

수행평가를 위한 공통과학(생물분야) 탐구실험 평가 도구의 개발 연구

정충덕* · 양순택*

I. 서 론

과학교육에서 탐구과정을 중요시하는 탐구중심의 교육과정은 1973년 2차 교육과정의 개정을 통해서 우리 나라에 처음 도입되었고(이원식, 유경로, 1983), 이후 5차 교육과정과(김미경 외, 1995) 6차 교육과정에서도 지속적으로 강조되었으며(교육부, 1992), 7차 교육과정에서도 탐구능력과 과학관련 태도 향상을 목표로 하고 있다(교육부, 1997).

오늘날 학교현장에서는 결과중심의 평가에서 과정을 중시하는 평가로 전환하기 위한 움직임이 확대되고 있으며, 수행평가를 도입하여 실시하고 있다(강순희 외, 2000). 그러나 수행평가와 같은 탐구교육과정은 많은 시간이 소요되고, 평가도구의 개발이 곤란하고, 교사에게 많은 부담을 준다는 점 등 여러 가지의 단점을 가지고 있다(허명, 1987). 그 동안 많은 연구를 통해서 이러한 단점을 극복하고, 탐구능력을 측정하기 위한 도구를 개발하려는 노력이 진행되었다. 과학탐구 평가표의 개발이나(허명, 1984), 과학탐구 능력 측정을 위한 평가도구의 개발(이종기, 1988; 권재술, 김범기, 1994; 김미경 외, 1995; 남정희, 1996; 강순희 외, 2000)이 여러 분야에서 다양하게 이뤄졌다. 또한 우리 나라 생물 교과서의 탐구지수 분석이나 학교 현장에서의 실험 실태에 대한 조사가 이뤄졌다(한기영, 1987; 정충덕, 박행신, 1995). 이외에도 과학 탐구능력 측정을 위한 행동요소 추출(우중옥 외, 1991; 이항로 외, 1994)과 탐구실험의 개발 등이 이루어졌다(조희영 외, 1994; 김현식, 1986).

이와 같은 노력은 모든 과학 교과에서 필요하며, 탐구력 신장이라는 과학 교과

* 제주대학교 과학교육과 교수

** 삼성여자고등학교 교사

의 목표를 생각해볼 때, 탐구 능력의 신장을 위한 탐구 실험을 개발하고, 이를 통하여 탐구 능력을 평가하기 위한 평가 도구의 개발이 필요한 실정이다. 또한 과정 중심의 평가 방법인 수행평가를 위하여 탐구능력 뿐만 아니라 과학적 태도와 실험기능 등 여러 가지 측면을 측정할 수 있는 도구로서 실험평가에 대한 연구는 매우 중요하다고 사료된다.

본 연구의 목적은 고등학교 과학 교과 과정 중 공통과학 생물분야에서 학생들의 탐구능력과 과학관련 태도 향상을 위한 탐구 실험 평가 도구를 개발하여 학생들의 탐구 과정 기능을 측정할 뿐만 아니라, 실험의 기능과 태도의 측면까지도 측정할 수 있는, 진정한 의미의 수행평가를 위한 평가 도구를 개발함에 있다.

II. 이론적 배경

1. 과학과 탐구 실험

과학의 과정으로서 탐구 과정의 지도는 과학실험과 밀접한 관계를 맺고 있다고 알려져 왔다. 하지만 학교의 과학교육에서 과학실험이 곧 탐구적 과학교육을 의미하지는 않는다(김현식, 1986). 교과서에 제시된 일반적 실험은 이미 알려진 과학적 원리나 현상 혹은 법칙을 먼저 학습하고 이를 확인하기 위한 확인실험으로 이뤄져 있다. 확인실험은 수동적인 학습활동이며 교사 중심의 활동이다(강순희 외, 2000). 대부분의 교사는 교과서에 내재된 실험의 내용이 비탐구적이라는 생각을 갖고 있으며, 실험활동 시에도 한 실험조의 인원이 7명 이상인 경우가 대부분이며, 생물과 교수학습 활동이 탐구과정과 실험 중심의 수업이 아니라 지식 중심의 수업으로 진행되고 있다(정건상, 1991). 이에 비해 탐구실험은 과학적 원리나 법칙을 학습하기 이전에 실시하고 얻어진 실험결과를 바탕으로 새로운 개념이나 법칙, 이론을 탐구하도록 하는 실험유형이다(강순희 외, 2000). 학생 스스로가 문제를 생각하고, 실험을 진행하고, 자료를 수집하고, 분석·종합하는 과정을 통해서, 과학자들이 어떻게 연구하는가, 또 새로운 과학 지식이 어떻게 얻어지는가를 이해 할 수 있다(허명, 1992). 또한 학습지도는 학생들의 지적수준과 인지적 과정에 맞추어 이뤄져야 할 것이다(조희영, 1992).

실험평가가 진정한 의미의 수행평가 도구로서 탐구력의 측정을 위해 사용될

수 있도록 하기 위해서는 평가를 위한 실험이 확인실험이 아닌 탐구실험이어야 하고, 실험의 수준이 학생의 수준에 맞는 것이어야 한다. 이를 위해 실험활동 전이나 중간, 실험 후에 교사나 혹은 동료 학생들과 토의를 할 수 있도록 하고, 이 경우 다른 사람과의 상호 작용이 일어나므로 좀 더 효과적인 과학학습을 할 수 있다(백순근, 1999). 그리고 학생들간의 수준과 특성에 맞는 탐구활동을 제공하여야 하고, 학생들이 모자라는 부분을 보충할 수 있는 자료를 제공해야 한다(조희영, 1992).

2. 탐구 실험의 평가

실험을 평가하는 방법에는 다양한 방법이 있으나, 어떠한 평가방법을 사용하더라도 가장 중요한 점은 어떤 영역을 평가할 것인지 그리고 학생들의 수준을 어떻게 구별할 것인지, 그렇게 구분하는 근거가 무엇인지 제시되어야 할 것이다(한국교육과정평가원, 1999). 이와 아울러 평가에 대한 신뢰도와 타당도에 대한 전문가의 검토와 수정·보완이 필요하다. 이러한 과정을 거침으로써 학생들의 탐구실험에 대한 유용한 평가기준이 마련될 수 있을 것이다.

1) 탐구 과정

탐구라는 말은 탐구과학, 탐구생활, 탐구수업, 탐구실험 등의 용어에서 많이 사용하고 있지만, 일반화된 정의를 찾아내기는 어렵다. 또한 과학에서 탐구는 특별한 방법으로 이뤄지는 것이 아니므로 탐구과정을 한마디로 정의하는 일은 어렵다. 그러나 학생들의 탐구능력을 측정하기 위한 평가표를 제작하기 위해서는 탐구과정의 상세화가 필요하고, 각각의 탐구과정에서 어떤 기능을 평가해야 할 것인가 하는 것은 매우 중요하다.

슈바브(Suhwab)는 탐구과정들을 여섯 단계로 제안하고 있으며, 미국의 과학진흥 위원회의 교육과정위원회에서 개발한 SAPA프로그램에서는 13단계로 구분하고 있다(김정화, 1993). 허명(1984)은 고등학교 생물 교과서에 나오는 실험을 평가하기 위한 과학적 탐구 평가목록에서, 탐구 기능을 평가할 수 있는 요소를 크게 4가지 영역으로 구분하였다. 김현식(1986)의 경우 4가지로, 이종기(1988)는 12가지 탐구 과정 요소를 선정하였다. 우종욱 등(1991)은 9개의 행동요소로, 이항 등(1994)은 11가지 탐구요소로 구분하였다. 권재술 등(1994)은 기초탐구능력파

통합적 탐구능력으로 구분하고 이를 각각 5개의 탐구요소로 구분하였으며, 김미경 등(1995)은 다섯 가지 영역으로 구분하고, 이를 평가하기 위한 16개의 세부 탐구 과정 요소를 선정하였다. 강순희 등(2000)은 12개의 탐구능력 요소로 구분하고 이를 기초적 탐구능력과 통합적 탐구능력으로 구분하였다. 이와 같이 각각의 연구자들은 연구 목적에 따라 다양하게 탐구과정을 분석하고 있음을 보여준다.

2) 평가 방법

이러한 탐구 과정 기능의 평가에 사용될 수 있는 평가 방법 중 실험 수행 과정을 평가할 수 있는 평가 방법은 다양하다(한국교육과정평가원, 1998). 첫째, 지필 평가로서 서술형 평가와 논술형 평가를 들 수 있다(백순근, 1997; 이범홍, 1997). 둘째, 보고서 평가로 보고서에는 실험 보고서, 관찰 보고서, 조사·연구 보고서 등 다양한 형태가 있을 수 있고(이돈희 외, 1997), 셋째, 학생들의 활동을 방해하지 않고 교사가 자연스럽게 관찰하는 것을 통해 이뤄지는 관찰 평가이며(백순근, 1997; 박호순 외, 1999), 넷째, 포트폴리오를 이용한 방법으로(박호순 외, 1999), 학습성과, 학습 진전도, 성장 발달 정도 등을 보다 타당하고 신뢰성 있게 평가할 수 있다는 것이다(김현재, 1998). 다섯째, 개념도 및 V도를 이용한 방법으로, 개념도를 통해 학생들이 가진 개념간의 관계를 부각시켜 오개념이 탐지될 수 있다(박승재, 조희영, 1995; 김현재, 1998). 또한 V도는 학생들의 탐방 내용이나 실험 내용에 대한 이해와 방법에 대한 이해 등을 평가하는데 적절하다(박승재, 조희영, 1995; 백순근, 1999).

3. 탐구 능력 평가에 대한 선행 연구

우리 나라에서의 과학 탐구능력의 평가 연구는 1973년 이후 제3차 교육과정기부터 관심의 대상이었으나 과학교육에 대한 평가는 주로 내용 성취도 평가 수준에 머물러 왔다(김미경 외, 1995). 그러다가 제4차 교육과정기부터는 초등 학교 교육과정 자연과의 지도 및 평가상의 유의점에서 탐구능력의 평가를 하도록 명문화하였다. 이에 따라 평가도구의 개발에 대한 연구도 이루어졌는데, 많은 연구를 통해 '무엇을 평가할 것인가' 및 '어떻게 탐구 능력을 평가할 것인가' 하는 문제 해결을 위한 연구가 진행되었다.

‘무엇을 평가할 것인가’에 대한 연구로서 과학탐구 평가표의 개발에 관한 연구가 이루어졌는데(허명, 1984), 이를 통해서 탐구과정 모델에 근거하여 하나 하나의 탐구과제를 분류하여 이를 평가할 수 있도록 하고, 탐구활동의 구조적 특성을 평가하고, 과학 교육 과정 전체를 종합적으로 평가할 수 있는 과학탐구 평가표(SIEI)를 개발하여 과학 교육 과정에 내재된 탐구학습 내용도 평가할 수 있도록 고안하였다. 이외에도 수학능력 시험의 평가를 위한 기준을 마련하기 위한 연구도 진행되어, 수학능력 시험 중 수리탐구 영역 중 탐구 사고력 평가를 위한 객관적 기준이 될 수 있도록 탐구영역의 평가목표 상세화를 시도하였다(우종옥 외, 1991). 또한 과학적 탐구의 본질에 대한 연구도 진행되어, 국내외의 연구를 분석하고, 물리 화학 생물 지구과학의 분야별 일반적 탐구의 특성에 대해 고찰하고(조희영 외, 1994), 물리 화학 생물 지구과학의 분야별 학습지도를 통해 탐구력을 길러줄 수 있는 학습지도 자료의 개발도 이뤄졌다(조희영 외, 1995).

이와 아울러 ‘어떻게 탐구 능력을 평가할 것인가’에 대한 연구도 다양하게 이루어졌다. 중학교 생물 분야에서 탐구 과제를 선정하고, 실험을 탐구 단계별로 탐구 기술을 평가할 수 있는 도구를 개발하였다(김현식, 1986). 이외에도 지필 평가를 이용한 고등학생의 탐구력 측정을 위한 평가도구가 개발되고(이종기, 1988), 지구과학 분야에서의 덜 발달된 과학 탐구능력 요소를 진단하고, 과학 탐구 능력 요소들의 신장 정도를 추적할 수 있는 도구가 개발되었다(이향로 외, 1994). 또한 지필 고사에 의한 표준화된 탐구능력 평가를 위한 평가문항의 개발에 관한 연구로 중학교 2학년용 표준화 검사지가 개발되었다(이연우, 우종옥, 1991). 이와 별도로 과학 탐구 능력을 기초탐구능력과 통합탐구능력으로 구분하여 탐구 능력을 측정하고자 하는 연구가 진행되어(권재술, 김범기, 1994), 초·중학생들의 과학 탐구능력 측정도구가 개발되었다.

하지만 개발된 탐구력 측정의 도구들은 지필 평가를 이용하여 학생들의 탐구력을 측정할 수는 있지만, 일회성의 평가이기 때문에 학교 수업시간에 탐구 과정을 평가를 위한 실험 활동을 지속적으로 평가하는 평가도구로 사용하기는 어렵다고 보여진다. 이러한 점을 보완한 연구들이 진행되었는데, 고등학교 물리 분야에서 실험을 실시하고 이를 통해 학생들의 탐구실험 능력을 평가할 수 있는 평가도구에 관한 연구가 진행되어(김미경 외, 1995) 개발된 실험 지도서를 통해 실험을 진행하고, 탐구능력 평가를 위해 지필 고사으로써 탐구과정 요소에 대한 평가를 할 수 있도록 고안했다. 또한 화학분야에서도 이러한 연구가 진행되어(남정희,

1996), 탐구능력 평가를 위한, 실험 수업을 통해 학생들의 탐구 능력을 평가할 수 있는 도구의 개발이 이루어졌다. 중학생을 대상으로 한 연구에서는 학생들을 대상으로 하여 실험을 수행하도록 하고 이를 통해서 학생들의 실험 수행과정을 평가하려는 연구가 진행되고(강순희 외, 2000), 중학교 과학실험에 적합한 실험모델과 실험을 수행하는 동안 보여지는 여러 가지 탐구과제 수행과정을 평가하는 평가도구를 개발하였다. 특히 이 연구에서는 실험 수행과정에서 탐구과정 요소에 대한 평가뿐만 아니라, 실험의 기능 및 태도까지도 평가할 수 있는 객관적 평가 기준을 제시하고 있다.

이러한 연구와는 별도로 교과서 및 교육현황에 대한 연구도 이루어졌다. 고등학교 생물 교과서의 탐구활동을 분석하고, 이를 개선하기 위한 연구가 진행되고(한기영, 1987), 생물 교육의 실태 및 문제점에 대한 연구도 진행되었는데(정건상, 1991), 학생들이 생물 학습에 흥미를 갖도록 수업 보조자료와 효과적인 탐구학습을 수행할 수 있는 별도의 프로그램이 체계적으로 개발·보급되어야 한다는 점을 지적하였다. 이러한 연구는 제주도내에서도 이루어져(정충덕 외, 1995), 도내의 12개 인문계 고등학교에서는 한정된 실험만을 실험수업으로 진행하고 있으며, 실험 계획 능력이나 결과 처리 능력보다는 1교시 수업 시간 내에 마칠 수 있는 기구의 조작·사용능력 관찰 및 측정능력 위주의 실험을 선호하고 있음을 확인하였고, 탐구학습을 위한 효율적 실험 수업을 진행하기 위해서 실험실 확충 및 실험의 2시간 연속수업 등을 제안하고 있다. 한편, 효과적인 실험의 수행에 대한 연구가 진행되어(이기종 외, 1989), 고등학교 화학 실험에서 가격이 저렴한 실험 및 컴퓨터 보조수업모형을 개발 이용하려는 연구가 진행되었다.

앞에서 살펴본 일련의 연구는 최근 강조되어지고있는 다양한 평가방법의 적용이나, 수행평가, 과정평가를 위해 필수적인 연구라고 판단된다. 그러나 선행연구에서 제시한 바와 같이(조희영 외, 1994, 1995) 물리, 화학, 지구과학 및 생물의 각 분야는 과목의 특성이 다르므로, 한 가지 기준으로 전체를 평가하기는 매우 곤란하다고 보여진다. 그러므로 각 분야에 맞는 적절한 평가 기준에 대한 연구가 필요하고, 특히 객관성을 확보할 수 있는 기준을 제시함으로써 오늘날 학교 현장에서 교사의 평가에 대한 신뢰도를 증가시키고, 교사와 학생간의 평가의 기준에 대한 불필요한 오해를 불식시키는데도 큰 기여를 할 것으로 판단된다.

III. 연구 방법

1. 연구의 대상

본 연구는 제주도내 고등학교 학생을 대상으로 실시하였다. 공통과학 과목 중 생물 분야에서 탐구 능력을 평가할 수 있는 주제를 선정하고, 실험 과정을 통해 측정하고자하는 탐구과정 기능에 대한 일반적인 평가 기준을 개발하였다. 공통과학 수업 시간을 이용하여, 학생들에게 개발된 실험을 실시하고, 실험 후에는 보고서를 제출하도록 하여 이를 평가하였으며, 실험이 진행되는 동안 교사는 학생들의 실험과정을 관찰하여 실험기능과 태도를 평가하였다. 객관적이며 신뢰도 높은 평가를 위하여 탐구 과정 기능에 대한 평가기준, 실험기능을 평가할 수 있는 평가기준, 실험태도에 대한 평가기준 등 세 가지 부분을 측정할 수 있는 평가기준 개발에 대한 연구를 진행하였다. 개발 과정에서 과학교육을 담당하는 교수 3명과 현장교사 8명이 협의하여 내용을 수정·개선하였다.

2. 실험 내용의 선정

실험의 내용은 고등학교에서 다루어지는 공통과학 교과서와 연관성이 있고, 탐구력 신장 및 측정에 용이한 실험 중에서 생물분야와 관련된 실험을 선정하였다. 실험 선정을 위해서 우리나라에서 발행되어 고등학교에서 선정된 공통과학 교과서의 생물분야 탐구 실험 주제를 확인하고, 이와 연관성이 큰 실험을 위주로 선정하였다. 실험 주제는 총 8개를 선정, 학기당 4회의 실험을 실시할 수 있도록 하여, 고등학교에서 실행되고 있는 수행평가를 위한 평가 자료를 제공하고, 학생들의 탐구력 신장을 위한 실험으로 이용될 수 있도록 하였다. 현재 고등학교에서 학생들이 교과서로 사용하고 있는 공통과학 11종 교과서는 <표 1>에서 제시하였다. 또한, 본 연구에서 선정한 8가지 실험 주제와 관련된 11종 교과서의 탐구·실험 주제를 <표 2>에서 제시하였다. 고등학교에서 선정한 공통과학 교과서 11종 교과서에 수록되어있는 생물 관련 실험 및 탐구 활동에 대한 내용은 부록에서 제시하였다.

<표 1> 고등학교에서 채택된 공통과학 교과서 현황

구분	출판사	저자	구분	출판사	저자
a	교학사	강만식 외 6명	g	한샘출판(주)	권재술 외 7명
b	(주)두산	장남기 외 12명	h	동아서적(주)	한수복 외 9명
c	(주)두산	강영희 외 13명	i	금성출판사	김시중 외 15명
d	(주)지학사	정해문 외 11명	j	(주)천재교육	우규환 외 11명
e	대한교과서	최돈형 외 8명	k	교학사	송인명 외 11명
f	학습개발사	김수용 외 10명			

<표 2> 선정된 실험 주제에 대한 11종 교과서의 탐구 실험 주제

실험 구분	탐구 실험 I	탐구 실험 II	탐구 실험 III	탐구 실험 IV
실험 주제	빛의 세기와 광합성	엽록체의 색소 분리	소화효소의 작용	자극에 대한 반응속도
단원	생물에너지	생물에너지	영양과 건강	자극과 반응
a	광합성과 호흡은 어떤 관계가 있을까? (p. 169)		침 아밀라제의 소화 작용에 영향을 미치는 요인은 무엇일까?(p.197)	자극에 대한 반사 운동은 얼마나 빨리 일어날까?(p. 207)
b		빛에너지와 엽록소 (p. 188)	소화 효소의 작용에 대한 실험(p. 207)	자극에 대한 반응 (p. 217)
c		광합성 색소의 분리실험 (p. 177)		자극과 반응의 경로 (p. 211)
d	빛의 세기와 광합성 (p76)	빛의 파장과 광합성 (p. 78)	영양소의 소화 (p. 230)	신경의 흥분전도 속도 측정(p. 248)
e	식물은 어떻게 빛에너지를 이용하고 있을까?(p. 212)		침에 의해 녹말은 어떻게 소화되는가? (p. 258)	
f	빛과 광합성 (p. 179)	광합성과 환경요인 (p.173)	침의 소화작용 (p. 204)	사람의 반응시간 조사 (p.223)
g	빛의 세기와 광합성 (p. 282)	빛의 파장에 따른 광합성의효과(p.284)	효소의 반응 속도와 온도와의 관계(p.133)	자극에 대한 반응의 빠르기(p.144)
h	빛의 세기에 따라 광합성량은 어떻게 변할까?(p.357)		침은 어떤 작용을 할까? (p. 109)	신경은 흥분을 얼마나 빨리 전달할까?(p. 128)
i	빛의 세기는 광합성률을 어떻게 변화시킬까?(p. 185)	식물은 녹색 빛을 싫어한다? (p. 187)	소화가 잘되려면? (p. 213)	시각 자극에 대한 반응 시간(p. 230)
j	광합성과 빛의 세기 (p. 172)	빛의 파장에 따른 광합성 효과(p. 174)	침의 소화작용 (p. 201)	
k	광합성과 빛의 세기 (p. 171)		침속에 있는 소화 효소의 작용(p.194)	반응의 소요시간 조사 (p. 215)

<표 3> (계속)

구분	실험	탐구 실험 V	탐구 실험 VI	탐구 실험 VII	탐구 실험 VIII
실험주제		식물의 군집 조사	촉각의 분포조사	혀의 미각 분포	사람의 유전형질 조사
단원			자극과 반응	자극과 반응	유전
a			피부의 촉각 수용체는 얼마나 조밀하게 분포되어 있을까? (p. 208)		눈으로 볼 수 있는 형질은 어떻게 유전될까?(p. 226) 눈으로 볼 수 없는 형질은 어떻게 유전될까?(p. 227)
b			자극 수용에 관한 실험(p. 216)		사람의 유전 형질 조사(p. 242)
c				혀의 미각 분포 (p. 208)	사람의 유전형질 (p. 239)
d					유전형질의 조사 (p. 288)
e					우리는 어떤 형질을 갖고 있을까?(p.295)
f			냉점과 온점의 분포조사(p. 221)		사람의 유전 형질 조사(p. 248)
g					사람의 유전형질 (p.167)갯불 형태의 유전 형질 (p.169)
h				맛자극은 하나의 감각기관에서만 받아 들일까? (p. 122)	사람의 유전 형질은 어떤 유전 형질에 의해 결정될까?(p.150) 왼손잡이의 유전은 어떻게 될까?(p. 153)
i					보조개는 어떻게 유전되는가? (p.258)색맹은 어떻게 유전되는가? (p. 259)
j			피부의 압점 분포 (p. 213)		사람의 유전형질(p. 234), 보조개의 유전(p. 236), 색맹 유전(p. 237)
k			피부의 압점 분포 조사(p. 221)		사람의 신체적 특성에 관한 유전형질(p. 212)

3. 평가 기준의 선정

탐구 과정에 대한 평가의 기준은 연구 대상에게 실험을 실시하는 과정 및 실험을 실시하고 난 후 실험에 과정에 대한 평가를 위해 ①실험을 실시하는 동안 및 실험을 실시하고 난 후 보여지는 탐구 과정 기능에 대한 평가기준, ②실험 기능에 대한 평가기준, ③실험 태도에 대한 평가기준 등 세 가지 기준을 설정하였다.

1) 탐구 과정 기능에 대한 평가기준

탐구 과정 기능은 탐구적 실험을 하는데 있어서 필요한 핵심적 능력을 뜻하는 것으로 이를 측정하는 데에 있어서 평가자의 주관적인 판단을 배제하고 객관성을 유지하기 위해서는 평가기준이 상세히 제시되어야 한다(우종옥 외, 1997). 본 연구에서는 먼저 평가해야 할 탐구과정 기능을 선정하고 이에 대한 일반적인 평가 기준을 개발하였다. 슈바브의 탐구과정, 허명의 과학 탐구 평가표(허명, 1984), SAPA II(1990), 이종기(1988)가 고등 학생용으로 개발한 과학 탐구 기능 검사, 남정희(1996)의 탐구과정 기능의 평가도구를 수정·보완하여 다음과 같이 6개의 탐구과정 기능 요소를 선정하였다.

(1) 가설 설정하기(Formulating hypotheses)

어떤 현상을 관찰하고 이를 통해 문제를 인식하고, 이러한 문제를 해결하기 위해 임시적인 답을 찾는데, 이러한 임시적 답이 가설이다. 종속변인과 독립변인이 주어진 실험이 주어지면, 이로부터 검증 가능한 가설을 확인 제안 할 수 있어야 하며, 이때 가설은 이전의 지식을 적용하여 어떤 과학의 원리나 개념과 일치하는 설명이 제시되어야 한다.

(2) 실험설계 및 변인통제하기(Designing experiment & Controlling variables)

어떤 문제가 주어지거나, 가설이 주어지면 그 문제나 가설을 검증할 수 있도록 실험이 설계되어야 한다. 이를 위하여 반드시 통제되어야 하는 변인을 확인하고, 통제할 수 있어야 한다.

(3) 관찰하기와 측정하기(Observing & Measuring)

자신이 세운 계획에 따라 실험이나 관찰을 진행하면서 자료를 얻는 과정으로 정확한 방법으로 관찰·측정되어야 하고, 관찰된 내용은 정확하게 기록되어야 한다.

(4) 자료 해석하기(Interpreting data)

실험을 수행하면서 얻어진 데이터 즉 관찰이나 측정된 자료를 자료의 성격이나 대상에 따라 적절한 방법으로 제시할 수 있어야 하고, 이러한 자료를 통해서 규칙성이나 경향성 유사성 등을 찾아내는 활동을 말한다. 즉 주어진 자료를 도표로 만들거나 그래프 등으로 변환 할 수 있어야 한다.

(5) 결론(Drawing conclusion)

실험을 통해 어떤 데이터가 주어지면, 이 자료만을 가지고 어떤 결론을 이끌어 낼 수 있어야 한다. 이때 자신의 실험 결과와 적절한 결론을 내릴 수 있어야 하고, 변인간의 관계도 인식하고 있어야 한다. 또한 자신이 세운 가설을 검증할 수 있는 형태로 제시하여야 한다.

(6) 평가(Evaluation)

탐구 활동의 과정이나 혹은 실험 결과에 대해 가설과 비교하여 그 적합성, 정확성, 타당성, 신뢰성, 적용 가능성, 일반화 가능성 등 가치 판단 기준을 마련하고, 그 기준에 맞추어 평가할 수 있어야 한다. 만일 자신이 세운 가설과 일치하지 않는 결론이 나왔을 때, 실험의 전 과정에 대해 탐구과정을 비판적으로 평가하고 그 이유를 밝힐 수 있어야 한다.

이들 탐구 과정 기능에 대한 평가의 척도는 <표 3>과 같다.

<표 3> 탐구과정 기능에 대한 일반적인 평가의 기준

탐구 과정 기능 요소	평 가 요 소	배 점
가 설 설 정 하 기	○ 전혀 가설을 세우지 못한다.	0
	○ 경험이나 자신의 단순한 추측에 의존 문제를 제기함으로써 문제를 정확하게 파악하지 못한다.	2
	○ 관찰에 기초하여 문제를 제기하나, 가설의 형태가 아닌 문제 제기이거나 추측을 한다.	4
	○ 관찰에 기초하여 가설을 세우지만 실험적으로 입증하지 못하는 가설을 세운다.	6
	○ 관찰에 기초하여, 과학적 지식을 토대로 가설을 세우나 정확하지 않거나 논리적 오류가 있다.	8
	○ 정확한 관찰과 과학적 지식을 근거로 논리적으로 완벽한 가설을 세운다.	10

<표 3> 계속

실 험 설 계 및 변 인 통 제 하 기	○ 전혀 실험을 설계하지 못한다.	0
	○ 실험 설계는 시도하나 변인통제가 전혀 이뤄지지 않는다.	2
	○ 일부 변인을 확인하기는 하나, 전체적 실험설계가 이뤄지지 않는다. 적절한 독립변인이 설정되지 않았다.	4
	○ 독립변인과 통제변인을 확인하나, 구체적으로 실험을 설계하지 못한다.	6
	○ 변인을 통제하고 실험을 설계하나, 실험 절차의 일부가 잘못 되었다.	8
	○ 변인을 적절하게 통제하고, 구체적인 실험절차를 고안한다.	10
관 찰 하 기 와 측 정 하 기	○ 전혀 자료 수집이 이뤄지지 않는다.	0
	○ 일부의 정성적 관찰만 이뤄지고, 가설 검증에 필요한 자료를 일부만 기술한다.	2
	○ 일부의 정량적 자료를 모으나, 전체적인 자료를 모으지 못한다.	4
	○ 가설 검증에 필요한 모든 자료를 모으나, 적절한 측정 기구를 사용하지 못하고, 측정방법과 측정단위가 바르지 못하다.	6
	○ 가설 검증에 필요한 모든 자료를 모으나, 하지만 적절한 측정 기구, 올바른 측정 방법, 정확한 측정단위의 세 가지 중 하나가 바르지 못하다.	8
	○ 가설 검증에 필요한 모든 자료를 모으며, 정확한 측정 기구와 정확한 방법을 통해 정확한 측정 단위까지 측정한다.	10
자 료 해 석 하 기	○ 전혀 자료 해석이 이뤄지지 않는다.	0
	○ 실험결과에 대한 단순한 결과만을 나열한다. 자료의 변환이 없다.	2
	○ 얻어진 자료를 그래프나 표 등 다른 형태로 전환시킬 수 있지만 변환된 자료가 바르지 못하다.	4
	○ 얻어진 자료를 다른 형태로 전환할 수 있으나 변환된 자료로부터의 해석이 정확하지 못하다.	6
	○ 변환된 자료로부터 경향성 규칙성 등을 확인하나, 변인과의 관계를 적절하게 설명하지 못한다.	8
	○ 변환된 자료가 완벽하고, 이로부터 변인과의 관계를 적절하게 규명하고 규칙성과 경향성을 찾아낸다.	10
결 론	○ 전혀 결론을 도출해 내지 못한다.	0
	○ 단순히 실험에서 나타난 현상이나 결과만을 제시한다.	2
	○ 실험의 결과를 변인과 연관시켜 해석 하나 논리적으로 부적절하다.	4
	○ 실험 결과를 변인과 연관시켜 해석하나, 특정 실험 결과에 국한된 결론을 도출한다.	6
	○ 변인과 연관시켜 결론을 도출하며, 변인관계도 적절하게 해석하나, 논리적으로 부적절한 용어가 사용된다.	8
	○ 논리적으로 적절한 용어를 사용하며, 변인들 사이의 관계를 과학적 용어를 사용하여 적절하게 제시한다.	10

<표 3> 계속

평 가	○ 자신의 세운 가설과 결론에 대한 비교가 이뤄지지 않는다.	0
	○ 자신의 결론을 가설과 비교하나, 옳고 그름을 판단하지 못한다.	2
	○ 결론을 가설과 비교하여, 맞지 않은 경우 그 이유를 찾기는 하나, 적절하지 못하다.	4
	○ 결론을 가설과 비교할 뿐 아니라, 가설설정, 실험설계, 자료 수집 및 해석, 등 실험과정별로 체계적으로 점검한다.	6
	○ 실험의 전 과정을 통해 구체적인 문제점을 제시한다. ○ 문제점과 관련된 추가실험 및 개선방안을 제시한다.	8 10

2) 실험 기능에 대한 평가 기준

실험 기능은 실험을 수행하는 과정에서 실험 기구나 기기를 다루고 실험 장치를 꾸미는 것과 같은 구체적 조작 활동이다. 여기에서 평가하고자 하는 기능은 실험 기구를 다루는 방법과 절차의 숙지, 다루는 솜씨 등이다. 이러한 신체적인 영역은 실험보고서를 통해서도 평가할 수 없고 직접 관찰하는 방법을 사용하여 평가하여야 한다(김창식 외, 1993). 관찰에 의한 평가 방법은 시간이 다소 많이 걸리는 단점이 있으나 실험 기능을 평가하는 가장 적절한 방법이다(김양현, 1999). 실험 기능을 평가하기 위해서 본 연구에서는 강순희 등(2000)이 개발한 실험기능 평가기준 등을 참고하여 점검표 형태로 만들었으며, 개발한 실험 기능에 대한 일반적 평가의 기준은 <표 4>와 같다.

<표 4> 실험 기능에 대한 일반적인 평가의 기준

평 가 요 소	양 호	미 흡
① 실험 기구나 기구를 바르게 사용한다.	<input type="checkbox"/>	<input type="checkbox"/>
② 실험 장치를 바르게 꾸민다.	<input type="checkbox"/>	<input type="checkbox"/>
③ 실험 장치를 바르게 사용한다.	<input type="checkbox"/>	<input type="checkbox"/>
④ 실험 장치를 사용할 때 안전하게 다룬다.	<input type="checkbox"/>	<input type="checkbox"/>

3) 실험 태도에 대한 평가 기준

과학교육에서 태도나 흥미, 관심 등의 정의적 영역에 대한 학습은 중요한 부분으로 인식되어 왔으나, 그 중요성에 비해 실제 지도와 평가 시 소홀히 취급되어 왔다(정완호 외, 1994). 이러한 현상은 정의적 영역을 측정하기 위한 타당하고 신뢰로운 도구가 부족하기 때문이다(우종욱, 이경훈, 1995). 일반적으로 과학 교육에서 태도라 함은 과학적 태도와 과학에 대한 태도로 나눌 수 있다. 과학적 태도

는 탐구하는 자세, 화학 정신과 관련된 것으로 문제를 해결하는 등 과학적 탐구 과정 차원에서의 태도이며, 과학에 대한 태도는 과학, 과학자 등을 포함하는 가치 차원에서의 태도를 일컫는 말이다(강순희 외, 2000). 이러한 태도의 평가는 실험 과정과 실험 전· 후에 이뤄지는 토론 참여도, 실험 수행 시 실험 참여에 대한 적극성, 주의사항에 대한 준수정도, 실험 마무리 및 정리정돈의 네 가지 평가 요소를 정하여 이를 각각 3단계로 나누어 평가하였다. 본 연구에서 적용한 일반적인 실험태도 평가기준은 다음의 <표 5>와 같다.

<표 5> 실험 태도에 대한 일반적인 평가의 기준

평가요소	양 호	보 통	미 흡
토론 참여	적극적 참여	소극적 참여	거의 참여 안함
실험 참여	능동적 참여	수동적 참여	거의 참여 안함
주의 사항	철저하게 지킴	대체적으로 지킴	거의 지키지 않음
정리 정돈	정리정돈 철저	정리 상태 양호함	정리정돈 안함

4. 탐구 실험 및 세부 채점 기준표 개발

학생들이 수행해야 할 실험은 고등학교 1학년 과정에서 다루어지는 공통과학 교과서에 실린 탐구활동 및 실험과 연관성이 있는 주제를 선정하였고, 단순하게 과학적 사실을 확인하는 확인 실험이 아닌 실제로 가설을 설정하고, 이를 해결하는 과정에서 탐구력을 신장할 수 있는 탐구실험을 개발하였다. 실험은 총 8개의 주제로 단원별로는 생물에너지 2, 영양과 건강 1, 자극과 반응 2, 유전 1, 및 기타 1 주제로 구성되어 있다. 각각의 주제에 대한 탐구실험을 우선 개발한 후, 실험 보고서를 통해 탐구능력을 평가할 수 있도록 탐구과정 기능을 6단계로 구분하고, 실험의 기능과 태도를 평가할 수 있는 일반적인 평가 기준을 개발한 후, 내용 타당도를 검증하였다. 이 과정에서 일차적으로 과학교육을 담당하는 교수와 일선 고등학교 생물 담당 교사들과 내용을 검토하고 수정·개선하여 완성하였고, 완성된 탐구 과정 기능, 실험기능 및 태도에 대한 일반적인 평가기준에 대한 내용 타당도를 설문을 통해 검증하였다. 검증 후 각각의 실험 주제에 대한 세부적인 채점 기준표의 개발이 이루어 졌는데, 개발된 탐구과정 기능에 따른 교사용 실험안내서 및 학생용 실험보고서를 개발하고 이를 적용하여 학생들로 하여금 실험을 실시하도록 하였다. 실험은 고등학교의 실험실에서 실시하였고, 실험시간은 50분

으로 조별로 실험을 실시하고, 보고서 작성을 위한 50분을 제공하여 2교시의 수업시간 동안에 실험과 보고서의 작성을 완성하도록 하고, 이를 통해 학생들의 탐구 과정 기능과 실험기능 및 태도를 평가하였다. 평가 과정에서 나타난 학생들의 답을 채점하는 과정에서 일반적인 평가기준을 적용하기 쉽게 세부 채점 기준표를 개발하고, 일선 교사들과 협의하여 수정·보완하였다. 완성된 세부 채점 기준표는 부록에서 제시하고 있으며, 각각의 실험에서 학생들의 반응을 일반적인 평가기준을 적용하여 평가 예를 제시함으로써, 채점이 용이하도록 하였다. 세부 채점 기준표는 일반적 평가기준을 근거로 실제 실험의 과정에서 도출되는 여러 가지 탐구 과정 기능 중에서 학생들이 제시하는 답을 실례로 그 기준으로 삼았기 때문에 평가자가 좀더 쉽게 채점할 수 있을 뿐만 아니라, 학생들이 제시하는 답에 따라서 혹은 교사에 따라서 적절한 조정이 가능하다.

5. 평가 기준에 대한 타당도 및 신뢰도 검사

1) 탐구과정에 대한 일반적인 평가기준, 실험태도 및 기능에 대한 일반적인 평가기준의 타당도 검사

개발된 탐구 과정 기능의 일반적 평가 기준에 대한 내용 타당도 검사를 실시하였다. 내용 타당도는 검사내용 전문가에 의하여 검사가 측정하고자 하는 속성을 제대로 측정하였는가를 주관적으로 판단하는 것이다(성태제, 2000). 개발된 평가 도구에 대한 내용 타당도에 대한 검사에 앞서 일차적으로 개발된 탐구과정 기능의 일반적 평가기준을 과학교육을 담당하는 교수와 일선 고등학교 교사들과 협의하여 내용을 선정하고, 선정된 내용에 대한 수정과정을 거쳐 일반적인 평가기준을 완성하였으며, 완성 후에는 과학교육을 담당하는 교수 3명과 고등학교에서 생물을 담당하는 교사 8명을 대상으로 내용 타당도 검증을 위한 설문을 실시하였다. 설문에 응한 사람들의 교육경력은 20년 이상이 4명, 20년 미만인 5명, 10년 미만이 1명, 5년 미만이 1명이었다. 설문지의 내용은 부록에서 제시하고 있다. 이 설문을 토대로 탐구과정 기능에 대한 일반적인 평가기준의 내용 타당도를 검증하였다. 내용 타당도는 상관계수와 같은 수량적 지수로 표현되지 않으므로(이종성 외, 1992), 타당도 검증의 결과를 수량적으로 표현하기 위해서 평가기준이 “매우 적절하다”고 판단되면 5점, “적절하다”고 판단되면 4점, “미흡하다”고 판단되면 3점, “부적절하다”고 판단되면 2점, “매우 부적절하다”고 판단되면 1점을 부여

한 후 이들 점수를 평균하여 5점 만점에 대한 퍼센트로 나타내었다. 이를 위하여 5점은 100점으로, 4점은 80점, 3점은 60점, 2점은 40점, 1점은 20점으로 환산하여 이들 값의 평균을 구하였다. 이와 아울러 탐구 실험 과정에서 나타나는 학생들의 실험 태도와 기능을 평가하는 기준이 되는 실험태도에 대한 일반적인 평가기준과 실험기능에 대한 일반적인 평가기준에 대한 타당도 검사도 설문지를 이용하여 탐구과정 기능에 대한 일반적인 평가기준과 동일한 방법으로 타당도를 검증하였다.

2) 세부 채점 기준표에 대한 신뢰도 검사

본 연구에서 개발된 평가 기준이 평가자에 따라 얼마나 유사하게 측정되는가는 중요하다. 오늘날 많은 학교에서 실시하는 수행평가의 어려움 중에서 객관성을 유지하는 문제는 매우 중요한데, 특히 대학 입시를 중요시하는 우리 나라의 현실에 비추어 볼 때, 대학 입학 기준으로서 고등학교의 내신 성적을 중요시하는 오늘날 수행평가의 객관성과 신뢰성 문제는 공교육의 신뢰성 문제와 관련하여 매우 중요성이 높다. 신뢰도의 측정 방법에는 재검사 신뢰도나 동형검사 신뢰도를 측정하는 방법과 내적 일관성 신뢰도를 측정하는 방법이 있으나, 본 연구에서는 탐구과정 기능에 대한 일반 평가기준을 각 탐구실험의 평가에 적용할 수 있도록 개발한 세부 채점 기준표에 대한 채점자간 신뢰도를 측정하기 위해서 Cronbach(1951)가 고안한 공식을 이용하여 cronbach α 를 측정하였으며, 채점자에 따라서 얼마나 동일하게 채점하였는가를 확인하기 위해서 Karl Pearson(1896)의 적률 상관 계수를 구하였다. 성태제(2000)는 상관계수에 따른 상관관계의 언어적 표현을 <표 6>과 같이 제시하고 있다.

<표 6> 상관계수에 따른 상관관계의 언어적 표현

상관계수의 범위	상관관계의 언어적 표현
.00 ~ .20	상관이 거의 없다.
.20 ~ .40	상관이 낮다.
.40 ~ .60	상관이 있다.
.60 ~ .80	상관이 높다.
.80 ~ 1.00	상관이 매우 높다.

신뢰도 측정을 위해서 선정된 실험주제와 교사용 실험안내서를 공통과학 실험 수업에 투입하고, 개발된 학생용 실험보고서를 이용하여 학생들 스스로 탐구실

험을 진행하도록 하고, 실험보고서를 작성하도록 하여 그 결과를 세부 채점 기준표를 이용하여 채점하도록 하였다. 실험은 서귀포 소재 S고등학교 1, 2학년 학생을 대상으로, 학교 실험실에서 교과 수업시간을 이용하여 실시하였다. 탐구 실험을 진행하면서 자유롭게 토의와 실험을 실시하도록 하였고, 실험이 끝난 후 정리 시간을 부여하여 조별로 실험과정에 대해 토의를 거친 후 보고서를 작성하도록 하였다. 실험 시 교사는 학생들의 실험과정에 개입하지 않고 학생들 스스로 실험을 진행하도록 기구 및 재료의 준비만을 도와주었다. 실험 후 학생들의 보고서에 대한 채점을 실시하였다. 실험에 대한 채점은 세부 채점 기준표 및 개발된 평가 도구를 이해하고, 실제로 고등학교에서 공통과학을 담당하는 생물교사 3명이 개발된 세부 채점 기준표를 이용하여 채점하도록 하였다. 적률 상관 계수나 cronbach α 의 측정은 이들의 채점결과를 이용 분석하였다. 각각의 교사는 채점 과정에서 기준이 모호하거나, 불합리하다고 판단되는 평가 기준에 대해서는 협의의 거쳐 채점기준을 수정하면서 각자 다른 채점자와는 무관한 상태에서 채점을 실시하였다. 채점자들의 교육경력은 모두 생물 교과를 5년 이상 담당한 일선 고등학교의 생물 교사들이다.

6. 개발된 평가 도구의 검사 및 수정

본 연구의 목적 중 가장 중요한 것은 학생들의 탐구 과정 기능을 측정함에 있어서 객관적이고 신뢰도 높은 평가 기준을 제공함에 있으므로, 개발된 평가기준이 서로 다른 채점자에 의해 채점되어질 때 얼마나 객관적으로 평가할 수 있는가 하는 것이 가장 큰 관건이다. 본 연구에서 개발되어진 세부 채점 기준표는 학생들이 작성한 보고서를 주관적으로 평가하는 주관식 평가의 방법이므로 채점자들 간의 신뢰도를 분석하였고, 이를 위해 채점되어진 결과를 PC를 이용 입력하고, SPSS 통계 프로그램과 Microsoft Excel 프로그램을 이용하여 처리하였다(노형진, 1999). 그리고 동일한 보고서에 대한 각기 다른 채점자간의 채점 결과가 일치되는 정도 즉, 채점자간 상관관계를 분석하기 위해서 Pearson의 적률 상관계수를 구하였는데, 8개의 주제에 대한 실험보고서에 대한 채점 결과를 이용하여 채점자간 상관계수를 구하였다.

개발된 평가기준에 대한 타당도와 신뢰도를 측정하는 과정에서 나타난 오류나 채점기준에 빠져있는 내용은 과학교육을 담당하는 교수의 자문을 구하고, 일선

고등학교 교사들과 협의하여 수정하였으며, 신뢰도와 채점자간 상관계수가 지나치게 낮은 문항의 경우, 이들과 논의하여 채점 기준을 명확하게 개선 및 수정하였다.

IV. 결과 및 고찰

1. 탐구과정 기능에 대한 일반적인 평가기준에 대한 내용타당도

탐구과정 기능을 평가하기 위해 개발되어진 탐구과정 기능에 대한 일반적인 평가 기준(표 3)이 적절한 평가의 기준이 되는지에 대한 내용 타당도를 검증한 결과는 <표 7>과 같이 나타났다.

<표 7> 탐구과정 기능에 대한 일반적인 평가기준에 대한 내용 타당도

탐 구 과 정 기 능	내 용 타 당 도 (%)
가설설정하기	89
실험설계 및 변인통제하기	89
관찰하기와 측정하기	91
자료 해석하기	87
결 론	95
평 가	91
평 균	90.3

내용 타당도를 항목별로 보면 탐구 과정 기능의 항목 중에서 결론이 95%로 가장 높았으며, 자료 해석하기 항목에서 87%로 가장 낮았다. 그리고 이들 평균값은 90.0%이상이었다. 이러한 결과는, 이종기(1988)의 연구에서 83%, 이연우 등(1991)의 85%, 이항로 등(1994)의 89% 및 강순희 등(2000)의 84%~98%와 비교해 볼 때 비슷하거나 비교적 높은 값을 보였으며, 이것은 탐구 과정 기능의 각 항목에 대한 내용 타당도가 높음을 알 수 있었고, 평가기준이 적절하게 평가하고 있음을 알 수 있었다.

2. 실험기능에 대한 일반적인 평가기준에 대한 내용 타당도

실험 기능에 대한 일반적인 평가기준이 학생들이 실험을 수행하는 과정에서

실험 기능을 적절하게 평가할 수 있도록 개발되었는가에 대한 내용 타당도를 검증한 결과 <표 8>과 같이 나타났다.

<표 8> 실험 기능에 대한 일반적인 평가기준에 대한 내용 타당도

평가자	A	B	C	D	E	F	G	H	I	J	K	평균
평가점수	5	5	5	5	5	5	4	4	3	3	2	4.18
환산점수	100	100	100	100	100	100	80	80	60	60	40	83.6

실험 기능에 대한 일반적인 평가기준을 평가하기 위하여 개발된 일반적인 평가 기준에 대한 내용타당도의 값은 83.6%이었으며, 강순희 등(2000)의 보고에서는 88%로 나타났다. 일반적으로 타당도가 80% 이상이면 적절하다고 받아들여진다(이종기, 1988; 이항로 외, 1994; 남정희, 1996). 따라서 본 연구의 실험 기능에 대한 일반적인 평가기준의 내용 타당도는 적절하다고 사료된다.

3. 실험 태도에 대한 일반적인 평가기준에 대한 내용 타당도

실험 과정을 통해 학생들의 실험 태도를 평가하기 위해 개발된 실험 태도에 대한 일반적인 평가기준에 대한 내용 타당도를 검증한 결과 <표 9>와 같이 나타났다.

<표 9> 실험 태도에 대한 일반적인 평가기준에 대한 내용 타당도

평가자	A	B	C	D	E	F	G	H	I	J	K	평균
평가점수	5	5	4	5	5	5	5	4	3	5	4	4.54
환산점수	100	100	80	100	100	100	100	80	60	100	80	90.9

본 연구에서 개발된 실험 태도에 대한 일반적인 평가기준의 내용 타당도는 90.9%로 나타났다. 이는 강순희 등(2000)의 연구에서 보여진 100%보다는 낮은 값이지만, 실험 기능에 대한 일반적인 평가기준에서와 마찬가지로, 내용 타당도가 80%보다 높기 때문에, 실험 태도에 대한 일반적인 평가기준도 학생들의 실험 태도를 평가하는데 적절한 타당도를 갖고 있다고 볼 수 있다.

4. 세부 채점 기준표에 대한 채점자간 신뢰도 및 상관계수

세부 채점 기준에 대한 채점자간 신뢰도와 채점자간 상관계수는 표 10에 제시되어 있으며, 세부 수치는 부록에서 제공하고 있다. 개발된 세부 채점 기준표에 의한 채점에서, 모든 항목의 채점자간 신뢰도가 0.82이상으로 나타나고 있으며, 탐구 과정 기능별 신뢰도의 평균값은 <표 10>에서 제시되고 있고 전체 채점자간 신뢰도의 평균값은 0.92였다. 이는 이연우 등(1991)의 연구에서 0.96, 남정희(1996)의 0.95보다는 낮으나, 이종기(1988)의 0.86, 이항로(1994)의 0.84 및 김미경 등(1996)의 0.75보다는 높은 값이며, 일반적으로 신뢰도는 0.90이상이면 높다고 생각하고, 0.80이상이면 비교적 높은 편으로, 0.70이상이면 만족스럽다고 하고 0.70이하이면 낮은 신뢰도를 나타낸다(남정희, 1996). 따라서 개발된 세부 채점 기준표에 대한 채점자간 신뢰도의 전체 평균값이 0.92이므로 본 연구에서 개발한 세부 채점 기준표에 대한 채점자간 신뢰도는 높다고 판단된다.

<표 10> 세부 채점 기준표에 대한 채점자간 신뢰도 및 상관계수

탐구과정 기능	평가자간 신뢰도 평균값 (Cronbach α)	채점자간 상관계수 평균값 (적률 상관 계수)
가설 설정하기	0.92	0.83
실험설계 및 변인통제	0.93	0.86
관찰하기 와 측정하기	0.91	0.79
자료 해석하기	0.91	0.81
결 론	0.93	0.84
평 가	0.92	0.81
평 균	0.92	0.82

성태제(2000)는 채점자간 상관계수가 0.40이상이면 상관이 있는 것으로, 0.60이상이면 상관이 높은 것으로, 0.80이상이면 상관이 매우 높은 것이라고 하였다.

본 연구에서 실시한 세부채점 기준표를 이용한 학생들의 보고서 채점 결과에 대한 채점자간 상관계수의 평균값은 실험설계 및 변인 통제가 0.86으로 가장 높았고, 관찰하기와 측정하기가 0.79로 가장 낮게 나타났으며, 평균값은 0.82이상이었다. 이는 남정희(1996)의 보고서에서 .81이상으로 나타나고, 강순희 등(2000)의 연구에서 나타나는 0.61이상의 값과 비교해서 비슷하거나 높은 수준으로 판단된다. 그러므로 세부 채점 기준표를 이용한 평가에서 채점자간 상관은 매우 높은 것으로 사료된다.

V. 적 요

본 연구는 수행평가를 위해 고등학교 공통과학 교과 중 생물분야에서 탐구 실험 평가를 통해 탐구능력을 평가하기 위한 객관적이고, 신뢰도 높은 평가 도구를 개발하는 것을 목적으로 하고 있다.

이를 위하여 고등학교에서 채택된 11종 공통과학 교과서의 탐구·실험 내용과 관련된 8가지 주제에 대한 탐구 실험을 개발하고, 실험 과정을 평가하기 위해 탐구 과정 기능을 6단계로 나누어 이를 평가하기 위한 일반적 평가기준, 실험 태도 및 기능에 대한 일반적인 평가 기준을 개발하고 타당도를 측정하였다. 또한 실제 탐구 실험에서 나타나는 다양한 학생들의 응답에 대한 평가의 기준으로서 세부 채점 기준표를 개발하여, 세부적인 평가 예를 제시하였으며, 채점자간 신뢰도와 상관계수를 측정하였으며, 학교 현장에서 개발된 평가도구를 쉽게 이용할 수 있도록 교사용 실험안내서와 학생용 실험 보고서를 개발하였다.

본 연구에서 개발된 수행평가를 위한 탐구 실험 평가도구는 공통과학 교과의 생물분야에서 실험시간을 이용하여 학생 스스로 가설을 세우고 이를 해결하기 위해 실험을 설계하고, 자료를 측정하고 해석하여 결론을 도출하는 실험과정을 평가할 수 있도록 하였다. 그리고 최근에 고등학교에서 강조되어지는 수행평가의 가장 큰 문제점인 객관성을 높임으로써, 학생과 학부모에게 신뢰받을 수 있는 평가의 기준을 마련했다는 점에서 중요하다고 생각된다. 또한 실험평가의 활성화는 학생들의 탐구력 향상에도 긍정적인 효과가 있으리라 예상되어지며 지속적인 연구가 필요하다고 사료된다.

참 고 문 헌

- 강순희·남정희·박종윤·최병순. 2000. 「중학교 과학 탐구 실험 수행평가 도구」. 이화여대 출판부.
- 교육부. 1992. 「교육과정」.
- 교육부. 1997. 「과학과 교육과정」.
- 권재술·김범기. 1994. 초·중학생들의 과학탐구능력 측정도구의 개발. 한국과학교육학회지, 14(3): 251~264.
- 김미경·오희균·박종원. 1995. 물리 탐구 실험의 평가를 위한 도구의 개발과 분석. 한국과학교육학회지, 16(1): 51~60.
- 김양현. 1999. 탐구적 일반화학 실험에 대한 평가준거 개발. 이화여자대학교 교육대학원

석사학위논문.

- 김정화. 1993. 「최신과학교육」. 학문사.
- 김창식·이화국·권재숙·김영수·김찬중. 1993. 「과학학습평가」. 2판. 교육과학사.
- 김현식. 1986. 중학교 과학 생물실험의 탐구적 지도 및 평가방안의 연구. 전북대학교 교육대학원 석사학위논문.
- 김현재. 1998. 「과학과 열린교육」. 교육과학사.
- 남정희. 1996. 화학실험에서 탐구 과정기능의 평가도구 개발. 한국교원대학교 대학원 박사학위논문.
- 노형진. 1999. 「한글SPSSWIN에 의한 조사방법 및 통계분석」. 형설출판사.
- 박승재·조희영. 1995. 「과학론과 과학교육」. 교육과학사.
- 박호순·홍후조. 1999. 「교육과정과 교육평가」. 문음사. pp. 480~496.
- 백순근. 1997. 수행평가의 이론과 실제. 한국교육평가연구회 학술세미나 논문 발표집. pp. 3~42.
- 백순근. 1999. 「수행평가의 이론과 실제」. 원미사.
- 성태제. 2000. 「타당도와 신뢰도」. 양서원.
- 우종옥·이경훈. 1995. 과학 관련 태도의 타당한 측정을 위한 연구(I). 한국과학교육학회지, 15(3): 332~348.
- 우종옥·이항로·김승훈. 1997. 과학실험 평가도구 개발을 통한 탐구능력 평가의 타당화에 관한 연구. 한국과학교육학회지, 17(1): 65~73.
- 우종옥·이항로·이경훈. 1991. 대학수학능력 시험의 수리·탐구 영역 중 지구과학 교과에 관련된 탐구능력 측정을 위한 행동 요소의 추출과 평가 목표의 상세화 연구 I. 한국과학교육학회지, 11(1): 83~96.
- 이기중·이광호·이종철·김자홍. 1989. 과학교육 개선을 위한 실험모델 제작에 관한 연구(I). 과학교육연구, 9(2): 13~28.
- 이돈희·허경철·백순근·김신영·채선희. 1997. 국가 공통 절대 평가기준 활용 방안 연구. 한국교육개발원.
- 이범홍. 1997. 「과학교과학연구」. 한국교육개발원.
- 이연우·우종옥. 1991. 과학탐구능력 측정을 위한 표준화 검사지 개발. 한국과학교육학회지, 11(1): 59~72.
- 이종기. 1988. 고등학생의 과학 탐구능력 측정을 위한 평가도구 개발. 한국교원대학교 대학원 석사학위논문.
- 이종성·강봉규·한종철. 1992. 「교육심리 측정·평가」. 종각출판사.
- 이항로·우종옥. 1994. 고등학생의 지구과학 탐구능력 측정을 위한 평가도구개발. 한국과학교육학회지, 15(1): 92~103.
- 정건상. 1991. 고등학교 생물과 탐구 학습의 실태 조사와 문제점 분석. 한국교원대학교 대학원 박사학위논문.
- 정완호·허명·윤병호. 1994. 국민학생의 과학적 태도 측정을 위한 도구의 개발. 한국과학교육학회지, 14(3): 265~271.
- 정충덕·박백신. 1995. 고등학교 생물과 실험실습의 현안과 개선방안에 관한 연구. 한국

- 생물교육학회지, 23(2): 157~172.
- 조희영. 1992. 과학적 탐구의 본질에 대한 분석 및 탐구력 신장을 위한 학습지도 방법에 관한 연구. 한국과학교육학회지, 12(1): 61~73.
- 조희영·이문원·조영신·지찬수·강순희·박종윤·허명·김찬중. 1994. 고등학교의 과학적 탐구력 신장을 위한 과학 학습지도 방법과 자료의 개발에 관한 연구 I. 한국과학교육학회지, 15(1): 54~67.
- 조희영·이문원·조영신·지찬수·강순희·박종윤·허명·김찬중. 1995. 고등학교의 과학적 탐구력 신장을 위한 과학 학습지도 방법과 자료의 개발에 관한 연구 II. 한국과학교육학회지, 15(2): 133~148.
- 한국교육과정평가원. 1998. 「국가 교육과정에 근거한 평가 기준 및 도구 개발 연구」. 연구보고 RRE 98-3-7.
- 한국교육과정평가원. 1999. 「고등학교 공통과학 평가 방법 개선 방안」. 연수자료 CRE 99-1-6.
- 한기영. 1987. 고등학교 생물교재의 탐구활동에 대한 분석 및 개선 방안에 관한 연구. 한국교원대학교 대학원 석사학위논문.
- 허명. 1984. 과학 탐구 평가표 개발. 과학교육학회지, 4(1): 57~65.
- 허명. 1987. 탐구학습의 이론과 실제. 과학교육. 4월호(24). pp. 22~27.
- 허명. 1992. 「탐구적 실험 실습지도」. 교육과학사. pp. 237~243.