과 운동 단원을 중심으로. 한국과학교육학회지, 20(4), 562-571.
- 백순근 (1998). 수행평가의 이론과 실제. 서울: 원미사.
- 백순근 (2000). 수행평가의 원리. 서울: 교육과학사.
- 서미령 (2004). 고등학교 지구과학I 수행평가에서 학급별 수행평가 시행 시일차에 따른 학업성취도. 이화여자대학교 대학원 석사학위논문.
- 서미영 (2007). 초등 교사의 과학 교수내용지식과 지필평가 문항과의 관계 연구. 청주교육대학교 대학원 석사학위논문.
- 서정권 (2008). 고등학교 과학 과목의 수행평가 운영 실태 조사: 실험, 실습법을 중심으로. 경상대학교 대학원 석사학위논문.
- 성을선, 남정희, 최병순 (2000). 중학교 과학수업에서 형성평가의 실제: 구성주의적 관점에서의 형성평가를 중심으로. 한국과학교육학회지, 20(3), 455-467.
- 성태제 (1999). 교육평가 방법의 변화와 결과타당도에 대한 고려. 한국교육학회, 37(1), 197-218.
- 성태제 (2002). 현대교육평가. 서울: 학지사.
- 성태제 (2012). 2012학년도 성취평가제 운영 매뉴얼 중학교용. 연구자료 ORM 2012-18. 서울: 한국교육과정평가원.
- 손정숙 (2011). 중등학교 실험실습 수행평가에 대한 과학 교사들의 인식 조사. 한국교원대학교 대학원 석사학위논문.
- 신옥순 (1997). 교사와 연구자의 협력연구를 위한 기초. 초등교육연구, 11(1), 329-344.

- 신정윤, 양일호 (2017). 과학과 수행평가의 문제점과 개선 방향. 예술인문사회융합 멀티미디어논문지, 7(1), 878-883.
- 신정윤, 유선아, 양일호 (2016). 초등학교 교사들의 과학과 수행평가 실시 현황 및 요구 분석. 청람과학교육연구논총, 22(2), 18-29.
- 신혜영 (2001). 과학 수행평가에 대한 실태와 중·고등학교 과학교사들의 인식조사. 강원대학교 대학원 석사학위논문.
- 심애순 (2008). 중학교 과학과 서술형 평가에 대한 교사와 학생의 인식 조사 및 실태 분석. 한국교원대학교 대학원 석사학위논문.
- 심주옥 (2005). 초등과학의 STS 주제에 대한 수행평가 도구의 개발과 적용. 부산교육대학교 대학원 석사학위논문.
- 심화선 (2006). 과학과 수행평가 실시에 관한 교사와 학생의 의견조사. 단국대학교 대학원 석사학위논문.
- 양일호, 나종철, 임성만, 임재근, 최현동 (2008). Klopfer의 교육 목표 분류 체계에 의한 초등학교 과학자 지필 평가 문항 분석: 5학년 1학기를 중심으로. 초등과학교육, 27(3), 221-232.
- 양찬호 (2015). 초임 과학영재 교사의 수업 전문성 향상을 위한 코칭과 멘토링 활용 전략의 개발과 적용. 서울대학교 대학원 박사학위논문.
- 오현석, 이기영 (2006). 현행 중등학교 과학 교과 지필평가 문항분석. 교육과정평가연구, 9(1), 405-424.
- 원효현 (2014). 성취평가 시행에 따른 수행평가 내실화 방안. 수산해양교육연구, 26(1), 2013-213.
- 윤민화 (2010). 중학교 물질 영역의 성취기준과 평가기준에 근거한 과학교사의 평가 문항 제작 능력 및 인식 조사. 한국교원대학교 대학원 석사학위논문.
- 이규민, 황경현 (2007). 초등학교 과학과 수행평가의 총체적 채점과 분석적 채점 방식에 대한 일반화가능도 분석. 한국아동교육학회지, 16(4), 169-184.
- 이근준, 정진우 (2004). 중등학교 과학실험수업의 탐구수준을 평가하기 위한 도구 개발 및 적용. 한국지구과학회지, 25(7), 507-518.
- 이기영, 안희수 (2005). 중등학교 과학 수행평가의 평가 유형과 채점 방식 및 신뢰도 분석. 한국과학교육학회지, 25(2), 173-183.

- 이선 (2017). 예비교사들의 모의수업 시연 후 자기성찰을 위한 자기점검표 개발. 교육논총, 54(4), 1-18.
- 이연숙 (2006). 교수학적 내용지식(PCK) 및 그 표상(PCKr)의 개념적 정의와 분석도구 개발: 예비 과학교사의 ‘힘과 에너지’ 수업 사례를 중심으로. 서울대학교 대학원 석사학위논문.
- 이은진, 김찬중, 이선경, 장신호, 권홍진, 유은정 (2007). 교사-연구자간 협력적 연수 프로그램에 참여한 과학 교사의 구성주의적 수업에 대한 내면화 과정. 한국과학교육학회지, 27(9), 854-869.
- 이인제, 김범기, 이범홍, 진재관, 박정, 김옥남, 서수현, 김신영, 강순희, 권재술, 김성원, 백성혜, 신동희, 이효녕, 조희형, 차희영 (2004). 과학과 교사의 학생 평가 전문성 신장 모형과 기준. 연구보고 RRE 2004-5-5. 서울: 한국교육과정평가원.
- 이인제, 이범홍, 박정, 진재관, 김옥남, 서수현, 김신영 (2004). 교사의 학생평가 실태 조사 및 전문성 신장에 대한 요구 분석. 연구 보고 RRE-2004-5-1. 서울: 한국교육과정평가원.
- 이재원, 류고운, 강석진, 노태희, 강훈식 (2018). 구성주의적 수행평가 워크숍에 참여한 예비 화학교사의 수행평가 과제 개발 과정에서 고려된 교과교육학 지식(PCK) 구성 요소 및 연계의 특징. 한국과학교육학회지, 38(4), 505-518.
- 이정애 (2003). 중학교 교사의 학생평가에 대한 인식과 평가 전문성 연구. 이화여자대학교 대학원 석사학위논문.
- 이종기, 임낙룡, 황성수 (2004). 과학영재의 과학탐구능력 신장을 위한 서술형 수행평가 문항 개발: 제7차 교육과정 고등학교 생물 2. 과학교육논총, 29, 115-133.
- 이택수 (2003). 창의적 문제해결력 신장을 위한 초등학교 과학과 수행평가 자료의 개발과 적용. 경인교육대학교 대학원 석사학위논문.
- 이현주, 최경희, 남정희 (2000). 형성평가의 피드백 유형이 학생들의 과학 성취와 태도, 교사-학생 상호작용에 미치는 영향. 한국과학교육학회지, 20(3), 479-490.
- 이효정 (2008). 중학교 과학과 수행평가 점수의 일반화가능도 분석. 계명대학교 대

- 학원 석사학위논문.
- 임청환 (2003). 과학 교과교육학 지식의 본질과 발달. 한국지구과학회지, 24(4), 325-249.
- 장수웅 (2002). 기술교육에서 창의적인 문제 해결 수업을 위한 수행평가 도구의 개발. 한국교원대학교 대학원 박사학위논문.
- 장효순, 최병순 (2010). CoRe 개발 과정을 통한 과학교사의 PCK 변화에 관한 사례 연구: 중학교 1학년 「분자의 운동」을 중심으로. 한국과학교육학회지, 30(6), 870-885.
- 전혜진 (2017). 과학적 핵심역량 기반 과정중심 수행평가 적용이 과학개념 및 과학 관련 태도에 미치는 효과. 부산교육대학교 대학원 석사학위논문.
- 정금순, 강훈식 (2011). 초등 과학영재수업에서 코칭의 활용에 대한 사례 연구. 한국과학교육학회지, 31(2), 239-255.
- 정시은 (2007). 중학교 과학교과의 수행평가 운영 실태와 개선 방안. 경상대학교 대학원 석사학위논문.
- 정영란 (2004). 제7차 교육과정에 근거한 준거지향적 수행평가 문항의 개발과 평가: 고등학교 과학 「생식」과 「생물 농축」 단원을 중심으로. 한국과학교육학회지, 24(3), 519-531.
- 정은영 (2018). 고등학교 생명 과학 I 평가에 대한 교사의 인식과 운영 실태. 생물 교육, 46(2), 176-186.
- 정은영, 최원호 (2014). 우리나라 초·중·고등학교 과학과의 평가 실태. 과학교육 연구지, 38(1), 168-181.
- 조성수 (2003). 형성평가에서 교정적 피드백 유형이 학업성취도와 학습태도에 미치는 효과. 서강대학교 대학원 석사학위논문.
- 조희형, 고영자 (2008). 과학교사 교수내용지식(PCK)의 재구성과 적용 방법. 한국과학교육학회지, 28(6), 618-632.
- 조희형, 조영신, 권석민, 박대식, 강영진, 김희경, 고영자 (2006). 중등 과학교사 양성 교육과정과 교수내용지식 연구 동향의 탐색. 교과교육학연구, 10(2), 281-301.
- 조희형, 최경희 (2002). 구성주의와 과학교육. 한국과학교육학회지, 22(4), 820-836.

- 지은림 (2011). 중고교 내신 절대평가 도입 추진방향과 과제. 연구 보고 2011-02-1. 서울: 한국교육개발원.
- 초한무 (2001). 구성주의 관점에서 평가도구. 교육논총, 18, 183-197.
- 최경희, 이현주, 남정희 (2000). 형성평가의 피드백 유형이 학생들의 과학 성취와 태도, 교사-학생 상호작용에 미치는 영향. 한국과학교육학회지, 20(3), 479-490.
- 최명숙 (2001). 구성주의에 대한 교사들의 인식과 수업에의 적용사례. 교육정보미디어연구, 7(1), 5-28.
- 최문수, 정인기 (2002). 웹을 활용한 초등학교 과학과 수행평가 측정 도구의 개발. 과학교육연구, 26, 69-83.
- 최은주 (2011). 중학교 과학과 서술형 평가에 대한 문항의 특성과 교사와 학생 인식 조사. 한국교원대학교 대학원 석사학위논문.
- 최진영, 송경오 (2005). 사회과 교수실제에 영향을 미치는 교사 연수 특징 분석: 교사연수의 내용 및 방법을 중심으로. 초등교육연구, 18(2), 411-430.
- 최혁준 (2013). 컴퓨터 기반 과학 수행평가 문항 개발: 전류계와 전압계 회로 연결 실험. 과학교육연구지, 37(2), 348-358.
- 최현숙, 김중복 (2013). 예비 물리 교사들의 학생평가 문항 제작 수행 수준 조사: 형성 평가를 위한 선다형 물리 개념 검사 문항 분석 중심으로. 과학교육연구지, 37(3), 458-475.
- 최희정, 임채성, 김은진 (2003). 과학 문제해결 과정을 활용한 초등 과학 수행평가 도구의 개발과 적용. 과학교육연구, 28, 139-151.
- 한재영, 윤지현 (2009). 중등 과학에서의 코칭 수업 과정 분석. 과학교육연구지, 33(1), 152-163.
- 홍수연 (2006). 자기 주도적 학습능력 신장을 위한 중학교 과학과 수행평가 과제 개발 및 효과. 연세대학교 대학원 석사학위논문.
- 홍정림, 최은경 (2002). 과학과 수행평가에 관한 중등학생의 인식 및 자아효능감 조사. 한국과학교육학회지, 22(2), 230-239.
- 홍혜정 (2012). 과학 교과에 적용되는 창의·인성 교육과 평가 방법에 대한 교사들의 인식. 이화여자대학교 대학원 석사학위논문.

- Ab Kadir, M. A. (2017). What teacher knowledge matters in effectively developing critical thinkers in the 21st century curriculum?. *Thinking Skills and Creativity*, 23, 79–90.
- Abell, S. K. (2008). Twenty years later: Does pedagogical content knowledge remain a useful idea?. *International Journal of Science Education*, 30(10), 1405–1416.
- Adeyemi, B. A. (2015). The efficacy of authentic assessment and portfolio assessment in the learning of social studies in junior secondary schools in Osun State, Nigeria. *IFE Psychologia: An International Journal*, 23(2), 125–132.
- Alonzo, A. C., & Aschbacher, P. R. (2005). *Factors which contribute to teachers' adoption of science notebooks as a formative assessment tool*. In A. Alonzo (Chair), Teachers' use of formative assessment in science classrooms: Factors in the development of expertise. Symposium conducted at the annual meeting of the American Educational Research Association, Montreal, Canada.
- American Federation of Teachers, National Council on Measurement In Education, & National Educational Association [AFT, NCME, & NEA] (1990). Standards for teacher competence in educational assessment of students. *Educational Measurement: Issues and Practice*, 9(4), 30–32.
- Appleton, K. (2003). *Pathways in professional development in primary science: Extending science PCK*. Paper presented at the 34th Australian Science Education Research Association(ASERA) Conference, Melbourne, VIC, Australia.
- Atjonen, P. (2014). Teachers' views of their assessment practice. *The Curriculum Journal*, 25(2), 238–259.
- Avraamidou, L., & Zembal-Saul, C. (2005). Giving priority to evidence in science teaching: A first-year elementary teacher's specialized practices and knowledge. *Journal of Research in Science Teaching*, 42(9), 965–986.

- Aydin, S., & Boz, Y. (2013). The nature of integration among PCK components: A case study of two experienced chemistry teachers. *Chemistry Education Research and Practice*, 14(4), 615-624.
- Banilower, E., Boyd, S., Pasley, J., & Weiss, I. (2006). *Lessons from a decade of mathematics and science reform: A capstone report for the local systemic change through teacher enhancement initiative*. Chapel Hill, NC: Horizon Research, Inc.
- Barnet, J., & Hodson, D. (2001). Pedagogical context knowledge: Toward a fuller understanding of what good teachers know. *Science Education*, 85(4), 426-453.
- Baron, J. (1991). Performance assessment: Blurring the edge of assessment, curriculum and instruction. In G. Kulm & S. Malcom (Eds.), *Science assessment in the service of reform*. (pp. 247-266). Washington, D.C.: American Association for the Advancement of Science.
- Beck, C. & Kosnik, C. (2006). *Innovations in teacher education: A social constructivist approach*, Albany, NY: State University of New York Press.
- Bell, B., & Cowie, B. (2001). *Formative assessment and science education*. Dordrecht, The Netherlands: Springer.
- Beyer, C. J., & Davis, E. A. (2009). Supporting preservice elementary teachers' critique and adaptation of science lesson plans using educative curriculum materials. *Journal of Science Teacher Education*, 20(6), 517-536.
- Beyer, C. J., & Davis, E. A. (2012). Learning to critique and adapt science curriculum materials: Examining the development of preservice elementary teachers' pedagogical content knowledge. *Science Education*, 96(1), 130-157.
- Black, P. J., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education*, 5(1), 7-73.
- Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2004). Working inside the black box: Assessment for learning in the classroom. *Phi Delta Kappan*,

- 86(1), 8-21.
- Bond-Robinson, J. (2005). Identifying pedagogical content knowledge (PCK) in the chemistry laboratory. *Chemistry Education Research and Practice*, 6(2), 83-103.
- Borgia, L. G. (2001). *Pedagogical content knowledge in project-based teaching*. Unpublished Doctoral Dissertation, University of Illinois at Urbana-Champaign, IL, USA.
- Bransford, J. D., Brown, A. L., & Cocking, R. R. (1999). *How people learn: Brain, mind, experience, and school*. Washington, D.C.: The National Academies Press.
- Brown, G. T. L. (2004). Teachers' conceptions of assessment: Implications for policy and professional development. *Assessment in Education Principles Policy and Practice*, 11(3), 301-318.
- Brown, J. H., & Shavelson, R. J. (1996). *Assessing hands-on science: A teacher's guide to performance assessment*. Thousand Oaks, CA: Corwin Press.
- Brown, M. (2009). The teacher-tool relationship: Theorizing the design and use of curriculum materials. In J. T. Remillard, B. A. Herbel-Eisenmann & G. M. Lloyd (Eds.), *Mathematics teachers at work: Connecting curriculum materials and classroom instruction*. (pp. 17-36). New York: Routledge.
- Carless, D. (2005). Prospects for the implementation of assessment for learning. *Assessment in Education: Principles, Policy & Practice*, 12(1), 39-54.
- Carlsen, W. S. (1999). Domains of teacher knowledge. In J. Gess-Newsome & N. G. Lederman (Eds.), *Examining pedagogical content knowledge*. (pp. 133-144). Boston, MA: Kluwer Academic Publisher.
- Carroll, C., & Mumme, J. (2007). *Learning to lead mathematics professional development*. Thousand Oaks, CA: Corwin & San Francisco, CA: WestEd.
- Carter, K. (1990). Teachers' knowledge and learning to teach. In W. R. Houston (Ed.), *Handbook of research on teacher education*. (pp. 291-310). New York: Macmillan.

- Chinn, C. A., & Malhotra, B. A. (2002). Epistemologically authentic inquiry in schools: A theoretical framework for evaluating inquiry tasks. *Science Education*, 86(2), 175-218.
- Clinch, J. J., & Keselman, H. J. (1982). Parametric alternatives to the analysis of variance. *Journal of Educational Statistics*, 7(3), 207-214.
- Cochran, K., DeRuiter, J., & King, R. (1993). Pedagogical content knowing: An integrative model for teacher preparation. *Journal of Teacher Education*, 44(4), 263-272.
- Dana, T. M., & Nichols, S. (1992). *Assessing the state of science education reform in Florida*. Paper presented at the annual meeting of the National Association for Research in Science Teaching, Boston, MA, USA.
- Davis, D. S., & Neitzel, C. (2011). A self-regulated learning perspective on middle grades classroom assessment. *The Journal of Educational Research*, 104(3), 202-215.
- De Jong, O., van Driel, J. H., & Verloop, N. (2005). Preservice teachers' pedagogical content knowledge of using particle models in teaching chemistry. *Journal of Research in Science Teaching*, 42(8), 947-964.
- Doran, R., Chan, F., & Tamir, P. (1998). *Science educator's guide to assessment*. Arlington, VA: National Science Teachers Association.
- Dorman, J. P., & Knightley, W. M. (2006). Development and validation of an instrument to assess secondary school students' perceptions of assessment tasks. *Educational Studies*, 32(1), 47-58.
- Driver, R. & Bell, B. (1986). Students' thinking and the learning of science: A constructivist view. *School Science Review*, 67, 443-456.
- Earl, L. M. (2012). *Assessment as learning, using classroom assessment to maximize student learning*. Thousand Oaks, CA: Corwin Press.
- Falk, A. (2012). Teachers learning from professional development in elementary science: Reciprocal relations between formative assessment and pedagogical content knowledge. *Science Education*, 96(2), 265-290.

- Fenstermacher, G. D., & Richardson, V. (1993). The elicitation and reconstruction of practical arguments in teaching. *Journal of Curriculum Studies*, 25(2), 101-114.
- Fernandez-Balboa, J. M., & Stiehl, J. (1995). The generic nature of pedagogical content knowledge among college professors. *Teaching and Teacher Education*, 11(3), 293-306.
- Fives, H., & Barnes, N. (2017). Informed and uninformed naïve assessment constructors' strategies for item selection. *Journal of Teacher Education*, 68(1), 85-101.
- Forbes, C. T., & Davis, E. A. (2008). The development of preservice elementary teachers' curricular role identity for science teaching. *Science Education*, 92(5), 909-940.
- Friedrichsen, P., Abell, S. K., Pareja, E. M., Brown, P. L., Lankford, D. M., & Volkman, M. J. (2009). Does teaching experience matter? Examining biology teachers' prior knowledge for teaching in an alternative certification program. *Journal of Research in Science Teaching*, 46(4), 357-383.
- Furtak, E. M., Thompson, J., Braaten, M., Windschitl, M. (2012). Learning progressions to support ambitious teaching practices. In A. C. Alonzo & A. W. Gotwals (Eds.), *Learning progressions in science: Current challenges and future directions*. (pp. 405-433). Berlin, Germany: Springer Science & Business Media.
- Gale, J., Wind, S., Koval, J., Dagosta, J., Ryan, M., & Usselman, M. (2016). Simulation-based performance assessment: An innovative approach to exploring understanding of physical science concepts. *International Journal of Science Education*, 38(14), 2284-2302.
- Garmston, R. J., & Wellman, B. M. (2009). *The adaptive school: A sourcebook for developing collaborative groups* (2nd ed.). Norwood, MA: Christopher-Gordon.
- Geddis, A. N., Onslow, B., Beynon, C., & Oesch, J. (1993). Transforming content

- knowledge: Learning to teach about isotopes. *Science Education*, 77(6), 575-591.
- Gess-Newsome, J. (1999). Pedagogical content knowledge: An introduction and orientation. In J. Gess-Newsome & N. G. Lederman (Eds.), *Examining pedagogical content knowledge*. (pp. 3-20). Boston, MA: Kluwer Academic Publisher.
- Glaser, R. & Pellegrino, J. W. (1987). Aptitudes for learning and cognitive processes. In F. Weinert & R. Kluwe (Eds.), *Metacognition, motivation and understanding*. (pp. 267-288). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Glaser, R. (1984). Education and thinking: The role of Knowledge. *American psychologist*, 39(2), 93-104.
- Gremler, D. D. (2004). The critical incident technique in service research. *Journal of Service Research*, 7(1), 65-89.
- Grossman, P. L. (1990). *The making of a teacher: Teacher knowledge and teacher education*. New York: Teachers College Press.
- Hanna, G. S., & Dettmer, P. A. (2004). *Assessment for effective teaching: Using context-adaptive planning*. Boston, MA: Pearson A&B.
- Hashweh, M. Z. (2005). Teacher pedagogical constructions: A reconfiguration of pedagogical content knowledge. *Teachers and Teaching: Theory and Practice*, 11(3), 273-292.
- Herrenkohl, L. R., Kawasaki, K., & Dewater, L. S. (2010). Inside and outside: Teacher-researcher collaboration. *The New Educator*, 6(1), 74-92.
- Hibbard, K. M., & Baron, J. B. (1990). *Assessing studies working in groups: Lessons from cooperative ad collaborative learning*. Paper presented at the annual meeting of the American Educational Research Association, Boston, MA.
- Hord, S. M., & Sommers, W. A. (2008). *Leading professional learning communities: Voices from research and practice*. Thousand Oaks, CA: Corwin Press; Reston, VA: National Association of Secondary School Principals; & Oxford,

- OH: National Staff Development Council.
- James, M. (2006). Assessment, teaching and theories of learning. In J. Gardner (Ed.), *Assessment and learning*. (pp. 47-60). London, UK: Sage Publications.
- James, M., & Pedder, D. (2006). Beyond method: Assessment and learning practices and values. *The Curriculum Journal*, 17(2), 109-138.
- Johnson, D. W., & Johnson, R. T. (1990). Cooperative learning in achievement. In S. Sharan (Ed.), *Cooperative learning: Theory and research*. (pp. 249-286). New York: Praeger.
- Johnson, S. M., & The Project on the Next Generation of Teachers. (2007). *Finders and keepers: Helping new teachers survive and thrive in our schools*. San Francisco, CA: Jossey-Bass.
- Jonassen, D. H., Peck, K. L., & Wilson, B. G. (1999). *Learning with technology: A constructivist perspective*. Upper Saddle River, NJ: Merrill.
- Kan, A., & Bulut, O. (2014). Crossed random-effect modeling: Examining the effects of teacher experience and rubric use in performance assessments. *Eurasian Journal of Educational Research*, 57, 1-28
- Katehi, L., Pearson, G., & Feder, M. (2009). *Engineering in K-12 education: Understanding the status and improving the prospects*. Washington, D.C.: The National Academies Press.
- Keller, M. M., Neumann, K., & Fischer, H. E. (2017). The impact of physics teachers' pedagogical content knowledge and motivation on students' achievement and interest. *Journal of Research in Science Teaching*, 54(5), 586-614.
- Kind, V. (2009). Pedagogical content knowledge in science education: Perspectives and potential for progress. *Studies in Science Education*, 45(2), 169-204.
- Knowles, M. S., Holton, E. F., & Swanson, R. A. (2000). *The adult learner: The definitive classic in adult education and human resource development* (5th ed). Houston, TX: Gulf Publishing Company.
- Koh, K. H., Tan, C., & Ng, P. T. (2012). Creating thinking schools through

- authentic assessment: The case in Singapore. *Educational Assessment, Evaluation and Accountability*, 24(2), 135-149.
- Krauss, S., Brunner, M., Kunter, M., Baumert, J., Blum, W., Neubrand, M., & Jordan, A. (2008). Pedagogical content knowledge and content knowledge of secondary mathematics teachers. *Journal of Educational Psychology*, 100(3), 716-725.
- Kulm, G., & Malcolm, S. M. (1991). *Science assessment in the service of reform*. Washington, D.C.: American Association for the Advancement of Science.
- Lee, E. (2005). *Conceptualizing pedagogical content knowledge from the perspective of experienced secondary science teachers*. Doctoral Dissertation, The University of Texas at Austin, TX, USA.
- Lee, M. H., Lin, T. J., & Tsai, C. C. (2013). Proving or improving science learning? Understanding high school students' conceptions of science assessment in Taiwan. *Science Education*, 97(2), 244-270.
- Loucks-Horsley, S. & Stiles, K. E. (2001). Professional development designed to change science teaching and learning. In J. Rhoton & P. Bowers (Eds.), *Professional development: Planning and design*. (pp. 13-24). Arlington, VA: NSTA Press.
- Loucks-Horsley, S., Love, N., Stiles, K. E., Mundry, S., & Hewson, P. W. (2003). *Designing professional development for teachers of science and mathematics* (2nd ed.). Thousand Oaks, CA: Corwin Press.
- Loucks-Horsley, S., Stiles, K. E., Mundry, S., Love, N., & Hewson, P. W. (2009). *Designing professional development for teachers of science and mathematics* (3rd ed.). Thousand Oaks, CA: Corwin Press.
- Loughran, J., Berry, A., & Mulhall, P. (2006). *Understanding and developing science teachers' pedagogical content knowledge*. Rotterdam, The Netherlands: Sense Publishers.
- Loughran, J., Milroy, P., Berry, A., Gunstone, R. F., & Mulhall, P. (2001). Documenting science teachers' pedagogical content knowledge through

- PaP-eRs. *Research in Science Education*, 31(2), 289-307.
- Loughran, J., Mulhall, P., & Berry, A. (2004). In search of pedagogical content knowledge in science: Developing ways of articulating and documenting professional practice. *Journal of Research in Science Teaching*, 41(4), 370-391.
- Love, N., Stiles, K. E., Mundry, S., & DiRanna, K. (2008). *The data coach's guide to improving learning for all students: Unleashing the power of collaborative inquiry*. Thousand Oaks, CA: Corwin.
- Magnusson, S., Krajcik, J., & Borko, H. (1999). Nature, sources, and development of pedagogical content knowledge for science teaching. In J. Gess-Newsome & N. G. Lederman (Eds.), *Examining pedagogical content knowledge*. (pp. 95-132). Dordrecht, The Netherlands: Kluwer Academic.
- Marable, M., & Raimondi, S. (2007). Teachers' perceptions of what was most (and least) supportive during their first year of teaching. *Mentoring and Tutoring: Partnership in Learning*, 15(1), 25-37.
- Marks, R. (1990). Pedagogical content knowledge: From a mathematical case to a modified conception. *Journal of Teacher Education*, 41(3), 3-11.
- McMillan, J. H. (2007). *Classroom assessment: Principle and practice for effective standards-based instruction* (6th ed.). Boston, MA: Allyn & Bacon.
- Mehrens, W. (1992). Using performance assessment for accountability purposes. *Educational Measurement: Issues and Practice*, 11(1), 3-9.
- Mintze, J. J., Wandersee, J. H., & Novak, J. D. (1998). *Teaching science for understanding: A human constructivist view*. Cambridge, MA: Academic Press.
- Mishra, P., & Koehler, M. J. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. *Teachers college record*, 108(6), 1017-1054.
- Morine-Dershimer, G., & Kent, T. (1999). The complex nature and sources of teachers' pedagogical knowledge. In J. Gess-Newsome & N. G. Lederman

- (Eds.), *Examining pedagogical content knowledge*. (pp 21-50). Boston, MA: Kluwer Academic Publisher.
- Mundry, S., & Stiles, K. E. (2009). *Professional learning communities for science teaching: Lessons from research and practice*. Arlington, VA: National Science Teachers Association Press.
- Oberg, C. (2010). Guiding classroom instruction through performance assessment. *Journal of Case Studies in Accreditation and Assessment*, 1, 1-11.
- Park, S., & Chen, Y.-C. (2012). Mapping out the integration of the components of pedagogical content knowledge (PCK): Examples from high school biology classrooms. *Journal of Research in Science Teaching*, 49(7), 922-941.
- Park, S., & Oliver, J. S. (2008). Revisiting the conceptualisation of pedagogical content knowledge (PCK): PCK as a conceptual tool to understand teachers as professionals. *Research in Science Education*, 38(3), 261-284.
- Plake, B. S., Impara, J. C., & Fager, J. J. (1993). Assessment competencies of teachers: A national survey. *Educational Measurement*, 12(4), 10-12.
- Popham, W. J. (2009). Unlearned lessons: Six stumbling blocks to our school's success. In C. T. Chauncey (Ed.), *Strategic priorities for school improvement*. (pp. 9-16). Cambridge, MA: Harvard Education Press.
- Remillard, J. T., & Geist, P. K. (2002). Supporting teacher's professional learning by navigating openings in the curriculum. *Journal of Mathematics Teacher Education*, 5(1), 7-34.
- Rhoton, J., & Wojnowski, B. (2006). Building on-going and sustained professional development. In J. Rhoton & P. Shane (Eds.), *Teaching science in the 21st century*. (pp. 113-126). Arlington, VA: NSTA Press.
- Roth, W. M. (1994). Experimenting in a constructivist high school physics laboratory. *Journal of Research in Science Teaching*, 31(2), 197-223.
- Roth, W. M., & Tobin, K. (2002). Redesigning an "urban" teacher education program: An activity theory perspective. *Mind, Culture, and Activity*, 9(2), 108-131.

- Roth, W. M., & Tobin, K. (2005). Coteaching: From praxis to theory. In W. M. Roth & K. Tobin (Eds.), *Teaching together, learning together*. (pp. 5-26). New York: Peter Lang.
- Sato, M., Coffey, J., & Moorthy, S. (2005). Two teachers making assessment for learning their own. *Curriculum Journal*, 16(2), 177-191.
- Schafer, W. D. (1993). Assessment literacy for teachers. *Theory into Practice*, 32(2), 118-126.
- Schwarz, C., Gunckel, K., Smith, E., Covitt, B., Enfield, M., Bae, M., & Tsurusaki, B. (2008). Helping elementary pre-service teachers learn to use science curriculum materials for effective science teaching. *Science Education*, 92(2), 345-377.
- Segall, A. (2004). Revisiting pedagogical content knowledge: The pedagogy of content/the content of pedagogy. *Teaching and Teacher Education*, 20(5), 489-504.
- Shapira-Lishchinsky, O. (2011). Teachers' critical incidents: Ethical dilemmas in teaching practice. *Teaching and Teacher Education*, 27(3), 648-656.
- Shavelson, R. J., Young, D. B., Ayala, C. C., Brandon, P. R., Furtak, E. M., Ruiz-Primo, M. A., Tomita, M. K., & Yin, Y. (2008). On the impact of curriculum-embedded formative assessment on learning: A collaboration between curriculum and assessment developers. *Applied Measurement in Education*, 21(4), 295-314.
- Short, J. B. (2006). Leading professional development for curriculum reform. In J. Rhoton & P. Shane (Eds.), *Teaching science in the 21st century*. (pp. 85-100). Arlington, VA: NSTA Press.
- Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, 15(2), 4-14.
- Shulman, L. S. (1987). Knowledge and teaching: Foundation of the new reform. *Harvard Educational Review*, 57(1), 1-21.
- Smith, D. C., & Neale, D. C. (1989). The construction of subject matter knowledge

- in primary science teaching. *Teaching and Teacher Education*, 5(1), 1-20.
- Stansbury, K., & Zimmerman, J. (2000). *Lifelines to the classroom: Designing support for beginning teachers*. San Francisco, CA: WestED.
- Stanulis, R. N., Little, S., & Wibbens, E. (2012). Intensive mentoring that contributes to change in beginning elementary teachers' learning to lead classroom discussions. *Teaching and Teacher Education*, 28(1), 32-43.
- Stiles, K. E., & Mundry, S. (2002). Professional development and how teachers learn: Developing expert science teachers. In R. W. Bybee (Ed.), *Learning science and the science of learning*. (pp. 137-151). Arlington, VA: National Science Teachers Association Press.
- Tamir, P. (1988). Subject matter and related pedagogical knowledge in teacher education. *Teaching and Teacher Education*, 4(2), 99-110.
- Tan, W. Y. (1982). Sampling distributions and robustness of t-ratio, f-ratio and variance-ratio in 2 samples and ANOVA models with respect to departure from normality. *Communications in Statistics Part A: Theory and Methods*, 11(22), 2485-2511.
- Thomas, L., Deaudelin, C., Desjardins, J., & Dezutter, O. (2011). Elementary teachers' formative evaluation practices in an era of curricular reform in Quebec, Canada. *Assessment in Education: Principles, Policy & Practice*, 18(4), 381-398.
- Thompson, C. L., & Zeuli, J. S. (1999). The frame and the tapestry: Standards-based reform and professional development. In L. Darling-Hammond & G. Sykes (Eds.), *Teaching as the learning profession: Handbook of policy and practice*. (pp. 341-375). San Francisco, CA: Jossey-Bass.
- Thorndike, R. M. (1997). *Measurement and evaluation in psychology and education* (6th ed.). Upper Saddle River, NJ: Prentice-Hall.
- Tierney, R. D. (2006). Changing practices: Influences on classroom assessment. *Assessment in Education: Principles, Policy & Practice*, 13(3), 239-264.

- Tobin, K., Roth, W. M., & Zimmermann, A. (2001). Learning to teach science in urban schools. *Journal of Research in Science Teaching*, 38(8), 941-964.
- Torrance, H. (1995). *Evaluating authentic assessment: Problems and possibilities in new approaches to assessment*. Buckingham, UK: Open University Press.
- van Driel, J. H., Beijard, D., & Verloop, N. (2001). Professional development and reform in science education: The role of teachers' practical knowledge. *Journal of Research in Science Teaching*, 38(2), 137-158.
- van Driel, J. H., Verloop, N., & De Vos, W. (1998). Developing science teachers' pedagogical content knowledge. *Journal of Research in Science Teaching*, 35(6), 673-695.
- Veal, W. R., & MaKinster, J. G. (1999). Pedagogical content knowledge taxonomies. *Electronic Journal of Science Education*, 3(4). Retrieved from <http://ejse.southwestern.edu/article/view/7615/5382/>
- von Glasersfeld, E. (1993). Questions and answers about radical constructivism. In K. Tobin (Ed.), *The practice of constructivism in science education*. (pp. 23-38). Washington, D.C.: Association for the Advancement of Science Press.
- Wang, J.-R., Kao, H.-L., & Lin, S.-W. (2010). Preservice teachers' initial conceptions about assessment of science learning: The coherence with their views of learning science. *Teaching and Teacher Education*, 26(3), 522-529.
- Weiss, I. R., & Pasley, J. D. (2009). *Mathematics and science for a change: How to design, implement, and sustain high-quality professional development*. Portsmouth, NH: Heinemann.
- Wilson, L. (1992). Children as evaluators. *Teaching Pre K-8*, 23(1), 64-67.
- Winterbottom, M., Brindley, S., Taber, K., Fisher, L., Finney, J., & Riga, R. (2008). Conceptions of assessment: Trainee teachers' practice and values. *The Curriculum Journal*, 19(3), 193-213.

<부록 I-1> 과학 평가 실태와 인식 조사 설문지

1. 다음은 선생님의 기본 정보와 평가 관련 교육 경험을 알아보기 위한 문항입니다. 질문의 해당하는 곳에 ○표 하거나 직접 기록해 주십시오.

근무 학교	① 중학교 () ② 고등학교 ()
성별	① 남 () ② 여 ()
교직 경력	()년 ()개월
학력	① 학사 () ② 석사(수료 포함) () ③ 박사(수료 포함) ()
전공	() 전공
과학 평가 관련 연수 경험	· 교원 임용 후, 대학이나 연수기관 등에서 과학 평가 관련 연수를 받은 경험이 있습니까? ① 예 () ② 아니오 ()
	· 연수받은 경험이 있다면, 최근 3년간 연수 시간은 대략 몇 시간입니까? 약 () 시간
	· 연수받은 경험이 있다면, 어떤 형태의 연수를 받으셨습니까? (복수응답 가능) ① 교육청 연수 ② 학교전달연수 ③ 인터넷 원격연수 ④ 평가컨설팅 연수 ⑤ 기타()

2. 다음은 현재 시행되는 과학 평가에 대해 알아보기 위한 문항입니다. 선생님께서 과학 과목에서 실시하셨던 대표적인 평가 사례에 대해, 예시와 같이 **평가 목적**과 **평가 방법**을 포함하여 자세히 상황을 묘사하여 주십시오.

(예시)

과학 영재교육원에 추천할 학생들을 선발하기 위하여, 교육과정 범위 내에서 문제를 출제하여 학생들의 과학 탐구력을 평가하였다.

사례 1	
사례 2	
사례 3	
사례 4	
사례 5	

3. 다음은 과학 과목의 평가에서 일어날 수 있는 상황에 대한 선생님의 의견을 묻는 문항입니다. 아래의 예시 상황을 읽고, 각 상황에서 선생님께서 어떻게 평가를 하실지 평가 목적과 평가 방법을 포함하여 자세히 적어 주십시오.

3-1. 예시 상황 1

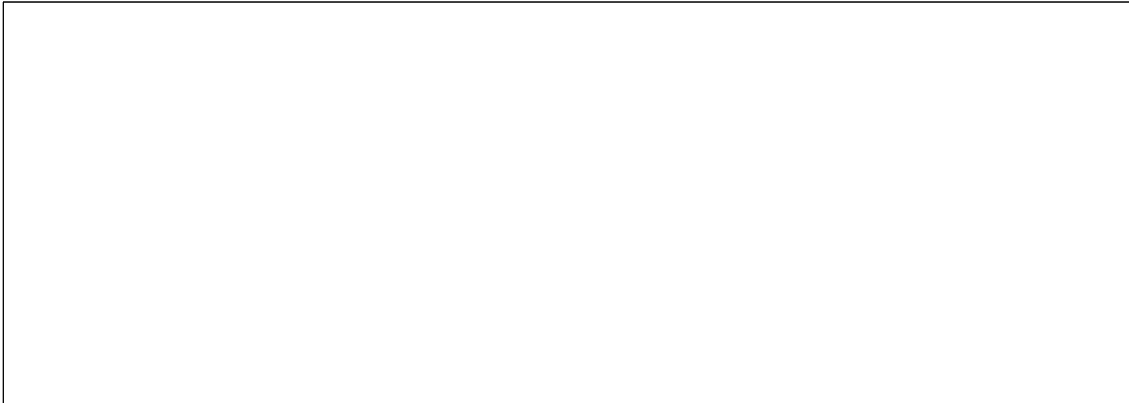
선생님께서서는 과학실에서 실험 수업을 합니다. 실험 주제는 혼합물의 분리이고, 학생들은 조별로 선생님이 만들어놓은 여러 가지 물질의 혼합물을 준비된 실험 도구를 사용하여 분리하는 활동을 할 예정입니다.

3-2. 예시 상황 2

선생님께서서는 과학 수업에서 사회적 이슈에 대한 토의 수업을 합니다. 토의 주제는 ‘지구 온난화 문제의 해결 방안’입니다. 학생들은 소집단별로 토론을 실시하고, 그 결과를 전체 학급 토론에서 발표하여 결론을 내릴 예정입니다.

3-3. 예시 상황 3

학생들은 한 학기 동안 자유탐구를 수행할 예정입니다. 학생들은 먼저 소집단별로 자유 탐구 주제를 논의할 것이고, 선생님은 학생들의 주제 선정 과정을 점검하여 조언을 할 것입니다. 이후 학생들은 소집단별로 탐구 활동을 수행할 것입니다.



4. 다음은 과학 평가에 대한 선생님의 인식을 묻는 객관식 문항입니다. 보기 중 선생님의 생각과 가장 일치하는 것을 하나만 선택하여 ○표를 해 주십시오.

문항	질문내용	전혀 그렇지 않다	그렇지 않다	보통이다	그렇다	매우 그렇다
1	다음 과학 수업을 계획할 때 학생들의 이해 정도에 대한 정보가 필요한데, 과학 평가 결과는 이 정보를 제공한다.	1	2	3	4	5
2	다음 과학 수업을 계획할 때, 학생들이 지난 과학 수업 시간에 잘 따라오지 못했더라도 정해진 교육과정의 순서를 따른다.	1	2	3	4	5
3	과학 평가를 할 때 학생들이 교육과정에 제시된 요소들을 알고, 이해하며, 할 수 있는지에 중점을 둔다.	1	2	3	4	5
4	학생들에게 주는 피드백은 학생들을 향상시키는데 도움이 된다.	1	2	3	4	5
5	나는 학생들에게 다른 학생에 비해 과학을 얼마나 잘 했는지 알려준다.	1	2	3	4	5
6	나는 학생들에게 자신의 과학 학습 목표를 스스로 결정할 수 있는 기회를 준다.	1	2	3	4	5
7	내가 질문을 사용하는 주요한 목적은, 학생들이 암기하고 있는 과학 지식을 이끌어내기 위함이다.	1	2	3	4	5
8	내가 실시하는 과학 평가가 가장 의미 있는 평가라고 생각한다.	1	2	3	4	5
9	내가 실시하는 과학 평가는 학생들이 자율적으로 과학 학습을 할 수 있도록 도와준다.	1	2	3	4	5
10	학생들에게 자신의 이전 과학 성적에 비해 얼마나 발전했는지 알려준다.	1	2	3	4	5
11	학생들이 과학 학습 목표를 이해할 수 있도록 같이 논의한다.	1	2	3	4	5
12	학생들의 과학 활동에 대한 평가는 주로 점수나 등급으로 매긴다.	1	2	3	4	5
13	나는 학생들이 자신의 과학 활동에 대해 스스로 평가해보도록 안내해준다.	1	2	3	4	5
14	나는 학생들의 장점을 파악하여, 그 장점을 더 계발하도록 조언해 준다.	1	2	3	4	5
15	학생들이 과학 공부를 할 때 생기는 문제점을 학생들이 스스로 파악할 수 있도록 도와준다.	1	2	3	4	5

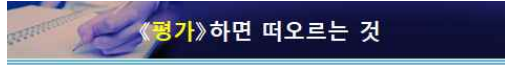
문 항	질 문 내 용	전혀 그렇지 않다	그렇지 않다	보통 이다	그렇다	매우 그렇다
16	실수를 했을 때, 학생들이 이것을 귀중한 학습 기회로 여길 수 있도록 격려해준다.	1	2	3	4	5
17	학생들이 자신에게 가장 효과적인 과학 학습 방법이 무엇인지 스스로 생각해 보도록 도와준다.	1	2	3	4	5
18	내가 질문을 사용하는 주요한 목적은, 학생들이 생각하는 이유와 설명을 이끌어내기 위함이다.	1	2	3	4	5
19	나는 학생들이 다른 학생들의 과학 활동을 평가해보도록 안내해준다.	1	2	3	4	5
20	학생들의 실수를 통해 학생들이 어떻게 생각하는지 알 수 있기 때문에 가치가 있다.	1	2	3	4	5
21	학생들이 과학 수업의 목표를 이해할 수 있도록 도와준다.	1	2	3	4	5
22	학생들의 과학 학업 평가는 주로 서술형으로 한다.	1	2	3	4	5
23	과학 학습 목표를 결정할 때 주로 정해진 교육과정을 따른다.	1	2	3	4	5
24	나는 학생들이 자신의 과학 학습에 대해 평가해보도록 안내해준다.	1	2	3	4	5
25	과학 평가를 할 때 학생들이 무엇을 알고, 이해하며, 그리고 무엇을 할 수 있는지에 중점을 둔다.	1	2	3	4	5
26	학생들이 과학 공부를 할 때 다음 단계를 계획할 수 있도록 도와준다.	1	2	3	4	5
27	학생들의 과학 학습을 평가할 때, 학생이 얼마나 노력했는지가 중요하다.	1	2	3	4	5
28	과학 평가 기준을 학생들이 이해할 수 있도록 같이 논의한다.	1	2	3	4	5
29	나는 학생들에게 다른 학생의 과학 활동을 평가할 수 있는 기회를 준다.	1	2	3	4	5
30	나는 과학을 공부하는 방법에 관해서 학생들과 정기적으로 논의한다.	1	2	3	4	5

구체화 활동지

단원명	분자의 운동과 상태 변화	소단원명	
과제명			
세부 계획안	<p>평가 계획이 드러날 수 있도록 <u>수행평가의 전체적인 상황과 흐름을 묘사</u>합니다. 이 과정에서 <u>교사는 어떤 활동을 하게 되는지, 학생은 무엇을 어떻게 수행하게 되는지</u> 설명합니다. <u>학생들에게 나누어 줄 안내문, 활동지, 자료의 양식</u>을 가능한 포함하여야 합니다.</p>		
평가 준거	<p>학생이 제출한 산출물이나 교사가 수집한 정보를 <u>어떤 준거로 평가할 것인지</u> <u>채점 기준</u> 등을 기록합니다. 선택한 평가 유형에 따라 필요한 <u>성취수준, 체크리스트, 평가 문항</u> 등을 가능한 구체적으로 제시하여야 합니다.</p>		
평가 결과 활용 방법	<p>수행평가를 통해 얻은 여러 가지 <u>평가 정보나 평가 결과를 어떻게 활용할지</u> 간략히 작성합니다.</p>		

※ B4 용지 크기에 맞게 인쇄하여 제공함.

<부록 III-1> 수행평가 전문성 향상 프로그램의 PPT 자료



[사진] 긴장감이 흐르는
수능 3일 전 고3 교실

[사진] 성적 통지표



사실...
평가(assessment) ≠ 시험(test)
평가(assessment) ≠ 줄 세우기 목적

오직 숫자(점수)로만 나의 수준을
평가한다.

음악, 체육, 미술도 지필 평가가 50%
비중을 차지한다.

지엽적인 내용들이 평가 문제로
나온다.

틀리도록 일부러 꼬아내는 문제들이
있다.

점수가 그 과목의 실력을 의미하지
않는 것 같다.

정해진 정답만을 알아내야 하는 평가가
실시된다.

무엇이
문제였을까?



❖교사는...

- 학생들이 알아야 할 것을 결정 → 필요한 정보를 학습자에게 전달하기 위한 전략을 설계 → 이 과정이 성공적으로 이루어졌는지 평가함.

❖학생은...

- 교사의 '평가 대상'
- 구조화된 실체, 속성, 관계를 익히는 것이 중요함.
- 구조화된 지식에서 개인 경험은 중요한 역할을 하지 못함.

구성주의란?

- 지식이 **개인의 인식과 경험**에 의하여 형성된다고 보는 인식론.
- 학생은 외부 현상을 적극적으로 해석하고 의미를 부여하여 **스스로/인지 구조를 구성**하는 능동적인 존재.

구성주의 평가란?

- **개인에게 유의미**하고, **실천적 맥락**을 존중하며, **개인별 차이**를 존중하는 질적인 평가
- 지식의 구성 과정에 관한 평가, 다양한 관점과 양식으로 이루어지는 평가

구성주의 평가의 목적

학습을 위한 평가(Assessment for learning) : 평가를 통하여 교사와 학생 개인의 교수·학습 개선에 도움을 주기 위함



구성주의 교수학습 및 평가 방법

❖ 학생의 개념 발달과 개념 변화를 유도

- 사실, 사건, 현상 등에 관한 학생들의 선행 개념을 확인.
- 학생들의 인지적 갈등을 유발하는 상황(학생들의 가정 또는 일상적 생각과 차이가 나는 내용) 제시.
- 학생의 생각이나 관념을 드러나게 함.
- 학생들이 오개념을 가지고 있을 경우, 그것의 타당성을 설명하고 대안적 개념을 제시.
- 일상생활 소재를 활용, 큰 개념 또는 대주제 중심으로 작성.

구성주의 교수학습 및 평가 방법

❖ 교사의 역할

- 학생의 자율성과 주도권을 인정.
- 조작적이고 상호작용적인 물리적 재료를 활용.
- 깊이 있는 질문을 하고 서로 묻게 함.
- 유추와 비유를 사용.

❖ 일반적인 전략

- 문제해결전략(과제, 모듈, 토의)
 - 순환학습을 통하여 학생들의 자연적 흥미를 함양.
 - 주어진 문제를 이해하고 해결하는 과정과 방법을 평가.
- (표준 절차를 따르게 하거나 제한된 점답만을 요구하는 평가는 지양)

평가 시기에 따른 분류

학습 전

학생들의 선수 학습 정도, 적성, 흥미 등 파악

진단 평가

학습 중

학습이 의도한 교육 목표에 부합하게 잘 진행되고 있는지 점검하여 학생에게 피드백 제공 (ex. 쪽지 시험, 퀴즈)

형성 평가

학습 후

교과 학습이 종료된 이후 교육 목표의 달성 정도를 종합적으로 판정하는 평가 (ex. 학기말 시험)

총괄 평가

평가 시기에 따른 분류

	진단 평가	형성 평가	총괄 평가
평가 시기	· 교수학습 시작 전	· 교수학습 진행 도중	· 교수학습 완료 후
평가 목적	· 적절한 교수 투입	· 교수학습진행의 적절성 확인, (교수법) 개선	· 교육목표 달성 교육프로그램 선택 및 결정 교육의 책임성 부여
평가 방법	· 비형식, 형식적 평가	· 수시평가 비형식, 형식적 평가	· 형식적 평가
평가 내용	· (학습 준비도)	· 학습 진도	· 학습 성취도
평가 유형	· 준거지향평가	· 준거지향평가	· 표준 혹은 준거지향평가
평가 문항의 난이도	· 준거에 부합하는 난이도의 문항	· 준거에 부합하는 난이도의 문항	· 준거지향평가: (다양한 난이도) 준거에 부합하는 난이도의 문항
평가 결과의 이용	· 보충학습 계획	· 교수학습 지도 계획 조정	· 성적 결정

평가 방법에 따른 분류

❖ 수행평가

- 논술형(글쓰기) 평가
- 실험 실습 평가
- 보고서 평가
- 관찰 평가
- 포트폴리오 평가
- 자기 평가 및 동료 평가
- 구술 / 면접/ 토론 평가

평가 시기와 평가 방법

평가 시기에 따라 적절한 평가 방법 선택!



구성주의와 수행평가

수행평가 : 평가자가 학습자들의 학습 과제 수행 과정 및 결과를 직접 관찰하고, 그 관찰-결과를 전문적으로 판단하는 평가 방식

실제 생활에서 보다 의미있고, 중요하고, 유용한 과제

단순히 답을 선택하는 것이 아니라, 학생 스스로 답을 구성하는 것, 산출물을 만들어내는 것, 행동으로 드러나는 것 등을 포함

학습자의 수행 과정과 결과를 평가자가 읽고, 보고, 듣고, 느끼는 모든 활동을 포함

수행평가의 지향 = 구성주의 평가의 특성
수행평가는 구성주의적 평가를 구현하는 바람직한 평가 체계가 될 수 있음

여러 가지 수행평가 방법

- 논술형(글쓰기) 평가
- 실험 실습 평가
- 보고서 평가
- 관찰 평가
- 포트폴리오 평가
- 자기평가
- 동료평가
- 토의, 토론 평가
- 구술, 면접 평가

어떻게 하면 구성주의적으로 활용할 수 있을까?

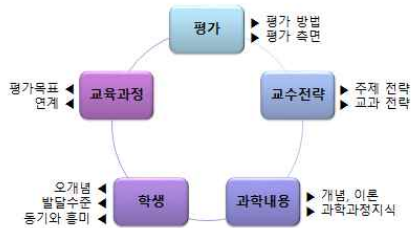
수행평가 과제 개발 및 시행절차

좋은 수행평가 과제를 개발하기 위해 생각해 볼 질문들

- 1) 전통적 평가에서 다루지 못하는 기본적 과제, 성취도, 기질, 중요한 전문적 지식 등이 무엇인가?
- 2) 학생들이 직면할, 또는 숙달해야 할 핵심적 수행, 역할, 상황이 무엇인가?
- 3) 실제의 수행 판단에 이용할 독특한 준거가 무엇인가?
- 4) 평가과제의 숙달이 실제로 무엇을 의미하는가?
- 5) 다른 사람의 도움, 자료의 수집, 시간, 보조자, 기준에 관한 사전 지식 등 검사에 필요한 조건들이 실제적인가?
- 6) 과제가 학생들의 능력을 충분히 평가할 수 있을 만큼 포괄적이고 심도 있는 것인가?

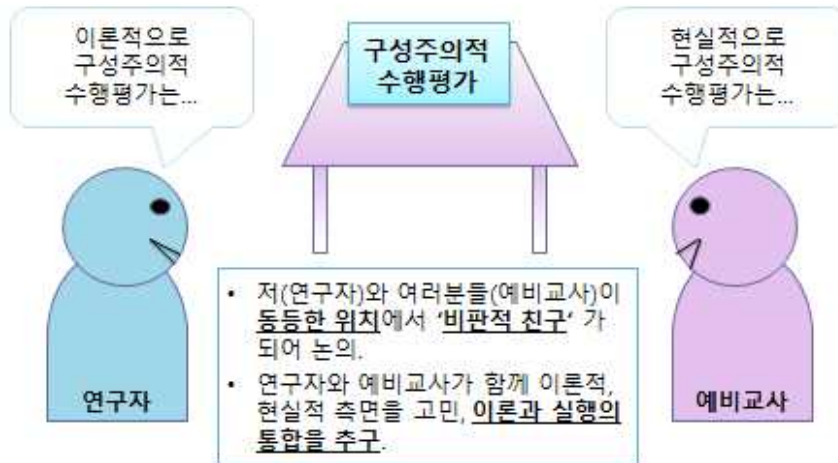
수행평가 과제 개발 및 시행절차

좋은 수행평가 과제 개발을 위한 요소



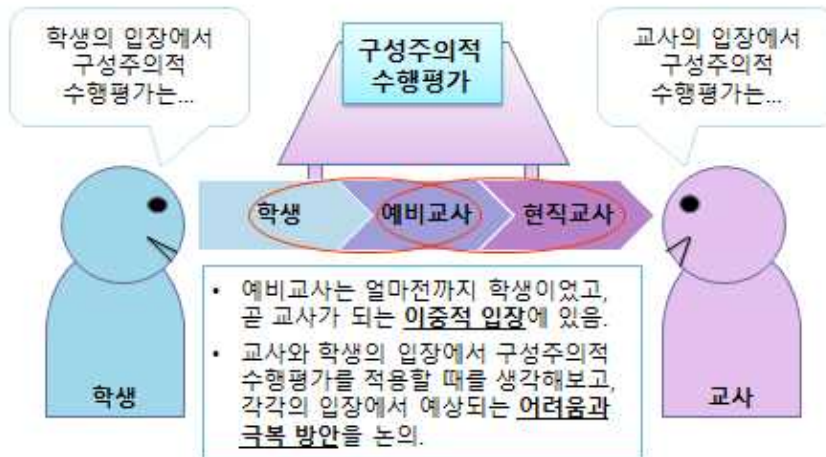
연구자-예비교사 협력적 논의

❖ 협력적 논의(collaborative discussion)란?



교사-학생의 이중적 입장에서 논의

❖ 이중적 입장(dual situation)에서의 논의란?



‘분자 운동과 상태변화’ 단원과 관련한 학생의 주요 오개념 유형

1. 확산

- ① 확산과 화학 변화를 혼동한다. (확산하는 분자들이 결합한다.) → 떨어져 있던 분자들이 확산하면서 접촉하면 결합하거나, 다른 분자로 변화한다는 생각이다.
- ② 확산하는 분자들이 불균일하게 분포한다.
→ 무거운 기체 분자가 가라앉는다는 형태로 자주 나타나며, 상자의 중앙에만 기체 분자들이 분포하거나 상자 한쪽에 모든 기체 분자들이 분포한다는 생각도 나타난다.
- ③ 공기의 운동에 따라 냄새를 내는 물질 분자가 퍼져나간다. (물의 운동에 따라 잉크 분자가 퍼져나간다)
→ 공기 또는 물이 확산에 중요한 역할을 하고, 냄새 분자 또는 잉크 분자는 수동적인 역할을 한다는 생각이다. 모든 분자가 스스로 운동하여 확산한다는 것을 설명하는 데 어려움을 겪는 경우이다.
- ④ 무거운 분자일수록 확산 속도가 더 빠르다.
→ 분자의 질량과 관계없이 확산 속도가 일정하다고 생각하는 예도 있다. 확산 속도와 분자량 사이의 관계를 배우지 않았거나, 분자 운동 에너지와 분자의 운동 속도를 혼동하여 생기는 오개념이다.

2. 기체

1) 압력과 부피의 관계

- ① 운동 속도가 변화한다. → 외부 압력이 증가하여 부피가 작아지면 충돌 횟수가 증가하여 분자 운동이 활발해진거나, 반대로 외부 압력이 증가하면 공간이 좁아져 활동이 제약을 받아 운동 속도가 감소한다는 생각이다.

2) 온도와 부피의 관계

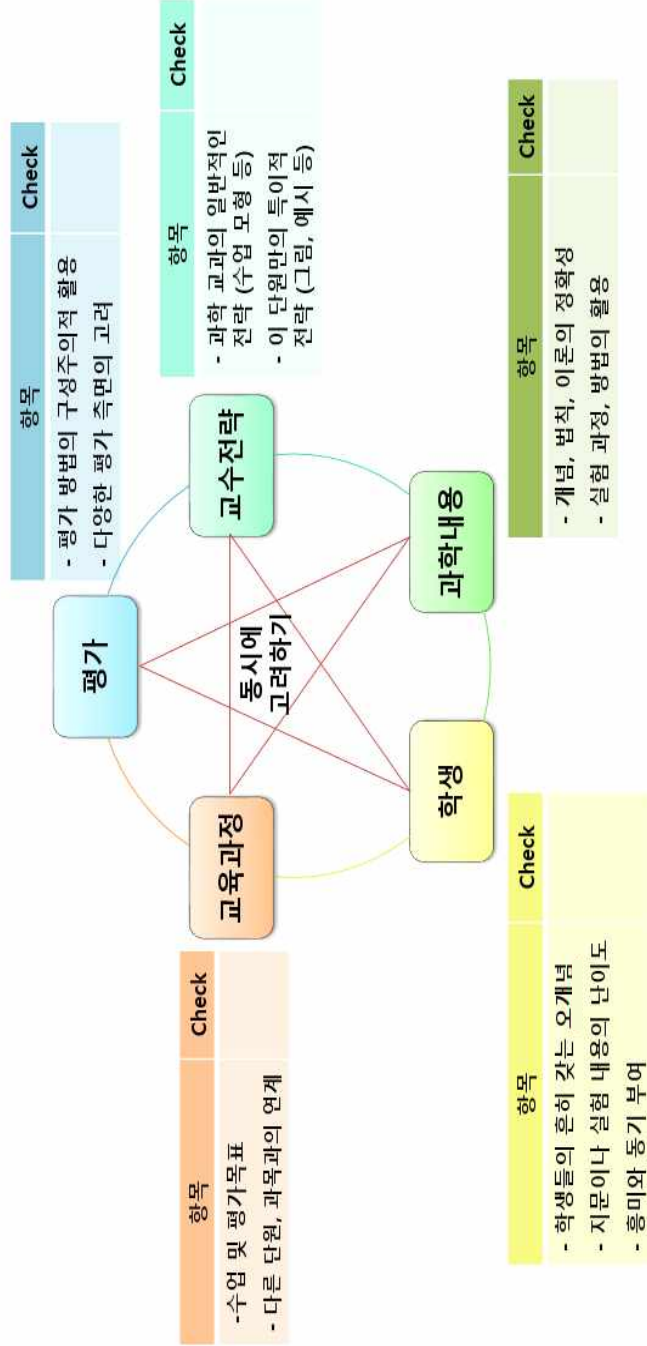
- ① 분자들이 불균일하게 분포한다. → 가열하기 전에는 분자 운동이 활발하지 않아서 기체 분자들이 주어진 공간에 균일하게 분포하지 못하고 바닥에서 잘 움직이지 못한다는 생각이 주로 나타난다.

3. 상태 변화

- ① 상태 변화 시 입자의 크기나 수가 변한다.
→ 기체 상태로 갈수록 부피가 커지기 때문에 같은 공간에 존재하는 입자의 수가 적어진다거나, 입자의 부피를 기체의 부피와 동일한 것으로 간주하여 입자의 크기가 커진다는 생각이다.
- ② 상태 변화 시 물질의 성질이 변한다. → 예를 들어, ‘고체가 단단해서 더 무겁다’ 고 생각하는 학생들이 있다.

좋은 수행평가 과제 개발을 위한 점검표

- 이 점검표는 지난주 워크숍에서 소개한 '좋은 수행평가 과제 개발을 위해 고려할 요소'를 바탕으로 제작하였습니다.
- 참고용 도구로 자유롭게 활용하세요.



ABSTRACT

Development and Application of the Program to Improve Pre-Service Chemistry Teachers' Performance Assessment Professionalism

Jaewon Lee
Chemistry Major
Department of Science Education
The Graduate School
Seoul National University

In order for a national curriculum that emphasizes the student-centered, level-based instruction to be settled appropriately in the actual school site, the assessment should also be changed correspondingly. Performance assessment can judge the level of individual students' achievement reasonably compared to paper-based assessment which estimates the relative rank of students. In addition, performance assessment is an ideal assessment system that can implement constructivist assessment, which is a paradigm of current science education. As the achievement standards-based assessment system has recently been introduced, improving teacher's performance assessment professionalism becomes more important. However, many teachers have difficulty in planning and implementing performance assessment constructively. Therefore, seeking effective strategies to improve science teachers' performance assessment professionalism is needed. In the development of performance assessment, various pedagogical content

knowledge (PCK) components such as science curriculum, science instructional strategies, students, subject matter, and orientation to science assessment are utilized with the knowledge of science assessment. Therefore, it can be useful to develop performance assessment professionalism through programs which provide assessment-related activities can use their PCK effectively.

As basic research for the development of the program to improve science teachers' performance assessment professionalism, the current status and problems of the science teachers' assessment professionalism were investigated, and the necessity of education at the level of pre-service teacher was derived (Study I). Therefore, the characteristics of the processes of developing performance assessment tasks considered by pre-service chemistry teachers were investigated in terms of the PCK components and their integrations (Study II). After analyzing the results of studies I and II to construct strategies and activities, the program to improve pre-service chemistry teachers' performance assessment professionalism was developed based on the principles of effective professional development. After applying this program to pre-service chemistry teachers, the characteristics of the PCK components and their integrations in the processes of developing performance assessment tasks were investigated and compared with the results of study II (Study III).

In study I, actual and preferred cases of assessment of 92 secondary school science teachers were investigated with the focus on the purpose and the method of assessment. The use of formative assessment according to the background information of the teachers was also investigated. The analysis of the actually implemented assessments in their responses revealed that most science teachers tended to aim at traditional summative assessment purposes, and that multiple-choice was the most prevailing assessment method followed by experiment, report, and essay. For hypothetical performance assessment situations, they exhibited their preferences for various types of constructivist assessment methods, whereas their purposes of the assessment still remained to be traditional.

In the experiment and open inquiry situations, those who had a relatively constructivist perspective toward assessment showed a statistically significant preference for using formative assessment than their counterparts ($p=.018$, $p=.023$). Through the study I, the importance of enhancing the perception of constructivist assessment of teachers was derived. However, it is difficult to achieve in a short period of time. In order to spread the constructivist performance assessment, education should begin at the pre-service teacher level, which can provide systematic and long-term education. As a result, the goal of this study could be embodied to improve pre-service chemistry teachers' performance assessment professionalism.

In study II, the characteristics of the PCK components and their integrations of pre-service teachers considered in the processes of developing performance assessment tasks were investigated. Eight pre-service chemistry teachers participated in this study. After brief workshop for performance assessment, they developed the performance assessment tasks through individual task development process and small group discussion process. Individual task development process consisted of the planning step and the embodying step. Their activities were recorded and videotaped. Semi-structured interviews were also conducted. It was found that the PCK component considered in individual task development process was concentrated on the knowledge of science assessment. The proportions of the knowledge of science curriculum, science instructional strategies and students were found to be relatively low. The feedback on students' performance was not considered at all. The orientation to traditional assessment which emphasized the objectivity of assessment, was stronger in the embodying step rather than in the planning step. During the planning step, the integrations among the PCK components mainly appeared with the knowledge of science instructional strategies. However, they hardly appeared in the embodying step and small group discussion process. The main characteristics of the PCK in small group discussion process were similar with those in the embodying step. Some PCK components

were considered more in small group discussion process, but most discussions were in the traditional perspectives. These results suggested that providing simple opportunities for developing performance assessment tasks had a significant limitation to improve pre-service teachers' performance assessment professionalism, and more substantive assistance should be provided in the aspect of utilization of the PCK components.

In study III, the program to improve pre-service chemistry teachers' performance assessment professionalism was developed by combining the results of studies I and II, based on the principles of effective professional development and professional development design framework. Considering the results of study II, the program provided the opportunity to utilize and develop PCK for them. Checklist and student misconception materials were also provided. Based on the results of study I, collaborative discussions between the researcher and pre-service teachers and discussions in the dual situation as a student and a teacher were conducted simultaneously to integrate theory and practice in constructivist performance assessment. Eight pre-service teachers who participated in the program for 3 weeks developed the performance assessment tasks in the same way as study II. Methods of data collection and analysis were also the same as study II to compare the results of two studies. In addition, the think-aloud method was used to investigate their developing processes of the performance assessment tasks. In individual task development process, the analysis of the results revealed that the PCK component considered in the planning step showed similar frequencies for the knowledge of science assessment, science instructional strategies, students, and subject matter components. In the embodying step, the knowledge of science assessment component was the most frequent. In small group discussion process, the frequencies for the knowledge of science assessment, science instructional strategies, and students components were similar. The integrations among the PCK components were found to be centered on the knowledge of science instructional strategies component in the planning step and

the knowledge of students component in the embodying step. In small group discussion process, the knowledge of students was actively considered and provided an opportunity to integrate various PCK components. They also tried to overcome the gap between theory and practice in constructivist performance assessment toward a constructivist way. However, the knowledge of science curriculum component and the subcomponent of science process skill were rarely used. The knowledge of science curriculum and subject matter components were also rarely integrated with other PCK components.

On the basis of the results, the role of researchers participating in collaborative discussions and specific ways for effective small group discussions were discussed in order to complement the program developed in this study. Also, specific plans for applying this program to the courses of the college of education and the need for a support system to connect a researcher and pre-service teachers were suggested in order to make effective use of the program.

Key words: performance assessment, constructivist assessment, pedagogical content knowledge, principles of effective professional development, program development, pre-service chemistry teacher
student number: 2013-21428