

과학 영재 선발 도구로서 지필 검사 의 적합성 탐색을 위한 질적 및 양적 문항 분석

이기영* · 동효관¹ · 홍준의² · 김현경² · 조봉제³

¹강원대학교 · 한국교육과정평가원 · ²한성과학고등학교 · ³서울전자고등학교

Qualitative and Quantitative Analysis of Paper-Pencil Test Items for Exploring its Appropriateness as a Selection Tool of the Gifted in Science

Lee, Kiyoung* · Dong, Hyokwan¹ · Hong, Juneui² · Kim, Hyunkyoung² · Jo, Bongjae³

¹Kangwon National University · Korea Institute of Curriculum & Evaluation ·

²Hansung Science High School · ³Seoul Electronics High School

Abstract: The purpose of this study was to analyse the qualitative and quantitative characteristics of paper-pencil tests for exploring its appropriateness as a selection tool of the gifted in science. For this purpose, we developed two (internal and external) item analysis frameworks, and applied these frameworks to analyse qualitative characteristics. Also, we analysed the relationship between two characteristics. The results of analysing qualitative characteristics revealed that the portion of items with acceleration context exceeding middle school curriculum level was relatively large, which caused low content validity. Furthermore, there was considerable deviation in content and context by subject matter and year, which caused test instability. Items measuring knowledge domain was the most prevalent, and too much weight on data interpretation & analysis domain in inquiry process skills. In case of creativity test, the portion of items measuring convergent thinking was much larger than that of divergent or associative thinking. Most of these items were represented by using pictures and tables rather than using graphs. Item types of multiple-choice and short answers were superior to essay types. Discrimination index, on the whole, was appropriate (above 0.3), but item difficulty showed a vast deviation (0.01~0.90). Correlation coefficients among subject matters and test tools were very low, and test reliabilities were also low. Low item difficulty & high discrimination index item types were distinguishable. Items with acceleration context were more discriminating than enrichment context. Implications of developing quality paper-pencil test items in the selection of gifted students are discussed.

Key words: gifted in science, paper-pencil test, item analysis, selection of gifted students

I. 서론

영재 교육은 도전적인 학습 환경의 제공을 통해 타고난 잠재 능력을 계발할 수 있도록 기회를 제공하는 측면에서 중요성을 가진다고 할 수 있다(조석희, 2001). 영재 교육을 계획함에 있어서 가장 먼저 이루어져야 할 일은 영재를 선발하는 일이다(Davis & Rimm, 1998). 영재 선발은 영재 교육을 받을 사람을 가려내는 것으로, 일반 학교 교육으로는 타고난 우수

잠재 능력의 개발이 어려운 학생들을 가려내어 학생들의 능력과 요구에 맞는 교육 프로그램을 제공하는 것을 목적으로 하고 있다(윤경미, 김정섭, 2006).

국내에 영재 교육이 도입되던 초기에 영재 선발은 영재의 정의를 바탕으로 영재 선발에 필요한 여러 가지 준거를 탐색하여 영재 판별 원칙을 세우고 이에 기초하여 과학 적성 검사나 과학 사고력 검사 등을 개발하고 이를 활용하는 방식으로 이루어졌다(조성은, 이화국, 2000; 조석희, 1990). 하지만 최근의 영재 선발은

*교신저자: 이기영(leeky@kangwon.ac.kr)
**2007.07.06(접수) 2007.08.28(1심통과) 2007.10.21(2심통과) 2008.01.15(최종통과)

다단계 선발을 통해 이루어지고 있다. 1단계는 여러 가지 기록과 관찰에 의한 교사의 추천을 통해 이루어지는데, 영재성에 대한 교사들의 이해 정도나 관점이 상이할 수 있어 추천에 의해서만 영재 선발이 이루어지는 어렵다(박성익 등, 2003). 교사의 추천은 우수한 잠재 능력을 가지고 있을 가능성이 있는 학생들에게 지필 검사나 문제 해결력 검사에 응시할 대상을 추천받는 단계로 활용되고 있다. 2단계에서는 과학 지식이나 기능을 기반으로 하는 지필 검사 형태의 창의적 사고능력 검사가 실시된다. 최근 과학 영재의 선발 과정에서 창의적 문제 해결력의 중요성이 강조됨에 따라(한기순 등, 2002) 많은 영재 교육 기관에서도 영재 선발 과정에서 지필 검사 형태의 창의적 사고 능력 검사를 시행하고 있다. 과학 지식과 탐구 기능을 기반으로 한 창의성 검사로는 한국 교육 개발원에서 개발한 과학 영역 창의적 문제 해결력 검사나 각급 영재 교육원에서 영재 선발 과정에서 개발하여 사용하는 창의성 검사가 있다(조석희 등, 1997; 배희병 등, 2004). 3단계는 탐구 활동을 기반으로 하는 문제 해결 과정을 전문가가 평가하는 과정을 통해 이루어지고 있다(최선영, 강호갑, 2006; 윤경미, 김정섭, 2006). 탐구 활동을 기반으로 하는 문제 해결 과정은 과학 탐구 활동을 통한 문제 해결 과정에 학생이 직접 참여하여 제시된 문제를 해결하도록 하는 것으로, 과학 영재 학교나 영재 교육원의 학생 선발에서 사용되고 있으며 활동 과정을 통해 개념의 이해와 적용 능력, 탐구 수행 능력, 창의적 문제 발견 및 해결 능력, 과제 집착력을 측정하고 있다(박인호, 2002; 배희병 등, 2004).

많은 예산과 시간, 노력, 전문 인력이 투입되는 문제 해결 과정의 전문가 평가를 통해 최종 영재 교육 대상자를 선발하는 것은 문제 해결 과정에 대한 전문가 평가가 영재성을 판별하는 데 보다 타당하다고 인정되기 때문이다. 그런데 이러한 검사는 시간이나 공간의 제약, 예산 확보의 어려움 등으로 인해 다수의 피험자를 평가할 수 없어 2단계인 지필 검사를 통해 대상자의 수를 크게 줄인 후 시행되고 있다. 따라서 영재 선발 과정에서 지필 검사는 최종적으로 영재 교육 대상자로 선발되는 데 있어서 중요한 역할을 한다고 할 수 있다. 지필검사의 특성에 따라서 영재성을 가진 학생이 최종 단계의 검사 기회를 얻을 수도 있고 그렇지 못할 수도 있기 때문이다.

이제까지 영재 선발의 타당성과 효율성을 높일 수 있는 여러 가지 연구가 이루어져 왔다. 영재 선발 과정에서 영재 선발의 준거가 될 수 있는 과학 창의성 평가

들의 개발(우중옥, 2003; 박종원, 박종석, 2003)과 다중 지능 이론을 영재 선발 과정에 적용하려는 탐색 등의 연구(류지영, 2004)가 이루어졌다. 특히 최근에는 영재와 일반 학생을 대상으로 지능과 같이 유전적 성향이 있는 기질 외에도 후천적 성격을 분석하여 지능이나 사고력 외에도 성격이나 태도 등과 같은 요인을 영재 선발에 도입하려는 연구도 이루어지고 있다(박주용, 오현숙, 2005). 또한, 최근에는 과학 영재와 일반 집단 간의 창의성 비교 연구를 통한 영재와 학업 성취 우수자의 특성을 비교한 연구도 이루어졌는데, 이 연구를 통해 영재 선발 방법이나 선발에 사용하는 문항이 영재로서의 잠재성보다는 학업 성취도 우수 학생을 선발할 가능성이 있음이 제기되고 있다(윤경미, 김정섭, 2006). 그러므로 영재 선발 과정의 타당성이 확보되기 위해서는 영재 선발 방법이나 선발에 사용되는 검사 도구를 구성하는 문항에 대한 체계적인 연구가 필요하다.

이러한 필요성에 비해 영재 선발에 사용된 검사 문항의 분석과 관련된 연구는 충분치 못하다. 영재 교육 기관에서 개발한 영재 선발용 검사 도구의 특성이나 적용 방식 등에 대한 비교 분석은 일부 이루어졌으나 영재 선발 결과가 사회적으로 민감한 사항이어서 문항의 공개가 이루어지지 않고 더욱이 문항의 적용 결과에 대한 체계적인 분석은 일부 영재 교육 센터나 과학 고등학교 선발 문항에 대한 분석, 일부 올림피아드 문항 분석 정도가 이루어져 있는 실정이다(동효관, 전영석, 2004; 이기영, 김찬중, 2005; 이상범, 2001; 이상범 등, 1999).

문항 분석은 문항의 내용 타당도를 알아보는 질적 평가와 피험자의 응답 결과를 분석하여 난이도, 변별도, 신뢰도, 답지 매력도 등을 분석하는 양적 평가로 이루어진다(성태제, 2004). 질적 평가에서 이루어지는 내용 타당도 분석은 검사가 측정하고자 하는 속성을 제대로 측정하는지를 알아보는 것이다. 과학 영재성을 측정하는 주요 준거로 과학 지식, 탐구 능력, 창의력이 제시되고 있으므로(권치순, 2003; 박종원, 박종석, 2003) 영재 교육 대상자가 영재 교육을 받을 만한 적성을 지니고 있는지를 알아보기 위해서는 영재 선발 도구가 갖는 내용 타당도를 분석할 필요가 있다. 양적 평가는 피험자의 응답 결과를 검사 이론에 근거하여 난이도, 변별도, 신뢰도 등을 분석하는 것이다. 일반적으로 영재 선발에는 학업 성취도나 지적 능력이 우수한 학생들이 지원하므로 영재 선발 과정에서 사용되는 검사 도구는 일반 학생들을 대상으로 개발된 검사 도

구보다 난이도를 높여 천정 효과(ceiling effect)가 나타나지 않도록 해야 한다. 천정 효과가 발생할 경우 상위권 학생들을 변별할 수 없기 때문이다. 따라서 우수한 능력을 지닌 학생들을 변별하기 위한 적절한 난이도와 변별도가 확보되어야 하며 검사를 통해 측정하려는 요소가 안정적으로 측정될 수 있는지를 알아보는 신뢰도 또한 분석되어야 한다.

지필 검사 형태의 창의성 검사가 영재 선발 결과에 영향을 미치는 주요 검사 중 하나라는 점과 영재 선발에 사용되는 검사의 문항들이 우리나라의 실정 상 예비 검사를 통해 양호도를 높이는 과정을 거쳐 만들어지지 못하고 선발에 앞서 영재 교육 전문가와 교과 내용 전문가들이 합숙을 통해 보안을 유지하는 상태에서 개발된다는 점을 고려할 때, 영재선발에 사용된 문항에 대한 질적 및 양적 측면에서의 체계적 분석은 영재 선발 문항의 질을 높이는데 반드시 필요하다고 할 수 있다.

이에 본 연구에서는 영재 교육원의 영재 선발에 사용된 지필 검사 문항의 질적 특성과 양적 특성을 살펴봄으로써 타당하고 신뢰로운 영재 선발 도구로서 지필 검사의 적합성을 탐색하는데 필요한 기초 자료를 얻고자 한다.

II. 연구 내용 및 방법

1. 연구 대상

본 연구는 Table 1에서 보는 바와 같이 서울특별시 소재 H 과학 고등학교 부설 영재 교육원 학생 선발에 사용된 2002년부터 2004년까지의 지필 검사(기초 지식 및 탐구 능력 검사, 창의성 검사) 문항을 대상으로 하였다. 과학 고등학교 부설 영재 교육원에서는 중학교 1학년을 대상으로 1차 서류 전형과 2차 지필 검사를 통해 80명의 학생들을 우선 선발한 후, 3차 과제수행 검사를 실시하여 2차와 3차 점수를 합산하여 최종 40명의 학생들을 선발하고 있다.

본 연구에서 분석된 문항들은 2차 지필 검사에 사용된 것이다. 지필 검사지는 수학과 과학(물리, 화학, 생물, 지구과학) 교과목의 문항들로 구성되며, 문항 비율은

수학 40%, 과학 60% 정도이다. 문항 출제는 과학 교과 담당 교사 모두(물리와 화학은 각 5명, 생물과 지구과학은 각 4명)가 각자 기초 지식 및 탐구 능력 2문항, 창의성 1문항씩을 출제하여 각 교과별로 검토 과정을 거쳐 2배수 이상의 예비 문항을 선제하였으며, 이후 각 교과 대표 출제진이 합숙을 통한 집중 작업으로 교과 간 상호 검토, 수정 및 보안을 거쳐 최종 문항을 선정하였다. 본 연구에서는 이 문항들 중에서 수학을 제외한 41개 과학 문항을 분석 대상으로 하였으며, 3년간 영재 선발에 응시한 674명의 문항별 점수 자료를 분석하였다.

2. 질적 문항 분석을 위한 분석틀 구안

본 연구에서는 과학 영재 선발 도구로 사용된 지필 검사 문항을 질적으로 분석하기 위해 2개의 문항 분석틀을 구안하였다. 이 문항 분석틀은 우종욱 등(1996)과 홍미영 등(2002)의 연구와 이기영, 김찬중(2005)의 지구과학 올림피아드 문항 분석 연구에서 사용한 문항 분석틀을 수정하여 구안된 내적 문항 분석틀(internal item analysis framework)과 외적 문항 분석틀(external item analysis framework)로 구분된다. 이 연구에서는 기존의 문항 분석틀과는 달리 외적 문항 분석틀을 추가하였으며, 내적 문항 분석틀 또한 영재 선발 문항에 적합한 차원(dimension)들로 재구성하였다.

내적 문항 분석틀은 문항이 포함하는 내적인 요소들을 분석하기 위한 것으로 Table 2와 같이 내용(content), 지식 및 탐구과정(knowledge & inquiry process), 맥락(context)의 3차원으로 구성하였다. 내용 차원은 물리, 화학, 지구과학, 생물의 내용 영역(domain)에 따라 분류하였으며, 각 영역들을 몇 개의 주제(topic)로 세분하였다(부록 참조). 지식 및 탐구 과정 차원은 4개의 일반적 탐구 과정 영역(문제 인식 및 가설 설정, 탐구 설계 및 수행, 자료 분석 및 해석, 결론 도출 및 일반화)에 지식 영역을 추가하여 하나의 차원으로 만들었으며, 각 영역들은 몇 개의 요소들로 나누었다. 맥락 차원은 문항의 내용이 해당 교육 과정 수준에의 부합 여부와 고차원 사고기능 요구 여부에 따라 일반(general), 심화(enrichment), 그리고 속진(acceleration) 3개의 영

Table 1
Details of study subjects

Type of test	Basic knowledge & Inquiry skills				Creativity				Total	
	Year	2002	2003	2004	Sum	2002	2003	2004		Sum
N of items		10	8	9	27	5	5	4	14	41
N of cases		218	215	241	-	218	215	241	-	674

Table 2
Internal item analysis framework

Dimension	Domain	Details
Content	Physics	Pa. Mechanics Pb. Work & Energy Pc. Electricity & Magnetism Pd. Light & Wave
	Chemistry	Ca. Material Science Cb. State of Matter & Solutions Cc. Atomic structure & Periodic table Cd. Chemical Bond Ce. Chemical Reaction
	Biology	Ba. Cell Bb. Metabolism Bc. Homeostasis Bd. Continuous of Life Be. Diversity of Life & Environment
	Earth science	Ea. Astronomy Eb. Geology Ec. Meteorology Ed. Oceanography Ee. Earth environment
Knowledge & Inquiry process (Basic knowledge & Inquiry skills test only)	I. Knowledge	I 1. Memory I 2. Understanding I 3. Application
	II. Identifying problems & formulating hypothesis	II 1. Identifying problems II 2. Formulating hypothesis
	III. Planning & performing inquiry	III 1. Formulating variable III 2. Devising & arranging apparatus III 3. Observation, measurement, classification, experiment III 4. Experiment procedure
	IV. Analyzing & interpreting data	IV 1. Quantitative analysis IV 2. Qualitative analysis IV 3. Transformation of data
	V. Making conclusion & generalization	V 1. Making conclusion V 2. Generalization
Creative thinking (Creativity test only)	a. Divergent	a1. Fluency a2. Flexibility a3. Unconventional thinking
	b. Convergent	b1. Coherency b2. Synthesis b3. Simplicity
	c. Associative	c1. Analogy c2. Metaphor c3. Abduction
Context	General	Correspond to appropriate curriculum, not require higher order thinking skills
	Enrichment	Correspond to appropriate curriculum, require higher order thinking skills
	Acceleration	Including contents(concepts) of upper curriculum

Table 3
External item analysis framework

Category	Sub-category	Description
Item representation	Picture	Data are presented with photograph or illustration
	Graph	Data are presented with graph
	Table	Data are presented with table
	Example	Options are presented with example
	Text	Item is presented with text only
Item type	Multiple-choice	Choose one or more options
	Short answer	Write answer using one or more words
	Essay	Describe explanation. cause, and procedure
Item form	Single	Single item without sub-items
	Multi-tied	Two or more sub-items are tied

역으로 구분하였다. 일반은 문항의 내용이 응시 대상인 중학교 1학년 교육 과정에 포함되면서 고차원 사고 기능을 요구하지 않는 것이며, 심화는 일반과는 달리 고차원 사고 기능을 요구하는 것이다. 속진은 중학교 1학

년 교육 과정의 이상의 상위 교육 과정을 포함하는 것을 말한다. 내용과 맥락 차원은 2가지 검사 도구 모두에 적용하였으며, 지식 및 탐구 과정 차원은 기초지식 및 탐구 능력 검사의 경우에만 적용하였다. 한편, 창의

성 검사의 경우는 문항의 특성상 지식 및 탐구 과정 차원 대신 창의적 사고(creative thinking) 차원을 추가하여 유형에 따라 발산적(divergent), 수렴적(convergent), 그리고 연관적(associative) 사고 영역으로 구분하였다(박종원, 박종석, 2003).

외적 문항 분석틀은 문항을 구성하는 외적 요소들을 분석하기 위한 것으로 Table 3과 같이 문항 표현(item representation), 문항 유형(item type), 문항 형식(item form)의 3개 범주(category)로 구분하였다. 문항 표현 범주는 무엇을 사용하여 발문을 구성하였느냐에 따라 그림, 그래프, 표, 보기로 구분하고, 이 중 어느 것도 사용하지 않고 오로지 글로만 표현된 문장(text)을 추가하였다. 문항 유형 범주는 피험자들이 답을 하는 방식에 따라 선다형(multiple-choice), 단답형(short answer), 서술형(essay) 유형으로 구분하였다. 문항 형식은 하부 문항의 유무에 따라 단독형(single)과 복합형(multi-tied)으로 구분하였다.

3. 질적 문항 분석 절차 및 방법

질적 문항 분석에 참여한 연구자는 모두 5명으로 영재 교육으로 박사 학위를 가진 전문가 1명, 대학원에서 과학 교육학 또는 내용학으로 박사 학위를 가졌거나 박사 과정 중에 있는 과학 교과 전문가 4명(전공별 1명)으로 구성하였다. 먼저, 영재 교육 전문가 주관 하에 4명의 교과 전문가들이 문헌 연구 및 수차례의 토의 과정을 거쳐 문항 분석틀을 구안하였다. 구안된 분석틀을 각 교과별로 적용하기 전에 분석자간 일치도를 높이기 위해 예비(pilot) 연습 과정을 거쳤다. 분석틀 적용 준거를 4명의 분석자가 공유한 다음, 각 교과의 2002년도 지필 검사 문항을 분석틀에 적용하여 그 결과를 가지고 서로의 의견을 교환하였다. 의견이 다를 경우는 분석틀을 다시 검토하여 이견을 조정하였으며, 분석틀 적용 준거가 애매하다고 판단된 경우는 분석준거에 대한 재합의를 통해 분석틀을 수정하였다.

4. 양적 문항 분석을 위한 난이도, 변별도, 상관계수, 신뢰도 산출

본 연구의 대상 문항 대부분이 부분 점수를 가진 서술형 지필 검사 문항이므로 문항의 난이도(item difficulty: P)는 선다형 문항 난이도 공식을 수정하여 사용하였으며, 문항의 변별도(discrimination index: DI)는 각 문항과 총점과의 상관계수를 구하였다. 난이도는 일반적으로 Cangelosi(1990)의 기준을 적용하여 0.25~

0.75를 적정한 것으로 판단하나, 영재 선발의 경우 일정 효과로 인해 문항의 난이도가 일반 학생의 선발 문항보다 더 높아야 적정하다는 주장(조석희 등, 1997; 배희병 등, 2004; 박성익 등, 2003)이 많기 때문에, 본 연구에서는 이러한 주장을 반영하여 난이도 0.3이내(± 0.05)를 적정한 난이도로 판단하였다. 변별도의 경우는 Ebel(1965)의 기준을 적용하여 0.3이상을 양호한 것으로 판단하였다.

한편, 영재 선발에 사용된 지필검사 도구 및 교과 간 상관을 알아보기 위해 상관계수를 산출하였으며, 문항의 내적 일관성 신뢰도인 Cronbach α 를 산출하였다(Cronbach, 1951).

III. 연구 결과 및 논의

1. 문항 분석틀 적용을 통한 질적 특성 분석

(1) 내적 문항 분석틀 적용 결과

Table 4는 기초 지식 및 탐구 능력 검사와 창의성 검사 문항을 각각 내적 문항 분석틀의 내용 차원에 적용한 결과를 나타낸 것이다.

화학 교과의 출제 비율이 전체 41문항 중 13문항으로 가장 높았고, 생물 교과의 비율이 8문항으로 가장 낮았으며, 4개 교과 모두 특정 영역에 출제가 편중되는 경향을 나타내었다. 이것은 영재 교육원 선발의 특성상 원칙적으로 문항 출제 범위가 중학교 1학년까지로 제한됨으로써 나타난 결과로 해석된다. 속진형 문항은 전체 41문항 중 11문항으로 27% 정도를 차지하는 것으로 나타났다. 기초 지식 및 탐구 기능 검사의 경우 속진형 문항 비율이 전체 27문항 중 9문항(33%)으로 14문항 중 2문항(14%)인 창의성 검사보다 더 높았으며, 교과별로도 상당한 차이가 있는 것으로 나타났다. 중학교 교육과정에서 다루지 않는 속진형 문항이 상당 부분 출제된 것은 내용 타당도에 문제가 있는 것으로 판단된다.

Table 5는 기초 지식 및 탐구 능력 검사 문항을 내적 문항 분석틀의 지식 및 탐구 과정 차원에 적용한 결과를 나타낸 것이다.

지식에 해당되는 문항이 전체 27문항 중 13문항(48%)으로 가장 많은 비중을 차지하였으며, 지식 영역 중에서는 기억이나 이해보다는 적용에 해당되는 문항이 대부분이었다. 탐구과정 영역에서는 14문항 중 11문항(79%)이 자료 분석 및 해석에 해당되는 문항으로 거의 대부분을 차지하고 있었으며, 문제 인식과 가설 설정, 탐구 설계 및 수행과 관련된 문항의 비율은 매우

Table 4
Result of item analysis for content dimension of internal framework

Type of test		Basic knowledge & Inquiry skills				Creativity				Total
Year		2002	2003	2004	Sum	2002	2003	2004	Sum	
Physics	Pa	1	1	1	3	1	1	1	3	6
	Pb									
	Pc									
	Pd	1	1	1	3	1			1	4
Chemistry	Ca			1	1		3		3	4
	Cb	2	2	2	6	1		1	2	8
	Cc									
	Cd									
	Ce	1			1					1
Biology	Ba									
	Bb	2(2)	2(2)	2(2)	6(6)					6(6)
	Bc					1			1	1
	Bd									
	Be							1(1)	1(1)	1(1)
Earth Science	Ea		1	1(1)	2(1)			1	1	3(1)
	Eb	2(1)	1		3(1)	1	1(1)		2(1)	5(2)
	Ec									
	Ed	1		1(1)	2(1)					2(1)
	Ee									
Sum		10(3)	8(2)	9(4)	27(9)	5	5(1)	4(1)	14(2)	41(11)

(): Number of acceleration item. a, b, c, d, e: Domains of each subject

Table 5
Result of item analysis for knowledge & inquiry process dimension of internal framework

Type of test		Basic knowledge & Inquiry skills			
Year		2002	2003	2004	Sum
Knowledge(I)	I 1				
	I 2	1		1	2
	I 3	5	3	3	11
Identifying problems & formulating hypothesis(II)	II 1				
	II 2		1		1
Planning & performing inquiry(III)	III1				
	III2				
	III3			2	2
	III4				
Analyzing & interpreting data(IV)	IV1	2	1	1	4
	IV2	2	1	2	5
	IV3		2		2
Making conclusion & generalization(V)	V 1				
	V 2				
Sum		10	8	9	27

낮았다. 특히, 결론 도출 및 일반화와 관련된 문항은 전혀 출제되지 않은 것으로 나타났다. 또한 자료 분석

및 해석 중에서는 자료의 변환보다는 정성적 및 정량적 해석이 주를 이루고 있는 것으로 나타났다.

이와 같은 결과는 기존의 대입 수능이나 올림피아드 문항 분석 결과와 일치하는 것으로(이기영, 김찬중, 2005; 홍미영 등, 2002), 지필 검사로 쉽게 평가할 수 있는 탐구 과정 영역이 자료 분석 및 해석이기 때문인 것으로 볼 수 있다.

Table 6은 창의성 검사 문항을 내적 문항 분석들의 창의적 사고 차원에 적용한 결과를 나타낸 것이다.

수렴적 사고에 해당되는 문항이 전체 14문항 중 11문항(79%)으로 거의 대부분을 차지하고 있었으며, 발산적 사고나 연관적 사고와 관련된 문항의 비율은 매우 낮았다. 수렴적 사고 중에서는 통합성이나 단순성보다는 정합성에 해당되는 문항이 주를 이루고 있는 것으로 나타났으며, 발산적 사고는 유창성이나 비관습적 사고보다는 주로 융통성을 측정하는 문항이었다.

이와 같이 발산적이거나 연관적 사고보다 수렴적 사고를 측정하는 문항이 대부분을 차지하며, 발산적 사고 중에서도 융통성을 측정하는 문항이 대부분이라는 것은 영재 선발에서 채점의 객관성과 신뢰도를 확보하기 위한 것으로 생각되나, 기초 지식 및 탐구 능력 검사와는 다른 창의성 검사의 고유 목적을 생각해 볼 때 바람직하지 못한 결과인 것으로 해석된다.

Table 6
Result of item analysis for creative thinking dimension of internal framework

Type of test		Creativity			
Year		2002	2003	2004	Sum
Divergent thinking(a)	a1				
	a2	1	1		2
	a3				
Convergent thinking(b)	b1	3	2	1	6
	b2		1	2	3
	b3		1	1	2
Associative thinking(c)	c1				
	c2				
	c3	1			1
Sum		5	5	4	14

Table 7
Result of item analysis for context dimension of internal framework

Type of test	Basic knowledge & Inquiry skills				Creativity				Total	
	Year	2002	2003	2004	Sum	2002	2003	2004		Sum
General		1			1		1		1	2
Enrichment		6	6	5	17	5	3	3	11	28
Acceleration		3	2	4	9		1	1	2	11
Sum		10	8	9	27	5	5	4	14	41

Table 7은 기초 지식 및 탐구 능력 검사와 창의성 검사 문항을 내적 문항 분석들의 맥락 차원에 적용한 결과를 나타낸 것이다.

전체 41문항 중 28문항(68%)이 심화 맥락에 해당되는 문항으로 가장 많은 부분을 차지하는 것으로 나타났으며, 11문항(27%)이 속진 맥락의 문항이었다. 일반 맥락에 해당되는 문항은 전체의 5% 정도로 가장 적은 부분을 차지하였다. 기초 지식 및 탐구 능력 검사와 창의성 검사 모두 심화 맥락의 비율이 각각 63%와 79%로 가장 높았다. 속진 맥락이 차지하는 비율은 기초 지식 및 탐구 능력 검사가 33%로 14%인 창의성 검사에 비해 상대적으로 높았다. 2004년도 기초지식 및 탐구 능력 검사의 경우는 속진형 문항의 비율이 전체 9문항 중 4문항(44%)으로 2002년과 2003년도에 비해 높게 나타나 연도별로도 큰 차이를 나타내었다.

영재 교육원 학생 선발 도구의 목적과 성격으로 보아 중학교 해당 교육과정을 벗어나지 않으면서도 고차원적 사고 기능을 요구하는 심화 맥락의 문항이 가장 많은 부분을 차지하는 매우 긍정적인 현상이라고 볼 수 있으나, 속진 맥락의 문항 비율이 높다는 것은 검사 문항의 내용 타당도를 낮추는 요인이 될 수 있다.

(2) 외적 문항 분석틀 적용 결과

Table 8은 외적 문항 분석틀을 적용한 결과를 나타낸 것이다.

문항 표현에서는 그림의 사용 비율이 가장 높았으며, 그래프의 사용 비율이 가장 낮은 것으로 나타났다. 이것을 앞의 탐구 과정 차원의 분석 결과와 연관 지어 보면, 탐구 과정에서 가장 많은 부분을 차지하는 자료 해석 및 분석이 그래프보다는 그림과 표를 이용하는 것으로 해석된다. 즉, 그림을 이용한 정성적 해석과 표를 이용한 정량적 해석을 요구하는 문항이 대부분이라는 것이다. 그래프는 각종 통계 자료나 변수들 간의 관계를 나타낸 표상(representation)으로, 그래프를 분석하여 그 안에 담긴 의미를 해석하는 능력은 전이력을 가지고 있어 더 높은 수준의 탐구능력인 그래프 작성

Table 8
Result of item analysis for external framework

Type of test		Basic knowledge & Inquiry skills				Creativity				Total
Year		2002	2003	2004	Sum	2002	2003	2004	Sum	
Item representation	Picture	3	5	7	15	4	5	3	12	27
	Graph	1	1		2					2
	Table	4	1	1	6					6
	Example	1	4	4	9	1		1	2	11
	Text	3		1	4	1		1	2	6
Item type	Multiple-choice	1	4	5	9	2		2	4	13
	Short answer	8	1	7	16	2	2	3	7	23
	Essay	2	3		5	2	4		6	11
Item form	Single	5	5	3	13	5	3	1	9	22
	Multi-tied	5	3	6	14		2	3	5	19

능력을 신장시킬 수 있다(Harris, 1999). 그러므로 그래프 해석 능력은 영재들이 갖추어야 할 필수적 탐구 능력 중의 하나이다. 그래프를 이용하여 문항을 제작하는 것이 그림과 표를 이용하는 것보다 어려운 일이긴 하지만 탐구 능력에서 그래프 해석 능력의 중요성을 감안할 때 영재 선발 문항에서 그래프를 이용한 문항이 더 많아져야 할 것이다.

문항 유형에서는 단답형 문항이 23개로 전체의 49% 정도를 차지하였으며, 선다형이 28%를, 서술형이 22%를 차지하는 것으로 나타났다. 기초 지식 및 탐구 능력 검사와 창의성 검사 모두 서술형이나 선다형에 비해 단답형의 비율이 가장 높았으며, 창의성 검사도 기초 지식 및 탐구 능력 검사에 비해 서술형과 선다형의 비율이 상대적으로 높았다. 창의성 검사 문항의 유형에서 선다형과 단답형의 비율이 서술형보다 더 높다는 것은 심각하게 고려되어야 할 부분이다. 창의성 검사에서는 기초 지식 및 탐구 능력 검사로 측정하는 것과는 다른 구인(construct)을 측정하고자 한다. 논리적 사고를 주로 측정하는 기초 지식 및 탐구 능력 검사와는 달리 창의성 검사는 창의적 사고에 기반한 문제 해결력을 측정한다(조석희, 2006). 하지만 창의성을 지필 검사를 통해 타당하기 평가하기란 매우 어려우며 예언도 또한 낮은 것으로 알려져 있다(Cohen & Ambrose, 1993). 지필 검사를 통해 창의성을 보다 타당하게 평가하기 위해서는 선다형과 같이 파편적 정보를 드러내느냐에 초점을 맞추는 것은 적절치 못하며, 해결해야 할 문제 상황에서 어떻게 정보를 조직하고 구조화하며 이를 이용할 수 있는지를 확인해내는 끝이 열린(open-ended) 형태의 서술형이 더 적절할 것이다(Herman et al., 1992).

문항 형식에서는 전체적으로 단독형과 복합형의 비율이 큰 차이가 없었으나, 창의성 검사에서는 단독형의 비율이 복합형에 비해 높게 나타났다. 복합형은 문항을 병렬적 또는 순차적으로 분절화한 것인데, 병렬적인 분절은 단답형으로만 구성된 경우가 가장 많았으며, 순차적 분절의 경우는 단답형에서 서술형으로 이어지는 경우가 대부분이었다.

2. 문항의 양적 특성 분석

(1) 문항 난이도와 변별도 분석

Table 9는 문항 난이도(P)와 변별도(DI)를 검사 도구에 따라 교과별 및 연도별로 비교하여 분석한 결과를 정리한 것이다.

평균 난이도에서는 검사 도구 간 큰 차이를 보이지 않았으며, 영재 선발 문항으로서 대체로 양호한 난이도를 나타내었다. 하지만 난이도 범위에서는 문항에 따라 매우 큰 차이가 있는 것으로 나타났으며, 연도별로도 상당한 차이가 있었다. 특히 2002년 창의성 검사의 경우는 난이도 범위가 0.01~0.90으로 가장 크게 나타나 문항에 따른 난이도 편차가 매우 심한 것으로 분석된다. 교과별로도 난이도에서 많은 차이를 나타내었다. 지구과학 교과와 난이도가 두 검사 도구 모두에서 가장 높아 다른 교과에 비해 어렵게 출제된 것으로 나타났으며, 창의성 검사에서는 생물 교과의 난이도가 유난히 낮아 매우 쉬웠던 것으로 나타났다.

평균 변별도는 두 검사 모두 0.30 이상으로 양호한 것으로 나타났으며, 기초 지식 및 탐구 능력 검사보다 창의성 검사의 변별도가 좀 더 높은 것으로 나타났다. 변별도 범위에서는 문항에 따른 변별력의 편차가 있긴 하지만 크게 문제가 될 정도는 아닌 것으로 판단된다.

Table 9
Summary of 2002-2004 Item difficulty(P) and Discrimination Index(DI) by subject

	Basic knowledge & Inquiry skills			Creativity			
	2002	2003	2004	2002	2003	2004	
Mean of P	Physics	0.47	0.47	0.56	0.40	0.29	0.24
	Chemistry	0.53	0.39	0.43	0.46	0.40	0.20
	Biology	0.48	0.44	0.30	0.90	-	0.82
	Earth science	0.43	0.17	0.25	0.01	0.23	0.14
	Mean	0.48	0.37	0.39	0.43	0.34	0.35
Range of P	0.09~0.84	0.06~0.63	0.08~0.58	0.01~0.90	0.23~0.49	0.14~0.82	
Mean of DI	Physics	0.46	0.39	0.50	0.58	0.45	0.66
	Chemistry	0.41	0.47	0.43	0.59	0.53	0.45
	Biology	0.33	0.49	0.36	0.20	-	0.63
	Earth science	0.37	0.23	0.32	0.17	0.44	0.33
	Mean	0.39	0.39	0.41	0.42	0.50	0.52
Range of DI	0.22~0.58	0.18~0.50	0.21~0.59	0.17~0.59	0.40~0.63	0.33~0.66	

교과별로는 난이도에 비해서는 그 차이가 적으나 지구 과학 교과와 변별도가 다른 교과에 비해 상대적으로 낮은 것으로 나타났다.

(2) 상관계수 및 신뢰도 분석

Table 10은 두 검사 도구에서 4개 교과 간 및 총점과의 상관계수를 산출한 것이다. 두 검사 모두에서 교과 간 상관관이 0.2 이하로 매우 낮거나 거의 없는 것으로 나타났으며, 부적 상관을 나타내는 경우도 있었다. 기초 지식 및 탐구 능력 검사에 비해 창의성 검사가 교

과 간 상관관이 매우 낮게 나타났는데, 이것은 문항 수가 작았기 때문인 것으로 분석된다. 4개 교과와 총점 간의 상관에서는 연도별 및 교과별로 많은 차이가 있었는데, 이것은 교과별로 총점에 기여하는 정도에 차이가 있음을 의미한다. 특히, 지구과학 교과와의 상관관이 다른 교과에 비해 낮게 나타났는데, 이것은 앞의 분석 결과를 토대로 할 때 부정적인 유형의 문항이 가장 많았기 때문인 것으로 판단된다.

Table 11은 두 검사도구간 및 총점과의 상관계수를 산출한 것이다. 두 검사도구간의 상관관은 0.27~0.33으

Table 10
Correlation between four subjects and sum

Year		Basic knowledge & Inquiry skills					Creativity				
		P	C	B	E	Sum	P	C	B	E	Sum
2002	P	1.00	0.19	0.24	0.12	0.64	1.00	0.09	-0.11	-0.04	0.81
	C	0.19	1.00	0.16	0.17	0.70	0.09	1.00	0.09	0.08	0.59
	B	0.24	0.16	1.00	0.09	0.43	-0.11	0.09	1.00	0.04	0.20
	E	0.12	0.17	0.09	1.00	0.61	-0.04	0.08	0.04	1.00	0.17
	Sum	0.64	0.70	0.43	0.61	1.00	0.81	0.59	0.20	0.17	1.00
2003	P	1.00	0.18	0.01	0.01	0.54	1.00	0.17	-	0.08	0.45
	C	0.18	1.00	0.17	0.07	0.64	0.17	1.00	-	0.11	0.90
	B	0.01	0.17	1.00	-0.04	0.64	-	-	-	-	-
	E	0.01	0.07	-0.04	1.00	0.34	0.08	0.11	-	1.00	0.44
	Sum	0.54	0.64	0.64	0.34	1.00	0.45	0.90	-	0.44	1.00
2004	P	1.00	0.24	0.10	0.10	0.66	1.00	0.01	0.07	0.03	0.66
	C	0.24	1.00	0.05	0.10	0.68	0.01	1.00	0.09	0.00	0.45
	B	0.10	0.05	1.00	0.16	0.53	0.07	0.09	1.00	0.08	0.63
	E	0.10	0.10	0.16	1.00	0.44	0.03	0.00	0.08	1.00	0.33
	Sum	0.66	0.68	0.53	0.44	1.00	0.66	0.45	0.63	0.33	1.00

P: physics C: chemistry B: biology E: earth science

로 낮게 나타나 두 검사도구가 측정하고자 하는 구인 (construct)이 중복되지 않음을 확인할 수 있었다. 총점과의 상관은 두 검사도구 0.74~0.88로 모두 높게 산출되었다. 하지만 총점과의 상관에서 두 검사도구간 차이는 없는 것으로 나타나 특정 검사도구가 선발에 더 큰 영향을 주지는 않은 것으로 판단된다.

Table 12는 두 검사도구의 문항 내적일관성 신뢰도인 Cronbach α 를 산출한 것이다. 2002년 창의성 검사를 제외하면 전체적으로 매우 낮은 신뢰도를 나타내었다. 이와 같은 낮은 신뢰도의 1차적인 요인으로 생각할 수 있는 것은 적은 문항 수이다. 일반적으로 문항의 수가 많을수록, 적절한 난이도와 높은 변별도를 가진 양호한 문항이 많을수록 검사 도구의 신뢰도는 높아진다

Table 11
Correlation between two tests and sum

Year	Basic knowledge & Inquiry skills			Sum
	Basic knowledge & Inquiry skills	Creativity		
2002	Basic knowledge & Inquiry skills	1.00	0.28	0.79
	Creativity	0.28	1.00	0.81
	Sum	0.79	0.81	1.00
2003	Basic knowledge & Inquiry skills	1.00	0.27	0.80
	Creativity	0.27	1.00	0.80
	Sum	0.80	0.80	1.00
2004	Basic knowledge & Inquiry skills	1.00	0.33	0.88
	Creativity	0.33	1.00	0.74
	Sum	0.88	0.74	1.00

Table 12
Estimated reliability of two tests

		Basic knowledge & Inquiry skills	
		Basic knowledge & Inquiry skills	Creativity
2002	N of Cases	218	218
	N of Items	10	5
	Cronbach α	0.420	0.688
2003	N of Cases	215	215
	N of Items	8	5
	Cronbach α	0.217	0.267
2004	N of Cases	241	241
	N of Items	9	4
	Cronbach α	0.368	0.163

(성태제, 2002). 하지만 본 연구 대상인 H 과학 고등학교 영재 교육원에서 3년 동안 매년 출제된 문항 수는 기초 지식 및 탐구 능력 검사의 경우 10문항 이하였으며, 창의성 검사의 경우는 5문항 이하였다. 또 하나의 요인으로는 문항 간 난이도의 큰 편차를 들 수 있다. 앞의 분석 결과에서도 나타났듯이 두 검사 도구에서 변별도는 대체적으로 양호한데 비해 너무 쉽거나 또는 너무 어려워 난이도가 적절하지 않은 문항이 많았다. 이와 같이 양호한 문항의 비율이 낮은 것이 신뢰도를 낮추는 결과를 가져온 것으로 분석된다.

3. 문항의 질적 특성과 양적 특성과의 연관성 분석

문항의 양적 특성과 질적 특성을 연관 지어 분석하기 위해 난이도와 변별도를 기준으로 Table 13과 같이 5가지 유형(A~E)으로 분류한 후, 이를 내적 문항 분석들의 3개 차원인 내용, 지식 및 탐구 과정, 그리고 맥락과의 연관성을 각각 분석하였다. A유형의 경우는 가장 바람직한 경우로 간주되며, B와 C유형은 난이도에서는 문제가 있긴 하지만 변별도가 높아 영재 선발 문항으로 긍정적인(positive) 문항으로 판단되며, D와 E유형은 부정적인(negative) 것으로 간주되는 문항이다.

Table 14는 내용 및 맥락 차원과 문항 유형(P & DI type)과의 연관성을 분석한 것이다. 기초 지식 및 탐구 능력 검사의 경우는 3년 동안 출제된 전체 27문항 중 6문항이 난이도와 변별도가 모두 양호한 A유형으로 나타났으며, 창의성 검사의 경우는 3년 동안 출제된 전체 14문항 중 3문항이 A유형에 해당하였다. 두 검사 모두 난이도는 낮지만 변별도가 양호한 B유형이 가장 많았으며, 특히 기초 지식 및 탐구 능력 검사의 경우는 전체 문항의 50% 정도(13/27)가 이 유형에 해당하였

Table 13
Analytic frame for classifying item types by using P & DI

Type of item	Description
A	Item difficulty & Discrimination index are good
B	Item difficulty is low, but Discrimination index is good
C	Item difficulty is high, but Discrimination index is good
D	Item difficulty & Discrimination index are not good
E	Item difficulty is good, but Discrimination index is not good

Table 14

Analysis of relationship between content(subject) and P & DI type, context and P & DI type

	Type of test	Basic knowledge & Inquiry skills												Creativity												Total										
		2002				2003				2004				Sum				2002				2003					2004				Sum					
		A	B	C	D	E	A	B	C	D	E	A	B	C	D	A	B	C	D	E	A	B	C	D	A		B	C	D	E	A	B	C	D		
Content	Physics	1	1			2				2	1	5			1	1			1					1	2	1	1	1	1			1	2	1	1	10
	Chemistry	2				1	1	1		1	1	1	2	4	1	1	1		1	2				1	1	3	1	1	3	1	1	13				
	Biology	1	1			1	1			2		3	2	1					1					1				1	1	1		8				
	Earth science	1	1	1			1	1		1	1	2	1	3	1				1	1				1				2	1			10				
Context	General		1									1							1									1				2				
	Enrichment	1	4			1	1	3		1	1	3	1	3	1	3	10		2	2	1	2		2	2	1		3	3	3	2	28				
	Acceleration	1	1	1		1	1			2	1	1	3	3	1	2					1			1		1		1	1			11				
	Sum	1	5	2	1	1	2	4		1	1	3	4	2	2	6	13	2	4	2	1	2		2	2	2	1	1	3	3	5	41				

A: P & DI are good B: P is low, DI is good C: P is high, DI is good D: P & DI are not good E: P is good, DI is not good

Table 15

Analysis of relationship between knowledge & inquiry process and P & DI type

	Type of test	Basic knowledge & Inquiry skills															Total					
		2002					2003					2004						Sum				
		A	B	C	D	E	A	B	C	D	E	A	B	C	D	E		A	B	C	D	E
Year	I		2	2	1	1					2		1	1	3		1	7	2	1	2	13
	II							1										1				1
	III												1		1		1	1	1	1		2
	IV	1	3					2	1		1		2		1	5	4	2				11
	V																					
Sum	1	5	2	1	1	2	4		1	1	3	4		2	6	13	2	4	2		27	

A: P & DI are good B: P is low, DI is good C: P is high, DI is good

D: P & DI are not good E: P is good, DI is not good

I: Knowledge II: Identifying problems & formulating hypothesis

III: Planning & performing inquiry IV: Analyzing & interpreting data

V: Making conclusion & generalization

다. 창의성 검사의 경우는 난이도가 높아 어렵지만 변별도는 높은 C 유형의 문항 비율이 28%(4/14)로 기초 지식 및 탐구 능력 검사 7%(2/27)에 비해 높게 나타난 것이 특징적이었다. 긍정적으로 판단되는 문항(A, B, C 유형)은 기초 지식 및 탐구 능력 검사가 27문항 중 21문항(78%), 창의성 검사가 14문항 중 12문항(86%)으로 부정적인 문항(D, E 유형)에 비해 매우 우세한 것으로 나타났다.

내용(교과)과의 연관성을 분석한 결과에서는 교과별로 유형별 문항 비율에서 상당한 차이를 나타내었다. 물리 교과와 비교하여 3년 동안 두 검사에 출제된 것을 모두 합친 전체 10문항 모두가 A, B, C 유형의 긍정적인 문항으로 나타나 4개 교과 중 가장 양호한 것으로 분석되었으나, 지구과학 교과와 비교하여 3년 동안 두 검사에 출제된 것을 모두 합친 전체 10문항 중 5문항만이 긍정적인 것으로 나타나 4개 교과 중 가장 양호하지 못한 것으로 분석되었다. 특히, 지구과

학 교과와 비교하여 3년 동안 두 검사에 출제된 것을 모두 합친 전체 13문항 중 2문항을 제외하고 모두 긍정적인 문항이었으며, 가장 적게 출제된 생물 교과와 비교하여 1문항을 제외하고 7문항 모두가 긍정적인 유형에 속하는 것으로 분석되었다.

맥락과의 연관성을 분석한 결과, 세 맥락 중 가장 많은 부분을 차지하는 심화 맥락의 경우는 B 유형의 문항이 가장 많았다. 특히 기초 지식 및 탐구 능력 검사의 심화 맥락 문항의 경우는 60% 정도(10/17)가 이 유형에 해당하였다. 또한 3년 동안 두 검사에 심화 맥락으로 출제된 것을 모두 합친 전체 28문항 중 6문항을 제외한 22문항(78%)이 긍정적인 유형의 문항인 것으로 분석되었다. 숙진 맥락의 경우도 B 유형의 문항이 가장 많았으며(36%), 전체 11문항 중 9문항(82%)이

긍정적인 유형으로 나타나 속진 맥락의 문항이 심화 맥락의 문항에 비해 부정적인 문항의 비율이 더 높을 것이라는 일반적인 예측과 반대로 나타났다.

Table 15는 지식 및 탐구과정과의 연관성을 분석한 결과이다. 지식 및 탐구능력 검사에 3년 동안 가장 많은 부분이 출제된 지식 영역의 경우, B유형이 13문항 중 7문항(54%)으로 가장 많았으며, 10문항(77%)이 긍정적인 문항 유형에 해당하였다. 탐구 과정 중 가장 많은 부분이 출제된 자료 해석 및 분석 영역은 난이도와 변별도가 모두 양호한 A유형의 문항(45%)이 가장 많았으며, 이 영역으로 3년 동안 출제된 11문항 중 9문항(82%)이 긍정적인 유형인 것으로 나타났다. 지식 영역 문항은 쉬우면서 변별력이 있는 문항이 많은 반면, 탐구과정 영역의 문항은 난이도가 적절하면서 변별력이 있는 문항이 많은 것으로 분석되는데, 이것은 지식에 비해 고차원적인 사고 기능이 요구되는 탐구 과정 영역의 특성이 반영된 것으로 생각되며 영재 선발 문항으로서 바람직한 지향점(orientation)인 것으로 판단된다.

IV. 결론 및 제언

현행 대학, 교육청, 그리고 과학 고등학교 부설 과학 영재교육원에서 영재를 판별(identification)하는 주된 방법은 정해진 인원을 다단계 검사를 통해 선발(selection)하는 것이다. 이 때 공통적으로 적용하는 단계가 지필검사를 이용한 선별(screening)이다(김경화, 2005). 이 연구에서는 과학 영재 선발 과정에서 사용된 지필 검사 문항의 질적 및 양적 특성 분석을 통하여 영재 선발 도구로서 지필검사의 적합성을 알아보려고 하였다.

문항의 질적 특성을 분석한 결과에서는 전체의 27% 정도의 문항이 중학교 교육과정을 벗어난 속진 맥락으로 출제되어 내용 타당도에 문제가 있는 것으로 나타났다. 속진 맥락이 차지하는 비율이 훨씬 높다는 것은 교육 기회 균등성(educational equity) 측면에서 생각해 보아야 문제이다. 다수의 연구 결과에서 영재로서의 잠재성은 사회 계층이나 인종 등과는 무관하지만, 영재 교육의 기회는 사회 계층과 매우 밀접하게 관련되어 있음을 밝히고 있다(Hansen & Feldhusen, 1994; Miliband, 2006; Richert *et al.*, 1982; Smutny, 1998). 선행 학습의 효과를 최소화하여 인위적으로 만들어진 영재성이 아닌 타고난 잠재적 영재성을 측정할 수 있는 선발 도구가 되기 위해서는 속진 맥락의 문항 비율을 최소화하는 것이 타당할 것으로 판단된다.

또한 출제 되는 내용이나 맥락에서 교과별 및 연도별로 상당한 차이가 있어 검사 도구의 안정성(stability)에도 문제가 있는 것으로 나타났다. 이와 같은 불안정성을 극복하기 위해서는 출제진들 간에 영재 선발에서 지필 검사로 측정하고자하는 구인에 대한 명확한 합의와 공유가 있어야 할 것이다.

기초 지식 및 탐구 능력 검사의 경우는 지식이 차지하는 비율이 가장 높으며, 탐구 과정 중에서는 자료 해석 및 분석에 치중됨으로써 고른 탐구 능력을 측정하고 있지 못한 것으로 드러났다. 이것은 영재 학생을 선발하는 본연의 목적에 부합하기 위해서는 실기 평가(practical test)를 통해 탐구 과정 영역 중 지필 검사로 평가하기 어려운 부분을 보완할 필요가 있음을 시사한다.

창의성 검사는 대다수의 문항이 발산적 사고나 연관적 사고보다는 수렴적 사고를 주로 측정하는 것으로 나타났다. 창의성 검사는 과학적 내용을 바탕으로 문제를 해결하는 과정에서 확산적 사고와 수렴적 사고 능력을 평가하는 것을 목적으로 한다(Isaksen & Treffinger, 1985). 이러한 창의적 문제 해결에는 확산적 사고와 수렴적 사고 모두 요구된다(김정섭, 2006). 그러므로 창의성 검사가 수렴적 사고를 주로 측정하게 된다면 창의성의 다양한 측면을 고루 평가하지 못하게 됨으로써 일반화가능성(generalizability)이 높은 영재를 선발하기 위한 검사 도구로서의 신뢰도를 떨어뜨리는 결과를 가져올 수 있다.

또한 그래프보다는 그림과 표를 이용하여 문항이 표현된 것으로 나타났는데, 그래프의 전이적 특성을 고려할 때(Harris, 1999) 영재 선발 문항의 외적 요소로서 비중이 커져야 할 것이다. 문항 유형에서는 특히, 창의성 검사에서 서술형보다는 선다형이나 단답형의 비율이 훨씬 높게 나타났다. 창의적 사고는 하나의 정확한 해결을 산출하는 논리적 사고와는 달리 많은 가능성이 있는 답을 산출한다(Hajcak & Garwood, 1993). 그러므로 창의성 검사의 경우는 선다형이나 단답형보다는 개방적인 형태의 서술형이 더 타당할 것이다.

한편, 문항의 양적 특성을 분석한 결과, 변별도는 두 검사 모두 0.30이상으로 양호하였으나, 문항별 난이도 편차는 0.01~0.90으로 매우 큰 것으로 나타났다. 교과 간 및 검사 도구 간 상관은 매우 낮았으며, 신뢰도 또한 낮은 것으로 나타났다. 문항별 난이도 편차는 검사 도구의 신뢰도 저하에 직접적인 영향을 미치게 되는 것이므로 편차를 줄이려는 노력이 있어야 할 것이다. 이 연구 결과에서와 같이 난이도가 너무 낮거나 너무

높은 문항은 출제된 의미가 없는 문항으로 전락할 수 있다.

문항의 양적 특성과 질적 특성과의 연관성에서는 문항의 난이도는 낮아 쉬운 문제였지만 변별도가 양호한 유형의 문항이 가장 많았으며, 속진 맥락의 문항이 심화 맥락에 비해 변별도가 더 양호한 것으로 나타났다. 속진 맥락의 문항에서 변별도가 높은 문항 비율이 높다는 것은 심화 맥락의 문항보다 영재 학생 선발에 더 큰 영향을 미치는 것으로 해석할 수 있다. 즉, 속진 맥락의 문항들이 상위 교육 과정을 선행하여 학습을 한 학생들과 그렇지 않은 학생들을 더욱 잘 변별함으로써 선행 학습을 한 학생들이 선발되는데 유리하게 작용한다고 볼 수 있다.

본 연구는 영재 교육 기관에서 영재성을 가진 학생을 신뢰롭게 선발하기 위해 적합한 지필 검사 도구를 개발하는데 있어 필요한 기초 자료를 제공하였다는 점에서 의미가 있을 것이다. 하지만 특정 영재 교육 기관의 경우만을 대상으로 하였으므로 연구 결과를 모든 경우로 일반화하기에는 무리가 있다. 앞으로 다른 영재 교육 기관의 경우를 대상으로 한 추가적인 연구가 있어야 할 것이다.

국문 요약

이 연구에서는 과학 영재 선발 도구로 사용된 지필 검사(기초 지식 및 탐구 능력 검사, 창의성 검사) 문항의 질적 특성과 양적 특성을 살펴봄으로써 타당하고 신뢰로운 영재 선발 도구로서 지필검사의 적합성을 탐색하는데 필요한 기초자료를 얻고자 하였다. 문항의 질적 특성을 분석하기 위해 내적 및 외적 문항 분석틀을 구안하여 적용하였으며, 문항의 양적 특성과 질적 특성의 연관성을 분석하였다.

연구 결과, 다수의 문항이 중학교 교육 과정을 벗어난 속진 맥락으로 출제되어 내용 타당도에 문제가 있는 것으로 분석되었다. 또한 출제되는 내용이나 맥락에서 교과별 및 연도별로 상당한 차이가 있어 검사 도구의 안정성에도 문제가 있는 것으로 나타났다. 기초 지식 및 탐구 능력 검사의 경우는 지식이 차지하는 비율이 가장 높으며, 탐구 과정 중에서는 자료 해석 및 분석에 치중됨으로써 고른 탐구 능력을 측정하고 있지 못한 것으로 드러났다. 창의성 검사에서는 대다수의 문항이 발산적 사고나 연관적 사고보다는 수렴적 사고를 주로 측정하고 있는 것으로 나타났다. 한편, 분석된 대부분의 문항이 그래프보다는 그림이나 표를 이용하여

표현되는 것으로 나타났으며, 문항 유형에서는 서술형보다는 선다형이나 단답형의 비율이 훨씬 높았다. 문항의 변별도는 두 검사 모두 양호하였으나, 문항별 난이도 편차는 매우 큰 것으로 나타났다. 교과 간 및 검사 도구 간 상관은 매우 낮았으며, 신뢰도 또한 낮은 것으로 나타났다. 또한 난이도는 낮아 쉬운 문제였지만 변별도가 양호한 유형의 문항이 가장 많았으며, 속진 맥락의 문항이 심화 맥락에 비해 변별도가 더 양호한 것으로 나타났다. 이와 같은 분석 결과들은 영재 교육 기관에서 학생 선발에 적합한 지필 검사 도구를 개발하는 데 기초 자료로 활용될 수 있을 것이다.

주요어: 과학 영재, 지필 검사, 문항 분석, 영재 선발

참고 문헌

- 권치순 (2003). 과학영재의 판별과 선발. 창의적 지식 생산자 양성을 위한 영재교육. KEDI 연수교재.
- 김경화 (2005). 뇌파검사를 활용한 과학영재의 판별 및 이에 기반한 선발 체계의 구안. 이화여자대학교 대학원 박사학위논문.
- 김정섭 (2006). 비판적 사고와 창의성 교육: 영재교육의 새로운 방향. 부산광역시영재교육진흥원 영재교육 직무연수 심화과정 교재.
- 동효관, 전영석 (2003). 한성과학고등학교 학생 선발과정의 현황 분석. 영재교육연구, 13(4), 65-92.
- 류지영 (2004). 다중지능이론의 영재교육에의 적용. 교육과정연구, 22(1), 147-169.
- 박성익, 조석희, 김홍원, 이지현, 유여홍, 진석언, 한기순 (2003). 영재교육학 원론. 교육과학사.
- 박인호 (2002). 과학영재 선발방안. 과학영재학교 교육학사운영에 관한 공청회 자료집.
- 박종원, 박종석 (2003). 과학적 창의성 평가를 위한 3차원 틀. 한국과학교육학회 제 43차 동계 학술대회 자료집.
- 박주용, 오현숙 (2005). 기질 및 성격 검사에서 밝혀진 영재기관 영재의 특성이 영재선발에 주는 시사점. 한국심리학회지, 11(1), 143-156.
- 배희병, 김승부, 오원섭, 백용주, 천태호, 박병철, 송인덕, 장기상, 양교석, 조규호, 윤남수, 서봉석, 김인관, 한학수, 연구탁 (2004). 과학영재 발굴·육성사업 결과 보고서. 한국과학재단, 66-70.
- 성태제 (2002). 타당도와 신뢰도. 학지사.
- 성태제 (2004). 문항 제작 및 분석의 이론과 실제. 서울: 학지사.
- 우중옥 (2003). 과학 영재의 평가틀. 한국과학교육학회 제 43차 동계 학술대회 자료집.
- 우중옥, 이항로, 구창현 (1996). 과학 탐구능력 평가 문항 유형 변화에 관한 종단적 연구. 한국과학교육학회

- 지, 16(3), 314-328.
- 윤경미, 김정섭 (2006). 영재판별의 새로운 변인: 문제해결력. *교육심리연구*, 20(3), 587-604.
- 이기영, 김찬중 (2005). 한국 지구과학 올림피아드 문항 분석을 통한 문항의 질 향상 방안. *한국지구과학학회지*, 26(6), 511-523.
- 이상법 (2001). 과학영재 선발 문항 성취도의 지역별 편차에 관한 연구. *한국과학교육학회지*, 21(1), 185-212.
- 이상법, 이광필, 최상돈, 황석근 (1999). 과학영재교육센터 학생선발문항 분석 및 선발방법에 대한 제언. *한국과학교육학회지*, 19(4), 604-621.
- 조석희 (1990). 과학영재 관별도구의 개발 및 타당화 연구. *초등교육연구*, 4, 55-69.
- 조석희 (2001). 창의적 지식 생산을 위한 영재교육. 영재교육담당자 직무연수 자료집. 한국교육개발원.
- 조석희 (2006). 창의적 문제해결력 계발을 위한 영재 교육. 한국교육개발원 제 7기 영재교육 교원 기초 연수 교재.
- 조석희, 시기자, 지은림 (1997). 과학영재관별도구 개발연구(II). *한국교육개발원*, 23-25, 39-41.
- 조성은, 이화국 (2000). 과학영재교육센터의 학생선발에 관한 연구. *전북대학교 과학교육논총*, 25, 25-52.
- 최선영, 강호감 (2006). 초등학교 과학영재학급 학생선발을 위한 과학 창의적 문제해결력 검사도구 개발. *초등과학교육*, 25(1), 27-38.
- 한기순, 배미란, 박인호 (2003). 과학영재들은 어떻게 사고하는가? *한국과학교육학회지*, 23(1), 21-34.
- 홍미영, 전경문, 이범홍, 이양락 (2002). 대학수학능력시험 화학II 문항에 대한 학생들의 응답 분석. *한국과학교육학회지*, 22(1), 204-213.
- Cangelosi, J. S. (1990). Designing tests for evaluating student achievement. Longman.
- Cohen, L. M. & Ambrose, D. (1993). Theories and practices for differentiated education for the gifted and talented. In K.A. Heller, F.J. Monks, & A.H. Passow (Eds.), *International handbook of research and development of giftedness and talent* (pp. 339-363). Oxford, UK: Pergamon.
- Cronbach L. J. (1951). Coefficient alpha and the internal structure of test. *Psychometrika*, 16, 297-334.
- Davis, G. A. & Rimm, S. B. (1998). *Education of the gifted and talented*. MA: Allyn & Bacon.
- Ebel, R. L. (1965). *Measuring educational achievement*. Englewood Cliffs, NJ: Prentice-Hall.
- Hajcak, F. & Garwood, T. (1993). *Expanding Creative Imagination*. West Chester, PA: Human Potential Press.
- Hansen, J. B. & Feldhusen, J. F. (1994). Comparison of trained and untrained teachers of gifted students. *Gifted Child Quarterly*, 38(3) 115-121.
- Harris, R. L. (1999). *Information Graphics: A comprehensive illustrated reference*. Atlanta, GA. Oxford University Press.
- Herman, J. L., Aschbacher, P. R., & Winters, L. (1992). *A practical guide to alternative assessment*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Isaksen, S. G. & Treffinger, D. J. (1985). *Creative problem solving: The basic course*. Buffalo, NY: Bearly Limited.
- Miliband, D. (2006). 'Choice and voice in personalised learning', *Schooling for tomorrow: personalising education*, Paris, OECD.
- Richert, E. S., Alvino, J., & McDonnel, R. (1982). *The national report on identification of gifted and talented youth. Assessment and recommendations for comprehensive identification of gifted and talented youth*. Sewell, NJ: Educational Improvement Centre South.
- Smutny, J. F. (1998). *The young gifted child: Potential and promise, an anthology*. Creskill, NJ: Hampton Press, Inc.

부록: 내적 문항 분석틀의 내용 차원에서 교과별 영역 및 주제

Subjects	Domains	Topics
Physics	Pa. Mechanics	Pa1. Various forces Pa2. Representation and composition of forces Pa3. Explanation of motion Pa4. Newton's law of motion Pa5. Momentum and impulse
	Pb. Work & Energy	Pb1. Work and power Pb2. Mechanical energy Pb3. Conservation of mechanical energy Pd4. Heat energy
	Pc. Electricity & Magnetism	Pc1. Static electricity Pc2. Electric circuits Pc3. Connection of resistances Pc4. Electric energy Pc5. Current and magnetic field Pc6. Electromagnetic induction Pc7. Alternating current and electromagnetic wave
	Pd. Light & Wave	Pd1. Property of wave Pd2. Reflection Pd3. Refraction Pd4. Dispersion and composition of light Pd5. Diffraction and interference Pd6. Wave energy
Chemistry	Ca. Material Science	Ca1. Material Ca2. Atom and Molecules Ca3. Formular weight Ca4. Chemical equation
	Cb. State of Matter & Solutions	Cb1. Three State Cb2. Gas Cb3. Liquid Cb4. Solid Cb5. Solutions
	Cc. Atomic structure & Periodic table	Cc1. Atomic structure Cc2. Periodic las Cc3. Metal complex and nonmetal complex
	Cd. Chemical Bond	Cd1. Class of chemical bond Cd2. Covalent bond and molecules Cd3. Intermolecular force Cd4. Organic compound
	Ce. Chemical Reaction	Ce1. Chemical reaction and energy Ce2. Reaction rate Ce3. Chemical equilibrium Ce4. Reaction of acid and base Ce5. Oxidation-reduction reaction
Biology	Ba. Cell	Ba1. Characteristic of Life Ba2. Structure and Function of Cell
	Bb. Metabolism	Bb1. Enzyme Bb2. Photosynthesis Bb3. Respiration Bb4. Digestion Bb5. Circulation Bb6. Excretion
	Bc. Homeostasis	Bc1. Sense Organ Bc2. Nervous System Bc3. Hormone Bc4. Animal Behavior
	Bd. Continuous of Life	Bd1. Reproduction Bd2. Genetic Bd3. Evolution
	Be. Diversity of Life & Environment	Be1. Texonomy Be2. Organism and Environment
Earth science	Ea. Astronomy	Ea1. Solar system Ea2. Motion of planets Ea3. Properties of stars Ea4. Galaxy & universe
	Eb. Geology	Eb1. Earth's structure Eb2. Crust forming materials Eb3. Plate tectonics Eb4. Earth's history Eb5. Field trip & geologic map
	Ec. Meteorology	Ec1. Weather change Ec2. Water in atmosphere Ec3. Atmospheric motion Ec4. Atmospheric circulation
	Ed. Oceanography	Ed1. Topography of ocean bottom Ed2. Properties of seawater Ed3. Sea current Ed4. Ocean wave Ed5. Tide
	Ee. Earth environment	Ee1. Greenhouse effect(global warming) Ee2. El Nino Ee3. Water cycle