

고등학생의 지구과학 탐구능력 측정을 위한 평가도구 개발

우종욱 · 이항로

(한국교원대학교)

(1994년 12월 6일 받음)

I. 서론

전통적인 과학 교육에서는 과학 하는 방법보다 과학 지식을 중요시하였으며, 과학 교사는 지식의 원천으로서 학생들에게 지식을 일방적으로 전달하는 역할을 주로 담당하였다. 그러나 1950년대까지 미국의 교육계를 지배해 온 듀이의 생활 중심 교육 사상은 달리 학문적인 본성과 본질을 강조해야 한다는 입장인 학문 중심 교육 사상이 등장하게 되었다. 이러한 교육 사상의 영향을 받은 과학 교육 개혁 운동을 지지해 준 이론은 과학 분야의 탐구론과 심리학 분야의 피아제의 인지 발달론, 그리고 이 이론을 원용한 부르너 나선형 교육과정론이라고 볼 수 있다(한중하, 1984).

이러한 사상과 이론을 근간으로 1960년대부터 추진해 온 과학 교육 개혁 운동의 특징을 살펴보면, 첫째로, 내용을 기본적인 개념적 사고 체계로 구조화하고 이를 탐구과정으로 재발견하도록 하는 형태로 재구성하는 지식의 구조화와, 둘째로, 과학의 본질 중의 하나를 '탐구'라고 본다면, 과학 지식은 이런 탐구과정의 산물이며, 과학은 어떤 문제의 해답을 찾는 과정 또는 탐구 방법이라 할 수 있는 과학적 탐구 방법 즉, 과학적 문제해결 방법의 도입으로 요약할 수 있다. 이러한 과학 교육의 시대적 요청에 따라 우리나라의 교육계에서도 전면적인 교육 과정의 개혁을 통하여 새로운 탐구 중심 교육 사조를 도입하게 되었다(정연태, 1984; 김동식, 1989; 정완호, 1989).

미국의 경우 과학교사 협의회(1982)에서는 1980년대의 과학 교육의 목표 중에서 그 첫째로 "과학 기술의 과정과 탐구능력의 개발"을 제시한 바 있으며, 우리나라에서도 새 교육 과정의 실현을 위해 "자연 현상을 과학적으로 탐구하는 능력의 배양"을 설정함으로써(문교부, 1973; 1981; 1985;

1988, 교육부, 1992) 과학 교육의 혁신을 실현하기 위한 노력을 경주해 오고 있다. 이러한 탐구 중심 교육을 도입함에 따라 국내외의 대다수 국가들이 탐구학습을 현장에 적용하기 위한 실험 연구를 진행한 결과에 의하면 학생들의 지적, 정의적 영역뿐만 아니라 상위 인지 능력을 포함한 탐구능력을 향상시키는데 탐구학습 방법이 전통적인 강의 중심 학습 방법보다 효과적임이 밝혀졌다.

이러한 장점을 지닌 탐구 위주의 과학 교육은 많은 과학 교육자들로부터 지지를 받고 있으며, 실제 과학 교육 현장에 적용하기 위한 다각적인 노력이 있었음에도 불구하고 국제 교육 성취도 평가 연구 기구(IEA: The International Association for the Evaluation Achievement, 1988)의 제 2차 과학 교육 성취도 평가 연구(SISS: Second IEA Science Study)의 국내 연구(임인재 등, 1986)와 국내의 한 연구단(박승재, 1988)의 연구 결과에 의하면 우리나라의 경우 학년이 높아질수록 실험 실습을 통한 학습의 기회가 감소하고, 실험을 통한 경험적 탐구적 방법보다는 교과서 지식 중심의 주입식 방법에 의하여 과학수업이 이루어지고 있는 것으로 나타나, 현재의 탐구적 과학 교육에 커다란 문제가 있음을 강력하게 시사해 주고 있다.

이러한 근본적인 원인에 대하여 Hurd 등(1980)은 "많은 교실에서 탐구학습이 실행되지 않고 있다"고 지적했으며, Costenson과 Lawson(1986)은 탐구 학습이 제대로 실시되지 못한 원인을 8가지로 지적과 함께 개선 방안도 제시하였다.

또한 허 명(1984)은 탐구 교육 과정이 현장에서 건들 수 밖에 없었던 원인을 ① 탐구 지도에 시간이 많이 든다, ② 단순한 개념을 전달하는데 비효율적이다, ③ 교사에게 많은 부담을 준다(자료 준비, 학습지도, 평가 등), ④ 타당도와 신뢰도가 높은 탐구능력 평가 방법의 개발이 어렵다 등을 지

적했다. 따라서 본 연구에서는 이러한 문제점들을 다소나마 해결하기 위하여 위에서 허 명(1984)이 지적한 네 번째 문제점인 타당도와 신뢰도 높은 과학탐구능력 평가 도구의 개발에 초점을 두고자 한다.

우리 나라의 초·중·고등학교 과학과 교육 과정에는 탐구 과정과 탐구 기술이 일관되게 강조되고 있으며, 제 5차 교육 과정에는 교과 목표와 더불어 평가상의 유의점을 설정하고, 평가의 관점 및 방법을 제시하여 인지적 영역인 지식과 이해, 탐구 수행 능력, 정의적 영역 및 실습 기능을 고루 평가하도록 되어 있어 교과와 성취 진단은 물론 학습지도와 평가의 일체화를 꾀하고 있음을 알 수 있다.

한편, 현행 대학 입시 제도인 대학수학능력시험의 과학 탐구 영역에서의 "과학탐구능력" 평가와 1996학년도부터 적용될 제 6 차 교육과정의 내용 체계에서 과학적 탐구를 한 영역으로 분리 독립시켜 그 동안 구호만 식상하던 탐구능력 배양을 더욱 강조해 놓고 있다(교육부, 1992). 이러한 현실을 감안해 볼 때 일선 학교 과학 교사들로 하여금 학생들이 학습하기 어려운 탐구능력 요소들을 진단하고 학생들의 탐구능력 신장 정도를 부분적으로나마 추적할 수 있도록 하기 위해서는 일반적인 과학 내용으로 구성된 탐구능력 측정을 위한 타당도와 신뢰도 높은 평가 도구의 개발이 필요함은 물론이고, 각기 다른 과학 교과 영역의 소재들을 중심으로 평가 문항을 작성한 후 현장 검증을 통해 타당도와 신뢰도가 입증된 다양한 과학탐구능력 평가 도구의 개발이 절실히 요청된다.

따라서 본 연구에서는 탐구학습 평가에 대한 이러한 현실을 감안하여 우리 나라 고등학생들에게 적용 가능하며, 문항 장면이 지구과학 교과 소재들로 구성된 신뢰도와 타당도가 높은 과학탐구능력 평가 도구를 개발하고자 한다.

II. 연구 방법 및 절차

본 연구는 평가 도구의 구비 조건을 만족시키는 과학탐구능력 평가 도구를 개발하는데 있다. 본 연구에서 채택한 평가 도구 개발 방법은 이론과 실제의 차이를 극소화시킬 수 있는 장점을 지닌 R&D 과정으로 교육 자료를 개발하는데 자주 이용되고 있으며, 본 연구는 다음과 같은 과정을 거쳐 수행되었다.

1. 개발하고자 하는 평가도구의 성격과 준거 설정

탐구능력 평가 방법에는 지필 평가, 실험 관찰에 의한 평

가, 실험 보고서에 의한 평가, 실험 과제에 의한 평가, 체크리스트법에 의한 평가 등 다양한 방법이 있으나, 본 연구자는 현실적인 연구 여건을 고려하여 지필 검사로 한다. 문항 유형은 5지 선다형, 문항 수는 33 문항, 문항 소재는 지구과학 소재로 하며 평가 시간은 50분으로 한다.

ASE의 정의에 의하면 탐구능력은 단순한 지시문이나 단기간 동안의 학습에 의하여 습득되는 것이 아니라 오랜 기간 동안의 반복된 훈련에 의해서만이 습득되는 상위 인지능력임을 알 수 있다. Cronbach(1984)는 학교에서 이루어지는 수업 내용에 영향을 받는 정도가 클수록 성취도 평가로 그 역할수록 소양 평가로 분류하였으며, 주로 학교에서 이루어지는 모든 평가는 성취도 평가의 성격과 소양 평가의 성격을 모두 지니고 있다고 지적하였다.

이와 같은 이론에 의하면 학교 학습의 결과에 의한 탐구능력 평가의 내용은 우리 생활 주변의 다양한 소재에서 평가 내용을 선정해야 함은 물론 각 과학 교과와 학습을 통해 습득한 개념들을 근거로 한 평가 내용을 선정하여야 한다.

2. 탐구과정 요소의 선정, 정의 및 평가 목표 설정

1) 과학 탐구 과정 요소의 선정

과학 탐구 평가들과 과학탐구능력 평가 도구 개발에 관한 광범위한 선행 연구를 근거로 하여, 고등학생들에게 적용 가능한 과학탐구능력 평가 도구에 포함될 11가지의 탐구능력 요소들을 선정하였다. 본 연구에서의 과학탐구능력 요소의 선정은 주로 SIEI(허 명, 1984)를 근간으로 했으며, 허 명이 개발한 탐구학습 과정의 순서에 따라 제시하면 다음 [표 1]과 같다.

[표 1] 본 연구에서 선정한 11가지 탐구과정 요소

(1) 가설 설정 (2) 변인 통제 (3) 실험 설계 (4) 자료의 변형 중 숫자 계산 (5) 실험 값을 그래프로 나타내기 (6) 추리 (7) 상관 관계 결정 (8) 인과 관계 결정 (9) 예상 : 내삽과 외삽 포함 (10) 결론 도출 (11) 일반화 또는 모델 형성

2) 탐구과정 요소의 정의

위에서 선정한 11개의 과학 탐구 과정 요소에 대한 정의를 내리기 위해, 허 명의 SIEI와 영국의 APU 평가틀에 제시된 평가 목표, NSTA, NAEP, SAPA가 제시한 평가 목표, Dillashaw 등이 과학 탐구 능력 평가 도구를 개발할 때 이용한 평가 목표, Sund와 Picard(1972)가 행동 목표와 평가

측정 소개한 평가목표, Bingman(1969)이 제시한 생물 교수의 탐구목표, Moliter와 George(1976), Nelson와 Abraham(1976), Padilla 등(1988), Small(1988), Matthesis 등(1988), Smith와 Welliver(1990) 등이 과학 탐구 능력 평가 도구를 개발할 때 사용한 평가 목표 등을 참고하고, 설문지, 세미나 등을 통하여 과학 교육 전문가의 자문을 얻어 본 연구자는 [표 2]와 같이 정의를 내렸다.

3) 탐구 과정 요소에 대한 평가 목표의 진술
평가 도구에 대한 내용 타당도를 높이기 위해서 본 연구자가 내린 11 가지 탐구 과정 요소에 대한 정의를 토대로 평가 문항 개발의 준거로 이용될 수 있도록 Gagne의 교육 목표 진술 방식인 평가 조건과 요구 행동으로 구성된 행동 목표 형태로 진술했으며, 그 내용은 다음[표 3]과 같다.

[표 2] 본 연구자가 내린 과학 탐구 과정 요소의 정의

탐구 과정 요소	조 작 적 정 의
1. 가설설정	관찰이나 경험한 사실을 근거로 하여 인식한 문제 상황 속에 내재해 있는 변인들 간의 관계에 대한 잠정적인 설명 또는 잠정적인 아이디어를 창출하는 능력으로써, 학습자의 경험을 토대로 한 논리적 사고 또는 직관적 사고 활동이다.
2. 변인통제	실험 결과에 관여하는 독립 변인과 종속 변인을 구별하고 두 변인 사이의 관계에 편파적인 효과를 기칠 수 있는 제3, 제4의 요인을 일정하게 유지해 주는 활동이다.
3. 실험설계	설정된 가설 속에 내재해 있는 독립 변인과 종속 변인 사이의 관계를 추출하기 위해 여러 가지 실험 방법 및 과정을 고안하는 능력으로써 원리적으로나 실제적으로 가설을 검증하기에 적합한 것이어야 한다. 필요한 경우에 조건 통제에 대한 조치가 포함된다.
4. 자료변형 숫자계산	중 실험, 관측, 관찰을 통해 얻은 자료를 실험 목적이나 문제해결을 위해 좀더 유용한 형태나 해석하기 쉬운 형태로 변환시키는 능력으로, 이러한 변환 과정에서는 단위계나 차원(dimension)에 대한 고려와 수학적 계산(가감승제)의 경우 반드시 오차 계산의 원칙에 따라야 한다.
5. 실험값을 그래프로 나타내기	실험, 관측, 관찰을 통해 얻은 자료를 실험의 성격이나 분석 방법에 따라 좀더 조직적이고 구조적인 형태로 표현하는 능력으로서, 표, 도표, 그래프, 천기도, 천체도, 벤-다이아그램(ben-diagram) 등으로 기록하는 경우가 이에 속한다.
6. 추리	관찰한 결과나 과거의 경험을 근거로, 이미 관찰한(발생한) 그 자체가 아니라 사실 뒤에 숨은 내용이나 사실을 뛰어 넘어 직접 지각할 수 없는 현상을 유추하는 과정으로, 여기서 유추된 사실은 실험에 의해 증명될 가능성이 거의 없다.
7. 상관관계 결정	관찰된 사실 내에 들어 있는 2 개 혹은 그 이상의 변인들 간의 관계를 알아내는 능력으로 관찰된 사실 내에 들어 있지 않은 부분을 유추하는 것이 아니라, 단지 주어진 사실 내에서 관계를 해석하는 것이다.
8. 인과관계 결정	이미 관찰된 사실이나 과거의 경험적 지식을 근거로 하여 실험, 관찰, 관측 결과나 자연적 결과와 효과에 대한 1가지 또는 그 이상의 원인을 확인하고 설명하는 것으로 보다 체계적이고 논리적이다.
9. 예상 (내삽, 외삽 포함)	일정한 경향성 또는 규칙성을 보이는 현재의 관찰 데이터를 근거로 아직 발생하지 않은 상황을 예측하거나, 사건의 일련성(sequence)을 유추하는 활동으로써, 앞으로의 실험이나 다른 사건에 의해 증명될 가능성이 많다.
10. 결론	실험, 관찰, 관측 사실을 일관된 체계로 설명(포괄적으로)하기 위해 수집된 데이터 군으로부터 중요한 모든 정보를 추출하고 합하여 종합적인 결론을 이끌어 내는 능력으로, 결론은 합리적으로 추출된 진술이고 논리적으로 완벽한 형태이다.
11. 일반화와 모델 형성	다양한 실험, 관찰, 관측 데이터를 서로 비교하거나 관련성을 조사해 공통된 특성이나 규칙성을 발견하여 이들(특성이나 규칙성)을 엮어서 보다 넓은 영역과 적용 범위를 갖는 일반 법칙 또는 자연 법칙까지로 확장하는 능력이다.

[표 3] 본 연구자가 내린 과학 탐구 과정 요소의 평가 목표

탐구 과정 요소	평가 목표
1. 가설설정	종속 변인과 독립 변인이 구체적으로 제시된 실험, 관찰, 관측 자료가 주어지면, 이로부터 검증 가능한 가설을 제안하거나 확인할 수 있다.
2. 변인통제	연구 문제나 가설이 주어지면, 주어진 문제나 가설을 검증하기 위해서 반드시 통제되어야 할 변인들을 확인하거나 통제할 수 있다.
3. 실험설계	연구 문제나 가설이 주어지면, 그 문제나 가설의 타당성을 검증하기 위해서 적절한 실험 설계를 할 수 있다.
4. 자료변형 중 숫자계산	실험, 관찰, 관측 데이터가 주어지면, 그 데이터를 실험, 관찰, 관측 문제를 해결하기 위해 간단한 계산을 통해서 좀더 유용하고 쉽게 설명할 수 있는 물리량으로 변형시킬 수 있다.
5. 실험값을 그래프로 나타내기	실험, 관찰, 관측 데이터가 표로 제시되면, 이것을 보다 명확하고 조직화된 자료 제시 형태인 그래프로 나타낼 수 있다.
6. 추리	실험, 관찰, 관측이나 자연 현상에서 관찰 가능한 사실이 표, 그래프, 그림, 언어 등으로 제시되면, 그 사실로부터 직접 관찰이 안되는 새로운 사실이나 성질을 유도할 수 있다.
7. 상관관계 결정	과학적 사실에 대한 결과가 표, 그래프, 그림, 언어 등으로 주어지면, 이들로부터 변인들 사이의 관계를 발견하여 언어나 수식으로 설명할 수 있다.
8. 인과관계 결정	실험적 또는 자연적 결과가 표, 그래프, 언어, 그림 등으로 주어지면, 이들 현상이 발생한 원인이나 증거들을 유추하여 설명할 수 있다.
9. 예상 (내삽, 외삽 포함)	다양한 실험, 관찰, 관측이나 자연 현상에서 관찰 가능한 데이터가 표, 그래프, 그림, 언어 등으로 분석될 경우, 데이터가 없는 중간 부분이나 밖의 영역까지 유추하여 설명할 수 있다. 혹은 특정한 실험, 관찰, 관측에 대한 결과나 사건 발생을 예상하여 기술할 수 있다.
10. 결론	실험, 관찰, 관측 내용이나 결과가 표, 그래프, 언어, 그림 등으로 주어지면, 실험 결과만을 근거로 이들을 종합하여 결론을 내릴 수 있다.
11. 일반화와 모델 형성	특정한 실험, 관찰, 관측으로부터 얻은 결론을 그 이외의 좀더 폭 넓은 상황에 적용할 수 있는지에 관한 실험, 관찰, 관측 내용과 결과가 주어지면, 이들 실험, 관찰, 관측 결과를 토대로 좀 더 일반화된 설명 체계로 제안할 수 있다.

4. 평가 문항의 개발 및 타당도, 객관도, 문항 명료성 점검

과학 탐구 과정 요소와 행동 목표 형태로 기술된 평가 목표에 근거하여 평가 문항을 개발하였다. 평가 문항 제작은 '선다형 문항의 개발 준거'로 Gronlund(1985)가 제시한 14가지 지침과 Nitko(1983), 이종성 등(1984), 황정규(1989)가 제시한 준거, 중앙교육평가원(1990, 1991a, b)이 제시한 문항 제작의 원칙과 방향 등을 바탕으로 이루어졌다. 총 90여 개의 문항을 개발한 후, 평가 목표에 합당하다고 판단되는 문항을 탐구 과정 요소별로 3 문항씩 선별하였다.

1) 평가 문항의 내용 타당도, 정답의 객관도, 문항의 명료성 점검 I

5명의 내용 전문가(3명의 대학 교수와 2명의 고등학교 지구과학 교사)에게 평가 목표, 개발된 평가 문항, 내용 타당도 점검표, 정답의 객관도 점검표, 문항의 명료성 점검표를 주어, 내용 타당도, 정답의 객관도, 문항의 명료성을 검토하도록 의뢰한 후, 이들의 점검 결과를 토대로 평가 문항을 수정·보완하였다.

(1) 내용 타당도 점검

본 연구에서는 내용 전문가에게 의뢰하여 판단하는 방법을 이용하였다. 내용 타당도를 판단하는 방법은 Hambleton, 등(1975)이 제안한 4단계 평정 방법(1=완전 부적절, 2=약간 부적절, 3=약간 적절, 4=대단히 적절)을 2단계 평정 방법(1=적절, 0=부적절)으로 수정하여 적용하였다. 점검표에 기록

한 결과를 토대로 구한 내용 타당도 지수(CVI:Index of Content Validity)는 총 165개의 응답(5명의 평정자×33 문항) 중 74%가 평가 목표와 일치했다. 널리 받아들여지고 있는 타당도 지수는 80% 이상(Doran, 1980; 황정규, 1989)에 비추어 볼 때, 이 값은 비교적 낮은 값이다.

(2) 정답의 객관도 점검

본 연구에서는 5 명의 내용 전문가에게 개발한 평가 문항과 정답의 객관도 점검표를 제시한 후 답안을 작성하도록 의뢰하여 그 결과를 분석한 결과 96.9%의 문항의 정답이 본 연구자와 일치했다. 정답의 객관도는 객관식의 경우 100%에 근접할수록 좋다고 판단할 수 있다. 그러므로 96.9%의 정답의 객관도는 객관도가 양호하다고 판단할 수 있다.

(3) 문항의 명료성 점검

3명의 내용 전문가에게 개발한 평가 문항과 평가 문항 명료성 점검표를 제시한 후 문제가 있다고 판단되는 점을 점검표에 직접 기록하도록 의뢰하여 그 결과를 근거로 평가 문항을 수정·보완하였다.

점검 결과 대부분의 문항에서 지적된 문제점을 수정·보완하였다. 위에 기술한 과정 (1), (2), (3)을 거쳐 수정·보완하여 개발한 지구과학 탐구 능력 평가 도구를 현장 검증 I의 연구 집단에 투입하였다.

2) 평가 문항의 내용 타당도, 정답의 객관도, 문항의 명료성 점검 II

개발한 평가 문항의 내용 타당도, 정답의 객관도, 문항의 명료성을 개선하기 위하여 현장 검증 I의 투입 결과를 근거로 수정·보완한 평가 문항을 탐구 과정 요소에 대한 평가 목표를 정확하게 이해하고 있으며 탐구 능력 평가 문항 개발에 관하여 다년간의 경험을 쌓은 과학 교육 전문가(대학교수) 4명과 현장 경험이 풍부(10년 이상)하고 입시 문제 출제 경험이 풍부한 현직 고등학교 과학 교사 2명 등 6명의 내용 전문가에게 I차 점검 때와 같은 방법으로 의뢰하여 점검한 결과를 토대로 I차 때와 동일한 방법으로 분석하였다.

(1) 내용 타당도 점검

총 198개 응답(6명×33 문항) 중 89%가 평가 문항 개발자와 일치하여 내용 타당도 지수가 89%로서 I차 점검 때 보다 향상된 값으로서 비교적 내용 타당도가 높은 평가 도구로 판단할 수 있다.

(2) 정답의 객관도 점검

총 198개의 응답(6명×33 문항) 중 196개가 본 평가 문항 개발자와 일치율을 보여 정답의 객관도는 99%가 되어 I차 점검 때 보다 많이 향상된 값으로써, 채점의 객관도가 비교적 높은 것으로 판단할 수 있다.

(3) 문항의 명료성 점검

내용 전문가의 점검 결과를 근거로 평가 문항을 수정·보완한 후 두 번째로 완성된 과학 탐구 능력 평가 도구를 현장 검증 II의 연구 대상 표본 집단에 투입하였다.

5. 개발된 과학 탐구 능력 평가 도구의 현장 검증 및 수정·보완

1) 현장 검증 I의 실시

완성된 평가 도구를 현장 검증 I의 연구 대상 표집에 투입하는 목적은 학생들이 개발한 평가 도구를 완성하는데 소요되는 시간의 적절성 여부를 결정하고, 평가 도구의 사용 방법에 관한 내용과 방향을 결정하며, 평가 도구를 구성하는 각 문항의 신뢰도, 난이도, 변별도, 문항 반응 분포에 근거한 오답의 효율성을 산출하여 그 결과를 토대로 평가 문항을 수정·보완함으로써, 우리 나라의 고등학생들에게 적용 가능한 지구과학탐구 능력 평가 도구를 개발하는데 있다.

(1) 연구 대상의 표집

모집단으로부터 연구 대상의 표집은 먼저 학교를 무선적으로 표집하고, 그 다음 학급 단위로 학생을 임의로 표집하는 집락 표집 방법(cluster sampling method)을 이용하였다. 인문계 고등학교 1.2.3 학년 학생을 대상 모집단으로 하여, 남·여 학교별로 1 개교를 표집한 후, 학년별로 1 개 학급씩 무작위로 표집하였다. 다음 [표 4]와 같이 304명의 학생을 표집하여 1991년 7월 중(7월 8일~7월 13일)에 현장 검증 I을 실시하였다.

[표 4] 현장 검증 I의 연구 대상의 표집

지역	학교	표집수			합계
		1학년	2학년	3학년	
중소도시	C 고등학교	50	52	51	153
읍 면	J 여고	49	51	51	151
합	2개교	99	103	102	304

(2) 문항 분석 및 평가 도구의 수정·보완

현장 검증 I에서 학생들의 반응을 분석한 결과 전체 평균이 18.71, 표준 편차 6.22, 신뢰도(K-R 20) 0.84, 변별도 지수 평균이 0.47, 정답율 평균이 56.9%, 난이도 지수 평균이 45.4%, 전체 측정의 표준 오차는 2.49이었고, 문항 반응 분포는 다양하게 나타났다.

이상적인 신뢰도의 범위에 대한 일치된 견해는 없지만, 보편적으로 받아들여지고 있는 값은 0.80이상이다 (Doran, 1980; Nitko, 1983). 그러므로 이 기준과 국내외에서 이미 개발된 평가 도구의 신뢰도 값에 견주어 볼 때 0.84란 신뢰도는 타당도의 선행 조건을 만족시키는 데 충분히 높은 편으로 판단할 수 있다.

어느 정도의 난이도가 이상적인가에 대해 일치된 견해는 없지만, 대개 20~80%의 문항 난이도를 가지고 전체 난이도의 평균치는 기재형(supply type)의 경우 50% 정도가 이상적이고(Doran, 1980; Aiken, 1987; 김창식 등, 1991), 본 연구에서는 5지 선택형이므로 60%가 이상적이다(Gronlund, 1985; 김창식 등, 1991).

이와 같은 증거에 비추어 전체 학생의 문항별 난이도 지수를 분석해 보면 전체적으로 허용 범위를 벗어난 6 문항 중 5 문항은 난이도가 낮게 나와 어렵다고 판단되어 쉽게 수정하였으며, 1 문항은 난이도가 높게 나와 쉽다고 판단되어 어렵게 수정하였다. 한편 전 학년에 걸쳐 난이도 지수가 0이하인 값을 나타낸 문항은 없었기 때문에 삭제해야 할 문항은 하나도 없었다.

널리 받아들여지고 있는 변별도 값은 Doran(1980)과 김창식 등(1991)은 + 0.3이상으로, Aiken(1987)과 Gronlund(1985)는 + 0.20 이상으로 보았다. 이 증거에 비추어 문항 변별도 지수를 살펴보면, 1 문항은 전체 학생의 변별도 지수가 너무 낮아 대폭적으로 수정·보완하였고 학년별로 문제가 있다고 판단되는 4 문항도 수정·보완하였다.

문항 반응 분포에서 오답에 대한 반응 분포를 보면 대체로 하위 집단에 속하는 학생들이 상위 집단에 속하는 학생들보다 더 많이 반응하고 있어 전반적으로 문항 구성의 조건을 만족하고 있으나 2 문항의 답지들 중 문항 구성의 조건에 부적합하다고 판단되어 수정하였다. 또한 각 문항의 5개 답지에 대한 피험자 전체의 반응 비율을 분석해 보면 약 50%의 피험자가 정답지에 반응하고 그 나머지의 피험자가 오답지에 고르게 분산 반응을 나타낼 때 문항이 학생의 능력 차를 예민하게 변별할 수 있고, 또한 오답은 오답으로서의 역할을 제대로 수행할 수 있다는 조건에 대부분의 문항은 만족하고 있으나 2 문항의 오답지들이 오답지로서의 역할이 부족하여 더욱 매력 있게 수정하였다.

2) 현장 검증 II의 실시

현장 검증 I의 결과를 분석하여 그를 근거로 수정·보완한 지구과학 탐구 능력 평가도구를 6명의 과학 교육 전문가에게 의뢰하여 타당도, 객관도를 점검한 결과(타당도, 객관도 점검 II)를 토대로 수정·보완한 지구과학 탐구 능력 평가 도구를 9월 중(1991, 9, 16~1991, 9, 20)에 현장 검증 II를 실시하였다. 현장 검증 II를 실시하는 목적은 현장 검증 I에서 어느 정도 문항의 효율성이 입증되었기 때문에, 첫째, 개발한 지구과학 탐구 능력 평가 문항이 우리 나라 고등학생들에게 적용 가능한가? 둘째, 신뢰도, 난이도, 변별도, 오답의 효율성을 산출하여 앞으로 평가 문항을 더 수정·보완할 필요가 있는지를 판단하는데 있다.

(1) 연구 대상의 표집

지역별, 성별, 학년별로 구분하여 [표 5]와 같이 총 872명을 무선 표집하였다. 연구 대상의 표집은 집락 군집 표집법을 이용하였다.

[표 5] 현장 검증 II의 연구 대상의 표집

지역	학교명	표집수(단위:명)				합계
		성별	1학년	2학년	3학년	
대 도시	대전 L고	남	47	52	52	151
	서울 S여고	여	51	51	52	154
중소도시	강원 춘천 K고	남	45	44	44	133
	충북 청주 C여고	여	50	52	51	153
읍 면	충남 연기군 J고	남	43	43	47	133
	충남 금산군 K여고	여	49	48	51	148
합	6개교		285	290	297	872

(2) 문항 분석 결과

현장 검증 I의 문항 분석에서 사용한 동일한 방법으로 현장 검증 II의 반응 결과에 근거하여 신뢰도, 정답율, 난이도, 변별도, 오답의 효율성을 산출하였다. 그 결과 전체의 평균이 19.29, 표준 편차 6.23, 신뢰도(K-R 20) 0.84, 정답율 평균 59.1%, 난이도 지수 평균은 49.4%, 변별도 지수 평균 0.47, 측정의 표준 오차는 2.49, 문항 반응 분포에 의한 오답의 효율성은 양호한 분포를 나타냈다. 전체의 신뢰도는 0.84로서 현장 검증 I의 값과 똑 같은 값으로써, 높은 신뢰도를 나타내고 있다.

문항 난이도 지수는 전체적으로 볼 때 문항 난이도 지수

의 허용 범위를 벗어난 문항이 3개였으나 범위를 크게 벗어나지 아니했고 학년별로 허용 범위를 벗어난 문항은 1문항(25번)이었다.

문항 변별도 지수 평균은 0.47이고, 전체적으로 볼 때, 문항 변별도 지수의 허용 범위를 벗어난 문항은 하나도 없었으며 학년별로 허용 범위를 벗어난 문항은 2문항이 있었다.

문항 반응 분포에서 오답에 대한 반응 분포를 보면 전반적으로 하위 집단에 속하는 학생들이 상위 집단에 속하는 학생들 보다 더 많이 반응하고 있어 문항 구성의 조건을 만족하고 있으며 각 문항의 5개 답지에 대한 피험자 전체의 반응 비율을 분석해 보면 약 50%의 피험자가 정답지에 반응하고 그 나머지가 오답지에 고르게 분산 반응해야 한다는 조건에 대부분 문항이 만족되고 있다.

이상과 같은 문항 반응 분석을 종합해 볼 때 신뢰도, 변별도, 정답율, 난이도, 문항 반응 분포에 근거한 오답의 효율성은 타당성이 있고, 신뢰로운 과학 탐구 능력 평가 도구로서 널리 수용할 수 있는 범위 이내에 속한다(Payne, 1974; Doran, 1980; 황정규, 1989; 김창식 등, 1991). 그러므로 상기한 결과를 근거로 판단해 볼 때 본 연구에서 개발된 지구 과학 탐구 능력 평가 도구는 더 이상의 수정을 고려하지 않아도 될 것으로 판단된다.

III. 개발한 지구과학 탐구 능력 평가 도구에 대한 기술

본 연구를 통하여 개발된 과학 탐구 능력 평가 도구에 대한 탐구 과정 요소별 문항 분포와 개관, 특징 그리고 도구를 구성하는 문항의 예를 제시하면 다음과 같다.

1. 지구과학 탐구능력 평가도구(Test of Earth Science Inquiry Skills)

본 연구에서 개발한 탐구 능력 측정을 위한 평가 도구의 명칭을 “지구과학 탐구 능력 검사(Test of Earth Science Inquiry Skills)”로 명명하였으며 영문 약자로는 “TESIS”라고 하였다. 다음 [표 6]에 제시된 바와 같이 11개의 탐구 과정 요소별로 각 3문항씩 총 33문항으로 구성되어 있다.

2. 지구과학 탐구능력 평가도구(TESIS)의 개관

본 연구에서 개발한 탐구 능력 검사의 특징은 [표 7]에 제

시되어 있다. 이 검사는 SIEI를 근거로 선정한 11가지의 탐구 요소에 대한 정의는 [표 2]에 기술되어 있으며, 평가 목표는 [표 3]에 기술되어 있다. 또한 [표 6]에 제시된 바와 같이 탐구 요소별로 각 3개 문항씩 총 33문항으로 구성되어 있다.

[표 6] 지구과학 탐구 능력 평가도구의 탐구 과정 요소별 문항 번호

과학 탐구 과정 요소	해당문항 번호
1. 가설 설정	4, 16, 28
2. 변인 통제	6, 18, 29
3. 실험 설계	5, 17, 30
4. 자료변형 중 숫자 계산	1, 13, 25
5. 실험 값을 그래프로 나타내기	2, 14, 26
6. 추 리	3, 15, 27
7. 상관 관계 결정	7, 20, 31
8. 인과 관계 결정	8, 11, 21
9. 예상(내삽, 외삽 포함)	9, 22, 32
10. 결론	10, 19, 24
11. 일반화 또는 모델 형성	12, 23, 33

[표 7]지구과학 탐구 능력 검사(TESIS)의 특징

1. 평가할 탐구 과정 요소	11가지 [표 1]
2. 총 문항 수	33
3. 문항 유형	객관식 5지 선다형
4. 적용 대상	고등학교 1, 2, 3학년 학생
5. 검사 시간	50분
6. 변별도 지수(D.I)	0.21 ~ 0.40 : 10문항
	0.41 ~ 0.60 : 20문항
	0.61 ~ 0.80 : 3문항 평균 : 0.47
7. 정답율(%)	21.0 ~ 40.0 : 6문항
	41.0 ~ 60.0 : 11문항
	61.0 ~ 80.0 : 15문항
	81.0 ~ 100 : 1문항 평균 : 59.1
8. 난이도 지수(%)	0.00 ~ 0.20 : 3문항
	0.21 ~ 0.40 : 5문항
	0.41 ~ 0.60 : 16문항
	0.61 ~ 0.80 : 9문항 평균 : 49.4
9. 신뢰도 지수(K-R 20)	0.84
10. 측정의 표준 오차(Sm)	2.49

평가 대상은 고등학교 1, 2, 3 학년, 문항 유형은 객관식 5지 선다형 지필 평가 문항, 검사 시간은 50분이다.

내용 타당도 지수(CVI)는 89%, 정답의 객관도는 99%, 전체 신뢰도(K-R 20)는 0.84, 전체 정답률 평균은 59.1%, 전체 난이도 지수 평균은 49.4%, 전체 변별도 지수 평균은 0.47, 전체 측정의 표준 오차는 2.49, 그리고 오답의 효율성은 전체적으로 양호하다.

이와 같은 결과는 본 연구에서 개발된 지구 과학 탐구능력 평가 도구가 타당하고 신뢰롭다는 증거가 된다.

IV. 결론 및 제언

1. 요약 및 결론

과학 교육 개혁 운동으로 도입된 탐구적 과학 교육에서는 과학 지식 자체보다는 과학 지식이 얻어지기까지의 과정 즉, 과학 하는 방법인 탐구 과정을 중요시하게 되었으며, 과학 교육 목표 역시 학생들로 하여금 탐구 과정 및 방법을 습득하여 탐구 능력 또는 과학적 문제 해결력을 향상시키는 데 두었다.

최근 국내외의 여러 가지 연구 결과에 의하면, 탐구 학습 방법이 탐구 능력뿐만 아니라 학업 성취도, 실험 기능 영역의 발달은 물론이고 과학적 태도 등에서 전통적인 학습 방법보다 높은 성취도를 보인다는 사실이 밝혀졌음에도 불구하고 과학 성취도가 기대 수준에 미치지 못하는 것으로 나타났다. 이는 탐구 능력 신장을 위한 현재의 탐구적 과학 교육에 큰 문제가 있다는 것을 강력하게 시사해 주는 것으로, 이러한 근본적인 원인은 여러 가지가 있겠으나, 탐구 학습이 현장에 제대로 투입되지 못한 원인 중의 하나는 탐구 능력 측정하는데 적합한 평가 도구가 제대로 개발되고 있지 못하다는 점이다.

따라서 본 연구에서는 과학 탐구 능력 평가 도구 개발을 위하여 SIEI에 근거하여 지필 평가가 가능한 11가지의 탐구 과정 요소를 선정하였으며, 각 탐구 과정 요소별로 그 의미를 명확히 하기 위하여, 선정된 탐구 능력 요소들에 대한 선행 정의를 토대로 각 요소들에 대한 정의를 내리고, 그 정의에 근거하여 평가 문항 개발의 준거로 이용될 수 있도록 Gagne의 교육 목표 진술 방식인 평가 조건과 요구 행동으로 이루어진 행동 목표 형태로 평가 목표를 설정한 후, 이를 근거로 평가 문항을 개발하였다. 평가 문항의 유형은 객관식 5지 선다형, 각 탐구 과정 요소별로 3개 문항씩 총 33문항, 검사 시간은 50분이며, 지구과학 교과의 상세한 지식이

없어도 응답할 수 있도록 문항 장면을 구성하였다.

개발된 평가 도구의 내용 타당도, 정답의 객관도, 문항의 명료성을 점검한 결과 내용 타당도면에서 74%, 정답의 객관도면에서 96.9%가 일치했다. 문제점들을 근거로 수정·보완하여 완성한 평가 도구를 검사에 소요되는 시간을 결정하고, 평가 도구의 사용 방법에 관한 내용과 방향을 결정하며, 우수한 평가 도구의 준거에 맞도록 수정·보완하기 위하여 중소도시·읍면의 고등학교 2개 학교 6개 학급 304명의 학생을 표집하여 현장 검증 I을 실시하였다. 그 결과를 토대로 신뢰도, 문항 변별도, 문항 정답률, 문항 난이도, 측정의 표준 오차를 산출하고, 문항 반응 분포에 근거하여 오답의 효율성을 구했으며, 이들을 근거로 평가 문항을 수정·보완하였다.

현장 검증 I의 결과를 근거로 수정·보완하여 완성한 평가 도구의 내용 타당도, 정답의 객관도, 문항의 명료성을 다시 점검한 결과 내용 타당도면에서 89%, 정답의 객관도면에서는 99%가 일치했다. 현장 검증 I의 결과와 과학 교육 전문가의 점검 결과를 근거로 수정·보완하여 완성한 평가 도구에 더 이상의 수정·보완해야 할 문제점이 있는가를 결정하기 위하여 대도시, 중소도시, 읍면의 6개 학교 18개 학급 총 872명을 무선 표집하여 현장 검증 II를 실시하였다. 그 결과를 토대로 신뢰도, 문항 변별도, 문항 정답률, 문항 난이도를 산출하고, 문항 반응 분포에 근거하여 오답의 효율성을 구하였다. 그 결과 전체 신뢰도(K-R 20)는 0.84, 문항 변별도 지수의 평균은 0.47, 문항 정답률의 평균은 59.1%, 문항 난이도 지수의 평균은 49.4%, 측정의 표준 오차의 평균은 2.49였으며, 문항 반응 분포에 근거한 오답의 효율성은 전반적으로 고르게 분포하는 양상을 나타냈다.

현장 검증 II의 결과 산출된 신뢰도, 변별도, 정답률, 난이도 지수의 학년별, 문항별 분포와 문항별 답지(정답과 오답 포함)의 분포는 우수한 평가 도구의 준거에 비추어 볼 때, 전반적으로 허용 범위 내에 있으며 양호하다고 판단된다. 이러한 현장 검증 II의 결과는 본 연구를 통하여 개발된 평가 도구가 신뢰롭고 타당하다는 강력한 증거가 된다.

본 연구에서 개발한 평가 도구는 지구과학 소재에 근거한 과학 탐구 능력 측정 자료로, 제한적으로는 탐구 학습 전반에 대한 진단 평가와 총괄 평가의 측정 자료로도 활용될 수 있을 것이며, 평가 도구의 개발 방법 및 과정은 과학 탐구 능력 평가 문항 개발의 지침으로 이용될 수 있을 것이다. 한편 일선 학교 과학 교사가 각 교과목 영역의 내용에 근거한 과학 탐구 능력 평가 문항을 개발할 때와 이와 유사한 과학 탐구 사고력 평가 문항을 개발하고자 할 때 기초 자료로 활용될 수 있을 것이다.

2. 제언

본 연구의 진행 과정에서 나타난 문제점과 결과를 바탕으로 더욱 효과적이고 객관적인 탐구 능력 평가 문항 개발을 위하여 다음과 같은 사항을 제언하고자 한다.

첫째, 과학 탐구 능력 측정은 과학 지식 내용과 관계없이 (content free) 탐구 능력을 평가해야 한다는 주장이 있으나, 이는 극히 제한적인 평가이기 때문에 학생들의 각 학년 수준에서 이미 학습된 개념과 관련된 탐구 능력 평가 문항 개발이 요청된다.

둘째, 첫 번째 문제점의 선행 조건으로 각 학년 수준에서 이미 학습된 개념에 대한 위계적인 연구가 필요하다.

셋째, 관찰-실험 기능을 제외한 탐구 사고력의 일부로서 문제 파악, 가설 형성, 변인 통제, 실험 장치 고안, 자료의 수집과 정리, 내삽, 외삽, 연역적, 귀납적 추론, 정성적, 정량적 추론, 종합 등과 같은 연계적 탐구 능력을 측정하는 탐구 능력 평가 도구의 개발이 요청된다.

넷째, 탐구 수행 능력 중 지필 검사로 측정이 어려운 심체적 영역에 속하는 "관찰, 측정 및 조작적 기능 등"은 실기 시험이나, 실험 방법 및 절차를 계획하고, 실험을 실시하여 결론을 도출하는 일련의 전 과정을 평가할 수 있는 내용으로 구성된 평가 문항 개발에 대한 연구가 요청된다.

다섯째, 본 연구를 기초로 하여 국민학교와 중학교 및 고등학교의 과목별 특성을 고려하여 각 교과에 관련된 내용만으로 구성된 탐구 능력 평가 도구의 개발이 필요하다.

참고문헌

김창식, 이화국, 권재술, 김찬중(1991). 과학 학습 평가. 서울: 교육과학사.
 문교부(1973). 제 2차 교육과정. 서울: 대한 교과서 주식회사.
 문교부(1981). 제 3차 교육과정. 서울: 대한 교과서 주식회사.
 문교부(1985). 제 4차 교육과정. 서울: 대한 교과서 주식회사.
 문교부(1988). 제 5차 고등학교 교육과정. 서울: 대한 교과서 주식회사.
 교육부(1992). 제 6차 고등학교 교육과정. 서울: 대한 교과서 주식회사.
 이종성, 강봉규, 한종철(1984). 교육심리 측정 평가. 서울: 중앙출판사.
 이종승(1989). 교육연구법. 서울: 배영사.
 임인재, 김영길, 유병웅(1986). 과학교육 성취도 평가 연구 (Ⅲ) 국내 연구 보고서. 중앙교육평가원 연구보고,

86-4.
 유갑열(1988). 한국과 일본 중학생의 과학적 사고력 및 탐구 과정 능력에 관한 연구. 물리교육, 6(2), pp.159~169.
 정완호(1989). 고등학교 교육 과정과 과학 교육, 과학교육. 박승재 편, 서울: 교육과학사, 'pp.157~192.
 정연태(1984). 과학 교육의 오늘과 내일. 서울: 한국 방송 사업단.
 중앙교육평가원(1990). 대학교육 적성 시험 1차 실험 평가 결과.
 _____(1991). 대학교육 적성 시험 2차 실험 평가 결과.
 _____(1991). 대학교육 적성 시험 3차 실험 평가 결과.
 한중하(1984). 미국 과학교육의 동향과 변화 : 1960~'80. 서울: 한국교육개발원 기타연구. OP 84-3.
 황정규(1989). 학교학습과 교육평가. 서울: 교육과학사.
 Aiken, L. R.(1987). Testing with Multiple - Choice Items. Journal of Research and Development in Education, 20, pp.44~55.
 APU(1980). Science in Schools, Age 15 : Report NO.1. Center for Studies in Science Education, University of Leeds.
 APU(1985). Department of Education and Science, In Science at Age 1. Sample Questions.
 Bingman, R. M.(1969). Inquiry Objectives in the Teaching of Biology. Boulder, Colorado: Biological Sciences Curriculum Study.
 Borg, W. R., and Gall, M. D.(1983), Educational Research ; An Introduction, 4th ed., New York: Longman Inc.
 Costenson, K., and Lawson, A.E.(1986). Why isn't Inquiry used in More Classrooms? The American Biology Teacher, 48, pp.150~158.
 Cronbach, L. J.(1984). Essential of Psychological Testing 4th ed., New York: Harper and Row, pp.75~135.
 Dillashaw, G. H., and Okey, J. R.(1980). A Test of the Integrated Science Process Skills for Secondary Science Students. Science Education 4, pp.601~608.
 Dillashaw, G. H., and Okey, J. R..(1980). Development and Validation of a Test of Enquiry Skills. Journal of Research in Science Teaching, 17(1). pp.7~16.
 Doran(1980). Basic Measurement and Evaluation of Science Instruction, Washington, D.C: National Science Teachers Association, pp.13~18.

- Gronlund, N. E.(1985). *Measurement and Evaluation in Teaching* 5th ed.. New York:Macmillan Publishing Co, pp.5~21, 169~212, 263~319, 346~378.
- Hambleton, P. K., Swaminathan, H., Algina, J., and Coulson, D.(1975). Criterion-Referenced Testing and Measurement:review of Technical Issues and Development. *An invited symposium presented at the annual Meeting of the American Educational Research Association*, Washington, D. C.
- Hur M.(1984). *Evaluation of Inquiry Activity in Science Curricula*, Doctoral Dissertation, Teachers College, Columbia University.
- Hurd, P. D., Bybee, R. W., Kahle, J. B., & Yager, R. E. (1980). Biology Educational in Secondary Schools of the United States. *American Biology Teacher*, 42, pp.338.
- Matthesis, F. E., Nakayama, G., Pottenger, F. M., and Jones, M.L.(1988). *Development of the Performance of Process Skills(POPS) Test For Middle Grades Students*, Microfilms, ED 365-252, pp.1~15.
- Moliter, L. L., and George, K. D.(1976). Development of a Test of Science Process Skills. *Journal of Research in Science Teaching*, 13(5), p.405~412.
- Nelson, M. A., and Abraham, E. C.(1973). Inquiry Skill Measures. *Journal of Research in Science Teaching*, 10(4), pp.291~297.
- Nitko, A. J.(1983). *Educational Test and Measurement an Introduction*, New York:Harcourt Brace Jovanovich, Inc.
- Padilla, M. J., Cronin, L. L., and 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.
- Payne, D.A.(1974). *The Assessment of Learning*, Lexington, Massachusetts: D.C. Heath and Co.
- Small, L.(1988). *Science Process Evaluation Model, Monograph*. ED 292 635, pp.1~19.
- Smith, K. A., Welliver.(1990). The Development of a Science Process Assessment for Fourth-Grade Students. *Journal of Research in Science Teaching*, 27(8), pp.717~738.
- Sund, R. B., and Picard, A.J.(1972). *Behavioral Objectives and Evaluational Measures: Science and Mathematics*, Columbus, Ohio:Charles E.Merrill Publishing Co.
- Welch, W.W., Klopfer, L.E., Aikenhead, G.S., & Robinson, J.T.(1981). The Role of Inquiry in Science Education :Analysis and Recommendations. *Science Education*, 65, pp.33~50.

(ABSTRACT)

The Development of a Test of Earth Science Inquiry Skills for High School Student

Jong-Ok Woo · Hang-Ro Lee
(Korea National University of Education)

Since the late of 1960,s, the improvement of science inquiry skills has been one of the most important goals in secondary science education.

To achieve this goal, it is essential to develop a valid and reliable instrument for evaluating science inquiry skills.

The purpose of this study is to develop a valid and reliable instrument for evaluating science inquiry skills for high school students.

The instrument is developed through R & D procedure, which includes two field trials of the instrument.

This study has formed a clear definition of the elements of science inquiry skills (formulating a hypothesis, controlling variables, designing an experiment, numeric calculation, graphing experimental data, inference, determining relationship, determining causalities, predicting experimental results <including interpolation and extrapolation>, drawing a conclusion, formulating a generalization or model), and established the goals of assessment and developed the items of assessment.

The instrument, Test of Earth Science Inquiry Skills, consists of 33 items which measure 11 science inquiry skills. There are content-free items for each science inquiry skills.

This study has selected 11 science inquiry skills which are considered appropriate for being evaluated by paper-and-pencil test with SIEI (Hur,1984).

The content validity of items, objectivity of the scoring keys and clarity of the items were checked by six experienced specialists in science education.

At the same time, the two field trials has been executed and produced the reliability of the instrument, item difficulty index, and the effectiveness of distracters.

The first field trial was performed using a sample of 304 high school students, and the second one using a sample of 872 high school students.

Because the content validity is 84 % and the reliability (K-R 20) is 0.84, the developed instrument in this study is considered valid and reliable.

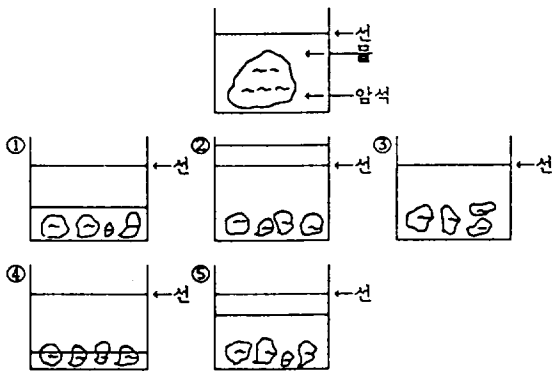
The difficulty index is 49.4 %, answer ratio 59.1 %, the discrimination index 0.47 and the effectiveness of distracters evenly distributed, which also suffice the criteria of good instrument.

The developed instrument in this study can diagnosis the well-developed science inquiry skills and the ill-developed science inquiry skills of the students, and trace the degree of the improvement of science inquiry skills.

<부록 > 지구과학 탐구능력 평가도구의 문항 예

본 연구에서 개발한 지구과학 탐구 능력 평가 도구를 구성하는 문항 9와 27 예를 제시하면 다음과 같다.

문제 9) 화강암을 물이 들어 있는 비이커 속에 넣었더니 물은 다음과 같이 선까지 올라 왔다. 만일 이 화강암을 4부분으로 조각 내어 넣는다면, 물의 높이는 어떻게 되겠는가? (단, 화강암은 틈이 없는 것으로 간주한다).



문제 27) 내륙 지방인 칠수네 마을에 도로 공사가 한창 진행 중이었다. 어느 날 산 중턱에서 지층과 암석을 깎아 낸 자리에 바다 고기 뼈의 화석이 대량으로 발견되었다. 이 바다 고기 뼈의 화석 모양은 현재 살고 있는 물고기 뼈와는 상당히 다른 형태를 띠고 있었다. 이로부터 이끌어 낼 수 있는 가장 적절한 설명은 ?

- ① 지하수에 살고 있던 물고기가 성장한 후 암석 속에 묻힌 것이다.
- ② 암석 틈의 물 속에 살던 물고기가 성장한 후 묻힌 것이다.
- ③ 과거에 살던 물고기가 흙 속에 묻힌 후 그곳이 상승한 것이다.
- ④ 육지와 연결된 동굴을 따라 올라온 물고기가 묻힌 것이다.
- ⑤ 누군가 이곳에 물고기의 뼈를 갖다 버렸다.