초등학교 고학년 학생의 과학 탐구능력 측정을 위한 평가 도구 개발

송경혜 · 이항로[1] · 임청환[2]
(대구비산초등학교) · (대전광역시교육청)[1] · (대구교육대학교)[2]

Development of a Test of Science Inquiry Skills for Elementary School Fifth and Sixth Graders

Song, Kyoung-Hye · Lee, Hang-Ro[1] · Lim, Cheong-Hwan[2]
(Daegu Beesan Elementary School) · (Daejeon Metropolitan Office of Education)[1]
· (Daegu National University of Education)[2]

ABSTRACT

The purpose of this study is to develop a valid and reliable evaluating instrument for elementary school fifth and sixth graders. The instrument is developed through R&D procedure, which includes two checks of science specialist and two field trials of the instrument. Evaluating items are content-free for each science inquiry skill. Each science inquiry element is based on SAPA and the 7th curriculum.

This study has selected 10 science inquiry skills(observing, classifying, measuring, predicting, inferring, recognizing of a problem, controlling variables, interpreting data, drawing a conclusion, designing an experiment), formulated a clear definition of the elements of science inquiry skills, and established the objectives of evaluation. The content areas are divided into three categories, material and energy, life and environment, and the earth and circulation. Each category contains 10 items. So the instrument consists of 30 items. The content validity of items, objectivity of the scoring keys, and clarity of the items has been checked twice by specialists in science education. At the same time, two field trials were performed to produce the reliability of the instruments, discrimination index, and item difficulty index. The instrument has the content validity is 91.6%, reliability 0.79, objectivity 93.3%, discrimination index 0.30, and item difficulty index 66.1%.

Key words: evaluating instrument, science inquiry skills, content validity, reliability

Ⅰ. 연구의 필요성 및 목적

1950년대까지 미국에서 지속되어 온 Dewey(1938)의 생활 중심 교육사조는 1957년 구소련의 스푸트니크 1호 인공 위성의 발사로 전환점을 맞게 되었다. Bruner(1960)를 중심으로 한 과학자, 철학자, 교육자들은 학문 중심 교육 사상을 전개하였으며 과학 교육 연구 중 많은 연구들이 이러한 교육사상을 반영하여 수행되었다. 1960년대부터 초등학교 과학 교육은 학생들이 그 과정에 직접 참여하여 활동함으로써 그 과정을 이해하게 하는 활동의 개발이 특징이라 할 수 있다(Molitor & George, 1976). SAPA(김기웅 등, 1971)도 이러한 과학교육사조의 배경에

*2004.8.24(접수) 2004.9.9(1심통과) 2004.10.28(2심통과) 2004.11.27(최종통과)
**송경혜(song003@hanmail.net)

의해 개발되었으며 그 후 SAPA의 교육과정을 근거로 하여 Padilla 등(1985)의 BAPS (A Test of Basic Process Skills), Mattheis 등(1988)의 POPS (Performance of Process Skills Tests), Dillashaw와 Okey(1980)의 TIPS (Test of Integrated Science Process Skills), APU (Assessment of Performance Unit, 1980), NAEP (The National Assessment of Educational Progress, 1989) 등 평가 대상의 학년과 학생, 그리고 학습 내용에 따라 여러 가지의 과학 탐구 능력 평가 도구가 개발되어 학생들의 과학 탐구 능력을 적절히 평가해 오고 있다(고혁민, 1994).

우리 나라의 과학과 교육과정에서도 이러한 경향이 잘 반영되고 있는데 3차 교육과정부터 SCIS(김상옥 등, 1968), ESS(김현재 등, 1970), SAPA(김기융 등, 1971) 등의 영향을 받아 지식의 구조, 기본개념, 탐구 방법 등을 강조하였다. 과학 교과서 구성도 설명이나 지식 전달보다는 단계별 활동 지시문이나 질문으로 구성되어 탐구 활동을 통한 과학의 개념이나 법칙을 알아내도록 되어 있다. 4차 교육과정(문교부, 1982)에서는 3차 교육과정(문교부, 1973)의 학문중심 교육과정을 토대로 하면서도 학습의 적절성을 감안하여 학습의 양과 수준을 조절하였고, 5차 교육과정(문교부, 1987), 6차 교육과정(교육부, 1992)에서는 실생활에서 접하는 문제를 아동들이 탐구활동을 통해 해결할 수 있는 능력을 길러주고자 하였다. 특히, 6차 교육과정부터는 내용 체계를 지식 영역과 과정 영역으로 나누어 과학 탐구 능력을 강조하였다. 이러한 교육과정의 변화는 7차 교육과정에 이르러 학습 주제별로 탐구 요소를 아동과 교사가 직접 선택하여 문제를 해결할 수 있도록 하여 더욱 창의성 있는 탐구 활동을 하도록 하였다(교육부, 1999). 또, 내용 체계를 지식과 탐구로 나누고 탐구를 탐구 과정과 탐구 활동으로 다시 나누었으며 각 분야를 학년에 적합한 활용 빈도를 제시하였다.

1960년대부터 시작되어 온 탐구 학습에 대한 노력에도 불구하고 탐구 활동 중심의 수업을 하는 데는 여러 가지 어려움이 있는 것으로 지적되었다. Costensons과 Lawson(1986)은 8가지의 원인과 그 해결책을 제시하였으며, 허 명(1984)은 탐구 지도에 시간이 많이 든다, 단순한 개념을 전달하는데 비효율적이다, 교사에게 많은 부담을 준다(자료 준비, 학습 지도, 평가 등), 타당도와 신뢰도가 높은 탐구 능력 평가 방법 개발이 어렵다 의 4가지 원인을 지적했다(이항로, 1991).

여러 가지 원인 중 평가의 방법이 교육 활동에 큰 영향을 준다는 것에 착안한 연구자들은 여러 종류의 표준화된 평가 문항을 개발하였다. Smith와 Welliver(1990)는 4학년 학생을 대상으로 하여 13가지 탐구 과정 요소를 선정하고 SPA (Science Process Assessment)를 개발하였다. Anderson 등(2002)은 대학의 생물 전공자들과 비전공자들을 대상으로 하여 CINS (Conceptual Inventory of Natural Selection)라는 생물의 자연 선택에 대한 문항을 개발하였고 Kim Chwee Daniel Tan 등(2002)은 고등학교 10학년 학생을 대상으로 하여 무기 화학의 정성적 분석을 이해하는지를 평가하는 선다형 문항을 개발하였다. 외국의 경우 APU, NAEP와 같이 국가 수준에서 탐구 능력을 측정 해 오고 있으며 문항 개발에 대한 연구가 꾸준히 계속되어 오고 있다.

우리 나라의 경우 이종기(1988), 이연우(1989), 이항로(1991), 정철(1997)이 중등을 대상으로, 은경용(1992)이 초등 5, 6학년을 대상으로 하여 과학 탐구 능력 평가 도구를 개발하였으며 권재술과 김범기(1994)가 초·중등을 대상으로 개발한 평가 문항을 고혁민(1994)이 개발과정과 타당성을 분석하였다. 정지숙(1996)은 초등 5, 6학년을 대상으로 3차원 평가틀을 제작하고 그에 기초한 과학탐구능력 문항 개발을 하였다. 위의 연구는 과학 수업 활동 전 과정을 반영하는 탐구 과정 요소를 대상으로 하였으며 R&D 과정을 거쳐 제작되었다. 그 외 남도식(1994)은 5학년 1학기 실험 단원, 노홍찬(1997)은 6학년 1학기 실험 단원, 안영균(1996)은 분류 능력, 정귀향(1996)은 길이, 넓이, 부피 측정 능력, 정정애(1996)는 저학년 5개 탐구 과정요소를 선정하여 평가 문항을 개발하였다.

이를 살펴보면 초등에서는 과학 수업 활동의 전 과정을 반영하는 평가 문항 개발 연구가 부족하고, 특히 현행 교육과정에서 제시한 탐구 과정 요소를 바탕으로 한 연구 결과는 찾아볼 수 없었다. 인지발달에 따른 탐구 능력의 차이가 많아 전 학년을 아우르는 연구가 없었고, 중학교와 초등학교 학생을 대상으로 한 고혁민(1994)의 연구에서도 학교급간의 탐구 능력 차이가 많았다고 하였다. 학년에 따라 탐구 과정 요소를 사용하는 빈도도 달라 탐구 과정 요소별로 문항을 개발하기도 하였으나 그 수는 많지 않았다. 따라서 초등에서 사용되는 기본 탐구 능력과 통합 탐구 능력을 모두 사용하는 고학년을 대상으로 하는 연구가 필요하였다.

본 연구자는 탐구 능력의 향상과 탐구 중심 수업의 확

대를 위해 과학탐구능력 평가 문항을 개발하고자 하였다. SAPA의 과정안을 바탕으로 하면서 우리나라 현행 제 7차 교육과정의 탐구 과정 요소를 바탕으로 하여 더욱 실용적인 과학탐구능력 평가 문항이 되도록 하였다.

본 연구에서 개발된 과학탐구능력 평가 문항은 아동들의 과학 탐구능력을 측정함은 물론 과학 탐구 수업에 대한 방향을 제시하는 방향타 역할을 하여 일선 교사와 아동들에게 과학 탐구 능력에 대한 소양을 높일 수 있을 것이다. 또한 현행 지식 중심의 평가 방법을 과학탐구능력 중심의 평가 방법으로의 전환을 통해 탐구능력 신장 중심으로의 과학과 교수 · 학습 방법의 활성화에 기여할 것으로 기대된다.

이러한 필요성에 의해 본 연구에서는 초등학교 고학년을 대상으로 하는 타당도와 신뢰도가 높은 과학 탐구 능력 검사지를 개발하고자 하였다.

Ⅱ. 연구 방법 및 절차

본 연구에서는 이론과 실제의 차이를 극소화시킬 수 있는 연구 · 개발(R&D) 과정을 이용하여 과학탐구능력 평가 문항을 개발하였으며, 구체적인 절차는 다음과 같다.

1. 개발하고자 하는 평가 도구의 특성

탐구 능력은 실험 실기, 보고서 작성, 포트폴리오 평가 등 다양한 방법으로 측정할 수 있다. 그러나 다인수 학급인 우리 나라 초등학교 현실을 고려할 때 위의 방법으로 과학 탐구 능력을 평가하기에는 어려움이 많아 지필평가의 방법을 사용하기로 하였다. 평가 결과의 일관성, 효율성, 객관성 등을 고려하여 선다형을 채택하였으며, 문항을 선택할 때 우연의 요인을 최소화하기 위해 5지 선다형 문항 형식을 사용하였다. 문항은 교과 내용에 대한 특별한 지식이 없어도 해결할 수 있도록(content-free) 하였으며 초등학생의 인지 발달 단계에 적합한 문항을 개발하였다. 평가 시간은 초등학교의 단위 수업 시간을 고려하여 40분으로 하였다.

2. 탐구 과정 요소의 선정, 조작적 정의 및 평가 목표 설정

1) 과학 탐구 과정 요소의 선정

평가 도구를 개발하기 위해 광범위한 평가틀과 평가 요소 선정에 관련한 참고 문헌을 조사하였다. 그 중 초등학생을 대상으로 하여 개발했던 SAPA의 탐구 과정 요소와 제 7차 교육과정의 탐구 과정 요소를 근간으로 하여 다음 Table 1과 같은 10가지 탐구 과정 요소를 선정하였다.

Table 1. Science inquiry elements selected in this study

① Observing	② Classifying
③ Measuring	④ Predicting
⑤ Inferring	⑥ Recognizing of a problem
⑦ Controlling variables	⑧ Interpreting data
⑨ Drawing a conclusion	⑩ Designing an experiment

2) 탐구 과정 요소의 조작적 정의

탐구 과정 요소에 대한 조작적 정의를 내리기 위해 이종기(1988), 이항로(1991), 은경용(1992)이 제시한 평가 목표, SAPA, NAEP가 제시한 평가 목표, Moliter와 George(1976), Nelson과 Abraham(1973), Padilla 등(1985), Mattheis 등(1988), Smith와 Welliver(1990) 등이 제시한 평가 목표를 참고하고, 과학탐구능력 관련 문헌과 과학 교육 전문가의 자문을 얻어 Table 2와 같이 정의를 내렸다.

3) 탐구 과정 요소에 대한 평가 목표 진술

본 연구자가 내린 조작적 정의에 근거하여 Mager(1962)의 진술방식에서 '기준'을 제외하고 평가 조건과 요구 행동으로 구성된 평가 목표를 Table 3과 같이 진술하였다.

3. 평가 문항의 개발 및 검증

1) 평가 문항의 개발

10가지 탐구 과정 요소를 물질과 에너지, 생명과 환경, 지구와 순환의 3가지 내용 영역별로 총 30문항을 개발하였다. 문항은 교과에 대한 특별한 지식이 없어도 해결할 수 있도록 제작하였으며 제 7차 교육과정에서 이수해야 하는 내용과 수준을 고려하여 다양한 내용으로 구성하였다. 물질과 에너지의 하위 영역은 에너지, 도구, 매질 이동, 물질 변화, 물질 성질이고, 생명과 환경의 하위 영역은 한살이, 생태계, 인체, 특성과 역할, 환경이다. 그리고

Table 2. Defining operationally of each science inquiry skill

Science inquiry skill	Defining operationally
Observing	Ability to accept external data through human sensory organs without any precedent knowledge or voluntary interpretation
Classifying	Ability of classifying objects based on the similarities, differences, and the results
Measuring	Ability of quantitatively showing process of the data size using appropriate instruments
Predicting	Ability of predicting future issues or results according to the given facts or observation
Inferring	Ability about process of recognizing observation, unmeasurable situation or facts based on observing or given facts
Recognizing of a problem	Ability about mental operating process which allows basic experiment designing after grasping relationships from given facts or situations
Controlling variables	Ability to maintain difference of the independent variable and the commonness of other variables, after selecting a decisive independent factor which affects dependent variables and establishing a hypothesis
Interpreting data	Ability to change the data from experiments and situations into tables or graphs, or concluding a meaningful conclusion. This includes the ability of mutually changing data into charts and statements.
Drawing a conclusion	Ability of concluding appropriate conclusions from inquiry process. This includes understanding data interrelationship and making clear of the cause and effect relations.
Designing an experiment	Ability to control variables of established hypothesis from external situation or data, and planning activity process to prove it

지구와 순환은 암석, 우주, 지구, 화석, 물로 구성하였다. 또 평가 문항은 황정규(1998)가 제시한 '선다형 문항 제작을 위한 제언'을 참고하여 제작하였다(부록 : 30문항 중 탐구과정 요소별 1문항씩 10문항 첨부).

2) 문항의 타당도, 객관도, 명료성 점검 I

평가 문항의 타당도, 객관도, 명료성을 점검 받기 위해 10명의 내용 전문가에게 의뢰하였다.

황정규(1998)는 평가 문항이 학습 목표 및 교수 목표를 얼마나 잘 대표하고 충실히 측정하느냐에 따라 타당도가 결정된다고 하였다. 본 연구에서는 10명의 내용 전문가에게 평가 목표와 문항을 나누어주고 2단계 평정 방법(1=적절, 0=부적절)으로 점검 받았다. 점검표에 기록한 결과를 토대로 구한 내용 타당도 지수(CVI : Content Validity Index)는 총 300문항 중 239문항이 일치하여 79.7%의 평가 목표가 일치하였다. 널리 받아들여지고 있는 타당도 지수는 80%이상(Doran, 1980; 이항로, 1991)으로 볼 때, 비교적 양호하나 더 향상시킬 필요가 있었다.

객관도는 10명의 내용 전문가가 작성해 준 답안을 분석해 본 결과 300문항 중 257문항이 정답으로 정답률이 85.7%였다. 정답의 객관도는 객관식의 경우 100%에 가까울수록 좋으므로 85.7%의 객관도는 더 향상시킬 필요가 있었다.

문항의 명료성은 10명의 점검자들에게 문항 명료성 점검표를 제시하고 점검표를 바탕으로 하여 문제가 있다고 판단되는 점을 직접 기록하도록 하여 점검받았다.

이런 과정의 타당도, 객관도, 명료성 점검을 바탕으로 하여 지적된 문제점을 바탕으로 수정·보완한 과학탐구능력 평가도구를 현장 검증 I 에 투입하였다.

Table 3. Evaluation objectives of each science inquiry skill

Science inquiry skill	Evaluation objectives
Observing	Accurate observation using eyesight for given situation images, appropriate description of observation results without anticipation and reasoning, showing qualitatively and qualitatively
Classifying	Determination of a standard based on similarities and differences of objects, and classifying them by that standard. Classifying the standard again into secondary, thirdly classification, and analogizing the standard from the results
Measuring	Precisely understanding experiment instrument using methods and scale reading methods. Knowing how to measure the number or size of the population of a large range in nature.
Predicting	Predicting future issues according to present situations with presenting the reasons
Inferring	Disclosing the causes for revealed results, deducing from surrounding situations
Recognizing of a problem	Understanding the purposes of given experiment situations and descriptions
Controlling variables	Searching for variables that can prove the experiment hypothesis. dividing variables into control variables and experiment variables, and controlling experiment variables.
Interpreting data	Indicating in graphs or appropriately explaining, according to the experiment result table or given data.
Drawing a conclusion	Drawing conclusion that can be generalized based on deducted data.
Designing an experiment	Recognizing the problem and designing experiment plans, considering the problem-solving hypothesis, variables, and preparation materials.

Table 4. Group sampling for the first field trial

Region	School	Number of students		
		5th grade	6th grade	Sum
Large city	C elementary school	84	80	164
Medium city	D elementary school	70	74	144
Sum	2	154	154	308

3) 현장 검증 I 의 실시

제작된 평가 도구를 현장 검증에 투입하여 평가도구를 구성하는 모든 문항을 해결하는데 소요되는 시간의 적절성과 평가 도구의 신뢰도, 변별도, 난이도를 검증 받을 수 있다. 초등학교는 각 반이 남녀공학이므로 성별은 고려하지 않았으며 특수반 아동은 제외하였다. 1차 현장 검증은 2003년 11월 29일~ 12월 6일에 실시하였다.

연구 대상의 표집

연구 대상은 대도시 학교 1개와 중 · 소도시 1개를 표집하고 5, 6학년 각 2학급씩 무작위로 표집하는 군집 표집 방법(cluster sampling method)을 이용하였으며 표집의 구체적 특성은 다음 Table 4와 같다.

문항의 신뢰도, 변별도, 난이도 검증

현장 검증 I에서 학생들의 반응을 분석한 결과 신뢰도(Cronbach α)는 0.80, 변별도 평균 지수는 0.31, 난이도 평균은 64.1%으로 나타났다.

개인 평가 도구로 널리 받아들여지고 있는 신뢰도 값은 0.80이상(Doran, 1980; Nitko, 1983)이므로 본 연구의 신

뢰도 값은 허용 범위 내에 있으나 더 향상될 필요가 있다.

문항의 변별도는 Doran(1980)과 김창식 등(1997)은 +0.3 이상으로, Aiken(1987)과 Gronlund(1985)는 +0.2 이상을 바람직하다고 보았다. 이 준거에 비추어 보면 전체 변별도 평균 지수는 0.30으로 양호한 값이다.

적절한 난이도는 대개 20~80%의 범위이다. 전체 난이도의 평균치는 기재형의 경우 50% (Doran, 1980; Aiken, 1987; 김창식 등, 1997), 5지 선다형일 경우는 60% (Gronlund, 1985; 김창식 등, 1997)가 바람직하다. 본 연구에서는 64.1%로 적절한 수치인 60%보다 약간 높게 나타났다.

변별도와 난이도가 허용 기준에서 벗어나는 문항은 수정하였다.

4) 문항의 타당도, 객관도, 명료성 점검 II

1차 현장 검증을 통해 신뢰도, 변별도, 난이도가 기준치에 미흡한 문항들을 수정·보완하여 2차 타당도, 객관도, 명료성 점검을 실시하였다. 2차 점검은 4인의 내용 전문가에게 의뢰하여 각 문항에 해당하는 과학탐구과정요소와 정답을 직접 적도록 하여 실시하였다.

총 120문항 중 110문항이 평가 목표와 일치하여 문항의 내용 타당도는 91.6%로 1차 점검때 보다 향상된 값으로 비교적 내용 타당도가 높다고 판단할 수 있었다.

또, 총 120문항 중 112문항의 정답이 일치하여 객관도가 93.3%로 비교적 양호하다.

5) 현장 검증 II의 실시

1차 전문가 점검 및 현장 검증 후 수정한 문항을 이용하여 현장 검증 II를 실시하였다. 2003년 12월 20일 ~ 12월 26일에 걸쳐 신뢰도, 변별도, 난이도를 검증하였다.

연구 대상의 표집

연구 대상은 대도시 4개교와 중·소도시 2개교를 표집하고 5, 6학년 각 4학급씩 무작위로 표집하는 군집 표집 방법(cluster sampling method)을 이용하였으며 표집의 구체적인 특성은 아래 Table 5와 같다.

문항의 신뢰도, 변별도, 난이도 검증

현장 검증 II의 분석은 현장 검증 I과 같은 방법으로 실시하였다. 그 결과 신뢰도(Cronbach α)는 0.79, 변별도 평균 지수는 0.30, 난이도 평균은 66.1%이다. 현장 검증 I과 비교해 볼 때 난이도는 2% 증가하였는데 이것은 문제를 명료화하여 문항의 정답률이 높아진 것으로 판단할 수 있다. 이와 함께 변별도 평균 지수는 0.31에서 0.30으로, 신뢰도는 0.80에서 0.79로 조금 낮아졌다. 그러나 이 수치들은 본 연구에서 정한 허용 범위에 포함되므로 문항의 신뢰도와 변별도는 양호하다고 볼 수 있다.

III. 개발한 탐구 능력 평가 도구에 대한 기술

본 검사 도구에서는 SAPA와 제 7차 교육과정의 탐구 과정 요소를 근간으로 하여 10가지의 탐구 과정 요소를 Table 1과 같이 선정하였다. 탐구 과정 요소에 대한 조작적인 정의를 Table 2에 나타내었고, 평가 목표를 Table 3에 나타내었다. 이를 바탕으로 10가지 탐구 과정 요소를 3가지 내용 영역에 해당되도록 총 30문항을 개발하여 2차에 걸친 전문가 점검과 현장 검증을 실시하여 본 검사 도구를 개발하였다. 개발한 과학탐구능력 평가도구의 명칭을 "초등학교 고학년 학생의 과학 탐구 능력 검사"로 명명하였으며, 그 특징은 아래 Table 6과 같다.

Table 5. Group sampling for the second field trial

Region	School	Number of students		
		5th grade	6th grade	Sum
Large city	K elementary school	0	88	328
	G elementary school	0	79	
	Y elementary school	82	0	
	B elementary school	79	0	
Medium city	J elementary school	60	65	250
	A elementary school	66	59	
Sum		287	291	578

Table 6. Characteristic of the developed instrument

Inquiry elements	10 Elements
Evaluating items	30 Items
Type	Selection type Five-options multiple choice item
Subject	Elementary 5th, 6th graders
Testing time	40 minutes
Content validity index	91.6%
Reliability(Cronbach α)	0.79
Objectivity	93.3%
Discrimination index	Average 0.30 0.00 ~ 0.20 : 4 items 0.21 ~ 0.40 : 23 items 0.41 ~ 0.60 : 3 items
Item difficulty index	Average 66.1% 0.21 ~ 0.40 : 1 items 0.41 ~ 0.60 : 7 items 0.61 ~ 0.80 : 21 items 0.81 ~ 1.00 : 1 items

Ⅳ. 결론 및 제언

본 연구에서는 초등학교 5, 6학년 학생의 과학 탐구능력을 평가할 수 있는 타당도와 신뢰도가 비교적 높은 평가도구를 개발하고자 하였다. 각 문항들에 대해 2차에 걸친 전문가들의 점검과 2차의 현장 검증을 통해 타당도 91.6%, 신뢰도 지수 Cronbach α값 0.79, 객관도 93.3%, 변별도 지수 0.30, 난이도 지수 66.1%인 과학탐구능력 평가도구를 개발하였다.

개발된 평가 도구는 과학 탐구 능력을 측정하는 도구로서 사용될 수 있으며, 탐구학습에 대한 진단 평가와 총괄 평가로도 이용될 수 있다. 탐구 과정 요소별, 내용 영역별로 따로 실시할 수도 있으며 탐구 능력 평가에 관련한 기초 연구 자료로도 이용될 수 있다.

지필 평가로만 개발된 본 평가 도구만으로는 과학 탐구 능력을 평가하기에 부족하므로 실험 · 실습 평가, 포트폴리오, 보고서 평가, 면담, 체크리스트 등 다양한 방법과 함께 실시하는 것이 바람직하다.

본 연구 과정과 결과로 얻어진 자료를 바탕으로 개선점이나 앞으로의 방향을 제시하면 다음과 같다. 첫째, 지필 평가만으로 과학탐구능력을 측정한다는 것은 많은 문제점을 가지고 있으므로 다양한 방법의 과학 탐구 능력 측정 방법을 개발하는 후속 연구가 이루어져야 한다. 둘째, 각 학교에서 과학 탐구 능력 평가 도구를 제작한다는 것은 어려운 일이므로 체계적이고 유기적인 연구가 지속적으로 이루어지는 것이 좋겠다. 셋째, 특정 교과 내용에 대한 지식이 필요없는 범교과적인 문항 개발 연구가 주종을 이루고 있으나 탐구 능력의 향상을 가져오기 위해서는 교과와 관련있는 탐구 능력 측정 도구도 필요하다. 넷째, 초등학생의 경우 인지 발달에 의한 탐구능력의 차이가 크기 때문에 각 학년에 맞는 개별적 탐구능력 측정 도구가 필요하다.

다섯째, 현재 과학 탐구 능력을 향상시킬 수 있는 프로그램은 거의 없으며 주로 영재 교육과 관련하여 개발되어 있다. 학생의 수준과 학년에 맞는 프로그램의 개발과 개발된 프로그램에 맞는 평가 문항의 개발이 유기적으로 이루어지는 것 또한 필요하다.

국 문 요 약

본 연구에서는 평가도구의 구비조건을 갖춘 초등학교 고학년 학생의 과학탐구능력을 측정할 수 있는 과학탐구능력 평가도구를 개발하고자 하였다. 평가도구를 구성하는 문항은 R&D 방법에 의해 개발되었으며, 2차례에 걸친 과학교육 전문가의 점검과 2번의 현장 검증을 통해 문항을 수정 · 보완하였다. 각 평가 문항은 교과 내용에 대한

지식이 없어도 해결할 수 있는 범교과적인 문항으로 제작하여 과학탐구능력의 하위 요소들을 측정하고자 하였다. SAPA 교육과정과 제7차 과학과 교육과정에서 제시한 과학탐구과정 요소를 근간으로 관찰, 분류, 측정, 예상, 추리, 문제인식, 변인통제, 자료해석, 결론도출, 실험설계의 10가지 하위 요소들을 선정하고 각 요소에 대한 조작적 정의와 평가 목표를 설정하였다. 물질과 에너지, 생명과 환경, 지구와 순환의 3가지 내용 영역의 소재를 중심으로 총 30문항을 개발하였다.

개발된 문항을 과학교육 전문가에게 2회 의뢰하여 타당도, 객관도, 문항의 명료성을 점검 받았으며 2번의 현장 검증을 통해 신뢰도, 변별도, 난이도를 검증하였다. 그 결과 타당도 91.6%, 신뢰도 지수 Cronbach α값 0.79, 객관도 93.3%, 변별도 지수 0.30, 난이도 지수 66.1%로 나타났다.

위와 같은 값들은 평가도구가 갖추어야 할 구비조건의 허용 범위내에 있는 것으로 보아 본 연구에서 개발한 과학탐구능력 평가도구는 양호한 것으로 볼 수 있다. 따라서 본 연구에서 개발한 과학탐구능력 평가도구는 초등학생들의 과학탐구능력 성취 수준의 진단, 과학과 교육과정과 교수?학습 자료 및 과학과 교수·학습 방법의 과학탐구능력 반영 수준을 평가하는 도구로도 이용될 수도 있을 것이다.

참 고 문 헌

고혁민(1994). 초·중학생들을 위한 과학 탐구 능력 측정 도구의 개발 과정 및 타당성 분석. 한국교원대학교 대학원 석사학위 논문.

교육부(1992). 국민학교 교육과정 해설. 대한교과서주식회사.

교육부(1999). 초등학교 교육과정 해설. 교육과학사.

권재술, 김범기(1994). 초·중학생들의 과학 탐구 능력 측정 도구의 개발. 과학교육학회지, 14(3), 251-264.

김기웅, 김성준, 김상해, 김창식 편저(1971). 교사용 지침서 SAPA(AAAS) 초등과학. 동화문화사.

김상옥, 김성희, 김영대, 김종성, 이우진, 임병기, 최병문 편저(1968). 교사용 SCIS 초등과학. 교육출판사.

김창식, 이화국, 권재술, 김영수, 김찬종(1997). 과학 학습 평가. 교육과학사.

김현재, 남상돈, 도태기, 임정환, 한안진 공력(1970). 교사용 지침서 ESS 초등과학. 동화문화사.

남도식(1994). 과학적 탐구능력 측정을 위한 자연과 실험 평가 도구 개발. 한국교원대학교 대학원 석사학위 논문.

노홍찬(1997). 초등 6학년 학생의 과학 실험 기능 측정을 위한 평가 도구 개발. 한국교원대학교 대학원 석사학위 논문.

문교부(1973). 국민학교 교육과정. 교학도서주식회사

문교부(1982). 국민학교 교육과정. 대한교과서주식회사.

문교부(1987). 국민학교 교육과정. 대한교과서주식회사.

안영균(1997). 초등학생들의 분류 능력 평가. 한국교원대학교 대학원 석사학위 논문.

은경용(1992). 국민학생의 과학 탐구능력 측정을 위한 평가도구 개발. 한국교원대학교 대학원 석사학위 논문.

이연우(1989). 과학 탐구 능력 측정을 위한 표준화 검사지 개발- 중학교 2학년의 자료 분석과 해석능력을 중심으로-. 한국교원대학교 대학원 석사학위 논문.

이종기(1988). 고등학생의 과학 탐구능력 측정을 위한 평가도구 개발. 한국교원대학교 대학원 석사 학위 논문.

이항로(1991). 고등학생의 과학 탐구능력 측정을 위한 평가도구 개발. 한국교원대학교 대학원 석사 학위 논문.

정귀향(1996). 국민학생들의 길이, 넓이 및 부피 측정 능력 평가. 한국교원대학교 대학원 석사학위 논문.

정정애(1996). 국민학교 저학년 과학 탐구 능력 측정을 위한 평가도구 개발. 한국교원대학교 대학원 석사학위 논문.

정지숙(1996). 삼차원 과학 탐구 평가틀을 이용한 국민 학생들의 과학 탐구능력 측정. 한국교원대학교 대학원 석사학위 논문.

정 철(1997). 구조화된 문항을 이용한 지구과학 탐구능력 평가도구 개발. 한국교원대학교 대학원 석사학위 논문.

허 명(1984). 과학 탐구 평가틀의 개발. 한국과학교육학회지, 4(2), 57-63

황정규(1998). 학교 학습과 교육평가. 서울, 교육과학사.

Aiken, L. R. (1987). Testing with Multiple-Choice Items. *Journal of Research and Development in Education*, 20(4), 44-45.

Anderson, D. L., Fisher, K. M., & Norman, G. J. (2002). Development and Evaluation of the Conceptual Inventory of Natural Selection. *Journal*

of Research in Science Teaching, 39(10), 952-978.

APU. (1980). *Science in Schools, Age 15: Report No. 1*. Center for Studies in Science Education, University of Leeds.

Bruner, J. (1960). *The Process of Education*. Cambridge: Harvard University Press.

Costensons, K. & Lawson, A. E. (1986). Why isn't Inquiry used in More Classrooms? *The Amercian Biology Teacher*, 48, 150-158.

Dewey, J. (1938). *Logic: The Theory of Inquiry*, New York: Henry Holt and Co.

Dillashaw, F. G. & Okey, J. R. (1980). Test of the Integrated Science Process Skills for Secondary Science Students. *Science Education*, 64(5), 601-608.

Doran, R. L. (1980). *Basic Measurement and Evaluation of Science Instruction*. Washington, D. C.: National Science Teachers Association, 13-18.

Gronlund, N. E. (1985). *Measurement and Evaluation in Teaching 5th ed.*, New York: Macmillan Publishing Co, 5-21, 169-212, 263-319, 346-378.

Kim Chwee Daniel Tan, Ngoh Khang Goh, Lian Sai Chia & David F. Treagust. (2002). Development and Application of a Two-Tier Multiple Choice Diagnostic Instrument to Assess High School Students' Understanding of Inorganic Chemistry Qualitative Analysis. *Journal of Research in Science Teaching*, 39(4), 283-301.

Mager, R. F. (1962). *Preparing Instructional Objectives*, Sanfrancisco: Rearon Pub.

Mattheis, F. E., Nakayama, G., Pottenger, F. M., & Jones, M. L. (1988). Development of the Performance of Process Skills (POPS) Test for Middle Grades Students, *Microfilms*, Ed 305-252, 1-15.

Molitor, L. L. & George. K. D. (1976). Development of a Test of Science Process Skills. *Journal of Research in Science Teaching*, 13(5), 405-412.

NAEP. (1989). Science objectives-1990 assessment.

Nelson, M. A. & Abraham, E. C. (1973). Inquiry Skill Measure. *Journal of Research in Science Teaching*, 10(4), 291-297.

Nitko, A. J. (1983). *Educational Test and Measurement an Introduction*. New York, Harcourt Brace Jovanovich, Inc.

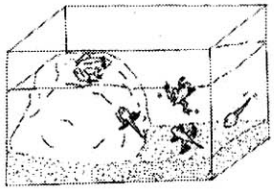
Padilla, M. J., Cronin, L. L., & Twiest, M. (1985). *The Development and Validation of a Test of Basic Process Skills*. Paper presented at the Annual Meeting of the National Association for Research in Science Teaching.

Smith. K. A. & Welliver, P. W. (1990). The Development of a Science Assessment for Fourth-Grade Students. *Journal of Research in Science Teaching*, 27(8), 727-738.

부 록

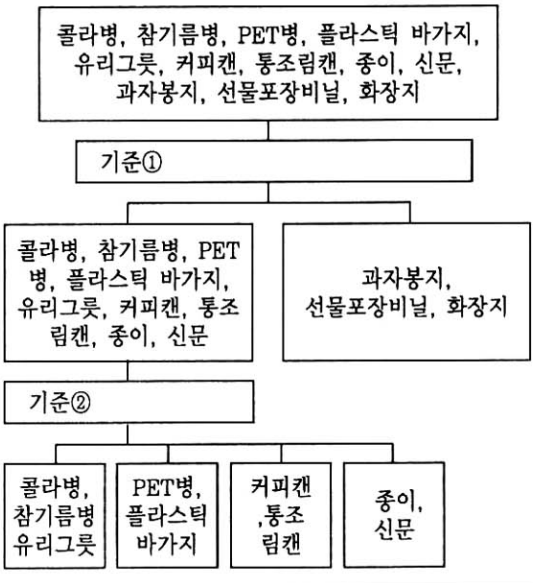
초등학생 탐구능력 검사지

[관찰] 11. 올챙이를 수조에 넣고 개구리로 키우고 있다. 아래 수조 안의 모습을 관찰한 내용으로 알맞은 것은?



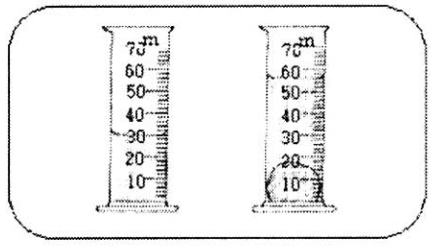
① 올챙이는 앞다리부터 생긴다.
② 올챙이는 물 밖에서 생활한다.
③ 올챙이는 달걀노른자를 주로 먹는다.
④ 개구리는 물 속, 물 밖에서 모두 생활한다.
⑤ 개구리는 울음주머니를 이용하여 소리를 낸다.

[분류] 12. 교실에서 생기는 여러 가지 쓰레기를 분리 수거하려고 한다. 아래와 같이 쓰레기를 분리하는 기준으로 바른 것은?



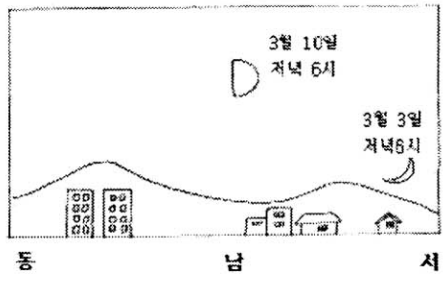
	기준①	기준②
①	태울 수 있는 것과 없는 것	모양
②	재활용 가능한 것과 불가능한 것	재질의 종류
③	태울 수 있는 것과 없는 것	썩는 시간
④	재활용 가능한 것과 불가능한 것	오래된 것의 순서
⑤	재활용 가능한 것과 불가능한 것	태울수 있는 것과 없는 것

[측정] 3. 물이 든 눈금실린더에 울퉁불퉁한 모양의 돌을 넣었더니 그림과 같이 물의 높이가 높아졌다.
이 돌이 물 속에서 차지하는 부피는 몇 mL일까?



① 23 mL ② 26 mL
③ 28 mL ④ 32 mL
⑤ 56 mL

[예상] 24. 3월 3일날 저녁 6시 서쪽 하늘에서 초승달을 보았다. 그 뒤 7일 후 3월 10일 저녁 6시에 하늘을 보니 반달이 남쪽 하늘에 떠 있었다. 다시 8일 후 3월 18일 저녁 6시에 하늘을 본다면 어떤 달이 어느 쪽 하늘에 더 있을까?



⬆(관찰자 위치)

① 동쪽하늘 반달 ② 동쪽하늘 보름달
③ 서쪽하늘 초승달 ④ 서쪽하늘 보름달
⑤ 남쪽하늘 보름달

[추리] 25. 아래의 설명으로 미루어 볼 때, '겨울에는 별이 더 잘 보인다'는 사실에 대한 증거로 맞는 것은?

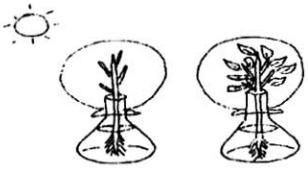
▶ 겨울에는 눈이 가끔 온다.
▶ 여름에는 비가 자주 많이 온다.
▶ 여름에는 기온이 높고 습도도 높다.
▶ 겨울에는 기온이 낮고 습도도 낮다.
▶ 여름에는 집중호우, 장마, 태풍이 온다.

▶ 여름에는 피부의 표면이 끈적거릴 때가 많다.
▶ 여름에는 구름이 잘 생기고 소나기가 자주 온다.
▶ 겨울에는 피부의 표면이 건조하여 살이 트기도 한다.

① 겨울에는 날씨가 더 춥기 때문이다.
② 겨울에는 하늘이 더 높기 때문이다.
③ 겨울철 별자리 종류가 더 많기 때문이다.
④ 겨울에는 공기 중에 수증기량이 더 적기 때문이다.
⑤ 겨울에는 공기 중에 오염 물질이 더 적기 때문이다.

[문제인식] 16. 다음의 실험은 무엇을 알아보기 위한 것일까?

실험 방법 : 똑같은 크기의 삼각 플라스크에 같은 양의 물을 넣고 같은 종류의 식물을 뿌리째 넣었다. 그리고 한쪽은 잎을 모두 따내고 다른 한쪽은 그대로 둔 채 비닐로 양쪽 모두 덮어 씌워서 햇빛이 잘 비치는 곳에 두었다. 다음날 비닐 봉지 안의 상황을 관찰한다.

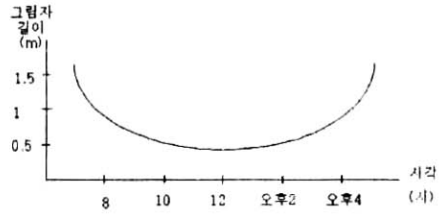


① 식물 뿌리가 어떤 역할을 하는지 알아보는 실험이다.
② 잎을 비닐봉지로 싸 놓으면 어떻게 되는지 알아보는 실험이다.
③ 나뭇잎이 없는 식물은 오래 못산다는 것을 알기 위한 실험이다.
④ 뿌리에서 흡수된 물이 잎을 통해 나간다는 것을 알아보기 위한 실험이다.
⑤ 식물도 숨을 쉰다는 것을 알아보기 위해 비닐봉지로 공기를 막는 실험이다.

[변인통제] 7. '소금 알갱이의 크기가 작을수록 빨리 녹을 것이다' 라는 예상을 한 후 실험 계획을 세웠다. 이 실험의 예상을 확인하기 위해 다르게 해 줘야 하는 조건은 무엇일까?

① 같은 양의 소금을 하나는 500mL의 물에, 다른 하나는 200mL의 물에 녹인다.
② 두 비커에 같은 양의 소금과 물을 넣은 후 한쪽은 저어주고 다른 한쪽은 젓지 않는다.
③ 한쪽은 소금을 막자사발에 곱게 갈고 다른 하나는 소금 알갱이 그대로 사용하여 녹인다.
④ 같은 양의 소금을 한쪽 비커는 상온(20℃)의 물에 다른쪽 비커는 알콜 램프로 가열한 물에 녹인다.
⑤ 같은 양의 물이 담긴 비커 두개에, 한쪽 비커에는 소금 20g을 다른쪽 비커에는 소금 50g을 녹인다.

[자료해석] 28. 운동장에 1m 막대를 세워놓고 하루 동안의 그림자의 길이를 재어 보았더니 다음 그래프와 같았다.



이 그래프에 대한 설명으로 옳은 것은?

① 하루 중 그림자의 방향은 항상 같다.
② 시간이 지날수록 그림자는 점점 길어진다.
③ 낮 12시를 중심으로 태양이 낮게 떴다가 높게 뜬다.
④ 그림자의 모양은 직선이 아니라 굽은 모양으로 생긴다.
⑤ 낮 12시를 중심으로 그림자가 짧아지다가 다시 길어진다.

[결론도출] 9. 다음의 사실을 바탕으로 결론을 바르게 정리한 것은?

가. 문 틈새로 빛이 곧게 들어온다.
나. 창문을 통해 아침 햇살이 곧게 들어온다.
다. 전구의 불빛이 물체를 지나 그림자를 만들었다.

① 빛은 물 속에서 곧게 나아가는 성질이 있다.
② 빛은 매끈한 표면에서 반사되는 성질이 있다.
③ 빛은 다른 물질을 만나면 꺾이는 성질이 있다.
④ 빛은 공기 중에서 곧게 나아가는 성질이 있다.
⑤ 빛은 다른 물질을 만나면 흩어지는 성질이 있다.

[실험설계] 30. 갈릴레이의 책을 읽다가 "무게가 다른 물체도 같은 높이에서 떨어뜨리면 같은 속도로 떨어진다"라는 내용을 보았다. 무거운 물체와 가벼운 물체를 동시에 떨어뜨리면 무거운 물체가 먼저 떨어질 것 같은 생각이 들어 실험을 해 보기로 하였다. 다음 중 올바른 실험 방법은 무엇일까?

① 20g짜리 추와 100g짜리 추를 학교 옥상에서 동시에 떨어뜨린다.
② 20g짜리 추와 20g짜리 나무토막을 학교 옥상에서 동시에 떨어뜨린다.
③ 20g짜리 추는 2층 창문에서, 100g짜리 추는 4층 창문에서 동시에 떨어뜨린다.
④ 20g짜리 추 중 하나는 4층 창문에서 다른 하나는 2층 창문에서 동시에 떨어뜨린다.
⑤ 스케치북 2장을 찢어 한 장은 그대로, 다른 한 장은 배를 접어 학교 옥상에서 동시에 떨어뜨린다.