

초·중학생들의 과학탐구능력 측정도구의 개발*

권재술·김범기

(한국고원대학교)

(1994년 2월 25일 받음)

I. 머리말

과학교육에서 탐구능력의 중요성은 오래 전부터 과학교육자들에게 인식되어 왔다. 과학탐구 능력은 여러가지 의미를 포함하고 있지만, 과학적 지식에 대응하는 것으로서 문제를 해결해 나가는 절차와 이와 관련된 사고과정을 의미한다. 과학교육에서 과학탐구능력(science process skill)과 유사한 말로는 문제 해결(problem solving), 과학적 방법(scientific method), 과학적 사고력(scientific thinking), 비판적 사고력(critical thinking)등이 있다. 이들은 서로 유사한 면이 있으나, 약간씩 다른 의미로 사용되고 있다. 본 연구에서는 과학탐구활동에서 절차적 지식과 밀접한 관계가 있다고 생각되는 과학탐구능력(science process skill)을 중심으로 연구하였다.

과학탐구능력을 역사적으로 고찰하여 보면, 멀리서 1910년 Dewey의 사고의 행동(Act of thought)연구에서 언급한 다섯 가지 논리적 단계에서 찾을 수 있으며(Helgeson, 1992), 가깝게는 Bruner(1961)의 교육의 과정과 과학자들이 문제 해결을 위하여 과학자들이 사용하는 기능에 초점을 맞춘 국민학교 프로그램인 SAPA(Science -A Process Approach)로 기원을 잡을 수 있다(Padilla et al., 1985).

과학탐구능력이란 과학자들이 조사하고 연구에 필요한 능력으로, 학생들이 어떤 문제에 부딪혔을 때 과학적 탐구 방법에 의해 스스로 문제를 해결하는 능력이다. 이 과학탐구능력은 범내용(content free)적으로 일상생활에서 높은

전이 효과가 있는 것으로 믿어진다. 과학탐구능력의 향상은 과학적 사실과 과학적 개념을 잘 이해할 수 있고 실험능력을 향상시킬 수 있는 것으로 기대된다.

전통적인 학교의 과학교육에서는 결과로서의 지식율 강조하여, 학교 과학교육에서는 원리, 법칙, 이론 체계를 학습하는 것에 중점을 두어 왔다. 그러나 1960년대 이후 과학교육은, 결과로 얻어진 이론의 전수보다는 자연을 탐구하는 과정으로 보는 경향이 강하게 대두되어 PSSC, CHEM Study, BSCS, ESCP, IPS, PS II, SAPA, ESS, SCIS 등 수많은 혁신적 교육과정이 개발되었고, 영국의 APU (Assessment of Performance Unit)와 TAPS (Techniques for Assessing Process Skills in Practical Science), 미국의 NAEP(The National Assessment of Educational Progress)의 과학학력 평가를, Hur(1984)의 SIEI(Scientific Inquiry Evaluation Inventory) 등의 과학탐구 평가틀이 만들어졌다. 과학탐구능력 측정도구는 주로 지필평가 형태로 많이 개발되었지만, APU의 실험기능 평가나 SISS(The Second IEA Science Study)와 같이 표준화된 실험기능 탐구능력 평가도구의 개발은 최근에 이루어지고 있다.

우리 나라의 과학교육은 1960년대의 외국의 혁신적 과학교육과정의 영향을 받았고, 제3차 교육과정부터 과학적 탐구력의 육성이 과학교육과정의 중요한 목표가 되었다. 이에 따라 학교현장에서 탐구학습이 과학교육의 대명사처럼 되어 탐구능력 신장을 강조하여왔다. 그러나, 학생들의 과학탐구능력을 측정하는 도구의 개발이 부진하였다. 따라서,

* 이 논문은 1993년도 한국학술진흥재단의 공모과제 연구비에 의하여 연구되었음.

우리 나라 학생들의 과학탐구능력을 바르게 측정하는 도구가 거의 없는 실정이다. 과학교육의 질을 향상시키기 위해서는 신뢰성 있고 타당성 높은 탐구능력 측정도구의 개발이 절실히 요구된다.

우리 나라에서 과학탐구능력 검사지로 국민학교 대상으로는 김창식 등(1992), 나종국(1993)의 연구, 중학생 대상으로는 이연우(1989), 고등학생 대상으로는 이종기(1988), 이항로(1991)의 연구가 있다. 이러한 평가 도구들은 변별도와 난이도에 의한 문항분석, 내용타당도, KR-20 신뢰도 등을 통하여 타당도와 신뢰도를 제시하였지만, 표집이 작거나 포함하는 탐구능력의 범위가 포괄적이지 못한 면이 없지 않다. 따라서, 보다 큰 표집을 대상으로 연구할 필요가 있으며 탐구 능력 요소에 대한 면밀한 분석을 통한 검사 도구의 개발이 필요하다.

본 연구에서는 과학 탐구능력을 탐구 요소별로 분류하고 이 탐구요소에 따라 문항을 개발한다. 개발된 과학탐구능력 측정도구를 예비 및 현장 검사를 실시하여 난이도, 변별도, 신뢰도, 내용타당도, 공인타당도, 구인타당도 등을 이용하여 표준화된 초·중학생들의 과학탐구능력 평가 도구를 개발하는데 그 목적이 있다.

II. 과학탐구능력 측정도구에 관한 선행 연구

현재까지 많이 사용되어온 과학탐구능력 측정도구를 고찰하여 보기로 한다. BAPS(Basic Process Skills in Science)는 4~8학년 학생들에게 적합한 기초탐구능력 관련되었으며, TIPS(The Test of Integrated Science Process Skills), TIPS II, MIPT(Middle Grades Integrated Science Process Skills Test), POPS(The Performance of Process Skills)등은 중학생들의 과학탐구능력을 측정하기 위한 도구로 개발되었다.

과학탐구능력 검사지에 따라 측정하는 탐구요소를 고찰하면 다음과 같다. SAPA는 기초탐구능력으로 관찰, 시/공간 관계의 사용, 수의 사용, 측정, 분류, 의사전달, 예상, 추리의 8가지 탐구요소가 있고, 통합탐구능력으로는 조작적 정의, 변인 통제, 자료 해석, 가설 설정, 실험의 5가지 탐구요소가 있어, 13가지의 탐구요소를 제시하였다. SAPA의 탐구요소는 전형적인 모델로 이용되고 있다.

Moliter와 George(1976)는 4~6학년 대상으로 탐구요소별로 9문항인 SPST(The Science Process Skills Test)를 개발하여 추리와 검증 탐구기능을 측정하였다. 신뢰도는 추리 탐구기능에서 0.539~0.589, 검증 탐구기능에서 0.714~0.838이었다.

Fraser(1979)는 과학적 비판 사고 기능을 포함한 세 가지 하위요인으로 이루어진 TOES(The Test of Enquiry Skills)를 개발하여 9가지 탐구기능을 측정하였다. TOES의 내부 일치신뢰도(KR-20)는 7세에서 0.72, 10세에서 0.59이며, 7세 학생의 검사-재검사 신뢰도는 0.65~0.82이었다.

Dillashaw와 Okey(1980)는 SAPA의 통합 탐구능력인 가설 설정, 조작적 정의, 변인 통제, 연구 계획 수립, 자료 해석의 탐구 기능을 측정하기 위해 TIPS를 개발하였다. TIPS는 4지 선다형의 36문항으로 변별도 0.40, 난이도 0.53, 신뢰도(α)는 0.89이다.

Burns 등(1983)은 변인 통제, 조작적 정의, 검증 가능한 가설 설정, 자료와 그래프 해석, 실험 설계의 5가지 탐구능력에 관한 변별도 0.35, 난이도 0.53, 신뢰도 0.86인 TIPS II를 개발하였다.

Ross와 Maynes(1983)는 전문가와 초심자의 능력 차이를 바탕으로 학습위계를 사용하여, 일곱 가지 과학문제 해결능력에 관한 선다형 검사지를 개발하였다. 일곱 가지 기능은, ① 초점 개발(가설 형성), ② 톱 만들기(실험 계획), ③ 수집된 자료의 적절성 판단, ④ 정보 기록, ⑤ 자료의 관계 찾기, ⑥ 결론 도출, ⑦ 일반화이다. 이 검사에 대한 내부일치 신뢰도는 0.58~0.69, 검사-재검사 신뢰도(α)는 0.722~0.789이다.

Cronin 등(1985)이 4~8학년용으로 개발한 BAPS는 관찰, 추리, 예상, 측정, 의사소통, 분류의 6가지 기초탐구능력을 측정한다. 평균은 22.4(표준 편차 5.3), 독해력 FOG 지수는 3.95, 신뢰도(KR-20)는 0.78, 문항 난이도의 평균은 0.62, 문항 변별도의 평균(양분점 상관)은 0.34으로, 36문항으로 구성되었다.

Padilla와 Cronin(1986)는 6~9학년 대상으로 신뢰도(KR-20)가 0.89, 난이도 평균이 0.68, 변별도(양분점 상관) 평균이 0.43인 40문항의 MIPT를 개발하였다.

Padilla와 McKenzie(1986)는 7~12학년을 대상으로 26문항의 그래프 그리는 능력을 측정하는 TOGS(The Test of Graphing in Science)를 개발하였다. TOGS는 변별도(양분점 상관) 평균이 0.43, 신뢰도(KR-20)는 0.81이며, 탐구기능의 목표는 적절한 축의 선택, 그래프에 점 찍기, 가장 적합한 선 그리기, 내삽, 외삽, 변인들 간의 관계를 기술하기, 탐색하고, 설명하고, 두 그래프에 나타난 자료를 상호관련 짓기 등이 있다.

Mattheis 등(1988)은 신뢰도(KR-20) 0.75, 난이도 0.47, 변별도 0.41인 21문항의 POPS를 개발하였다.

국내의 과학탐구능력 검사지로서, 김창식 등(1992)이 문제 인식하기, 조사계획 세우기, 관찰하기, 측정하기, 실험하

기, 조사결과 정리하기, 조사결과 해석하기, 조사결과 발표하기 등의 8개 하위 기능별로 국민학교 4~6학년 대상의 지필 및 실기 평가 도구를 개발하였고, 이연우(1989)는 중학교 2학년을 대상으로 자료 분석과 해석 능력에 관한 탐구기능을 중심으로 검사지를 개발하였다. 이종기(1988)는 고등학생을 대상으로 가설 설정, 변인 통제, 실험 설계, 자료 변형, 그래프로 나타내기, 추리, 상관관계 결정, 인과관계 결정, 예상, 결론, 일반화와 모델 형성, 평가 등의 12가지 탐구기능에 대한 36문항을 개발하였다.

논리적 사고력 평가와 관련된 연구들은 주로 과학탐구능력과 관련지어 접근하는 연구들이 있다. Tobin과 Capie(1981)는 중학교 이상의 학생들을 대상으로 Lawson이 개발한 검사지의 문항을 골라서 변인 통제, 비례, 상관, 확률, 조합 논리의 다섯 가지 논리적 사고력을 측정하는 10문항의 TOLT(Test of Logical Thinking)를 만들었다. 신뢰도(α)는 0.84, 예언 타당도는 0.74, 요인 분석에 의한 구인타당도는 변량의 33%를 설명하고 있음을 제시하였다. 이 도구로 Tobin과 Capie(1982)는 6~8학년을 대상으로 TOLT와 TIPS의 상관관계가 $r=0.60$ 임을 보고하였다.

Roadranka 등(1983)은 보존 논리, 비례 논리, 변인 통제, 확률 논리, 상관 논리, 조합 논리의 여섯 가지 논리 사고력을 측정하는 GALT(Group Assessment of Logical Thinking)를 개발하였다. 21문항의 GALT 신뢰도(α)는 0.89, 각 하위 검사의 내부 일치도는 0.13~0.85이다. 문항 난이도는 평균 0.55, 변별도는 평균 0.49이다. 요인분석을 통하여 논리 사고력 요인에 관한 검사지임을 검증하였다. Baird와 Borich(1987)는 대학생을 대상으로 TIPS와의 GALT와의 상관을 조사한 결과 $r=0.62$ 임을 보고하였고, 한국, 미국, 일본, 필리핀의 중학생을 대상으로 TIPSⅡ와 GALT(12문항)를 투입하여 조사한 결과 각각 $r=0.56$, $r=0.63$, $r=0.64$, $r=0.45$ 의 상관관계가 있음을 보고하였다(김범기, 1989; Takemura, 1989; Manzano, 1987).

국내에서 취학전 아동에서부터 7학년의 우수집단을 대상으로 연구한 정연태와 김범기(1990)는 TIPSⅡ와 GALT의 상관 관계를 연구한 결과 $r=0.73$ 이었다. 일반 고등학생을 대상으로 연구한 임청환(1992)은 $r=0.48$ 을, 중학생을 대상으로 한 소원주(1993)는 $r=0.65$ 이었다.

Ⅲ. 연구 방법 및 절차

연구자들은 세미나를 통하여 10개의 탐구요소를 선정하고, 탐구요소에 따라 각각 3개의 문항을 만들어 검사지를 개발하였다. 개발된 문항을 예비 및 현장 검사를 통하여 문

항을 수정·보완하여 초·중학생들의 과학탐구능력 측정도구를 개발하였다.

1. 연구 대상

초·중학생의 과학탐구능력 검사지를 개발하기 위한 표집 대상은 [표 1]과 같이 대도시(특별시 및 직할시), 중·소도시(지방 도시), 읍·면 소재 지역별로 유층표집하였으며, 중학생은 성별을 고려하여 선정하였다.

[표 1] 예비 검사 및 현장 검사의 표집 대상

구 분	예비검사		현장검사 I				현장검사 II		
	국민 중학교	국민학교	국민학교	중학교	국민 중학교	국민 중학교	국민 중학교		
4 6 학년 2학년		(597)	5학년 6학년 1학년 2학년 3학년	(1184)		5 6 학년 1-3 학년			
	N=205	N=102	N=286	N=301	N=405	N=391	N=381	N=159	N=248
대도시	140	-	140	140	210	203	216	-	-
중 소 도시	-	102	90	93	92	96	91	-	-
읍 면 지역	65	-	66	68	103	92	81	159	248
전 체	307				1781			407	

2. 탐구요소의 선정 및 측정도구의 개발 근거

과학탐구능력 측정 도구를 개발하기 위하여 연구자들이 측정 도구 개발을 위한 많은 문헌들을 검토하고 협의한 결과 측정 도구 개발의 근거는 다음과 같다.

① 기초 탐구능력으로 관찰, 분류, 측정, 예상, 추리력, 통합 탐구능력으로 자료 변환, 자료 해석, 가설 설정, 변인 통제, 일반화 등 탐구요소 10개를 선정하였다.

② 범교과적(content free) 생활 소재나 익히 알고 있는 내용을 선정하여 암기 지식에 영향 받지 않도록 유의하여 개발한다.

③ 측정 도구는 4지 선다형으로 제작하며, 검사 시간은 40분 이내로 하여, 국민학교와 중학교에서 1시간 이내에 실시할 수 있도록 한다.

④ 문항 내용의 이해를 돕기 위하여 사진과 그림을 가급적 많이 사용하도록 한다.

⑤ 문항의 내용을 이해하기 위한 독해력의 수준은 국민학교는 5학년을 기준으로 작성한다.

⑥ 탐구요소별로 3~4 문항을 제작하고, 전체의 문항수는 40문항 이내로 한다.

위와 같은 준거에 따라 초·중학생에 맞는 문항을 개발하기 위하여 탐구요소별로 평가목표를 정의하였다. 각각의 탐구요소에 해당하는 평가목표는 [부록 1]에 제시되어 있다. 탐구기능의 평가목표에 따라 제작한 1~2개 문항을 검토·수정하고, 탐구요소별로 3개 문항을 최종적으로 선정하여 30문항의 초·중학생의 과학 탐구능력 측정도구 시안을 만들었다.

개발된 시안을 예비 및 현장 검사를 실시하여 문항을 수정하였다. 예비 검사를 실시하고 변별도, 난이도, 신뢰도, 타당도를 통하여 9개의 문항을 수정하고 4개의 문항을 교체하였으며, 현장검사 I를 실시하고 3개 문항을 수정하여 표준화된 과학탐구능력 측정도구를 개발하여 현장검사 II를 실시하였다. 예시 문항은 [부록 2]에 제시하였다.

3. 연구 결과의 분석 및 통계 처리

과학탐구능력 측정 도구를 개발하기 위하여 사용된 분석 방법으로는 난이도와 변별도 검증을 통한 문항 분석, 내부일치 신뢰도와 반분 신뢰도를 통한 신뢰도 검증, 내용타당도, 구인타당도, 공인타당도를 통한 타당도 검증 등이 사용되었다.

분석하였다. 내부일치 신뢰도는 KR-20의 공식을 이용하였으며, 반분신뢰도는 Spearman-Brown 신뢰도 공식을 이용하였다.

내용타당도는 본 연구에 참여하지 않은 과학교육 전문가 2명과 초·중등학교 과학 교사 11명에게 측정 문항과 목표 간의 적절성 여부의 평가를 의뢰하여 나타난 CVI(Index of Content Validity) 지수로 구하였다. 구인타당도는 요인분석과 군집 분석 등을 이용하였다. 그리고 최종 개발된 초·중학생들의 과학탐구능력 검사지를 GALT 검사지와 상관계수를 이용하여 공인타당도를 구하였다.

IV. 연구 결과 및 논의

과학탐구능력 측정도구를 개발하기 위하여 예비검사에 307명, 현장 검사 I에 1781명, 현장 검사 II에 407명에 대해서 조사하였다. 예비 및 현장 검사를 통하여 변별도, 신뢰도, 타당도를 조사·분석하여 문항을 개발하고, 본 연구에서 개발한 측정도구를 이용하여 초·중학생들을 대상으로 학교급별, 남녀별, 지역별 과학탐구능력을 조사하였다.

1. 과학탐구능력 검사지에 대한 문항 분석

과학탐구능력 검사지의 예비 및 현장검사의 정답률에 의한 문항의 난이도의 분포는 [표 2]와 같다. 예비 및 현장 검

[표 2] 과학탐구능력 측정도구의 난이도

난이도 지수	예비 검사			현장 검사 I			현장 검사 II		
	국민학교 N=205	중학교 N=102	전체 N=307	국민학교 N=597	중학교 N=1184	전체 N=1781	국민학교 N=159	중학교 N=248	전체 N=407
0.10 이하 문항수	-	-	-	-	-	-	-	-	-
0.11 ~ 0.29 문항수	3	2	3	1	1	1	2	2	2
0.30 ~ 0.79 문항수	23	18	20	25	22	24	24	24	24
0.80 ~ 0.89 문항수	1	7	4	3	6	4	3	3	3
0.90 이상 문항수	3	3	3	1	1	1	1	1	1
전체 평균	0.583	0.681	0.615	0.582	0.673	0.639	0.577	0.638	0.589

문항 분석에서는 문항정답률에 의한 난이도, Johnson의 DI 변별도 지수, 양분점(point-biserial) 변별 지수, 초·중학생들의 차이 검증을 통한 문항 변별력 검증을 실시하였으며, 상위집단과 하위집단의 비율은 각각 27%를 두그룹으로 나누어

사 I의 난이도의 평균값은 유사하나 현장 검사 I의 문항수의 분포가 고르게 배치되어 있다. 난이도의 평균은 국민학생이 0.58, 중학생은 0.67이고 전체적으로 0.64이다. Cronin 등(1985)이 개발한 BAPS의 문항 난이도 평균은 0.62이며, Dillashaw와

Okey(1980)의 TIPS가 0.53, Padilla와 Cronin (1986)의 MIPT가 0.68으로서 본 연구결과와 유사하게 나타났다.

문항의 변별도는 [표 3]과 같이 예비 검사보다 현장 검사 I에서 변별력있는 문항들이 많아졌다. 현장 검사 I에서 국민학생은 1번, 3번, 7번, 12번, 23번, 28번 문항에서, 중학생들은 1번, 3번, 7번, 12번, 23번, 26번, 27번, 28번 문항에서 변별도가 낮았다. 1번, 3번, 7번은 기초탐구기능이므로 정답률이 높아서 변별도가 낮고, 23번, 26번, 27번 문항은 통합 탐구기능이므로 난이도가 높아 변별도가 낮은 것으로 판단되어 12번과 28번 문항만을 수정하였다.

현장 검사 I의 자료 결과에서 국민학생과 중학생의 탐구능력을 통계적으로 유의미한 차이에 의한 문항의 변별 정도는 7번($t=-2.49, p<0.05$), 20번 문항($t=-3.63, p<0.001$)에서 국민학생이 중학생보다 정답률이 높은 유의미한 차이를 나타내고 있었다. 하지만 유의미한 평균의 차이가 0.05, 0.08로 그다지 크지 않으며, 난이도, 변별도, 내용타당도에서 문제가 없으므로 그대로 사용하기로 결정하였다.

과학탐구능력 검사지의 예비 및 현장 검사에서의 탐구기능별 평균 및 표준편차는 [표 4]와 같이 전체적으로 거의 유사하게 나타났다. 기초탐구능력에서 살펴볼 때, 현장검사 I

[표 3] 과학탐구능력의 문항별 변별도

변별도 지수	예비 검사		현장 검사 I			현장 검사 II		
	국민학교 N=205	중학교 N=102	국민학교 N=597	중학교 N=1184	전체 N=1781	국민학교 N=159	중학교 N=248	전체 N=407
0.40 이상 문항수	9	12	9	12	10	5	13	12
0.30 - 0.39 문항수	5	3	12	9	11	12	10	10
0.20 - 0.29 문항수	10	6	5	6	5	8	5	6
0.19 이하 문항수	6	9	4	3	4	5	2	2

[표 4] 예비 및 현장 검사의 탐구요소별 평균 및 표준편차

과학 탐구 능력	예비 검사		현장 검사 I				현장 검사 II						
	국민학교 4, 6학년 N=205		중학교 2학년 N=102		국민학교 5, 6학년 N=597		중학교 1~3학년 N=1184		국민학교 5, 6학년 N=159		중학교 1~3학년 N=248		
	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	
기초탐 구능력	관찰	2.77	0.47	2.82	0.38	2.24	0.68	2.44	0.66	2.29	0.71	2.21	0.74
	분류	1.33	0.74	1.76	0.79	1.62	0.88	2.02	0.84	2.08	0.83	2.35	0.76
	측정	2.22	0.80	2.49	0.64	2.01	0.83	2.30	0.78	1.99	0.94	2.18	0.83
	추리	1.79	0.86	1.95	0.73	1.83	0.93	2.06	0.80	1.83	0.82	2.03	0.79
통합탐 구능력	예상	2.25	0.80	2.37	0.70	2.00	0.90	2.22	0.85	1.86	0.94	2.15	0.90
	자료변환	1.80	0.96	2.28	0.80	1.32	0.91	2.20	0.84	1.61	0.89	2.07	0.91
	자료해석	1.24	0.89	1.59	0.99	1.39	0.88	1.73	0.92	1.31	0.87	1.65	0.92
	가설설정	1.21	0.77	1.56	0.86	1.20	0.84	1.58	0.88	1.20	0.78	1.49	0.83
전체	변인통제	1.43	0.84	1.83	0.87	1.69	0.81	1.88	0.83	1.74	0.88	1.88	0.86
	일반화	1.42	0.90	1.77	0.77	1.46	0.95	1.76	0.84	1.40	0.87	1.55	0.88
전체	17.48	4.15	20.44	3.91	17.47	4.53	20.19	4.42	17.32	4.26	19.55	4.64	

의 분류탐구기능의 평균이 국민학생 1.62, 중학생 2.02으로, 다른 기초탐구능력과 비교할 때 매우 낮은 점수를 나타내어 4번 문항을 쉬운 문장으로 수정하였다.

분석 결과를 바탕으로 4번, 12번, 28번을 수정하여 최종적인 초·중학생들의 과학탐구능력 측정 도구를 개발하였다. 개발된 검사지를 현장 검사 I 을 실시한 읍·면 소재지 학교의 다른 학생들을 대상으로 현장 검사 II 를 실시하였다.

현장 검사 II 의 내용에서 수정한 3문항을 중심으로 살펴보면, 4번 문항의 정답률은 현장검사 I 에서 국민학생이 30%, 중학생이 49%의 정답률이었으나, 현장 검사 II 에서는 국민학생이 80.9%, 중학생이 84.6%이었다. 또한, 변별도가 0.24~0.32로 적합한 변별력을 보여주었으며, 분류 탐구기능은 다른 기초탐구능력의 점수와 유사한 결과를 나타내었다. 12번 문항은 변별도의 평균이 0.19에서 0.29로 향상되어 적합한 변별력이 있었으며, 28번 문항의 변별도는 -0.05~0.25이나 학년이 낮은 국민학교 5, 6학년과 중1학년이 -0.05~0.09로 아주 낮은 변별도를 보인 반면에, 학년이 높은 중 2, 3학년은 0.14~0.25의 낮은 변별도를 나타내고 있으므로 학년의 차이에 따른 변별력이 있는 문항으로 생각된다. 그리고 현장 검사 II 가 과학탐구능력이 낮은 읍·면소재지 학교를 통하여 검증하였으므로, 대도시 및 중·소도시의 학생들을 대상으로 검증하면 집단 검사지로서 적합한 변별도가 나타난다고 생각된다.

반분신뢰도인 Spearman-Brown의 신뢰도 계수는 국민학생이 0.68, 중학생이 0.69, 전체적으로 0.71로 양호하였다.

본 연구의 과학탐구능력 검사지의 내부일치신뢰도는 Dillashaw와 Okey(1980)의 TIPS(0.89)나 Padilla와 Cronin(1986)의 MIPT(0.89)보다 약간 낮으나, Cronin 등(1985)의 BAPS(0.78)나 Mattheis 등(1988)의 POPS(0.75)와 유사하며, 형식 조작적 사고력 검사지를 개발한 허형(1993)의 신뢰도(α) 0.71보다 높다. 따라서 신뢰로운 과학탐구능력 측정 도구라고 할 수 있다.

3. 과학탐구능력 측정 도구에 대한 타당도 검증

본 연구에서 개발한 과학탐구능력 측정 도구에 관하여 타당도를 검증하였다. 내용 타당도는 94.1%이었고, 공인 타당도 및 구인 타당도는 다음과 같았다.

1) 내용타당도

과학 탐구 능력에 관한 내용 타당도는 본 연구에 참여하지 않은 과학교육 전문가 2명과 현장 초·중등 과학교사 11명에게 측정 문항과 목표 간의 적절성 여부의 평가를 의뢰하였다. 현장검사 I 에서 개발된 검사지의 내용타당도의 CVI 지수는 과학교육전문가는 100%, 중등과학교사 5명은 94.7%, 초등교사 6명은 91.7%였으며, 정답률은 과학교육전

[표 5] 과학 탐구능력의 예비 및 현장 검사의 신뢰도

신뢰도	예비 검사		현장 검사 I						현장 검사 II		
	국민학교	중학교	국민학교		중 학교			전 체 (1752)	국민 학교 (159)	중학교 (248)	전체 (407)
	4, 6년 (205)	2년 (102)	5년 (296)	6년 (301)	1년 (405)	2년 (391)	3년 (388)				
내부일치 신뢰도 (KR-20)	0.687	0.667	0.662	0.744	0.698	0.747	0.733	0.743	0.683	0.751	0.740
반분 신뢰도 (Spearman-Brown)	-		0.679		0.686			0.707	0.562	0.639	0.630

()속의 숫자는 조사 학생수

2. 과학탐구능력 측정도구의 신뢰도

측정 도구에 대한 신뢰도는 검사의 동질성의 정도를 나타내는 KR-20 신뢰도와 반분 신뢰도인 Spearman-Brown을 구하였다. [표 5]와 같이 KR-20은 예비 검사보다 현장 검사 I 에서 향상된 0.74을 나타내고 있다. 현장 검사 I 에 대한

문가 98.3%, 중등과학교사 4명이 93.3%, 초등교사 6명은 87.2%으로 나타났다.

예비 검사를 통하여 수정하여 개발된 검사지의 어휘 수준을 5학년에 알맞는지의 여부를 과학교육전문가 및 현장 과학교사를 대상으로 조사한 결과는 95.1%가 적절한 것으로 응답하였다.

현장 검사 I의 연구 결과를 통하여 수정한 새 문항에 대한 내용타당도는 본 연구에 참여하지 않은 두 명의 과학교육 전문가의 평가 목표와 문항 사이의 일치율과 정답률이 모두 100.0%였다. 그리고 수정된 새문항이 국민학교 5학년 학생에게 맞는 문장을 사용하였는지를 과학교육 대학원생 18명에게 설문한 결과 88.9%의 응답이 보통 수준 이상의 어휘라고 응답하여 적절한 것으로 사료된다.

2) 공인타당도

본 검사지의 공인타당도(concurrent validity)는 현장검사 II에서 구하였다. 본 검사지와 GALT 검사지와 상관은, 국민학교 6학년이 0.43, 중학교 1학년이 0.51, 중학교 2학년이 0.64로 나타났다. GALT와 TIPS II의 상관관계를 분석한 Baird와 Borich(1987)의 연구는 0.62, 중학생을 대상으로 조사한 소원주(1993)의 연구는 0.65이었다.

이러한 연구의 상관관계는 본 연구의 중학교 2학년의 0.64와 유사하고, 김범기(1989)의 연구에서 한국의 중학생을 대상으로 한 연구결과는 0.56으로, 중학생 1, 2학년에 대한 본 연구의 공인타당도와 거의 같음을 알 수 있다.

3) 구인 타당도

요인 분석(factor analysis)을 이용하여 구인 타당도를 조사하였다. 과학탐구능력 측정도구의 현장 검사 I 자료를 요인분석하여 문항을 검증하였다. 30문항에 대한 회전하기 이전의 요인행렬에서 하나의 요인에 아이겐 값이 3.88이고, 전체 변량의 13%를 설명하였다. 주성분 분석한 결과는 9개 요인의 아이겐 값이 1이상이며 10번째 요인은 0.988이다. 커

뮤날리티가 0.4 이상인 요인이 6개이며, 나머지 4요인도 0.30이상으로 나타났으며, 10개 요인이 전체 변량의 45.9%를 설명하였다. 형식조작적 사고력 검사지를 개발한 허형(1993)의 연구에서 9가지 요인들이 전체 변량의 46.8%를 설명하고 있어 본 연구결과와 유사하다.

현장검사 II에 대한 자료에서 10가지 탐구요소들을 직교 회전한 결과는 [그림 1]과 같다. [그림 1]에서 기초탐구능력에 관찰 분류, 측정, 예상, 추리의 다섯 가지 탐구요소가 0.48~0.74으로 부하되었으며, 통합탐구능력에 자료 변환, 자료 해석, 가설 설정, 변인 통제, 일반화의 다섯 가지 탐구요소가 0.44~0.66으로 부하되었다.

허형(1993)은 형식조작적 사고력 검사지의 개발에서 요인별 부하량을 0.30 이상으로 설정하였고, Zohar, Weinberger와 Tamir(1994)는 BCT(Biology Critical Thinking) 검사지 개발에서 0.35이상 부하되는 요인을 선택하였다.

본 연구에서 10가지 탐구요소들이 기초 및 통합탐구능력에 0.4이상으로 부하되었으므로 기초 및 통합탐구능력에 각각 5가지 탐구요소들로 이루어져 있음이 검증되었다. 따라서 본 연구에서 개발한 과학탐구능력 검사지는 초·중학생들의 10가지 탐구기능과 기초 및 통합탐구능력을 측정하는 타당성 있는 측정 도구라고 할 수 있다.

과학탐구능력과 탐구기능 사이에 얼마나 상호독립적인가를 알아보기 위하여 상관계수를 알아 본 결과는 [표 6]과 같다. 과학탐구능력과 각 하위 탐구요소 간의 상관은 모두 통계적으로 유의있는 높은 상관관계를 나타내고 있다. 그 중에서 자료 변환과의 상관 계수($r=0.65$)가 가장 높고, 관찰과의 상관계수($r=0.44$)가 가장 낮았다. 각 탐구기능들 간의 상

요인 분석		변인	요인 1	요인 2
9 10 6 8 7 5 4 2 3 1	1.	관찰	0.745	-0.070
	2.	분류	0.597	0.148
	3.	측정	0.612	0.100
	4.	예상	0.616	0.267
	5.	추리	0.485	0.303
	6.	자료변환	0.278	0.623
	7.	자료해석	0.346	0.440
	8.	가설설정	0.031	0.559
	9.	변인통제	0.083	0.657
	10.	일반화	0.105	0.663

[그림 1] 과학탐구능력의 직교회전 후 요인 분석 결과(현장 검사 II)

관계수는 모두 $p < 0.001$ 수준에서 $r = 0.098 \sim 0.316$ 의 상관 계수를 나타내었다. 이러한 결과는 각 탐구요소들이 어느 정도 서로 독립적임을 암시하고 있다.

현장검사 I의 자료에서, 국민학생들의 과학탐구 능력의 하위 탐구요소들의 군집 정도를 알아 보기 위하여 상관계수를 알아 본 결과 $r = 0.306 \sim 0.059$ 으로 분포되어 있다.

탐구능력의 평균 점수는 11.44(76.3%)으로, TIPS를 이용하여 7~12학년의 통합탐구능력을 조사한 Dillashaw와 Okey(1980)의 연구 결과(18.99, 52.8%), MIPT를 이용하여 중학생의 통합탐구능력을 검사한 Padilla와 Cronin(1986)의 연구 결과(27.14, 67.9%), 우리 나라의 중학생을 대상으로 TIPS II를 이용한 소원주(1993)의 연구 결과 57.3% 보다는

[표 6] 과학탐구능력과 탐구기능별 단순상관관계

탐구기능	관찰	분류	측정	추리	예상	자료변환	자료해석	가설설정	변인통제	일반화
관찰	1.000									
분류	.183*	1.000								
측정	.213*	.198*	1.000							
추리	.148*	.165*	.174*	1.000						
예상	.206*	.231*	.198*	.232*	1.000					
자료해석	.213*	.299*	.235*	.247*	.316*	1.000				
자료변환	.197*	.227*	.246*	.173*	.164*	.225*	1.000			
가설설정	.157*	.236*	.177*	.164*	.225*	.285*	.259*	1.000		
변인통제	.101*	.163*	.098*	.098*	.136*	.205*	.138*	.145*	1.000	
일반화	.156*	.235*	.175*	.202*	.248*	.270*	.232*	.237*	.209*	1.000
전체	.443*	.551*	.493*	.483*	.577*	.648*	.586*	.545*	.449*	.558*

* $p < 0.001$

4. 초·중학생들의 과학 탐구 능력의 실태

초·중학생들의 과학탐구능력의 실태를 현장 검사 I의 자료를 분석하여 학교급별, 성별, 지역별로 알아보고자 한다.

초·중학교별로 탐구기능에 따른 과학탐구능력을 t-검정한 결과는 [표 7]과 같이 모든 탐구기능에서 중학생들이 국민학생보다 통계적으로 유의미하게 높은 것을 볼 수 있고, 특히 자료 변환 기능에서 중학생들($M=2.21$)이 국민학생($M=1.45$)보다 큰 차이($t=16.66, p < 0.001$)가 있다.

초·중학생들의 기초 및 통합탐구능력의 차이를 비교한 결과는 기초탐구능력의 평균은 1.09, 통합탐구능력의 평균은 1.68의 차이가 있었으며, 통합탐구능력에서 더 큰 차이가 있었다. 따라서 중학생들이 국민학생보다 통합탐구능력이 높음을 알 수 있다.

본 연구의 기초탐구능력의 평균 점수는 국민학생이 9.95(66.3%), 중학생이 11.4(73.6%)으로서, Cronin 등(1985)의 BAPS검사서에서 5학년이 23.1(64.1%), 전체 평균 27(75%)과 거의 유사함을 알 수 있다. 본 연구에서 중학생의 통합

[표 7] 국민학교와 중학교 학생들의 과학탐구능력 비교

탐구요소	국민학교(N=597)		중학교(N=1184)		t	
	평균	표준편차	평균	표준편차		
기초 탐구 능력	관찰	2.24	0.68	2.44	0.66	4.70*
	분류	1.67	0.88	2.02	0.84	8.29*
	측정	2.01	0.83	2.30	0.78	3.55*
	추리	1.83	0.93	2.06	0.80	5.29*
	예상	2.00	0.90	2.22	0.85	4.90*
소계	9.95	2.38	11.04	2.36	9.22*	
통합 탐구 능력	자료변환	1.32	0.91	2.20	0.84	16.66*
	자료해석	1.39	0.88	1.73	0.92	4.28*
	가설설정	1.20	0.84	1.58	0.88	6.21*
	변인통제	1.69	0.81	1.88	0.83	3.47*
	일반화	1.46	0.95	1.76	0.84	5.75*
소계	9.80	2.84	11.44	2.36	11.94*	
전체	17.47	4.53	20.19	4.42	12.04*	

* $p < 0.001$

대체로 높았다. 이러한 결과는 본 연구의 검사지가 중학생 뿐만 아니라 국민학생들의 통합탐구능력도 함께 측정할 수 있도록 문항들을 만들었기 때문이라고 생각된다.

중학생들의 성별에 따른 탐구능력을 비교하면 자료 변환, 자료 해석, 가설 설정 탐구요소를 제외한 모든 요소에서 남자가 여자보다 높은 유의미한 차이를 나타내고 있다. 전체적으로 남자(M=20.67)가 여자(M=19.74)보다 높았다($t=3.63$, $p<0.001$). 중학교 학생들의 남녀별 탐구능력 및 요소별 평균 및 표준편차는 [부록 3]과 같다.

국민학생들의 과학탐구능력의 지역별 차이를 알아본 결과는 분류, 측정, 예상, 자료 변환, 가설 설정, 변인 통제에서 통계적으로 유의미한 차이를 보여주고 있다.

이를 다시 Scheffe 사후검증을 한 결과, 분류·측정에서 대도시 아동들이 읍·면지역 아동들보다 높은 차이를, 예상·가설 설정·변인 통제에서 대도시 아동이 중·소도시 아동보다 높은 차이를, 자료 변환에서 중·소도시 아동들이 대도시와 읍·면지역 아동들보다 낮은 유의미한 차이를 보이고 있었다. 전체적인 과학탐구능력은 대도시 지역의 국민학생(M=18.35)이 중·소 도시 지역(M=16.49)과 읍·면 지역(M=16.99)의 국민학생들보다 높다($F=10.61$, $p<0.001$).

[표 8] 초·중학생의 과학탐구능력의 지역별, 성별, 학년별 변량분석

변량의 원인	자승화	자유도	평균 자승화	F
주효과	3608.38	4	902.09	46.44***
학교급	2841.77	1	2841.77	146.29***
지역	409.16	2	204.58	13.13***
성별	254.98	1	254.98	10.53***
2-way 상호작용	291.75	5	58.35	3.00*
지역-학교급	201.22	2	100.61	5.18**
성별-학교급	73.01	1	36.51	1.88
지역-성별	25.87	2	25.87	1.33
3-way 상호작용				
지역-성별-학교급	104.76	2	52.38	2.70*
설명된 변량	4004.89	11	364.08	
잔여효과	34363.86	1769	19.43	18.74***
합계	38368.75	1780	21.52	

*** $p<0.001$, ** $p<0.01$, * $p<0.05$

이러한 지역별 차이를 초·중학생의 기초 및 통합탐구능력을 비교한 결과는 대도시의 초·중학생들이 다른 지역의 학생들보다 기초 및 통합탐구능력이 높은 점수를, 중·소도시의 국민학생이 통합탐구능력에서 다른 지역의 학생들보다 낮은 점수를 나타내고 있다. 국민학교 학생들의 탐구능력 및 요소별 평균 및 표준편차는 [부록 4]와 같다.

초·중학생들의 과학탐구 능력에 대한 지역별, 학교급별, 성별의 상호작용 효과를 알아 보기 위하여 삼원 변량 분석(3-way ANOVA)을 한 결과는 [표 8]과 같이 지역-학교급별($F=5.18$, $p<0.01$), 지역-성별-학교급별($F=2.70$, $p<0.05$)은 의미있는 상호작용을 나타내고 있지만, 지역-성별, 성-학교급별의 상호작용 효과는 없었다.

V. 결론

본 연구는 학생들의 탐구능력을 진단하고, 과학학습에서 탐구학습의 효과를 알아보는 데 도움을 주기 위하여 초·중학생들의 과학탐구능력을 측정하는 도구 개발에 그 목적이 있다.

과학탐구능력은 기초탐구능력과 통합탐구능력으로 구분하였다. 기초탐구능력으로는 관찰, 분류, 측정, 추리, 예상 등 5개의 탐구요소로, 통합탐구능력으로는, 자료변환, 자료해석, 가설설정, 변인통제, 일반화 등 5개의 탐구요소로 구분하였다. 그리고, 각각의 탐구능력을 측정하기 위하여 각각 탐구요소마다 3개의 문항을 개발하였다. 탐구능력을 측정하는 도구는 범교과적인 내용으로, 문항 내용의 이해를 돕기 위하여 그림을 가능한 많이 이용하였으며, 측정의 편리성을 위하여 객관식의 4지 선택형으로 30문항이며, 측정 시간은 40분이내를 목표로 개발하였다.

본 연구에서 개발한 과학탐구능력 측정 도구는 예비 및 현장 검사를 실시하였고, 이를 통하여 얻은 자료를 이용하여 난이도, 변별도, 내용타당도, 구인타당도, 공인타당도 등으로 분석하였고, 문항을 수정 개발하였다.

예비 검사와 2차에 걸친 현장 검사를 통하여 수정 보완하여 최종적으로 확정된 과학탐구능력 측정도구의 난이도는 평균 58.9이었고, 양분점 변별도는 0.34이었다. 또 요인 분석 결과 기초탐구능력에 관찰, 분류, 측정, 예상 및 추리가, 통합탐구능력에 자료변환, 자료해석, 가설설정, 변인통제 및 일반화가 0.4이상 부하되어 과학탐구요소로 확인되었다. 따라서, 각 탐구요소들은 각각 의미있는 탐구능력의 하위 요소로 인정할 수 있었다. 또한, 본 도구를 사용하여 학교별, 성별, 지역별 차이를 조사한 결과 의미있는 차이를

나타낸 것으로 보아서 본 도구가 초·중학생들의 과학탐구 능력을 평가하는 데 유용하다는 것을 보일 수 있었다.

참 고 문 헌

- 권재술(1990). 과학적 탐구능력과 그 평가, 연구보고(90-1), 중앙교육평가원
- 권재술·김창식·이화국·김영수·김찬중(1991). 과학학습 평가, 교육과학사.
- 김창식·권치순·이화국·한안진(1992). 과학 탐구실험 능력 평가 모형 및 평가 도구의 개발에 관한 연구, 한국과학교육단체총연합회.
- 나종국(1993). 생활소재를 이용한 과학 탐구 능력 평가문항 개발, 한국교원대학교 대학원 석사학위 논문.
- 소원주(1993). 중학생의 논리적 사고력과 통합적 탐구능력에 관한 한 일 비교 연구, 한국교원대학교 대학원 석사학위논문.
- 이연우(1989). 과학 탐구 능력 측정을 위한 표준화 검사지 개발-중학교 2학년의 자료분석과 해석 능력을 중심으로-, 한국교원대학교 대학원 석사학위논문.
- 이종기(1988). 고등학생의 과학 탐구 능력 측정을 위한 평가 도구 개발, 한국교원대학교 대학원 석사학위논문.
- 이항로(1991). 고등학생의 과학 탐구 능력 측정을 위한 평가 도구 개발-지구과학소재를 중심으로- 한국교원대학교 대학원 석사학위논문.
- 임정환(1992). 논리적 사고력과 과학탐구 기능 요소의 위계적 분석, 한국교원대학교 대학원 박사학위논문.
- 정연태·김범기(1990). 사고훈련에 의한 우수집단의 논리적 사고력 및 프로세스 스킬의 발달에 관한 연구, 물리교육 8(1)
- 허 형(1993). 형식조작적 사고력 검사 개발 연구, 교육학연구, 31(3)
- 金範基(1989). 物理教育の韓·日比較研究, 廣島大學博士學位論文.
- AAAS(1970) Science-A Process Approach, Washington D. C.: AAAS/Xerox Co.
- Adey, P. S. & Harten, W.(1986). A Piagetian Analysis of Process Skill Test Items, *Journal of Research in Science Teaching*, 23(8): 707-726.
- Baird, W. E. & Borich, G. D.(1987). Validity Considerations for Research on Integrated Science Process Skills and Formal Reasoning Ability, *Science Education*, 71(2): 259-269.
- Burns, J. C., Okey, J. R. & Wise, K. C.(1983) Integrated Process Skills Test II, Department of Science Education University of Georgia Athens, Georgia 30602
- Dillashaw, G. H. & Okey, J. R.(1980). Test of the Integrated Science Process Skills for Secondary Science Students, *Science Education*, 64: 601-608.
- Doran, R. L. & Meng, E.(1993). Improving Instruction and Learning through Evaluation Elementary School Science, ERIC.
- Fraser, B. J.(1979). Test of Enquiry Skills Handbook. Hawthorne, Victoria: Australian Council Educational Research.
- Helgeson, S. L.(1992). Problem Solving Research in Middle/Junior High School Science Education, ERIC.
- Hur, M.(1984). Evaluation of Inquiry Activity in Science Curricula, Ed. D. Thesis, Columbia University.
- Manzano, V. U.(1987). Factors in Science Education Correlates of Adolescents Students' Logical Thinking Skills in the Philippines and Japan. Doctoral Dissertation, Hiroshima University.
- Mattheis, F. E., Jones, M. L. & Nakayama, G.(1988) A U.S.-Japan Cooperative Study of Science Process Skills in Middle School Students Participating in the Fast Program(Results of Pre-Pilot-Tests in North Carolina).
- Molter, L. L. & George, K. D.(1976). Development of a Test of Science Process Skills, *Journal of Research in Science Teaching*, 13(5): 405-412.
- Ostlund, K. L.(1992). Science Process Skills(Assessing Hands-on Student Performance), California, Addison-Wesley Publishing Company.
- Padilla, M. J., Okey, J. R. & Dillashaw, G. H.(1983). The Relationship Between Science Process Skill and Formal Thinking Abilities, *Journal of Research in Science Teaching*, 20(3): 239-246.
- Padilla, M. J. & Cronin, L. L.(1986) Middle Grades Integrated Science Process Skills Test(MIPT). Department of Science Education, University of Georgia, Athens, GA.
- Padilla, M. J., Cronin, L. L. & Twiest, N.(1985) Test of Basic Process Skills(BAPS), Department of Science

- Education University of Georgia Athens, GA 30602
- Padilla, M. J. & Mckenzie, D. L.(1986). The Construction and Validation of the Test of Graphing in Science(TOGS), *Journal of Research in Science Teaching*, 23(7): 571-579.
- Roadrangka, V., Yeany, R. H, & Padilla, M. J.(1983). The Construction of a Group Assessment of Logical Thinking(GALT). Paper presented at the 56th annual meeting of the National Association for Research in Science Teaching, Dallas, TX, April 5-8.
- Ross, J. A. & Maynes, F. J.(1983). Development of a Test of Experimental Problem-Solving Skills, *Journal of Research in Science Teaching*, 20(1): 63-75.
- Takemura, S.(1989). A Study of Cognitive Skills and Other Related Factors in Science Education, Department of Science Education, Faculty of Education, Hiroshima University, Research Project Report, Grant-in-Aid for scientific during 1986-1988 Project Number: 61450098.
- Tobin, K. G. & Capie, W.(1981). The Development and Validation of a Group Test of Logical Thinking, *Educational and Psychological Measurement*, 41(2): 413-423.
- Tobin, K. G. & Capie, W.(1982a). Relationships between Formal Reasoning Abilities, Locus of Control, Academic Engagement and Integrated Process Skills Achievement, *Journal of Research in Science Teaching*, 19(2): 113-121.
- Tobin, K. G. & Capie, W.(1982b). Development and Validation of a Group Test of Integrated Science Processes, *Journal of Research in Science Teaching*, 19(2): 133-141.
- Wellington, J.(1989). Skill and Process in Science Education-An Critical Analysis-, The Falmer Press.
- Zohar, A., Weinberger, Y. & Tamir, P.(1994). The Effect of the Biology Critical Thinking Project on the Development of Critical Thinking, *Journal of Research in Science Teaching*, 31(2): 183-196.

(ABSTRACT)

The Development of an Instrument for the Measurement of Science Process Skills of the Korean Elementary and Middle School Students

Kwon, Jae-Sool and Kim, Beom-Ki
(Korea National University of Education)

There are a considerable number of instruments in testing science process skills in Korea as well as America and Europe. However, it has been difficult to find the instruments encompass the wide range of scientific process skills. Most instruments developed and used in Korea were focused on a school level or a few components of scientific process skills.

In this study, the researchers examined the whole body of scientific process and identified 10 component skills. Three items for each component process skills were developed and revised by two pilot tests.

The final instruments showed moderate difficulty, discrimination index, reliability, and validity. The instrument was also tested the usability by finding differences of students' ability on science process skills by grade levels, sexes, and city size. The results showed that the instrument was proved to be useful to discriminates the differences.

[부록 1] 과학탐구능력 하위 요소별 평가 목표

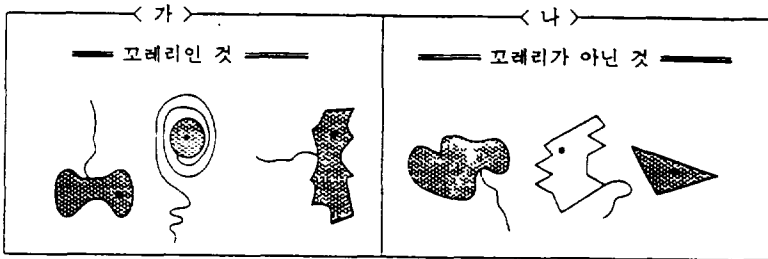
탐구요소	평 가 목 표
관찰	① 주어진 그림을 관찰하고 차이점을 바르게 찾아낼 수 있다. ② 사물, 사진이나 사진이 주어졌을 때, 여러 가지 감각을 이용하여 올바르게 관찰할 수 있다. ③ 변화하는 사물, 사진이나 사진이 주어졌을 때, 변화하는 정도를 올바르게 관찰 할 수 있다.
분류	① 주어진 여러 가지 사물을 관찰하고, 한가지 특성에 따라 사물을 분류할 수 있다 ② 주어진 여러 가지 사물을 관찰하고, 복합 특성에 따라 사물을 분류할 수 있다 ③ 주어진 여러 가지 사물의 분류된 특징을 관찰하고 복합 특성을 바르게 찾을 수 있다.
측정	① 계측기를 바르게 사용할 수 있다. ② 막대자를 이용하여 사물을 바르게 잴 수 있다. ③ 격자(모눈 눈금)를 이용하여 넓이를 계산할 수 있다.
추리	① 관찰에 의해서 얻어진 자료를 바탕으로 이미 일어난 사건을 돌이켜 생각해보는 것 ② 측정에 의해서 얻어진 자료를 바탕으로 이미 일어난 사건을 돌이켜 생각해보는 것
예상	① 관찰에 의해서 얻어진 자료를 바탕으로 아직 일어나지 않은 사건을 미리 생각 해보는 것 ② 측정에 의해서 얻어진 자료를 바탕으로 아직 일어나지 않은 사건을 미리 생각해 보는 것
자료변환	① 실험이나 조사를 통하여 얻어진 자료를 정리하여 도표를 만들 수 있다. ② 실험이나 조사를 통하여 얻어진 비단순 관계의 도표를 보고 그래프로 표현할 수 있다. ③ 정성적으로 얻어진 자료를 변환할 수 있다.
자료해석	① 실험이나 조사를 통하여 얻어진 비례 관계의 그래프를 해석하여 두 변인 사이의 관계를 알 수 있다. ② 실험이나 조사 활동을 통하여 얻어진 비단순 관계의 그래프를 해석하여 두 변인 사이의 관계를 알 수 있다. ③ 실험이나 조사를 통하여 얻어진 복합 관계의 그래프를 해석하여 두 변인 사이의 관계를 알 수 있다.
가설설정	① 검증 가능한 가설을 찾을 수 있다. ② 주어진 실험 상황에서 가설을 설정할 수 있다. ③ 가설을 검증할 수 있는 실험을 설계할 수 있다.
변인통제	① 주어진 실험 상황에서 독립변인을 찾을 수 있다. ② 주어진 실험 상황에서 종속변인을 찾을 수 있다. ③ 주어진 실험 상황에서 통제된 변인을 찾을 수 있다.
일반화	① 주어진 자료의 공통성, 경향성, 규칙성 등을 잘 파악할 수 있다. ② 관찰 측정된 사실을 가장 잘 설명할 수 있는 모델을 설정할 수 있다. ③ 주어진 사실로부터 일반화할 수 있는 영역을 잘 찾아낼 수 있다.

[부록 2] 과학탐구능력 평가문항의 예

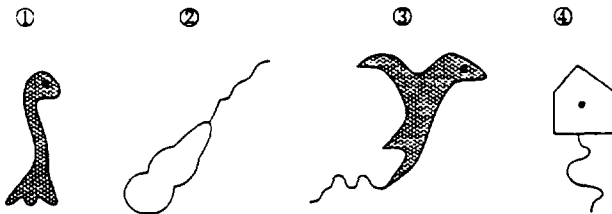
<5번 문항>

탐구 요소	평가 목표	정답	내용 타당도	난이도 (1 차)	변별도(1차)
분 류	사물의 분류된 복합 특성 찾기	②	100(%)	0.72	0.35

5. 그림 <가>는 꼬레리의 모양이고, 그림 <나>는 꼬레리가 아닌 것이다.



다음 중에서 꼬레리인 것은? -----()



<25번 문항>

탐구 요소	평가 목표	정답	내용 타당도	난이도 (1 차)	변별도(1차)
가설설정	검증 가능한 가설 찾기	④	69.2(%)	0.56	0.42

25. 몇몇 학생들이 모여서 실험을 하려고 한다. 다음 중에서 실험을 통하여 알아보기에 가장 적절한 것은? -----()

- ① 어떤 치약이 더 좋은가?
- ② 식물은 여름에 왜 잘 자랄까?
- ③ 머리가 좋은 사람은 공부를 잘하는가?
- ④ 온도의 변화는 바퀴벌레의 움직임에 영향을 줄까?

<부록 3> 중학생들의 남녀별 과학탐구능력

과학탐구능력		남(N=580)		여(N=604)		t
		M	SD	M	SD	
기초탐구능력	관찰	2.54	0.63	2.35	0.67	5.19***
	분류	2.08	0.82	1.95	0.85	2.73**
	측정	2.36	0.78	2.25	0.77	2.52*
	추리	2.12	0.77	2.01	0.83	2.31*
	예상	2.28	0.86	2.16	0.84	2.35*
소 계		11.38	2.37	10.72	2.30	4.89
통합탐구능력	자료변환	2.28	0.83	2.14	0.84	2.78**
	자료해석	1.72	0.94	1.74	0.90	-0.24
	가설설정	1.59	0.85	1.57	0.90	0.33
	변인통제	1.88	0.83	1.88	0.83	-0.17
	일반화	1.82	0.85	1.69	0.82	2.79*
소 계		11.60	2.75	11.29	2.62	2.01*
전 체		20.67	4.55	19.74	4.26	3.63***

***p<0.001, **p<0.01, *p<0.05

[부록 4] 국민학생들의 지역별 과학탐구능력

과학탐구능력		대 도시 (N=280)		중·소도시 (N=183)		읍·면 지역 (N=134)		F
		M	SD	M	SD	M	SD	
기초탐구능력	관찰	2.30	0.64	2.26	0.72	2.29	0.65	0.28
	분류	1.80	0.82	1.62	0.90	1.42	0.91	9.29***
	측정	2.26	0.80	2.10	0.81	2.03	0.88	4.38*
	추리	1.89	0.90	1.81	0.91	1.73	0.93	1.35
	예상	2.15	0.89	1.81	0.93	1.98	0.89	8.02***
소 계		10.41	2.34	9.61	2.23	9.45	2.49	10.39***
통합탐구능력	자료 변환	1.52	0.99	1.26	0.86	1.57	0.90	5.81**
	자료 해석	1.63	0.95	1.43	0.85	1.46	0.96	2.99
	가설 설정	1.41	0.88	1.13	0.81	1.34	0.91	5.93**
	변인 통제	1.81	0.83	1.58	0.91	1.77	0.75	4.33*
	일반화	1.57	0.88	1.48	1.00	1.41	0.86	1.45
소 계		10.12	2.86	9.24	2.76	9.92	2.81	5.51*
전 체		18.35	4.57	16.49	4.25	16.99	4.52	10.61***

***p<0.001, **p<0.01, *p<0.05