

## 고등학교 수리영역 시험의 난이도 예측 요인 분석

고호경<sup>1)</sup> · 이현숙<sup>2)</sup>

본고는 5년간의 고등학교 연합학력평가자료 분석을 통하여 수리영역의 난이도 예측요인을 분석하였다. 난이도 예측을 위한 통계적 모형을 산출하기 위하여 먼저 문항 분석을 통해 수리영역의 난이도 예측에 영향을 미칠 것으로 판단되는 주요 변인들을 '내용영역', '행동영역', '문항의 형식' 등의 범주에 따라 추출하였다. 추출된 독립 변인들에 대하여 단계선택방법을 사용한 다중회귀분석을 실시함으로써 정답률 예측에 유의미한 변수들을 선택하였으며, 교차 타당도를 통하여 최종적으로 선택된 예측 모형이 독립적으로 수집된 자료에 대하여 어느 정도의 설명력을 보이는지 검증하였다. 본 연구는 대학수능시험 출제나 현장에서 수리영역의 평가문항을 개발하는데 있어서 사전 정답률을 예측하는 데 있어 고려해야 할 요인을 제시함으로써 보다 정확한 정답률 예측에 필요한 기초정보를 제공하는데 그 의의를 두고 있다.

주요용어: 대학수학능력시험, 수리영역의 평가, 난이도, 난이도 요인

### I. 수리영역에서 난이도 예측의 필요성 및 목적

교육인적자원부가 발표한 「학교교육 정상화를 위한 2008학년도 이후 대학입학제도 개선안」(이하 2008 대입 개선안, 2004.10.28)에 따라 금년부터 대학수학능력시험(이하 수능)의 점수 체제가 등급제로 변화하게 된다. 즉, 2005학년도 이후 영역/과목별 표준점수와 그에 근거한 백분위 및 등급을 제공하던 점수 체제에서 영역/과목별 등급만 제공하는 점수 체제로 전환되는 것이다. 이러한 변화의 취지는 대학 입시에서 수능의 비중을 낮추고 학교생활기록부의 반영 비중을 늘려 고등학교 교육의 정상화를 기하겠다는 정책 하에 고려된 사항이다.

2008학년도의 수능은 9등급제로의 전환이라는 점을 제외하면, 시험 체제나 출제체제에서의 변화는 없을 것으로 보고 있다. 그런데 이러한 수능 체제하에서는 우선적으로 영역/과목별로 9개 등급이 본래 산정된 백분율에 따라 산출되는 것이 중요하므로, 안정적인 등급 산출을 위한 다각도의 연구가 진행 되어야 한다. 그 중에서도 문항의 난이도를 고르게 분포시킴으로써 다양한 성취 수준의 학생들을 변별할 수 있는 문항을 출제하는 것은 수능 문항 출제의 성패를 좌우하는 가장 중요한 요인 중 하나이다. 사전에 예비검사 없이 시험의 난이도, 특히 문항의 난이도를 정확하게 예측하고 조정하는 일은 매우 어려운 일이나 수능시험은 사회적 민감성이 매우 큰 시험이니만큼 출제과정에서 난이도를 가능한 한 일관되게 유지하기

1) 한국교육과정평가원 (koho@kice.re.kr)  
2) 한국교육과정평가원 (hyunysi@kice.re.kr)

위한 지속적인 노력은 필수적일 수밖에 없다.

영역/과목별로 9개의 등급을 안정적으로 산출하기 위해서는 출제 과정에서 검사 및 문항의 난이도 예측에 도움이 되는 정보를 세밀하게 분석·제공할 필요가 있다. 그러나 현실적으로 문항 제작자들이 내리는 예상 정답률과 실제 학생들이 보이는 정답률에는 많은 차이가 있다. 또한 정답률을 정확히 예상하는 것에는 많은 어려움이 있을 뿐 아니라 이에 대한 연구도 많지 않은 실정이다. 고부담시험(high-stakes test)인 수능이 우리 교육 현장과 사회에 미치는 영향을 조명해볼 때, 매년 적절하고 안정적인 난이도 수준 유지가 필수적일 뿐 아니라, 고부담시험이 피험자에게 요구되어지는 능력을 정확하고 신뢰성 있게 측정되도록 개발될 때 교육사회에 바람직한 공헌을 할 수 있다고 볼 수 있다(Miller, 2003).

이에 따라서, 이미 다른 영역에서는 문항 난이도 관련 변인에 관한 연구가 활발히 진행되고 있는 상황이며(예, Freedle과 Kostine, 1991, 1993; Scheuneman, Gerritz & Embretson, 1991; Alderson, 2000; Baumann 과 Serra, 1984; 장경숙, 2005), 이에 따른 학생들의 시험전략을 질적으로 분석한 연구(예, Lee, 2003)들도 있어서 문항 난이도를 설명하고 예측하는 데 중요한 자료들을 제시하고 있다. 수리영역의 연구에서도 1999년 이후의 수능 문항을 내용 영역과 행동 영역으로 나누어 정답률을 비교해 본 결과 내용 영역은 수학적 기초, 대수, 기하, 해석, 확률과 통계 순으로 정답률이 높았고, 행동 영역의 정답률은 계산, 간단한 이해, 외적 문제해결, 증명, 발견적 추론, 복잡한 이해, 내적 문제해결의 순이었다(이종승 외, 2003). 그러나 이러한 분석은 내용 영역에 있어서 5개의 큰 범주로 구분이 되어 있어서 상세한 정보를 얻는 것은 어렵다고 볼 수 있다. 뿐만 아니라 내용 영역과 행동 영역이 각각 독립 상태에서의 정답률을 집계함으로써 인제, 요인들 간의 종합적인 분석과 다른 관련된 요인과의 관련성 등은 고려되어 있지 않았다.

이에 본 연구는 난이도에 영향을 미치는 객관적인 요인들을 추출해내어 수능 문항 출제에 유용한 정보를 제공함으로써 문항 개발 또는 검사지 구성을 위한 문항 추출 등을 원활하게 하는 데 목적을 두어 다음과 같은 연구를 수행하고자 한다.

2002학년도 이후 수능의 모의 평가 형태인 연합학력고사 자료를 토대로 정답률에 영향을 미치는 요인을 크게 세 가지 영역 즉, '내용 영역', '행동 영역', '문항 유형' 하에 각각 요인들을 세분화하여 난이도에 미치는 요인들 추출하고 상관관계를 분석한다. 또한 각 요인들을 독립변인으로 한 다중회귀분석을 실시함으로써 난이도 예측을 위한 통계적인 모형을 산출하고 이 모형의 타당성 및 적용 가능성에 대해 검토하고자 한다.

## II. 연구방법

### 1. 분석 자료

본 연구는 2002년부터 2006년에 걸쳐 실시된 전국연합학력평가(이하 연합평가) 수리 영역의 결과에 대해 연도별로 1개 혹은 2개의 응시 자료를 가지고 분석하였다. 모든 자료는 한국교육과정평가원에서 분석한 문항과 정답률을 계열별로 구분하여 사용하였으며, 각 연합평가에 응시한 학생 수 현황과 사용한 자료의 연도는 아래 표에 제시하였다.

고등학교 수리영역 시험의 난이도 예측 요인 분석

<표 1> 고등학교 3학년 연합평가(한국교육과정평가원, 2003~2006) 수리영역 응시 인원 (단위: 명)

실시년도 계열	2002/06		2003/10		2004/10		2005/10		2006/10	
	인문계	자연계	인문계	자연계	가형	나형	가형	나형	가형	나형
고 3	248,404	150,575	232,471	133,520	94,617	261,382	81,023	283,506	162,462	304,120

<표 2> 고등학교 2학년 연합평가(한국교육과정평가원, 2003~2006) 수리영역 응시 인원 (단위: 명)

실시년도 계열	2002/06		2003/06		2004/06		2005/06		2006/06	
	인문계	자연계	가형	나형	가형	나형	가형	나형	가형	나형
고 2	248,404	450,575	163,298	300,749	158,561	309,400	157,912	309,635	162,462	304,120

<표 3> 고등학교 1학년 연합평가(한국교육과정평가원, 2003~2006) 수리영역 응시 인원 (단위: 명)

실시년도	2002/06	2002/11	2003/06	2003/11	2004/06	2004/09	2005/06	2006/06	2006/09
고 1	466,467	446,806	478,558	381,162	487,812	489,709	500,680	497,154	479,001

2. 문항 분석 및 독립 변인 추출

정답률에 영향을 미치는 주요 독립 변인들을 추출하기 위하여 우선 2002년부터 2006년까지 고등학교 1, 2, 3학년 연합평가에서 나타난 수리 영역의 문항 분석표에 근거하여 문항의 내용 영역, 행동 영역, 그리고 문항의 형식 등의 범주에 따라 세부 항목을 분류하였다. 내용 영역 분류로는, 1학년은 고등학교 전 범위를 내용적 유사성에 따라 7개 영역으로, 2학년은 8개 영역, 3학년은 16개 영역으로 나누어 분석하였다. 행동 영역은 계산, 이해, 추론, 문제해결의 4개 영역으로 구분하고, 다시 추론 영역을 귀납, 유추, 추측을 포함한 발견적 추론과 증명을 다루는 연역적 추론으로 세분하였으며, 문제해결 영역은 수학 내적 문제해결과 수학 외적 문제해결로 세분하여 모두 6개의 행동 영역으로 분류하였다. 또한 본 고 연구자들은 문항의 형식에 따라 정답률에 영향을 미칠 것이란 가정 하에, 문항의 형식을 정답형, 합답형, 부정형, 완성형, 단답형으로 나누어 분석을 시도하였다. 최종적으로 추출된 독립 변인은 <표 4>에 제시되어 있다.

<표 4> 학년별 최종 추출된 독립변인

학년	독립변수		
	내용 영역	행동 영역	문항 유형

고 1	수학의 기초, 수와 식, 도형의 방정식, 통계, 평면도형, 함수, 삼각함수	계산, 이해, 발견적 추론, 연역적 추론, 내적문제해결, 외적문제해결	정답형, 합답형, 부정형, 완성형, 단답형
고 2 이과/가형	수학의 기초, 수와 식, 도형의 방정식, 행렬 연산 및 역행렬, 지수 로그, 수열, 수학적 귀납법, 확률	계산, 이해, 발견적 추론, 연역적 추론, 내적문제해결, 외적문제해결	정답형, 합답형, 부정형, 완성형, 단답형
고 2 문과/나형	수학의 기초, 수와 식, 도형의 방정식, 행렬 연산 및 역행렬, 지수 로그, 수열, 수학적 귀납법, 확률	계산, 이해, 발견적 추론, 연역적 추론, 내적문제해결, 외적문제해결	정답형, 합답형, 부정형, 완성형, 단답형
고 3 이과/가형	수와 식, 도형의 방정식, 행렬 연산 및 역행렬, 지수 로그, 지수 로그 함수, 수열, 수학적 귀납법, 무한 수열의 극한, 확률, 통계, 함수의 극한과 연속성, 이차 곡선, 공간좌표와 벡터, 삼각함수의 덧셈정리, 삼각함수와 함수의 극한, 미분법과 적분법	계산, 이해, 발견적 추론, 연역적 추론, 내적문제해결, 외적문제해결	정답형, 합답형, 부정형, 완성형, 단답형
고 3 문과/나형	수와 식, 도형의 방정식, 행렬 연산 및 역행렬, 지수 로그, 지수 로그 함수, 수열, 수학적 귀납법, 무한 수열의 극한, 확률, 통계, 함수의 극한과 연속성, 이차 곡선	계산, 이해, 발견적 추론, 연역적 추론, 내적문제해결, 외적문제해결	정답형, 합답형, 부정형, 완성형, 단답형

### 3. 상관 및 회귀 분석

각 문항에 대한 내용 및 행동 영역, 문항 유형 등의 분석을 통해 최종적으로 선정된 독립 변인 각각에 대하여, 고전검사이론에서 난이도의 지표로 흔히 사용되고 있는 정답률과의 이 변량 상관 분석을 실시하였다. 상관 및 회귀 분석은 학년별로 구분하여 실시하였으며, 고등학교 2학년과 3학년의 경우 수리 가형과 수리 나형을 선택하는 응시자 집단의 특성에 있어 매우 차이가 있을 것으로 예상되어 두 유형 간에 동일 문항이 포함되어 있음에도 불구하고 두 집단을 구분하여 분석을 실시하였다. 각 집단 내에서 연도별로도 응시자 집단에 있어 특성이 다소 차이를 보일 것으로 예상되나, 현재의 수능에서 문항에 대한 난이도를 예측할 때 연도별 응시자 집단의 특성 차이를 고려하고 있지 않음을 감안하여 2002년도부터 2006년도 까지 수집된 모든 자료를 통합하여 분석을 실시하였다. 이는 이종승 외(2003)의 연구 결과에서도 뒷받침되는데, 수능 자료를 이용하여 수리 영역의 문항 난이도에 관한 추정 모형을 개발하는 이 연구에서 '통합회귀방정식모형'이 안정성과 예측의 정확성 측면에서 가장 우수한 모형이라고 한 것과 일관된다. 여기서 '통합'의 의미는 2002년과 2003년 자료를 묶었음(통합)을 의미하는 바로 연도 구분 없이 자료를 통합하여 정답률에 대한 예측 회귀 분석을 실시했다는 의미이다. 따라서 본 연구에서는 2002년부터 2006년까지 연합평가에서 사용된 문항을

동시에 사용하여 정답률에 대한 통합회귀방정식모형을 산출하였다. 다시 말해, 각각의 문항에 대한 정답률에 대한 예측 변인으로서 문항 분석을 통하여 추출된 독립 변인에 대해 예측 모형을 설정하기 위하여 다중 선형 회귀 분석을 실시하였다.

수능에 출제되는 문항들은 대개 2개 이상의 내용 및 행동 영역을 복합적으로 측정하는 경우가 많으므로 각 문항에 대하여 해당되는 주요 독립 변인에 모두 표기하는 방식으로 자료를 부호화하였다. 상관 계수의 통계적 유의미성을 검증하기 위하여 유의수준 0.05에서 t-검정을 실시하였는데, 상관의 방향에 대한 사전 정보를 고려하지 않았기 때문에 양측 검정을 적용하였다. 정답률에 대한 예측 모형을 설정하기 위해 실시한 다중 선형 회귀 분석에서는 정답률을 종속 변인으로 하여 단계 선택(stepwise selection) 방법을 사용하여 유의미한 변인을 추출하였다. 즉 독립 변인 중 종속 변인과 상관이 높은 독립 변인부터 예측 모형에 순차적으로 포함시키되, 동시에 이미 모형에 포함된 독립 변인들 중 종속 변인에 대한 설명력이 약한 변인을 단계별로 제거해 나가는 방법을 사용하여 예측 모형에 포함시킬 독립 변인들을 추출하였다. 독립 변인이 하나씩 추가될 때마다 종속 변인에 대한 설명력은 당연히 증가하게 되어 있지만 모형의 간명성 원리(principle of parsimony)를 고려하여 설명력의 증가폭이 더 이상 통계적으로 유의미하지 않는 단계에서 최종 모형을 선택하였다. 마지막으로, 다중 선형 회귀 분석을 통해 최종적으로 설정된 예측 모형은 분석에 이용된 자료에 최적화된 모형이므로, 이 모형을 독립적으로 수집된 다른 자료에 적용할 경우 설명력이 저하되는 정도를 산출하여 교차 타당도를 검증하였다.

### III. 연구 결과

#### 1. 상관 분석 결과

<표 5> 최종 선정된 독립 변인에 대한 이원상관분석 결과. [( ) 안의 수는 문항 수]

독립변인		고3이과/가형		고3문과/나형		고2이과/가형		고2문과/나형		고1	
		r (N)	p	r (N)	p	r (N)	p	r (N)	p	r (N)	p
내용영역	수학의 기초					-0.097 (23)	0.236	-0.085 (23)	0.301	-0.043 (52)	0.484
	수와식	0.067 (15)	0.419	0.07 (16)	0.394	0.143 (26)	0.081	0.186 (28)	0.023*	0.126 (133)	0.038*
	도형의 방정식	-0.171 (24)	0.036*	-0.185 (24)	0.023*	-0.23 (37)	0.005*	-0.204 (43)	0.012*	-0.084 (45)	0.167
	행렬 연산 및 역행렬	0.312 (7)	0.000*	0.23 (13)	0.005*	0.173 (27)	0.034*	0.212 (7)	0.009*		
	지수 로그	0.225 (7)	0.006*	0.151 (15)	0.066	0.086 (29)	0.296	-0.015 (35)	0.858		
	지수 로그 함수	0.022 (8)	0.790	-0.068 (13)	0.405						
	수열	0.04 (6)	0.623	0.082 (9)	0.320	-0.013 (2)	0.874	0.12 (1)	0.143		

	수학적 귀납법	0.023 (3)	0.784	-0.042 (5)	0.608	-0.129 (3)	0.116	-0.013 (3)	0.876		
	무한수열의 극한	-0.084 (5)	0.305	0.077 (13)	0.347						
	확률	-0.147 (9)	0.073	-0.136 (24)	0.098	0.002 (2)	0.980	0.047 (2)	0.570		
	통계	0.111 (5)	0.176	-0.057 (10)	0.490					0.039 (9)	0.518
	함수의극한과 연속성	0.061 (21)	0.458	-0.153 (9)	0.062						
	이차 곡선	-0.125 (13)	0.129	0.062 (1)	0.448						
	공간좌표와 벡터	-0.048 (10)	0.556								
	삼각함수의 덧셈정리	0.027 (3)	0.742								
	삼각함수와 함수의 극한	-0.059 (6)	0.475								
	미분법과 적분법	-0.086 (10)	0.298								
	평면도형									-0.14 (22)	0.021*
	여러 가지 함수									-0.088 (14)	0.146
	삼각함수									-0.109 (4)	0.073
행 동 영 역	계산	0.428 (26)	0.000*	0.532 (23)	0.000*	0.393 (31)	0.000*	0.24 (29)	0.003*	0.26 (53)	0.000*
	이해	0.097 (42)	0.237	-0.008 (49)	0.918	-0.086 (49)	0.295	0.072 (52)	0.384	0.118 (81)	0.053
	발견적 추론	-0.101 (18)	0.217	-0.079 (18)	0.339	-0.086 (18)	0.294	-0.098 (17)	0.232	0.007 (22)	0.905
	연역적 추론	0.078 (9)	0.346	0.131 (9)	0.111	0.14 (10)	0.088	0.125 (11)	0.129	-0.015 (22)	0.804
	내적문제해결	-0.285 (31)	0.000*	-0.3 (29)	0.000*	-0.278 (22)	0.001*	-0.316 (12)	0.000*	-0.222 (52)	0.000*
	외적문제해결	-0.206 (21)	0.011*	-0.203 (21)	0.013*	-0.08 (20)	0.332	-0.055 (19)	0.500	-0.178 (43)	0.003*
문 항	정답형	0.221 (84)	0.006*	0.285 (86)	0.000*	0.349 (81)	0.000*	0.409 (83)	0.000*	0.426 (148)	0.000*
	합답형	0.03 (14)	0.716	0.071 (12)	0.387	-0.106 (16)	0.196	-0.014 (14)	0.863	-0.074 (28)	0.225

고등학교 수리영역 시험의 난이도 예측 요인 분석

유형	부정형					-0.101 (1)	0.220			0.086 (2)	0.156
	완성형	0.08 (12)	0.333	0.166 (11)	0.043*	0.082 (12)	0.319	0.11 (13)	0.18	0 (20)	0.995
	단답형	-0.333 (36)	0.000*	-0.435 (39)	0.000*	-0.351 (40)	0.000*	-0.521 (40)	0.000*	-0.443 (74)	0.000*

2002년부터 2006년까지 실시된 연합평가 중에서 표집된 수리 영역의 870개의 문항에 대해 학년별, 시험 유형별로 집단을 구분하여 상관 분석을 실시한 결과는 <표 5>와 같다. 집단별로 학년별로 최종 선정된 독립 변인에 있어서 다소 차이를 보이고 있지만 두 집단 이상에 공통적으로 포함된 독립 변인에 대한 결과를 살펴보면, 우선 내용 영역에 있어 '수와 식'과 '행렬 연산 및 역행렬'은 정답률과 정적 상관을 보이고 있고, '수학적 기초'와 '도형의 방정식'은 부적 상관을 보이는 공통된 특징을 발견할 수 있다. 정답률과 정적 상관을 보이는 경우, 어떤 문항이 해당 독립 변인을 측정하고 있는 경우 대체로 높은 정답률을 보인다고 볼 수 있으며, 부적 상관을 보이는 경우는 해당 독립 변인이 포함된 경우 대체로 낮은 정답률을 보이는 것으로 해석할 수 있다. 내용 영역의 다른 공통 변인들에 대해서는 일부 집단에서는 정적 상관을, 다른 집단에서는 부적 상관을 보여 특정한 패턴을 나타내지 않았다. 행동 영역의 6개 변인 중에는 '계산'이 정답률과 정적 상관을 나타내었고, '내적문제해결'과 '외적문제해결'은 부적 상관을 나타내었다. 이것은 기존에 조사된 수능에 대한 정답률과 크게 차이를 보이지 않는 결과이다. 문항의 형식 변인 중에서는 '정답형'과 '완성형'의 경우 정답률과 정적 상관을 보였고, '단답형'의 경우 부적 상관을 나타내었다.

유의 수준 0.05에서 유의미한 상관을 나타낸 독립 변인을 집단별로 살펴보면, 고등학교 1학년의 경우 '수와 식', '계산', '정답형' 등의 변인에서 유의미한 정적 상관을 나타내었고, '도형', '내적문제해결', '외적문제해결', '단답형' 등의 변인에서 부적 상관을 나타내었다. 2학년의 경우 수리 가형과 나형 집단에서 공통적으로 '행렬 연산 및 역행렬'과 '계산', '정답형' 등의 변인에서 정적 상관을 보였고, '도형의 방정식', '내적문제해결', '단답형' 등에서 부적 상관을 보였다. 2학년 나형 집단의 경우 '수와 식'에서 정답률과 정적 상관을 나타내었다. 3학년의 경우 두 집단에서 공통적으로 '행렬 연산 및 역행렬'과 '계산', '정답형' 등의 변인에서 정적 상관을 보였고, '도형의 방정식', '내적문제해결', '외적문제해결', '단답형' 등에서 부적 상관을 보였다. 3학년 가형 집단의 경우 이에 더하여 '지수와 로그'에서 정적 상관을 보였고, 나형 집단의 경우 '완성형'에서 정적 상관을 나타내는 차이가 있었다.

내용 및 행동 영역, 그리고 문항의 형식 등의 항목별 독립 변인의 수와 유의미한 상관을 보이는 독립 변인의 수를 비교해보면, 내용 영역을 구성하는 독립 변인에 비해 행동 영역 및 문항의 형식을 구성하는 변인들이 정답률과 보다 큰 상관을 보이는 것을 볼 수 있다. 다시 말하면, 어떤 내용으로 문항을 구성하는가와 더불어 그 문항에서 측정하고자 하는 행동 영역 역시 문항의 난이도를 결정하는 주요한 요인이 된다는 것이다. 가령 어떤 문항이 학생의 '계산'능력을 측정하고자 한다면 '내적 문제해결력'이나 '외적 문제해결력'을 묻는 문항보다는 정답률이 높을 것으로 기대한다는 의미가 된다. 또한 같은 내용의 문항에 대하여 '내적·외적 문제 해결력'을 묻는 문항에 대하여 단답형으로 묻는다면 정답형이나 완성형으로 묻는 문항에 비해 정답률이 낮아질 것으로 기대할 수 있다.

한편, 상관 분석의 결과로 구한 상관 계수의 수치가 전반적으로 크지 않음을 볼 수 있었다. 정답률과 유의미한 상관을 보인 36개의 상관 계수 중에서 .30이하는 23개, .30과 .50 사이는 11개, .50이상은 2개로 전반적으로 높지 않은 편이나 이전의 선행 연구 결과(이종승 외, 2003; 장경숙, 2005; Kostin, 2004; Jin & Park, 2005)와 비교할 때 다소 높은 편이다. 난이도 예측에 관한 연구에서 대체로 상관 계수가 낮게 나타나는 것은 Kostin(2004)이 지적한 바와 같이 여러 개의 변인들이 동시에 복합적으로 난이도를 결정하기 때문으로 해석될 수 있으며, 선행 연구에 비해 본 연구 결과에서 나타난 상관 계수가 대체로 높은 것은 비교적 많은 수의 문항으로 분석함으로써 기인한 자료 분석의 안정성 때문인 것으로 유추된다.

## 2. 회귀 분석 결과

<표 6> 단계 선택 방법을 통한 다중선형 회귀분석

		최종 회귀모형에 포함된 독립 변인	자유도	F	p	R <sup>2</sup>
고1		단답형, 계산, 이해, 합답형, 내적문제해결	5/266	29.171	.000	.354
고2	가형	계산, 단답형, 수학 10-나, 합답형	4/145	20.314	.000	.359
	나형	단답형, 계산, 내적문제해결, 행렬, 수학 수와식, 발견적 추론	6/143	21.891	.000	.479
고3	가형	계산, 단답형, 이해, 행렬, 연역적 추론, 확률	6/143	14.616	.000	.380
	나형	계산, 단답형, 완성형, 행렬, 수열, 이해, 도형의 방정식, 함수의 극한과 연속성	8/141	23.179	.000	.568

단계 선택 방법을 통한 다중선형 회귀분석 결과는 <표 6>에 제시하였다. 집단별로 예측 모형에 포함된 독립 변인의 수와 종류, 그리고 종속 변인에 대한 설명력에 있어서 다소 차이를 보였다. 설명력에 대한 대표적인 지수로 사용되는 R2는 고등학교 1학년 예측 모형에서 .354로 가장 낮은 값을 나타냈고, 3학년 수리 나형 집단에서 .568으로 가장 높은 값을 보였다. 독립 변인의 경우, ‘계산’과 ‘단답형’이 모든 집단에 대해 공통적으로 가장 설명력이 높은 변인으로 선정되었고 내용 영역의 경우에는 집단별로 다소 차이를 보였다.

고등학교 1학년에 대해 실시한 회귀 분석 결과를 살펴보면, 종속 변인에 대한 설명력이 가장 큰 변수는 ‘단답형’이고, ‘계산’, ‘이해’, ‘합답형’, 그리고 ‘내적문제해결’ 등의 순으로 모형에 포함된 바, 난이도에 가장 큰 영향을 미치는 요인들이라 할 수 있을 것이다. 상관 분석 결과 ‘수와 식’과 ‘도형’이 내용 영역을 구성하는 변인들 중에서 정답률과 가장 높은 상관을 보이는 변인으로 나타났으나 회귀 모형에는 포함되지 않았다. 이는 이 두 변인이 이미 모형에 포함된 행동 영역과 문항의 형식 변인들의 일부와 매우 높은 상관을 보이기 때문으로 판단된다. 최종적으로 선택된 예측 모형에 대한 F 검정 결과는 F(5, 266)=29.171로 유의 수준 .001에서 유의미하였다. R2는 .354로 다소 낮은 수치를 보이는데, 이는 모형에 포함된 다섯 개의 변인이 종속 변인에 대한 전체 변산 중 약 35.4% 정도를 설명할 수 있는 것으로 해석할 수 있다.

고등학교 2학년 수리 가형 집단의 경우, ‘계산’, ‘단답형’, ‘도형의 방정식’, ‘합답형’ 등의 순



고등학교 수리영역 시험의 난이도 예측 요인 분석

으로 모형에 포함되었다. ‘합답형’의 경우, 고등학교 1학년의 분석 결과와 마찬가지로, 상관 분석에서는 유의미한 결과를 나타내지 않았으나 회귀 모형에는 포함되었다. 따라서 합답형 역시 난이도에 미치는 영향이 무관하다고 할 수 없을 것이다. 최종 예측 모형에 대한 F 검정 결과는  $F(4, 145)=20.314$ 로 유의 수준 .001에서 유의미하였다.  $R^2$ 는 .359로 역시 낮은 수치가 산출되었다. 수리 나형 집단의 경우, ‘단답형’, ‘계산’, ‘내적문제해결’, ‘행렬 연산과 역행렬’, ‘수와 식’, ‘발견적 추론’의 순으로 모형에 추가되어 수리 가형 집단과 비교하여 다소 차이를 보이고 있다. 수리 가형과 나형 집단 모두에서 공통적으로 강한 상관을 나타내었던 세 개의 내용 영역 변인 중 ‘도형의 방정식’은 가형 집단에 대한 모형에만 포함되었고, ‘수와 식’과 ‘행렬 연산과 역행렬’은 나형 집단에만 포함되었다. 수리 나형 집단의 최종 모형을 구성하는 여섯 개의 변인 중 세 개가 행동 영역 변인으로 나타난 것으로 보아 이 집단의 경우 행동 영역이 중요한 예측 변인의 역할을 함을 알 수 있다. 최종 모형에 대한 F 검정 결과는  $F(6, 143)=21.891$ 로 유의 수준 .001에서 유의미하였다.  $R^2$ 는 .479로 상대적으로 높은 설명력을 나타내었다.

고등학교 3학년 수리 가형 집단의 경우, ‘계산’, ‘단답형’, ‘이해’, ‘행렬 연산과 역행렬’, ‘연역적 추론’, ‘확률’의 순으로 모형에 포함되었다. 이 집단 역시 행동 영역의 설명력이 다른 영역보다 높은 것으로 나타났다. 내용 영역 중 정답률과 높은 상관을 보였던 ‘도형의 방정식’과 ‘지수와 로그’는 모형에서 제외되었다. 따라서 이들 영역은 단독으로는 의미가 있는 것처럼 보였으나 다른 영역과 함께하면 난이도에 큰 영향을 주지 않을 수 있다는 것이다. 또한 최종 모형에 대한 F 검정 결과는  $F(6, 143)=14.616$ 으로 유의 수준 .001에서 유의미하였으나,  $R^2$ 는 .380으로 비교적 낮은 설명력을 나타내었다. 수리 나형의 경우 ‘계산’, ‘단답형’, ‘완성형’, ‘행렬 연산과 역행렬’, ‘수열’, ‘이해’, ‘도형의 방정식’, ‘함수의 극한과 연속성’의 순으로 모형에 포함되었으며, 통계적으로 유의미한 상관을 보이지 않았던 ‘수열’과 ‘함수의 극한과 연속성’ 등이 회귀 모형에 포함되었다. 최종 모형에 대한 F 검정 결과는  $F(8, 141)=23.179$ 로 유의 수준 .001에서 유의미하였다.  $R^2$ 는 .568로 다섯 개의 집단 중 가장 높은 설명력을 나타내었는데, 이는 종속 변인의 전체 변산 중 56.8%를 예측 회귀 모형으로 설명할 수 있다고 해석될 수 있다. 2학년의 경우와 마찬가지로 가형 집단에 비해 나형 집단에 있어서 회귀 모형의 설명력이 비교적 큰 것으로 나타났다.

회귀 분석의 결과 역시, 전반적으로 내용 영역보다는 행동 영역 및 문항의 형식을 구성하는 독립 변인이 종속 변인에 대해 보다 큰 설명력을 나타내는 것을 알 수 있었으며 특히 행동 영역의 독립 변인이 가장 많은 설명력을 나타내었다.

<표 7> 최종 회귀 모형에 포함된 독립 변인에 대한 표준화 계수 및 t-검정 결과

	비표준화계수		표준화 계수	t	p	R <sup>2</sup>	
	B	표준오차					
고1	(상수)	48.118	1.717		28.031	.000	.354
	단답형	-19.958	2.053	-.491	-9.724	.000	
	계산	12.538	2.557	.274	4.903	.000	
	이해	8.351	2.268	.211	3.681	.000	
	합답형	-11.715	3.098	-.197	-3.781	.000	
	내적문제해결	-5.423	2.590	-.118	-2.094	.037	

고2 가형	(상수)	50.838	1.969		25.823	.000	.359
	계산	18.421	3.242	.388	5.681	.000	
	단답형	-18.018	2.962	-.414	-6.082	.000	
	도형의 방정식	-8.767	3.024	-.197	-2.899	.004	
	합답형	-9.540	4.346	-.153	-2.195	.030	
고2 나형	(상수)	34.126	1.353		25.229	.000	.479
	단답형	-17.502	2.041	-.532	-8.576	.000	
	계산	7.649	2.364	.208	3.236	.002	
	내적문제해결	-8.303	2.633	-.202	-3.153	.002	
	행렬	13.044	4.211	.189	3.098	.002	
	수와식	6.369	2.307	.171	2.761	.007	
	발견적 추론	-7.686	2.880	-.167	-2.669	.008	
고3 가형	(상수)	50.383	2.222		22.674	.000	.380
	계산	23.712	4.120	.429	5.767	.000	
	단답형	-12.406	3.299	-.253	-3.760	.000	
	이해	12.050	3.291	.259	3.661	.000	
	행렬	16.152	6.926	.163	2.332	.021	
	연역적 추론	14.844	6.137	.169	2.419	.017	
	확률	-14.021	5.904	-.159	-2.375	.019	
고3 나형	(상수)	38.212	2.031		18.815	.000	.568
	계산	29.537	3.384	.532	8.727	.000	
	단답형	-17.699	2.629	-.388	-6.732	.000	
	완성형	14.331	4.474	.187	3.203	.002	
	행렬	10.100	4.075	.142	2.479	.014	
	수열	10.213	4.752	.121	2.149	.033	
	이해	5.366	2.571	.126	2.087	.039	
	도형의 방정식	-7.781	3.320	-.132	-2.344	.020	
	함수의 극한과 연속성	-10.683	4.735	-.127	-2.256	.026	

<표 7>은 각 집단에 대한 예측 모형을 구성하는 각각의 독립 변인에 대한 비표준화 계수 및 표준오차, 그리고 표준화 회귀 계수와 t-검정 결과를 나타낸다. 비표준화 계수를 표준오차로 나눈 값에 대해 t-검정을 실시한 결과 각 예측 모형에 포함된 변인 각각에 대해 모두 .05에서 유의미한 결과를 나타내어, 모든 회귀 계수가 안정적으로 추정되었음을 알 수 있다.

### 3. 예측 모형의 교차 타당도 검증

마지막으로 회귀 분석에 사용된 자료를 통해 구한 예측 모형이 독립적으로 표집된 다른 표본 자료에서 어느 정도의 설명력을 가지는 지를 알아보기 위해 교차타당도 검증을 실시하였다. 교차 타당도를 위하여 2학년은 2006년 9월 자료로, 3학년은 2002년 10월 자료 등 새로이 수집된 자료에 정답률 예측을 위하여 최종적으로 산출한 회귀 모형을 적용하였으며, 1학년의 경우 독립된 자료를 구하는 데 어려움이 있어서 분석에 사용했던 자료 중 가장 최근 자료인 2006년 9월 자료를 사용하여 교차 타당도를 구하였다.

<표 8>은 교차 타당도 표본에서 구한 R<sup>2</sup>와 회귀 분석 표본에서 구한 R<sup>2</sup>, 그리고 설명력 저하율 등을 나타낸다. 먼저 회귀 분석에서 구한 예측 모형을 교차 타당도 표본에 적용하여 구한 정답률과 교차 타당도 표본에서 얻은 실제 정답률과의 상관은 고3 나형의 경우 .72로 가장 높게 나타났고 고3 가형의 경우 .503으로 가장 낮게 나타났다. 회귀 분석에서 구한 R<sup>2</sup>를 교차 타당도 표본에서 구한 R<sup>2</sup>와 비교하여 설명력의 저하율을 계산한 결과, 고2 나형에서는 2.8%의 저하율을 보여 예측 모형이 고2 나형 전체 집단에 대한 정답률을 상당히 안정적으로 예측하고 있음을 알 수 있다. 반면, 고3 가형의 경우 33.5%의 저하율을 나타내어, 고3 가형 전체 집단에 대한 정답률 예측이 다소 불안정함을 알 수 있다. 그러나 전반적인 R<sup>2</sup> 값을 볼 때, 정답률의 예측에 있어 가장 높은 설명력을 가지는 고3 나형 집단의 경우도 역시 전체 정답률 변산의 약 52% 정도밖에 설명하지 못하고 있고 다른 집단의 경우는 이보다 현저히 떨어지는 값을 보임으로써 정답률의 예측에 있어서 회귀 분석을 통한 예측 모형을 사용하는 데 상당한 제약이 따르는 것으로 판단된다.

<표 8> 정답률 예측 모형에 대한 교차 타당도 분석 결과

	예측모형의 정답률과 실제정답률과의 상관계수	교차 타당도 표본에서 구한 R <sup>2</sup>	회귀 분석의 R <sup>2</sup>	설명력 저하율
고1	.568	.323	.354	8.7%
고2 가형	.555	.308	.359	14.2%
고2 나형	.683	.466	.479	2.8%
고3 가형	.503	.253	.380	33.5%
고3 나형	.720	.519	.568	8.6%

#### IV. 마치는 말

이종승 외(2003)는 문항 난이도 추정모형을 개발하기 위하여 난이도 관련변인으로 2002 수능 회귀방정식모형, 2003 수능 회귀방정식 모형, 통합 구조방정식 모형, 통합 회귀방정식 모형 등 네 가지 예측 함수를 만들어 비교하였는데, 통합 회귀방정식 모형의 경우 수리영역의 난이도를 분석하기 가장 좋은 문항이라 하였다. 이 연구에 따르면, 통합 회귀방정식 모형은 문항의 형식에 따른 난이도를 분석하는 데 설명력이 가장 크고, 그 다음으로는 내용 영역과 행동 영역으로 보고하였으며 영역별 난이도 예측에 대한 분석적 접근과 객관적 접근의 가능성을 보여주었다.

본 연구에서는 폭을 더 넓은 2002년부터 2006년까지의 연합평가의 결과로 통합 회귀방정식 모형을 산출하고, 독립변인에 대한 예측모형 설정을 위하여 다중선형회귀분석을 실시하였다. 유의미한 변인만을 추출하기 위하여 정답률을 종속 변인으로 하여 단계 선택 방법을 택하여 최적화된 최종 모형을 결정하였으며, 독립적으로 수집된 다른 자료에 의한 설명력 저하 정도를 산출하여 교차타당도를 검증하였다.

본 연구 결과 5개의 집단에 대해 최종적으로 산출된 회귀 모형 각각에 대한 R<sup>2</sup> 값에 근거할 때, 정답률에 대해 비교적 안정된 설명이 가능한 것으로 나타났으며, 특히 고3 나형의 경우 가장 높은 설명력을 보이고 있어서, 정답률에 대한 전체 변산 중 56.8%를 예측 회귀

모형으로 설명할 수 있는 것으로 나타났다. 또한 본 연구에서는 내용 영역, 행동 영역, 문항 유형 등 각 영역별로 난이도와와의 관련성이 분석되었을 뿐 아니라 난이도를 예측하는 데 있어서 각 영역을 구성하는 하위 요인 간의 관련까지 보여주었다. 이는 같은 내용 영역이라 할지라도 행동 영역이나 문항의 형식에 따라서 난이도가 달라질 수 있음을 암시하고 있다. 가령 고등학교 1학년인 경우, 정답률에 미치는 영향은 단답형, 계산, 이해, 합답형, 내적문제 해결의 순이며 이들 중 계산과 이해 문제는 높은 정답률에, 나머지 3 영역은 낮은 정답률에 영향을 주는 것으로 분석되었다. 또한 고 2 가형의 경우도 단답형이 낮은 정답률에 가장 큰 요인을, 그 다음은 도형의 방정식과 합답형이 영향을 미치고 계산 문제가 높은 정답률에 미치는 영향이 가장 큰 것으로 나타났고, 나형인 경우 가형보다는 더 많은 요인이 정답률에 영향을 주는 것으로 나타났는데, 단답형, 내적문제해결, 발견적 추론이 낮은 정답률에 그리고 계산, 행렬, 수와식이 높은 정답률에 영향을 미치는 순으로 분석되었다. 마지막으로 고 3 인 경우, 계산, 이해, 단답형, 연역적 추론, 행렬, 확률의 순이며, 그중 단답형과 확률이 낮은 정답률에 유의미한 영향을 주는 것으로, 나형인 경우 2학년과 마찬가지로 가형보다는 더 많은 요인에 따라 영향을 받았는데, 계산, 단답형, 완성형, 행렬, 도형의 방정식, 함수의 극한과 연속성, 이해, 수열 순이며 이중 단답형과 도형의 방정식, 함수의 극한과 연속성은 낮은 정답률에 미치는 영향임을 알 수 있었다. 이것은 각 영역에 대한 각각의 영향 뿐 아니라 보다 종합적인 관련성을 설명하고 있다.

그러나 본 연구는 다음과 같은 측면에서 제한점을 가지고 있다. 첫째로 각 독립 변인들이 정답률을 예측하는 정도를 부호화하는 데 있어 어떤 문항에 대해 해당 변인이 측정 요인으로 포함되는 지의 여부로만 부호화함으로써 각 독립 변인들의 예측도에 대한 정교한 지표로 사용하는 데 한계가 있었다. 둘째, 분석에 사용된 총 문항수는 870개이나 각 독립 변인에 해당되는 문항수는 1에서 148까지 다소 큰 폭의 변화를 보이고 있어서 상관 분석 결과와 예측 모형 추정의 안정성에 있어 다소 제약을 받았다는 점이다. 셋째, 본 연구에서는 연도별로 응시자의 특성에 큰 변화가 없다는 가정 하에 통합회귀방정식모형을 사용하였으므로 응시자의 특성에 있어서 변화가 큰 경우 결과의 적용에 있어 한계가 있다는 점이다.

또한 난이도 변인을 추출하기 위해 이와 같이 내용 영역과 행동 영역, 문항의 형식 등의 요인들을 고려하는 것은 나름대로의 의미가 있다고 판단되나, 같은 내용 영역에 속한다고 하더라도 난이도는 매우 다르게 나타날 수 있다. 예를 들어, 본 연구에서는 ‘함수의 극한과 연속성’이 정답률과 부적 상관관계를 보이는 것으로 나타났으나 사실 그것은 어떠한 내용으로 어떠한 문제해결력을 요하냐에 따라 충분히 달라질 수 있을 것이다. 이러한 이유에서 어느 정도의 범주를 설정하여 목록에 따른 난이도 관련 변인을 내용 영역, 행동 영역, 문항 형태와 관련된 요인으로 구분하여 탐색한 후 보다 구체적인 문항 내적인 요인을 파악하는 것이 필요할 것이다. 앞에서 나온 결과에서 보여주듯이, 정답률의 예측에 있어서 회귀분석을 통한 예측 모형을 사용하는 것은 다소 제약이 따르며 집단별로 상당한 차이를 보이고 있다. 실제로 문항의 난이도에 미치는 영향은 앞에서 제시한 3가지 영역 외에도, ‘내용이나 제재의 생소성 정도’나 ‘문제 해결에 필요한 개념의 수’, ‘계산의 복잡성 정도’, ‘문제 해결에 필요한 추론의 정도나 걸리는 시간’, ‘매력적인 오답’ 등 상당히 복잡하고 다양한 요소에 의한 것이어서 객관화된 영역에 의한 분류만으로 측정하기에는 다소 한계가 있는 것으로 판단된다. 따라서 정성적인 방법에 의하여 학생들이 어려움을 느끼는 요인들과 교사가 어렵다고 판정하는 원인들을 보다 정밀하게 파악해 볼 연구의 필요성이 강하게 대두된다고 할 수 있다.

## 참고문헌

- 이종승·김성훈·김재철·송현정·박문환·장경숙 (2003). 대학수학능력시험 문항난이도 추정모형개발-언어 영역, 수리 영역, 영어 영역을 중심으로. *교육평가연구*, 16(2), 1-23.
- 장경숙 (2005). 대학수학능력시험 외국어(영어)영역 읽기 난이도 예측 모형 개발. *Foreign Language Education*, 11(1), 111-130.
- 한국교육과정평가원 (2002). 고1연합학력평가(6월) 결과분석-수리 영역. 대수능 CAT 2002-17. 서울.
- 한국교육과정평가원 (2002). 고1연합학력평가(11월) 결과분석-수리 영역. 대수능 CAT 2002-19. 서울.
- 한국교육과정평가원 (2003). 고1연합학력평가(6월) 결과분석-수리 영역. 대수능 CAT 2003-32. 서울.
- 한국교육과정평가원 (2003). 고1연합학력평가(12월) 결과분석-수리 영역. 대수능 CAT 2004-1-2. 서울.
- 한국교육과정평가원 (2004). 고1연합학력평가(6월) 결과분석-수리 영역. 대수능 CAT 2004-6-2. 서울.
- 한국교육과정평가원 (2004). 고1연합학력평가(9월) 결과분석-수리 영역. 대수능 CAT 2004-9-2. 서울.
- 한국교육과정평가원 (2005). 고1연합학력평가(6월) 결과분석-수리 영역. 대수능 CAT 2005-4-2. 서울.
- 한국교육과정평가원 (2006). 고1연합학력평가(6월) 결과분석-수리 영역. 대수능 CAT 2006-11-2. 서울.
- 한국교육과정평가원 (2006). 고1연합학력평가(9월) 결과분석-수리 영역. 대수능 CAT 2006-12-2. 서울.
- 한국교육과정평가원 (2002). 고2연합학력평가(6월) 결과분석-수리 영역. 대수능 CAT 2002-20. 서울.
- 한국교육과정평가원 (2003). 고2연합학력평가(6월) 결과분석-수리 영역. 대수능 CAT 2003-29. 서울.
- 한국교육과정평가원 (2004). 고2연합학력평가(6월) 결과분석-수리 영역. 대수능 CAT 2004-7-2. 서울.
- 한국교육과정평가원 (2005). 고2연합학력평가(6월) 결과분석-수리 영역. 대수능 CAT 2005-5-2. 서울.
- 한국교육과정평가원 (2006). 고2연합학력평가(6월) 결과분석-수리 영역. 대수능 CAT 2006-14-2. 서울.
- 한국교육과정평가원 (2006). 고2연합학력평가(9월) 결과분석-수리 영역. 대수능 CAT 2006-16-2. 서울.
- 한국교육과정평가원 (2002). 고3연합학력평가(6월) 결과분석-수리 영역. 대수능 CAT 2002-7. 서울.
- 한국교육과정평가원 (2002). 고3연합학력평가(10월) 결과분석-수리 영역. 대수능 CAT

- 2002-27. 서울.
- 한국교육과정평가원 (2003). 고3연합학력평가(10월) 결과분석-수리 영역. 대수능 CAT 2003-42. 서울.
- 한국교육과정평가원 (2004). 고3연합학력평가(10월) 결과분석-수리 영역. 대수능 CAT 2004-23-2. 서울.
- 한국교육과정평가원 (2005). 고3연합학력평가(10월) 결과분석-수리 영역. 대수능 CAT 2006-12-2. 서울.
- 한국교육과정평가원 (2006). 고3연합학력평가(10월) 결과분석-수리 영역. 대수능 CAT 2006-13-2. 서울.
- Alderson, J. C. (2000). *Assessing reading*. Cambridge: Cambridge University Press.
- Baumann, J. F., & Serra, J. K. (1984). The frequency and placement of main ideas in children's social studies textbooks: A modified replication of Braddock's research on topic sentences. *Journal of Reading Behavior*, 16, 27-40.
- Freedle, R., & Kostin, I. (1991). *The prediction of SAT reading comprehension item difficulty for expository prose passages*. ETS Research Report RR 91-29. Princeton, NJ: Educational Testing Service
- Freedle, R., & Kostin, I. (1993). *The prediction of TOEFL reading comprehension item difficulty for expository prose passages for three item types: Main idea, inference, and supporting idea items* (Report No. 44). Princeton, NJ: Educational Testing Service.
- Jin, K. & Park, J. (2005). Factors related to the difficulty of college scholastic ability test English items. *English Teaching*, 60(3), 17-25.
- Kostin, I. (2004). *Exploring items characteristics that are related to the difficulty of TOEFL dialogue items* (ETS Research Report PR-04-11). Princeton, NJ: Educational Testing Service.
- Lee, J.-W. (2002). An exploratory study on reading comprehension test-taking processes and strategies in the EFL context. *English Teaching*, 57(4), 177-195.
- Miller, G. (2003). Analyzing the minority gap in achievement scores: Issues for states and federal government. *Educational Measurement: issues and Practice*, 22(3). 30-36.
- Scheuneman, J. D., Gerritz K., & Embretson, S. E. (1991). *Effects of Prose Complexity on Achievement Test Item Difficulty* (RR 91-43). Educational Testing Service, Princeton

## Factors of Predicting Difficulty of Mathematics Test Items in College Scholastic Ability Test

Ko, Ho Kyoung<sup>3)</sup> · Yi, Hyun Sook<sup>4)</sup>

### Abstract

This study explored the possibility of building a statistical model predicting difficulty of mathematics test items through the analysis of nation-wide scholastic ability test results for the past 5 years. Multiple linear regression analysis was conducted in predicting difficulty of mathematics test items. We adopted three major areas for independent variables: the content area, the behavior area, and the test item format area, each of which was categorized into more detailed sub-areas. For the dependent variable, the proportion of correct answer was used to represent the item difficulty. Statistically significant independent variables were included in the regression model based on the stepwise selection method. Several important factors affecting difficulty of mathematics test items for each area were identified. R-squares for the final regression model were fairly high, implying that the regression equation can be used to predict difficulty of test items at an acceptable level. Lastly, the regression model was cross-validated using independently collected data. We believe that this study will provide basic but very critical information for predicting the proportion of correct answer by showing the factors that should be considered for developing mathematics test items for the college entrance examination or high school classroom test.

Key Words : College Scholastic Ability Test, High school mathematics test, Item difficulty, Factors of item difficulty,

---

3) Korea Institute of Curriculum and Evaluation (koho@kice.re.kr)

4) Korea Institute of Curriculum and Evaluation (hyunsyi@kice.re.kr)