

초등학교 과학탐구력 측정을 위한 수행평가 도구 개발

한광래 · 김정길 · 김해경 · 남철우 · 송판섭 · 은경용*

광주교육대학교 과학교육과, *광주양산초등학교

Development of Performance Assessment of Scientific Inquiring Ability in Elementary School

Han Kwang-Lae, Jeong-Kil Kim, Hai-Gyoung Kim, Chul-Woo Nam,
Pan-Seob Song and Kyong-Yong Eun*

*Kwangju National University of Education, *Kwangju Yangsan Elementary School*

ABSTRACT

The present educational and social circumstances require the improvement of science education and the promotion of scientific technology simultaneously. Under this situation, it is necessary to develop the performance assessment evaluating the ability of scientific inquiry. The purpose of this study is to develop a valid and reliable instrument of the performance assessment that is can evaluate the scientific inquiring ability.

The characteristics of the instrument developed through this study, are as follows,

1. The performance assessment can be impartially achieved for all the elements of scientific inquiry, which are required in the units of elementary science from 3rd to 6th grade of elementary school.
2. To maintain the objectivity of performance assessment, the detailed standpoints and standards are established.
3. The instruments are devised to evaluate the thinking skills with the experimental reports of student, the operation skills by the check list of evaluation that the teacher's observation for the student are recorded.
4. Considering the level of elementary school student, the items required the complex thinking and the investigative skills are exclude as much as possible.
5. The items are arranged according to the learning steps of elementary school, so that the evaluation may be achieved in parallel with the process of the real teaching -learning in class.
6. According to circumstances, make use of these materials for the evaluation as the practical teaching-learning materials instead of the normal teaching-learning materials

The first field trial with the instrument was carried out, using a sample of 20 students from the 3rd to 6th at K and Y elementary school located in Kwangju city. The results of mean achievement quotient for each grade are as follows, the first term of 3rd grade (experiment and exercise skills ; 85%, inquiry thinking skills ; 74%), the second term of 3rd grade (experiment and exercise skills ; 81%, inquiry thinking skills ; 76%), the first term of 4th grade (experiment and exercise skills ; 70%, inquiry thinking skills ; 59%), the second term of 4th grade (experiment and exercise skills ; 61%, inquiry thinking skills ; 71%), the first term of 5th grade (experiment and exercise skills ; 84%, inquiry thinking skills ; 67%), the second term of 5th grade (experiment and exercise skills ; 73%, inquiry thinking skills ; 70%), the first term of 6th grade (experiment and exercise skills ; 83%, inquiry thinking skills ; 84%), the second term of 6th grade (experiment and exercise skills ; 87%, inquiry thinking skills ; 81%).

1. 서론

최근에 기초과학 및 첨단기술 그리고 국가간 지적 소유권 협정에 대한 사회적 관심이 고조되면서, 과학기술 진흥을 위한 과학교육의 책임과 역할을 강조하고 있다. 과학교육은 의도적이고 조직화된 인간의 활동이며 이러한 활동의 방향을 잡아주는 것이 과학교육의 목표이다. 20세기 후반부터 과학교육의 목표는 과학의 내용보다는 과정을 중시하고, 과학의 본질 및 과학의 사회적 역할을 중요시하는 방향으로 바뀌어 왔다. 이와 같은 세계적 과학 목표의 변화 추세와 더불어 과학 교육목표에 관한 연구들이 활발하게 진행되어 많은 종류의 과학교육 목표들이 개발되었고, 나라마다 탐구학습을 위한 다양한 자료를 개발하여 일선 과학 교육 현장에 적용되어 왔다.

평가란 본질적으로 교육과정이나 프로그램에 의하여 의도한 교육목표가 어느 정도 실현되었는가를 밝히는 과정이며, 프로그램에 관한 결정을 내리기 위해 정보를 수집하고 사용하는 과정이라고 정의할 수 있다(Tyler, 1949; Cronbach, 1964). 이러한 정의에 비추어 평가의 관점은, 학습목표에 제시된 행동 목표의 세부적인 탐구능력을 평가할 수 있도록, 가능한 어떤 행동의 변화를 기대할 수 있는가를 명확하고, 구체적으로 진술하여야 한다. 이와 같이 학습목표와 평가의

밀접한 관련성으로 인하여, 교육목표의 분류체계의 개발은 흔히 평가 준거의 필요성에 의해 추진되었다. 이러한 현상은 Bloom의 교육목표분류학(Taxonomy of Educational Objectives)이 발표된 이후에, Klopfer의 과학교육목표 분류체계로 이어졌고, 행동주의 심리학에 바탕을 둔 SAPA교육과정, 그리고 국가적 수준의 지원을 받은 영국의 APU의 평가 틀이나, 미국의 NAEP의 평가를 개발에도 계속되어 왔다. 우리나라의 경우는 교육부에서, 과학학력 평가가 과학교육 전반에 미치는 지대한 영향을 고려하여 6차 IBRD 교육차관 사업의 일환으로 전국 초·중등학교 과학 교육의 실태를 파악하고, 그 진흥방향을 모색하기 위한 연구를 과학 교육 진흥 자문단에게 위탁하여 1986년부터 3년간에 걸쳐 수행하였고, 초·중등 학생들의 과학 성취도를 측정하기 위한 과학 학습 평가틀을 개발하였다(박승재 외, 1986).

근래의 연구 보고에 따르면, 탐구학습에 대한 기대 수준과 실제 수준사이에는 상당한 차이가 있어, 탐구학습을 현장에 적용한 것이 그다지 성공적이지 못하였다는 결론을 내리고 있다. 과학교육자들은 이러한 탐구학습의 빈약한 결과가 탐구학습에 대한 혼돈, 탐구학습에 익숙하지 않은 과학교사, 결과주의에 만연된 사회 풍조와 암기 위주의 입시교육, 탐구교육에 필요한 행정적, 재정적 지원 부족 및 시설 부족 등에서 비롯된 것이라고 생각하고 있다. 그러나 이들보다 본질

적인 장애 요인의 하나는 탐구과정에 걸맞은 탐구학습 자료 및 이에 대한 구체적인 평가자료의 부족, 현장 평가 여건의 미비 등으로 인하여, 교육과정에서도 의도하는 목표에 따라 적절한 평가가 이루어지지 못하고 있는 실정에서 그 원인을 찾을 수 있다. 더구나 일선 교사들에 의해서 이루어지고 있는 탐구력 신장을 위한 수업연구도 그 효과를 검증할 수 있는 평가자료가 없다는 점에서 혼란이 일어날 수밖에 없고 막연히 그렇게 하는 것이 좋을 것 같다는 가정 아래 이루어지고 있다. 그러므로 정상적인 과학 교육을 위해서는 무엇보다도 입시위주의 교육이 벗어난 점수 만능주의와 결과주의에서 벗어나, 과학 학습 평가가 과학 교육을 위한 평가로서 원래의 기능을 찾도록 해야 된다.

과학학습에서의 평가는 탐구결과로서의 지식의 내용과 지식에 이르는 과정, 즉 자연을 조사하는 과정에 초점을 맞추고 있다. 그러나 전통적으로 써왔던 지필 평가에 대하여, 편협한 지식이나 단순한 수준의 평가에 그치며, 학습에 의해 성취된 사고와 문제 해결 능력에 대한 통찰력이 결여되어 있고, 학습자가 실세계를 이해하기 위하여 적용할 수 있는 능력을 측정하지 못하는 등 여러 가지 문제점이 제기되고 있다. 위와 같은 비판에 대응하여 미국 과학교육계에서는 지필 평가나 객관식을 지양하고, 실제평가(Authentic Assessment), 대안 평가 방법(Alternative Assessment Formats) 또는 “학교밖의 생활과 관련된 특정과제에 대해 학생들의 수행 실력을 직접 시험”하는 방안이 촉구되고 있다(Kauchak, 1995). 그러나 우리 나라의 경우, 실제 현장 연구 결과에 의하면 진도확보와 실험학습을 균형 있게 조화시키는 일은 현실적으로 어려움이 많고 일선 학교에서 정기고사시 문항을 출제하여 평가하는 경우 문항 작성상의 어려움이 많으므로 실험실습 및 탐구능력을 측정할 수 있는 평가 문항을 다양하게 많이 개발하여야만 한다고 보고하였다. 이제까지 개발된 탐구능력을 측정할 수 있는 평가도구들도 모의실험 상황에서 학생들의 탐구능력을 측정하는 지필평가

의 범주를 벗어나지 못하고 있는 실정이다(민경덕 외, 1985 ; 최병순 외, 1993). 이러한 상황을 감안할 때 학생들의 과학적 탐구능력을 종합적으로 평가할 수 있고, 타당도와 신뢰도가 높은 수행 평가도구의 개발은 과학교육과 과학기술의 진흥을 위하여 반드시 필요하다고 생각한다. 본 연구는 과학교육 진흥 자문단이 제시한 과학학력 분류체계 중 탐구력의 범주에 속하는 탐구사고력, 실험기능을 측정할 수 있고, 신뢰도와 타당도가 높아 초등학교 상급학년에 적용할 수 있는 현장 위주의 수행 평가도구를 개발하기 위한 것이다.

II. 선행연구

1. 수행평가의 이론적 배경

오늘날의 대다수의 과학교육 연구자들은 관념주의 실체론 및 상대주의 인식론과 구성주의 심리학을 받아들여, 학생들이 일상적인 경험을 통해서 갖게된 직관적인 관념에 대한 연구를 통하여 과학 학습지도의 전략과 방법을 제시할 뿐만 아니라 전통적인 학습의 한계와 문제점을 극복할 수 있는 새로운 과학 학습론을 추구하고 있다. 한편, 1980년대에 미국의 과학교육에 대한 연구 결과들은 과학교육의 심각성을 지적하면서, 미국이 위기에 처한 상태이며 교육개혁이 절실히 필요하다고 역설하고 있다(National Commission on Excellence in Education, 1983 ; National Science Board, 1989 ; National Assessment of Educational Progress, 1990).

우리 나라의 경우, 초등학교에서 중학교 그리고 고등학교로 올라갈수록 과학시간 중 실험시간이 점점 줄어들고 그나마 넉넉하지 않은 과학시간이 내용위주로 전개되며, 지식암기 위주의 과학교육이 되고 있다. 또한 성적 평가에서 상대적인 비교를 강조하기 때문에 동료들간의 지나친 경쟁을 조장하거나 선택형(객관식)시험 때문에 학생들이 창의성을 발휘할 기회가 거의 없으며, 입시 위주의 교육 활동 때문에 입시에서 취급되기 어려운 정의적 영역 등에 대한 교육은

소홀히 하는 등의 문제를 야기하고 있다. 위와 같은 비판에 대응하여 앞으로 교육평가는 전인 교육을 조장할 수 있고, 학생의 창의성이나 문제 해결능력 등, 고등 사고 기능을 신장시킬 수 있는 질적 평가로 전환해야 하며, 이러한 질적 평가를 위해서 대안 평가방법을 사용하거나, 학교 밖의 생활과 관련 있는 특정 과제에 대해 학생들의 수행 실력을 직접 시험하는 방안이 촉구되고 있다. 전통적으로 학생의 선발이나 배치를 목적으로 학생들의 학습 결과를 양적으로 측정해 온 평가 방식을 '양적 평가'라 한다면, 그와 반대로 수행평가는 "학생 스스로가 자신의 지식(knowledge)이나 기능(skills)을 나타낼 수 있도록 산출물을 만들거나, 행동 등으로 나타내거나, 답을 작성(구성)하도록 요구하는 평가방식"으로서, 교수-학습 과정의 질적인 측면을 평가하고 그 과정의 개선을 도움으로써 교사 개개인의 자질을 향상시키고 학생 개개인의 교육적 성장을 돕기 위한 평가를 '질적 평가'라고 할 수 있다(백순근, 1996). 대부분 대안 평가라는 용어가 통칭적으로 사용되고 있지만(Worthen, 1993) 통상적으로 대안평가, 실제평가, 참평가, 직접적평가, 실행위주의 평가 그리고 수행평가라는 용어들이 사실상 구분 없이 쓰여지고 있다. 그러나 최근에 우리 나라에서는 수행평가라는 용어를 많이 사용하는 경향을 보인다.

Shepard(1989)는 바람직한 수행평가가 갖추어야 될 요소를 다음과 같이 제시하고 있다. 1) 평가과제는 실제학습과제와 밀접해야 한다. 2) 복합적이고 도전적인 사고과정을 요구해야 한다. 3) 한가지 이상의 접근 방식을 고려해야 하며 평가 전에 제시한 증거에 입각해서 평가해야 한다. 이 증거는 학생들에게 평가 기준을 알려주며 교사에게는 학생의 실행 평가를 고르게 할 수 있도록 안내해 준다.

이상에서 알 수 있듯이, 수행평가란 실제학습 과정에서 학생이 수행하는 것을 직접 측정하므로써, 그들의 능력이나 숙달의 정도를 평가하는 것이며, 이와 같은 수행평가의 결과로써 교사들은 실생활 현장에 적용되는 학생들의 능력에 대

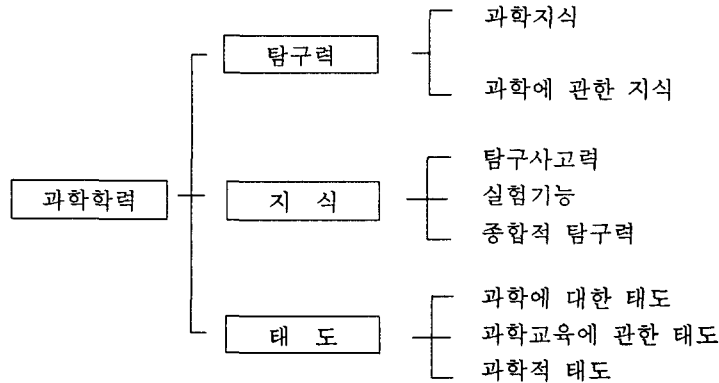
해 좀 더 정확한 정보를 얻을 수 있게 된다. 현재 널리 사용되고 있는 수행 평가의 종류로는 서술형(주관식) 검사, 논술형 검사, 구술 시험, 찬반 토론법, 실기 시험, 실험·실습법, 면접법, 관찰법, 연구보고서, 포트폴리오(portfolio) 등이 있다. 이러한 방법들은 새롭게 개발되었다기 보다는 과거에도 있었던 것이지만, 학습에 있어서 기존의 권위나 가치관에 대한 수동적인 수용보다는 자기 나름의 세계를 재창조하는 과정 중심의 인지 심리학을 근거로 하여 창의성이나 문제 해결력 등 고등 사고 기능을 파악하고 신장하기 위한 교육 평가의 방법으로 각광받고 있다(국립교육평가원, 1996 ; 백순근, 1995 ; 김민호, 1996).

2. 과학적 탐구력과 탐구요소

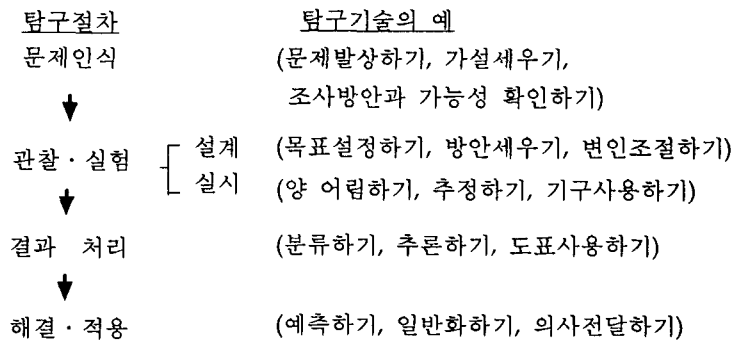
'사고'와 '사고력'은 다양한 의미와 수준으로 정의할 수 있겠지만, 사고란, '추론의 요소와 직관 창의의 요소가 긴밀한 상호작용을 통하여 문제를 해결해 나가는 체계적인 정신활동' 또는 '문제를 발견하거나 형성하기 위하여, 이렇게 발견하고, 형성된 문제의 본질을 이해하기 위하여 문제의 본질을 외적준거에 비추어 평가하고, 그 문제의 의의를 파악하기 위하여, 그렇게 파악된 문제를 해결하기 위하여 온갖 종류의 지적능력이 동원되어지고 활동되어지는 과정'이라 할 수 있다(성일제 외, 1988 ; 허경철 외, 1989). 이러한 의미에서 볼 때 사고는 문제를 해결하는 능력이며 보다 효율적으로 문제를 해결하려고 할 때, 문제의 대상, 한 개인이 가지고 있는 배경지식, 기본사고기능(추론, 직관, 창의의 요소)의 적절한 조작, 그리고 그 사람이 문제를 해결하기 위해 갖추어진 바람직한 성향에 크게 좌우되며, 이 밖에 호기심, 흥미, 태도 등의 요소도 사고에 영향을 주게 된다. 여기에서 성향이란 사고활동의 전개에 필수적인 사고하려는 개인의 욕구, 문제가 해결될 때까지 사고를 계속하려는 의지와 사고를 철저하게 하려는 마음가짐을 말한다. 효율적인 사고에 필요한 일반적 성향으로 제시된 것은 지적열성(intellectual dynamism), 지적성실(intellectual good faith), 지적인내(intellectual

perseverance), 지적정직(intellectual honesty), 그리고 지적겸손(intellectual humility)이다. 그리고 과학적 탐구력 (또는 사고력)은 자연현상이나 사물에 관련 인지된 문제를 해결하는 능력으로 정의할 수 있다. 이러한 과학적 탐구력은 일반적인 사고에 포함되지만, 문제가 자연현상과 사물에 관련하여 인지되고, 사고의 주체자가 가지고 있는 과학적 지식에 따라 논리적 실증적으로 처리되며 과학에 대한 흥미, 태도에 따라 과학적 문제 해결의 정도가 달라진다는 특성을 지니고 있다(이범홍 외, 1989). 보통 과학적 문제를 해결하기 위해서는 문제 발상, 가설 설정, 실험 설계, 자료의 수집과 정리, 자료의 해석 및 분석, 결과의 종합, 새로운 개념, 이론 및 법칙의 창출

단계를 순환하는 과정을 거친다. 이 때 각각의 단계를 수행하는데 필요한 요소를 탐구요소라 부르며, 그 동안 과학 탐구 요소에 관한 분류 체계는 여러 가지가 발표되었으나, 그 중에서 과학교육진흥자문단의 분류체계만을 소개하겠다. 우리 나라의 교육부는 제6차 IBRD 교육차관 사업의 일환으로 1986년부터 3년간에 걸쳐 우리나라 초·중등학교의 과학교육 실태를 과학교육진흥자문단에 위촉 조사하였다. 이러한 조사 결과를 바탕으로 학생들의 과학학력 성취도를 측정하기 위한 분류체계(그림1)를 마련하였다. 이러한 분류체계는 그림2의 과학적 탐구절차 모형과 탐구기술을 배경으로 작성되었다(박승재 외, 1986).



<그림 1> 과학교육진흥자문단의 과학 학력 분류체계



<그림 2> 탐구력 평가를 작성에 이용된 탐구모형

3. 최근 우리 나라의 과학탐구능력 평가

학생의 과학 탐구 실험 능력 개발은 현대 과학교육에서 가장 중요한 교육 목표로 간주되고 있으며 1992년에 개정 고시된 제 6차 교육과정에서도 과학 지식의 습득과 탐구 기능의 형성, 그리고 자연계에 대한 탐구활동 통해서 형성되는 과학적 태도의 함양은 자연과 교육의 중요한 목표로 되어 있다(교육부, 1992). 과학교육에 과학적 탐구력이 강조되는 이유는 과학의 본질과 과학교육의 목표, 그리고 교육의 실재를 함께 생각하여 봄으로서 그 실마리를 찾을 수 있다. 일반적으로 과학적 탐구력은 자연현상이나 사물에 관련되어 과학적 문제를 해결하여 가는 과정에서 요구되는 능력들을 의미한다. 따라서 문제해결 과정의 모형으로부터 탐구능력의 세부요소를 추출할 수 있다. 과학교육에서 이러한 탐구능력의 평가는 매우 중요함에도 불구하고, 우리 나라 학교 현장에서 탐구능력의 평가가 제대로 이루어지지 않는 실정이다. 지금까지 우리 나라에서 이루어진 과학탐구능력 평가에 관한 연구내용을 살펴보면 다음과 같다.

김명환 외(1981)는 '초등 과학 교육에서 탐구능력 평정 척도 작성과 그 적용에 관한 연구'에서 주로 국민학교 4, 5, 6학년 교과 내용을 중심으로 탐구요소를 선정하고 평가 문항을 개발하였다. 송형호 외(1984)는 '과학적 탐구학습을 위한 탐구과정 요소의 평가에 관한 연구'에서 국민학교 1, 4, 6학년을 대상으로 9개의 탐구요소(관찰, 분류, 기록전달, 측정, 예상, 추리, 변인조절, 데이터 해석, 모델형성 및 가설설정)에 대한 실험 평가 모델을 교과서 실험내용 중심으로 개발하였다. 김주훈 외(1984)는 '국민학교 자연과 평가의 원리와 실제'에서 국민학교 5, 6학년의 자연과 전 단원에 대한 내용 분석을 거쳐 지식·이해영역, 정의적 영역, 탐구과정 영역의 지필평가 문항을 개발하였다. 민경덕 외(1985)는 주로 중·고등학교 교과서 내용을 중심으로 탐구능력을 평가하기 위한 지필평가, 실험실기평가, 실험 보고서 및 실험 태도 평가, 학생자신 및 학생 상호평가 점검표 등의 평가 모형을 정립하고

평가 문항 및 평가 모델을 제시하였다. 이종기(1988)는 허명의 과학탐구표(SIEI)를 기준으로 우리 나라 고등학생을 대상으로한 TSIS(Test of Science Inquiry Skills)를 개발하였다. 이 평가 도구는 가설설정, 변인통제, 실험설계, 자료의 변형 중 숫자 계산, 실험 값을 그래프로 나타내기, 추리, 상관관계 결정, 인과관계 결정, 예상(내삽과 외삽도 포함), 결론, 일반화 또는 모델 형성, 평가 등 12개 탐구요소를 선정하고 각 요소당 3 문항씩 모두 36문항으로 구성되었다. 이항노(1991)는 지구과학 소재를 중심으로 한 고등학생의 과학 탐구능력측정을 위한 평가도구(Test of Earth Science Inquiry Skills : TESIS)를 개발하였다. 이 검사는 11개의 탐구 과정 요소마다 3문항씩 모두 33문항으로 구성되었으며, 검사 시간은 50분 간이다. 대상은 고등학교 1, 2, 3학년이며 문항의 형태는 객관식 5지 선다형이다. 박승재는 문교부 과학기술 교육차관 사업 중의 하나인 초·중등학교 과학교육의 실태조사와 진흥방안의 연구 보고서 중 과학학습 평가(국민학교)에서 과학적 탐구 사고력의 평가 범주를 첫째, 관찰·분류 및 측정 둘째, 문제 발견 및 해결방안 셋째, 데이터 해석 및 일반화로 정하고 탐구 사고력 평가 예시 문항을 개발하였다. 또한 그는 과학적 탐구 사고력 평가를 탐구 사고력 요소, 과학적 탐구 상황, 자연과학 기본 개념의 3차원으로 고찰하여 평가틀을 만들고 탐구 사고력 범주와 요소를 문제파악 및 정보수집, 관찰 실험 및 자료수집, 정보자료, 정보자료 종합해결, 검증 비판과 평가판단 등 5개의 대범주에서 다시 15개의 소범주로 구분하고 대학 수학능력 시험에 적용할 것을 제의하였다(박승재, 1989, 1991). 이무(1992)는 과학적 탐구 사고력 평가 문항 형태에 관한 연구를 하였다. 이 연구는 3단계로 이루어졌는데 1단계에서는 1990년 중앙 교육 평가원에서 실시한 대학 교육 적성 시험 1차 실험 평가 문항 중 수리 탐구 영역에서 과학 관계의 문항을 변형하여 각각 7가지의 평가문항 형태로 개발하고 각각의 문항 형태가 갖는 특징을 비교하였다. 2단계에서는 한 문항당 두 개의 소 문항

을 연결하여 개발하고 첫 번째 소 문항에서는 선택의 결과를 물었으며, 두 번째 소 문항에서는 그 답을 택한 과정을 묻는 문항을 개발하였다. 3 단계에서는 한 문항당 세 개의 소 문항을 연결하여 작성하고 첫 번째 소 문항에서는 선택의 결과, 두 번째 소 문항에서는 과정을, 세 번째 소 문항에서는 관련된 과학적 개념을 묻는 문항을 개발하였다. 한국과학교육단체 총연합회의 연구진이 교육부의 위탁을 받아 많은 교사들이 과학 탐구실험 평가에 관한 기초지식과 기술을 용이하게 익히고, 이러한 지식과 기술을 과학 수업에 활용하는데 이용할 수 있도록 하기 위하여 아주 자세히 기술한 과학 탐구 실험 평가 모형과 평가 도구를 개발하였다(이화국 외, 1992). 최근에 걸쳐 과학탐구능력 측정을 위한 평가도구의 개발과 관련된 연구를 살펴보면, 정완호 등(1993)은 국민학생의 과학 탐구능력 측정을 위한 평가도구를 개발하였고, 최병순 외(1993)은 지필 평가에서 나타난 학생들의 탐구능력과 실험과정에서 보여주는 탐구능력과 관계연구, 권재술, 김범기(1994)는 초·중학생들의 과학탐구 능력 측정도구의 개발하였다. 남정희(1996)는 화학실험에서 탐구과정 기능의 평가도구 개발, 김미경 외(1996)은 물리 탐구 실험의 평가를 위한 도구의 개발과 분석을 실시하였고, 조선형, 정수길(1996)은 초등학교 학생들을 위한 과학탐구능력 측정 도구 개발하였다. 최영철(1997)은 초등학교 자연과 평가문항 분석 및 실태조사를 하였다. 수행평가와 관련된 최근의 연구로는 남명호(1995)가 수행평가의 타당성 연구에서 과학 실기 평가, 실험보고서 평가, 그리고 컴퓨터 시뮬레이션 평가를 비교하였고, 박도순(1995)은 교육평가분야 연구의 최근 동향과 수행평가에서의 문제점을 제시하였다. 그리고 정귀향과 김범기(1997)는 초등학교 학생들의 측정 수행 능력을 평가하였다. 이외에도 수행평가라는 주제로 국내에서 발행된 문헌이 약간 있다(국립교육평가원, 1996 ; 석문주 외, 1997 ; 한국초등교육평가연구회, 1997).

III. 연구내용 및 방법

지금까지 초등학교에서 이루어진 자연과의 평가는 과학의 지식영역에 못미쳐 개념의 암기 정도를 알아보는 것이었고, 실험평가를 한다해도 이미 습득되어진 탐구기능이나 기술을 확인하는 정도로 그쳐, 학생들의 탐구수행능력을 평가하는데 소홀히 하였다. 교사들은 평가의 문제점으로 탐구능력에 대해서는 평가의 관점과 기준, 신뢰도의 객관성 등에 대하여 자신이 없고, 과학적 태도에 대해서는 평가도구가 부족하고, 평가기준이 모호하여 평가자의 주관이 개입되기 쉽다는 점등을 들고 있다. 특히 실기평가의 관점과 평가의 척도에 대해 많은 어려움이 있다고 보고된 바 있다(한국교육개발원, 1984 ; 권치순, 1988). 이러한 점들을 고려하여, 본 연구에서는 평가도구의 개발 방향을 다음과 같이 정하였다.

- 1) 자연과의 평가 영역 중 지식과 태도 영역을 제외하고, 탐구사고력과 실험·실습 기능에 포함되어 있는 탐구영역만을 평가 도구 개발 대상으로 한다.
- 2) 초등학교 자연과 3, 4, 5, 6학년의 단원에 대한 각각의 내용과 4개의 과학분야 영역에서 골고루 평가가 이루어지도록 평가도구를 개발하고, 이것을 종합하면 자연과의 탐구력에 대한 총괄 평가 성격을 띠 수 있도록 한다.
- 3) 수행평가의 객관성을 유지하기 위하여 평가관점과 그 기준을 상세화하여 기술하도록 한다.
- 4) 탐구능력 영역에서 탐구사고력의 평가는 주로 실험 보고서를 통하여 이루어지도록 하고, 실험·실습기능의 평가는 학생들이 실험·실습에 임하는 모습을 직접 관찰하여 평가관점표에 기록하므로써 이루어지도록 한다. 이 경우, 실험·실습 기능에 대한 평가 관점이나 기준을 일목요연하게 정리하여 평가점점표와 함께 제공한다.
- 5) 평가의 척도를 너무 세분화하면 실행상 어려움이 뒤따르므로, 삼단계(상·중·하)로 단순화하여 평가한다.

6) 초등학교의 학생들의 경우 복합탐구사고력을 측정하는데는 어려움이 뒤따르므로 될 수 있는 한 복합탐구사고력을 요구하는 평가문항은 지양하고, 개개의 탐구요소를 측정할 수 있도록 평가도구를 개발하는 한편, 상황에 따라서는 교사의 안내도 가능하다는 것을 전제로 한다.

7) 자연과의 단원 내용대로 수업을 진행하면서, 수행평가가 자연스럽게 이루어질 수 있도록, 학습의 단계에 맞추어서 평가문항을 순서대로 배열한다.

8) 수행평가의 목적이 반드시 학생의 성취도를 정량화 하는 것이 주된 것은 아니지만, 평가도구의 신뢰도와 타당성을 높이기 위하여, 직접 수행평가를 실시하게 되는 현장 교사들의 의견을 충분히 반영하고, 과학교육 전문가의 조언을 통하여 수정 보완이 이루어지도록 한다.

위에서 언급한 방향에 따라 다음과 같이 본 연구의 목표와 내용을 선정하였다.

첫째, 과학 평가에 관련된 국내외 문헌을 수집하여, 과학 평가에 대한 세계적 동향 및 주요 논쟁점과 문제점을 파악한다.

둘째, 이와 같은 문헌 조사 결과를 토대로 과학적 탐구능력의 본질과 과학 학습평가의 기초를 이해하는데 필요한 이론적 배경을 정리하여 제공한다.

셋째, 교과서의 단원분석을 통하여 학습과제를 추출하여, 세부적인 학습목표와 학습요소를 작성

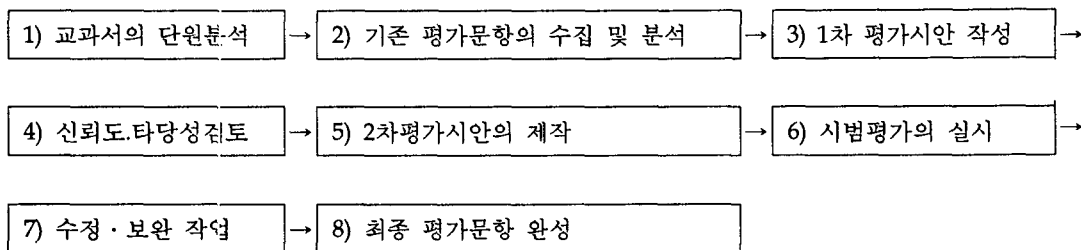
한다. 이렇게 작성된 세부학습 목표를 달성하기 위해 아동들에게 요구되는 탐구요소를 파악한다.

넷째, 현재 국내외에서 사용되고 있는 자연과에 관련된 평가문항과 탐구능력 검사지를 수집하여, 각 단원별, 영역별로 평가문항을 분리하고, 각각의 문항들이 측정할 수 있는 평가요소를 분석한다.

다섯째, 과학교육진흥 자문단이 제시한 과학학력 분류체계 중 탐구력 범주에 속하는 탐구사고력, 실험기능 및 종합 탐구력을 총괄적 측정할 수 있는 실험 또는 실험 주제를 학년별, 세부영역별로 선정하고, 측정 가능한 탐구요소에 대한 상세한 평가목표를 정한다.

여섯째, 위 단계에서 얻어진 자료를 기초로 수행 평가시안을 학년별, 영역별로 작성한다.

이렇게 제작된 1차 평가시안을 일선 자연과 교사와 과학교육 전문가에게 의뢰하여 그 신뢰도와 타당성에 대한 검토를 거친다. 이 결과를 토대로 1차 평가 시안을 수정 보완하여 2차 평가 시안을 제작한다. 그 다음, 광주권에 위치한 2개 초등학교 어린이들을 대상으로 시범평가를 실시하고, 시범 평가 결과를 바탕으로 과정에서 발생한 문제점이나 보완 해야할 점들을 일선 교사와 교육전문가와 논의하여 새로운 평가 문항을 제작한다. 본 연구에서 과학 탐구력 측정을 위한 수행평가 도구 개발을 위해 거치게 되는 전 과정을 그림 3에 세부 단계별로 나타내었다.



<그림 3> 초등학교 자연과 수행 평가도구의 개발을 위한 과정

IV. 연구결과

시를 표1에 나타내었다.

1. 평가문항 개발을 위한 이원분류표 작성

자연과 평가의 타당도를 높이기 위해서는 어느 특정한 교육목표와 지도 내용의 일부에 치우치거나 국한되지 않도록 평가틀이 마련되어야 한다. 이러한 평가틀로서 우리 나라 교육계에서 널리 이용되고 있는 것은 '행동 (behavior)'과 '내용 (content)'를 평가의 목표와 분류 기준으로 이용하는 이원분류표이다. 이번 연구에서 초등학교 자연과 수행평가를 위한 문항제작에 사용한 이원분류표는 평가 대상이 되는 각 학년과 학기별 단원의 내용에 대하여 학습목표 도달에 필요한 탐구요소로 작성되어 있다. 본 연구에서 사용한 이원분류표는 부록에 정리되어 있으며, 그 예

2. 평가문항 제작

앞에서 언급한 평가 도구와 개발의 방향, 평가의 목표와 탐구요소의 선정, 그리고 작성된 이원목적 분류표를 바탕으로 과학 탐구력 측정을 위한 수행평가 문항을 제작하였다. 제작한 수행평가 도구의 시안은 초등학교 자연과 담당교사와 광주교육대학교 과학교수들의 신뢰도와 타당성 검토를 통하여 수정 보완되었고, 시범평가를 한 결과를 토대로 다시 한번 초등학교 자연과 교사와 광주교대 과학과 교수들의 의견을 반영하여 다시 수정 보완이 이루어졌다.

본 연구에서 개발한 평가 문항의 특징은 다음과 같다.

첫째 ; 초등학교 자연과 3, 4, 5, 6학년의 전

<표 1> 초등학교 자연과 수행평가 도구 제작을 위한 이원분류표

학년 학기	내 용			탐구수행능력														비 고					
				실험실습기능						탐구사고력													
	단 원	소 단 원	평가 문항 선정 주제	관 찰	측 정	기 구 조 작	만 들 기	시 약 사 용	사 육 재 배	기 타	분 류	가 설 설 정	실 험 설 계	예 상	추 리	의 사 소 통	모 형 사 용		자 료 수 집 및 설 계	자 료 해 석	기 타		
3-1	1.시소 와 저울	(3)물체 의 무 게 재 기	양팔저 울로 물 체의 무 게 재기		○	○								○					○				
	2.개구 리와 배추 흰나 비	(2)배추 흰나 비의 한살 이	애벌레 가 자라 면서 변 하는 모 양관찰			○			○								○	○					
	3.식물 의 자람	(2)잎과 줄기	강낭콩 키우기	○					○										○				
	4.날씨	(1)온도 재기	교실안 의 온도 재기		○											○			○	○			

단원 내용과 과학분야 4개의 영역에서 요구되는 탐구요소에 대한 평가가 고르게 이루어질 수 있도록 고려하였다.

둘째 ; 지금까지 평가도구 개발에서 문제점으로 파악된 수행평가의 객관성을 확보하기 위하여, 평가관점과 그 기준을 상세화하여 제시하였다. 이 때, 평가의 척도를 너무 세분화하면 실행상 어려움이 뒤따르므로, 삼단계 (상·중·하)로 단순화하여 평가하도록 하였다.

셋째 ; 탐구능력 영역에서 탐구사고력의 평가는 주로 실험 보고서를 통하여 이루어지도록 하고, 실험·실습기능의 평가는 학생들이 실험·실습에 임하는 모습을 직접 관찰하여 평가관점표에 기록함으로써 이루어지도록 하였다. 이 경우, 실험·실습 기능에 대한 평가 관점이나 기준을 일목요연하게 정리하여 평가점검표와 함께 제시하였다.

넷째 ; 초등학교의 학생들의 경우 복합탐구사고력을 측정하는데는 어려움이 뒤따르므로 될 수 있는 한 복합탐구사고력을 요구하는 평가문항은 지양하고, 개개의 탐구요소를 측정할 수 있도록 하였다.

다섯째 ; 실제로 교실에서 진행되는 자연과 학습과 수행평가가 병행되어 이루어질 수 있도록 학습단계에 맞추어 수행평가 문항을 배열하였고, 상황에 따라서는 교사의 안내도 가능하다는 것을 전제로 하였다.

여섯째 ; 학생용 실험 보고서, 교사용 모범답안과 실험·실습 기능 평가점검표를 직접 학습자료로 사용 가능하도록 배려하였고, 평가문항 제작표에 학습의 준비와 평가에 도움이 될 만한 내용들을 포함시켰다.

3. 시범평가 결과

본 연구에서 2차 평가시안으로 시범평가를 실시한 대상은 광주 시내에 위치한 광주교육대학교 부속초등학교와 양산초등학교에 재학중인 3, 4, 5, 6학년 학생들이며, 평가 대상 인원은 20명 정도로 하였고, 실시기간은 1998년 3월초에서 1998년 5월말까지이다.

시범평가 결과 학년과 학기별 전체적인 평균 학업성취도는 다음과 같았다.

3학년 1학기 (실험·실습기능 ; 85 %, 탐구사고력 ; 74 %), 3학년 2학기 (실험·실습기능 ; 81 %, 탐구사고력 ; 76 %), 4학년 1학기 (실험·실습기능 ; 70 %, 탐구사고력 ; 59 %), 4학년 2학기 (실험·실습기능 ; 61 %, 탐구사고력 ; 71 %), 5학년 1학기 (실험·실습기능 ; 84 %, 탐구사고력 ; 67 %), 5학년 2학기 (실험·실습기능 ; 73 %, 탐구사고력 ; 70 %), 6학년 1학기 (실험·실습기능 ; 83 %, 탐구사고력 ; 84 %), 6학년 2학기 (실험·실습기능 ; 87 %, 탐구사고력 ; 81 %).

V. 결론 및 제언

과학교육에서 평가는 탐구결과로서의 지식의 내용과 지식에 이르는 과정, 즉 자연을 조사하는 과정에 초점을 맞추고 있다. 그러나 전통적으로 써왔던 지필 평가에 대하여 편협한 지식이나 단순한 수준의 평가에 그치며 학습에 의해 성취된 사고와 문제 해결 능력에 대한 통찰력이 결여되어 있고, 학습자가 실세계를 이해하기 위하여 적용할 수 있는 능력을 측정하지 못하는 등 여러 가지 문제점이 제기되고 있다. 또한 성적 평가에서 상대적인 비교를 강조하기 때문에 동료들 간의 지나친 경쟁을 조장하고, 입시 위주의 교육 활동 때문에 정의적 영역 등에 대한 교육을 소홀히 하는 등의 문제를 야기하고 있다. 이러한 상황을 감안할 때, 학생들의 과학적 탐구능력을 종합적으로 평가할 수 있고 타당도와 신뢰도가 높은 수행 평가도구의 개발은 과학교육의 정상화와 과학기술의 진흥을 위하여 반드시 필요하다고 생각한다.

지금까지 자연과 수행평가 도구를 개발하는 과정에서 나타난 문제점은 다음과 같다.

첫째 ; 수행평가의 목적이 학생의 성취도를 반드시 정량화 하려는 것은 아니지만, 평가 관점과 평가 기준을 상세화 할 때, 연구자의 주관적인 견해가 작용하였으며, 문항의 타당도와 신뢰도는 광범위한 현장 적용과 전문가의 의견을 바탕으

로 보완되어야 한다고 생각된다. 둘째 ; 수행평가의 특성에 뒤따르는 문제이지만, 많은 시간과 여러 명의 평가 요원이나 실험 조교가 동원되어야 했다. 셋째 ; 수행평가에서 지역적 특성이 고려되지 않았고, 조별 탐구능력에 대한 평가를 소홀히 하였다. 넷째 ; 수행평가의 도구 개발에는 많은 시간이 소요되나 연구기간이 너무 짧아서 시범 평가의 결과를 바탕으로 평가시안을 수정 보완하여 재투입을 해 볼 시간적인 여유가 없었다. 이상의 논의를 통하여 몇 가지 제언을 하겠다.

첫째 ; 평가의 관점은 그 시대의 교육관에 따라 달라지게 마련이다. 수행평가의 정착을 위해서는 먼저 쉽게 변할 것 같지 않는 우리 사회의 제도적, 사회적 가치관이 달라져야 한다.

둘째 ; 수행평가는 학습의 일부분으로 이루어져야 하며, 학생 개인의 특성과 수준을 파악하여 학생들의 창의성과 문제해결 능력을 키워주는 방향으로 이루어져야 한다.

셋째 ; 수행평가 도구의 개발을 위하여 개인보다는 국가적 수준이나 연구 단체별로 체계적인 연구사업이 전개되어야 한다. 특히 수행평가 도구를 개발할 때, 모든 단계마다 현장교사들을 적극적으로 참여시켜, 그 결과에 대해서 전문적인 분석과 함께 현장 적용 가능성 여부에 대한 검토가 이루어져야 한다.

넷째 ; 수행평가의 실시는 교사들에게 시간과 노력면에서 부담이 될 수밖에 없다. 이러한 문제를 해소하기 위해서는 연구보조원을 배치하거나, 최근에 논의되고 있는 수습교사 또는 과학교과 전담제의 실시가 바람직하다고 생각한다.

끝으로, 위에서 논의한 수행평가를 위한 전제 조건을 만족하지 않는 한, 과학교육의 개선은 어려운 것으로 예상하며, 과학교육 관련단체와 연구소를 중심으로 정책적인 판단에 관심을 기울여야만 할 것이다.

* 본 연구는 1997년 교육부 교과교육공동연구 학술지원비에 의하여 수행되었음.

참고 문헌

- 교육부(1992). 국민학교 교육과정. 서울.
- 국립교육평가원(1996). 수행평가의 이론과 실제. 대한교과서 주식회사, 서울, 39-50.
- 권재술, 김범기(1994). 초·중학생들의 과학탐구능력 측정 도구의 개발. 한국교원대 과학교육 논문집, 4 (1) : 301-314.
- 권치순(1988). 바람직한 자연과 학습평가 방안의 탐색, 서울교육대학 과학교육 연구, 제 5집.
- 김명환 외(1981). 초등학교 과학교육에서 탐구능력의 평정척도 작성과 그 적용에 관한 연구. 대구교대 논문집, 1-86.
- 김미경 외(1996). 물리 탐구 실험의 평가를 위한 도구의 개발과 분석. 한국과학교육학회지, 16 (1) : 51-60.
- 김민호(1996). 국민학교 6학년 학생의 수학 수행 평가 결과에 관한 분석. 한국교원대. 석사학위논문.
- 김주훈, 이양락(1984). 국민학교 자연과 평가의 원리와 실제. 한국교육개발원 연구 보고.
- 남명호(1995). 수행평가의 타당성 연구 : 과학실기 평가, 실험보고서 평가, 컴퓨터 시뮬레이션 평가의 비교, 고려대.
- 남정희(1996). 화학실험에서 탐구과정 기능의 평가도구 개발. 한국교원대. 박사학위논문.
- 민경덕, 권병규, 여환진, 박덕규, 안병호, 정원우, 이병교(1985). 중학 과학 교육의 탐구 학습에 대한 평가 문항 개발 (Ⅲ). 경북대학교 과학교육 연구지, 6 : 1-49.
- 민경덕, 양홍분, 이선행, 정원우, 이병교, 김유한, 장훈(1985). 중학 과학 교육의 탐구 학습에 대한 평가 문항 개발 (Ⅱ). 경북대학교 과학교육 연구지, 9: 1-39.
- 민경덕, 이병교, 정원우(1985). 고등학교 과학교육의 실험 실습 평가 실태 및 개선 방안. 경북대학교 과학교육연구지.
- 박도순(1995). 교육평가 분야 ; 연구의 최근 동향과 수행평가의 문제, 교육 진흥 27 : 120-128.
- 박승재(1989). 과학학습 평가 - 국민학교- 연구 보고서. 문교부.
- 박승재(1991). 과학적 탐구 사고력 평가, 서울

- 대학교 물리교육과, 물리학습연구실.
- 박승재 외(1986). 고등학교 과학교육의 실제 분석과 진흥 방안 및 점검 체제 연구.
- 백순근(1995). 학력평가를 위한 새로운 대안, 교육개발, 97, 44-48.
- 백순근(1996). 교수-학습 과정을 개선하기 위한 질적 평가 방안. 교육개발, 99, 85-89.
- 석문주 외(1997). 수행평가, 교육과학사, 서울.
- 성일제 외(1988). 사고력 신장을 위한 프로그램 개발 연구 (II), 한국교육개발원.
- 송영호, 채남식(1984). 과학적 탐구학습을 위한 탐구과정 요소의 평가에 관한 연구, 전주교대 과학교육연구소 논문집 10 : 1-37.
- 이무(1992). 과학적 탐구 사고력 평가 문항 형태에 관한 연구. 서울대 박사 학위 논문.
- 이범홍 외(1989). 과학적 사고력 신장프로그램 개발을 위한 방안 탐색, 한국교육개발원.
- 이종기(1988). 고등학생의 과학탐구능력 측정을 위한 평가도구개발, 한국교원대, 석사학위 논문.
- 이항노(1991). 고등학생의 과학탐구 능력 측정을 위한 평가도구 개발, 한국교원대 석사학위 논문.
- 이화국, 권치순, 김창식, 한안진(1992). 과학탐구 실험 능력 평가 모형 및 평가 도구 개발에 관한 연구, 한국교원단체총연합회.
- 정귀향, 김범기(1997). 초등학생들의 측정 수행능력 평가, 한국과학교육학회지 17 (2) : 127-138.
- 정완호 외(1993). 국민학생의 과학 탐구 능력 측정을 위한 평가 도구 개발, 한국교원대 과학교육논문집 3 (1) : 87-98.
- 조선형, 정수길(1996). 초등학교 학생들을 위한 과학탐구능력 측정 도구 개발, 청주교대 과학과수학교육 논문집 17 : 91-114.
- 최병순 외(1993). 지필평가에서 나타난 학생들의 탐구능력과 실험과정에서 보여주는 탐구능력과의 관계 연구, 한국교원대 과학교원논문집 3 (1) : 218-227.
- 최영철(1997). 초등학교 자연과 평가문항 분석 및 실태조사. 한국교원대, 석사학위 논문.
- 한국교육개발원(1984). 국민학교 자연과 평가의 원리와 실제, 한국교육개발원 개발 연구, TR 84-7.
- 한국초등교육 평가 연구회(1997). 수행평가. 교학사. 서울.
- 허경철 외(1989). 사고력 신장을 위한 프로그램 개발 연구 (III), 한국교육개발원.
- Cronbach, L. J.(1964) Evaluation for course improvement. In Robert W. Health (Ed), New curricula, New York ; Harper & Row.
- Kauchak, D.(1995). Trends in Elementary Science Education in the United States , 한국초등과학교육학회지, 14. 104-111.
- Klopfer, L. E.(1971). Evaluation of Learning in Science, In, B. S. Bloom, J. T. Hasting & C. F. Madaus (ed). Handbook of Formative and Summative Evaluation of Student Learning, New York : McGraw-Hill.
- NAEP(1990). Accelerating Academic Achievement -A summary of findings from 20 years of NAEP, Educational Tests Service, Princeton, NewJersey .
- National Commission on Excellence in Education(1983). A Nation at Rasle - The Imperative for Education Reform ; U. S. Dep. of Edu. Washington, D. C.
- National Science Board(1989). Science and Engineering Indicator ; U. S. Government printing office, Accelerating Academic Achievement - A Summary of Findings From 20 Years of NAEP.
- Shepard, L.(1989). Why we need better assessment, Educational Leadership, 46(7) : 4-6.
- Tyler, R. W.(1949). Basic principles of curriculum and instruction, Chicago : The University of Chicago press.
- Worthen, B.(1993). Critical issues hat will determine the furture of alternative assessment, Ph ; Delta Kappan, 74: 444-454.

(1998년 12월 14일 접수)