

대학수학능력시험 난이도 관련 변인 탐색

박 문 환*

본 연구의 목적은 대학수학능력시험에서 출제되는 수리 영역의 문항에 대한 난이도를 결정짓는 변인을 탐색하고, 난이도를 측정할 수 있는 구체적인 검토�도구를 개발하는 것이다. 이를 위해 문헌 분석과 기출 수능 문항에 대한 내용 및 난이도 분석을 시도하였다.

여기서 추출된 변인을 토대로 1차 검토투구를 만들고 오답률과 추출된 변인들 사이의 상관분석을 시도하여 상관도가 비교적 낮은 변인은 삭제하거나 수정 보완하여 2차 검토투구를 만들었다. 2차 검토투구를 이용하여 여러 가지 예측 모형을 개발하고, 각 모형에 대하여 회귀계수를 비교하고, 교차타당도를 분석한 결과 안정성과 정확성이 가장 높은 최적의 예측모형은 '통합 회귀방정식 모형'으로 나타났다. 또한 통합 회귀방정식 모형을 토대로 분석하였을 때 오답률에 대한 설명력은 답지의 형식, 내용 영역, 행동 영역, 문항의 생소성 등의 순으로 나타났다.

본 연구는 난이도 관련 변인에 대한 연구가 전무한 상태에서 이루어졌기 때문에 많은 한계가 있으며 추후에 난이도를 보다 정확하게 추정할 수 있는 연구 방법이 기대된다.

1. 서 론

난이도는 문항의 특성을 결정하는 중요한 요인이다. 특히 대학수학능력시험(이하, 수능)과 같이 사회적으로 관심도가 높은 시험의 경우 난이도 수준의 안정성과 적정성은 중요한 역할을 한다. 예컨대, 수능이 너무 쉽게 출제되면 최상위 학생들을 변별하기가 어려워져서 비슷한 점수대의 학생들을 선발해야 하는 대학의 입장이 곤란하게 되고, 수능이 너무 어렵게 출제되면 과열 과외를 부추기게 되어 학부모들의 경제적 부담과 학생들의 학습 부담을 가중시키게 된다.

이러한 점을 고려하여 교육인적자원부와 한

국교육과정평가원은 수능의 난이도가 매년 비슷한 수준에서 적정 수준을 유지하도록 하겠다는 방침을 세워 놓고 있으며, 수능의 난이도를 적정 수준으로 유지하기 위한 방안으로써 출제 위원에 현직교사 비율을 늘이고 수능 연구를 위한 상시체제를 구축하고는 있으나, 출제 과정에서 난이도를 평가하는 전문가 집단의 대응 방안은 아직 이론이나 체계적 분석에 근거하지 못하고 있다. 즉 '모든 난이도 요소를 고려했을 때 이 문항의 정답률은 대략 몇 퍼센트 정도'라는 통합적이고 주관적인 전문가의 감에 전적으로 의존하고 있다. 이에 본 연구는 난이도 관련변인을 몇 개의 하위영역으로 나누고, 동시에 각 하위영역에 서너 개의 평가 기준을 제시하여 각 문항에 대한 난이도를 평가하게 함

* 한국교육과정평가원, e-mail: sorada@kice.re.kr

으로써 통합적이고 주관적인 접근에 치우치고 있는 난이도 평가를 보완하려 하였다.

본 연구에서는 문항 난이도가 어떠한 변인에 영향을 받는지 탐색하기 위해서 선행 연구 개관, 기존 수능 문항의 내용분석, 전문가 의견 수렴 등의 방법을 이용하였다. 가능한 많은 변인으로 1차 검토도구를 만들었으며, 1차 검토 도구에 대하여 상관분석 및 회귀분석을 통하여 1차 예측모형을 만들었다. 그리고 1차 예측모형에서 다소 중복된다고 판단되는 변인을 삭제하여 2차 검토도구를 만들고, 다시 2차 검토도구에 대하여 몇 가지 예측모형을 구안하고, 이들 예측모형 가운데 최적의 예측모형을 선택하여 수리영역에 대한 난이도 예측모형을 제안하고자 한다.

결국 본 연구는 수능 난이도에 영향을 미칠 수 있는 변인들을 체계적이고 심층적으로 분석하여 탐색함으로써 난이도 예측에 유의미한 정보를 제공하는 것을 그 목적으로 하고 있으며, 그 목적을 구체적으로 제시하면 다음과 같다.

첫째, 난이도에 영향을 미칠 수 있는 변인이 무엇인지 탐색하고 그것을 측정할 수 있는 구체적인 검토도구를 개발하는 것이다.

둘째, 인문계 상위 50%의 예상 정답률을 예측하는 모형을 개발하는 것이다. 예측 모형의 안정성과 정확성을 보기 위해서 '중다상관분석'과 '실제 오답률·예측 오답률 간 상관' 및 '실제 정답률 평균'과 '예측 정답률 평균'을 비교하고자 한다.

셋째, 안정성과 예측력에서 가장 탁월한 모형을 최종 예측 모형으로 선정한 다음, 최종 예측 모형의 설명력과 영역에 따른 난이도 관련변인의 차이를 비교하고자 한다.

II. 수리영역 난이도 변인 추출 및 예측 모형 개발

1. 1차 검토 도구의 변인 추출 근거

수리영역의 난이도 관련 변인에 대한 선행 연구나 수학 시험에서 어떠한 변인이 전체 학생의 정답률에 어느 정도 영향을 미치는지에 대한 구체적인 국내 연구 결과를 찾아보기 힘들다. 그러나 정답률에 영향을 미치는 변인에 대한 시사점을 받을 수 있는 문헌은 찾아볼 수 있다. 본 연구에서는 이러한 문헌 연구를 통해 관련변인을 추출하고 이러한 변인들이 기출 수능 문항에서 어느 정도 상관관계를 가지는지를 밝힘으로서, 각 문항에 대한 정답률 산출에 대한 근거를 찾고자 한다.

먼저 현행 수능에서의 평가목표 분류체계에 따르면, 내용 영역은 각 과목(공통수학, 수학 I, 수학 II)의 대단원별로 분류되고 있으며, 행동 영역은 계산, 이해, 추론, 문제해결로 분류된다.

그러나 내용영역과 행동영역만을 난이도에 영향을 미치는 변인으로 판단하기는 다소 무리가 있기 때문에 또 다른 난이도 관련 변인을 추출하기 위해 관련된 참고문헌을 살펴보았다. 이를 위해 우정호(1992, 1998), 강시중(1993), 강옥기 외(1989), 강완·백석윤(1998), Boodoo(1997), Standards(1989), TIMSS 보고서(Travers, 1989) 등의 문헌을 조사하였으나 대략적으로는 현행 수능에서의 분류방식과 유사한 것으로 나타났다. 수능의 난이도 관련 변인 가운데 내용 영역과 행동 영역이 중요한 요소임을 확인할 수 있었다. 하지만 같은 내용영역에 속한다

1) 단순상관분석이란 하나의 준거변수와 하나의 예측변수간의 관련성을 분석하는 것이며, 중다상관분석이란 하나의 준거변수와 두개 이상의 예측변수간의 관련성을 분석하는 것이다.

고 하더라도 난이도는 매우 상이하게 나타날 수 있다. 예를 들면, 삼각함수에 대한 내용이라고 하더라도 도입부분에서 다루는 내용과 관련된 문제와 사인정리나 코사인정리와 같이 상당한 정도의 학습이 이루어진 이후에 다루어지는 내용과 관련된 문제는 난이도가 유사하다고는 할 수 없을 것이다. 그렇다고 해서 내용 영역을 세분하여 분류하면 거의 100여개 이상이 나타나는데, 이렇게 할 경우 각 요소를 난이도 별로 분류해야 하는 등의 문제점이 발생하게 되어 분석 자체가 매우 힘들어지게 된다. 행동영역에 있어서도 '이해' 영역에 속하는 문제 중에는 기본적인 개념을 적용하는 문제가 있는가 하면, 몇 가지 개념 사이의 관련성을 파악하고 이를 적용하는 문제가 있을 수 있다. 이러한 이유에서 내용영역은 목록의 수를 줄여서 보다 큰 범주로 구분하고, 행동영역 가운데 이해 영역은 기본적인 개념을 적용하는 간단한 이해와 다소 복잡한 이해로 구분하였다. 대신 정답률에 대한 보다 정확한 예측을 위해 난이도와 관련된 것으로 판단되는 또 다른 범주를 찾고자 하였다. 이를 위해 Freedle(1991)의 문헌 등을 참고하여 수학 문항의 난이도와 관련된 것으로 판단되는 요인으로 문제의 형식, 제재의 생소성, 답지의 형식, 오답의 매력도, 문제의 배열 위치, 문제의 익숙도 등을 추출하였다. 또한 기출 대수능의 수학 문항을 검토하는 과정에서 문제의 복잡도, 풀이 절차의 복잡도, 풀이 방법의 다양성, 보조적 자료의 사용 여부도 난이도에 영향을 미칠 것으로 판단되어 이들 변인도 난이도 관련 변인에 첨가하였다.

이상의 난이도 관련 변인을 크게 내용영역, 행동영역, 문항 내적인 요인, 문항 형태와 관련된 외적인 요인으로 구분하고, 이들 범주에 따라 난이도 관련 변인을 <표 II-1>과 같이 분류하여 수능 출제 경험이 풍부한 교수 2명과 교

사 1명으로 구성된 협의회를 통해 적절성 및 타당성을 검증 받았다.

전문가 협의회에서는 다음과 같은 의견이 제기되었다.

먼저 첨가되어야 할 변인으로 답지 선택의 단순성, 문장 표현의 명료성, 답의 추정 가능성, 합답형, 계산의 복잡성, 수험생 변인 등이 고려되어야 한다는 지적이 있었으며, 제재의 생소성과 문제의 익숙도는 서로 반비례 관계에 있다고 판단되기 때문에 중복되는 변인으로 판단되고, 또한 현행 수능에서 행동영역별로 문제의 위치가 정해지기 때문에 문제의 배열 위치와 중복된다는 의견이 있었다. 또한 풀이 절차의 복잡도, 풀이 방법의 다양성, 보조적 자료의 사용, 내용 영역의 분류 등에 있어서는 다소 수정이 필요하다는 언급이 있었다. 또한 제재의 생소성은 다양한 수준으로 분류될 수 있기 때문에 연속적인 척도를 사용하는 것이 합당할 것이라는 지적도 있었다.

이러한 의견을 바탕으로 중복된다고 판단되는 변인은 다시 조정하고 난이도 관련 변인으로 부적합하다고 판단되는 변인은 삭제하여 다음과 같이 재분류하였다.

또한 예측변인의 수가 너무 많으면 코딩하는 시간이 늘어나게 되어 검토 교사의 부담이 가중될 뿐만 아니라 빠른 시간 내에 문제지 전체에 대한 난이도를 추정할 수 없다는 문제점이 제기됨에 따라 검토 교사가 검토해야할 변인의 수를 줄일 필요성이 생겼다. 그래서 교육과정에서 제시한 분류, 출제 과정에서 제시되는 이원목적 분류표나 문항분석표를 통해서 확인 가능한 변인과 문제지를 통해 판단 가능한 변인 등 객관적으로 코딩 가능한 문항은 이러한 자료를 이용하기로 하고, 나머지 변인에 대해서만 검토 교사에게 코딩을 받기로 하였다.

이러한 논의 결과를 바탕으로 크게 교사의

검토가 필요한 변인과 객관적인 변인으로 분류하고 다음과 같이 정리하였다. 각 변인에 대해서는 중간 수준에 해당하는 예시문항이 포함되어 있으며, 이를 기준으로 하여 각 검토 교사는 문항의 변인별 수준을 측정하게 된다.

- 교사의 검토가 필요한 변인 : 제재의 생소성, 문제의 복잡도, 답의 추정 가능성, 계산의 복잡도, 보조적 자료, 풀이 방법의 다양성
- 객관적인 변인 : 내용영역, 행동영역, 답지의 형식

2. 1차 검토 도구의 척도화

본 연구에서는 각 변인의 수준 즉 변인별 척도를 구성하는데 있어서 수준이 높을수록 어려운 문항과 관련되도록 한다는 원칙을 세웠다. 그렇게 함으로서 기출 대수능 문항의 오답률과 각 측정변인 사이의 상관도를 계산하기 위함이며, 최종적으로 이러한 상관도를 바탕으로 측정변인들은 문항에 대한 오답률을 예측하는데 활용하기 위함이다.

<표 II-1> 난이도 예측 변인과 척도

예측변인	개념적 정의	측정 문항	척도
내용영역	이원목적분류표의 내용 영역	이 문항의 내용 영역은 다음 중 어디에 속하는가?	① 수학적 기초 ② 대수 ③ 기하 ④ 해석 ⑤ 확률과 통계
행동영역	이원목적분류표의 행동 영역	이 문항의 행동 영역은?	① 계산 ② 이해/추론 ③ 문제해결
내적 요인	문제의 형식	문장제 문제 여부	① 아니다. ② 그렇다
	제재의 생소성	학습자 수준, 개념의 전문성 등을 고려한 내용이나 제재의 낯선 정도	① 매우 익숙하다. ② 이해할 수 있는 보통 수준이다. ③ 생소하고 이해하기 어렵다.
	문제의 복잡도	문제해결을 위해 필요한 개념의 수	① 1개 ② 2개 ③ 3개 이상
	답지의 형식	5지선다형/단답형 여부	① 5지선다형 ② 단답형
	풀이 절차의 복잡도	풀이 절차의 복잡성 여부	① 단순하다. ② 보통이다. ③ 복잡하다.
	풀이 방법의 다양성	다양한 풀이 방법의 가능성 여부	① 그렇다. ② 아니다.
	보조적 자료	표나 그래프, 그림 등의 사용 여부	① 그렇다. ② 아니다.
	오답의 매력도	정답에 영향을 주는 오답지의 혼동 정도	① 특이한 오답요인은 없다. ② 보통 수준의 오답 요인이 있다. ③ 높은 수준의 오답 요인이 있다.
외적 요인	문제의 배열 위치	문제가 속한 배열의 위치는?	① 초반(1번-10번 사이) ② 후반(21번-30번 사이) ③ 중반(11번-20번 사이)
	문제의 익숙도	교과서 및 참고서에서 다루어진 정도	① 충분히 다루어진다. ② 부분적으로 다루어진다. ③ 전혀 생소하다.

가. 객관적인 변인

객관적인 변인이란 교육과정에서 제시한 분류, 이원분류표에 의한 분류, 문제지 등을 통해 비교적 객관적으로 판단 가능한 변인을 의미하며, 크게 내용영역, 행동영역, 답지의 형식으로 구분하였다.

① 내용 영역

앞에서도 언급하였듯이 내용영역을 세분화하면 그 수가 너무 많아지게 되어 이를 척도로 사용하기에는 무리가 있기 때문에 교육과정 및 대수능에서의 분류에 따라 크게 수학적 기초, 대수, 기하, 해석, 확률과 통계로 나누고 각 영역별로 역대 대수능(2000년부터 2003년)에서의 평균 정답률을 조사하였으며, 그 계산한 결과는 <표 II-2>와 같다.

이러한 결과에 근거하여 내용 영역별 수준을 수학적 기초, 대수, 해석, 기하, 확률과 통계의 순으로 배열하였다.

② 행동영역

현행 대수능에서는 행동영역을 크게 계산, 이해, 추론(발견적 추론, 증명), 문제해결(내적,

외적)으로 구분하고 있다. 내용영역에서와 마찬가지로 역대 대수능에서의 각 행동영역별 정답률을 조사하였다.

2000년부터 2003년까지의 대수능의 평균정답률은 <표 II-3>과 같다.

이해 문제는 매우 간단한 지식만을 적용하는 문제뿐만 아니라 상당히 고급 지식을 적용하여야 하는 문제도 있으며, 또한 위의 표를 통해 이해와 발견적 추론 사이의 정답률의 차이가 없다는 점 때문에 이해를 '간단한 이해문제'와 '복잡한 이해문제'로 구분하였다²⁾. 이때 '간단한 이해문제'는 계산 다음으로 정답률이 높을 것으로 예상되며, '복잡한 이해문제'는 내적 문제해결 다음으로 정답률이 낮을 것으로 예상된다.

이러한 이유로 행동영역에 대한 척도를 계산, 간단한 이해, 증명, 외적문제해결, 발견적 추론, 복잡한 이해, 내적문제해결의 순으로 정하였다.

③ 답지의 형식

답지의 형식에 대한 척도는 기출 수능 문항을 조사하여 크게 합답형³⁾이 아닌 5지선다형, 합답형, 단답형으로 구분하였다. 여기서 합답형

<표 II-2> 2000년부터 2003년까지의 내용영역별 평균정답률

내용영역	수학적 기초	대수	해석	기하	확률과 통계
평균정답률	75.5	67.7	55.9	50.7	41.9

<표 II-3> 2000년부터 2003년까지의 행동영역별 평균정답률

행동영역	계산	이해	추론		문제해결	
			발견적 추론	증명	내적	외적
평균정답률	85.94	56.55	56.00	74.49	37.20	58.29

2) 본 연구에서는 기본적인 개념을 적용하는 문제를 '간단한 이해'로, 그리고 몇 가지 개념 사이의 관련성을 파악하고 이를 적용하는 문제를 '복잡한 이해'로 구분하였다.

3) 합답형 문항이란 몇 개의 명제를 <보기> 형태로 제시하고, 그 중에서 참인 것을 모두 고르거나 거짓인 것을 모두 고르는 형태의 문제를 의미한다.

문항은 보통 2개 이상의 부분 문제로 분류할 수 있기 때문에 합답형이 아닌 5지선다형 문제 보다는 더 어려울 것으로 판단된다. 이러한 이유로 답지의 형식에 대한 척도를 합답형이 아닌 5지선다형, 합답형, 단답형의 순으로 정하였다.

나. 교사의 검토가 필요한 변인

교사의 검토가 필요한 변인은 다소는 개인의 주관적인 판단에 의해 달라질 수도 있는 변인을 의미하며, 제재의 생소성, 문제의 복잡도, 답의 추정 가능성, 계산의 복잡도, 보조적 자료, 풀이 방법의 다양성으로 구분하였다. 그리고 각 변인에 대한 개념적 정의와 척도는 다음과 같다.

① 제재의 생소성

학습자 수준, 개념의 전문성 등을 고려한 내용이나 제재의 낯선 정도를 말한다.

제재의 생소성에 대한 척도는 전문가 협의회에서 지적된 사항을 반영하여 0에서 10까지의 연속적인 척도로 제시하였으며, 다음의 기준에 의한다.

- 0 : 매우 초보적인 수준의 문항
- 2 : 교과서의 기본예제 등에서 다루어지는 정도의 문항
- 4 : 교과서의 응용문제 등에서 다루어지는 정도의 문항
- 6 : 익숙한 기호, 개념 등을 사용하고 있으나, 이러한 기호나 개념 등에 대한 관계를 종합하여 새로운 사실을 묻고 있는 문항
- 8 : 낯선 제재(용어에 대한 새로운 정의 등)를 사용하고 있으나, 그 의미를 이해한다면 해결의 실마리를 쉽게 찾을 수 있는 문항
- 9이상 : 낯선 제재(용어에 대한 새로운 정의 등)를 사용하고 있으며, 해결의 실마리를 찾기 위해서 상당한 사고력이 요구되는 문항

② 문제의 복잡도

문제해결을 위해 필요한 개념의 수를 말한다.

문제의 복잡도에 대한 척도를 정하는데 있어서 '나머지 정리'를 이용하여 '최대공약수'를 구하는 경우 혹은 확률밀도함수를 구하기 위해 적분을 이용하는 경우와 같이 두 개 이상의 개념이 서로 밀접한 관련이 있을 때에는 단일한 개념으로 분류하도록 하였다. 그러나 '가우스 기호'와 '이차방정식'과 같이 서로 밀접한 관련이 없지만 문제를 해결하기 위해서는 두 개 이상의 개념을 활용할 때에는 필요한 개념의 수는 2개 이상으로 분류하도록 하였다.

③ 답의 추정 가능성

답을 추정하는 가능성의 정도를 말한다.

답의 추정 가능성이란 특정한 값을 대입하여 답을 구할 수 있거나 직관적으로 정답이 아닌 것이 분명한 오답의 경우 배제하고 답을 추측하는 정도를 의미한다.

④ 계산의 복잡도

정상적인 계산 절차(또는 풀이과정)에 대한 복잡성 여부를 의미한다.

계산의 복잡도는 다음과 같이 기준을 정하였다.

- 단순하다 : 1~2 줄의 간단한 풀이가 필요
- 보통이다 : 3~4 줄의 풀이가 필요
- 복잡하다 : 5줄 이상의 풀이가 필요

⑤ 보조적 자료

문제에 사용된 표나 그래프, 그림 등이 문제의 이해에 도움이 되는 정도를 의미한다.

보조적 자료란 문제를 보다 쉽게 이해하기 위한 자료이거나, 그 자료를 사용하지 않으면 문제의 진술이 어려워져서 사용된 자료를 의미한

다. 표, 그래프, 그림 등이 문제 자체와 관련이 있으면, 보조적 자료가 아니다. 보조적 자료에 대한 척도는 크게 5단계로 구분하였으며, 직관적 판단에 의해 척도를 정하기는 하였지만 연구를 진행하면서 수정이 필요할 것으로 판단되는 예측변인이다.

⑥ 풀이 방법의 다양성

다양한 풀이 방법의 존재 여부를 의미한다.

다양한 풀이방법이 있는 경우, 그렇지 않을 때보다는 학생들이 더 쉽게 문제를 풀 수 있을 것으로 판단된다. 그러나 다양한 풀이방법에 대한 가능성은 교사의 판단에 의존할 수 밖에 없을 것이다.

이상을 정리하여 <표 II-4>와 같이 예측변인을 정하였다.

3. 1차 예측 모형 개발

최종적으로 선택된 9개의 변인을 3명의 교사에게 검토를 의뢰하고, 이를 2001년, 2002년, 2003년 수능 문항과 2003년 모의 수능 문항 중에서 인문계 문제만을 대상으로 하여 오답률과 각 예측변인과의 상관관계를 분석하였다. 여기서 상관관계가 1에 가깝다는 것은 두 변인 사이의 종속성이 강하다는 것을 의미하며, 0에

가깝다는 것은 두 변인이 서로 독립적이라는 것을 의미한다. 본 연구에서는 요인분석을 하는데 있어서 상관관계가 0.3 이상인 경우 두 변인이 종속적이라고 판단하였으며, 0.3 미만인 경우 두 변인이 독립적이라고 판단하였다. 이는 요인분석을 하는데 있어서 상관관계수 0.3을 기준으로 삼고 있기 때문이다(김계수, 2001). 그러나 연도별로 각 변인별 상관관계수가 큰 차이를 보이고 있기 때문에 변인 사이의 종속성이 강하다고 하더라도 수학의 난이도 관련 변인으로 상당한 의미가 있다고 판단된 경우에는 그 변인을 수용하기도 하였으며, 이러한 과정에서 연구자의 주관적 판단도 상당히 개입되어 있음을 밝힌다. 3명의 검토 교사를 통해 코딩 작업은 2001년, 2002년, 2003년 수능 문항과 2003년 모의 수능 문항에 대해 실시하였으나 2001년도 수능이 소위 ‘물수능’이라는 비판을 받을 정도로 다른 수능에 비해 점수가 매우 높게 나타났다는 점과 실제로 연도별 상관분석을 한 결과가 매우 특이하게 나타났다는 점에서 2001학년도 상관분석 결과는 배제하였으며, 또한 2003년 모의 수능에 많은 실업계 고등학생이 배제되어 있다는 점 때문에 본 수능 결과와는 다소 다른 결과를 보일 수 있다는 지적에 따라 2003년 모의 수능에 대한 상관분석 결과는 배제하기로 하고 2002년과 2002년 수능 결과를 중심

<표 II-4> 1차 검토도구

분류	번호	예측변인	척도
교사의 검토가 필요한 변인	v1	제재의 생소성	[0, 10]의 연속척도
	v2	문제의 복잡도	3단계 척도
	v3	답의 추정 가능성	5단계 척도
	v4	계산의 복잡도	3단계 척도
	v5	보조적 자료	5단계 척도
	v6	풀이 방법의 다양성	2단계 척도
객관적인 변인	v7	내용영역	5단계 척도
	v8	행동영역	7단계 척도
	v9	답지의 형식	3단계 척도

으로 상관분석 결과만을 이용하기로 하였으며, 그 결과는 <표 II-5>~<표 II-7>과 같다.

가. 상관분석 결과

앞에서 언급한 9개의 변인을 독립변인으로 하고, 오답률(=100-정답률)을 종속변인으로 하

여 총 10개의 변인에 대한 상관도를 분석하였다. 또한 수능 문제에 대한 각 변인의 설명력을 조사하기 위하여 연도별로 상관도를 구하고, 2002년과 2003년 수능의 문항을 대상으로 상관도를 조사하였다. 이러한 분석을 통하여 독립변인 사이에 특히 상관관계가 높은 경우

<표 II-5> 2002학년도 상관분석 결과

	오답률	V1	V2	V3	V4	V5	V6	V7	V8	V9
오답률	1									
V1	.473**	1								
V2	.193	.403**	1							
V3	.167	-.168	.096	1						
V4	.413**	.523**	.443**	-.118	1					
V5	.141	.284**	.075	.050	-.094	1				
V6	-.137	-.002	-.021	.017	-.027	-.147	1			
V7	.397**	.137	.000	.180	.077	-.020	-.061	1		
V8	.460**	.604**	.409**	-.082	.354**	.208*	-.216*	.146	1	
V9	.364**	-.110	-.129	.046	.136	-.169	.094	.053	-.219*	1

*p<.05 **p<.01

<표 II-6> 2003학년도 상관분석 결과

	오답률	V1	V2	V3	V4	V5	V6	V7	V8	V9
오답률	1									
V1	.565**	1								
V2	.307**	.325**	1							
V3	-.122	-.246*	-.312**	1						
V4	.452**	.497**	.256*	.164	1					
V5	.114	.143	-.010	-.049	.100	1				
V6	-.227*	-.192	.000	.146	-.119	-.196	1			
V7	.507**	.142	.137	.031	.210*	.176	-.228*	1		
V8	.450**	.303**	.197	-.016	.422**	.172	-.100	.309**	1	
V9	.305**	-.026	-.056	.056	.204	-.301**	.114	.125	.153	1

*p<.05 **p<.01

<표 II-7> 두 개의 시험지를 통합적으로 상관분석한 결과

	오답률	V1	V2	V3	V4	V5	V6	V7	V8	V9
오답률	1									
V1	.523**	1								
V2	.244**	.380**	1							
V3	.005	-.199**	-.077	1						
V4	.433**	.515**	.378**	.035	1					
V5	.127	.212**	.040	-.006	.002	1				
V6	-.183*	-.095	-.009	.090	-.069	-.173*	1			
V7	.456**	.159*	.083	.108	.151*	.080	-.138	1		
V8	.457**	.465**	.331**	-.041	.389**	.190*	-.157*	.233**	1	
V9	.332**	-.068	-.096	.051	.167*	-.237**	.104	.086	-.039	1

*p<.05 **p<.01

문제에 대한 각 변인 사이의 설명력이 중복된다는 것을 의미하기 때문에 이들 변인에 대해서는 다른 문항과 통합하여 수정 보완하였으며, 독립변인과 종속변인 사이에 상관관계가 낮은 문항에 대해서도 보완을 하였다. 각 연도별로 오답률과 9개의 예측변인에 대한 상관관계의 분석 결과는 <표 II-8>과 같이 나타났다.

앞에서도 언급하였듯이 독립변인 사이의 상관도는 0에 가까울수록 서로 독립적이라고 할 수 있기 때문에 문항 난이도를 예측할 때 설명력에 대한 중복도를 줄일 수 있다. 또한 종속변인(오답률)과 독립변인 사이의 상관도는 1에 가까울수록 밀접히 관련되어 있다고 할 수 있으며, 문항 난이도를 예측할 때, 매우 중요한 역할을 한다고 설명된다.

그러므로 독립변인 사이의 상관도는 0에 가까울수록, 그리고 종속변인과 독립변인 사이의 상관도는 1에 가까울수록 좋다고 할 수 있다.

먼저 각 연도별 오답률과 독립변인 사이의 상관도를 조사하면 다음과 같다.

변인 v1, v4, v7, v8, v9은 연도에 관계없이 매우 높게 나타났다. v3은 2003 수능에서 음의 상관도를 보이고 있으며, v6은 모든 연도에서 음의 상관도를 보이고 있다. 결국 예측변인 v3과 v6은 문항 난이도에 대한 설명력이 낮을 뿐만 아니라 내용 분석에 의하여 논리적으로 기대할 수 있는 방향으로 작용하지 않기 때문

에 독립변인 군에서 제외하였다.

나. 빈도 분석 결과

9개의 독립변인은 각각 2개에서 9개의 응답치를 가지며, 각 응답자의 번호가 커질수록 어려워진다고 해석되도록 검토 도구를 개발하였다. 각 독립변수가 문항 난이도 예측에 공헌할 수 있으려면, 기대한 대로 기능해야 할 뿐만 아니라 각 응답지에 반응빈도가 충분히 분산되어 있어야 한다. 전자는 앞 절의 상관분석에서 검토되었으며, 후자 즉 반응수준별 반응빈도의 분산은 이 절에서 검토하고자 하는 바이다. 2002년도 수능과 2003년도 수능 자료를 3명의 검토자가 부호화한 결과 v2의 1번 척도에 대한 비율이 76.7%로 다른 척도에 비해 매우 높게 나타났으며, v4의 3번 척도가 11.7%로 다른 척도에 비해 매우 낮게 나타났기 때문에 척도를 약간 수정할 필요가 있었다.

또한 v1의 각 척도에 대한 비율이 적절하게 분산되어 있기 때문에 척도의 수를 줄이더라도 큰 문제가 없을 것으로 판단되어 이들 측정문항에 대한 척도를 다소 수정하였다. 전체 측정문항 9문항에 대한 α 계수는 0.5150, v3과 v6을 제외한 7개 문항에 대한 α 계수는 0.5729로 나타났다. 1차 예측 모형에 대한 결과는 부록에 있으며, 수정된 검토 도구는 다음 절에 제시되었다.

<표 II-8> 오답률과 독립변인간의 상관도

연도	오답률	V1	V2	V3	V4	V5	V6	V7	V8	V9
2002 수능	1	.473**	.193	.167	.413**	.141	-.137	.397**	.460**	.364**
2003 수능	1	.565**	.307**	-.122	.452**	.114	-.227*	.507**	.450**	.305**
통합	1	.523**	.244**	.005	.433**	.127	-.183*	.456**	.457**	.332**

4. 2차 검토 도구

앞에서 논의된 결과를 바탕으로 측정 문항을 수정하고 보완하여 2차 검토 도구를 <표 II-9>와 같이 구성하였다.

<표 II-10>은 2002년 수능과 2003년 수능을 통합하여 오답률 및 7개의 예측변인 사이의 상관도를 나타낸 것이다. 7개의 예측변인 각각에 대하여 오답률과의 상관도를 살펴보면 유의수준 5% 범위에서 7개의 변인 모두 오답률과 상관관계가 있음을 보여주고 있으며, 유의수준 1% 범위에서 MII-2를 제외한 나머지 6개의 변

인 모두 오답률과 상관관계가 있음을 보여주고 있다. 오답률과 예측문항 MII-6의 상관도는 0.509으로 가장 높으며, 가장 낮은 상관도를 보이는 예측문항은 MII-2로 그 상관도는 0.226이다. 또한 독립변인 가운데 내용 영역과 답지의 형식은 다른 독립변인과는 서로 독립적인 것으로 나타났다. 그러나 제재의 생소성, 문제의 복잡도, 계산의 복잡도, 행동 영역 사이의 상관도는 유의한 수준으로 나타나고 있어서 서로 종속적인 요소들이 중첩되어 있다는 것을 알 수 있다.

그 원인에 대해서는 후속연구에서 보다 세밀

<표 II-9> 2차 검토도구

분류	번호	예측변인	척도
교사의 검토가 필요한 변인	MII-1	제재의 생소성	4단계 척도
	MII-2	문제의 복잡도	4단계 척도
	MII-3	계산의 복잡도	4단계 척도
	MII-4	보조적 자료	5단계 척도
객관적인 변인	MII-5	내용영역	5단계 척도
	MII-6	행동영역	7단계 척도
	MII-7	답지의 형식	3단계 척도

<표 II-10> 제2차 예측모형 변인과 2002·2003학년도 수리영역 실제 오답률 간 상관

오답률	MII-1	MII-2	MII-3	MII-4	MII-5	MII-6	MII-7	
오답률	1							
MII-1	.442**	1						
MII-2	.226*	.263**	1					
MII-3	.384**	.477**	.419**	1				
MII-4	.305**	.533**	.200*	.359**	1			
MII-5	.456**	.118	-.057	.129	.107	1		
MII-6	.509**	.379**	.412**	.394**	.315**	.285**	1	
MII-7	.332**	-.049	-.089	.111	-.025	.086	-.043	1

*p<.05 **p<.01

한 분석이 이루어질 필요가 있지만, 본 연구자의 판단으로는 현행 수능 문항의 특성과 밀접한 관련이 있을 것으로 생각된다. 예컨대, 문항이 생소하다면 개념의 복잡도는 증가할 가능성이 높으며 계산의 복잡도도 증가할 가능성이 높다. 또한 생소한 문항은 이해 영역에서 복잡한 이해나 문제해결의 문항으로 분류될 가능성이 높아지게 된다. 그러나 이러한 가능성은 개연성이 높다는 것이지 필연적으로 반드시 그렇게 된다는 것은 아니다. 결국 후속연구에서는 기출 문항을 보다 면밀하게 분석함으로써 증점된 부분을 제거할 수 있도록 연구의 방향을 정해야 할 것이다.

III. 최적 예측 모형 탐색

2장의 검토도구를 바탕으로 비표준화 회귀계수를 구한 결과, MII-4의 비표준화 회귀계수는 음수로 나타났다. 비표준화 회귀계수란 추정회귀식을 구하였을 때, 각 변인의 계수를 의미한다. 변인의 계수가 양수가 된다면 오답률을 추정할 때 그 변인은 설명력이 있다고 판단되지만, 변인의 계수가 음수가 된다면 오답률을 추정할 때 설명력이 다른 변인과 중복된다는 것을 의미한다. 결국, MII-4는 최종적으로 선택

된 난이도 예측 변인에서 제외하였으며, 최종 검토도구는 <표 III-1>과 같다. 교사의 검토가 필요한 변인은 제재의 생소성, 문제의 복잡도, 계산의 복잡도로 구성되어 있으며 그 척도는 각각 4단계, 4단계, 3단계이다. 객관적인 변인은 내용 영역, 행동 영역, 답지의 형식으로 구성되어 있으며 그 척도는 각각 5단계, 7단계, 3단계이다.

본 장에서는 최종적으로 선택된 검토도구에 대해 여러 가지 모형을 적용하고 회귀분석 결과, 예측 오답률과 실제 오답률 사이의 상관도, 예측 평균과 실제 평균 사이의 관계 등을 이용하여 최적화된 예측 모형을 탐색하고자 한다.

가. 회귀분석

<표 III-2>는 최종적으로 선정된 여섯 개의 수리 영역 난이도 관련변인으로 2002 수능 회귀방정식 모형, 2003 수능 회귀방정식 모형, 통합 구조방정식 모형, 통합 회귀방정식 모형 등 네 가지 예측 함수를 만들어서 네 모형에 대한 회귀분석 결과를 나타낸 것이다.

B계수는 비표준화 회귀계수를 의미하는 것으로, 각 문항의 오답률을 계산하는데 이용된다. S.E는 분산을 나타내며, β 계수는 표준화 회귀계수를 의미한다. 표준화 회귀계수가 높다는 것은 오답률에 대한 설명력이 높다는 것을

<표 III-1> 최종 검토도구

분류	독립변인	예측변인	척도
교사의 검토가 필요한 변인	MII-1 (1)	제재의 생소성	4단계 척도
	MII-2 (2)	문제의 복잡도	4단계 척도
	MII-3 (3)	계산의 복잡도	4단계 척도
객관적인 변인	MII-5 (4)	내용영역	5단계 척도
	MII-6 (5)	행동영역	7단계 척도
	MII-7 (6)	답지의 형식	3단계 척도

의미한다. 변인별로 본다면 통합 회귀방정식 모형의 경우 '답지의 형식'이 가장 설명력이 컸고 '내용 영역', '제재의 생소성', '행동 영역', '문제 해결에 필요한 개념의 수', '계산의 복잡성 정도' 등이 그 뒤를 따랐다.

R^2 는 6개의 측정문항에 의한 설명력을 나타내는 것으로 2002 수능 회귀방정식 모형이 0.586, 2003 수능 회귀방정식 모형이 0.606, 통합 구조방정식 모형이 0.550, 통합 회귀방정식 모형이 0.550으로 비교적 높게 나타나고 있다.

바꾸어서 말하면 6개의 예측변인을 이용한 모형의 설명력이 높게 나타나고 있어서 비교적 정확한 결과를 산출할 것으로 기대된다는 것을 의미한다.

나. 예측 오답률·실제 오답률 간 상관
<표 III-3>은 수리 영역의 네 모형에 대한 중다상관 및 예측·실제 오답률 사이의 상관도를 나타낸 것이다. 수리 영역의 2002수능 회귀방정식 모형의 경우 중다상관은 .765였으나 그

<표 III-2> 수리 영역의 네 모형에 대한 회귀분석 결과

모형	변인	B계수	S.E.	β 계수	t	R^2
2002수능 회귀방정식 모형	상수	-22.879	10.065		-2.273 [*]	.586
	MII-1 (1)	2.970	2.610	.119	1.138	
	MII-2 (2)	2.598	3.805	.078	.683	
	MII-3 (3)	.289	3.536	.010	.082	
	MII-5 (4)	4.999	1.877	.252	2.664 ^{**}	
	MII-6 (5)	5.464	1.257	.479	4.348 ^{**}	
	MII-7 (6)	13.389	2.786	.471	4.806 ^{**}	
2003수능 회귀방정식 모형	상수	-38.577	10.650		-3.622 ^{**}	.606
	MII-1 (1)	13.607	3.207	.450	4.243 ^{**}	
	MII-2 (2)	5.348	4.586	.114	1.166	
	MII-3 (3)	-1.009	4.038	-.027	-.250	
	MII-5 (4)	9.219	2.119	.404	4.350 ^{**}	
	MII-6 (5)	1.875	1.409	.142	1.331	
	MII-7 (6)	7.617	2.715	.247	2.806 ^{**}	
통합 구조방정식 모형	상수	-26.437	7.098		-3.725 ^{**}	.550
	MII-1 (1)	7.350	1.970	.249	3.731 ^{**}	
	MII-2 (2)	2.559	2.819	.056	0.908	
	MII-3 (3)	1.059	2.518	.029	0.421	
	MII-5 (4)	6.399	1.372	.287	4.666 ^{**}	
	MII-6 (5)	3.645	0.917	.283	3.974 ^{**}	
	MII-7 (6)	10.112	1.871	.335	5.404 ^{**}	
통합 회귀방정식 모형	상수	-26.981	7.330		-3.681 ^{**}	.550
	MII-1 (1)	7.702	2.045	.280	3.767 ^{**}	
	MII-2 (2)	2.658	2.924	.068	.909	
	MII-3 (3)	1.015	2.606	.031	.390	
	MII-5 (4)	6.544	1.420	.312	4.609 ^{**}	
	MII-6 (5)	3.523	.949	.287	3.711 ^{**}	
	MII-7 (6)	9.930	1.927	.334	5.153 ^{**}	

^{*}p<.05

^{**}p<.01

회귀계수를 2003학년도 수능 및 2004학년도 9월 모의평가에 그대로 적용했을 때 예측 오답률·실제 오답률 간의 상관은 각각 .662와 .717이었다. 2003수능 회귀방정식 모형의 경우 중다상관은 .779였으나 그 회귀계수를 2002학년도 수능 및 2004학년도 9월 모의평가에 그대로 적용했을 때 예측 오답률·실제 오답률 간의 상관은 각각 .648과 .764였다. 통합 구조방정식 모형과 통합 회귀방정식 모형 경우에도 중다상관은 모두 .742였으나 그 회귀계수를 2004학년도 9월 모의평가에 적용한 결과 예측 오답률·실제 오답률 간의 상관은 각각 .789와 .791로써 거의 변화가 없음을 알 수 있었다.

요컨대 수리 영역의 경우 다른 수능 문항에 대한 설명력을 근거로 네 가지 예측 모형을 평가한다면 '통합 회귀방정식 모형'이 가장 안정적인 모형이라고 볼 수 있었다.

다. 예측 평균과 실제 평균

<표 III-4>는 수리 영역의 네 모형에 대한 예측 평균과 실제 평균을 나타낸 것이다. 수리 영역의 경우 2002학년도 수능, 2003학년도 수능, 2004학년도 9월 모의평가의 실제 평균은 각각 50.88, 54.11, 51.74였다. 2002수능 회귀방정식 모형으로 예측한 경우 2003학년도 수능과 2004학년도 9월 모의평가의 예측 평균은 54.89와 55.93이었다. 2003수능 회귀방정식 모형으로 예측한 경우 2002학년도 수능과 2004학년도 9월 모의평가의 예측 평균은 50.08과 55.32이었다. 이에 비해서 통합 구조방정식 모형과 통합 회귀방정식 모형으로 예측한 경우 2004학년도 9월 모의평가의 예측 평균은 각각 55.31과 55.28이었다.

요컨대 수리 영역의 경우 실제 평균과의 차이를 근거로 네 가지 예측 모형을 평가한다면

<표 III-3> 수리 영역의 네 모형에 대한 중다상관 및 예측·실제 오답률 간 상관

모형	2002년	2003년	2004년 9월	전체
2002수능 회귀방정식 모형	.765	.662	.717	.711
2003수능 회귀방정식 모형	.648	.779	.764	.713
통합 구조방정식 모형	.737	.744	.789	.742
통합 회귀방정식 모형	.732	.749	.791	.742

음영 : 각 모형의 중다상관

<표 III-4> 수리 영역의 네 모형에 대한 예측 평균과 실제 평균

모형	2002년		2003년		2004년 9월		전체	
	\bar{y}	\bar{y}'	\bar{y}	\bar{y}'	\bar{y}	\bar{y}'	\bar{y}	\bar{y}'
2002수능 회귀방정식 모형		50.88		54.89		55.93		55.88
2003수능 회귀방정식 모형	50.88	50.08	54.11	54.11	51.74	55.32	52.50	52.10
통합 구조방정식 모형		50.51		54.54		55.31		52.52
통합 회귀방정식 모형		50.48		54.51		55.28		52.50

\bar{y} : 실제 평균

\bar{y}' : 모형 예측 평균

모형 간 차이가 거의 없었지만 '통합 회귀방정식 모형'이 가장 예측력이 높다고 볼 수 있었다.

라. 수리 영역 최종 예측 모형

이상의 결과를 토대로 수리 영역에서는 2002 수능 회귀방정식 모형, 2003 수능 회귀방정식 모형, 통합 구조방정식 모형, 통합 회귀방정식 모형 중에서 모형의 안정성과 예측의 정확성 측면에서 가장 우수하다고 판단할 수 있는 모형은 '통합 회귀방정식 모형'으로 볼 수 있다. 이를 근거로 최종 예측 모형을 제안하면 <표 III-5>와 같다.

IV. 결론 및 제언

수능과 같은 사회적으로 관심도가 높은 시험에서는 난이도 수준의 안정성과 적정성이 매우 중요하기 때문에 수능 출제 과정에서 수능 난이도를 예측하기 위한 노력이 이루어지고 있다. 그러나 이제까지의 수능 출제 과정에서 난이도 예측은 수능 출제위원과 검토위원들의 전문적 판단에 전적으로 의존하고 있기 때문에 객관성이 떨어지는 문제점을 안고 있다. 이에 본 연구는 전문가의 주관적이고 통합적인 판단에 의존하던 기존의 통합적 접근을 보완하

여 보다 객관적이고 분석적이며, 보다 과학적인 방법으로 난이도를 예측할 수 있는 토대를 마련하려는 것이다.

본 연구에서는 문헌 분석과 기출 수능 문항에 대한 내용과 관련문헌을 조사하여 난이도 관련 변인을 탐색하였으며, 이에 대하여 수능 출제 경험이 많은 있는 교수와 교사가 참여한 전문가 협의회에서 검토를 받았다.

이를 바탕으로 난이도 관련변인을 몇 개의 구성변인으로 나누어 검토도구를 제작하였다. 검토도구 타당화는 두 단계로 이루어졌다. 먼저 1차 검토도구를 제작하여 2002·2003학년도 수능 문항에 적용해 봄으로써 난이도에 대한 설명력이 낮은 변인을 제외하였다. 1차 검토도구의 문제점을 보완한 2차 검토도구도 1차 검토도구와 동일한 절차를 밟으면서 난이도 관련 변인 검토도구를 타당화하였다.

즉, 1차 검토도구 중에서 난이도에 대한 설명력이 높은 변인을 이용하여 제1차 예측모형을 만들었고, 2차 검토도구에서도 마찬가지로 제2차 예측모형을 만들었다. 예측모형은 중다 회귀방정식과 구조방정식을 활용한 것으로서 2002수능 회귀방정식 모형, 2003수능 회귀방정식 모형, 통합 구조방정식 모형, 통합 회귀방정식 모형 등 4개의 모형 중에서 안정성과 정확성이 가장 높은 모형을 최종 예측모형

<표 III-1> 수리 영역의 최종 예측 모형

수리 오답 백분율 추정치	= 7.702·(내용이나 제재의 생소성 정도) + 2.658·(문제 해결에 필요한 개념의 수) + 1.015·(계산의 복잡성 정도) + 6.544·(내용 영역) + 3.523·(행동 영역) + 9.930·(문항의 형식) - 26.981
---------------------	---

으로 제안하였다.

그 결과를 요약하면 다음과 같다.

첫째, 최종 예측 모형에서 사용된 예측 변인으로는 제재의 생소성, 문제의 복잡도, 계산의 복잡도, 내용 영역, 행동 영역, 답지의 형식 등이 있다. 각 변인은 오답률과의 상관도가 유의한 것으로 나타났으며, 내용 영역과 답지의 형식은 다른 독립변인과는 서로 독립적인 것으로 나타났다. 그러나 제재의 생소성, 문제의 복잡도, 계산의 복잡도, 행동 영역 사이의 상관도는 유의한 수준으로서 서로 종속적인 요소들이 중첩되어 있는 것으로 나타났다. 그 원인은 현행 수능 문항의 특성과 밀접한 관련이 있을 것으로 판단되며, 후속연구에서 그 원인에 대한 세밀한 분석이 요구된다.

둘째, <표 III-2>에서 오답률에 대한 설명력을 알 수 있는 표준화 회귀계수를 살펴보면, 문항의 생소성, 내용영역, 행동영역, 답지형식 등이 높게 나타나고 있다. 특히 통합 회귀방정식 모형을 기준으로 살펴보면, 답지의 형식이 가장 높고, 다음으로 내용 영역, 행동 영역, 문항의 생소성 등의 순으로 나타나고 있다.

셋째, 네 모형에 대한 중다상관 및 예측·실제 오답률 사이의 상관도를 분석한 결과와 예측 평균과 실제 평균 사이의 관계 등을 토대로 네 가지 예측 모형을 평가한다면 수리 영역에서는 '통합 회귀방정식 모형'이 가장 예측력이 높은 모형으로 나타났다.

본 연구는 수리 영역의 난이도 관련 변인에 대한 연구가 거의 전무한 상태에서 이루어졌기 때문에 많은 한계가 있을 것으로 생각되지만 본 연구 결과는 앞으로의 난이도 관련 변인에 대한 연구의 기초를 제공할 수 있을 것이다.

본 연구의 한계 및 제안점은 다음과 같다.

첫째, 난이도를 설명하거나 결정하는 더 많

은 변인을 찾아내고 이를 분류하는 노력이 필요하다. 이론적 분석 결과 외에도 요인분석 등의 방법을 이용함으로써 난이도 관련변인이 어떠한 요인으로 구성되어 있는지 탐색할 필요도 있다.

둘째, 실제 난이도와 난이도 관련 변인 간의 관계에 대한 추가적인 분석이 필요하다. 본 연구에서는 난이도 관련변인을 선정할 때 실제 난이도와 피어슨 상관 계수를 중요한 기준으로 활용하였다. 피어슨 상관은 변인 간 관계가 직선적이며 두 변인 모두 동간변인 또는 비율 변인이어야 한다는 가정이 필요하다. 그러나 본 연구에서는 독립변인과 종속변인 간 관계가 직선적이라는 가정 하에 분석을 하였기 때문에 실제 상관을 과소추정했을 가능성이 있다.

셋째, 난이도 관련 변인으로 학생에 대한 변인을 고려할 필요가 있다. 본 연구에서는 매해 수험생 집단의 동질성을 가정하고 있다. 그러나 해마다 수험생의 특성이 바뀔 수 있다. 다만 현행 수능의 특성 때문에 매해 수험생의 특성을 객관적으로 확인할 수 있는 장치가 마련되어 있지 않다는 문제점이 지적될 수 있다.

넷째, 본 연구에서는 개별 문항에 대한 난이도를 파악하는데 주안점을 둔 것이 아니라, 수리 영역 문항 전체에 대한 난이도를 파악하는데 주안점을 두고 있다.

그러므로 개별 문항에 대한 실제와 예측 난이도에 대한 차이는 다소 크게 나타날 수 있다. 그러나 수능에서 난이도를 조절하기 위해서는 개별 문항에 대한 난이도 정보가 더 유용할 수 있다. 그러므로 개별 문항에 대한 난이도를 보다 정확하게 추정할 수 있는 연구 방법이 모색될 필요가 있다.

위와 같은 한계를 고려하여 후속연구가 이루어지길 기대한다.

참고문헌

- 강시중(1993). 수학교육론. 서울: 교육출판사.
- 강육기·정은실·박교식·강문봉(1989). 수학적 사고력 신장 프로그램 개발을 위한 방안 탐색 - 국민학교 산수과를 중심으로 -. 서울: 한국교육개발원.
- 강완·백석윤(1998). 초등수학교육론. 서울: 동명사.
- 교육부(1995). 고등 학교 수학과 교육 과정 해설(교육부 고시 제1999-19호('92. 10. 30.)). 서울: 대한교과서 주식회사.
- 김계수(2001). AMOS 구조방정식 모형분석. 서울: 고려정보산업(주)
- 우정호(1992). 수리·탐구영역과 교수-학습방향 - 수리분야 -. 대학수학능력시험과 교수-학습방향. 전국교육평가심포지엄보고서 제9집 92-2. 서울: 국립교육평가원.
- 우정호(1998). 학교수학의 교육적 기초. 서울: 서울대학교출판부.
- Boodoo, G., Tatsuoka, K. & Yan, D. (1997). *Investigating the relationships between the verbal and the mathematics attributes on the SAT-preliminary results*. Paper presented at the annual meeting of National Council on Measurement in Education, Chicago, Illinois, 1997.
- Travers K. J. & Westbury I. (1989). *The IEA Study of Mathematics I: Analysis of Mathematics Curricula*. Pergamon press.
- Freedle R. & Kostin I. (1991). *The Prediction of SAT Reading Comprehension Item Difficulty for Expository Prose Passages*. Educational Testing Service. Princeton, New Jersey.

A Study of Variables Related to Item Difficulty in College Scholastic Ability Test

Park, Moon Hwan (KICE)

The purpose of this study was to examine particular variables that play a significant role in the difficulty of math test items in College Scholastic Ability Test (CSAT). The study also aimed to develop a model of measuring the item difficulty. Variables correlated to item difficulty were drawn from the review of the related literature and the analysis of the content and difficulty of the past test items of CSAT. The first instrument was designed by using the correlated variables. According to the results of correlation analysis, the second instrument was made by deleting the variables which showed relatively low correlation with item

difficulty and by refining some variables.

Several models were proposed by using the revised instrument. The comparison of the R square and cross validity of each model reveals that integrated regression model was the most stable and accurate among the proposed models. The study also showed that statistically significant predictors were choice format, content domain, behavior domain, and the degree of item familiarity in the order of proportion of variance accounted by the predictors. Despite the limited scope of the present research, it can be suggested that its findings provide useful insights into predicting math test item difficulty.

* **Key word:** item difficulty(문항 난이도), College Scholastic Ability Test(대학수학능력시험), correlation coefficient(상관계수), regression analysis(회귀분석), regression equation model(회귀방정식모형)

<부록 1> 1차 예측모형 분석 결과

모형	변인	비 표준화 회귀계수	S.E.	표준화 회귀계수	t	R ²
2002모형	상수	-27.633	8.158	-	-3.387**	.588
	V1	2.086	1.183	.183	1.764	
	V2	-1.487	3.577	-.035	-.416	
	V4	4.526	3.069	.139	1.475	
	V5	2.789	1.933	.112	1.442	
	V7	5.749	1.435	.290	4.007**	
	V8	3.904	1.057	.346	3.695**	
	V9	12.505	2.161	.440	5.788**	
2003모형	상수	-36.474	8.490	-	-4.296**	.608
	V1	5.414	1.047	.431	5.172**	
	V2	6.833	4.986	.103	1.370	
	V4	.911	3.480	.023	.262	
	V5	1.210	1.920	.048	.630	
	V7	7.835	1.700	.343	4.608**	
	V8	1.767	1.055	.135	1.675	
	V9	8.299	2.392	.269	3.469**	
LISREL 통합모형	상수	-29.164	5.640	-	-5.171**	.570 (2002) .552 (2003)
	V1	4.094	.754	.335	5.431**	
	V2	1.258	2.799	.019	.450	
	V4	2.393	2.217	.061	1.079	
	V5	1.866	1.305	.076	1.430	
	V7	6.407	1.060	.288	6.044**	
	V8	2.341	.710	.183	3.298**	
V9	10.418	1.553	.347	6.709**		
회귀분석 통합모형	상수	-29.389	5.774	-	-5.090**	.569
	V1	4.188	.768	.352	5.450**	
	V2	1.460	2.890	.029	.505	
	V4	2.290	2.277	.064	1.005	
	V5	1.827	1.335	.073	1.369	
	V7	6.471	1.090	.308	5.937**	
	V8	2.278	.726	.187	3.138**	
V9	10.300	1.593	.346	6.464**		

*p<.05 **p<.01