

ORIGINAL ARTICLE

# 초등 예비교사가 평가 문항 제작 시 주목하는 것은 무엇인가? : 날씨와 우리 생활 단원을 중심으로

임성만<sup>1</sup> · 김성운<sup>2\*</sup>

(<sup>1</sup>한국교원대학교 교수, <sup>2</sup>충북탄소중립지원센터 팀장)

## What Pre-service Elementary School Teachers Focus on When Developing Assessment Items: Focusing on the Unit ‘Weather and Our Lives’

Sung-Man Lim<sup>1</sup> · Seong-Un Kim<sup>2\*</sup>

(<sup>1</sup>Korea National University of Education, <sup>2</sup>Chungbuk Carbon Neutrality Support Center)

### ABSTRACT

Summative assessment provides information on how well students have achieved learning objectives, making the development of high-quality assessment items essential for accurate evaluation. This is one of the competencies that teachers must possess. This study aims to analyze summative assessment items created by pre-service elementary teachers, examining their intentions and the difficulties encountered in the item development process. The study involved 45 second-year students enrolled in an elementary teacher training university. They were grouped into teams of three and tasked with developing ten items, documenting the purpose of each item, the answer key, and the challenges faced during item creation. The collected summative assessment items were analyzed using a two-dimensional purpose classification table that includes Klopfer's taxonomy of educational objectives. The intentions behind the summative assessments and the difficulties faced during item development were inductively organized and analyzed through qualitative data analysis. The results revealed that pre-service elementary teachers adequately reflected scientific content elements but did not evenly cover assessment domains. The most challenging aspect for them was adjusting the difficulty level. Although they considered most factors that should be taken into account during item development, these considerations were not reflected in the actual items. These findings suggest that knowledge and experience are crucial in developing summative assessment items, and systematic lectures are necessary for pre-service elementary teachers.

**Key words** : elementary pre-service teacher, weather and our life, summative assessment, assessment item development

### I. 서론

평가는 단순히 학생들의 학업 성취를 측정하는 것을 넘어서 학습 결손의 파악, 학업 진전 상황 파악, 교

수 방법의 개선, 평가 방법의 개선, 평가 도구의 개선, 진로 지도 등과 같은 것에 활용할 수 있는 다양한 목적을 반영한 것이라고 할 수 있다(손연아 외, 2018). 이러한 평가의 다면적 기능은 학교 교육의 질적 향상을

Received 24 July, 2024; Revised 12 August, 2024; Accepted 22 August, 2024

\*Corresponding author : Seong-Un Kim, 250, Taeseongtabyeon-ro, Gangnae-myeon, Heungdeok-gu, Cheongju-si, Chungcheongbuk-do, Republic of Korea

E-mail : aaul@naver.com

© The Korean Society of Earth Sciences Education. All rights reserved.

This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License(<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

위해 필수적인 것으로 볼 수 있다. 그러나 학교 현장에서는 이러한 평가의 본질적인 측면보다는 학생들을 성적에 따라 서열화하는 평가의 부분적 기능에 집중하고 있다(교육부, 2014). 더불어 학교 현장의 교사들은 교수학습 방법에 비해 평가에 대한 관심이 상대적으로 낮아 평가를 교육에 충분히 활용하지 못하는 것으로 보여진다(정은영과 최원호, 2014). 이러한 측면에서 학교 현장의 교사가 평가의 본질적인 목적을 이해하는 것은 매우 중요하다.

특히 다른 교과와 달리 과학 평가는 과학학습의 관점에서 학생들이 지닌 지식뿐만 아니라 탐구, 태도 등에 관한 영역까지 고려하여 평가해야 하므로(교육부, 2015) 과학 평가의 본질적인 목적을 이해하는 것은 더욱 중요하다고 할 수 있다. 이러한 평가의 목적을 반영한 평가 문항을 개발할 때에는 평가의 목적과 더불어 평가하고자 하는 능력이나 내용을 고려해야 한다(교육부, 2015). 학교 현장에서는 이러한 평가해야 할 능력과 내용을 구성하기 위해 일반적으로 이원목적분류표를 작성하여 평가 문항을 개발한다. 이러한 부분에서도 과학 교과는 일반 교과와 달리 이원목적분류표를 작성할 때 일반 교과에서는 Bloom의 교육 목표 분류를 주로 이용하나(Krathwohl, 2002), 과학과에서는 과학의 특수성을 반영한 Klopfer의 교육 목표 분류 체계를 활용한다(권계술, 1984; 김상달 외, 2005).

평가는 앞서 서술한 바와 같이 교과에서 요구하는 다양한 목표의 도달 정도를 측정해야 하므로 평가 문항 개발이 필연적으로 요구된다. 평가 문항에는 평가자의 의도가 반영되어 있다. 이러한 점에서 평가자가 개발한 문항을 분석하는 것은 학교 현장의 평가 상황을 파악하는데 중요한 척도가 될 수 있다. 평가는 교사가 지녀야 할 전문성 중 하나이다. 그러나 평가와 관련된 교사의 인식 및 전문성은 그리 높지 않으므로(김신영, 2012; 정은영과 최원호, 2014; 박현주, 2016), 평가에 대한 교사의 전문성 신장이 필요하다(Abell & Siegle, 2011; Carless, 2015; DeLuca & Klinger, 2010; Engelsen & Smith, 2014). 이와 더불어 실제 예비교사들의 평가 문항을 분석한 김정훈과 박영신(2012)의 연구와 최현숙과 김중복(2013)의 연구에서는 예비교사들이 개발한 문항의 타당도, 오류, 어려운 지문 등의 문제점을 지적하며 예비교사의 문제 개발 경험과 능력 향상이 중요하다고 지적하고 있다.

이러한 평가의 관점에서 본 연구는 예비교사들의

평가 문항 개발과 관련된 실질적인 연구를 진행하였다. 예비교사들로 하여금 지정된 단원과 관련한 총괄 평가 문항을 개발하게 한 후 이 문항을 앞서 언급한 Klopfer의 교육 목표 분류 체계가 포함된 이원목적분류표에 따라 평가 문항을 분석하였다. 교사가 개발한 평가 문항을 분석하여 평가 의도를 파악하는 것은 학교 현장의 평가 상황을 파악하는 데 매우 중요하다. 또한 예비교사들의 평가 관점을 점검하여 예비교사들이 평가를 교수학습을 위한 도구로써 어떻게 인식하고 과학교육의 특수성을 평가에 어떻게 반영하고 있는지를 이해하는 것을 바탕으로 미래 과학 교육과 과학 평가의 방향성을 예측하고 준비할 수 있을 것이다. 그러므로 예비교사들이 제작한 평가 문항을 분석하여 예비교사들이 과학 평가에서 중요하게 생각하는 부분이 무엇이며, 나아가 예비교사들이 가지고 있는 과학교육의 관점을 파악하는 것은 과학교육의 질적 향상을 위한 기초 자료가 될 수 있을 것이다.

## II. 연구방법

### 1. 연구 대상

이번 연구는 현재 우리나라 초등 교사 양성 대학에 재학 중인 2학년 학생 45명을 대상으로 실시되었다. 연구 대상인 초등 예비교사들은 과학교육과 관련하여 2학년 1학기에 초등학교 3~6학년 과학과 교과서에 나오는 전체 단원에 대한 과학 개념을 학습하였으며, 2015 개정 과학과 교육과정에서 제시된 필수 탐구 활동을 경험한 상태였다. 더불어 연구 당시인 2학년 2학기에는 과학과 교수 방법을 학습하는 ‘과학교육방법론’을 수강 중이었다. 교육학과 관련해서는 1학년 1학기에 교육학개론, 초등교육론, 2학기에 교육철학 및 교육사, 2학년 1학기에 교육평가를 수강하였다.

### 2. 자료 수집

자료 수집은 예비교사들이 ‘과학교육방법론’ 강의 를 통해 과학의 본성, 과학 탐구과정, 주요 학습 이론 등을 학습한 후 본격적인 ‘과학 교과 평가 방법’ 강의 에 들어가기 전에 이루어졌다. 예비교사들이 과학 교

과의 평가에 대한 구체적인 소양이 갖춰지기 전 과학 교과와 총괄평가를 작성하도록 하는 것에는 두 가지 목적이 있었다. 연구의 목적은 첫째, 예비교사가 현재 인지하고 있는 과학 교과 평가의 주안점은 무엇인지 파악하기 위함이었다. 둘째, 예비교사 자신들이 현재 인지하고 있는 과학 교과 평가의 주안점을 확인한 후 과학 교과와 구체적인 평가 관점을 익히는 것이 학습에 도움을 줄 것이라 기대했기 때문이다. 실제로 의도했던 두 번째 목적대로본 연구의 연구자는 예비교사들이 자신이 작성한 문항을 보면서 구체적인 과학 교과

의 평가 관점을 확인해감으로써 평가의 관점이 문항 구성 전과 다르게 변화해감을 확인할 수 있었다.

예비교사들이 개발한 총괄평가 문항의 단원은 2015 개정 과학과 교육과정의 5~6학년 군에 속하는 ‘(6) 날씨와 우리 생활’이었다. Fig. 1은 ‘(6) 날씨와 우리 생활’의 교육과정 성취기준과 탐구 활동이다.

예비교사들은 ‘날씨와 우리 생활’ 단원의 교과서와 실험관찰을 보면서 총괄문항을 개발하였다. 예비교사들은 3명 1조로 구성되어 총괄평가 10문항을 개발하였다. 본 연구에 참여한 예비교사들은 총 45명이었음

- [6과06-01] 습도를 측정하고 습도가 우리 생활에 영향을 주는 사례를 조사할 수 있다.
- [6과06-02] 이슬, 안개, 구름의 공통점과 차이점을 이해하고 비와 눈이 내리는 과정을 설명할 수 있다.
- [6과06-03] 고기압과 저기압이 무엇인지 알고 바람이 부는 이유를 설명할 수 있다.
- [6과06-04] 계절별 날씨의 특징을 우리나라에 영향을 주는 공기의 성질과 관련지을 수 있다.

〈탐구 활동〉

- 진습구 온도계로 습도 측정하기
- 이슬, 안개 발생 실험하기
- 바람 발생에 대한 모형 실험하기

Fig. 1. ‘(6) 날씨와 우리 생활’의 교육과정 성취기준과 탐구 활동

1. 다음 습도 측정기의 사용법을 보시오. 다음 습도 측정기를 보시오.

- ① 습도 측정기 사용법을 보시오.
- ② 습도 측정기 사용법을 보시오.
- ③ 습도 측정기 사용법을 보시오.
- ④ 습도 측정기 사용법을 보시오.

2. 다음 습도 측정기 사용법을 보시오. 다음 습도 측정기를 보시오.

시간	1	2	3	4	5	6	7	8	9	10
10	80	80	70	60	50	44	39	37	35	32
11	80	80	70	60	50	44	39	37	35	32
12	80	80	70	60	50	44	39	37	35	32
13	80	80	70	60	50	44	39	37	35	32
14	80	80	70	60	50	44	39	37	35	32
15	80	80	70	60	50	44	39	37	35	32
16	80	80	70	60	50	44	39	37	35	32
17	80	80	70	60	50	44	39	37	35	32
18	80	80	70	60	50	44	39	37	35	32
19	80	80	70	60	50	44	39	37	35	32
20	80	80	70	60	50	44	39	37	35	32
21	80	80	70	60	50	44	39	37	35	32
22	80	80	70	60	50	44	39	37	35	32
23	80	80	70	60	50	44	39	37	35	32
24	80	80	70	60	50	44	39	37	35	32
25	80	80	70	60	50	44	39	37	35	32
26	80	80	70	60	50	44	39	37	35	32
27	80	80	70	60	50	44	39	37	35	32
28	80	80	70	60	50	44	39	37	35	32
29	80	80	70	60	50	44	39	37	35	32
30	80	80	70	60	50	44	39	37	35	32

3. 다음 습도 측정기 사용법을 보시오. 다음 습도 측정기를 보시오.

7. 습도 측정기 사용법을 보시오.

1. 습도 측정기 사용법을 보시오.
2. 습도 측정기 사용법을 보시오.

4. 다음 습도 측정기 사용법을 보시오. 다음 습도 측정기를 보시오.

5. 다음 습도 측정기 사용법을 보시오. 다음 습도 측정기를 보시오.

- ① 습도 측정기 사용법을 보시오.
- ② 습도 측정기 사용법을 보시오.
- ③ 습도 측정기 사용법을 보시오.
- ④ 습도 측정기 사용법을 보시오.

6. 다음 습도 측정기 사용법을 보시오. 다음 습도 측정기를 보시오.

8. 습도 측정기 사용법을 보시오. 다음 습도 측정기를 보시오.

9. 다음 습도 측정기 사용법을 보시오. 다음 습도 측정기를 보시오.

10. 다음 습도 측정기 사용법을 보시오. 다음 습도 측정기를 보시오.

Fig. 2. 예비교사들이 개발한 총괄평가 문항 예시

로 총 15조로 구성되어 활동하였다. 수집된 문항은 총 15세트 150문항이었다. Fig. 2는 예비교사들이 개발한 총괄평가 문항 예시이다.

예비교사들은 총괄평가 문항을 개발하면서 추가적으로 문항의 출제 의도와 답안지, 문항을 개발하면서 어려웠던 점을 추가로 작성하였다. ‘문항 출제 의도’는 예비교사들이 평가에서 어떤 점에 주안점을 두고 있는지 추가적으로 파악하기 위함이었으며, ‘문항 개발 시 어려웠던 점’은 예비교사들이 평가 문항 개발에서 어려워하는 점이 무엇인지를 파악하기 위한 것이었다. 총괄평가 문항은 3명의 조원이 문항의 유형, 문항 구성, 평가해야 할 과학 개념 등에 대해서 자유롭게 토의하면서 개발하였다. 문항 개발한 필요한 표나 그래프, 그림은 교과서 나 실험관찰에 제시된 내용을 오려 붙이거나 간단하게 그려서 활용하도록 안내하였다. 또한 총괄평가 문항 개발에는 충분한 시간을 제공하여 조원들이 모두 만족할 때까지 개발 및 검토할 수 있게 하였다. 평균적으로 대략 1시간 30분 정도가 소요되었으며, 문항 개발에 저해되는 시간으로 생각되지는 않았다.

### 3. 자료 분석

연구를 통해 수집된 자료는 총괄평가 15세트 150문항, 15조의 총괄평가 출제 의도, 15조의 평가 문항 개발 시 어려웠던 점에 대한 설문지로 총 3가지였다. 먼저 총괄평가 문항은 과학 내용 요소(과학 개념), 평가 영역(기억/이해/적용/기초탐구/통합탐구), 문항유형(선택형/서답형), 문항 난이도 등을 반영한 이원목적분류표를 작성하여 분석하였다(Fig. 3). 이원목적분류표의 작성에서는 연구자들이 개별적으로 작성한 것을 비교하였고, 합의를 통해 일치시키는 과정을 거쳤다.

총괄평가 출제 의도와 평가 문항 개발 시 어려웠던 점은 질적 자료 분석과정을 거쳐 귀납적으로 정리·분석하였다. 분석은 질적 분석의 일반적인 절차인 전사를 제외한 코딩 그리고 주제의 발견이라는 과정으로 이루어졌다(Creswell, 2007). 코딩 과정에서 질적 자료 사이의 의미를 부여할 수 있도록 반복적으로 살펴보았다(Huberman & Miles, 1994). 분석의 전과정에서는 과학교육전문가인 공동연구자와 함께 협의하며 이루어졌다. 특히, 문항 난이도는 초등학교 근무 경력 10년

과학 5-2		2. 날씨와 우리 생활		(1)분반-(1)조										
		문항이원분류 및 답안		김OO, 김□□, 박OO										
문항 번호	내용	평가 영역			문항유형					난이도				
		기억	이해	적용	탐구		선택형				서답형			
					기초 탐구	통합 탐구	선다	진위	배합	단답	괄호	서술		
1	습도 측정하기		○							○				하
2	습도가 우리 생활에 미치는 영향			○			○							하
3	이슬, 안개, 구름의 발생 원리	○								○				하
4	이슬, 안개, 구름의 특징	○							○					하
5	(1) 비가 내리는 과정	○						○						하
	(2) 눈이 내리는 과정	○						○						하
6	바람이 부는 까닭					○				○				중
7	계절별 날씨의 특징	○							○					하
8	날씨가 우리 생활에 미치는 영향			○								○		중
9	날씨가 우리 생활에 미치는 영향			○								○		중
10	이슬이 발생하는 과정	○										○		하

Fig. 3. 초등 예비교사들이 제출한 총괄평가 문항을 분석한 이원목적분류표

이상의 석사학위를 소지한 현장 교사 1인, 초등학교 15년 근무 경력의 대학원생과 함께 협의하는 과정을 거쳤다.

### Ⅲ. 연구 결과 및 논의

#### 1. 초등 예비교사가 개발한 총괄평가 문항 분석

초등 예비교사들에게는 ‘날씨와 우리 생활’단원의 교과서와 실험관찰이 주어졌으며, 이를 이용해 총괄평가 10문항을 개발하는 과제가 주어졌다. 초등 예비교사들은 3명 1조로 구성되었으며, 예비교사들은 조별로 협의하면서 자신들이 생각하는 평가의 관점이 반영하여 총괄평가 10문항을 출제하였으며 출제를 마친 후 출제 의도를 서술하였다. 분석 초기에는 초등 예비교사들이 제출한 총괄평가 문항을 분석하여 Fig. 3에서 볼 수 있는 것과 같이 과학 내용 요소, 과학 내용 요소(과학 개념), 평가 영역(기억/이해/적용/기초탐구/통합탐구), 문항유형(선택형/서답형), 문항 난이도 등을 반영한 이원목적분류표를 작성하였다. 이원목적분류표를 작성하여 분석한 구체적인 내용은 다음과 같다.

##### 가. 초등 예비교사들이 개발한 총괄평가의 과학 내용 요소 분석

2015 개정 과학과 교육과정의 ‘3. 내용 체계 및 성취기준’에 제시된 ‘날씨와 우리 생활’ 단원 관련 내용 요소는 습도, 이슬, 구름, 저기압과 고기압, 계절별 날씨이다(교육부, 2015). 그리고 학습 요소도 이와 비슷한데, 습도, 이슬, 안개, 구름, 비와 눈, 기압, 저기압과

고기압, 바람, 계절별 날씨이다(교육부, 2015). 내용 요소에 비해 학습 요소가 조금 더 구체화되어 제시된 것을 알 수 있다. 아울러 ‘날씨와 우리 생활’ 단원 관련 교육과정 성취 기준은 4개가 제시되어 있으며, 탐구활동은 3개가 제시되어 있었다. Fig. 4는 2015 개정 과학과 교육과정에 제시된 ‘날씨와 우리 생활’ 단원의 성취기준과 탐구활동이다.

교육과정에 제시된 ‘날씨와 우리 생활’ 단원의 성취기준을 자세히 살펴보면, 첫 번째 성취기준은 습도 측정과 습도가 우리 생활에 미치는 영향에 대한 것이며, 두 번째 성취기준은 이슬, 안개, 구름의 공통점과 차이점을 이해하고 비와 눈이 내리는 과정을 설명하는 것이었다. 세 번째 성취기준은 저기압과 고기압, 그리고 바람이 부는 이유에 대한 것이었으며, 네 번째 성취기준은 계절별 날씨의 특징과 우리나라 날씨와 공기의 성질의 관련성을 이해하는 것이었다.

이번 연구에서 초등 예비교사들이 개발한 총괄평가의 과학 내용 요소는 교육과정의 내용 요소와 학습 요소와 관련하여 분류해보면, Fig. 5와 같다. Fig. 5에서 볼 수 있듯이 습도, 응결, 이슬/안개/구름, 기압, 공기의 성질, 바람, 비와 눈, 계절별 날씨, 날씨가 우리 생활에 미치는 영향으로 총 9요소로 분석되었다. 분석된 요소들은 구체적으로 살펴보면, 이슬/안개/구름과 관련된 문항이 총 155문항(1문항을 딸린 문항인 2문항으로 쪼개어 출제된 것을 포함) 중 36문항으로 가장 많이 출제되었다. 교육과정의 내용 요소나 학습 요소와 관련이 없는 것은 응결, 바람, 날씨가 우리 생활에 미치는 영향으로 3개 요소였다. 이중 ‘바람’은 교육과정의 탐구활동에 제시된 것으로 교육과정과 관련이 있다고 할 수 있다. 하지만 ‘응결’의 개념은 3~4학년군의 물의 상태 변화와 물의 여행 단원에 학습해야 할 요소이다. 더

[6과06-01] 습도를 측정하고 습도가 우리 생활에 영향을 주는 사례를 조사할 수 있다.

[6과06-02] 이슬, 안개, 구름의 공통점과 차이점을 이해하고 비와 눈이 내리는 과정을 설명할 수 있다.

[6과06-03] 고기압과 저기압이 무엇인지 알고 바람이 부는 이유를 설명할 수 있다.

[6과06-04] 계절별 날씨의 특징을 우리나라에 영향을 주는 공기의 성질과 관련지을 수 있다.

#### 〈탐구 활동〉

- 진습구 온도계로 습도 측정하기
- 이슬, 안개 발생 실험하기
- 바람 발생에 대한 모형 실험하기

Fig. 4. 2015 개정 과학과 교육과정에 제시된 ‘날씨와 우리 생활’ 단원의 성취기준과 탐구활동

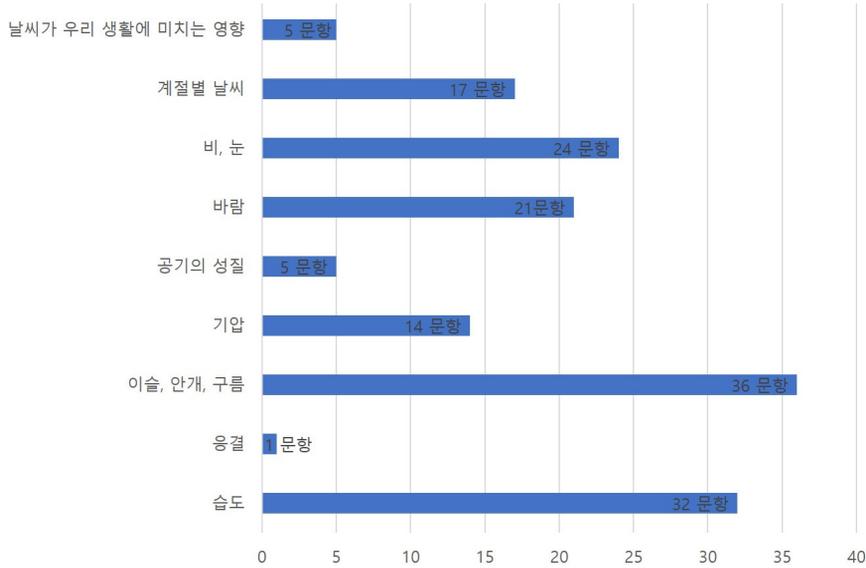


Fig. 5. 초등 예비교사들이 개발한 총괄평가의 과학 내용 요소별 문항 수

불어 ‘날씨가 우리 생활에 미치는 영향’은 2009 개정 과학과 교육과정의 5~6학년군의 ‘(3) 날씨와 우리 생활’ 단원의 ‘학습 내용 성취 기준’의 ‘(사) 날씨가 우리 생활에 많은 영향을 주고 있음을 안다.’에 해당하는 내용으로 2015 개정 과학과 교육과정과 직접적으로 관련이 있다고는 할 수 없다. 그러나 실제 2015 개정 교육 과정에 맞춰 개발된 교과서에는 다루어지고 있는 내용으로 초등 예비교사들은 이점을 반영하여 출제한 것으로 분석된다.

초등 예비교사들이 개발한 문항을 교육과정의 내용 요소와 학습 요소에 따라 구체적으로 분석해보면, 이슬/안개/구름(36문항), 습도(32문항), 비와 눈(24문항), 바람(21문항), 계절별 날씨(17문항), 기압(14문항), 공기의 성질(5문항) 순으로 출제되었으며, 다른 요소에 비해 이슬/안개/구름, 습도, 그리고 비와 눈의 출제가 많은 것을 알 수 있었다. 이것은 단원의 전반부에 배치된 차시들에서 다루는 개념으로 초등 예비교사들의 문항이 계획적으로 출제되지 않았음을 확인할 수 있었다. 덧붙여 뒷 차시로 갈수록 출제된 문항의 수가 줄어들고 문항 출제 수의 편차가 큼을 통해 이렇게 추론할 수 있겠다.

초등 예비교사들이 개발한 문항을 과학 내용 요소별로 분석해보면, 습도와 관련된 문항은 습도의 측정(17문항), 습도가 우리 생활에 미치는 영향(14문항), 습도의 개념(1문항)으로 습도의 측정에 관한 문항이 가

장 많았으며, 습도의 개념은 묻는 문항은 단 1문항에 불과했다. 다음으로 이슬/안개/구름에 관련된 문항을 보면, 이슬, 안개, 구름의 발생 원리나 과정(23문항), 발생 실험(11문항), 이슬, 안개, 구름의 특징(2문항) 순으로 문항이 개발됨을 확인할 수 있었다. 초등 예비교사들은 이슬, 안개, 구름의 차이점인 ‘발생’에 대한 부분을 주요 학습 개념으로 파악하고 있었다.

기압과 관련하여 개발된 문항을 분석해보면, 공기의 온도에 따른 공기의 무게 비교를 포함한 기압의 종류(8문항), 기압의 개념(6문항)에 대한 문항으로 분석되었다. 그리고 바람에 대한 문항은 바람이 부는 까닭과 원리(13문항), 바람 발생 모형 실험(8문항)으로 분석되었다. 비와 눈에 대한 문항은 24문항 모두 비나 눈이 내리는 과정에 대한 문항으로 구성됨을 확인할 수 있었다. 나머지 응결(1문항), 공기의 성질(5문항), 계절별 날씨(17문항), 날씨가 우리 생활에 미치는 영향(5문항)은 내용 요소명과 관련하여 문항이 개발되었음을 확인할 수 있었다.

이상의 문항의 과학 내용 요소 분석을 종합해보면, 앞서 서술한 바와 같이 2015 개정 교육과정에 제시된 내용 요소가 습도, 이슬, 구름, 저기압과 고기압, 계절별 날씨, 총 5개 요소이며, 교육과정에 제시된 학습 요소가 습도, 이슬/안개/구름, 비와 눈, 기압, 저기압과 고기압, 바람, 계절별 날씨로 7개 요소라는 점을 생각하면 교육과정의 내용 요소와 학습 요소가 골고루 출제

되었다고 할 수 있다. 그러나 앞서 서술한 세부 문항의 문항 수를 확인해보면, 10문항이 되지 않는 문항은 이번 연구 참여 조인 15조에서 모두 출제된 것은 아님을 확인할 수 있다. 더불어 Fig. 5에 제시한 과학 내용 요소가 총 9개 요소로 정리된 것에 비추어보면, 같은 내용 요소가 중복 되어 출제되었음을 추론할 수 있다. 즉 앞서 서술한 것과 같이 초등 예비교사들이 문항을 출제할 때 단원의 과학 내용 요소를 골고루 분배하여 출제하지 않았음을 다시 한번 확인할 수 있었다.

**나. 초등 예비교사들이 개발한 총괄평가의 평가 영역 분석**

초등 예비교사들이 개발한 총괄평가 문항을 평가 영역별로 분석해보았다. 평가 영역은 Klopfer의 목표 분류 체계를 이용하여 1차 분석을 위해 작성한 이원목적분류표에 제시된 평가 영역을 중심으로 분석하였다. 분석한 결과는 Table 1과 같다. 초등 예비교사들의 총괄 평가 문항은 Table 1에서 보는 것과 같이 기억에 관한 문항이 65문항(41.10%)으로 가장 많았으며, 그 다음으로 이해 32문항(20.38%), 적용 31문항(19.75%), 통합 탐구 28문항(17.83%), 기초 탐구 1문항(0.64%) 순으로 분석되었다.

탐구 영역(29문항, 18.47%)에 비해 지식 영역(128문

항, 81.53%)에 해당하는 문항의 출제가 월등히 많았음을 확인할 수 있었다. 하지만 이러한 결과가 고무적인 이유는 초등 예비교사들이 지식 위주의 문항 출제가 아닌 탐구 영역의 문항도 출제했다는 것은 탐구의 중요성을 인식하고 있다는 점을 보여주는 방증이라고 할 수 있겠다. 특히 ‘날씨와 우리 생활’ 단원이 5학년 단원으로 기초 탐구 보다는 통합 탐구에 대한 문항 개발이 많이 이루어진 것은 교육과정에서 요구하는 과학 탐구 과정의 강조와 맥을 같이한다고 할 수 있겠다.

**다. 초등 예비교사들이 개발한 총괄평가의 문항 유형 분석**

문항 유형에서는 서답형 문항 출제가 선택형과 비교해 다소 많았음을 확인할 수 있다. 하위 유형으로는 단답형(57문항, 36.96%), 선다형(49문항, 29.35%), 서술형(35문항, 21.74%), 배합형(14문항, 8.70%)순으로 많았고, 진위형과 괄호형은 소수에 그쳤다.

**라. 초등 예비교사들이 개발한 총괄평가의 문항 난이도 분석**

문항의 난이도에서는 하 수준(87문항, 56.13%)이 가장 많았고 중 수준(50문항, 32.26%), 상 수준(18문항,

Table 1. 초등 예비교사들이 개발한 총괄평가의 평가 영역 분석

구분	기억	이해	적용	탐구		합계
				기초 탐구	통합 탐구	
출제 문항 수	65 (41.10%)	32 (20.38%)	31 (19.75%)	1 (0.64%)	28 (17.83%)	157 (100%)

※ 5조의 10번 문항의 경우 문항 유형은 두 개가 통합된 문제이나 평가 영역은 기억 영역으로 1개 처리됨.

Table 2. 초등 예비교사들이 개발한 총괄평가의 문항 유형 분석

구분	선택형			서답형			합계
	선다형	진위형	배합형	단답형	괄호형	서술형	
출제 문항 수	49 (29.35%)	2 (2.17%)	14 (8.70%)	57 (36.96%)	1 (1.09%)	35 (21.74%)	158 (100%)

※ 하위 문제가 있는 경우(5문제)와 한 문제 안에 2가지 유형이 있는 경우를 포함하였음.

Table 3. 초등 예비교사들이 개발한 총괄평가의 문항 난이도 분석

구분	상	중	하	합계
출제 문항 수	18 (11.61%)	50 (32.26%)	87 (56.13%)	155 (100%)

※ 하위 문제가 있는 경우(5문제)를 포함하였음.

11.61%) 의 순으로 나타나 대체로 예비교사들은 평가 문항을 쉽게 출제하는 것으로 나타났다.

## 2. 초등 예비교사의 총괄평가 출제 의도 분석

이번 연구에서는 총괄평가 문항 외에 초등 예비교사들은 어떤 출제 의도를 가지고 총괄 평가 문항을 개발하였는지 조사하였다. Fig. 6은 문항을 개발한 3조의 문항 출제 의도이다. Fig. 6에서도 볼 수 있듯이 초등 예비교사들은 통상적으로 교과서 단원 표지에 제시되는 성취기준을 바탕으로 다양한 문제 유형을 고려하여 문제를 출제하였음을 확인할 수 있었다. 또한 총괄평가이기 때문에 기본 개념과 원리를 중심으로 실생활과 연계할 수 있는 문제를 출제해야 한다는 출제 의도를 파악할 수 있었다. Table 4는 초등 예비교사들의 총괄평가 출제 의도를 질적 분석하여 정리한 것이다.

Table 4에서 볼 수 있듯이 초등 예비교사의 총괄평가 출제 의도는 과학 내용 및 평가 요소, 문제 유형, 문제 난이도의 3개의 범주와 8개의 주제 묶음으로 분석되

었다. 먼저 과학 내용 및 평가 요소의 범주를 보면, 초등 예비교사들은 과학 개념의 정확한 이해, 과학 개념과 생활과의 관계 이해, 차시별 균형 있는 출제, 성취 기준 고려와 같은 내용을 반영하여 문항을 출제하였음을 확인할 수 있었다. 이러한 경향은 총괄평가가 학습이 모두 끝나고 학습의 결과를 검증하기 위한 평가라는 점에서 초등 예비교사들이 총괄평가의 성격을 잘 파악하고 있으며, 이를 평가 문항 출제에 잘 반영하고 있는 것으로 판단할 수 있었다.

두 번째 범주로 문제 유형이 분석되었는데, 초등 예비교사들은 다양한 평가 유형을 반영하고자 노력한 것으로 분석되었다. 하지만 출제 의도와는 달리 앞서 Table 2에서 평가 유형을 분석한 내용에서 확인할 수 있지만, 평가 유형이 골고루 나타나지 않음을 확인할 수 있었다.

세 번째 범주는 문항 난이도로 분석되었다. 초등 예비교사들은 총괄평가를 출제할 때 문항 난이도를 고려한 것으로 확인되었다. 총괄평가의 평균을 7~80점 정

성취기준을 바탕으로, 다양한 문제 형식을 구성하여, 기본개념과 원리를  
구하기 위하여 적용해 볼 수 있는지.  
(실생활)

Fig. 6. 초등 예비교사들의 총괄평가 문항 출제 의도

Table 4. 초등 예비교사의 총괄평가 출제 의도 분석

범주	주제 묶음	주제
과학 내용 및 평가 요소	과학 개념의 정확한 이해	<ul style="list-style-type: none"> <li>문제를 통해 개념 정리 및 적용을 할 수 있도록 함.</li> <li>이론과 답구를 적절히 섞어서 내고자 했다.</li> <li>날씨와 관련된 용어를 복습하고 각각의 날씨가 발생하는 과정과 원리를 제대로 이해했는지 평가합니다.</li> </ul>
	과학 개념과 생활과의 관계 이해	<ul style="list-style-type: none"> <li>과학 개념을 실생활과 연관 지어 서술할 수 있다.</li> <li>“날씨와 우리 생활”이라는 단원명만큼 날씨와 우리 생활과의 연관성을 학생들이 떠올릴 수 있도록 문항을 제작했습니다.</li> </ul>
	차시별 균형 있는 출제	<ul style="list-style-type: none"> <li>소단원(차시)별 균형 있게 출제함.</li> <li>소단원(차시)에 1문제 이상 출제하여 모든 개념을 이해했는지 확인할 수 있도록 한다.</li> </ul>
	성취 기준 고려	<ul style="list-style-type: none"> <li>전체적인 내용을 담기 위해 성취기준별 2문항씩 구성하도록 함.</li> <li>성취기준별로 고르게 내고자 했다.</li> <li>성취기준에 근거하여 꼭 알아야 하는 개념을 숙지하고 있는지를 검사하기 위한 문제로 구성하였다.</li> </ul>
문제 유형	다양한 평가 유형 반영	<ul style="list-style-type: none"> <li>다양한 평가 유형 반영(객관식/주관식, 서술형/논술형)</li> <li>다양한 문제의 형태를 갖추려고 노력하였다.</li> </ul>
문항 난이도	문항의 난이도를 조절하여 출제	<ul style="list-style-type: none"> <li>난이도가 다른 문제를 출제하기가 까다로웠다.</li> <li>쉬운 문항을 앞부분에 배치</li> </ul>

도로 본다면, 두 번째 범주인 문제 유형과 달리 초등 예비교사들이 출제 의도한 문항 난이도는 대체적으로 골고루 반영된 것으로 볼 수 있다. 앞서 Table 3에서 확인할 수 있듯이 상 11.61%, 중 32.26%, 하 56.13%로 중과 하를 합하면 88.39%로 어느 정도 평균과 유사한 점수를 나타낸다고 할 수 있다. 더불어 초등 예비교사들은 문항 난이도에 대해 쉬운 문제를 앞에 배치하여 학생들로 하여금 평가를 조금 더 수월하게 치르도록 배려하는 부분도 출제 의도에 반영하고 있었다.

### 3. 평가 문항 개발 과정에서 초등 예비교사가 겪은 어려움

초등 예비교사에게 평가 문항 제작의 경험을 제공하는 것은 평가의 중요성을 인식시켜주는 계기를 마련해줄 뿐만아니라 교육에서 평가가 차지하는 비중을 인지하도록 하는 계기가 될 수 있다(임성만, 2017). 그렇다면 초등교사는 평가 문항 개발 과정에서 어떠한 어려움을 겪었을까? 초등 예비교사 4조에서는 Fig. 7 과 같이 학생들의 수준을 파악하기 힘들어 평가 문항 개발이 어려웠다고 응답하였다. 더불어 많은 초등 예비교사들이 지적했던 것과 같이 초등 예비교사 4조에서도 성취 기준별 문항 구성과 객관식 문항 개발 시 선지 구성의 어려움을 지적하였다.

선다형 문항의 선지 구성을 비롯한 난이도 조절의 어려움은 Fig. 8에서 보는 것과 같이 초등 예비교사들이 총괄평가 문항 개발 시 어려워하는 항목임을 확인할 수 있었다. 초등 예비교사들이 총괄평가 문항 개발 시의 어려움을 종합해보면 Table 5와 같다.

Table 5를 보면, 초등 예비교사들이 문항 개발 과정에서 겪는 어려움으로는 ‘과학 내용 및 평가 요소’와 관련해서는 성취기준 고려, 이전 내용과의 연계성 고려, 실생활 연계 문제의 구성이 분석되었다. ‘문제 유형 및 난이도’와 관련해서는 초등 예비교사들은 문제의 유형을 다양하게 고려하는 것, 난이도 조절을 어려워하고 있었다. 마지막으로 ‘경험 부족’ 범주에서는 가장 많은 초등 예비교사들이 지적했던 선지의 구성을 비롯해 학생의 수준 파악, 정확한 개념 이해 부족, 서술형 문제 출제 경험의 부족, 학생들의 오개념 고려, 중복 문항 배제, 문제 서술, 알맞은 자료 찾기가 있었다. 이번 연구를 통해 무엇보다도 초등 예비교사들은 평가 문항 개발의 경험이 부족하여 다양한 어려움을 겪고 있는 것을 확인할 수 있었다. 임성만(2017), Lam(2015), Siegel과 Wisschr(2011)이 연구에서 지적했던 것과 같이 초등 예비교사들에게는 평가 문항 제작을 비롯해 다양한 경험을 제공하는 것이 매우 우선시 되어야 할 것으로 생각된다.

학생들이 어떤 부분을 어려워하고 어떤 부분을 쉽게 이해할 수 있는지 파악하는 것이 어려웠고, 성취기준별로 어떤 문항 유형을 활용하는 것이 좋은지 고민이 되었다. 그리고 객관식 문항에서 선지를 구성하는 것이 어렵다.

Fig. 7. 총괄평가 문항 개발 과정에서 초등 예비교사 4조가 겪은 어려움

난이도 조절이 힘들었다.  
선다형 문항을 만들 때, 선지 구성 순서결  
정하는 것이 어려웠다.  
선다형 문항을 만들 때, 선지가 매우 중복을  
피하는 것이 어려웠다.  
선다형 문항이 만들어 지는데 구성형 문항의  
개수가 많아서 난이도가 올라가는 것.

Fig. 8. 총괄평가 문항 개발 과정에서 초등 예비교사 12조가 겪은 어려움

Table 5. 초등 예비교사의 문항 개발 과정에서 겪은 어려움

범주	주제 묶음	주제
과학 내용 및 평가 요소	성취기준 고려	<ul style="list-style-type: none"> <li>성취기준을 고려하다보니, 문항 제작이 더욱 어려웠던 것 같다.</li> <li>각 성취기준에 따라 쉬운 문항 1개, 어려운 문항 1개씩 구성하여 학생들을 변별하고자 하였다.</li> </ul>
	이전 내용과의 연계성 고려	<ul style="list-style-type: none"> <li>아이들이 이전에 배운 내용, 성취기준에 어긋나지는 않는지 파악하는 것이 어려웠다.</li> </ul>
	실생활 연계 문항의 구성	<ul style="list-style-type: none"> <li>생각보다 과학적 내용을 실생활과 연결해서 문항을 낼 때 사례가 잘 생각나지 않아 어려웠다.</li> <li>아이들이 실생활에 적용할 수 있도록 하는 문항을 만드는 것도 어려웠다.</li> </ul>
문항 유형 및 난이도	문항의 유형을 다양하게 고려하는 것	<ul style="list-style-type: none"> <li>내용이 겹치지 않게 구성하기 어려웠다.</li> <li>각 개념을 학습하기에 적합한 문항 유형(선다형, 서술형)을 생각하기 어려웠다.</li> </ul>
	난이도 조절	<ul style="list-style-type: none"> <li>초등학생에게 적합한가?</li> <li>학생들의 수준에 맞춰 문항의 난이도를 조절하는 것이 어려웠다.</li> </ul>
경험 부족	선지의 구성	<ul style="list-style-type: none"> <li>객관식을 만들 때 정답을 제외한 나머지 문항(오답)을 만드는 것이 고차적인 사고를 요구하고 있어서 어려웠다.</li> <li>선지의 명확한 표현 고려의 어려움</li> </ul>
	학생의 수준 파악	<ul style="list-style-type: none"> <li>학생의 수준을 파악하기 어려워 문항의 난이도를 조절하기 어려웠음</li> <li>학생 수준에 대한 파악의 어려움</li> </ul>
	정확한 개념 이해 부족	<ul style="list-style-type: none"> <li>정확한 개념이해가 필요함을 깨닫게 되었다.</li> </ul>
	서술형 문항 출제 경험의 부족	<ul style="list-style-type: none"> <li>서술형 문항을 쓸 때 문항의 어휘가 길어지면 학생들이 어려워하지 않을까 신경쓰느라 힘들었다.</li> <li>창의적인 문항을 만들고자 했는데 오히려 성취기준에 벗어나기도 했다.</li> </ul>
	학생들의 오개념 고려	<ul style="list-style-type: none"> <li>학생들이 헷갈릴 만한 개념, 혹은 오개념이 무엇이 있을지 생각해보는 것이 어려웠다.</li> </ul>
	중복 문항 배제	<ul style="list-style-type: none"> <li>한정된 내용 내 중복되지 않는 문항 구성의 어려움</li> </ul>
	문항 서술	<ul style="list-style-type: none"> <li>문항을 명확하게 서술하는 것</li> </ul>
	알맞은 자료 찾기	<ul style="list-style-type: none"> <li>10문항이라는 제한된 문항 수에 교과서의 주요 탐구과정을 선별해서 담는 과정에서 어려움이 있었다.</li> <li>문항에 알맞은 자료를 찾는 것이 어려웠다.</li> </ul>

#### IV. 종합논의

총괄평가는 학생들이 학습목표를 얼마나 잘 달성했는지에 대한 정보를 제공한다. 즉, 학생들의 평가결과를 교육활동에 대한 피드백으로 추후 교육에 대한 합리적인 의사결정을 위한 중요한 정보이다(O'Malley, 2010). 정확한 평가를 위해서 평가 문항의 품질은 중요하며 우수한 평가 문항을 개발하는 것은 교사가 갖추어야 할 중요한 역량 중 하나로 볼 수 있다(Engelsen & Smith, 2014). 우수한 평가 문항은 적절한 난이도, 목표와 문항 사이의 일관성, 목표에 적합한 유형, 명확한 지시 제공, 적절한 언어의 사용과 같은 특징을 갖는다(Chen & Chen, 2016; Gareis & Grant, 2015; Osterlind, 2002). 총괄평가 출제 의도에서 드러난 것과 같이 초등 예비교사들은 우수한 평가 문항의 특징에 대해서 잘

이해하고 있는 것으로 보여진다. 그러나 노태희 외(2015)가 보고한 것과 같이 예비교사들이 갖는 인식과 평가의 실행은 차이가 있는 것으로 여겨진다.

여러 선행연구에서 총괄평가의 문항 개발에서 교사와 예비교사들이 어려움을 겪고 있다고 보고하고 있다. 문항 개발에 대해서 불안 및 부정적인 태도가 나타나며(Ebinye, 2001; Quansah & Amoako 2018), 문항 개발에 대한 지식과 기술이 부족한 것으로 나타났다(김신영, 2012; Agu et al., 2013; Annan, 2020; Onyechere, 2000; West, 2005). 교사와 예비교사들은 타당하고 신뢰로운 시험을 구성하는 것이 어려우므로 이미 개발된 평가 문항을 주로 이용하게 된다(Onyechere, 2000). 선행연구에서는 개발된 평가 문항이 타당함에 대한 양적연구가 주로 연구되어 왔으므로 총괄평가 문항 개발에서 필요한 지식과 경험은 무엇인지에 대해서 구체적으로 밝혀진 바는 없다. 이에 본 연구에서는 초등 예비

교사의 문항 개발에 대한 인식으로부터 질적분석을 통해 예비교사들에게 필요한 지식과 경험이 무엇인지 밝혀냈으며 이는 추후 예비교사의 문항 개발을 위한 교육자료를 개발하는데 중요한 아이디어가 될 것이다.

우수한 문항을 개발하는 것은 지식과 경험을 바탕으로 여러 상황을 고려하여야 하는 전문성이 필요하다(성태제, 2010). 또한, 이러한 지식은 정적인 것이 아니라 복잡하고 역동적이다(Xu & Liu, 2009). 문항 개발을 위해서는 교과목에 대한 이해, 학생에 대한 이해, 문항 제작에 대한 이해 뿐만 아니라 문항 제작과 검토의 경험 역시 갖춰야한다(Engelsen & Smith, 2014). 문항 개발의 전문성을 갖추기 위해서는 문항 개발의 경험과 훈련, 개발된 문항의 분석에 대한 명시적인 훈련이 필요하다(성태제, 2004; 임성만, 2017; Kinyua & Okunya, 2014; Lam, 2015; Quagrain & Arhin, 2017). 따라서 교사 및 예비교사는 문항 개발의 효과적인 방법에 대한 정기적인 교육을 조직하는 것이 필요하다고 생각된다.

## V. 결론 및 제언

이번 연구는 초등 예비교사들이 총괄평가 문항을 개발 과정에서 평가 문항과 관련하여 어떠한 양상을 보이고 평가 문항 개발과 관련한 출제 의도와 어려움을 분석하는 데 목적이 있었다. 연구 결과를 요약하면 다음과 같다. 첫째, 초등 예비교사들은 총괄평가 문항 개발 시 과학 내용 요소를 골고루 반영하고 있었다. 둘째, 초등 예비교사들은 문항 평가 영역과 관련해서는 영역별로 골고루 반영하지 못하고 있었으며, 이러한 경향은 문항 유형에서도 나타났다. 셋째, 초등 예비교사들이 총괄평가 문항 개발 시 가장 어려워했던 부분은 문항 난이도를 조절하는 부분이었으며, 이러한 부분은 실제 문항 난이도를 분석하는 과정에서도 나타났다. 넷째, 초등 예비교사들의 문항 출제 의도를 파악한 결과, 초등 예비교사들은 문항 개발 시 고려해야 할 사항을 거의 대부분 고려하여 출제한 것을 확인할 수 있었으나 실제 개발된 문항은 그렇지 못하였다. 즉 과학 내용 요소는 골고루 반영되었으나, 평가 영역을 비롯해 문항 유형과 난이도의 조절에서는 실패하였다고 할 수 있었다. 다섯째, 초등 예비교사들이 총괄 평가 문항

개발 시 가장 큰 어려움은 문항 난이도 조절과 객관식 문항의 선지 구성, 학생의 수준 이해 부족이 가장 큰 것으로 분석되었다.

연구 결과를 바탕으로 다음과 같이 본 연구의 결론을 도출하였다. 첫째, 초등 예비교사들은 과학 내용 요소를 균형 있게 반영하는 능력을 갖추고 있음을 확인하였다. 이는 예비교사들이 과학 교과의 전반적인 내용을 이해하고 있음을 시사한다. 그러나 평가 영역과 문항 유형의 균형 있는 구성에는 어려움을 겪고 있어, 이 부분에 대한 교육과 훈련이 필요함을 알 수 있다. 둘째, 초등 예비교사들의 문항 개발 과정에서 겪는 어려움을 비롯해 문항 유형 및 난이도 조절 실패는 평가 문항 제작 경험 및 학생에 대한 이해 부족에서 오는 것으로 판단할 수 있다. 이는 이론적 지식과 실제적 적용 사이의 간극을 보여주는 것으로, 예비교사 교육에서 더 많은 실습 기회와 피드백이 필요함을 시사한다. 셋째, 객관식 문항의 선지 구성에 대한 어려움은 예비교사들이 학생들의 다양한 사고 과정과 오개념을 예측하고 활용하는 능력이 부족함을 나타낸다. 이를 개선하기 위해 학생들이 갖는 일반적인 오개념과 사고 과정에 대한 이해를 높여야 할 것이다.

이러한 결론은 예비교사 교육 과정에서 평가 문항 개발에 대한 더 체계적이고 실제적인 교육이 필요함을 시사한다. 특히 과학 교과의 특성을 반영한 다양한 평가 영역과 문항 유형에 대한 교육, 학생들의 수준을 이해하고 이를 문항 개발에 반영하는 능력을 기르는 교육이 강화되어야 할 것이다. 이를 위해서는 교사 양성 과정 및 실습 과정에서 예비교사들에게 다양한 경험의 기회를 제공해야 한다. 다른 직업과 달리 교직은 인턴 과정과 같은 적응 기간이 없다. 임용고시를 합격하면 바로 교사가 되는 시스템으로 교직에 대한 다양한 경험은 교사 양성 과정에서 충분히 제공되어야 한다. 또한 평가에 대한 중요성을 다시금 인식하여 초등 예비교사들에게 체계적인 강의가 이루어져야 함을 명심해야 할 것이다.

## 국문요약

총괄평가는 학생들이 학습목표를 얼마나 잘 달성했는지에 대한 정보를 제공하므로, 정확한 평가를 위해

서는 우수한 평가 문항을 개발해야 하며 이는 교사가 갖추어야 할 역량 중 하나이다. 본 연구는 초등 예비교사들이 총괄평가 문항을 분석하고 출제 의도와 문항 개발 과정의 어려움을 분석하고자 한다. 연구는 초등 교사 양성 대학에 재학 중인 2학년 학생 45명을 대상으로 실시되었다. 3인 1조로 10개 문항을 개발하도록 하였고, 문항의 출제의도, 답안지, 문항 개발시 어려웠던 점을 작성하도록 요구하였다. 수집된 자료 중 총괄평가 문항은 Klopfer의 교육 목표 분류 체계가 포함된 이원목적분류표에 따라 평가 문항을 분석하였고, 총괄평가 출제 의도와 평가 문항 개발 시 어려웠던 점은 질적 자료 분석과정을 거쳐 귀납적으로 정리·분석하였다. 연구결과, 초등 예비교사들은 과학 내용 요소는 골고루 반영하고 있었으나, 문항 평가 영역에서는 골고루 반영하지 못했다. 가장 어려워했던 점은 난이도 조절이었다. 출제 의도에서는 문항 개발 시 고려해야 할 사항을 거의 대부분 고려하였으나 실제 개발된 문항에는 반영되지 않았다. 이러한 결과로부터 총괄평가 문항개발에는 지식과 경험이 중요하며 초등 예비교사들에게 체계적인 강의가 필요하다는 것을 제언한다.

주제어: 초등 예비교사, 날씨와 우리 생활, 총괄평가, 평가 문항 제작

## References

- 교육부(2014). 초등학교 과학 6-1 교사용 지도서. 한국과학창의재단 국정도서편찬위원회.
- 교육부(2015). 과학과 교육과정. 교육부 고시 제205-74호.
- 권재술(1984). Klopfer의 과학교육 목표 분류의 본질과 문제점. *과학교육논총*, 9, 67-72.
- 김상달, 이용섭, 최성봉(2005). Klopfer의 교육 목표 분류에 따른 제7차 교육과정의 중학교 과학 교육 목표 분석-7학년을 중심으로, *한국지구과학교육학회지*, 26(7), 640-651.
- 김신영(2002). 현장교사의 평가전문성 연구. *교육평가연구*, 15(1), 67-85.
- 김신영(2012). 교사의 학생평가와 성취평가제. *교육평가연구*, 25(4), 655-677.
- 김정훈, 박영신(2012). 중등 예비 과학교사들의 지구과학 영역 탐구문제 개발 능력 분석. *한국지구과학회지*, 33(3), 294-305.
- 노태희, 이재원, 강석진, 강훈식(2015). 중등 과학교사의 과학 평가 실태와 지향. *한국과학교육학회지*, 35(4), 725-733.
- 박현주(2016). 중학교 과학교사의 지필평가의 역할 및 실태에 대한 인식 조사. *교사교육연구*, 55(3), 389-398.
- 손연아, 윤지현, 맹희주(2018). 대학교 전공 계열별 교수자들의 학습 평가 목적에 대한 인식과 수업 유형에 따른 평가 방법 활용 현황. *교육문화연구*, 24(3), 5-24.
- 임성만(2017). '지층과 화석'단원에 대해 초등 예비 교사가 개발한 평가 문항 분석. *학습자중심교과교육연구*, 17(7), 129-146.
- 정은영, 최원호(2014). 우리나라 초·중·고등학교 과학의 평가 실태. *과학교육연구지*, 38(1), 168-181.
- 최현숙, 김중복(2013). 예비 물리 교사들의 학생평가 문항 제작 수행 수준 조사 & #8211; 형성 평가를 위한 신다형 물리 개념 검사 문항 분석 중심으로. *과학교육연구지*, 37(3), 458-475.
- Abell, S. K., & Siegel, M. A. (2011). Assessment literacy: What science teachers need to know and be able to do. In D. Corrigan, J. Dillon, & R. Gunstone (Eds.), *The professional knowledge base of science teaching* (pp. 205-221). London, UK: Springer.
- Agu, N. N., Onyekuba, C., & Anyichie, A. C. (2013). Measuring teachers' competencies in constructing classroom-based tests in Nigerian secondary schools: Need for a test construction skill inventory. *Educational Research and Reviews*, 8(8), 431.
- Annan, J. K. (2020). Preparing globally competent teachers: A paradigm shift for teacher education in Ghana. *Education Research International*, 2020(1), 8841653.
- Carless, D. (2015). Exploring learning-oriented assessment processes. *Higher Education*, 69(6), 963-976.
- Chen, H., & Chen, J. (2016). Retrofitting non-cognitive-diagnostic reading assessment under the generalized DINA model framework. *Language Assessment Quarterly*, 13(3), 218-230.
- DeLuca, C., & Klinger, D. A. (2010). Assessment literacy development: Identifying gaps in teacher candidates' learning. *Assessment in Education: Principles, Policy*

- & Practice, 17(4), 419-438.
- Ebinye, P. O. (2001). Problems of testing under the continuous assessment programme. *Journal of Qualitative Education*, 4(1), 12-19.
- Engelsen, K. S., & Smith, K. (2014). Assessment literacy. In *Designing assessment for quality learning* (pp. 91-107). Dordrecht: Springer Netherlands.
- Gareis, C., & Grant, L. W. (2015). Teacher-made assessments: How to connect curriculum, instruction, and student learning. Routledge.
- Kinyua, K., & Okunya, L. O. (2014). Validity and reliability of teacher-made tests: Case study of year 11 physics in Nyahuru district of Kenya. *African Educational Research Journal*, 2(2), 61-71.
- Krathwhol, D. R. (2002). A revision of Bloom's taxonomy. *Theory into Practice*, 41(4), 212-219.
- Lam, R. (2015). Language assessment training in Hong Kong: Implications for language assessment literacy. *Language Testing*, 32(2), 169-197.
- O'Malley, P. (2010). Students evaluation: Steps for creating teacher-made test. In *Assessment Group Conference-School programme*. Maryland: Kennedy Krieger Institute.
- Onyechere, I. (2000). New face of examination malpractice among Nigerian youths. *The Guardian Newspaper* July 16.
- Osterlind, S. J. (2002). *Constructing test items: Multiple-choice, constructed response, performance, and other formats*. New York: Kluwer Academic Publishers.
- Quaigrain, K., & Arhin, A. K. (2017). Using reliability and item analysis to evaluate a teacher-developed test in educational measurement and evaluation. *Cogent Education*, 4(1), 1301013.
- Quansah, F., & Amoako, I. (2018). Attitude of Senior High School (SHS) teachers towards test construction: Developing and validating a standardised instrument. *Research on Humanities and Social Sciences*, 8(1), 25-30.
- Siegel, M. A., & Wissehr, C. (2011) Preparing for the plunge: Preservice teachers' assessment literacy. *Journal of Science Teacher Education*, 22(4), 371-391.
- West, K. I. (2005). Japanese high school mathematics teacher competence in real world problem solving and its implication for the use of modeling to improve Japanese mathematics education. Columbia University.
- Xu, Y., & Liu, Y. (2009). Teacher assessment knowledge and practice: A narrative inquiry of a Chinese college EFL teacher's experience. *Tesol Quarterly*, 43(3), 492-513.