

초등학생의 과학핵심역량 측정을 위한 컴퓨터 기반 평가도구(TCSC) 개발

채희인 · 노석구[†]

Development of Test of Computer-based Tools for Measurement of Science core Competencies (TCSC) of Elementary School Students

Chae, Heein · Noh, Sukgoo[†]

국문 초록

본 연구는 과학핵심역량을 평가할 수 있는 표준화된 도구가 개발되어 있지 않아 학교에서 지필 위주의 수행평가가 이루어지고 있고, 평가 주체의 경우에도 교사에 의한 평가에만 머물러 있는 현실을 개선하기 위하여 표준화된 과학핵심역량 평가도구를 개발하고자 하였다. 이를 위해 과학교육 전문가와 초등학교 교사 5명의 검토를 바탕으로 초안을 마련하였으며, 과학교육을 전공한 교사 3명의 내용타당도 검토, 1차(52명)와 2차(54명) 난이도 및 변별도 검사를 시행하였고, 25명의 학생을 대상으로 예비 검사를 시행하여 내적 신뢰도 분석을 하였다. 최종 확정된 문항은 4~6학년 278명을 대상으로 본 검사를 진행하였으며, 7명의 교사가 평가자로 참여하였다. 이 과정에서 기초통계량 분석(평균, 표준편차, 최댓값, 최솟값, 왜도, 첨도)과 2차 측정모형을 구안하여 확인적 요인분석(구성타당도-표준화회귀계수, 모형적합도-절대적합도, 증분적합도, 간명적합도 지수)을 시행하였다. 마지막으로 공인타당도 분석과 신뢰도 분석을 하였다. 이를 바탕으로 과학핵심역량평가의 활용 방법을 제시하기 위하여 문항 체계표 및 Z점수로 환산한 평정기준을 제시하였다. 최종적으로 13명의 과학교육 전문가와 교사, 그리고 409명의 학생을 대상으로 평가도구 개발이 이루어졌다. 이와 같은 과학핵심역량평가(TCSC) 개발은 평가의 주체를 교사에서 확장하며, 다양한 방식과 도구를 통해 보다 실제적인 평가를 가능하게 하고, 기존의 방식으로는 평가가 어려운 부분을 평가할 수 있었고, 컴퓨터 기반 평가로 인하여 교사의 편의성을 증진할 수 있었다.

주제어: 과학핵심역량, 초등학교, 평가, 컴퓨터 기반 평가

ABSTRACT

This study aimed to develop a standardized evaluation tool of science core competencies to improve the reality that paper-oriented performance evaluation is being conducted in schools due to the lack of such a tool. Toward this end, we prepared a draft on the basis of a review of a science education expert and five elementary school teachers; moreover, we conducted first and second tests to determine difficulty and discrimination. Lastly, we conducted reliability analysis on 25 students. The final questions were administered on 278 students in fourth to sixth grades, and seven teachers participated as evaluators. In this process, the study conducted confirmatory factor analysis by devising a basic statistical analysis and a secondary measurement model. Finally, the study performed certified validity and reliability analyses. Based on the results and to present the use of the test of computer-based tools for the measurement of science core competencies (TCSC), the evaluation criteria were converted into a question system table, and Z-scores were presented. Finally, the evaluation tools were developed for 13 science education experts, teachers, and 409 students. The development of the TCSC was meaningful as

2023.03.15(접수), 2023.03.29(1심통과), 2023.04.17(2심통과), 2023.04.24(최종통과)

E-mail: sgnoh@ginue.ac.kr(노석구)

an effective evaluation tool while expanding the subject of evaluation by teachers, which enables practical evaluation using various methods and tools, evaluates the difficult parts of the existing method, and enhances teacher convenience.

Key words: science core competency, elementary school, assessment, computer-based assessment

I. 서 론

1. 연구의 목적 및 필요성

2015 개정 교육과정에 학교 교육의 전 과정을 통해 중점적으로 기르고자 하는 핵심역량(자기관리 역량, 지식정보처리 역량, 창의적 사고 역량, 심미적 감성 역량, 의사소통 역량, 공동체 역량)을 명시하면서 우리나라 교육과정에 역량이라는 개념이 처음으로 등장하게 되었다. 이를 바탕으로 각 교과에서도 교과별 특성에 맞는 핵심역량을 제시하였고, 그에 따라 과학 교과에서도 과학적 사고력, 과학적 탐구 능력, 과학적 문제해결력, 과학적 의사소통능력, 과학적 참여 및 평생학습능력과 같은 5가지의 과학핵심역량(ScC: Science core Competencies)을 제시하게 되었다(교육부, 2015a, 2015b).

하지만 총론 수준의 역량과 각 교과 수준의 역량이 상호관련성이 낮고, 각각 개발되어 위계 관계가 역전되는 등 많은 문제점이 제기되었다. 또한 교육과정은 핵심역량 중심으로 변화되었지만 핵심역량 평가에 대한 구체적인 도구 및 지침이 부족하여 학교 현장에서는 기존의 평가 방법을 그대로 활용하는 등 많은 혼란이 있었다. 따라서 역량 평가의 어려움이 학교 현장에서 발생하였다(김현정과 김성기, 2021; 박종임 등, 2021b). 이와 같은 문제점을 해결하기 위하여 핵심역량의 평가를 위한 많은 연구가 이루어졌지만 기존의 평가와 큰 차이점을 갖지 못하는 경우가 많았고, 핵심역량의 실제적인 측정에 대한 많은 논쟁이 이어졌다.

학교 현장에서 형성평가라는 이름으로 실행되고 있는 많은 평가가 구조화된 맥락에서 폐쇄적인 질문으로 시행되는 전통적인 방식의 총괄평가에 치우쳐있었으며, 그 목적 또한 학생들의 지적 학습 수준 특징과 선발을 위한 것에 불과했다. 따라서 평가의 결과가 학습자의 지속적인 성찰과 학습의 개선을 위한 피드백 자료로 활용되는 경우도 거의 없는 것으로 나타났다(강훈식과 강석진, 2015; 김은숙과 김현경, 2021; 박영석, 2009; 여상인 등, 2007; Brown,

2004; Wang *et al.*, 2010).

교사가 구성주의적 평가관을 갖고 있더라도 실제로 구성주의적 평가를 시행하지 않고 전통적 평가를 시행하는 괴리가 있으며, 많은 교사가 평가에 대한 인식은 구성주의적인 형성평가를 지향하고 있었지만, 실제로 교사들이 실행했던 평가의 경우 대부분 전통적인 총괄평가를 실행하는 경우가 많아 평가에 대한 인식과 실행에 많은 괴리가 있었다(김민환 등, 2022; 김성기 등, 2020; 노태희 등, 2015; Thomas *et al.*, 2011). 즉 과학 평가 활동이 학생의 학습을 지속적으로 성찰하고 개선하며, 교수·학습의 질을 재고하기 위해 서술형과 논술형 및 수행평가 등의 다양한 방식으로 실행되어야 한다는 인식과 다르게 실제로는 학생의 인지적 학습능력만을 평가하고 평가 결과는 학생들의 서열을 나열하는 것에만 활용되고 있었다.

학교 현장에서 이루어지고 있는 평가 활동의 문제점을 바탕으로 국가 수준에서 제시하고 있는 평가 방향에 대하여 파악하기 위해 핵심역량 중심의 교육과정 개혁을 통해 만들어진 2015 개정 교육과정의 총론과 과학 교과의 평가 방향을 살펴본 결과 교육과정에서는 구성주의적 평가 및 다양한 평가를 명시하고 있었고, 앞으로 적용될 2022 개정 교육과정에 평가가 학생의 학습과 성장을 도울 수 있도록 계획하여 실시되어야 하고, 지식·이해 중심의 평가를 지양하며, 학습의 결과뿐만 아니라 학습의 과정도 함께 평가하도록 하였다(교육부, 2022). 하지만 실제 학교 현장에서 이루어지는 과학 평가 활동은 이와 많은 차이를 보이는 것이 현실이다(신정운과 박상우, 2021; 심보경과 유미현, 2020). 앞으로 적용될 2022 개정 교육과정에도 평가가 학생의 학습과 성장을 도울 수 있도록 계획하여 실시되어야 하고, 지식·이해 중심의 평가를 지양하며, 학습의 결과뿐만 아니라 학습의 과정도 함께 평가하도록 하였다(교육부, 2022).

특히 교육과정에서는 성취기준에 근거하여 학교에서 중요하게 지도한 내용과 기능을 평가하며 수행평가, 정의적 영역 평가, 실험·실습 등의 평가가

이루어져야 한다고 명시하고 있지만, 구체적인 평가 방법이 제시되어 있지 않기 때문에 기존에 시행되어 온 실험보고서와 지필 형식의 수행평가 위주로 이루어지고 있다.

즉 과학 평가의 경우 형식적으로 실시되고 있으며, 실제로는 지필평가를 중심으로 평가되고 있고, 탐구 영역보다 지식 영역에 한정하여 평가하는 경향이 큰 것으로 나타났다. 그 이유로는 구성주의적 평가 활동에 시간과 노력이 많이 필요하고, 능력 부족과 평가 영역의 제한 및 국가 수준에서 마련된 평가 관련 자료의 부족 등이 있었다(강남화와 김민지, 2020; 김성기 등, 2021; 양일호 등, 2008; 오현석과 이기영, 2006; 정은영과 최원호, 2014). 또한 곽영순과 하지훈(2020)은 교사의 과학 학습평가에 관한 연구를 통하여 평가에 대한 이해 부족에 따른 어려움, 업무 과중 등에 따른 평가 실행의 어려움, 교사의 평가 업무 과중과 실효성에 대한 의문이 있다고 하였다.

따라서 평가가 교과 지식 위주로 이루어지기 때문에 사교육을 받는 비중이 크며, 사교육은 대부분 교과 내용 숙지와 이를 반복하기 위한 선행학습 위주로 진행되어 정답을 맞히기 위한 암기 위주의 시험공부를 하는 경향이 있었다. 크게 확장하여 살펴보면 학생들은 12년간의 학업의 결과를 한 번 치러지는 수능 시험으로 평가받으며, 그로 인해 인생이 바뀌거나 올해 입시가 실패로 끝날 경우 다시 365일을 학원에서 지내게 되는 사회전반적인 문제를 야기하게 되었다. 이러한 방식은 학생의 진정한 학습을 위한 것이 아닌 입시만을 목적으로 하는 교육 평가로 변질되게 하여 결국에는 미래의 지식정보화 사회에서 필요로 하는 창의·융합형 인재 육성과 거리가 먼 산업사회의 표준화된 인재 교육에만 머물러 있게 될 것이다(김유진, 2015; 김현주, 2015; 정종오, 2015; 조준형, 2015).

따라서 이론과 교육과정 문서에만 제시되는 수준에서 벗어나 학교 현장에서 구성주의적 형성평가가 적용되기 위하여 많은 노력과 전반적인 평가 시스템의 개혁이 필요하다고 할 수 있고, 과학핵심역량을 실제적으로 평가할 수 있는 도구가 개발되어 교사의 핵심역량 평가에 대한 부담을 줄이고, 컴퓨터 기반 평가와 같은 방법을 통하여 다양한 정보를 수집할 필요가 있다. 하지만 대부분의 연구가 과학핵심역량이 아닌 일반적인 역량을 지수화하는 것에만 머물러 있었고(권희경 등, 2020; 남궁지영 등, 2016),

과학핵심역량을 표준화할 수 있는 검사 도구가 아닌 과학핵심역량 조사 문항 개발(하민수 등, 2018)과 같은 연구만 진행되었다. 또한 컴퓨터 기반으로 과학핵심역량평가를 위한 연구가 진행되었지만 특정 단위이나 차시, 또는 특정한 탐구능력을 파악하기 위해 컴퓨터 기반으로 제작된 탐구 활동을 수행하는 것에 초점을 두어 과학핵심역량의 전반적인 평가는 어려웠다(김인숙 등, 2020; 박종임 등, 2021a).

따라서 본 연구에서는 범교과적 핵심역량과 과학핵심역량의 개념적 정의에 대한 국내·외 연구 동향 분석과 같은 채희인과 노석구(2015)의 사전 연구를 바탕으로 채희인(2017)이 총론 수준의 범교과적 핵심역량에 대한 91개 요소 사이의 479개의 연결과 과학핵심역량에 대한 78개 요소 사이의 387개의 연결을 네트워크 분석기법을 활용하여 분석하고, 11명의 전문가를 대상으로 델파이 연구를 시행하여 정립한 ‘핵심역량 구조모형’과 ‘과학핵심역량 모형 - 윤리 · 인성역량(과학적 태도 및 정서, 과학적 가치관, 과학적 윤리성), 개인적 관리역량(과학적 진로개발, 평생학습능력), 상호작용역량(과학적 의사소통능력, 과학적 책임과 참여, 과학적 협업능력), 인지적 학습역량(과학적 탐구능력, 과학적 사고력, 과학적 지식학습능력, 과학적 문제해결력, 과학적 정보처리능력)’을 바탕으로 과학핵심역량 측정을 위한 평가도구(TCSC)를 개발하고자 하였다.

2. 평가도구 설계의 원리

교육의 평가에 있어서 실제적인 평가를 지향하는 평가관(assessment)의 기본 원리를 살펴보면 첫째, 학생과 이를 둘러싼 다양한 환경을 파악하기 위해서는 개인, 환경, 그리고 이들의 상호작용에 대한 포괄적인 자료수집이 필요하다(박도순 등, 2007). 즉 오랜 시간 동안 다양한 방법으로 다양한 맥락적 요소에 대한 평가를 지속적으로 실시하는 것을 의미한다고 할 수 있다. 둘째, 학생과 맥락에 대한 자료수집과 해석에 있어서 체계적인 접근을 필요로 한다(박도순 등, 2007). 이는 평가에 활용되는 다양한 자료들이 체계적인 절차와 방법을 통하여 수집되고 해석되어야 하며, 종합적으로 해석되어야 함을 의미한다. 셋째, 타당도와 신뢰도의 조화가 중요시된다. 기존에 전통적인 평가관에서는 신뢰도와 효율성이 핵심적인 가치로 인식되었지만, 구성주의적 평가관에서는 타당도와 효과성도 매우 중요한 가치로 인정

받고 있다.

이와 같은 학습활동에 대한 기본적인 평가 원리를 바탕으로 본 연구에서는 과학핵심역량평가의 원리를 다음과 같이 제시하였다.

첫째, 과학핵심역량평가는 다양한 요소에 대한 평가를 시행한다. 특히 과학핵심역량 모델에서 제시하는 4가지 대영역의 13개 과학핵심역량 요소에 대한 평가에 중점을 두며, 이는 학생의 윤리·인성역량, 개인적 관리역량, 상호작용역량, 인지적 학습역량을 모두 포함하는 것으로 학생의 역량을 구성하는 종합적인 부분에 대한 평가라고 할 수 있다. 또한 다양한 요소에 대한 평가는 기존에 우리나라의 과학교육에서 실행된 평가 활동이 대부분 과학적 지식과 같은 인지적인 영역에 치중되어 있던 점을 개선하고 지식이 아닌 과학핵심역량을 평가하기 위해서는 지식 이외의 다양한 영역에 대한 평가도 실행되어야 함을 의미한다. 그에 따라 본 연구에서는 과학핵심역량 모델을 바탕으로 ‘조직화의 원리’로 이를 설명하고자 한다.

과학핵심역량평가의 조직화 원리는 학생의 역량을 평가할 때 부분이 아닌 역량을 구성하는 전체 요소에 대한 평가가 총체적으로 이루어져야 한다는 것이다. 하지만 실제 학교 현장에서는 가장 파악하기 쉽고 개발하기 쉬운 인지적 학습역량만 평가하고 있으며, 그중에서도 특히 과학적 지식학습역량만을 평가하는 경향이 크게 나타났다. 이는 파악이 어렵고 평가하기 어려운 요소인 윤리·인성역량과 상호작용역량과 같은 요소들을 간과하는 결과를 초래하게 되었다. 따라서 겉으로 보기에는 모든 요소에 대하여 평가하고 있다고 생각할 수 있지만, 실제로는 역량의 가장 중요한 부분들을 간과하고 있었다.

우리는 파악하기 쉽고 평가에 큰 노력을 들이지 않아도 측정이 가능한 인지적 학습역량만을 평가해왔으며, 인지적 학습역량과 같은 지식의 암기와 이해에 대한 평가를 통해 학생들을 서열화하고 선발하였다. 즉 이러한 것들이 학생들의 능력을 대표할 수 있고 때론 전부라고 판단하였다.

하지만 조금만 다른 관점에서 살펴보면 우리가 간과하고 있던 많은 부분이 학생들의 역량을 구성하는 중요한 요소임을 파악할 수 있다. 특히 본 연구의 사전 연구로 진행되었던 과학핵심역량 모델에 대한 연구 결과 역량에서 지식과 같은 인지적 학습역량뿐만 아니라 다양한 역량들이 존재하며 인지적

학습역량은 다양한 역량 중 하나에 불과하기 때문에 하나의 요소만으로 그 학생의 역량을 평가했던 기존의 평가방식에 개선이 필요하다는 점을 인식시켜 주었다.

둘째, 과학핵심역량평가는 다양한 교육 주체에 의한 평가를 시행하기 위해 ‘다차원의 원리’를 바탕으로 한다. 특히 학생을 둘러싼 학습의 맥락적 요소를 공유하고 있는 교사, 동료 그리고 학생의 자기평가 활동은 다양한 주체에 의한 평가라는 점에서 필수적인 부분이라고 할 수 있다. 물론 동료평가의 경우 그 실효성에 대한 상당한 논란이 있는 것은 사실이나 논란을 피하기보다 지속적인 인식 개선과 교육을 통하여 동료평가가 정착된다면 기존에 학생에 대해서 파악하지 못한 새로운 측면의 평가가 가능하다는 점에서 매우 중요하다고 할 수 있다.

동료평가는 교사의 평가 부담을 덜어줄 수 있으며, 학습 결과가 아닌 과정의 평가에서는 동료 학생들이 학생을 둘러싼 교육적 맥락의 필수적인 주체이며, 다양한 관점에서 평가하는 것을 가능하게 한다는 장점이 있다(채수진, 2006; Topping & Ehly, 1998), 하지만 이러한 장점에도 동료평가는 Ballantyne *et al.*(2002)과 Falchikov(1995)의 견해와 같이 동료평가의 공정성과 같은 문제점이 있다. 따라서 과학핵심역량평가에서는 동료평가의 문제점을 해결하기 위해 지속적인 동료평가에 대한 인식 개선과 컴퓨터 기반 평가를 바탕으로 하는 통계적인 방법을 통하여 최소 3인 이상의 평가를 지향하며, 최대한으로는 컴퓨터 기반 평가가 빅데이터의 처리가 가능하다는 장점을 활용하여 학생이 소속되어 있는 반의 모든 학생을 동료평가의 주체로 활용하여 동료평가의 문제점을 해소하고자 하였다.

셋째, 과학핵심역량평가는 다양한 방법에 의한 평가를 실행하기 위해 ‘다양성의 원리’를 바탕으로 시행한다. 특히 역량의 각 요소에 맞는 다양한 평가방식을 통하여 학생의 역량을 파악하고자 하였으며, 자기평가의 경우 문항에 대한 해결 능력의 측정을 바탕으로 과학핵심역량의 함양 정도를 파악하기 위하여 선다형 문항과 자기 설문 중심의 설문형 평가를 시행하고, 교사 및 동료평가의 경우 관찰, 면담, 보고서 및 포트폴리오 평가를 기반으로 시행되는 설문 중심의 평가를 시행하여 하나의 평가 활동에서 다양한 방법이 활용될 수 있도록 하였다.

넷째, 과학핵심역량평가는 지속적인 평가를 위해

‘지속성의 원리’를 바탕으로 시행한다. 이를 위해 평가 활동이 교수·학습에 부담을 주지 않도록 컴퓨터 기반 평가도 함께 제작하였으며, 이는 평가에 대한 교사의 부담을 덜어주고, 다양한 역량에 대한 평가를 수행하기 위해 각 역량별 평가지를 활용하여 매번 평가하는 부담을 덜어줄 수 있을 것이다. 특히 지속적인 평가를 통해 누적된 결과는 학생의 과학핵심역량의 함양 정도를 파악할 수 있는 중요한 자료로써 활용될 수 있을 것이다.

다섯째, 과학핵심역량평가는 학생의 과학핵심역량의 함양 정도를 효과적이고 효율적으로 측정할 수 있는 도구적 특성인 평가의 ‘도구성의 원리’를 바탕으로 컴퓨터 기반 평가를 시행한다. ICT를 활용한 평가는 기존의 전통적인 평가의 도구뿐만 아니라 기존에 측정하기 힘들었던 평가를 가능하게 하는 가능성을 열어줄 것이다(Griffin et al., 2012). 따라서 컴퓨터 기반 평가 시스템의 구축은 평가와 평가 자료의 저장 그리고 활용의 유의성으로 인해 국제적인 평가뿐만 아니라 다양한 평가 활동에서 활용되고 있으며, 본 과학핵심역량평가에서는 위의 네 가지 평가 원리를 실현하기 위한 방안으로 컴퓨터 기반 평가를 시행하고자 하였으며, 이는 본 평가의 현장 적용 가능성을 높이는 중요한 요소가 될 것이다.

평가를 한자와 영어로 살펴보면 ‘評價’와 ‘e-valuation’으로 이는 “가치를 드러내거나 매김”을 뜻한다고 할 수 있으며, 국어사전에도 “사물의 가치나 수준 따위를 평하는 것”을 평가라고 하였다(박도순 등, 2007). 즉 교육에서는 학생의 가치를 드러내는 활동이 평가하고 할 수 있으며, 이는 교수·학습과 함께 교육 활동의 중요한 과정이라고 할 수 있다. 그러므로 학생의 학업적인 가치는 단편적인 지식의 암기를 평가하는 활동이 되어서는 안 되며, 진정한 가치를 발견할 수 있는 활동이 되어야 한다. 그러므로 과학핵심역량평가(TCSC)에서는 과학핵심역량 모델을 바탕으로 학생의 역량을 이루는 다양한 요소에 대한 평가를 시행하는 것을 목표로 하였다.

II. 연구 방법

1. 연구 대상

과학핵심역량평가(TCSC: Test of Computer-based tools for measurement of Science core Competencies)를

개발하기 위한 초기 단계에서 과학교육 전문가인 교육대학 교수와 초등학교 교사 5명의 검토를 바탕으로 평가체제를 구성하여 초안을 제작하였다. 개발된 문항은 전문가(교수)의 검토를 바탕으로 과학교육을 전공하고 현장에서 과학 교육과 관련된 다양한 활동을 하고 있는 3명의 현직교사를 대상으로 내용타당도(CVI) 검토를 하였고, 각각 52명과 54명의 4~6학년 학생을 대상으로 1·2차 난이도 및 변별도 점검을 시행하였으며, 25명의 학생을 대상으로 예비 검사를 시행하여 문항의 내적 신뢰도를 분석하였다.

본 검사 및 확인적 요인분석 단계에서는 대도시/중·소도시/농어촌을 고려하여 경기도 소재 초등학교 12개 학급에 재학 중인 4~6학년 총 278명의 학생을 대상으로 평가를 시행하였고, 교사평가의 과정에서 7명의 현직 교사가 참여하였다. 최종적으로 TCSC를 개발하기 위한 과정에서 13명의 과학교육 전문가와 현직 교사가 참여하였으며, 초등학교 4~6학년 학생 409명이 참여하였다.

2. 검사 도구

TCSC를 개발하기 위한 과정에서 최종적으로 개발된 문항의 공인타당도를 검증하기 위하여 표준화된 검사도구를 특정하기 어려운 평생학습능력을 제외한 과학적 태도 및 정서, 과학적 의사소통능력, 과학적 탐구능력의 세 영역에 대한 상관계수를 분석하여 공인타당도 검증을 하였다. 이는 새롭게 개발된 평가도구의 경우 기존에 개발되고 많은 연구를 통해 활용되고 검증되어 온 평가도구와의 상관관계를 살펴보는 것이 새로 개발된 평가도구의 공인타당도 검증에 가장 효과적인 방법이기 때문이다(김동현, 2014). 이를 위해 본 검사의 세 영역에서 표준화된 평가도구로 개발되고, 많은 연구에 활용되어 본 연구를 통해 개발된 평가도구의 공인타당도 검증에 적합하다고 판단되는 김효남 등(1998)의 과학적 태도 검사, 전성수(2013)의 과학적 의사소통능력 검사, 권재술과 김범기(1994)의 과학탐구능력 검사를 활용하였다.

3. 평가도구 개발 절차

TCSC를 개발하기 위하여 사전 연구, 교육과정, 과학교육 전문가 협의 결과 그리고 평가의 기본 원리를 바탕으로 문항을 구성하였으며, 문항의 초안에 대한 내용타당도, 신뢰도, 난이도 및 변별도 검증을 실시하였다. 위와 같은 단계를 통하여 개발된 본 검사 문항

은 278명의 학생을 대상으로 본 검사를 시행하여 기초통계량 분석(평균, 표준편차, 최댓값, 최솟값, 왜도, 첨도)을 하였고, 2차 측정모형을 구성하여 확인적 요인분석을 하였다.

2차 측정모형으로 구안된 연구모형에 대한 구성타당도와 모형적합도 분석을 시행한 이후 전체 및 각 영역에 대한 상관분석, 공인타당도 검증, 최종적인 검사 문항의 신뢰도 분석, 평정기준의 제시를 통한 활용방안 마련의 단계를 거쳐 최종적인 문항 개발을 마무리하였으며(Fig. 1), 개발의 단계에 따른 세부적인 내용은 연구 결과 및 논의에서 제시하였다.

III. 연구 결과 및 논의

1. 과학핵심역량평가(TCSC)의 개발

1) 평가 문항 제작

학생의 진정한 과학핵심역량을 평가하기 위하여

본 연구에서는 채희인(2017)의 과학핵심역량 모델의 13가지 역량 중 각 영역별로 데이터 분석과 델파이 연구를 통하여 그 중요성이 높은 네 가지 역량을 평가하기 위한 TCSC를 개발하고 이를 바탕으로 현장 적용 가능성 및 실현 가능성에 관한 논의를 하고자 한다.

본 연구에서 채희인(2017)의 사전 연구를 바탕으로 본 평가도구를 개발한 이유는 2015 개정 초등학교 교육과정의 역량중심 교육과정을 추구하며 처음 역량이 도입되었지만 총론 수준의 역량과 각 교과 수준의 역량이 하나의 유기적인 연구를 통해 역량요소를 제시한 것이 아니어서 역량의 위계가 역전되거나 연관성이 부족하고, 과학핵심역량의 경우에도 총론에서 제시된 역량과 유기적인 관련성을 갖는 과학 교과의 역량을 심층적인 연구를 통해 제시한 것이 아니었기 때문에 많은 문제점이 제기되었고 2022 개정 과학과 교육과정에서는 역량을 명시적으로 제시하지 않게 되었다. 따라서 국내·외에서 진행된 국가적 수준의 역량 규명 연구와 기관 및 개

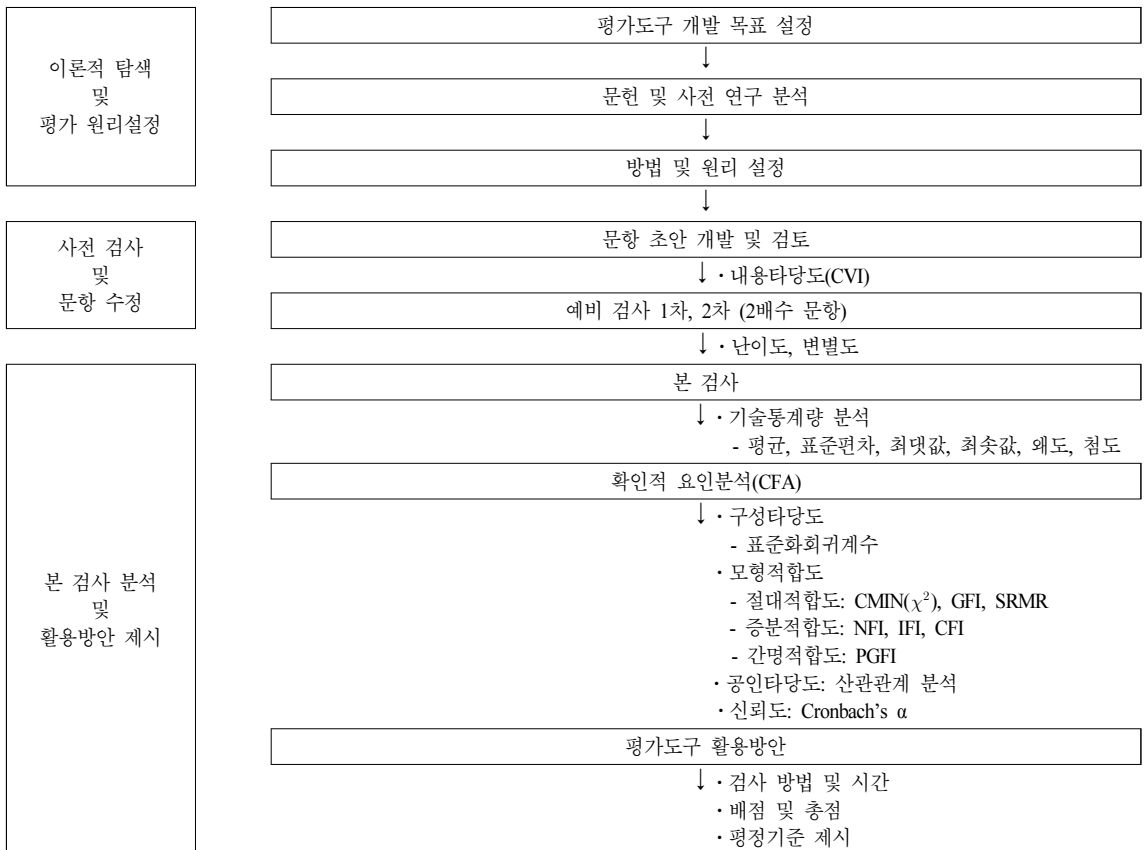


Fig. 1. The process of research

인 수준의 연구에서 제시하고 있는 핵심역량에 관한 빅데이터를 키워드 네트워크 분석을 통하여 상호 유기적인 관련성을 갖는 역량의 체계를 규명하고, 델파이 연구를 통하여 제시한 채희인(2017)의 연구는 과학핵심역량에 관한 체계적이고 학술적인 접근을 하였다는 점에서 과학핵심역량평가 개발을 위한 사전 연구로 적합하다고 판단하였다.

특히 과학핵심역량 모델의 13가지 역량 중 각 영역별로 한 가지씩 총 4가지 영역만을 선정한 이유는 평가 문항 개발을 위한 물리적인 제약을 고려한 것으로 각 요소 모두에 대한 표준화된 평가도구를 개발하는 것은 본 연구의 수준과 규모를 고려하였을 때 어렵다고 판단하였고, 추후 국가 주도의 과학핵심역량평가 개발을 통하여 각 영역의 전문가 및 평가 전문가 그리고 프로그래머 등이 참여하는 거시적인 연구가 필요하다고 할 수 있다.

본 연구에서는 채희인(2017)의 과학핵심역량 모델 중 윤리·인성역량의 ‘과학적 태도 및 정서’, 개인적 관리역량의 ‘평생학습능력’, 상호작용역량의 ‘과학적 의사소통능력’, 인지적 학습역량의 ‘과학적 탐구능력’에 대한 사전 평가 문항을 구성하였다. 평가 문항의 경우 초등학교 학생들에게 적합한 용어와 서술로 작성하였으며, 자기평가의 경우 40분 이내로 평가할 수 있는 문항 구성을 원칙으로 하였다. 특히 TCSC의 다섯 가지 평가 원리를 바탕으로 문항을 구성하여 체계적인 평가가 되도록 하였다. 각 평가 항목들의 경우 과학교육 전문가와 현직 교사 5명의 검토를 바탕으로 자기평가/교사평가/동료평가 체계를 구성하였으며, 과학적 의사소통능력과 과학적 탐구능력의 경우 설문 방식의 평가뿐만 아니라 문항에 대한 해결 능력의 측정을 바탕으로 과학핵심역량의 함양 정도를 파악하기 위하여 선다형 문항을 도입하여 다양한 방식의 평가가 이루어질 수 있도록 구성하였다.

네 가지 영역 모두 자기평가와 교사평가를 기본으로 하였으며, ‘과학적 태도 및 정서’와 ‘과학적 의사소통능력’의 경우 과학교육 전문가 집단의 협의 결과 동료평가가 가능하며 다양한 주체에 의한 평가가 필요하다고 판단하여 동료평가 형식을 도입하였다. 반면 ‘평생학습능력’과 ‘과학적 탐구능력’의 경우 동료들이 평가하기에 어려운 부분이 있을 수 있다는 판단에 따라 동료평가를 시행하지 않았다. 따라서 동료평가가 필요한 요소와 불필요한 요소에

대하여 선별적인 접근을 하였다. 특히 동료평가의 경우 최소 3인 이상의 학생들이 평가를 시행하여 동료평가의 단점을 보완할 수 있는 체계를 마련하였으며, 컴퓨터 기반 평가방식의 구성으로 인하여 동료평가의 인원을 더 늘리는 것이 가능하도록 하였다.

이를 바탕으로 4가지 역량의 하위요소별로 검사 문항의 약 2배수인 77문항을 예비문항으로 개발하였다. 세부 요소의 선정에 있어서 기존의 검사 문항 및 이론적 배경을 참고하였으며, 최종적으로 전문가 집단에 의한 협의 과정을 통하여 문항을 선정하였다.

첫째, 과학적 태도 및 정서의 경우 Fraser(1981)가 개발한 TOSRA(Test of Science-Related Attitudes)를 바탕으로 세부 요소를 선정하기 위하여 과학적 태도 및 정서와 관련된 요인들을 탐색하였다. 먼저 일반적인 심리학의 범주에서는 태도와 정서를 다양한 감정의 요소들로 세분화하였다.

Roseman *et al.*(1990)은 주요 정서의 요소를 기쁨, 안도감, 희망, 열정, 자부심, 놀람, 혐오, 괴로움, 슬픔, 두려움, 좌절, 불친절함, 분노, 부끄러움, 죄책감, 유감의 16가지로 구분하였다. 김경희(1997)는 흥미, 즐거움, 슬픔, 우울, 분노, 혐오, 공포·불안, 수줍음, 수치심, 죄책감, 사랑의 11개 정서 요소로 분류하였다.

과학교육과 관련된 연구의 경우에는 일반적인 태도 및 정서와는 차별화되는 요소들이 포함되어 있었다. 대표적으로 Fraser(1981)의 TOSRA에서는 ‘과학탐구에 대한 태도’, ‘과학적 태도의 수용’, ‘과학 수업의 즐거움’, ‘과학에 대한 취미로서의 즐거움’, ‘과학에 대한 직업으로서의 즐거움’ 등을 하위 요소로 제시하고 있다. 김효남 등(1998)은 국가 수준의 과학에 관련된 정의적 특성의 평가체제 개발을 통하여 과학과 관련된 정의적 영역을 인식, 흥미, 과학적 태도의 3개 범주로 구분하였으며, 인식 영역의 경우 과학에 대한 인식, 과학교육에 대한 인식, 과학자와 과학 관련 직업에 대한 인식, 과학-기술-사회의 상호 관련성에 대한 인식으로 구분하였다. 흥미 영역의 경우 과학에 대한 흥미, 과학 학습에 대한 흥미, 과학과 관련된 활동에 대한 흥미, 과학과 관련된 직업에 대한 흥미, 과학 불안으로 세분화하였다. 마지막으로 과학적 태도 영역의 경우 호기심, 개방성, 비판성, 협동성, 자진성, 끈기성, 창의성으로 구분하였다. Misitomu(2012)는 SMASEC(Students’ Motivation, Attitude and Self-Efficacy Toward Learning Chemistry Questionnaire)를 개발하고 과학 학습과 관련된 희망,

즐거움, 흥미, 만족의 정서를 측정하고자 하였다. 광영순 등(2006)도 흥미, 즐거움, 자신감 등의 과학 학습과 관련된 정서에 관하여 연구하였다.

과학적 태도 및 정서에 관한 사전 연구들을 바탕으로 교육과정에 포함된 태도 및 정서의 요소들을 종합적으로 고려하여 호기심, 즐거움, 만족, 수용성, 끈기, 자신감, 객관성, 개방성의 8가지 하위요소를 최종적으로 선별하였다.

둘째, 평생학습능력의 경우 UNESCO에서 제시한 평생학습에 관한 이론적 배경을 바탕으로 하였으며, Uzunboylu & Hursen(2011)의 LLLCS(Life-Long Learning Competence Scale)에 제시된 평생학습의 6가지 영역인 자기관리역량(self-management competencies), 학습하는 방법에 관한 역량(competencies of learning how to learn), 새로운 것에 도전하고 개척하는 역량(competencies of initiative and entrepreneurship), 정보습득역량(competencies on acquiring information), 디지털 역량(digital competencies), 결정과 실행역량(competencies of decision-taking)과 그 하위 요소들을 참고하였다. 김진화 등(2014)은 앎을 위한 평생학습역량, 실존을 위한 평생학습역량, 행함을 위한 평생학습역량, 공생을 위한 평생학습역량, 생성을 위한 평생학습역량으로 구분하고 그 하위요소들을 바탕으로 평생학습역량 측정 도구를 개발하였다. 백수정과 이희수(2012)는 평생교육자의 퍼실리테이션 역량에 관한 연구를 바탕으로 평생교육과 관련된 역량 요소들을 선별하여 진단척도를 개발하였다.

또한 평생학습능력은 자기주도적 학습능력과 관련되며(김지자 등, 1996; 이석재 등, 2003; Guglielmino, 1977), 이와 관련한 연구를 살펴보면 이석재 등(2003)은 생애능력의 하위 요소로서 자기주도적 학습능력을 포함시키며 학습욕구진단, 목표설정, 학습을 위한 자원파악, 기본적 자기관리능력, 학습전략의 선택, 학습실행의 지속성, 결과에 대한 노력기인, 자기성찰의 8가지 하위요소로 구분하였다. Guglielmino(1977)는 SDLRS(Self-Directed Learning Readiness Survey)를 제작하며 학습기회에 대한 개방성, 효과적인 학습자로서의 자아개념, 학습에 대한 주도성과 독립성, 자기 자신의 학습에 대한 책임 수용, 학습사랑, 창의성, 기본 학습기능과 문제해결 기능을 가용하는 능력, 미래에 대한 긍정적 지향으로 하위요소를 구분하였다. 김지자 등(1996)은 SDLRD-K-96을 개발하며 독창적 접근, 탐구적 특성, 자발적 계획, 학습의 책

임성 수용, 학습에 대한 사랑, 미래 지향성, 학습자적 신념의 7개 요소로 구분하였다.

본 연구에서는 평생학습능력에 관한 문헌연구와 교육과정에 포함된 핵심역량의 평생학습능력을 종합적으로 고려하여 평생학습능력을 자신성, 주도성, 자기이해, 자기관리, 미래지향, 인식, 열정, 기초학습능력, 학습자원 파악, 공동체 활용의 10개 요소로 구분하였다.

셋째, 과학적 의사소통능력의 경우 전성수와 박종호(2014)의 과학적 의사소통능력 검사 도구에 포함된 과학적 설명(서술, 설명)과 과학적 주장(근거, 정당화)을 바탕으로 다양한 연구에 포함된 의사소통능력의 요소들을 참고하였으며, 기본적으로 BAPS에서 과학적 의사소통을 단어나 기호(상징)를 활용하여 사건, 그림 또는 물체를 표현하는 것으로 규명한 것을 바탕으로 하였다(Padilla *et al.*, 1983).

구분용 등(1994)은 의사소통의 요소로 경청, 공감, 수용적 존중, 솔직성, 즉시성, 구체성을 주장하였으며, 유현숙 등(2004)은 정보수집 및 종합, 경청, 고정관념적 사고 극복, 창의적 의사소통, 자기 드러내기, 주도적 의사소통, 정서표현, 타인관점 이해로, MacKlin & Rossiter(1976)는 표현력, 자아노출, 이해력, Ruben(1976)은 존중 표현, 상호작용 태도, 지식정향, 인내, 타인역할정향적 행동, 상호작용관리로 구분하였다. 또한 Bienvenu & Stewart(1976)는 자기표현, 주의, 명료성, 감정처리, 회피, 지배, 차이 조정, 지각된 적응성, 이석재 등(2003)은 정보수집, 경청, 고정관념적 사고 극복, 창의적 의사소통, 자기 드러내기, 주도적 의사소통, 타인관점 이해로 의사소통능력의 요소들을 추출하였다. 마지막으로 이현주 등(2015)은 과학기술 사회에서 요구되는 시민 역량에 대한 검사를 제작하기 위하여 요인분석을 시행하고 의사소통능력을 의사표현 및 의견조정, 적극적 청취의 두 가지 요소로 구분하였다.

과학적 의사소통능력에 대한 문헌연구와 교육과정에 포함된 핵심역량의 과학적 의사소통능력을 종합적으로 고려하여 본 연구에서는 과학적 의사소통능력을 경청, 명료성, 협력적 대화, 설득, 판단, 의견조정, 이해 및 적용, 설명, 표와 그래프 이해, 시각화의 10개 요소로 구분하였다.

넷째, 과학적 탐구능력의 경우 SAPA 프로그램에서 제시하는 기초 탐구 과정(관찰, 분류, 시·공간 관계 사용, 의사소통, 수 이용, 측정 예상, 추리), 통

합 탐구 과정(가설 설정, 변인통제, 실험 수행, 모델 설정, 자료해석, 조작적 정의)과 이를 바탕으로 제시된 우리나라 교육과정의 기초 탐구 과정(관찰, 분류, 측정, 예상, 추리, 의사소통 등), 통합 탐구 과정(문제 인식, 가설 설정, 변인 통제, 자료 해석, 결론 도출, 일반화 등)을 바탕으로 하였다(교육부, 2015b; AAAS, 1989).

또한 Klopfer(1971)는 과학탐구의 요소로 관찰과 측정, 문제 발견과 해결방안 모색, 자료의 해석 및 일반화, 이론적 모델의 형성, 검증 및 수정을 제시하였으며, NAEP(1984)는 관찰, 측정, 실험, 의사소통, 자료해석, 일반화, 내삽과 외삽, 가설설정, 비유에 의한 논증, 종합, 귀납과 연역에 의한 논증, 모델 설정 그리고 허명(1984)은 자료의 수집과 정리(기구조작, 관찰, 측정, 자료의 기록, 분류, 자료의 변형), 자료 해석 및 계산(추론, 상관관계 결정, 인과관계 설명, 외삽, 예언), 자료의 종합 및 평가(요약, 결론, 일반화, 평가), 가설설정 및 실험설계(문제발상, 가설설정, 조건통제, 실험설계), 권재술과 김범기(1994)는 기초탐구능력(관찰, 분류, 측정, 추리, 예상), 통합탐구능력(자료변환, 자료해석, 가설설정, 변인통제, 일반화)로 과학탐구의 요소를 구분하여 제시하였다.

과학적 탐구능력에 관한 문헌연구와 교육과정에 포함된 탐구 과정 요소를 종합적으로 고려하여 본 연구에서는 과학적 탐구능력을 관찰, 분류, 측정, 추리, 예상, 가설설정, 변인통제, 자료변환, 자료해석, 결론 도출의 10개 요소로 구분하였다.

2) 전문가 검토 및 수정·보완

평가의 경우 의도한 바를 제대로 평가하고 있는지에 대한 타당화 작업이 필요하며, 그에 따라 전문가 검토를 바탕으로 과학교육을 전공한 3명의 현직 초등학교 교사를 대상으로 내용타당도 검토를 시행하였다. 특히 내용타당도(Content Validity Index: CVI) 평가를 시행하는 것은 전문가 그룹의 판단이 매우 중요하며(성태제, 2002), 이를 수치화하여 판단의 척도로 활용하는 방법도 매우 중요하다고 할 수 있다. 따라서 본 연구에서는 Fehring(1987)이 제시한 CVI 산출 방식을 활용하였으며, 이는 Likert 5점 척도의 경우 각 점수별로 가중치(1점=0, 2점=.25, 3점=.50, 4점=.75, 5점=1.00)를 부여하는 방식이다. 판단 준거의 경우 .80 이상은 매우 타당한 것(major)으로 보았으며, .50 이상 .80 미만은 수정 및 연구자의 판단에

따라 선택적으로 타당한 것(minor)으로 보았다.

과학교육 전문가 3인에 의한 내용타당도 검토 결과 대부분 .80 이상으로 내용타당도가 높게 나온 것으로 나타났으며, 과학적 태도 및 정서 영역의 개방성 12번 문항(.75)과 평생학습능력 영역의 주도성 20번 문항(.75), 자기이해 22번 문항(.58)의 경우 .50에서 .85의 범위에 포함되어 수정 및 연구자의 판단에 따라 선택적으로 타당한 것(minor)으로 나타났으며, 추후 최종 검사 문항의 선정에 있어서 다른 지표들과 함께 다각적으로 판단할 필요가 있는 것으로 파악되었다.

3) 예비 검사

예비문항 중 49~52, 63~67번 문항의 경우 선다형 문항으로 난이도와 변별도를 판단할 필요가 있기 때문에 52명의 학생을 대상으로 1차 난이도와 변별도를 점검하였으며, 결과를 바탕으로 문항을 수정·보완하여 54명의 학생을 대상으로 2차 난이도와 변별도를 점검하였다. 또한 문항의 적절성을 높이기 위하여 25명의 학생을 대상으로 예비 검사를 실행하여 문항의 내적 신뢰도를 분석하였다.

문항 난이도(item difficulty)는 문항의 쉽고 어려운 정도를 의미하는 지수로, 전체 피험자 중에서 정답자의 비율로 나타낸다. 문항 난이도를 판단하는 절대적인 기준은 없지만 Cangelosi(1990)가 제시한 기준과 일반적인 연구 및 평가 도구 개발에서 활용되는 판단 준거를 바탕으로 본 연구에서는 .30 미만이면 매우 어려운 문항, .30 이상 .80 미만이면 적절한 문항, .80 이상이면 매우 쉬운 문항으로 판단하였다.

난이도를 정확히 판단하는 것은 매우 어려운 문제이지만 일반적으로 .50일 때 해당 문항은 변별도가 가장 높은 분포를 보이므로 난이도는 .50이 가장 적합하다고 할 수 있다. 본 연구에서는 문항 난이도의 허용 범위를 .30 이상 .80 미만으로 설정하고 문항 개발을 하였다.

문항 변별도(item discrimination)는 문항이 피험자의 능력을 알맞게 구분하는 정도를 나타내는 지수로, 변별도 지수는 해당 문항의 점수와 피험자의 총 점과의 상관계수에 의하여 추정한다. 예를 들어 검사의 총점을 기준으로 피검사자 집단을 상·하위 집단으로 구분하였을 때, 해당 문항의 상위집단 정답률이 하위집단 정답률보다 높았다면 그 문항은 변별도가 높은 것으로 판단할 수 있다.

본 연구에서는 Ebel(1965)의 변별도 지수의 판단 준거를 바탕으로 .30 이상 .40 미만은 변별력이 있는 문항, .40 이상을 변별력이 높은 문항으로 판단하였다. 하지만 문항 변별도의 경우 문항 난이도에 민감하게 반응하므로 두 가지 지수 모두를 종합적으로 판단하고자 하였다.

내용타당도 결과와 1차 난이도 및 변별도 점검 결과 52, 66, 67번 문항의 경우 변별도 허용 범위를 벗어나 이를 바탕으로 과학교육 전문가 1인과 현직 교사 2인의 협의를 통하여 2배수의 예비문항 초안을 수정 및 보완하였으며, 확정된 2배수의 예비문항 중 49~52, 63~67번 문항의 경우 54명의 학생을 대상으로 2차 난이도 및 변별도 점검을 시행하였고, 본 연구에서 설정한 난이도와 변별도의 범위를 모두 충족하였다.

2배수의 예비문항에 대한 신뢰도 검사 결과 문항 전체의 신뢰도는 .955로 매우 높은 것으로 나타났으며, 세부 영역별로 살펴보면 과학적 태도 및 정서 영역이 .898, 평생학습능력 영역이 .914, 과학적 의사소통능력 영역이 .807, 과학적 탐구능력 영역이 .772로 나타났다.

4) 본 검사

약 2배수의 예비문항에 대한 전문가의 내용타당도 점검 결과와 예비 검사를 통한 난이도 및 변별도, 신뢰도 분석 결과를 바탕으로 1배수의 자기평가, 교사평가, 학생평가 문항을 Fig. 2와 같이 구성하였다.

본 연구에서는 문헌 및 연구 데이터에 대한 분석과 이론을 바탕으로 검사 문항을 구성하였기 때문에 탐색적 요인분석을 생략하고 확인적 요인분석을 실시하였으며(노경섭, 2015; Arbuckle, 2016), 확인적 요인분석의 결과는 다음과 같다.

(1) 기술통계

검사 대상인 총 278명의 학생을 대상으로 시행한

	자기평가(S)	교사평가(T)	동료평가(P)
과학적 태도 및 정서(A)	AS1-AS8 40점	AT1-AT6 30점	AP1-AP6 30점
평생학습능력 (L)	LS1-LS10 50점	LT1-LT10 50점	-
과학적 의사소통능력 (C)	CS1-CS4 20점 CS5-CS8 20점	CT1-CT6 30점	CP1-CP6 30점
과학적 탐구능력(I)	IS1-IS5 25점 IS6-IS10 25점	IT1-IT10 50점	-

Fig. 2. Composition and allocation of TCSC

본 검사의 각 영역 및 문항 전체에 대한 기술통계 결과를 살펴보면 Table 1과 같이 각 영역별 100점 만점의 문항에 대하여 과학적 태도 및 정서의 경우 31.00부터 97.50의 점수 분포를 보였으며, 평균은 72.69, 표준편차는 10.67로 나타났다. 평생학습능력의 경우 23.00에서 100.00의 점수 분포를 보였으며, 평균은 69.99, 표준편차는 13.00으로 나타났다. 과학적 의사소통능력의 경우 23.50에서 90.00의 점수 분포를 보였으며, 평균은 65.79, 표준편차는 10.17로 나타났다. 과학적 탐구능력의 경우 30.00에서 92.00의 점수 분포를 보였으며, 평균은 67.00, 표준편차는 10.95로 나타났다.

각 영역별 총합인 400점 만점의 전체 결과를 살펴보면 112.50에서 372.50의 점수 분포를 보였으며, 평균은 275.47, 표준편차는 41.43으로 나타났다. 또한 각 영역 및 전체 문항에 대한 왜도와 첨도를 살펴보면 ±2.00의 적합 기준을 만족하는 것으로 나타났다(배정민, 2012; 신건권, 2016).

(2) 확인적 요인분석

본 연구를 통하여 개발된 TCSC에 대한 확인적 요인분석(CFA: Confirmatory Factor Analysis)을 시행하기 위하여 TCSC의 구조모형(structural model)을 구

Table 1. The results of statistics

N=278

항 목	Min	Max	M	SD	왜도	첨도
과학적 태도 및 정서	31.00	97.50	72.69	10.67	-.34	.52
평생학습능력	23.00	100.00	69.99	13.00	-.16	.21
과학적 의사소통능력	23.50	90.00	65.79	10.17	-.54	1.18
과학적 탐구능력	30.00	92.00	67.00	10.95	-.20	.01
전체	112.50	372.50	275.47	41.43	-.28	.48

성하였으며, TCSC의 특성에 맞게 2차 측정모형(second-order measurement model)으로 구성하고 구성타당도와 모형적합도 분석을 시행하였다. 2차 측정모형을 사용한 이유는 TCSC의 경우 ‘과학적 태도 및 정서’, ‘평생학습능력’, ‘과학적 의사소통능력’, ‘과학적 탐구능력’ 등의 하위 요소로 구성되어 있으므로 2차 측정모형이 개념을 더 명확하게 파악할 수 있기 때문이며, 이는 연구자의 판단 및 이론적 근거를 바탕으로 2차 측정모형이 특정 개념을 보다 정확하게 연결할 수 있을 때 2차 측정모형을 사용할 수 있기 때문이다(신건권, 2016).

특히 2차 측정모형의 경우 확인적 요인분석의 결과 구성타당도와 모형적합도가 인정된다면 각각의 1차 개념들은 관측변수로 변환하여도 연구 결과는 왜곡되지 않으며, 이 경우 잠재변수인 1차 개념들을 구성하는 측정변수들을 합하여 단일 지표로 만드는 항목합산(item parceling)을 하여 관측변수인 측정변수로 변환한다. 이러한 과정은 구조모형의 간명도(parsimony)를 높일 수 있으며(신건권, 2016), 연구자의 판단과 이론적 배경을 바탕으로 과학핵심역량의 요소를 선별하고 유목화하는 과정을 통해 하위 문항을 구성한 본 연구의 취지를 가장 잘 살리면서도 효율적인 방법이라고 할 수 있다.

따라서 본 연구에서는 TCSC의 구조모형을 Fig. 3의 좌측 모형과 같이 2차 측정모형으로 구성하였으며, 이를 바탕으로 1차 잠재변수의 측정변수들을 항목합산하여 간명화 작업을 시행하고 Fig. 3의 우측 모형과 같이 간명화된 과학핵심역량평가 구조모형으로 구성타당도와 모형적합도 점검을 하였다.

구안된 구조모형의 구성타당도(construct validity)는 표준화회귀계수(standardized regression weight)와 평균분산추출지수(AVE: Average Variance Extracted) 그리고 개념신뢰도(CR: Composite Reliability)의 측정치로 평가하며, 이론을 기반으로 하는 사회과학 분야의 연구 및 보고서의 경우 세 가지 중 한 가지 지수만을 계산하여 제시하여도 가능하기 때문에 표준화회귀계수를 산출하였다(노경섭, 2015; 신건권, 2016; Arbuckle, 2016).

그 결과 Fig. 2의 우측 그림에 표시된 것과 같이 과학핵심역량평가의 네 가지 영역 모두 표준화회귀계수의 판단 기준인 .7 이상인 것으로 나타나 기본적인 구성타당도가 충족되었다고 할 수 있다.

모형적합도의 경우 이를 판단하기 위한 판단 지수를 선정하기 위하여 문헌 및 연구 자료들에 대한 검토를 실시하였고(김대업, 2008; 노경섭, 2015; 문수백, 2009; 신건권, 2016; Arbuckle, 2016), Table 2와 같이 절대적합도 지수(CMIN(χ^2), GFI, SRMR), 증분적합도 지수(NFI, IFI, CFI), 간명적합도 지수(PGFI)를 선정하여 그 결과를 제시하였다. 위와 같이 다양한 판단 지수를 활용한 것은 모형의 적합도 판단에 있어서 절대적인 지수는 없기 때문에 연구자는 여러 지수들을 종합적으로 고려하여 모형의 적합성을 판단해야 하기 때문이다.

절대적합도지수(absolute fit index)는 연구모형의 전반적인 적합도를 절대적으로 평가하는 지수로 CMIN(χ^2), GFI, SRMR 등이 있다. CMIN(χ^2)은 표본의 수와 표본모형오차의 곱으로 산출한 기본 통계량으로 구조방정식이 개발된 초기부터 가장 많이

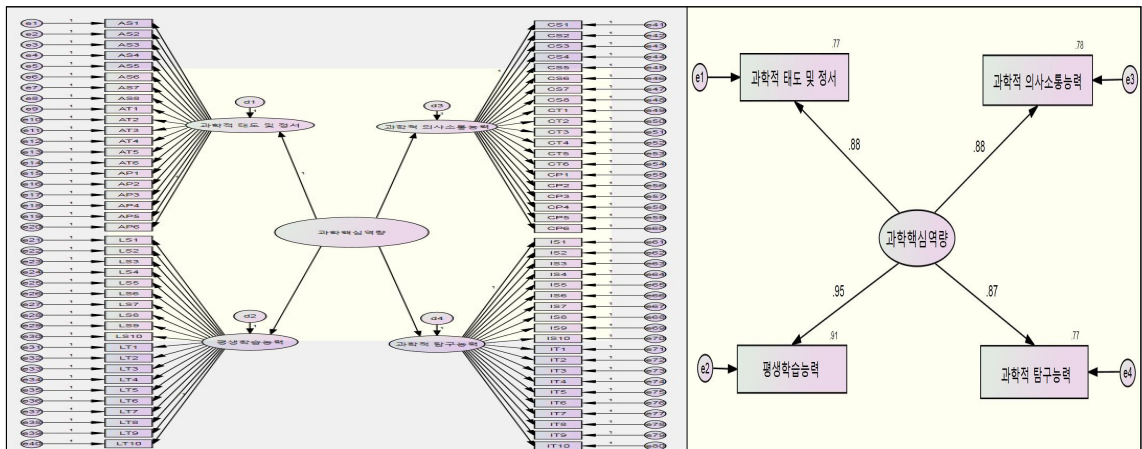


Fig. 3. Structural model

쓰이던 방법이지만 표본의 수가 많을수록 불리하다는 단점이 있어, 표본의 수에 민감하게 반응할 수 있기 때문에 $CMIN(\chi^2)$ 만으로 적합도를 판단하는 것은 바람직하지 않다. 따라서 $CMIN(\chi^2)$ 지수가 부적합으로 판정되더라도 다른 지수들과 함께 다각도로 고려하여 모형적합도를 판단하는 것이 필요하다고 할 수 있다. GFI(Goodness of Fit Index)는 기초부합 지수로 분석된 데이터와 원래 데이터 간의 차이를 나타낸 비율을 뜻하며, 표본의 크기에 영향을 받고, .9 이상이면 적합하다고 할 수 있다. SRMR(Standardized Root Mean Square Residual)은 표준잔차공분산지수로 된 데이터에 대한 분석 지수인 RMR(Root Mean Square Residual)을 표준화한 값이며, .05 이하이면 적합하다고 할 수 있다.

증분적합지수(incremental fit index)는 연구모형이 측정변수 간에 공분산이 존재하지 않는다고 가정하는 기초모형에 비해 어느 정도 향상되었는가 또는 얼마나 정확하게 측정하였는가를 평가하는 지수로 NFI, IFI, CFI 등이 있다. NFI(Normed Fit Index)는 연구모형을 기초모형과 비교한 지수로 .9 이상이며 적합하다고 할 수 있다. IFI(Incremental Fit Index)는 연구모형과 기초모형을 비교한 개선 정도는 나타내며, .9 이상이면 적합하다고 할 수 있다. CFI(Comparative Fit Index)는 비교부합지수로 모집단의 모수와 분포를 고려하여 NFI의 단점을 개선하기 위해 개발된 지수이며, .9 이상이면 적합하다고 할 수 있다.

간명적합도지수(Parsimonious Fit Index)는 모형의 간명성을 나타내는 지표로, 변수가 많아질수록 연구모형이 복잡해지면서 적합도 지수가 올라가는 것을 방지하기 위한 목적으로 개발된 지수이며, PGFI 등이 있다. PGFI (Parsimony Goodness of Fit Index)는 0에서 1의 범위를 가지며 수치가 높을수록 양호하다고 할 수 있다.

모형적합도 분석에 관한 결과를 종합하여 살펴보면, 절대적합도의 $CMIN(\chi^2)$ 의 경우 적합도 기준에 미치지 못하였지만 $CMIN(\chi^2)$ 의 특성상 표본의 수가 많을수록 불리하며, 표본의 수에 민감하게 반응하기 때문에 다른 지표들과 함께 고려하여 판단할 필요가 있었다. 절대적합도의 다른 지표들의 경우 GFI가 .93, SRMR이 .02로 적합한 결과를 보였다. 또한 증분적합도의 경우 NFI, IFI, CFI 모두 .96으로 분석되었고, 간명적합도의 PGFI의 경우 .19로 나타났다.

따라서 적합도 지수를 종합적으로 고려한 결과 과학핵심역량평가에 대한 연구모형이 적합성이 있는 것으로 판단되었으며, 초등학생들의 과학핵심역량을 평가하기 위한 검사 도구로서 적합한 모형이라고 할 수 있다.

(3) 공인타당도

최종적으로 구안된 TCSC의 공인타당도를 확인하기 위하여 활용도가 높으며 표준화된 검사 도구를 특정하기 어려운 평생학습능력을 제외한 과학적 태도 및 정서, 과학적 의사소통능력, 과학적 탐구능력의 세 영역에 대한 공인타당도 검증을 실행하였으며, 그 결과는 다음과 같다.

과학적 태도 및 정서 영역의 공인타당도를 살펴보기 위하여 현재 가장 많이 활용되고 있으며, 오랫동안 사용되고 있는 김효남 등(1998)의 과학적 태도 검사를 활용하고 그 결과를 분석하였다. 분석 결과 과학핵심역량평가의 과학적 태도 및 정서 영역 검사 결과와 과학적 태도 검사 결과의 상관관계수가 .56으로 파악되었으며, 상관관계가 있는 것으로 나타났다.

과학적 의사소통능력 영역의 공인타당도를 살펴보기 위하여 전성수(2013)의 과학적 의사소통능력 검사를 활용하여 상관관계를 파악하였으며, 그 결과 과학핵심역량평가의 과학적 의사소통능력 영역 검

Table 2. Fit index

항 목	부합도 지수	허용범위	적합성 여부	
절대적합도	$\chi^2(p)$	48.09(.000)	$\chi^2(p)$ 가 .05 이상 ($p>.05$)	부적합
	GFI	.93	.9 이상	적합
	SRMR	.02	.05 이하	적합
증분적합도	NFI	.96	.9 이상	적합
	IFI	.96	.9 이상	적합
	CFI	.96	.9 이상	적합
간명적합도	PGFI	.19	0~1	적합

사 결과와 전성수(2013)의 과학적 의사소통능력 검사 결과의 상관관계수가 .51로 나타났으며, 상관관계가 있는 것으로 파악되었다.

과학적 탐구능력 영역의 공인타당도를 살펴보기 위하여 권재술과 김범기(1994)의 과학탐구능력 검사를 활용하여 상관관계를 파악하였으며, 그 결과 과학핵심역량평가의 과학적 탐구능력 영역 검사 결과와 권재술과 김범기(1994)의 과학탐구능력 검사 결과의 상관관계수가 .41로 나타났으며, 상관관계가 있는 것으로 파악되었다.

따라서 많이 활용되거나 공인된 문항을 특정할 수 없는 평생학습능력을 제외한 과학적 태도 및 정서, 과학적 의사소통능력, 과학적 탐구능력의 세 영역에 대한 공인타당도 검증을 실행한 결과 세 영역 모두 활용도가 높았던 표준화 검사들과의 상관관계가 있는 것으로 나타나 공인타당도가 확보되었다고 판단하였다.

(4) 최종 검사 문항의 신뢰도

최종검사 문항의 신뢰도를 분석하여 살펴본 결과 Table 3과 같이 전체문항의 신뢰도는 .97로 매우 높은 신뢰도를 보이는 것으로 분석되었으며, 각 영역별 신뢰도는 과학적 태도 및 정서(.89), 평생학습능력(.92), 과학적 의사소통능력(.87), 과학적 탐구능력(.89)의 네 영역 모두 신뢰도가 매우 높게 나온 것으로 파악되었다.

Table 3. Reliability index

항 목	Cronbach' α
과학적 태도 및 정서	.89
평생학습능력	.92
과학적 의사소통능력	.87
과학적 탐구능력	.89
전 체	.97

2. 과학핵심역량평가(TCSC)의 활용

본 연구에서는 TCSC를 개발하기 위하여 핵심역량에 대한 개념적 정의에 관한 연구에서부터 출발하여, 핵심역량과 과학핵심역량에 대한 분석 및 유목화, 과학핵심역량 모델 구안, 과학핵심역량중심 수업모형 개발을 진행한 채희인(2017)의 사전 연구를 바탕으로 과학핵심역량평가 예비 문항 개발 및

적용을 통한 통계적 분석의 과정을 통하여 최종적인 TCSC를 개발하였다.

최종적인 검사 문항은 자기평가 문항, 교사평가 문항, 동료평가 문항으로 구성되어 있으며, 개발된 검사 도구가 과학핵심역량중심 수업모형과 함께 초등학교 현장에 적용되어 학생들의 핵심역량 함양에 도움이 될 수 있도록 다음과 같은 활용방안을 제시하였다.

1) 문항 구성

TCSC는 학생의 과학핵심역량 함양 정도를 평가하기 위하여 사전 연구를 통해 제안된 과학핵심역량 모델의 대표적인 요소인 과학적 태도 및 정서, 평생학습능력, 과학적 의사소통능력 그리고 과학적 탐구능력을 측정하기 위한 목적으로 개발된 평가이다.

이를 종합적으로 측정하기 위하여 본 평가에서는 여러 평가 주체에 의한 다양한 방식의 평가를 구성하였으며, 각 영역별 세부 구성 요소와 학생, 교사, 동료평가의 문항 번호를 포함하는 문항 체계표는 Table 4와 같다.

2) 검사 방법 및 시간

TCSC는 초등학교 4~6학년 학생들의 과학핵심역량의 함양 정도를 평가하기 위한 검사 도구로 이를 효과적으로 실행하기 위하여 교사는 평가를 이해하고 방법을 숙지할 필요가 있다. 특히 학생의 자기평가뿐만 아니라 동료 학생 3명에 의한 평가가 이루어진다는 점에서 학생과 과학 수업에 함께 참여한 모둠 단위로 평가하는 것이 효과적이라고 할 수 있다.

검사 시간의 경우 학생평가 문항의 경우 40분, 동료평가 문항의 경우 20분으로 총 60분이며, 최초 평가의 경우 학생들이 동료평가 방법에 대한 이해가 부족할 수 있으므로 충분히 설명해주는 시간이 필요하다 할 수 있다. 또한 교사 평가의 경우 지속적인 지도와 관찰 과정에서 수시로 평가를 시행하여 교수·학습과 평가가 효율적으로 이루어질 수 있도록 한다.

3) 배점 및 총점

TCSC의 문항별 배점은 5점이며, 과학적 태도 및 정서, 평생학습능력, 과학적 의사소통능력, 과학적 탐구능력의 영역별 점수는 100점이고, 총 400점으로

Table 4. The composition of TCSC

영역	하위 요소	학생(S)	교사(T)	동료(P)
과학적 태도 및 정서 (A)	호기심	AS1	AT1	AP1
	즐거움	AS2	AT2	AP2
	객관성	AS3	AT3	AP3
	수용성	AS4	AT4	AP4
	끈기	AS5	AT5	AP5
	개방성	AS6	AT6	AP6
	만족	AS7		
	자신감	AS8		
평생학습능력 (L)	자진성	LS1	LT1	
	주도성	LS2	LT2	
	자기이해	LS3	LT3	
	자기관리	LS4	LT4	
	미래지향	LS5	LT5	
	인식	LS6	LT6	
	열정	LS7	LT7	
	기초학습능력	LS8	LT8	
	학습자원 파악	LS9	LT9	
	공동체 활용	LS10	LT10	
과학적 의사소통능력 (C)	경청	CS1	CT1	CP1
	명료성	CS2	CT2	CP2
	협력적 대화	CS3	CT3	CP3
	설득	CS4	CT4	CP4
	이해 및 적용	CS5		
	설명	CS6		
	표와 그래프 이해	CS7		
	시각화	CS8		
	판단		CT5	CP5
	의견조정		CT6	CP6
과학적 탐구능력 (I)	관찰	IS1	IT1	
	분류	IS2	IT2	
	측정	IS3	IT3	
	추리	IS4	IT4	
	예상	IS5	IT5	
	가설설정	IS6	IT6	
	변인통제	IS7	IT7	
	자료변환	IS8	IT8	
	자료해석	IS9	IT9	
	결론도출	IS10	IT10	

구성하였다. 과학적 태도 및 정서와 과학적 의사소통능력의 동료평가 점수의 경우 학생 3명의 점수를 1/3하여 합산하였다.

4) 검사 점수의 평정기준

TCSC의 결과를 교육평가 본연의 목적에 맞게 학생의 과학핵심역량 함양 정도를 파악하고, 이를 바탕으로 동기 유발 및 자기주도적 탐구학습을 촉진

하며, 진로지도와 교수·학습의 개선에 보다 효과적이고 효율적으로 활용될 수 있도록 본 검사 결과를 바탕으로 표준 점수를 활용한 평정기준을 제시하고자 하였다.

이를 위해 TCSC 결과가 정규분포의 형태를 나타내는지 확인할 필요가 있으므로 기초통계량을 바탕으로 최저·최고 점수의 포함범위인 110점부터 380점까지 10점 단위로 빈도를 추출하여 과학핵심역량 평가 결과의 히스토그램을 작성하였으며, Fig. 4와 같이 정규분포 곡선과 비교하였다. 먼저 278명의 초등학생을 대상으로 시행한 과학핵심역량평가의 본 검사 결과는 평균이 275.47, 표준편차가 41.43이었으며, 최솟값은 112.50, 최댓값은 372.50이었다. 또한 왜도와 첨도가 -.28과 .48로 기술통계량을 통해서도 해당 점수 분포가 정규분포의 형태를 유추할 수 있는 기준인 ± 2.00 의 범위를 만족한 것으로 나타났다.

TCSC 결과에 대한 히스토그램과 정규분포 곡선과의 비교를 통하여 보다 세부적으로 살펴보면 과학핵심역량평가 결과가 정규분포 곡선과 거의 유사한 형태를 보이는 것으로 나타났다. 따라서 과학핵심역량평가의 결과 분포는 정규분포를 따르고 있다고 할 수 있다.

평정기준의 제시를 위한 판단 지표는 성태제(2004)의 Z점수와 규준점수들 간의 관계를 활용하였다. 평정기준의 제시를 위해 과학핵심역량평가의 결과를 평균이 0, 표준편차가 1인 Z점수로 환산하여 전체 및 각 영역별 평정기준 점수를 산출한 결과 Table 5와 같았다.

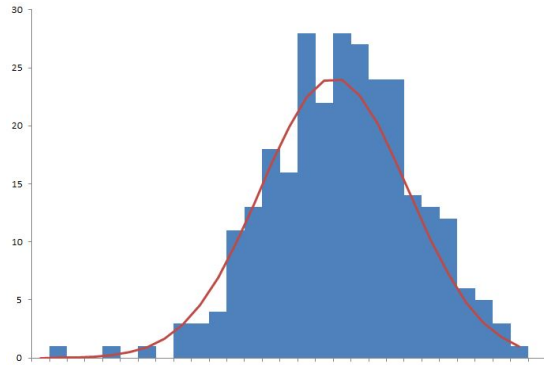


Fig. 4. The histogram of TCSC

TCSC의 총점을 기준으로 살펴보면 400점 만점 중 348점 이상은 ‘1. 극히 높은’ 단계, 327점 이상은 ‘2. 매우 높은’ 단계에 해당하는 것을 알 수 있었다. 또한 224점 미만은 ‘8. 매우 낮은’ 단계, 203점 미만은 ‘9. 극히 낮은’ 단계에 해당하였다.

영역별로 살펴보면 과학적 태도 및 정서의 경우 100점 만점 중 91점 이상은 ‘1. 극히 높은’ 단계, 86점 이상은 ‘2. 매우 높은’ 단계에 해당하는 것을 알 수 있었다. 또한 59점 미만은 ‘8. 매우 낮은’ 단계, 54점 미만은 ‘9. 극히 낮은’ 단계에 해당하였다.

평생학습능력의 경우 100점 만점 중 93점 이상은 ‘1. 극히 높은’ 단계, 86점 이상은 ‘2. 매우 높은’ 단계에 해당하는 것을 알 수 있었다. 또한 54점 미만은 ‘8. 매우 낮은’ 단계, 47점 미만은 ‘9. 극히 낮은’ 단계에 해당하였다.

Table 5. The criterion score of TCSC

평정	평정기준 점수				
	과학적 태도 및 정서	평생학습능력	과학적 의사소통능력	과학적 탐구능력	전체
1	91	93	84	86	348
2	86	86	79	81	327
3	81	80	73	75	307
4	75	73	68	70	286
5	70	67	63	64	265
6	65	60	58	59	244
7	59	54	53	53	224
8	54	47	48	48	203
9					

과학적 의사소통능력의 경우 100점 만점 중 84점 이상은 '1. 극히 높은' 단계, 79점 이상은 '2. 매우 높은' 단계에 해당하는 것을 알 수 있었다. 또한 53점 미만은 '8. 매우 낮은' 단계, 48점 미만은 '9. 극히 낮은' 단계에 해당하였다.

과학적 탐구능력의 경우 100점 만점 중 86점 이상은 '1. 극히 높은' 단계, 81점 이상은 '2. 매우 높은' 단계에 해당하는 것을 알 수 있었다. 또한 53점 미만은 '8. 매우 낮은' 단계, 48점 미만은 '9. 극히 낮은' 단계에 해당하였다.

본 연구에서는 위와 같이 과학핵심역량평가의 점수만으로 학생의 과학핵심역량 함양 정도와 상대적인 위치를 파악할 수 있도록 평정기준을 제시하여 손쉽게 활용할 수 있도록 하였다.

IV. 결론 및 제언

본 연구는 2015 개정 교육과정이 도입된 이후로 우리나라 교육이 역량중심으로 전환되었음에도 불구하고 핵심역량을 평가할 수 있는 정확한 도구나 지침이 부족하여 학교 현장에서 평가에 대한 혼란이 많으며, 기존에 이루어지던 평가 방식으로만 진행되어 진정한 의미의 역량평가가 이루어지지 못하고 있는 현실을 개선하는데 시사점을 주고자 하였다.

이를 위해 평가와 관련된 사전 연구를 바탕으로 조직화의 원리, 다차원의 원리, 다양성의 원리, 지속성의 원리, 도구성의 원리와 같은 다섯 가지의 평가 원리를 설정하고 과학핵심역량을 평가할 수 있는 문항을 개발하고자 하였다. 평가도구 개발의 과정에서는 과학교육 전문가와 초등학교 교사 5명의 검토를 바탕으로 초안을 마련하였으며, 과학교육을 전공한 교사 3명의 내용타당도 검토 및 1차(52명)와 2차(54명) 난이도 및 변별도 검사를 시행하였고, 25명의 학생을 대상으로 예비 검사를 시행하여 내적 신뢰도 분석을 하였다. 최종확정된 문항은 4~6학년 278명을 대상으로 본 검사를 진행하였으며, 7명의 교사가 평가자로 참여하였다. 이 과정에서 기초통계량 분석(평균, 표준편차, 최댓값, 최솟값, 왜도, 첨도)과 2차 측정모형을 구안하여 확인적 요인분석(구성타당도-표준화회귀계수, 모형적합도-정대적합도, 증분적합도, 간명적합도 지수)을 시행하였다. 마지막으로 공인타당도 분석과 신뢰도 분석을 하였다. 이를 바탕으로 과학핵심역량평가의 활용 방법을 제시하기 위

하여 문항 체계표 및 Z점수로 환산한 평정기준을 제시하였다. 최종적으로 13명의 과학교육 전문가와 교사, 그리고 409명의 학생을 대상으로 평가도구 개발이 이루어졌다.

이와 같은 TCSC 개발을 바탕으로 본 연구에서는 다음과 같은 결론을 얻을 수 있었다.

첫째, TCSC는 기존에 핵심역량을 평가하기 위하여 개발된 많은 평가가 지금까지 학교 현장에서 계속 이루어졌던 지필과 보고서 위주의 수행평가를 기본으로 한다는 점과는 많은 차이가 있다고 할 수 있다. 특히 과학핵심역량평가는 교사평가와 더불어 학생의 자기평가와 탐구 수행을 함께한 동료평가를 함께 한다는 점에서도 기존의 평가와 많은 차이가 있다고 할 수 있다. 또한 과학핵심역량평가는 표준화된 평가도구로서 본 연구를 통해 4~6학년 278명의 과학핵심역량평가 지수를 산출한 결과 평균은 275.47이었고, 표준편차는 41.43이었다. 특히 학생들은 400점 만점 중 112.50에서 372.50의 점수 분포를 보였으며, 과학적 태도 및 정서 영역의 경우 평균과 표준편차가 72.67점(10.67)이었으며, 평생학습능력의 경우 69.99점(13.00), 과학적 의사소통능력의 경우 65.79점(10.17), 과학적 탐구능력의 경우 67.00점(10.95)이었다. 이와 같은 결과를 바탕으로 Z점수로 환산하여 평정기준 점수를 산출하였으며, 348점 이상은 '1. 극히 높은' 단계에 해당하는 것으로 파악되었다.

둘째, TCSC는 기존의 평가에서 측정하기 어려운 부분에 대한 평가가 가능한 것으로 나타났다. 특히 기존에 학교 현장에서 이루어지던 과학학습평가는 교사평가 위주로 이루어지면서 모든 평가의 부담과 관점이 교사 중심으로 이루어져 왔다. 즉 교사는 과학 교육 평가 활동에 어려움을 갖고 있었다(김성기 등, 2021). 하지만 과학핵심역량평가는 평가의 주체를 교사평가, 자기평가, 동료평가로 분산함에 따라 다양한 관점에서 평가가 이루어질 수 있었으며, 교사가 관찰하지 못하는 부분에 대한 평가도 가능한 것으로 나타났다.

셋째, TCSC는 컴퓨터 기반 평가로 개발되었다. 그 이유는 기존에 실험보고서 및 지필 평가 결과물 중심으로 이루어지는 평가의 경우 평가의 주체가 교사 한 명이며, 교수·학습에 대한 부담으로 인해 평가가 단순하게 이루어지지 않는다면 교사는 과도한 부담으로 인해 제대로 된 평가를 할 수 없으므로 최대한 단순하게 이루어져 온 것이 현실이다. 하지

만 과학핵심역량평가는 교사의 평가 부담을 줄이면서도 다양한 측면에서 최대한 많은 정보를 바탕으로 평가를 할 수 있도록 컴퓨터기반으로 제작되었다. 특히 학년, 반, 번호를 학생 고유의 개인 ID로 활용하여 평가를 하면 모든 데이터가 종합되어 최종적인 결과물을 얻을 수 있기 때문에 기존의 방식보다 간편하면서 보다 심층적이고 다양한 평가가 가능하다는 장점이 있다. 따라서 교사의 평가 업무 부담을 줄여주어야 한다는 권영순과 하지훈(2020)의 연구 결과에 부합하는 평가도구라고 할 수 있다.

넷째, 지금까지 학생의 역량을 측정하기 위한 평가도구 및 지표를 개발한 연구들은 대부분 과학핵심역량을 측정하기 위한 것이 아니라 일반적인 역량을 측정하기 위한 것이었다(권희경 등, 2020; 남궁지영 등, 2016). 또한 과학핵심역량을 측정하기 위한 평가도구의 경우에도 특정한 단원이나 탐구능력을 측정하기 위한 목적으로만 개발되었으며(김인숙 등, 2020; 박종임, 2021a), 표준화된 문항의 경우에도 설문을 위한 조사 문항에 머물러 있었다(하민수 등, 2018). 따라서 본 연구에서 개발한 TCSC의 경우 학생의 과학핵심역량의 평가를 목적으로 하는 표준화된 평가도구의 개발이라는 점에서 의의가 있다고 할 수 있다.

이와 같은 결론을 바탕으로 본 연구에서는 다음과 같은 제언을 하고자 한다. 본 연구가 수행되는 과정에도 기존의 컴퓨터 기반 평가에서 더 발전하여 인공지능기반 평가가 매우 빠르게 발전하고 있었다. 특히 ChatGPT로 대표되는 인공지능의 발전은 교육에도 많은 영향을 미치고 있는 것이 현실이다. 따라서 본 연구를 통해 개발된 TCSC를 보다 발전시켜 인공지능 기반의 과학핵심역량 평가도구 개발로 확장할 수 있을 것이다.

탐구 활동 과정에서 생성되는 수많은 언어·시각 데이터, 탐구 활동 결과물 및 상호작용과 관련된 실제적인 데이터는 효과적이고 효율적인 방법으로 과학핵심역량을 측정할 수 있을 것이라고 생각하며, 기존의 지필 형태의 평가와 컴퓨터를 단순히 도구적으로 활용하는 평가에서는 측정하기 어려운 학생의 실제적인 역량을 측정할 수 있을 것이라고 판단된다. 따라서 탐구 활동 중에 표현되는 학생의 다양한 긍정적·부정적 언어 사용은 ‘과학적 태도 및 정서역량’의 측정에 활용될 수 있으며, 모둠 탐구의 과정에서 생성되는 수많은 상호작용 데이터는 ‘과학적

의사소통능력’을 평가하는 실제적이고 중요한 자료이고, 탐구 실험의 장치를 바르게 설계하고 탐구 기기를 적절하게 활용하는 시각적 분석 데이터는 ‘과학적 탐구능력’ 평가에 활용될 수 있을 것이다. 또한 학생의 최종적인 실험보고서에 작성된 많은 글과 그림의 자료는 AI에 의한 즉각적인 판단을 통해 ‘과학적 지식학습능력’을 평가할 수 있고, 그에 따른 학생 개별적인 피드백 제공에 가장 효과적인 판단 근거가 될 수 있을 것이므로 이와 관련한 후속 연구들이 계속 이루어지길 바란다.

참고문헌

- 강남화, 김민지(2020). 고등학교 통합과학 수행평가 사례를 통해 탐색한 교사의 수행평가 실천 특성과 PCK 사이의 관련성. 한국과학교육학회지, 40(3), 291-305.
- 강훈식, 강석진(2015). 초등학교 교사의 과학 평가 전문성에 대한 자가진단 내용 분석. 초등과학교육, 34(2), 153-163.
- 곽영순, 김찬중, 이양락, 정득실(2006). 초, 중등 학생들의 과학 흥미도 조사. 한국지구과학회지, 27(3), 260-268.
- 곽영순, 하지훈(2020). 초등학교 과학 수업에서 과정중심평가 적용 실태 분석. 한국과학교육학회지, 40(5), 465-471.
- 교육부(2015a). 2015 개정 초등학교 교육과정(제2015-74호 [별책 1]). 세종: 교육부.
- 교육부(2015b). 2015 개정 과학과 교육과정(제2015-74호 [별책 9]). 세종: 교육부.
- 교육부(2022). 2022 개정 과학과 교육과정(제2022-33호 [별책 9]). 세종: 교육부.
- 구본용, 구혜영, 이명우(1994). 또래상담자 훈련프로그램 개발연구. 서울: 청소년 대화의 광장.
- 권재술, 김범기(1994). 초·중학생들의 과학탐구능력 측정도구의 개발. 한국과학교육학회지, 14(3), 251-264.
- 권희경, 김주아, 박경호, 안혜정, 최인희, 김명섭(2020). 2020KEDI 학생역량 조사 연구(RR2020-29). 인천: 한국교육개발원.
- 김경희(1997). 한국 아동의 정서에 관한 심리학적 연구. 한국심리학회지 발달, 10(1), 43-56.
- 김대업(2008). Amos A to Z 논문작성절차에 따른 구조방정식 모형분석. 파주: 학현사.
- 김동현(2014). 초등학생 과학 학습정서 검사 도구 개발. 한국교원대학교 대학원 박사학위논문.
- 김민환, 박다혜, 노태희(2022). 과학탐구실험의 ‘역사 속의 과학 탐구’에서 과학교사의 평가 실태와 평가 지향 조사. 한국과학교육학회지, 42(6), 597-610.
- 김성기, 민희정, 백성혜(2020). 학습공동체를 통한 과학교

- 사의 평가에 대한 신념에 관한 연구. *교원교육*, 36(1), 1-21.
- 김성기, 민희정, 하민수, 백성혜(2021). 과학교사의 수행평가에 대한 평가 전문성 탐색. *교육과정평가연구*, 24(1), 231-151.
- 김유진(2015). 수능 마친 고3 수험생 “수능 점수로 대학가는게 복잡한 고차방정식 같아요”. <http://m.news.naver.com/read.nhn?mode=LSD&sid1=102&sid2=256&oid=014&aid=0003547530>
- 김은숙, 김현경(2021). 중학교 과학 학업성취도 평가 개선 방안 연구: 국가수준 학업성취도 평가 결과를 중심으로. *학습자중심교과교육연구*, 21(6), 151-162.
- 김인숙, 이문복, 박종임, 송민호, 김성기, 곽민호, 김성숙, 김창환, 임은진(2020). 핵심역량 평가를 위한 컴퓨터 기반 평가 시스템(KICE-eAssessment) 개발 연구(Ⅱ). 진천: 한국교육과정평가원.
- 김지자, 김경성, 유귀옥, 유길한(1996). 초등학교 교사를 위한 자기주도학습 준비도 측정도구의 개발과 활용방안. *사회교육학연구*, 2(1), 1-25.
- 김진화, 전은선, 박선경(2014). 직장인의 평생학습역량에 대한 탐색적 실증 연구. *농업교육과 인적자원개발*, 46(1), 181-205.
- 김현정, 김성기(2021). 탐구보고서에 기반한 화학교사의 과학 역량 평가 실태 분석. *대한화학회지*, 65(3), 209-218.
- 김현주(2015). 성적이 ‘인생의 등급’ 결정한다고?. <http://m.media.daum.net/m/media/society/newsview/20150719050508130>
- 김효남, 정완호, 정진우(1998). 국가수준의 과학에 관련된 정의적 특성의 평가체제 개발. *한국과학교육학회지*, 18(3), 357-369.
- 남궁지영, 김양분, 박경호, 최인희, 박현정(2016). 2016 KEDI 학생역량 조사 연구(RR2016-13). 진천: 한국교육개발원.
- 노경섭(2015). 제대로 알고 쓰는 논문 통계분석. 서울: 한빛아카데미.
- 노태희, 이재원, 강서진, 강훈식(2015). 중등 과학교사의 과학 평가 실태와 지향. *한국과학교육학회지*, 35(4), 725-733.
- 문수백(2009). 구조방정식 모델링의 이해와 적용. 서울: 학지사.
- 박도순, 권순달, 김명화, 김석우, 김영애, 김옥남, 김종필, 김진규, 김진하, 김형주, 남명호, 남현우, 박완성, 백순근, 성태제, 양길석, 어윤경, 이기중, 이나현, 이원석, 이종승, 이현주, 장은아, 조지민, 최대호(2007). *교육평가: 이해와 적용*. 서울: 교육과학사.
- 박영석(2009). 초등학교 사회 수업에서 교사의 형성평가 실행에 관한 연구. *시민교육연구*, 41(4), 51-79.
- 박종임, 이문복, 송민호, 김광규, 한정아(2021a). 핵심역량 평가를 위한 컴퓨터 기반 평가 시스템(KICE-eAssessment) 개발 연구(Ⅲ). 진천: 한국교육과정평가원.
- 박종임, 이문복, 송민호, 김광규, 한정아(2021b). 핵심역량 평가를 위한 컴퓨터 기반 평가도구 개발 및 현장 적용 방안(한국교육과정평가원 이슈페이퍼 vol. 5). 진천: 한국교육과정평가원.
- 배정민(2012). 그림으로 이해하는 닥터 배의 술술 보건의 학통계. 서울: 한나래 아카데미.
- 백수정, 이희수(2012). 성인교육 퍼실리테이터 역량 개발: 필요 수준과 현재 수준 간의 차이 분석을 중심으로. *평생학습사회*, 8(3), 59-85.
- 성태제(2002). 타당도와 신뢰도. 서울: 학지사.
- 성태제(2004). 현대 기초통계학의 이해와 적용. 서울: 교육과학사.
- 신건권(2016). *Amos23 통계분석 따라하기*. 서울: 청람.
- 신정운, 박상우(2021). 초등 과학과 과정중심 평가문항 예시 자료의 특징. *초등과학교육*, 40(2), 191-209.
- 심보경, 유미현(2020). 2015 개정 교육과정 통합과학 및 과학탐구실험의 수행평가에 반영된 과학과 핵심역량 분석. *현장과학교육*, 14(4), 481-500.
- 양일호, 나종철, 임성만, 임재근, 최현동(2008). Klopfer의 교육 목표 분류 체계에 의한 초등학교 과학과 지필평가 문항 분석: 5학년 1학기를 중심으로. *초등과학교육*, 27(3), 221-232.
- 여상인, 박애자, 임희준(2007). 초등 과학수업에서 학습자 중심의 수행평가준거 설정이 학습동기, 학습몰입, 학업성취도에 미치는 효과. *초등교육연구*, 20(1), 395-412.
- 오현석, 이기영(2006). 현행 중등학교 과학 교과 지필평가 문항분석. *교육과정평가연구*, 9(1), 405-424.
- 유현숙, 김태준, 이석재, 송선영(2004). 국가수준의 생애능력 표준설정 및 학습체제 질 관리 방안 연구(Ⅲ). 서울: 한국교육개발원.
- 이석재, 장유경, 이현남, 박광엽(2003). 생애능력 측정도구 개발 연구: 의사소통능력, 문제해결능력, 자기주도적 학습능력을 중심으로. 서울: 한국교육개발원.
- 이현주, 최윤희, 고연주(2015). 집단지성을 강조한 과학기술 관련 사회쟁점 수업이 중학교 영재학급 학생들의 역량 함양에 미치는 효과. *한국과학교육학회지*, 35(3), 431-442.
- 전성수(2013). 초등학생의 과학적 의사소통능력 검사도구 개발. *한국교원대학교 대학원 박사학위논문*.
- 전성수, 박종호(2014). 초등학생의 과학적 의사소통능력 과 과학 탐구능력, 논리적 사고력, 학업 성취도 수준과의 관계 분석. *한국과학교육학회지*, 34(7), 647-655.
- 정은영, 최원호(2014). 우리나라 초·중·고등학교 과학과의 평가 실태. *과학교육연구지*, 38(1), 168-181.
- 정종오(2015). 노벨물리학상 나카무라 “아시아교육은 시간 낭비”. <http://view.asia.co.kr/news/view.htm?idxno=>

20150120 09285961944

- 조준형(2015). 노벨상 6명 배출 미국토대 확장 “영어실력은 도구 중 하나일 뿐, 중요한 건 사고력 배양, 정해진 답 없는 문제 풀 능력 키워야”. http://m.media.daum.net/m/media/world/new_sview/20151021180024633
- 채수진(2006). 의과대학 문제중심학습에서 학생평가방법의 일관성 검토: 튜터평가, 동료평가, 자기평가를 중심으로. *열린교육연구*, 14(1), 331-349.
- 채희인(2017). 과학핵심역량의 재구조화에 기반한 초등과학 수업전략 개발 및 적용. 경인교육대학교 대학원 박사학위논문.
- 채희인, 노석구(2015). 범교과적 핵심역량과 과학핵심역량의 개념적 정의에 대한 국내외 연구 동향의 네트워크 분석. *학습자중심교과교육연구*, 15(7), 23-40.
- 하민수, 박현주, 김용진, 강남화, 오필석, 김미점, 민재식, 이윤형, 한효정, 김무경, 고성우, 손미현(2018). 2015 개정 과학과 교육과정에 기초한 과학과 핵심역량 조사 문항의 개발 및 적용. *한국과학교육학회지*, 38(4), 495-504.
- 허명(1984). 과학탐구 평가표의 개발. *한국과학교육학회지*, 4(2), 57-63.
- AAAS. (1989). *Science for all Americans*. Washington D.C.: AAAS.
- Arbuckle, J. (2016). *IBM SPSS Amos user's guide*. Chicago, IL: IBM Amos Development Corporation.
- Ballantyne, R., Hughies, K., & Mylonas, A. (2002). Developing procedures for implementing peer assessment in large classes using an action research process. *Assessment and Evaluation in Higher Education*, 27(5), 427-441.
- Bienvenu, M. J., Sr., & Stewart, D. W. (1976). Dimensions of interpersonal communication. *Journal of Psychology*, 65, 60-66.
- 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.
- Cangelosi, J. (1990). *Designing tests for evaluating student achievement*. New York: Longman.
- Ebel, R. L. (1965). *Measuring educational achievement*. Englewood Cliffs, N. J.: Prentice-Hall.
- Falchikov, N. (1995). Peer feedback marking: Developing peer assessment. *Innovations in Education and Training International*, 32(2), 175-187.
- Fehring, R. J. (1987). Methods to validate nursing diagnoses. *Heart and Lung*, 16(6), 625-629.
- Fraser, B. J. (1981). TOSRA: Test of science related attitudes. Hawthorne, VIC: Australian Council for Educational Research.
- Griffin, F., Care, E., & McGaw, B. (2012). The changing role of education and schools. In F. Griffin, B. McGaw & E. Care (Eds.), *Assessment and teaching of 21st century skills* (pp. 1-15). New York: Springer.
- Guglielmino, L. M. (1977). *Development of the Self-Directed learning readiness scale* (Unpublished doctoral dissertation). University of Georgia, US.
- Klopfer, L. E. (1971). *Evaluation of learning in science*. New York: McGraw-Hill, Inc.
- Macklin, T. T., & Rossiter, C. M. (1976). Interpersonal communication and self-actualization. *Communication Quarterly*, 24, 45-50.
- Misitomu, A. I. (2012). *An exploration of year 10 students' motivation, attitude and self-efficacy toward learning chemistry* (Unpublished doctoral dissertation). University of Waikato, New Zealand.
- NAEP. (1984). *Science objectives: Fifth national assessment 3rd draft*. Washington, D.C.: NAEP.
- Padilla, M. J., Okey J. R., & Dillashaw F. G. (1983). The relationship between science process skill and formal thinking abilities. *Journal of Research in Science Teaching*, 20(3), 239-246.
- Roseman, I. J., Spindel, M. S., & Jose, P. E. (1990). Appraisals of emotion-eliciting events: Testing a theory of discrete emotions. *Journal of Personality and Social Psychology*, 59(5), 899-915.
- Ruben, B. D. (1976). Assessing communication competency for intercultural adaptation. *Group and Organization Studies*, 1, 334-354.
- 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.
- Topping, K., & Ehly, S. (1998). *Peer assisted learning*. Mahwah: Lawrence Erlbaum Associates.
- Uzunboylu, H., & Hursen, C. (2011). Lifelong Learning Competence Scale (LLCS): The study of validity and reliability. *H. U. Journal of Education*, 41, 449-460.
- 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, 522-529.

채희인, 평택청아초등학교 교사(Heein Chae; Teacher, Cheonga Elementary School).

† 노석구, 경인교육대학교 교수(Sukgoo Noh; Professor, Gyeongin National University of Education).

[부록1] 과학핵심역량평가(TCSC)

학생평가 문항

과학핵심역량평가지

본 검사지는 여러분의 과학핵심역량을 측정하기 위한 도구입니다. 또한 본 검사는 시험이 아니기 때문에 결과는 학교 성적과 무관합니다. 각 문항을 잘 읽고 한 문제도 빼놓지 말고 기록해주시기 바랍니다.

()학년 ()반 ()번 남/여

이름 :

ID :

【주의점 및 작성 요령】

- * 하나도 빼놓지 말고 모든 질문에 √ 또는 ○표시를 해주시기 바랍니다.
- * 시험을 보는 것이 아니므로 편하게 해주시면 됩니다.
- * ID는 학년, 반, 번호를 연속해서 적으면 됩니다.
예) 4학년 5반 7번 -> 4507
예) 6학년 3반 14번 -> 6314

질문	정해 아니다 (1점)	아니다 (2점)	보통 이다 (3점)	그렇다 (4점)	매우 그렇다 (5점)
----	-------------------	-------------	------------------	-------------	-------------------

과학적 태도 및 정서

1. 나는 과학적으로 탐구할 수 있는 일들에 대해 호기심이 많다.					
2. 나는 과학 실험이나 과학과 관련된 활동을 통해 새로운 사실을 알아가는 것이 즐겁다.					
3. 나는 객관적인 자료와 사실을 바탕으로 과학 탐구활동의 결론을 내린다.					
4. 나는 다른 사람들의 실험 결과가 옳다고 판단되면 내 주장과 다르더라도 받아들일 수 있다.					
5. 나는 과학 실험을 끝까지 하려고 노력한다.					
6. 나는 과학 탐구활동을 통해서 새로운 발견을 할 수 있다고 생각한다.					
7. 나는 우리 주변의 문제를 과학적으로 탐구하는 활동에 만족을 느낀다.					
8. 나는 우리 주변의 문제를 과학적 탐구활동을 통해 해결할 수 있다는 자신감이 있다.					

평생학습능력

1. 나는 과학 실험이나 과학과 관련된 활동을 할 때 궁금한 것이 생기면 스스로 찾아본다.					
2. 나는 과학 탐구활동을 할 때 친구들을 이끌어 가며 주도적으로 하는 편이다.					
3. 나는 과학 실험이나 활동을 위해 어떤 지식이나 지원이 필요한지를 잘 알고 있다.					

질문	정해 아니다 (1점)	아니다 (2점)	보통 이다 (3점)	그렇다 (4점)	매우 그렇다 (5점)
4. 나는 탐구활동을 미리 계획하고, 얼마나 진행되었는지 파악할 수 있다.					
5. 나는 과학 탐구활동이 나의 발전에 도움이 될 것이라고 생각한다.					
6. 나는 과학 지식뿐만 아니라 과학을 탐구하는 방법을 배우는 것이 우리 일상생활의 과학적 문제를 해결하는데 도움이 될 것이라고 생각한다.					
7. 나는 지속적으로 과학 탐구활동을 하고 싶다는 생각과 열정을 갖고 있다.					
8. 나는 과학 탐구활동을 시작하기 위한 기초적인 준비가 되어 있다.					
9. 나는 과학 탐구활동에 필요한 정보 및 자료들이 무엇인지 알고 있다.					
10. 나는 과학 탐구활동에 필요한 선생님, 부모님 그리고 주변 사람들의 도움과 지원을 잘 활용할 수 있다.					

과학적 의사소통능력

1. 나는 과학 탐구활동을 할 때 다른 사람의 의견을 주의 깊게 듣는다.					
2. 나는 과학 탐구활동의 결과를 뚜렷하고 분명하게 이야기한다.					
3. 나는 과학 탐구활동을 할 때 다른 사람들과 협력적으로 이야기한다.					
4. 나는 서로의 주장이 다를 때 과학 탐구활동의 결과를 바탕으로 다른 사람을 설득할 수 있다.					

5. 시험이내 모듬은 과학시간에 물이 얼 때의 부피 변화를 알아보기 위한 실험을 하였습니다. 교과서에 제시된 내용과 학생들의 대화를 바탕으로 실험 장치를 알맞게 준비한 것을 하나 고르시오.....()

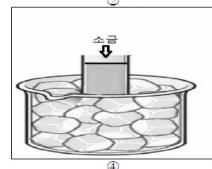
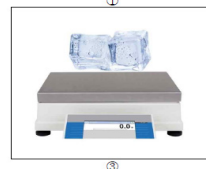
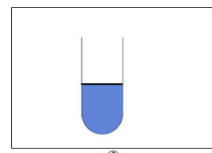
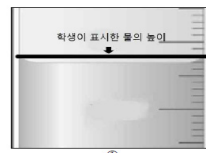
어떻게 할까요?

ㄱ. 플라스틱 시험관에 물을 한 정도 넣고, 물의 높이를 유성 펜으로 표시하자.

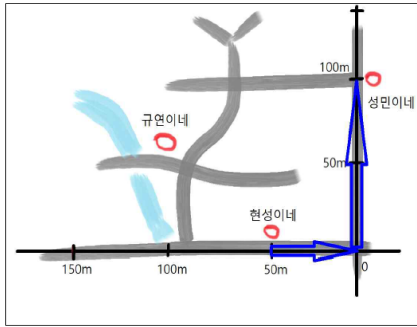
ㄴ. 비커에 잘게 부순 얼음과 소금을 넣고 잘 섞은 다음에 ㄱ의 플라스틱 시험관을 얼음이 든 비커에 꽂습니다.

* 얼음이 담긴 비커에 소금을 넣으면 시험관 속의 물이 빨리 얼어요.

시험: 플라스틱 시험관에 물을 한 정도 넣고 물의 높이를 유성 펜으로 표시하자.
시험: 그러니까, 시험관의 물을 자세히 보니까 물의 높이가 일직선으로 되어 있지 않아.
우현: 액체의 부피를 측정하기 위해서는 액체의 오목한 면과 눈금의 높이가 수평이 되도록 유지하고 읽어야해.
영재: 과학 준비실에서 전자저울을 가져와서 얼음이 얼마나 무거운지 재어볼까?
시험: 우리 실험은 물이 얼 때의 부피변화를 알아보려는 것이잖아.

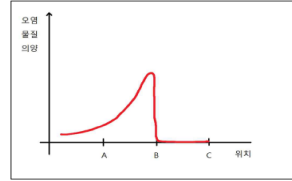


6. 보람초등학교의 스쿨버스는 현성이네 집에서 08시 18분에 출발해서 사거리로 08시 25분에 지나 성민이네 집에 08시 41분에 도착하였습니다. 다음 중 스쿨버스의 위치 변화를 바르게 설명한 것을 하나 고르시오.....()

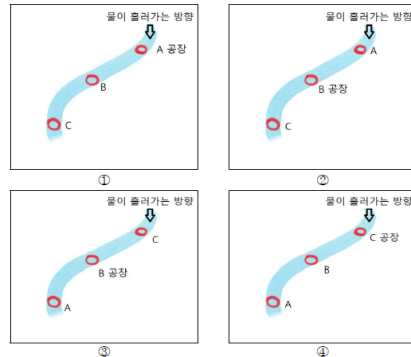


- ① 스쿨버스는 현성이네 집을 출발한 후 6분 뒤에 사거리에 도착하였다.
- ② 스쿨버스는 사거리에서 성민이네 집까지 100m를 7분 동안 이동하였다.
- ③ 스쿨버스는 현성이네 집에서 성민이네 집까지 150m를 23분 동안 이동하였다.
- ④ 스쿨버스는 현성이네 집에서 성민이네 집까지 200m를 16분 동안 이동하였다.

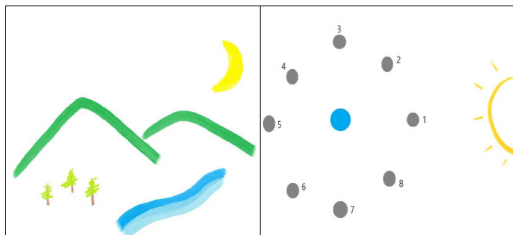
7. 유승이가 살고 있는 지역에서는 강의 오염으로 인해서 물고기가 죽고 심한 냄새가 나는 일에 대하여 회의를 하였습니다. 하지만 어디에서 이런 오염물질이 강으로 흘러 들어가는지 아무도 몰랐습니다. 그래서 유승이는 같은 학교 과학반 친구들과 함께 이 문제를 해결할 방법을 찾기 위하여 책을 찾아보다가 강물에 오염물질이 얼마나 많은지 확인하면 오염물질이 강으로 흘러들어가는 곳을 찾을 수 있다는 내용을 보고 친구들과 함께 탐구활동을 시작하여 다음과 같은 결과를 얻었습니다.



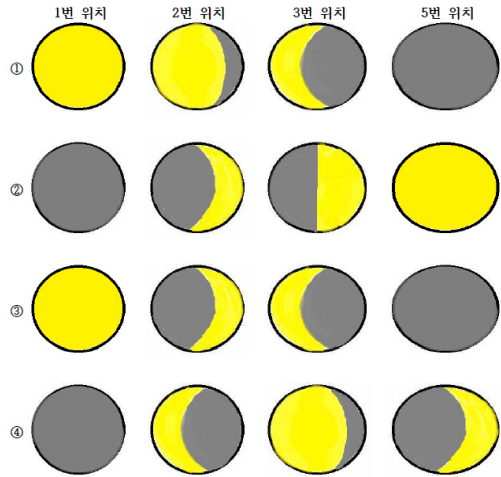
유승이네 학교 과학반 학생들이 얻은 결과를 바탕으로 오염물질이 강으로 흘러들어가는 공장의 위치를 그림으로 바르게 나타낸 것을 하나 고르시오.....()



8. 다인이와 유나는 어제 저녁에 공원을 산책하다가 배 모양처럼 생긴 달을 보았습니다. 이를 신기해하던 다인은 매일 달을 관찰하다가 달의 모양이 변하는 것을 알게 되었습니다. 다인이와 유나의 대화를 바탕으로 그림의 1, 2, 3, 5에 알맞은 달의 모양을 제시한 것을 하나 고르시오.....()



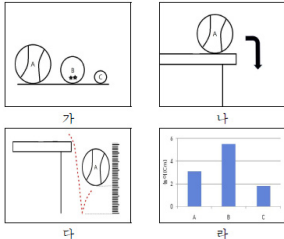
다인: 달은 빛을 스스로 내지 못하기 때문에 우리들이 보는 달빛은 태양 빛이 달에 반사된 거야. 그럼에서 빛을 내는 태양, 번호가 적혀있는 작은 원은 달. 달 가운데에 있는 큰 원은 지구(우리)라고 볼 수 있지.
 유나: 그래서 달이 그림과 같이 1번에서 8번까지 지구를 돌 때마다 우리는 태양 빛을 반사하는 달의 밝은 부분을 보기 때문에 달의 모양이 바뀌는 것이구나!
 다인: 맞아. 그래서 초승달, 상현달(반달), 보름달과 같이 모양이 다른 달을 볼 수 있는 거야.
 유나: 그렇다면 1번의 위치에 달이 있다면 달은 어떤 모양으로 보일까? 달에서 태양 빛을 반사하는 부분이 지구의 반대편에 있네.
 다인: 응, 5번의 위치에 달이 있으면 달에서 태양 빛을 반사하는 부분이 지구를 향하고 있겠네.
 유나: 그렇다면 1번, 2번, 3번, 5번의 위치에 달이 있을 때 각각 달은 어떤 모양으로 보일까?
 다인:



※ 노란색(밝은색)이 빛나는 부분

질문	전혀 아니다 (1점)	아니다 (2점)	보통이다 (3점)	그렇다 (4점)	매우 그렇다 (5점)
과학적 탐구능력					
1. 나는 관찰을 통하여 사물과 현상에 대한 정보를 얻어낼 수 있다.					
2. 나는 공통적인 특성과 조건을 바탕으로 사물을 나눌 수 있다.					
3. 나는 길이, 무게, 부피, 시간 등을 정확하게 잴 수 있다.					
4. 나는 관찰한 사실과 증거를 바탕으로 결과를 바르게 해석하고 설명할 수 있다.					
5. 나는 과학 탐구활동을 할 때 관찰한 사실을 바탕으로 앞으로 어떤 일이 일어날지 예상할 수 있다.					

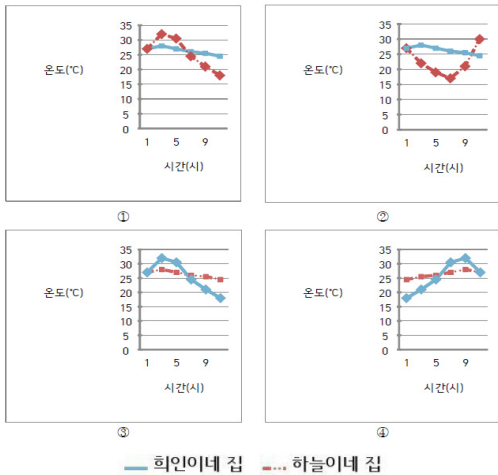
6. 다음은 진화가 실험한 장면을 순서대로 나타낸 그림입니다. 다음 그림을 보고 진화가 이 실험을 통하여 알아보고자 한 것은 무엇인지 가장 알맞은 것을 하나 고르시오.....()



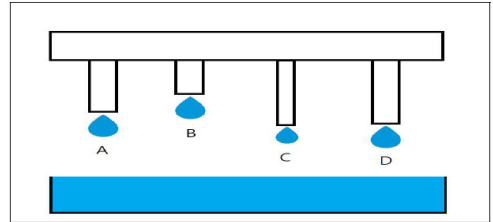
- ① 공의 종류에 따라서 멀리 날아가는 정도가 다를 것이다.
- ② 공을 떨어뜨리는 높이가 공의 종류가 다르면 튀어오르는 높이도 다를 것이다.
- ③ 공을 떨어뜨렸을 때 가벼운 공이 더 잘 튀어오를 것이다.
- ④ 공을 떨어뜨리는 높이가 같더라도 공의 종류에 따라 튀어오르는 높이는 다를 것이다.

8. 희인이는 하늘이는 집의 실내 온도를 비교하기 위하여 각자 오후 1시부터 11시까지 두 시간 마다 온도계를 사용하여 집의 온도를 측정하였습니다. 그 결과를 다음과 같이 표로 작성하였지만 시간에 따라 두 집의 온도 변화를 더 쉽게 파악하기 위해서 하늘이는 학교에서 배운 꺾은선 그래프를 그리기로 하였습니다. 오후 1시부터 11시까지 희인이나 하늘이네 집의 실내 온도를 그래프로 바르게 나타낸 것을 하나 고르시오.....()

	1시	3시	5시	7시	9시	11시
희인이네 집(℃)	27.0	32.0	30.5	24.5	21.0	18.0
하늘이네 집(℃)	27.0	28.0	27.0	26.0	25.5	24.5



7. 도연이는 수영장에 갔을 때 물방울이 떨어지면 물이 튀면서 물결이 치는 것을 보았습니다. 집으로 돌아온 도연이는 물방울을 떨어뜨리는 높이가 다르면 물이 튀는 양과 물결의 크기가 다를 것이라고 생각하여 다음 그림과 같이 실험을 준비하였습니다. 도연이가 실험을 하기 위하여 '갈게 해주어야 할 조건'과 '다르게 해주어야 할 조건'으로 바른 것을 고르고, 다음 그림의 A, B, C, D 중 도연이가 알아 보려는 문제와 다르게 준비된 실험 장치를 하나 고르시오.....()



갈게 해주어야 할 조건 다르게 해주어야 할 조건 잘못 준비된 실험 장치

- ① 물이 떨어지는 높이 물의 온도 A
- ② 물이 떨어지는 높이 떨어뜨리는 물의 양 B
- ③ 떨어뜨리는 물의 양 물이 떨어지는 높이 C
- ④ 물의 온도 물이 떨어지는 높이 D

9. 재현이는 오케스트라 공연을 보기 위하여 음악회에 갔는데 공연장의 벽에 스피커와 같은 것들이 많이 붙어있는 것을 보았습니다. 그 이유를 궁금해 하던 재현이는 선생님께 질문을 하였고, 선생님께서는 다음과 같은 실험을 준비해주셨습니다.

① 비어있는 비커

② 수건을 넣은 비커

③ 구겨진 종이를 넣은 비커

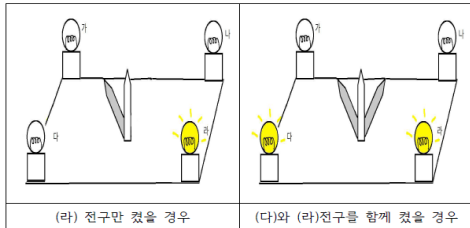
비커의 안쪽을 향해 같은 세기로 소리를 내어본 재현이는 귀로 들었을 때 들리는 소리의 세기를 다음과 같이 상, 중, 하로 표시하여 결과를 정리하였습니다.

번호	①	②	③
소리의 세기	상	하	중

위와 같은 재현이의 실험 결과를 알맞게 해석한 것을 하나 고르시오...()

- ① 비커 속에 물건을 넣고 소리를 내면 소리가 들리지 않는다.
- ② 구겨진 종이를 넣은 비커에 대고 소리를 내었더니 비어있는 비커에 대고 소리를 낼 때보다 조금 크게 들렸다.
- ③ 수건을 넣은 비커에 대고 소리를 내었더니 비어있는 비커에 대고 소리를 낼 때보다 작게 들렸다.
- ④ 구겨진 종이를 넣은 비커는 다른 어떤 방법으로 실험해도 수건을 넣은 비커보다 소리의 세기가 클 것이다.

10. 축구를 좋아하는 지성이는 부모님과 함께 야간 축구 경기를 보러갔습니다. 저녁에 시작되는 경기였지만 축구장에는 사람들이 많이 있었고, 운동장에는 선수들이 열심히 경기를 하고 있었습니다. 경기를 열심히 지켜보던 지성이는 한 가지 특이한 장면을 보았습니다. 선수들 주변으로 여러 개의 그림자가 있었기 때문입니다. 왜 그런지 궁금해 하던 지성이는 집에 돌아와 아버지께 여쭙어 보았고, 아버지는 연필과 전구를 이용해서 그림과 같은 실험 장치를 만들어주셨습니다. 지성이와 아버지의 실험을 통해서 알 수 있는 결론으로 바른 것을 하나 고르시오.....()



- ㉠ 그림자는 불이 켜진 전구와 물체(연필)의 가운데 부분에 생긴다.
- ㉡ (가) 전구와 (라) 전구를 함께 켜면 그림자는 양쪽으로 더 길어질 것이다.
- ㉢ (가) 전구를 켜면 (다) 전구가 있는 방향으로 연필의 그림자가 생길 것이다.
- ㉣ 그림자는 빛이 곧게 나아가다가 물체를 만나 빛의 일부 또는 전부가 막혀 생긴다고 할 수 있다.

교사평가 문항

과학핵심역량평가지

본 검사지는 해당 학생의 과학핵심역량을 측정하기 위한 도구입니다. 각 문항을 잘 읽고 한 문제도 빼놓지 말고 기록해주시기 바랍니다.

학생 이름 :

학생 ID :

【주의점 및 작성 요령】

- * 하나도 빼놓지 말고 모든 질문에 √ 또는 ○표시를 해주시기 바랍니다.
- * ID는 학년, 반, 번호를 연속해서 적으면 됩니다.
예) 4학년 5반 7번 -> 4507
예) 6학년 3반 14번 -> 6314

질문	전혀 아니다 (1점)	아니다 (2점)	보통이다 (3점)	그렇다 (4점)	매우 그렇다 (5점)
과학적 태도 및 정서					
1. 학생은 과학적으로 탐구할 수 있는 일들에 대해 호기심이 많다.					
2. 학생은 과학 실험이나 과학과 관련된 활동을 통해 새로운 사실을 알아가는 것을 즐거워한다.					
3. 학생은 객관적인 자료와 사실을 바탕으로 과학 탐구활동의 결론을 내린다.					
4. 학생은 다른 사람들의 실험 결과가 옳다고 판단되면 자기주장과 다르더라도 받아들일 수 있다.					
5. 학생은 과학 실험을 끝까지 하려고 노력한다.					
6. 학생은 과학 탐구활동을 통해서 새로운 발견을 할 수 있다고 생각한다.					
평생학습능력					
1. 학생은 과학 실험이나 과학과 관련된 활동을 할 때 궁금한 것이 생기면 스스로 찾아본다.					
2. 학생은 과학 탐구활동을 할 때 친구들을 이끌어 가며 주도적으로 하는 편이다.					
3. 학생은 과학 실험이나 활동을 위해 어떤 지식이 나 지원이 필요한지를 잘 알고 있다.					
4. 학생은 탐구활동을 미리 계획하고, 얼마나 진행되었는지 파악할 수 있다.					
5. 학생은 과학 탐구활동이 자기의 발전에 도움이 될 것이라고 생각한다.					

질문	전혀 아니다 (1점)	아니다 (2점)	보통이다 (3점)	그렇다 (4점)	매우 그렇다 (5점)
6. 학생은 과학 지식뿐만 아니라 과학을 탐구하는 방법을 배우는 것이 우리 일상생활의 과학적 문제를 해결하는데 도움이 될 것이라고 생각한다.					
7. 학생은 지속적으로 과학 탐구활동을 하고 있다는 생각과 열정을 갖고 있다.					
8. 학생은 과학 탐구활동을 시작하기 위한 기초적인 공부가 되어 있다.					
9. 학생은 과학 탐구활동에 필요한 정보 및 자료들이 무엇인지 알고 있다.					
10. 학생은 과학 탐구활동에 필요한 선생님, 부모님 그리고 주변 사람들의 도움과 지원을 잘 활용할 수 있다.					
과학적 의사소통능력					
1. 학생은 과학 탐구활동을 할 때 다른 사람의 의견을 주의 깊게 듣는다.					
2. 학생은 과학 탐구활동의 결과를 뚜렷하고 분명하게 이야기한다.					
3. 학생은 과학 탐구활동을 할 때 다른 사람들과 협력적으로 이야기한다.					
4. 학생은 서로의 주장이 다를 때 과학 탐구활동의 결과를 바탕으로 다른 사람을 설득할 수 있다.					
5. 학생은 상대방의 주장을 과학적 근거를 바탕으로 판단할 수 있다.					
6. 학생은 과학 탐구활동에 대한 모동원의 의견이 서로 다를 때 이를 조정할 수 있다.					

질문	전혀 아니다 (1점)	아니다 (2점)	보통이다 (3점)	그렇다 (4점)	매우 그렇다 (5점)
과학적 탐구능력					
1. 학생은 관찰을 통하여 사물과 현상에 대한 정보를 얻어낼 수 있다.					
2. 학생은 공통적인 특성과 조건을 바탕으로 사물을 나눌 수 있다.					
3. 학생은 길이, 무게, 부피, 시간 등을 정확하게 잴 수 있다.					
4. 학생은 관찰한 사실과 증거를 바탕으로 결과를 바르게 해석하고 설명할 수 있다.					
5. 학생은 과학 탐구활동을 할 때 관찰한 사실을 바탕으로 앞으로 어떤 일이 일어날지 예상할 수 있다.					
6. 학생은 어떤 과학적 현상이 왜 일어났는지에 대한 잠정적인 까닭과 설명인 가설을 세울 수 있다.					
7. 학생은 다르게 해주어야 할 조건과 함께 해주어야 할 조건을 구별하여 변인 통제 활동을 할 수 있다.					
8. 학생은 과학 탐구활동을 통해 얻은 자료를 피나 그래프 등으로 나타내거나 변환할 수 있다.					
9. 학생은 과학 탐구활동을 통해 얻은 자료를 분석하여 그에 담긴 의미를 파악하고 설명할 수 있다.					
10. 학생은 자료에 대한 해석을 바탕으로 구체적인 결론을 이끌어낼 수 있다.					

동료평가 문항

과학핵심역량평가지

본 검사지는 해당 학생(친구)의 과학핵심역량을 측정하기 위한 도구입니다. 또한 본 검사는 시험이 아니기 때문에 결과는 학교 성적과 무관합니다. 각 문항을 잘 읽고 한 문제도 빼놓지 말고 기록해주시기 바랍니다.

학생(친구) 이름 :

학생(친구) I D :

【주의점 및 작성 요령】

* 하나도 빼놓지 말고 모든 질문에 √ 또는 ○표시를 해주시기 바랍니다.

* ID는 학년, 반, 번호를 연속해서 적으시면 됩니다.

예) 4학년 5반 7번 -> 4507

예) 6학년 3반 14번 -> 6314

질문	전혀 아니다 (1점)	아니다 (2점)	보통이다 (3점)	그렇다 (4점)	매우 그렇다 (5점)
과학적 태도 및 정서					
1. 학생(친구)은 과학적으로 탐구할 수 있는 일들에 대해 호기심이 많다.					
2. 학생(친구)은 과학 실험이나 과학과 관련된 활동을 통해 새로운 사실을 알아가는 것을 즐거워한다.					
3. 학생(친구)은 객관적인 자료와 사실을 바탕으로 과학 탐구활동의 결론을 내린다.					
4. 학생(친구)은 다른 사람들의 실험 결과가 옳다고 판단되면 자기주장과 다르더라도 받아들일 수 있다.					
5. 학생(친구)은 과학 실험을 끝까지 하려고 노력한다.					
6. 학생(친구)은 과학 탐구활동을 통해서 새로운 발견을 할 수 있다고 생각한다.					
과학적 의사소통능력					
1. 학생(친구)은 과학 탐구활동을 할 때 다른 사람의 의견을 주의 깊게 듣는다.					
2. 학생(친구)은 과학 탐구활동의 결과를 뚜렷하고 분명하게 이야기한다.					
3. 학생(친구)은 과학 탐구활동을 할 때 다른 사람들과 협력적으로 이야기한다.					
4. 학생(친구)은 서로의 주장이 다를 때 과학 탐구활동의 결과를 바탕으로 다른 사람을 설득할 수 있다.					
5. 학생(친구)은 상대방의 주장을 과학적 근거를 바탕으로 판단할 수 있다.					
6. 학생(친구)은 과학 탐구활동에 대한 모동원의 의견이 서로 다를 때 이를 조정할 수 있다.					