

碩士學位論文

共通科學教科 總括評價의 問項分析

指導教授 金 奎 用



110112

濟州大學校 教育大學院

物 理 教 育 專 攻

吳 文 植

2000年 8月

共通科學教科 總括評價의 問項分析

指導教授 金 奎 用

이 論文을 教育學 碩士學位 論文으로 提出함.

2000 年 6 月 日

濟州大學校 教育大學院 物理教育專攻

JEJU NATIONAL UNIVERSITY LIBRARY

提出者 吳 文 植

吳文植의 教育學 碩士學位論文을 認准함.

2000 年 7 月 日

審査委員長 _____ (印)

審査委員 _____ (印)

審査委員 _____ (印)

共通科學教科 總括評價의 問項 分析

吳 文 植

濟州大學校 教育大學院 物理教育專攻

指導教授 金 奎 用

본 연구에서는 ASC의 ITEMAN과 XCALIBRE 프로그램을 이용하여 고등학교 공통과학교과의 총괄평가 검사와 문항의 특성을 분석하였다.

연구자료는 제주도내 7 개 고등학교의 공통과학 교과에서 실시했던 중간고사와 기말고사의 응답 자료이며, 분석에는 18개 검사, 400문항을 사용하여, 다음의 결과를 얻었다.

첫째, 고전검사이론에 의하면 쉬운 문항이 어려운 문항보다 많이 출제되고 있고, 변별력이 높은 문항이 전체의 55%를 차지하고 있다.

둘째, 문항반응이론의 3-모수 로지스틱 모형에 의하면 75.2%가 적절한 변별도지수를 갖는 문항들을 출제하고 있어, 상·하위집단을 매우 잘 변별해 주고 있으며, 또한 문항난이도는 매우 쉬운 문항에서 어려운 문항까지 정규분포를 이루고 있으며, 문항추측도는 0.2 이상과 0.2이하의 문항이 각각 41.0, 59.0%의 비율로 출제되고 있는 것으로 나타났다.

셋째, 검사에 대해 문항반응이론의 기본 가정인 단일차원성의 가정은 어느 정도는 만족하고 있는 것으로 나타났다.

넷째, 고전검사이론에 의한 검사점수와 문항반응이론에 의한 능력 모수 사이의 상관은 매우 높은 상관을 보이고 있다.

다섯째, 인문·실업계 고등학교 사이에 검사 문항의 특성은 문항추측도를 제외하고 통계적으로 유의미한 차이가 있는 것으로 나타났다.

* 본 논문은 2000년 8월 제주대학교 교육대학원 위원회에 제출된 교육학 석사 학위 논문임.

目 次

초 록	i
I. 序 論	1
II. 理論的 背景	4
1. 教育評價의 概念	4
2. 古典檢査理論	6
3. 問項反應理論	10
4. 問項分析을 위한 컴퓨터 프로그램	28
III. 研究의 對象 및 方法	31
1. 研究의 對象	31
2. 研究의 方法	31
IV. 研究의 結果	34
1. 問項特性的 推定 結果 및 解析	34
2. 檢査 特性的 分析	38
3. 問項反應理論의 適用可能性 分析	40
4. 古典檢査理論에 의한 檢査點數와 3-母數 로지스틱 模型에 의한 能力 母數 推定值의 相關關係	41
5. 學校 類型別 問項 特性的 比較	43
V. 結 論	47
參考文獻	49
<Abstract>	52
附 錄	54

I. 序 論

학교의 교육활동은 교과내용, 교수-학습, 평가의 세 가지 요소로 구성되어 있는데, 평가는 나머지 두 요소의 함수이면서 동시에 두 요소에 영향을 미치는 변수이다.

학생들은 평가를 통해 학습 결과에 대해 관심을 가지며, 학생들의 학습은 장려된다. 어떤 교과목의 평가는 학교에서 학생들이 학습해야 할 내용을 곧 평가내용으로 삼고 있어 학습과 평가는 항상 밀접한 관련을 맺고 있으며, 학습자에게 바람직한 학습 결과를 가져올 수 있도록 유도하는 활동이다.

학습 결과에 대한 평가는 학생들의 학습에 대한 오류를 교정할 수 있는 방향을 제시해 주고, 교사의 교수-학습 방법 개선의 방향을 제시하여 준다. 한편 학생들의 평가의 결과가 진행되어졌던 수업을 재고하게 되고 더 나은 수업방법을 찾아내게 하여 앞으로의 수업방향을 바로 잡을 수 있는 근거를 제공하게 된다.

오늘날 학교 현장은 학생들의 능력을 평가하기 위하여 정기적으로 시험이나 검사가 시행되고 있다. 이 때 사용하는 평가도구가 얼마만큼의 특성을 지녔는가 하는 것을 비교하고 고려하는 것은 학생들의 교육내용을 이해하고 더 나은 교수-학습 방법을 익히는데 중요하다.

검사나 시험문제가 갖는 난이도, 변별도 등의 특성을 올바르게 평가하고 이 특성을 감안한 평가도구의 구안이나 적용은 필수 불가결함에도 불구하고 이를 평가하고자 하는 데는 많은 문제점들을 내포하고 있다 하겠다. 효율적이고 가치 있는 평가가 이루어지기 위해서는 검사 또는 문항마다의 난이도, 변별도 등의 특성을 고려하여 피험자의 능력을 평가하는 것이 올바른 것이며, 이를 위한 검사의 타당도, 신뢰도, 그리고 문항의 특성이 논리적으로 올바른 것이어야 한다.

일반적으로 하나의 검사는 개개의 문항들로 이루어지며, 문항은 각각 난이

도, 변별도 등의 특성을 지닌다. 이러한 문항의 특성을 분석하는 이론으로 검사문항을 유사한 피험자들에게 여러 번 반복 시행함으로써 학생들의 능력을 나타내는 진점수를 산출할 수 있는 古典檢査理論(classical test theory)과, 검사의 특성과 측정하고자 하는 피험자의 특성을 미리 예측할 수 있는 問項反應理論(item response theory)을 들 수 있다.

고전검사이론은 과거 수 십 년간 학교현장에서 적용되어져 왔고, 평가 문항에 대한 특성의 분석, 그리고 피험자의 능력을 나타내는 진점수의 산출 등에 활용되어져 왔다. 그러나 검사를 시행한 적이 없는 경우 피험자들이 그 문항에 어떻게 반응할 것인지를 예측하지 못하고, 문항이나 검사의 통계치가 피험자의 집단에 따라 다르게 추정된다는 약점이 있고, 피험자의 능력을 추정할 때 점수오차를 동일하게 적용하는 적절치 못한 개념을 갖고 있다¹⁾.

김석우²⁾는 학생들의 학습능력을 올바르게 평가하기 위해서는 논리적으로 타당성을 지닌 문항반응이론 적용의 필요성이 점차 인식되고 있음을 강조하고 있다. 문항반응이론은 사전에 유사한 피험자에게 검사를 시행하지 않고도 검사문항에 대한 피험자의 반응을 확률적으로 예측할 수 있다. 이미 시행한 검사문항의 모수치가 주어진 문항풀(pool)에서 문항모수치를 기준으로 다음에 시행할 검사의 문항을 선택하는 경우 피험자의 능력을 근사적으로 추정할 수 있다³⁾.

문항반응이론은 1980년대 이래로 국내의 여러 검사의 문항분석 연구에 적용되어 그 타당성을 입증하여 왔다. 그러나 국내에서 시행되었던 선행 연구는 대학입학 학력고사 및 대학입학수학능력시험의 문항에 대한 연구가 대다수이고, 영어 및 수학 교과에 대한 분석과 심리측정분야에 대한 연구가 진행되었으나, 과학 교과 특히 공통과학 교과에 대해 문항반응이론을 적용한 문항분석 연구는 거의 없는 실정이다. 특히 학교에서 실시되었던 총괄평가 문항에 대한 분석 연구는 드물다. 더욱이 문항반응이론에 대한 문항분석용 컴퓨터 프로그램으로 LOGIST와 BILOG, 또는 MULTILOG를 이용한 문항의 분석에 대한 연구는 많이 있으나, Assessment Systems Corporation사

(ASC)의 Item and Test Analysis Package를 이용한 문항분석 연구는 거의 없는 실정이다.

따라서 본 연구에서는 공통과학 교과목의 총괄평가 검사와 문항에 대해 ASC Item and Test Analysis Package의 고전검사이론을 적용할 수 있는 ITEMAN 프로그램과 문항반응이론을 적용할 수 있는 XCALIBRE 프로그램을 이용하여 다음의 연구를 수행하고자 한다.

첫째, 이론적 배경을 통해 고전검사이론과 문항반응이론을 고찰하며,

둘째, 고등학교 공통과학 교과목의 총괄평가에서 시행되었던 학생의 응답 자료를 이용, 고전검사이론과 문항반응이론의 3-모수 로지스틱 모형에 의해 추정된 문항난이도, 문항변별도 및 문항추측도의 통계치를 알아보고,

셋째, 고전검사이론과 문항반응이론의 3-모수 로지스틱 모형에 의해 총괄평가 검사의 특성인 검사의 신뢰도 및 정보량의 관계를 알아보며,

넷째, 문항반응이론의 적용가능성을 알아보기 위하여 학생들의 검사결과인 응답자료의 단일차원성의 만족여부를 탐색하며,

다섯째, 고전검사이론에 의하여 측정된 학생들의 검사점수와 문항반응이론의 3-모수 로지스틱 모형에 의한 피험자의 능력 추정치사이의 상관관계를 분석하며,

여섯째, 고전검사이론에 의해 계산되고, 문항반응이론의 3-모수 로지스틱 모형에 의해 추정된 문항난이도, 문항변별도 및 문항추측도의 평균값이 학교형태(인문, 실업)에 따라 차이가 있는지를 알아보고자 한다.

II. 理論的 背景

1. 總括評價의 概念

황정규⁴⁾는 교육평가를 ‘교수 프로그램이 교육효과에 관한 의사결정을 하기 위해서 학습자의 행동변화 및 학습과정에 관한 정보를 수집하고 이용하며 교육적 의사결정을 내리는데 도움을 주는 과정’이라고 정의하고 있다. 이러한 시각에서 교육평가는 미래의 형성적 목적에 이바지해야 하며, 인간을 이해하는 평가가 되어야 하고, 그 과정은 언제나 종합적 시각 위에서 실마리를 찾아야 한다.

Tyler⁴⁾는 교육의 과정을 교육목표의 선정과 세분화, 학습경험의 선정과 조직, 교수-학습과정의 투입, 교육평가의 과정으로 요약하고 교육평가 자체는 이들 모두에 기여하고 이바지하는 바가 크다고 하였다.

교수-학습과정의 여러 단계와 측면에서 교육평가는 그것 자체가 교수방법으로 이용되기도 하고, 또 전략의 선택 및 적정화를 위해 필요한 증거를 제공하는데 도움을 준다. 수업과정의 초기단계에서의 診斷評價, 교수-학습과정에 관련해서는 形成評價, 그리고 교육목표의 달성, 즉 성취수준 도달여부를 판단하는 總括評價의 세 가지 형태의 평가가 필요하다.

여기에서 형성평가와 총괄평가는 개념상으로 다를 뿐 아니라 평가에 필요한 자료수집의 계획이나 자료분석 방법에서도 다르게 된다.

성태제⁵⁾는 총괄평가를 총합평가(summative evaluation)라고 하고 ‘교수-학습이 끝난 다음 교수목표의 달성, 성취 여부를 종합적으로 판단하는 평가 형태’를 말하고 있다.

Scriven⁶⁾은 총괄평가를 형성평가와 구분하여 ‘教育課程의 總體(curriculum package)가 충분할 만큼의 중요한 교육적 진보를 가져왔는지를 확인하는 행위’라 하였다.

Bloom⁷⁾ 등은 총괄평가를 ‘학과목, 학기, 그리고 교육프로그램의 끝에 실시하는 평가로 성취 혹은 숙달정도와 교육목표 획득여부를 결정하는 행위’라 정의하였다.

총괄평가는 자격증 부여, 당락결정, 점수에 의한 서열화, 그리고 집단 비교 등의 목적을 지니고 있으므로 교사 이외에 교과전문가, 교육과정 전문가, 교육평가 전문가들이 공동으로 참여하여 실시하고 있다. 대표적인 예로 간호사나 의사자격시험, 고시, 교사임용고사, 그리고 전국단위 학력평가를 위한 검사를 들 수 있으며, 학교에서 실시하는 중간시험이나 기말시험 등도 이에 포함된다고 할 수 있다.

총괄평가를 위하여 검사를 제작할 경우는 검사의 목적이 어디에 있느냐에 따라 출제 지침이 달라진다. 어떤 준거 혹은 목표가 중심일 경우는 그 준거에 알맞은 수준의 문항들을 제작하여 지식 수준이 목표 혹은 준거에 도달하였느냐를 확인하는 데 중점을 둔다. 그러나 피험자, 학급, 학교 등을 상대 비교에 의하여 서열화하는데 관심을 두는 경우는 매우 쉬운 문항부터 매우 어려운 문항까지 골고루 출제하여야 한다.

이런 관점에서 성태제⁵⁾는 총괄평가는 평가수혜자, 교육평가 전문가, 그리고 교사들에게 익숙해 있어야 하며, 책임성을 부여하기 위하여 총괄평가가 이용된다고 주장하고 있다.

황정규⁴⁾는 총괄평가 문항의 선정기준으로 난이도를 들고 있다. 난이도의 변산은 넓게 퍼져 있는 것이 보통이나 30-70%정도의 문항난이도를 보이는 문항이 대부분을 차지하며, 난이도 평균이 50%정도가 되도록 하는 것이 이상적이라고 하고 있다. 그러나 이는 평가의 유형에 따라 달라짐을 말하고 있다. 중요한 것은 수업이 시작되기 전에 계획한 교수목표에 맞는 모집단을 가장 잘 대표할 수 있도록 표본해야 한다는 것이다.

성태제⁸⁾는 좋은 문항의 수리적 요건으로 문항난이도의 적절성과 신뢰도가 높아야 한다고 하였다.

총괄평가 문항의 구성은 규준지향평가나 혹은 준거지향평가나 하는 평가

관점에 따라 여러 가지 문항 특성들 즉 문항난이도, 문항변별도, 검사의 신뢰도 등을 고려하여야 할 것이다. 규준지향 평가의 경우 특히 문항의 변별도 기준은 필수적이며, 신뢰도 또한 검사가 지녀야 할 요소이다. 신뢰도를 높이기 위한 방법으로는 문항의 표본이 문항의 모집단을 잘 대표하고 있어야 하며, 문항이 수가 많아야 하고, 검사의 길이가 길어야 한다. 즉 검사의 변산도가 크면 클수록 바람직한 것이다. 한편 준거지향평가에서는 변산도가 그리 큰 문제가 되지는 않는다. 또한 문항의 유형에 따라, 선택형 문항인 경우 답지의 매력을 충분히 고려하여야 한다.

2. 古典檢査理論

고전검사이론은 검사도구의 총점에 의하여 분석되는 이론으로, 검사에 의한 관찰점수는 피험자의 실제 능력을 나타내는 진점수와 검사를 시행할 때마다 나타나는 오차점수의 합으로 나타내짐을 가정한다. 또한 피험자의 진점수를 알 수 없기 때문에 이론적으로 동일한 검사를 동일 피험자에게 무한히 반복 실시하여 얻은 점수들의 평균점수로 추정한다.

1) 古典檢査理論의 基本 假定과 問項 特性

고전검사이론은 다섯 가지 기본 가정과 평행검사를 제작할 수 있다는 가정에 기초를 두고 있다⁹⁾.

첫째, 관찰 점수는 진점수와 측정오차의 합이라는 덧셈 모형을 가정한다.

진점수란 동일한 검사를 반복하여 독립적으로 시행할 때 관찰한 점수의 평균이라는 수학적 개념이다.

둘째, 관찰 점수의 기대값이 진점수라는 가정이다. 이는 진점수의 오차의 기대값은 '0'이라는 결론이다.

셋째, 측정 오차와 진점수의 상관은 '0'이라는 가정이다. 이는 진점수의 정의로부터 나오는 귀결이다.

넷째, 측정 오차는 상호간에 상관이 없으며 다른 검사의 진점수와도 상관이 없다고 가정한다.

다섯째, 임의의 두 검사를 실시했을 때 두 검사점수의 평균과 분산이 동일한 평행검사를 제작할 수 있다는 가정이다.

피험자의 진점수를 알 수 없기 때문에 이론적으로 동일검사를 동일 피험자에게 무한히 반복 실시하여 얻은 점수들의 평균점수로 추정한다¹⁰⁾.

(1)問項難易度

문항난이도는 문항이 얼마나 어려운가의 정도를 말하며 문항에 정답을 한 학생의 비율 P로 정의한다. P가 크면 쉬운 문항이고 낮으면 어려운 문항이다⁵⁾.

문항난이도는

$$P = \frac{r}{N} \dots\dots\dots(1)$$

으로 표시한다. 여기서 r는 문항의 정답자 수, N는 전체 피험자수이다.

상태제⁵⁾는 문항난이도 값에 대하여 표 1과 같이 범위별로 표현하고 있다.

표 1. 문항난이도의 범위별 표현

문항난이도	문항난이도의 범위별 표현
0.00 ~ 0.25	어려운 문항
0.25 ~ 0.75	적절한 문항
0.75 ~ 1.00	쉬운 문항

문항난이도는 0.00 ~ 1.00에 이르기까지의 값의 범위를 가지고 있다. 그러나 이는 피험자 집단이 다름에 따라 변한다. 개인에 대한 문항난이도란 존재하지 않으며, 피험자 집단의 표준편차와 평균에 의해 결정되는 통계치이다. 이를 계산하는데는 반드시 각 문항에 대한 정답을 모르고 답하는 추측요인을 고려하여야 하며, 검사에 대해 끝까지 정답을 제시하지 못한 미달 항이나

중간에 비워두고 지나간 항을 고려하여야 한다⁴⁾.

추측에 의한 정답자들을 정답자 수에서 제외하면 교정난이도 P' 는 식 (2) 와 같다.

$$P' = \frac{R - \frac{W}{K-1}}{N - NR} \dots\dots\dots(2)$$

여기서 R는 정답 수, W는 오답 수, K는 문항의 답지의 수, NR는 각 문항의 미달항 학생수이다.

일반적으로 규준지향검사는 난이도가 0.20 ~ 0.80인 문항으로 구성되며 0.20 ~ 0.80 사이에서 정규분포를 하도록 하는 것이 적절하다. 즉 문항의 난이도가 0.50인 것이 가장 많고 차츰 적은 수의 문항이 0.20 ~ 0.80 수준의 난이도를 가지고 평균 난이도가 0.50가 되도록 하는 것이 바람직한 것으로 알려져 있다¹¹⁾.

준거지향 검사일 때는 시험의 목적에 따라 문항의 난이도를 매우 낮게 또는 높게 정할 수 있다.

(2) 問項辨別度

문항타당도의 개념 중에 한 문항이 피험자 능력의 상하를 얼마나 잘 변별하는 능력이 있는냐를 보는 것을 問項辨別度(item discrimination)라고 하고 계산되어 나온 수치를 辨別度 指數(discrimination index :D.I)라고 부른다.

문항이 능력에 따라 피험자를 변별하는 정도를 나타내는 수로서 능력이 높은 피험자가 문항의 답을 맞히고 능력이 낮은 피험자는 문항의 답을 맞히지 못하였다면 이 문항은 피험자를 변별하는 기능을 가진 문항이라 할 수 있다.

이종성⁹⁾은 문항변별도 지수를 문항에 정답을 한 상위 집단에 속하는 피험자의 비율 π_U 과 하위 집단에 속하는 피험자의 비율 π_L 의 차이로서 식 (3) 과 같이 정의하고 있다.

$$D.I = \pi_U - \pi_L \dots\dots\dots(3)$$

Johnson⁴⁾은 변별도 지수를 식(4)와 같이 나타내고 있다.

$$D.I = \frac{R_U - R_L}{f} \dots\dots\dots(4)$$

여기서 D.I는 문항변별도 지수, R_U, R_L는 각각 상부집단, 하부집단의 정답자 수, f는 각 집단의 학생 수이다.

식 (4)에 따르면 D.I=0.20이면 그 문항의 변별도가 있다고 해석할 수 있으며, 지수 0.20은 그 변량이 약 4%로서 통계적으로 4%수준에서 유의있는 정도이다⁴⁾.

성태제⁵⁾는 문항변별도 지수를 범위에 따라 표 2와 같이 표현하고 있다.

표 2. 문항변별도 지수의 범위별 표현

문항변별도 지수	문항변별도지수의 범위별 표현
0.40 이상	변별력이 높은 문항
0.30 ~ 0.39	변별력이 있는 문항
0.20 ~ 0.29	변별력이 낮은 문항
0.10 ~ 0.19	변별력이 매우 낮은 문항
0.10 이하	변별력이 없는 문항

2) 古典檢査理論의 問題點

고전검사이론은 피험자의 능력을 추정하는 절차가 비교적 용이하고, 문항에 대한 특성으로 난이도, 변별도 등을 추정하는 것이 쉽기 때문에 활용의 기회가 많다. 그러나 많은 문제점들을 내포하고 있다.

고전검사이론의 단점과 그 적용의 제한점을 정리하면 다음과 같다^{12),13),14),3)}

첫째, 난이도, 변별도 등의 문항 특성치와 피험자 능력 추정치는 피험자의 검사총점에 의한 분석이다.

둘째, 진점수 추정을 위한 반복 측정은 실제적으로 불가능하다.

셋째, 난이도, 변별도 등의 문항 특성치는 피험자 집단의 표본 특성에 따라 변한다. 즉 표본에 종속적이다.

넷째, 검사도구의 특성에 따라 능력 추정치가 변한다. 즉 피험자 능력이

검사에 종속된다.

다섯째, 피험자 능력 수준에 관계없이 모든 피험자의 검사에 대한 신뢰도가 같아진다, 즉 문항 특성 또는 피험자 능력 추정치의 오차변량이 모든 피험자에게 동일하다고 가정하고 이를 적용하는 것이 납득하기 어렵다.

여섯째, 극단적으로 우수한 능력 또는 낮은 능력의 피험자에게는 정확한 능력 추정치를 제공하지 못한다.

일곱째, 피험자가 검사 문항에 접했을 때 그 수행의 정도를 나타내는 어떤 정보도 사전에 제시해 주지 못한다.

여덟째, 고전검사이론의 기본 가정인 평행 검사의 제작이 불가능하다.

3. 問項反應理論

1) 問項反應理論의 意義와 長點

일반적으로 하나의 검사는 개개의 문항들로 구성되며, 검사점수는 문항점수의 합이 된다. 검사를 제작하는 경우 제작된 검사를 목표집단에 시행할 때, 그 검사의 통계적 특성치와 측정적 특성을 미리 예측할 수 있어야 한다. 사전에 유사한 피험자에게 검사를 시행하지 않고도, 검사문항에 대한 피험자의 반응을 확률적으로 예측할 수 있는 문항 모수와 피험자 능력 모수를 기술할 필요가 있다. 이미 시행한 검사문항에 대한 피험자의 반응으로부터 다음에 시행할 문항을 선택한다. 반복해서 문항을 선택할 때 가능한 한 피험자의 능력을 정확하게 추정할 수 있도록 문항을 선택하여야 한다. 피험자에게 시행할 일련의 문항으로부터 피험자의 능력을 근사적으로라도 추정할 수 있어야 한다. 또한 문제은행에 있는 검사문항들이 모든 능력수준을 측정하는데 얼마나 효과적인가를 알아야 한다¹⁾.

이러한 일은 고전검사이론으로는 불가능하다. 대부분의 검사는 피험자의 능력수준이나 기능을 추정하는 일이 주요한 과제가 되며 문항반응이론은 피험자의 능력수준과 반응의 관계를 수학적으로 진술하는 데서부터 출발한다.

문항반응이론이 수리적으로 복잡함에도 불구하고 교육, 심리, 측정분야에서 연구, 응용되고있는 이유는 다음과 같은 장점이 있기 때문이다⁸⁾.

첫째, 문항특성의 불변성이다. 즉 문항의 모수치를 추정하는데 피험자 집단과 독립적이다. 문항은 문항이 지니고 있는 고유한 특성이 있기 때문에 피험자집단의 특성에 의해 문항의 특성이 달리 추정되지 않는다는 점이다. 능력이 높은 집단에서 검사를 실시한 후 문항의 특성을 추정하는 능력이 낮은 집단에서 검사를 실시한 후 문항의 특성을 추정하든 문항특성이 같게 추정된다는 점이다.

둘째, 피험자 능력의 불변성이다. 피험자의 능력을 추정하는데 검사의 문항과 독립적이다. 피험자는 개인이 지니고 있는 고유한 능력 수준이 있기 때문에 다른 검사도구를 택하였다 하여도 피험자의 능력이 달리 추정되지 않는다는 점이다. 문항반응이론은 피험자가 쉬운 검사도구를 택하든 어려운 검사도구를 택하든 그 피험자의 능력이 같게 추정되는 장점을 가지고 있다.

셋째, 학교현장과 관련해서 문항반응이론에 의한 문항의 정확한 평가와 관리가 가능해져, 문제은행의 제작이 용이하다.

2) 潛在的 特性 母型과 問項反應理論의 假定

임의의 검사에서 피험자의 검사 점수에 영향을 미칠 수 있지만 관찰할 수 없는 특성이 있다고 가정하고, 이를 잠재적 특성이라 하여 피험자의 검사 점수로부터 잠재적 특성을 추정하는 절차와 관련된 일련의 이론을 잠재적 특성 모형(latent trait model) 또는 문항반응이론이라고 한다.

잠재적 특성 모형과 문항반응이론은 따로 발전되어 왔으나, 최근에 와서는 잠재적 특성 모형을 일반적으로 문항반응이론이라 하며¹⁵⁾¹⁶⁾, 잠재적 공간의 차원성, 국소독립성, 문항특성곡선이라는 세 가지 기본 개념을 도입한다.

문용린¹⁷⁾은 문항반응이론을 검사문항의 분석에 올바르게 적용하기 위해서는 대체로 세 가지 측면의 전제가 만족되어야 함을 강조하고 있다.

첫째, 검사 문항의 수와 사례 수에 대한 전제이다. 문항특성곡선의 여러

모수의 측정오차에 영향을 주는 요인으로 검사문항의 수와 표집 사례수가 가장 큰 영향을 미치는 것으로 밝혀지고 있다. Lord의 연구¹⁸⁾에 의하면 문항의 특성으로 문항의 난이도, 변별도 및 추측도를 고려하는 3-모수 로지스틱 모형을 적용할 경우 검사 문항 수는 최소한 50, 사례 수는 1,000명 이상은 되어야 한다. 1 혹은 2-모수 로지스틱 모형인 경우에는 그 보다 더 적은 문항 수와 사례 수로도 가능하다.

둘째, 모든 검사문항들이 단일 잠재 특성을 재고 있다는 전제의 만족이다. 만약 한 검사내의 문항들이 두 개 이상의 잠재 특성을 재고 있다고 하면 문항반응이론 기법의 적용은 원칙상 불가능하다. 다만 Drasgow 와 Parsons¹⁹⁾에 의하면 문항반응이론의 기법은 단일 차원성의 가정에 극도로 예민한 것은 아니라고 한다. 그러나 단일 차원성이 크게 의심스러울 때는 이에 대한 검증은 꼭 필요하다.

셋째, 검사의 각 문항들은 국소적으로 독립적이어야 한다는 전제의 만족이다. 즉 한 문항에 정답할 확률이 다른 문항에 정답할 확률에 영향을 미치지 않아야 한다. 이 상호 독립성을 검증하는 방법으로 흔히 단일차원성의 가정이 만족되면 대체로 이 조건도 만족되는 것으로 간주하고 한다.

(1) 潛在的 空間의 次元性

잠재적 공간의 차원성이란 피험자의 검사 점수에 영향을 미치는 잠재적 특성의 수를 말한다. 잠재적 공간은 잠재적 특성에 대응하는 일련의 직교 차원으로 정의되며 잠재적 공간에서 피험자의 위치는 잠재적 특성 점수로 결정된다.

잠재적 특성 모형에서 잠재적 공간은 단일차원이라고 가정하는데 검사는 단일 능력이나 특성을 측정하는 동질적인 문항으로 구성된다는 것이다. 즉, 한 개 이상의 잠재적 특성을 측정하는 검사문항을 고려할 수 있다.

어떤 한 검사를 문항반응이론에 의해 분석하고자 한다면 잠재적 공간의 단일차원성 가정의 만족여부를 확인하는 과정이 필요하다. 단일차원성의 적

합도 검증은 흔히 요인분석 방법의 사용을 권하는데 대부분의 검사들이 한 개의 요인만을 얻는 경우는 없다. 따라서 지배적인 요인과 그 하위 요인들간의 총분산의 비율에 의해서 단일차원성의 지표로 삼자는 방법과, 1차 및 2차 요인까지를 포함시키거나, 하나의 지배적 요인과 작은 여러 개의 요인으로 구성된 검사를 통해서, 또한 다차원적 검사자료를 사용해 검증하는 방법이 제안되고 있다.

검사 문항의 요인분석 결과, 두 개 이상의 요인으로 분석된다면, 동일한 요인으로 묶을 수 있는 독립적인 문항에 대하여 잠재적 특성 모형을 적용할 수 있다¹⁹⁾.

안창규와 신석기²⁰⁾는 “Rasch 모형에 의한 MMPI(Minnesota Multiphasic Personality Inventory) 우울증 척도의 문항 양호도 분석” 연구에서 문항반응이론의 자료 적합도 검증 방법으로 주요인분석방법을 사용하여 각 요인이 가지는 전체 설명력의 비율인 고유치가 1.628 이상인 8개 요인에 대한 공통변량의 값으로서 검증하여, 우울증 척도의 60문항은 대체로 일차원성의 가정과 국소독립성의 가정을 만족한다고 하였다. 이순화²¹⁾는 “문항반응이론을 적용한 일반지능검사의 문항분석”의 연구에서 요인분석을 통해 언어추리 검사와 도형추리 검사의 문항이 공통변량 0.5이상인 문항이 각각 2개, 3개 등이며, 검사점수의 분산 중에서 제 1요인에 의해 설명될 수 있는 분산의 비율이 14.6%, 19.3%로 20%에 미달되지만 문항반응이론을 실행하기 위한 자료로 적합하다고 판단하고 있다. 정재임²²⁾은 “교사 출제 영어시험 문제 분석을 위한 문항반응이론의 적용 가능성” 연구에서 각각 20, 21 및 26문항으로 구성된 영어시험문제가 요인분석결과 제 1 인자에 의해 설명할 수 있는 변량이 17.3, 17.2 및 17.2 %로 단일차원성의 가정을 어느 정도 만족한다고 하고 있다.

(2) 局所 獨立性 假定

국소 독립성 가정은 강한 형태와 약한 형태로 분류할 수 있다¹⁾.

강한 형태는, 피험자의 검사문항 점수는 통계적으로 독립이라는 가정이다. 이 가정은 검사문항에 대한 피험자의 점수는 문항에 대한 피험자의 점수에 영향을 미치지 않음을 의미한다.

문항에 정답을 할 확률은 세 개의 문항 모수와 피험자 능력 모수에 좌우되며, 다른 어떠한 것에도 좌우되지 않는다. 만약, 그 모형이 참이라고 하면 어느 피험자가 특정 문항에 정답을 할 확률을 결정하는데 필요한 것은 그 피험자의 능력 모수 뿐이다. 피험자의 능력을 안다고 하면, 다른 문항에 정답을 했는지 오답을 했는지를 아는 것이 그 피험자가 특정 문항에 정답을 할 확률을 결정하는데 도움이 되지 않는다.

문항 i 의 점수를 $u_i=0$ 또는 1이라고 하면 국소 독립성의 가정은 다음 식 (5)로 나타낼 수 있다.

$$P(u_i=1 | \theta) = P(u_i=1 | \theta, u_j, u_k, \dots) \quad (i \neq j, k, \dots) \quad \dots\dots\dots (5)$$

국소 독립성을 수학적으로 달리 진술하면 모든 문항에 정답을 할 확률은 각 문항에 정답을 할 확률의 곱과 같다고 표현할 수 있다. 3개의 문항 i, j, k 를 예로 들면,

$$P(u_i=1, u_j=1, u_k=1 | \theta) = P(u_i=1 | \theta)P(u_j=1 | \theta)P(u_k=1 | \theta) \dots (6)$$

약한 형태의 가정으로 국소 독립성은 θ 가 고정되어 있을 때 두 문항의 상관성이 없어야 만족된다. 동일한 능력을 가진 피험자에 대하여 두 문항의 점수간에 상관성이 없다고 가정하면 만족된다. 그러나 상이한 집단 내에서 문항이 상관성이 없어야 한다는 것은 아니다.

국소 독립성은 단일차원성으로부터 당연히 귀결되는 가정이며, 추가되는 가정은 아니다. 단일차원성과 동일한 형태의 가정이므로 요인분석으로 그 적합성을 검증할 수 있다²³⁾.

(3) 問項特性曲線

문항반응이론은 문항 하나 하나에 근거하여 분석하는 이론이다. 그러므로 문항반응이론에서 가장 중요한 개념의 하나가 문항특성곡선이다.

임의의 한 검사 문항에서 전체 피험자의 검사점수와 문항에 정답한 비율과의 관계를 그림 1로 나타내며 이를 문항-관찰점수 회귀곡선(item - observed score regression curve)¹⁾이라 하는데, 이 곡선에서 측정하려는 잠재적 특성치는 알 수 없기 때문에 관찰한 검사 총점을 특성치의 추정치로 사용하여 문항특성곡선을 나타낸다.

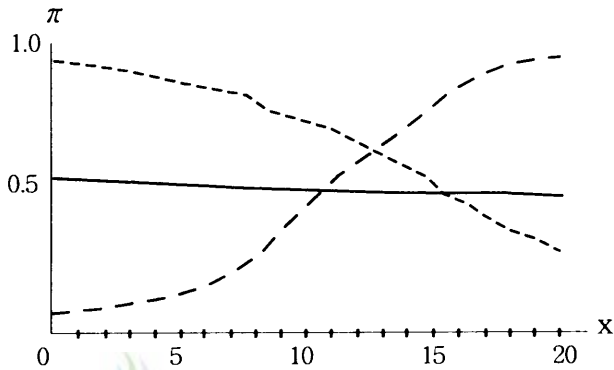


그림 1. 문항-관찰점수 회귀 곡선

문항특성곡선은 그림 2와 같이 검사가 측정하려는 잠재적 능력 특성치 θ 를 검사 총점 x 로, 능력 특성치 θ 를 소지한 피험자가 문항에 정답을 할 확률 $P(\theta)$ 를 문항에 정답을 한 피험자의 비율 π 로 나타냈다. 문항특성곡선은 능력 특성치 θ 와 문항 모수로 규정되는 수학적 함수를 이론적으로 가정한다.

문항에 정답을 확률과 검사가 측정하려는 특성 사이의 관계를 나타내는 수학적 함수로, S자 형태를 지니며, 가로축은 그림 1의 검사총점 대신에 능력 θ 로, 세로축은 정답을 한 피험자의 비율을 문항에 정답할 확률 $P(\theta)$ 로 변환하여 나타낸다.

인간의 능력은 거의 없는 상태에서 무한히 높은 능력수준의 범위를 지나나, 이를 표준화하여 평균이 0이고 표준편차가 1인 점수로 변환하였다. 세로축은 능력수준에 따라 문항의 답을 맞출 확률을 나타내고 범위는 0에서 1이다²⁴⁾.

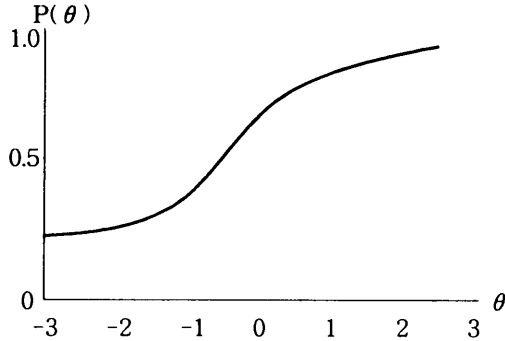


그림 2. 문항특성곡선

문항반응이론은 다양한 문항 반응을 나타내는 문항곡선을 다룰 수 있으나, 특성치 θ 가 증가함에 따라 그 문항에 정답을 할 확률 $P(\theta)$ 이 증가하는 특성 곡선만을 다룬다⁹⁾.

3) 問項反應模型

검사를 구성하는 문항에 대한 피험자의 응답 결과를 0과 1로만 채점하는 하는 문항반응모형을 이분 반응모형²⁴⁾이라 하고, 일반적으로 인지적 능력을 측정하는데 사용되며, 수학적 함수의 형태로 나타낸 정규오자이브 모형과 로지스틱 모형이 있다.

정규오자이브 모형은 정규오자이브 함수에 의하여 문항의 답을 맞출 확률이 계산되며, 문항난이도만 고려한 문항반응모형은 1-모수 정규오자이브 모형, 문항난이도와 문항변별도를 고려하면 2-모수, 문항추측도까지 고려한 문항반응모형은 3-모수 정규오자이브 모형이라 하며, 식 (7), (8) 및 (9)로 나타낼 수 있다.

$$P(\theta) = \int_{-\infty}^{(\theta-b)} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz \quad \dots\dots\dots (7)$$

$$P(\theta) = \int_{-\infty}^{a(\theta-b)} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz \quad \dots\dots\dots (8)$$

$$P(\theta) = c + (1-c) \int_{-\infty}^{a(\theta-b)} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz \quad \dots\dots\dots (9)$$

$P(\theta)$ 는 문항의 답을 맞출 확률, a^* 는 문항변별도, b 는 문항난이도, c 는 문항추측도, θ 는 피험자 능력, 그리고 z 는 $(\theta - \mu)/\sigma$ 로 피험자 능력의 표준화 점수이다. 여기서 μ 는 피험자 능력의 평균치이다.

정규오자이브 함수에 의해 문항의 답을 맞출 확률을 계산할 때, 적분의 어려움이 있다. 이를 해결하기 위하여 로지스틱 함수를 사용한다.

Harley⁸⁾는 정규 오자이브 모형과 로지스틱 모형에 의한 $P(\theta)$ 값의 차이를 줄이기 위하여 로지스틱 함수에 의하여 문항의 답을 맞출 확률을 계산할 때, 두 모형에 의해 추정된 문항변별도 사이에 $a=1.702a^*$ 라는 관계를 발견하였다. a^* 는 정규오자이브 모형에 의한 문항변별도, a 는 로지스틱 모형에 의한 문항변별도이다. 따라서 로지스틱 모형에 대한 문항반응모형 식은 문항변별도에 상수 1.7을 곱하여 나타낸다. 로지스틱 모형에 대한 문항반응모형의 식은 다음 식 (10), (11) 및 (12)와 같이 나타낼 수 있다.

$$P(\theta) = c + (1-c) \frac{1}{1 + e^{-a(\theta-b)}} \dots\dots\dots (10)$$

$$P(\theta) = \frac{1}{1 + e^{-a(\theta-b)}} \dots\dots\dots (11)$$

$$P(\theta) = \frac{1}{1 + e^{-(\theta-b)}} \dots\dots\dots (12)$$

(1) 3-모수 로지스틱 모형

능력 특성치 θ 인 피험자가 한 문항에 정답을 할 확률을 $p(\theta)$ 라고 하면 3-모수 로지스틱 모형의 함수는 식 (10)으로 나타내어진다. 여기에서 a, b, c 는 문항의 특성을 나타내는 문항 모수이다.

그림 3에서 c 는 문항추측도 모수(guessing parameter)이다. 모수 c 는 능력특성치가 완전히 낮은 피험자 즉 $\theta = -\infty$ 인 피험자가 문항에 정답을 할 확률이다. 추측에 의해서 정답을 할 수 없는 문항이라고 하면 $c=0$ 이다¹⁾.

그림 3에서 b 는 문항난이도 모수를 나타낸다. 능력 특성치와 동일한 척도 상에서 $-\infty$ 와 $+\infty$ 사이의 값으로 정의되어지나, 실제로는 특성치의 분포가 -3부터 +3까지의 범위라고 가정하면 b 는 약 -2부터 +2까지의 범위를 갖는

다. $b = -2$ 라고 하면 대단히 쉬운 문항이고, $b = +2$ 라고 하면 대단히 어려운 문항이다. 로지스틱 곡선은 $\theta = b$ 에서 변곡점을 갖는다. $c = 0$ 라고 하면 b 는 $P(\theta) = 0.5$ 에 대응하는 θ 값과 같다. $c \neq 0$ 라고 하면 b 는 $P(\theta) = (1+c)/2$ 에 대응하는 능력수준 θ 위의 점을 말한다. 이 확률은 c 의 값과 1.0 사이의 중간 점을 말한다. 여기서 알 수 있는 것은 문항의 답을 맞힐 확률이 가장 낮은 경우의 값으로 추측도 모수 c 가 정의되어진다. 따라서 난이도 모수는 문항의 답을 맞힐 확률이 c 와 1.0 사이의 중간 능력 척도상의 점으로 정의할 수 있다.

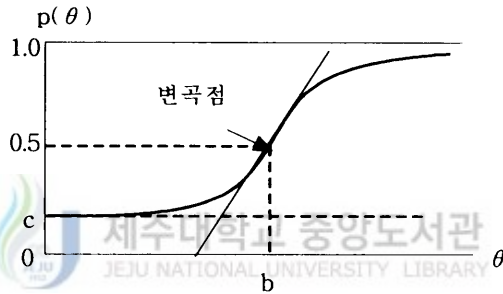


그림 3. 문항 모수

그림 3에서 a 는 특성 수준에 따라 문항 반응이 변하는 정도를 나타내는 문항의 변별도 모수이다. a 는 변곡점($\theta = b$)에서 곡선의 기울기에 비례하고 능력 특성이 정규적으로 분포되어 있다면 a 는 식 (13)으로 추정할 수 있다.

$$a = \frac{\rho_{jx}}{\sqrt{1 - \rho_{jx}^2}} \dots\dots\dots (13)$$

이 식에서 ρ_{jx} 는 문항 점수와 검사 총점간의 양분점 상관계수이다.

(2) 2-모수 로지스틱 모형

추측 요인이 작용할 수 없는 문항($c=0$)이라면, 3-모수 문항 반응 함수는 2-모수 모형이 된다. 식 (11)은 두 개의 모수 a 와 b 만을 갖는 함수이다.

이 식에서 a 와 b 는 3-모수 모형에서와 같이 문항변별도와 난이도이다.

(3) 1-모수 로지스틱 모형(Rasch 母型)

확률이론의 관점에서 문항난이도만을 고려하여 전개한 1-모수 모형은 2-모수 모형의 가정에, 문항 변별력 a 는 모든 문항에 대해 $a = 1.0$ 의 값으로 고정되어지며, 단지 난이도 모수만 다른 값을 가지게 되는 모형으로 1960년 Georg Rasch에 의해 개발된 모형이며 Rasch 모형이라고도 한다⁸⁾.

식 (10)에서 b 는 문항의 특성을 나타내는 문항 난이도 모수이다.

1-모수모형(Rasch 모형)은 2 또는 3-모수 모형보다 추정하여야 할 모수가 적고 모수를 추정하는 문제가 용이하다는 이점을 갖고 있다. 반면에, 추가적인 가정을 하기 때문에 Rasch 모형이 자료에 적합하기는 더 어렵다. 따라서 다른 모형과 마찬가지로 일련의 검사 문항에 대하여 모형의 적합도 검정을 하여야 한다.

4) 問項反應理論에 依한 問項의 母數

(1) 問項辨別度

문항변별도는 문항이 문항의 위치지수 아래의 능력을 가진 피험자들을 얼마나 잘 변별할 수 있는가를 서술한다. 이 속성은 본질적으로 문항특성곡선의 기울기를 말하며, 그 기울기가 클수록 문항은 피험자들을 더욱 잘 변별할 수 있다.

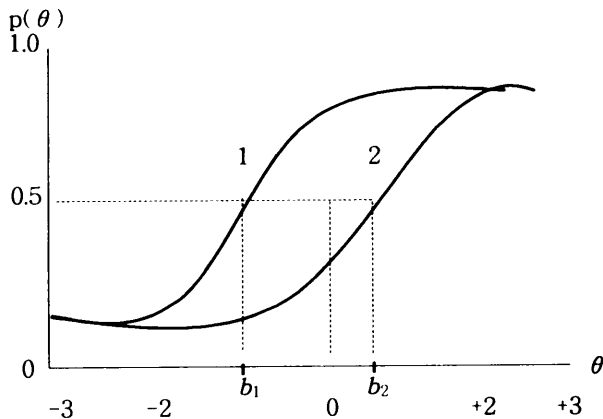


그림 4. 문항변별도

문항변별도는 문항특성곡선 상에서 문항에 정답 할 확률 0.5에 대응하는 능력척도상의 점이다. 위의 그림 4의 1과 2의 경우 변별도는 각각 b_1 과 b_2 로 곡선 2에 해당하는 문항이 변별력이 높다.

Baker⁸⁾는 로지스틱 모형 하에서 문항변별도를 설명하기 위하여 표 3과 같이 문항변별도 모수의 범위에 대한 표현을 구체화 하였다.

표 3. 문항변별도 모수에 대한 범위별 표현

문항변별도 모수	변별도 모수에 대한 범위별 표현
0.00	없다
0.01 - 0.34	매우 낮다
0.35 - 0.64	낮다
0.65 - 1.34	적절하다
1.35 - 1.69	높다
1.70 이상	매우 높다
+ ∞	완벽하다

이러한 기준은 문항특성곡선을 위한 로지스틱 모형 하에서 변별도 모수치를 해석할 때 적용된다.

모든 모형에서 음의 변별도를 가진 문항특성곡선은 문항난이도가 같고 양의 변별도를 가진 문항특성곡선의 형태와 문항난이도를 나타내는 점을 중심으로 좌우대칭 형태를 나타낸다. 문항변별도가 음수 값을 지닌다면 능력수준이 낮은 피험자가 능력수준이 높은 피험자에 비해서 정답을 맞힐 확률이 높은 문항이라는 의미로 이런 문항은 있어서는 안될 문항이며, 문항변별도가 $b=0$ 이란 의미는 피험자의 능력차이에도 불구하고 피험자들이 그 문항의 답을 맞힐 확률이 같음을 의미하므로 이런 문항도 좋지 못한 문항이다²⁵.

(2) 問項難易度

문항난이도 모수치는 그 문항의 능력척도상의 어디에서 기능하는지를 나타내 주는 것이다. 문항특성곡선의 기울기는 문항난이도에 일치하는 능력수준에서 가장 가파르다⁸⁾.

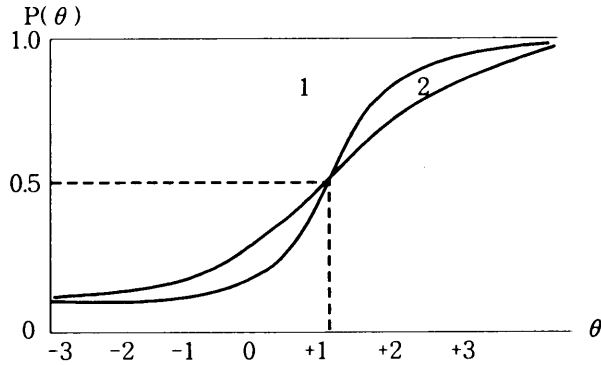


그림 5. 문항난이도

문항의 난이도는 문항특성곡선상의 문항에 정답할 확률 0.5에 대응하는 점에서의 곡선의 기울기로 정의한다. 그림 5의 1과 2의 경우 2문항의 경우가 1의 경우보다 문항이 난이도가 어렵다는 것을 말해준다.

문항반응이론에서 문항난이도는 1-모수, 2-모수 모형에 의하면 문항의 답을 맞힐 확률 0.5에 대응하는 능력척도상의 점이고, 3-모수 모형에 의하면 $(1+c)/2$ 에 대응하는 능력척도상의 점이다. 문항의 난이도 모수 a 는 Z점수 척도와 같으므로 이론적으로 $-\infty$ 에서 $+\infty$ 까지의 값을 가질 수 있으나, 실제로는 -3.0에서 +3.0까지의 값을 갖는다.

문항난이도 모수에 대하여 범위별로 다음의 표 4와 같이 표현할 수 있다.

표 4. 문항난이도 모수에 대한 범위별 표현

문항난이도 모수	난이도 모수에 대한 범위별 표현
-2.0 이하	매우 쉽다
-2.0 - -0.5	쉽다
-0.5 - +0.5	중간이다
+0.5 - +2.0	어렵다
+2.0 이상	매우 어렵다

표 4의 표현은 절대적이 아니며, 어느 정도 비전문가의 문항 평가에 적절한 지침이 될 수 있다. 그러나 문항변별도와 같이 문항난이도 모수치에 의하여 문항난이도를 언어적으로 표현하는데는 몇 가지 문제가 있다. “쉽다”와

“어렵다”는 용어는 어떤 견해에 의존하는 상대적인 용어이다.

문항난이도는 검사의 목적과 검사가 의도하는 집단의 예견되는 능력분산에 의해 선택되어야 한다. 규준지향검사의 경우 문항난이도가 모든 능력범위를 통하여 균일하게 퍼져 있어야 하고, 돌출부형 검사나 목표지향검사인 경우에는 특정한 위치에서 기능할 수 있는 난이도를 가진 문항을 선정하는 것이 바람직하다⁸⁾.

(3) 問項推測度

문항이 선택형 문항일 경우 정답을 모르는 피험자는 추측을 하게 된다. 이론적으로 능력이 전혀 없는 피험자가 문항의 답을 맞힐 확률이 $P(\theta)=0$ 이 되나, 실제적으로 우연에 의해 문항의 답을 맞힐 수 있다. 능력이 전혀 없는, 즉 $-\infty$ 의 능력수준을 갖는 피험자가 문항의 답을 맞힐 확률을 문항추측도 c 로 표기하며, 추측 모수(pseudoguessing parameter)라고도 부른다²⁶⁾²⁷⁾.

<그림 6.>의 c 는 문항추측도로 능력척도 상 $-\infty$ 에 대응하는 문항에 정답할 확률이다. Lord(1968)는 이를 문항특성곡선의 최저한계(lower bound)라 명명하였으나, 최근에는 문항추측도라 부른다.

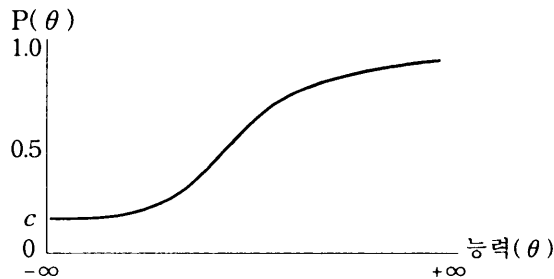


그림 6. 문항추측도

고전검사이론에서는 총 피험자중 문항의 답을 알지 못하고 우연에 의해 문항의 답을 맞힌 피험자수의 비율로 계산되나 문항반응이론에서는 문항특성곡선의 최저한계에 의해 계산한다. 문항추측도 값이 높을수록 그 문항은 좋지 않는 문항이라 할 수 있다. 이는 능력이 전혀 없는 학생도 문항을 맞힐

수 있는 확률이 높다는 것으로서, 이는 문항 안에 정답을 암시하는 요소들이 많기 때문이라 할 수 있다. 또한 문항의 추측도가 c 인 문항에서 문항을 맞힐 수 있는 확률의 범위는 c 에서 1.0이므로, 문항추측도가 높을수록 문항변별력은 떨어진다고 할 수 있다⁸⁾.

문항추측도에 의한 문항의 평가 시 범위별로 표현하면 표 5와 같다⁸⁾.

표 5. 문항추측도 모수의 범위별 표현

문항추측도(5지 선다형인 경우) 모수	문항추측도 모수의 범위별 표현
$c < 2.0$	양호한 문항
$2.0 \leq c < 3.0$	추측도가 높은 문항
$3.0 \leq c$	검사에서 수정하거나 삭제할 문항

이 것은 5-지 선다형 문항인 경우이며, 답지 수에 따라 그 기준이 달라질 수 있다.

5) 문항반응이론의 문항 모수와 능력 추정

문항반응이론에 의한 모형에서의 능력 및 문항모수치의 추정방법은 대체로 log 방법, 최대 우도(maximum likelihood: ML)방법 및 Bayesian 방법 등으로 나뉘어지는데, 여기에서는 주로 후자의 두 방법을 기술하고자 한다¹²⁾.

최대 우도 방법은 문항과 원점수의 행렬표에서, 반응 패턴인 관찰된 자료는 가장 가능한 경우라는 가정 하에서 관찰된 표집 자료를 얻을 우도(likelihood)가 최대가 되도록 모수치를 추정하는 방법이다. 이러한 최대우도 방법은 모수치에 대한 우도를 조건화하지 않는 무조건적 최대우도방법(Unconditional Maximum Likelihood: UML)과 이 방법의 추정의 오차를 개선하기 위해서 제안된 조건적 최대우도방법(Conditional Maximum Likelihood: CML)이 나왔는데 이는 Rasch 모형에서 원 점수는 잠재적 특성 추정의 충분 통계치라는 사실에 근거하여 파악되지 못한 능력 모수치 대신에 원 점수를 사용함으로써 자료의 우도를 조건화해서 문항 모수치를 피험자 표집에 독립적으로 추정할 수 있는 방법이다. 이러한 CML 방법에서의

독립성을 갖는 방식에 비하여 주변 최대 우도 방법(marginal maximum likelihood :MML)은 표집에서 능력 수준은 관련 전집에서 무선 표집 되었다는 가정 위에서 같은 점수를 가지는 모든 반응 벡터는 개인 모수치에 대하여 비율적 우도를 가진다는 사실에 근거하여 표집에 독립하는 문항 모수치를 구하고는 것이다.

Rasch 모형의 Bayesian 추정 방법은 만일 θ_j 의 분포가 무선적으로 또 독립적으로 분포되어 있다는 사전정보나 신념을 가지고 있다면 그 사전분포에 문항 반응 우도(item response likelihood)를 결합한 사후의 분포는 사전 분포와 같아진다는 사실에 입각하여 문항 모수치를 구한다.

Thissen²⁸⁾은 주변 최대 우도 방법이 CML과 동일한 결과를 얻으면서도 편리한 방법이라고 하며 CML은 1-모수 모형에만 적용되지만 MML은 어떤 모형에도 적용성 내지 적합도 검정을 할 수 있는 장점이 있다고 했다.

(1) 최대우도추정



정확한 추정을 위해서는 표본의 크기가 크고 문항이 많은 검사가 좋다. u_j 를 문항 j 의 점수라고 하면 주어진 θ (특성치)에 대한 j 문항 반응의 조건 분포는 다음 식 (14)와 같다.

$$L(u_j | \theta) = P_j^{u_j} Q_j^{1-u_j} \dots\dots\dots (14)$$

잠재적 모형의 국소 독립성이라는 가정에 따라 n 개의 문항 반응의 조건 분포는 각 문항에 대한 식 (14) 분포의 곱과 같다.

$$L(u_1, u_2, \dots, u_n | \theta) = \prod_{j=1}^n P_j^{u_j} Q_j^{1-u_j} \dots\dots\dots (15)$$

N 명의 피험자(특성치가 $\theta_1, \theta_2, \theta_3, \dots, \theta_N$ 인 피험자)가 n 개의 문항에 대한 반응은 다음 행렬 식 (16)으로 나타낼 수 있다.

$$U = \{u_{ij}\}, \begin{pmatrix} j=1, 2, \dots, n \\ i=1, 2, \dots, N \end{pmatrix} \dots\dots\dots (16)$$

N 명의 피험자의 반응은 독립이라고 가정할 수 있으므로, N 명의 피험자에 대한 n 개의 문항반응의 조건분포는 N 명의 피험자에 대한 각 분포 식 (15)의

곱과 같다.

$$L(\mathbf{u} | \boldsymbol{\theta}; \mathbf{a}, \mathbf{b}, \mathbf{c}) = \prod_{i=1}^N \prod_{j=1}^n P_{ji}^{u_{ji}} Q_{ji}^{1-u_{ji}} \dots \dots \dots (17)$$

식 (17)에서 $P_{ji} \equiv P_j(\theta_i)$ 는 i 번째 피험자(특성치 θ_i)가 j 문항에 정답을 할 확률이고 $Q_{ji} \equiv 1 - P_{ji}$ 이다.

$$\begin{aligned} \boldsymbol{\theta} &= \{\theta_1, \theta_2, \dots, \theta_i, \dots, \theta_N\}' \\ \mathbf{a} &= \{a_1, a_2, \dots, a_j, \dots, a_n\}' \\ \mathbf{b} &= \{b_1, b_2, \dots, b_j, \dots, b_n\}' \\ \mathbf{c} &= \{c_1, c_2, \dots, c_j, \dots, c_n\}' \end{aligned} \dots \dots \dots (18)$$

최대우도 추정치(maximum likelihood estimates) θ_i, a_j, b_j, c_j 는 식 (17) 이 최대값을 갖도록 취하는 모수치이다¹⁾. 식 (17)의 양변에 자연대수를 취하면,

$$\ln L = \sum_{i=1}^N \sum_{j=1}^n [u_{ji} \ln P_{ji} + (1 - u_{ji}) \ln Q_{ji}] \dots \dots \dots (19)$$

우도방정식을 구하기 위하여 식 (19)를 θ_i, a_j, b_j, c_j 에 대하여 미분한 식을 0으로 놓는다.

$$\frac{\partial \ln L}{\partial \theta_i} = \sum_{j=1}^n \frac{u_{ji} - P_{ji}}{P_{ji} Q_{ji}} \frac{\partial P_{ji}}{\partial \theta_i} = 0, (i=1, 2, \dots, N) \dots \dots \dots (20)$$

$$\left. \begin{aligned} \frac{\partial \ln L}{\partial a_j} &= \sum_{i=1}^N \frac{u_{ji} - P_{ji}}{P_{ji} Q_{ji}} \frac{\partial P_{ji}}{\partial a_j} = 0, (j=1, 2, \dots, n) \\ \frac{\partial \ln L}{\partial b_j} &= \sum_{i=1}^N \frac{u_{ji} - P_{ji}}{P_{ji} Q_{ji}} \frac{\partial P_{ji}}{\partial b_j} = 0, (j=1, 2, \dots, n) \\ \frac{\partial \ln L}{\partial c_j} &= \sum_{i=1}^N \frac{u_{ji} - P_{ji}}{P_{ji} Q_{ji}} \frac{\partial P_{ji}}{\partial c_j} = 0, (j=1, 2, \dots, n) \end{aligned} \right\} \dots \dots \dots (21)$$

3-모수 모형에 대하여 나타내면 다음의 식 (22)와 같다.

$$\left. \begin{aligned} \frac{\partial P_{ji}}{\partial \theta_i} &= \frac{1.7 a_j Q_{ji} (P_{ji} - c_j)}{1 - c_j} \\ \frac{\partial P_{ji}}{\partial a_j} &= \frac{1.7 (\theta_i - b_j) Q_{ji} (P_{ji} - c_j)}{1 - c_j} \\ \frac{\partial P_{ji}}{\partial b_j} &= \frac{-1.7 a_j Q_{ji} (P_{ji} - c_j)}{1 - c_j} \\ \frac{\partial P_{ji}}{\partial c_j} &= \frac{Q_{ji}}{1 - c_j} \end{aligned} \right\} \dots \dots \dots (22)$$

문항 모수를 알고 있다고 하면 식 (22)의 N 개의 방정식으로부터 피험자 i 의 특성의 추정치 θ_i 를 구할 수 있다. 추정치는 다른 θ 에 대해 독립이다. 또한 피험자 능력모수(θ)를 알고 있다고 하면 식 (22)로부터 문항 j 의 모수를 구할 수 있다. 문항 j 의 모수는 다른 문항의 모수와 독립이다.

(2) 二分 反應 模型의 問項 母數值 推定

검사에서 문항 모수의 실제 값을 모르기 때문에 문항반응이론에서 검사를 분석할 때 중요한 과제 중의 하나는 문항 모수치를 추정하는 일이다. 문항 모수 추정치는 문항들의 기술적인 속성들에 관한 정보를 제공한다⁸⁾.

문항 모수를 추정하기 위하여 피험자의 능력수준을 알고 있다고 가정한다. 검사에 응답한 피험자들을 능력 수준에 따라 J 개의 피험자 집단으로 나눌 때, j 집단의 피험자 n_j 명 중에서 r_j 명이 문항의 답을 맞았다면 능력 수준 θ_j 에서 관찰된 정답비율은 $P(\theta_j) = r_j/n_j$ 이다. 관찰된 정답비율에 가장 알맞은 문항특성곡선을 찾기 위하여 가장 보편적으로 사용되는 방법은 최대 우도 추정법(maximum likelihood estimation)이다. 각 능력집단에서 n_j 명중 r_j 명이 문항의 답을 맞힐 확률은 식 (23)으로 나타내며, 우도함수(likelihood function)라고 한다.

$$\text{Prob}(r_1 r_2 \cdots r_J) = \prod_{j=1}^J \frac{n_j!}{(n_j - r_j)! r_j!} [P(\theta_j)]^{r_j} [1 - P(\theta_j)]^{n_j - r_j} \cdots \quad (23)$$

관찰자료를 가장 잘 설명하는 문항특성곡선을 이루는 문항 모수 추정치는 식 (23)식의 값을 최대로 하는 값을 선택한다.

최대 우도 추정법에 의하여 문항 모수치를 추정하는 절차는 다음과 같다. 문항 모수에 초기 값들을 부여한 후 각 능력 수준별로 $P(\theta_j)$ 를 계산하여 관찰된 $P(\theta_j)$ 와 이론적인 $P(\theta_j)$ 의 불일치 정도가 최소화 될 때까지 문항 모수 추정을 반복한다. 다시 말하여 문항 모수 추정오차가 무시하여도 좋을 만큼 작아질 때까지 문항 모수 추정치를 반복 교정하여 문항 모수의 추정치를 얻는다.

(3) 피험자 능력의 數와 推定

검사에 대한 피험자의 응답을 이분적으로 채점하여 각 문항에 0점 혹은 1점을 부여한 리스트 $U=(u_1, u_2, \dots, u_n)$ 를 문항반응벡터(item response vector)라고 한다. 피험자가 각 문항에서 u_1, u_2, \dots, u_n 점을 받을 확률은 식 (24)로 주어진다.

$$\text{Prob}(u_1, u_2, \dots, u_n) = \prod_{i=1}^n P_i^{u_i} Q_i^{1-u_i} \dots\dots\dots (24)$$

피험자의 능력을 추정하는 기본원리는 검사의 모든 문항을 통하여 실제 반응과 문항의 답을 맞힐 이론적 확률과의 차를 최소화하기 위해 우도함수를 최대로 하는 능력 모수 추정치를 찾아야 한다.

최대 우도 추정법에 의하여 피험자의 능력 모수를 추정하는 절차는 다음과 같다. 피험자의 능력 모수의 초기 값과 알고 있는 문항 모수치를 가지고 피험자가 각 문항의 답을 맞힐 확률을 계산한다. 그 다음 피험자가 실제로 응답한 문항반응벡터와 이론적인 문항의 답을 맞힐 확률과의 일치성을 증가시키기 위하여 피험자 능력 모수의 추정치를 교정한다. 실제 문항 반응 벡터와 이론적인 확률과의 차이가 무시하여도 좋을 만큼 매우 작아질 때까지 이러한 과정을 반복하여 피험자 능력 모수의 추정치를 얻는다. 2-모수 로지스틱 이분반응모형에 있어서 피험자 능력 모수를 추정하는 공식은 식 (25)와 같다.

$$\theta_{s+1} = \theta_s + \frac{\sum_{i=1}^n a_i (u_i - P_i)}{\sum_{i=1}^n a_i^2 P_i Q_i} \dots\dots\dots (25)$$

θ_s 는 s 번째 반복추정에서의 능력 모수 추정치, u 는 문항에 대한 피험자의 응답(문항점수), P 는 능력수준이 θ_s 인 피험자가 문항의 답을 맞힐 이론적 확률이다.

피험자의 능력을 추정하는 식 (25)를 설명하면 우변에 있는 θ_s 는 능력을

추정하기 위하여 부여하는 초기 값이고, +기호 오른쪽에 있는 부분은 $\Delta\theta_s$ 로 표기할 수 있으며 능력추정의 교정 값이다. 즉 좌변의 θ_{s+1} 는 s 번째 반복추정에서의 능력모수 추정치 θ_s 에 $\Delta\theta_s$ 를 더한 값으로써 다음단계의 능력모수 추정치이다. $\Delta\theta_s$ 부분을 보면 능력모수 추정을 위한 반복추정의 본질을 알 수 있다. $u_i - P_i$ 는 피험자의 실제 반응과 능력수준 θ_s 를 가지고 있는 피험자가 문항의 답을 맞힐 이론적 확률과의 차이이다. 만약 피험자의 능력모수를 제대로 측정하였다면 u_i 와 P_i 의 차이가 작고 나아가 검사 전체에서 이 차이의 합은 매우 작아진다. 실제 문항 반응과 문항의 답을 맞힐 확률과의 차가 최소화되면 $\Delta\theta_s$ 값은 매우 작아 0에 가까워지므로 능력추정을 위한 반복 교정은 끝나고 피험자의 능력모수 추정치는 θ_{s+1} 값이 된다.

피험자의 능력추정에 결합 최대 우도 추정법 사용 시 다음 두 경우는 피험자의 능력추정이 곤란하다. 첫째, 피험자가 문항을 하나도 맞추지 못하였을 때 피험자의 능력이 음의 무한대이다. 둘째, 피험자가 모든 문항에 정답을 하였을 때 피험자의 능력은 양의 무한대이다. 그러나 최근에 제안된 주변 최대 우도 추정법과 베이지안 추정법을 사용하면 0점을 받은 피험자의 능력모수치는 물론 만점을 받은 피험자의 능력모수치도 모두 추정할 수 있다²⁹⁾.

4. 문항분석을 위한 컴퓨터 프로그램

고전검사이론은 문항분석에 적용할 때 그 계산과정이 간편하나 많은 적용상의 문제점을 갖고 있다. 그러나 문항반응이론은 많은 장점을 가지고 있음에도 계산과정의 복잡함으로 인하여 컴퓨터를 활용한 분석이 필요하다. 여기서 개인용 컴퓨터에 사용 가능한 프로그램을 중심으로 고찰하고자 한다.

1) BICAL, BILOG, LOGIST

BICAL은 Rasch 모형 즉 1-모수 로지스틱 모형에 의하여 문항과 능력의 모수치를 추정하는 프로그램으로 Wright와 Mead에 의하여 제작되었으며, 추정된 문항모수 추정치는 문항변별도가 1.0으로 고정되어 있기 때문에 실제의 문항응답자료와 부합되지 않는 문제점을 가지고 있다⁸⁾.

LOGIST 프로그램은 3-모수 로지스틱 모형은 물론 2-모수 로지스틱 모형, 1-모수 로지스틱 모형에 의한 문항 특성 추정이 가능하다. LOGIST는 결합 최대우도 추정법에 의하여 문항을 분석하므로 모든 피험자가 정답을 하거나, 어느 피험자도 정답을 하지 못한 문항의 특성 추정이 불가능하며, 모든 문항에 정답을 한 피험자 또는 어느 한 문항도 정답을 하지 못한 피험자의 능력 추정이 불가능하다.

BILOG 프로그램은 LOGIST의 문제점을 해결할 수 있는 주변 최대 우도 추정법에 의하여 문항 특성 및 피험자의 능력을 추정하는 프로그램이다. LOGIST보다 미래 지향적이기는 하나, 프로그램의 실행을 위해서는 부 프로그램의 작성과 이에 따르는 옵션 사항의 입력을 위한 프로그램 작성이 필요하고, 짧은 검사도구와 소수의 피험자에 의한 추정은 정확하지 않다.

2) ASC의 ITEMAN과 XCALIBRE

1995년에 Wisconsin University에서 발표된 이 프로그램은 Assessment System Corporation의 Item and Test Analysis Package에 포함되어 있다.

윈도우즈용 문항분석 프로그램으로 ITENAN, XCALIBRE 등을 포함하며, 검사와 문항에 대하여 그 특성을 추정할 수 있다³⁰⁾.

ITEMAN은 고전검사이론에 의한 문항특성 프로그램으로 자료의 입력 설정과 출력 설정, 그리고 출력물의 자료 형태를 지정하는 부문으로 되어 있다. 출력물을 지정할 때, 피험자의 검사점수와 문항 특성치를 동시에 계산할

수 있도록 지정 할 수 있으며, 이분 상관 계수, 이분점 상관 계수, 피험자의 점수별 통계 등을 출력하도록 정할 수 있다.

XCALIBRE는 문항반응이론을 적용한 프로그램으로 주변 최대 우도 추정법에 의하여 2-모수, 3-모수 로지스틱 모형에 의한 추정이 가능하며, 추정의 실행에 앞서 문항 특성의 추정에 필요한 여러 가지 초기치를 지정할 수 있다. 피험자 능력의 추정을 위한 주변 최대 우도 추정법과 베이저안 통계를 위한 EAP(expected a posterior)방법과 MAP(maximum a posterior) 방법을 선택할 수 있으며, 문항 모수 추정의 사전 분포 능력을 지정할 수 있도록 되어 있다.

추정의 실행 결과 첫째, 피험자의 평균점수, 표준편차, 추정의 신뢰도, 추정된 문항 모수치의 통계치, 모수치의 오차, 적용한 문항 반응 모형의 적합도, 검사 정보 곡선, 검사 특성 곡선 등을 산출할 수 있고, 둘째, 피험자의 능력 모수 추정치가 산출되어진다.

III. 研究의 對象 및 方法

1. 研究의 對象

본 연구에서 분석한 자료는 1999년도 제주도내 7개 고등학교 1학년 공통 과학 교과목의 중간고사 및 기말고사에서 실시하였던 총괄평가의 학생 응답 자료로, 학교 유형별로는 인문계 5개교, 실업계 2개교이다.

분석대상 학교의 분석대상 검사수 및 문항수는 표 6과 같다.

표 6. 자료 분석 대상 및 분석 문항 수

학교별	분석대상 검사수	분석대상 문항수
ja*	4	80
ji*	4	80
jm**	1	24
nj*	2	47
nn*	2	56
sa**	4	88
sb*	1	25
계	18	400

*:인문계, **:실업계

검사를 구성하는 문항의 수와 피험자의 수는 제한을 두지 않았고, 검사 중에서 5-지 선다형 문항만을 본 연구의 분석대상으로 사용하였으며, 이분반응 모형으로만 분석하였다. 한편 주요인 분석에 필요한 데이터로 변환이 용이한 ji 학교의 4개 검사를 대상으로 요인분석을 실시하였다.

2. 연구의 方法

고전검사이론에 의한 문항의 특성과 피험자의 검사점수는 ASC의 Item

and Test Analysis Package에 포함되어 있는 ITEMAN(v 3.50d)을 이용하여 계산하였으며, 문항반응이론의 3-모수 로지스틱 모형에 의한 문항특성과 피험자능력 모수는 XCALIBRE(v 1.10c)를 이용하여 주변 최대 우도 추정 방법으로 추정하였다.

문항과 검사에 대한 자료의 분석 과정은 다음과 같다.

단계1. 고등학교 총괄평가의 피험자 응답자료는 학교에서 실시하였던 평가의 결과 응답한 OMR(Optical Mark Reader)용 응답지를 OMR에 의해 읽어들이고, 응답 결과를 텍스트 파일 또는 엑셀 파일로 저장한 후 이를 프로그램에서 요구하는 컴퓨터 언어인 ASCII 데이터 파일로 변환한 후 이를 사용하였다.

단계2. ITEMAN에서 프로그램 실행을 통하여 고전검사이론의 통계량인 문항난이도와 문항변별도 그리고 검사신뢰도(Cronbach α)를 산출하고, 피험자 검사점수를 계산하였다. 문항변별도의 계산에서 상위그룹은 검사점수 기준 상위 27%, 하위그룹은 검사점수 기준 하위 27%로 정하여 계산하였다.

단계3. 문항반응이론 중 3-모수 로지스틱 모형을 적용할 수 있는 XCALIBRE 프로그램을 이용하여 주변 최대 우도 추정 방법으로 문항별 변별도 모수(a), 난이도 모수(b), 추측도 모수(c) 및 피험자별 능력 모수 추정치(θ), 검사의 신뢰도(K-R 21)를 산출하였다. 이 때 모수 추정에 사용한 초기치 값은 다음 표 7과 같다.

표 7. 문항모수 추정의 초기치

통계량 \ 모수	a	b	c
평균	0.75	0.00	1/#Alt
표준편차	0.12	2.00	0.5c ~ 1.5c

#Alt : 각 문항에 포함된 선택항목의 수

단계4. 고전검사이론과 문항반응이론의 3-모수 로지스틱 모형에 의한 검사의 특성을 비교하기 위해 고전검사이론에 의한 검사 신뢰도(Cronbach α)와 문항반응이론에 의한 검사의 신뢰도(K-R 21) 및 평균정보량, 기대정보량

을 산출하였고, 이들 사이의 상관을 알아보기 위해 Pearson 상관계수를 산출하였다.

단계5. 검사가 문항반응이론의 기본 가정인 단일차원성의 만족 여부를 확인하기 위해 SPSS PC+를 이용 주요인분석을 통해 요인별로 설명된 총분산을 계산하였다.

단계6. 두 평가모형에 의한 문항 분석의 결과 고전검사이론의 검사점수와 문항반응이론에 의한 피험자 능력 모수치의 상관 관계는 SPSS PC+를 이용하여 kendall의 순위상관계수로 산출하여 나타내었다.

단계7. 단계 2와 단계 3의 결과 계산된 문항난이도, 문항변별도 및 문항추측도에 대해 인문계와 실업계로 나누어 학교 형태에 따른 평균값의 차이를 T검증 방법으로 비교하였다.



IV. 研究의 結果

1. 問項 特性의 推定 結果 및 解析

1) 古典檢査理論에 의한 分析

(1) 問項難易度

고전검사이론에 의해 고등학교 공통과학 교과서의 총괄평가 검사의 문항에 대해 문항난이도를 계산한 결과 표 1의 기준에 따라 학교별로 그 분포를 나타내면 표 8과 같다(부록 1 참조).

표 8. 고전검사이론에 의한 문항난이도 분석

학교별	검사 수	문항수	난이도분포						난이도 평균	표준 편차	최소값	최대값
			0.25이하		0.25~0.75		0.75이상					
			n	%	n	%	n	%				
ja	4	80	3	3.8	55	68.8	22	27.5	0.64	0.16	0.16	0.91
ji	4	80	2	2.5	52	65.0	26	32.5	0.64	0.20	0.14	0.96
jm	1	24	0	0	16	66.7	8	33.3	0.71	0.11	0.49	0.88
nj	2	47	1	2.1	38	80.9	8	17.0	0.59	0.18	0.23	0.94
nn	2	56	2	3.6	42	75.0	12	21.4	0.60	0.19	0.15	0.92
sa	4	88	11	12.5	71	80.7	6	6.8	0.48	0.19	0.08	0.85
sb	1	25	2	8.0	18	72.0	5	20.0	0.56	0.23	0.21	0.89
전체	18	400	21	5.3	292	73.0	87	21.7				

표 8에서 고전검사이론에 의한 검사 문항의 난이도는 전체적으로 쉬운 문항이 어려운 문항보다 4.09배 정도로 많이 출제되고 있음을 볼 수 있다. 한편 검사 당 쉬운 문항은 5.11문항, 어려운 문항은 1.16문항을 출제하고 있어 대체적으로 고전검사이론에 의한 문항분석의 결과 쉬운 문항은 약 21.7%, 어려운 문항은 약 5.3%로 출제되고 있다. 대체적으로 표본조사 결과 고등학교 공통과학 교과서의 총괄평가의 경우 그 문항은 쉬운 편이며, 쉬운 문항은 전체 문항 중에서 21.7%를 출제하고 있는 것으로 나타났다.

(2) 問項辨別度

고전검사이론에 의해 고등학교 총괄평가 검사의 문항에 대해 문항변별도를 계산한 결과 표 2의 기준에 따른 분포는 표 9와 같다.

표 9. 고전검사이론에 의한 문항변별도 분석

학교	문항 수	변별도분포										변별도 평균	표준 편차	최소값	최대값
		0.40 이상		0.30 ~ 0.39		0.20 ~ 0.29		0.10 ~ 0.19		0.10미만					
		n	%	n	%	n	%	n	%	n	%				
ja	80	44	55.0	24	30.0	9	10.3	3	4.7	0	0	0.42	0.12	0.12	0.71
ji	80	45	56.2	20	25.0	8	10.0	6	7.5	1	1.3	0.37	0.13	0.03	0.69
im	24	21	87.5	3	12.5	0	0	0	0	0	0	0.62	0.13	0.39	0.81
nj	47	34	72.4	3	6.4	5	10.6	5	10.6	0	0	0.47	0.18	0.14	0.82
nn	56	28	50.0	15	26.8	9	16.1	4	7.1	0	0	0.41	0.14	0.11	0.74
sa	88	48	54.5	19	21.6	10	11.4	7	8.0	4	4.5	0.41	0.17	-0.31	0.76
sb	25	13	52.0	6	24.0	5	20	1	4.0	0	0	0.41	0.13	0.10	0.62
전체	400	233	58.0	90	22.5	46	11.5	26	6.5	5	1.25				

표 9에서 고전검사이론에 의해 계산된 문항변별도는 전체적으로 변별력이 높은 문항을 많이 출제하고 있는 것으로 나타났다. 변별력이 0.40이상인 문항의 비율이 55.0%이고, 변별력이 떨어지는 문항으로 변별도가 0.20미만인 문항의 비율은 전체 의 7.0%로 나타났다. 변별력이 0.30 이상인 문항이 76.0%로 전체 중 차지하는 비율이 매우 높다. 즉 고등학교의 총괄평가 문항의 변별도는 매우 높은 것으로 나타나 상위집단과 하위집단을 매우 잘 변별해 주는 것으로 볼 수 있다.

2) 問項反應理論에 依한 分析

(1) 問項辨別度

문항반응이론의 3-모수 로지스틱 모형에 의해 고등학교 공통과학 교과의 총괄평가 검사의 문항에 대해 문항변별도를 추정한 결과를 표 3의 기준에 따라 나타내면 표 10과 같다.

표 10. 문항반응이론에 의한 문항변별도 모수 분석

학교	문항수	변별도분포								평균	표준 편차	최소 값	최대값
		0.35 ~ 0.64		0.65 ~ 1.34		1.35 ~ 1.69		1.70 이상					
		n	%	n	%	n	%	n	%				
ja	80	28	35.0	52	65.0	0	0	0	0	0.68	0.11	0.44	0.99
ji	80	25	31.3	55	68.7	0	0	0	0	0.70	0.09	0.51	0.90
jm	24	0	0	17	70.8	7	29.2	0	0	1.25	0.03	0.97	1.60
nj	47	5	10.6	42	89.4	0	0	0	0	0.83	0.14	0.60	1.26
nn	56	31	55.4	25	44.6	0	0	0	0	0.63	0.09	0.41	0.86
sa	88	1	1.1	87	98.9	0	0	0	0	0.79	0.08	0.69	1.07
sb	25	2	8.0	23	82.0	0	0	0	0	0.78	0.02	0.01	0.96
전체	400	92	23.0	301	75.2	7	1.8	0	0				

표 10의 고등학교 총괄평가의 문항변별도 지수에 대한 분포에서 적절한 수준인 0.65~1.34의 변별도 값을 갖는 문항이 전체 400문항 중에서 301문항을 출제하여, 적절한 문항이 75.2%로 매우 높은 비율로 출제되고 있다.

문항반응이론의 3-모수 로지스틱 모형에 의한 문항변별도를 기준으로 하여 문항을 선정하여 검사를 실시할 때 검사에 적합한 문항은 0.65이상의 변별도 값을 갖는 문항으로 308문항으로 77.0%를 차지하였다.

한편 문항의 변별도가 낮은 0.65미만 문항은 92문항으로 23.0%를 차지하고 있고, 이런 문항은 삭제 또는 수정을 통해 사용하는 것이 바람직하다.

(2) 問項難易度

문항반응이론의 3-모수 로지스틱 모형에 의한 고등학교 고통과학 교과와 총괄평가 검사의 문항에 대해 문항난이도 모수를 추정된 결과를 표 4의 기준에 따라 나타내면 표 11과 같다.

표 11에서 문항반응이론의 3-모수 로지스틱 모형에 의한 문항난이도 모수치 분포는 전체적으로 쉽거나 매우 쉬운 문항이 32.3%, 어렵거나 매우 어려운 문항은 33.5%, 중간인 문항은 34.2%를 출제하고 있는 것으로 나타나고 있어, 문항반응이론의 3-모수 로지스틱 모형에 의한 문항분석의 결과 문항난이도 모수는 매우 고른 분포를 나타내고 있다. 이는 분석과정에서 사전 모

수 능력의 지정에서 평균값이 0이 되도록 하고, 표준편차는 2.00이 되도록 정한데서 기인한다고 할 수 있다.

표 11. 문항반응이론에 의한 문항난이도 모수 분석

학교	문항 수	난이도분포										난이도 평균	표준 편차	최소값	최대값
		-2.00 이하		-2.00 ~ -0.50		-0.50 ~ 0.50		0.50 ~ 2.00		+2.00 이상					
		n	%	n	%	n	%	n	%	n	%				
ja	80	1	1.2	26	32.5	36	45.0	12	15.0	5	6.3	-0.14	1.04	-2.24	3.00
ji	80	5	6.3	27	33.8	24	30.0	19	23.8	5	6.3	-0.18	1.43	-3.00	3.00
jm	24	0	0	15	62.5	8	33.3	1	4.2	0	0	-0.57	0.58	-1.48	0.84
nj	47	2	4.3	11	23.4	20	42.6	13	27.6	1	2.1	0.02	1.06	-2.25	2.09
nn	56	4	7.1	16	28.6	18	32.1	11	19.6	7	12.5	-0.03	1.37	-2.90	3.00
sa	88	0	0	10	11.4	26	29.5	41	46.6	11	12.5	0.76	1.11	-1.46	3.00
sb	25	0	0	10	40.0	5	20.0	8	32.0	2	8.0	0.22	1.36	-1.85	2.58
전체	400	12	2.3	115	30.0	137	34.2	105	26.5	31	7.0				

학교별로 세분화하였을 때, jm, sb의 경우 난이도 -0.5이하인 쉬운 문항이 각각 62.5, 40.0%를 차지하고 있는 것으로 나타났으며, sa의 경우 난이도 0.5이상인 어렵거나 매우 어려운 문항의 비율이 각각 59.1%를 나타내고 있어 어려운 문항의 출제 비율이 높게 나타나고 있다.

한편 문항 모수치 추정 과정에서 문제점이 발견되는 문항인 난이도 3.0이상인 문항이 4개 학교에서 15문항이 출제되어 있다. 이러한 문항은 검사를 시행할 때 문항의 선정에서 제외하거나 수정 과정을 거쳐 선정되어야 한다.

(3) 問項推測度

문항반응이론의 3-모수 로지스틱 모형에 의한 고등학교 공통과학 교과와 총괄평가 검사의 문항에 대한 문항추측도의 추정 결과를 표 5의 기준에 따라 나타내면 표 12와 같다.

표 5의 표현에 의하면 추측도 0.20이하의 양호한 문항이 전체 400문항 중 41.0%로 165문항을 차지하고 있으며, 추측도가 0.20~0.30의 높은 문항은 전

체 400문항 중 51.3%인 205문항의 분포를 보이고 있다.

표 12. 문항반응이론에 의한 문항추측도 모수 분석

학교	문항수	추측도분포						추측도 평균	표준 편차	최소값	최대값
		0.20이하		0.20~0.30		0.30이상					
		n	%	n	%	n	%				
ja	80	14	17.4	61	76.2	5	6.3	0.23	0.05	0.11	0.38
ji	80	36	45.0	32	40.0	12	15.0	0.22	0.05	0.12	0.38
jm	24	22	91.7	2	8.3	0	0	0.14	0.04	0.09	0.26
nj	47	30	63.8	15	31.9	2	4.3	0.19	0.05	0.12	0.38
nn	56	10	17.9	42	75.0	5	8.9	0.23	0.08	0.06	0.37
sa	88	40	45.5	41	46.6	7	7.9	0.22	0.07	0.12	0.58
sb	25	13	52.0	12	48.0	0	0	0.20	0.04	0.14	0.28
전체	400	165	41.0	205	51.3	31	7.7				

한편 검사에서 수정하거나 삭제하여 출제에서 제외시켜야 할 추측도 0.30 이상의 문항이 7.7%로, 31문항이 출제되고 있다. 31문항은 추측에 의해 피험자가 답을 할 수 있는 확률이 매우 높은 문항으로 검사에서 거의 효용이 없는 문항이다. 추측도 모수의 평균값이 0.20이상인 검사가 절반 이상을 차지하고 있어 고등학교 총괄평가의 문항의 선정은 주의를 기울일 필요가 있다.

추측도 모수는 능력이 전혀 없는 피험자가 문항에 정답을 할 수 있는 확률로 난이도 모수가 낮으면 추측도 모수가 높게 나타난다⁸⁾는 연구결과에 비추어 볼 때 문항난이도와 문항추측도를 함께 고려하여 문항이 작성·제시되어야 하며, 문항추측도가 지나치게 높은 문항은 출제 과정에서 수정되거나, 삭제함으로써 검사의 질을 높일 수 있어야 한다.

2. 檢査 特性의 分析

고전검사이론과 문항반응이론에 의한 검사의 신뢰도를 나타내는 값으로 고전검사이론에 의한 분석의 결과 나타난 Cronbach α 와 문항반응이론에 의한 분석의 결과 나타난 K-R 21, 그리고 문항반응이론에 의한 분석과정에서 나타난 검사의 기대 또는 평균정보량을 검사별로 나타내면 표 13과 같고 표 14는 검사의 신뢰도와 정보량과의 상관 관계를 분석한 결과이다.

표 13. 검사의 신뢰도와 정보량

검사	Cronbach α	K-R 21	기대정보량	평균정보량
ja1	0.706	0.704	2.252	1.729
ja2	0.733	0.733	2.815	2.012
ja3	0.752	0.752	2.984	2.081
ja4	0.769	0.769	2.935	2.061
ji1	0.720	0.721	2.630	1.992
ji2	0.718	0.718	2.724	2.043
ji3	0.724	0.725	2.636	2.046
ji4	0.663	0.686	2.112	1.830
jm1	0.926	0.927	9.879	5.697
nj1	0.780	0.780	3.811	2.849
nj2	0.835	0.835	5.284	3.423
nn1	0.742	0.742	3.088	2.423
nn2	0.868	0.750	3.151	2.418
sa1	0.713	0.716	2.220	1.803
sa2	0.775	0.777	3.475	2.494
sa3	0.641	0.749	3.212	2.754
sa4	0.796	0.796	3.857	2.790
sb1	0.761	0.768	3.642	2.912
평균	0.750	0.758	3.643	2.643

표 14. 검사의 신뢰도와 정보량 상관관계 분석

		상관계수		
		K-R 21	기대정보량	평균정보량
Pearson 상관	K-R 21	1.000	0.958**	0.953**
	기대정보량	0.958**	1.000	0.981**
	평균정보량	0.953**	0.981**	1.000
유의확률(양쪽)	K-R 21	.	0.000	0.000
	기대정보량	0.000	.	0.000
	평균정보량	0.000	0.000	.
제곱합 및 교차곱	K-R 21	0.063	1.888	1.036
	기대정보량	1.888	61.450	33.237
	평균정보량	1.036	33.237	18.683
공분산	K-R 21	0.004	0.111	0.061
	기대정보량	0.111	3.615	1.955
	평균정보량	0.061	1.955	1.099
N	K-R 21	18	18	18
	기대정보량	18	18	18
	평균정보량	18	18	18

** 상관계수는 0.01 수준에서 유의함

표 13에서 고전검사이론에 의한 검사의 신뢰도 측정치인 Cronbach α 는 문항반응이론에 의한 검사의 신뢰도 K-R 21과 비슷한 결과를 보였으며, 표 14에서 보는 바와 같이 문항반응이론에 의한 검사의 신뢰도 K-R 21과 평균 정보량 및 기대정보량사이의 상관은 상관계수가 0.953, 0.958로서 매우 높은 상관을 보여주고 있으며, 모든 검사가 유의확률 0.01 수준에서 유의미한 것으로 나타났다.

표 14의 결과는 검사의 신뢰도와 기대되는 정보량, 그리고 평균 검사정보량 사이에 매우 높은 상관이 있음을 보여주고 있어, 검사에 대한 특성을 나타낼 때 검사의 신뢰도를 문항반응이론의 정보량으로 대신할 수 있음을 보여준다.

3. 問項反應理論의 適用可能性 分析

문항반응이론의 기본 가정인 단일차원성의 만족 여부를 확인하기 위하여 주요인 분석을 실시하였다. 본 연구의 주요인 분석에 사용한 자료는 주요인 분석에 적합하도록 변환이 용이한 ji의 4개 검사를 선정하였으며, SPSS PC+ 프로그램의 요인분석을 통해 각 검사의 단일차원성의 만족 여부를 확인한 결과 성분점수의 분산 중에서 제 1요인에 의해 설명할 수 있는 분산의 비율을 나타내면 표 15와 같다(부록 2 참조).

표 15. 요인분석의 결과

검사	문항수	고유치 1.0 이상의 성분 수	성분점수의 분산 중에서 제 1요인에 의해 설명할 수 있는 분산의 비율
1	20	8	15.06
2	20	7	17.40
3	20	6	18.05
4	20	7	15.43

표 15에서 고유치가 1.0이상인 성분 요인의 수는 8, 7, 6 및 7개 요인으로 각각 선택되어졌고 성분점수의 분산 중에서 제1요인에 의해 설명할 수 있는 분산 비율은 각각 15.06, 17.40, 18.05 및 15.43% 로 나타났다.

Reckase²¹⁾는 검사의 성분점수의 분산 중에서 제 1요인에 의해 설명할 수 있는 분산의 비율이 20%는 되어야 한다고 하였다. 그리고 10%가 안될지라도 능력 모수는 안정된 추정치를 얻을 수 있지만 문항 모수의 추정치는 안정적이지 못하다고 하였다. 이와 같은 선행연구의 결과로 볼 때 본 연구의 결과중 성분점수의 분산중에서 제 1요인에 의해 설명할 수 있는 비율은 미달된다고 볼 수 있다. 그러나 Hambleton와 Swaminathan은 Reckase의 연구 결과(61)는 임의적 준거라고 하고 있다. 한편 박순옥³¹⁾은 대학입학학력고사 문항의 단일차원성의 만족에 대한 연구 결과 15% 정도로도 단일 차원성을 만족한다고 하고 있으며, 정재임²²⁾은 중학교의 교사출제 영어과 평가 문항의 단일 차원성 검증에서 요인분석의 결과 제 1인자에 의해서 설명할 수 있는 변량이 17.3%, 17.2%로 단일 차원성을 만족한다고 하고 있다.

따라서 Hambleton의 17%, 박순옥의 15%, 정재임의 결과와 비교할 때, 본 연구의 결과는 문항반응이론의 기본 가정인 단일차원성 가정에 어느 정도는 만족하는 결과로 나타났으며, 고등학교 공통과학 교과의 총괄평가는 문항반응이론의 적용이 가능한 검사라고 판단할 수 있다.

한편 국소독립성의 가정은 단일차원성에 대한 당연한 귀결이며, 본 연구의 문항반응이론에 의한 국소독립성의 가정은 만족한다고 할 수 있다.

4. 古典檢査理論에 의한 檢査點數와 3-母數 로지스틱 模型에 의한 能力 母數 推定值의 相關關係

문항 분석의 결과 고전검사이론에 의한 피험자의 검사점수와 문항반응이론의 3-모수 로지스틱 모형에 의한 피험자의 능력 모수 추정치를 비교하였다. 각 검사마다 고전검사이론에 의한 문항 분석 결과 주어지는 피험자 검사점수와 문항반응이론의 3-모수 로지스틱 모형에 의한 피험자 능력 추정치를 산출하여 이 사이의 Kendall의 순위 상관계수 τ 를 산출한 결과는 다음의 표 16과 같다.

표 16. 피험자의 검사점수와 능력 모수추정치의 상관관계

검 사	Kendall τ
ja1	0.938**
ja2	0.918**
ja3	0.937**
ja4	0.930**
ji1	0.914**
ji2	0.908**
ji3	0.918**
ji4	0.862**
jm1	0.954**
nj1	0.910**
nj2	0.937**
nn1	0.914**
nn2	0.960**
sa1	0.866**
sa2	0.887**
sa3	0.714**
sa4	0.929**
sb1	0.872**
평균치	0.904**

** : .01수준에서 유의함

평균 상관계수 0.904로서 매우 높은 상관을 보이고 있으며, 유의도는 0.000으로 통계적으로 유의미한 것으로 나타났다. 이것은 김혜기³²⁾의 연구 결과 두 이론에 의한 검사 점수의 상관관계는 상관계수 0.876으로 유의수준 0.05에서 통계적으로 유의미한 것으로 나타났다는 것과 일치한다.

따라서 전통적인 고전검사이론에 의한 검사 점수를 대신하여 고전검사이론이 갖는 단점을 보완하고 문항반응이론이 갖는 장점을 살리는 측면에서 ASC 프로그램을 사용한 문항 분석의 결과 주어지는 학생들의 능력 추정치를 표준화함으로서 학생들의 능력을 새롭게 표시할 수 있다는 가능성을 말해 준다.

5. 學校 類型別 問項 特性의 比較

1) 古典檢査理論에 의한 問項難易度 比較

고전검사이론에 의한 문항 분석의 결과 문항난이도에 대하여 표본 조사한 연구대상을 인문계 고등학교와 실업계고등학교로 나누어 문항난이도의 평균값의 차이가 있는지 알아본 결과는 표 17과 같다.

표 17. 고전검사이론에 의한 문항난이도의 비교

구분		M	SD	t	df	p
고전난이도	인문	0.62	0.19	4.277	398	0.000
	실업	0.53	0.20			

M:평균, SD:표준편차, t: T검정 통계량(t점수), df:자유도

표 17에서 등분산 검정결과 $F=0.683$ 으로 5% 수준에서 등분산이 가정된다는 가정하에 T검정을 실시한 결과 인문계고등학교와 실업계고등학교에서 실시했던 공통과학교과의 총괄평가 문항의 난이도의 평균값은 0.62, 0.53이며, 그 차이 0.09는 5% 수준에서 통계적으로 유의하다. 즉 인문계고등학교와 실업계 고등학교의 공통과학교과의 총괄평가 문항은 고전검사이론에 의해 분석한 결과 문항난이도에는 인문계고등학교가 실업계고등학교보다 더 쉬운 문항을 검사에 제시하고 있다고 할 수 있다.

2) 古典檢査理論에 의한 問項辨別度の 比較

고전검사이론에 의한 문항분석의 결과 인문계고등학교와 실업계 고등학교의 총괄평가 문항의 변별도의 평균값에 대하여 차이가 있는지를 알아본 결과는 표 18과 같다.

표 18. 고전검사이론에 의한 문항변별도의 비교

구분		M	SD	t	df	p
고전변별도	인문	0.41	0.14	-2.387	398	0.018
	실업	0.46	0.18			

표 18에서 T 검정을 위해 먼저 등분산이 가정여부를 검증한 결과 $F=0.789$ 로 유의확률 $0.005 < 0.05$ 로 두 집단의 변별도의 등분산은 가정되지 않았으며, 이에 따라 변별도의 평균의 차이에 대한 T검정 결과 $t=-2.387$ 이며, 평균값이 각각 0.41, 0.46로 그 차이 0.05은 0.02의 표준오차 범위 내에서 유의확률 $0.018 < 0.05$ 로 통계적으로 유의하다. 즉 실업계고등학교의 공통과학과의 총괄평가 문항의 변별도가 인문계고등학교의 경우보다 변별력이 높다고 할 수 있다.

3) 問項反應理論에 의한 問項辨別度の 비교

문항반응이론의 3-모수 로지스틱 모형에 의한 문항변별도의 추정결과 인문계고등학교와 실업계고등학교의 문항변별도의 평균값에 대한 비교 결과는 표 19와 같다.

표 19. 3-모수 로지스틱 모형에 의해 추정된 문항변별도의 비교

구분		M	SD	t	df	p
문항변별도	인문	0.72	0.13	-7.908	398	.000
	실업	0.90	0.22			

표 19에서 문항반응이론에 의한 분석결과 인문계고등학교와 실업계고등학교의 공통과학 총괄평가 문항의 변별도는 $F=35.394$ 로 유의확률 $0.000 < 0.05$ 의 수준에서 분산이 동일하지 않으며, 이에 따른 T검정 결과는 $t=-7.908$ 로 평균차이 0.18만큼 0.02의 오차범위 내에서 $0.000 < 0.05$ 수준에서 통계적으로 유의하다. 즉 실업계고등학교의 경우가 인문계고등학교의 경우보다 변별도의 평균값이 크다고 할 수 있어, 변별도의 평균은 유의미한 차이가 있다고 할 수 있다.

즉 문항반응이론에 의한 문항분석의 결과 실업계고등학교가 인문계고등학교보다 변별력이 높은 문항을 출제하고 있는 것으로 나타났다.

4) 問項反應理論에 의한 問項難易度の 비교

문항반응이론의 3-모수 로지스틱 모형에 의한 문항난이도의 추정결과 인문계고등학교와 실업계고등학교의 문항난이도의 평균값에 대한 비교 결과는 표 19와 같다.

표 20. 3-모수 로지스틱 모형에 의해 추정된 문항난이도의 비교

구분		M	SD	t	df	p
문항난이도	인문	-0.08	1.21	-4.198	398	0.000
	실업	0.48	1.16			

표 20에서 문항반응이론에 의한 문항분석 결과 인문계고등학교와 실업계고등학교 총괄평가 문항의 문항난이도에 대해 차이를 검증한 결과는 $F=0.042$ 로 유의확률 $0.838 > 0.05$ 로 분산이 같음을 가정할 수 있고, 따라서 T검정 결과는 $t=-4.198$ 로 $0.000 < 0.05$ 의 수준에서 문항난이도의 평균값은 통계적으로 유의한 차이가 있다고 할 수 있다. 즉 실업계고등학교에서 출제되는 문항은 인문계고등학교의 경우보다 어려운 문항을 출제하고 있는 것으로 나타났다.

5) 問項反應理論에 의한 問項推測度の 비교

문항반응이론의 3-모수 로지스틱 모형에 의한 문항추측도의 추정결과 인문계고등학교와 실업계고등학교의 문항추측도의 평균값에 대한 비교 결과는 표 21과 같다.

표 21. 3-모수 로지스틱 모형에 의해 추정된 문항추측도의 비교

구분		M	SD	t	df	p
문항추측도	인문	.20	0.07	-0.200	398	0.841
	실업	.20	0.07			

표 21에서 문항반응이론에 의한 문항분석의 결과 인문계고등학교의 추측도 평균값은 0.20이고, 실업계고등학교의 경우는 0.20으로 그 차이는 0.00이

었다. 이 차이에 대해 통계적으로 의미가 있는지를 확인한 결과 $F=0.199$, 유의확률 $0.655 > 0.05$ 수준에서 등분산을 가정할 수 있었으며, 이에 따라 두 경우의 문항반응이론에 의한 문항추측도의 평균의 차에 대한 T검정 결과는 $t=-0.200$ 이며, 유의확률 $0.841 > 0.05$ 로 통계적으로 유의하지 않았으며, 두 경우의 문항추측도의 평균값은 차이가 없다고 할 수 있다.



V. 結 論

본 연구는 Assessment System Corporation의 Item and Test Analysis Package를 이용하여 고등학교 공통과학 총괄평가 문항에 대해 분석하였다.

고전검사이론에 따라 총괄평가 문항의 문항난이도, 문항변별도를 계산하였으며, 문항반응이론의 3-모수 로지스틱 모형에 의해 문항난이도, 문항변별도, 문항추측도 등을 계산하여, 문항의 양호한 정도를 파악하였다.

주요인분석을 통해 문항반응이론의 기본 가정인 단일차원성의 만족 여부를 확인하였으며, 추정된 피험자의 검사점수와 능력 추정치의 상관 관계를 알아보았고, 인문·실업계고등학교의 총괄평가 문항의 모수치의 평균사이에는 차이가 있는지를 확인하여 다음과 같은 결과를 얻었다.

첫째, 고전검사이론에 의한 문항난이도 분석의 결과 난이도 0.75이상의 쉬운 문항이 난이도 0.25이하인 어려운 문항보다 4.09배의 많은 빈도를 보여 쉬운 문항을 많이 출제하고 있으며, 문항변별도의 지수 0.40이상의 변별력이 높은 문항이 분석 대상 문항 중에서 55.0%로 변별력이 높은 문항을 많이 출제하고 있는 것으로 나타났다.

둘째, 문항반응이론에 의한 문항변별도는 75.2%가 변별도 0.65~1.34의 적절한 변별도지수를 갖는 문항을 출제하고 있어, 상·하위집단을 매우 잘 변별해 주고 있으며, 문항난이도는 매우 쉬운 문항에서 매우 어려운 문항까지 정규적 분포에 가깝게 출제가 되고 있고, 문항추측도는 0.20이하의 문항과 0.20이상의 문항이 41.0%, 59.0%의 비율로 출제되고 있었으며, 또 0.30이상의 수정하거나 검사에서 제외하여야 할 문항도 7.7%정도로 31문항이 출제되고 있는 것으로 나타났다.

셋째, 연구에 사용되었던 검사의 응답 자료 중 4개 검사에 대해 문항반응이론의 기본 가정인 단일차원성의 만족 여부를 확인한 결과 어느 정도는 만족하고 있어 문항반응이론이 고등학교 공통과학 교과와 평가 문항의 분석에

적용이 가능함을 보여 주고 있다.

넷째, 피험자의 고전검사이론에 의한 검사점수와 문항반응이론에 의한 능력추정치 사이의 상관은 상관계수의 평균이 0.904로 매우 높은 상관을 보이고 있어 문항반응이론의 3-모수 로지스틱 모형에 의한 피험자의 능력 추정치가 고전검사이론에 의한 피험자의 검사점수를 대신할 수 있음을 보여주고 있다.

다섯째, 고전검사이론과 문항반응이론의 3-모수 로지스틱 모형에 의한 문항 모수치를 인문계고등학교와 실업계고등학교에 대해 비교한 결과 문항 반항반응이론의 3-모수 로지스틱 모형에 의한 문항추측도 모수를 제외한 모든 경우 통계적으로 유의미한 차이가 있는 것으로 나타났다.

본 연구의 결과 적절한 검사 문항의 선정을 위해 필요한 기준이 될 수 있는 문항특성을 제시하였고, 학생들의 능력 정도를 정확하게 추정할 수 있는 신뢰도가 높은 검사와 문항구성으로 교육의 효과 정도를 파악하는 것이 필요하다고 할 수 있다. 따라서 고등학교 공통과학 교과서의 총괄평가에서 이미 출제되었던 문항들의 특성을 기준으로 양호한 문항을 선정하고, 다음에 시행할 총괄평가에 활용이 가능하다는 결론을 내릴 수 있다.

고등학교 공통과학 교과서의 총괄평가의 표본이 단일 차원성을 어느 정도 만족시키는 것으로 나타나, 문항반응이론의 적용가능성을 시사해 준다.

본 연구는 문항반응이론이 학생들의 능력을 평가하는 과정에 적용되어야 하며, 이를 위한 문항분석용 컴퓨터 프로그램의 개발과 보급이 이루어져야 함을 제안하고자 한다. 또한 검사 시행 후 문항반응이론에 의한 문항분석의 결과로부터 양호한 문항을 추출해 내고, 이를 토대로 한 문제은행 작성과 이를 활용한 총괄평가의 시행을 제안한다.

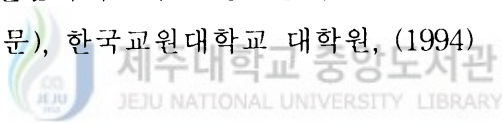
참 고 문 헌

- 1) 이종성 : 문항반응이론과 응용, 대광문화사, (1991)
- 2) 송미랑 : 고전검사 이론을 이용한 중학교 과학 참고서의 평가 문항 분석에 관한 연구(석사학위논문), 연세대 교육대학원, (1996)
- 3) 김석우 : 고전적 검사이론의 모순성과 문항반응이론의 타당성, 교육월보 1991년 8월호, (1991)
- 4) 황정규 : 학교학습과 교육평가, 교육과학사, (1998)
- 5) 성태제 : 문항제작 및 분석의 이론과 실제, 학지사, (1991)
- 6) Scriven, M. : The methodology of evaluation, In R. Tyler, R. Gagne & M. Scriven(Eds.) *Perspectives on curricular evaluation*, No. 1., Chicago; Rand McNally.(1976)
- 7) Madaus, G.F. : Handbook on formative and summative evaluation of student learning, New York; McGraw-Hill, (1971)
- 8) 성태제 : 문항제작 및 분석의 이론과 실제, 학지사, (1996).
- 9) 이종성 : 고전검사이론과 문항반응이론, 연세교육과학 제 30집, (1986)
- 10) 성태제 : 중등학교 교육평가 일반연수 교재, 탐라교육원, (1998)
- 11) 임 형 : 고전검사이론에 의한 문항분석, 교육평가 제 3호, (1993)
- 12) 안창규 : 교육 및 심리 검사에 있어서 문항반응이론의 성격과 그 적용성, 교육평가연구 제2권 제2호, (1987)
- 13) 김정환 : 잠재적 특성이론 고찰, 한국교원대학교 교수논총 제7집 제1호, (1991)
- 14) 조병호 : 준거지향검사의 문항선택에 대한 문항반응이론과 전통적 접근방법의 비교 연구(석사학위논문), 한국교원대학교 대학원,(1991)
- 15) Ryan, J. P. : Introduction to latent trait analysis and item response theory. In W.E.hathaway(Ed), *Testing in the schools*. (pp.49-64). Sanfrancisco ; Jossey-Bass. (1983)
- 16) Hambleton, R.K. & Linden, W.J. : Advances in item response theory

- and applications: An introduction. *Applied Psychological Measurement*, 6, pp.373-378, (1982)
- 17) 문용린 : 문항반응이론을 이용한 문항편과성의 추정방법, *교육평가연구* 제 2권 제 2호, (1987)
 - 18) Lord, F.M. : An analysis of the verbal scholastic aptitude test using Birnbaum's three-parameter logistic model. *Educational and Psychological Measurement*, 28, pp. 289-378, (1968)
 - 19) Dragsow, F. & Parsons, C. K. : Application of Unidimensional Item Response Theory Models to Multimimensional Datas. *Applied Psychological Measurement*, 7, pp.189-200, (1983)
 - 20) 안창규, 신석기 : Rasch 모형에 의한 MMPI우울증 척도의 문항 양호도 분석, *교육평가연구*, 제 2권 제 1호, (1987)
 - 21) 이순화 : 문항반응이론을 적용한 일반지능검사의 문항분석(석사학위논문), 숙명여자대학교 대학원, (1989)
 - 22) 정재임 : 교사 출제 영어시험 문항 분석을 위한 문항반응이론의 적용 가능성(석사학위논문), 경북대학교 교육대학원, (1993)
 - 23) Lord, F. M. & Novick, M. R. : *Statistical Theories of Mental Test Scores*. Reading MA; Addison-Wesley, (1968)
 - 24) 서울대학교 교육연구소 : *교육학 대백과사전*, 하우동설, pp.1162-1170, (1998)
 - 25) 윤혜경 : 모든 것이 정답' 혹은 '정답 없음'답지를 포함하는 선다형 문항의 문항특성과 검사특성 분석(석사학위논문), 이화여자대학교 대학원, (1996)
 - 26) 이용환 : 고전적 평가모형과 3-모수 로지스틱 평가모형을 활용한 문항 분석(석사학위논문), 한국교원대학교 대학원 , (1993)
 - 27) Hambleton, R.K.(Ed.). : *Application of item response theory*.
Vancouver : Educational Research Institute of British

Columbia. (1983)

- 28) Thissen, D. : Marginal Maximum Likelihood Estimation for One-parameter Logistic Model. *Psychometrika*, 47, 2, pp.185-186, (1982)
- 29) 송미영 : 이분반응모형과 등급반응모형에 의한 문항특성과 피험자 능력 추정의 정확성(석사학위논문), 이화여자대학교 대학원, (1994)
- 30) University of Wisconsin : *User's Manual for the ASC Item and Test Analysis Package*, Assessment Systems Coporation, Madison, (1995)
- 31) 박순옥 : 고전검사이론과 문항반응이론의 문항분석 연구, 연세대 대학원(석사학위논문), (1989)
- 32) 김혜기 : 고전검사이론과 문항반응이론에 의한 검사결과의 분석(석사학위논문), 한국교원대학교 대학원, (1994)



<Abstract>

An Item Analysis of the Common Science Summative Evaluation

Oh, Moon-Sik

Physics Education Major

Graduate School of Education, Cheju National University

Cheju, Korea

Supervised by Professor **Kim, Kyu-Yong**

In this thesis, analyzed the characteristic of the high school common science summative evaluation test and items by using the ASC's ITEMAN and XCALIBRE program.

The data are obtained from the response material for the achievement taken in the 7 high schools in Cheju-do, using 18 tests and 400 items.

The conclusion is as follows ;

First, by the classical test theory, the number of easy items has been setting more than the number of hard items, and the items of the high discrimination power amount to 55% of the whole.

※ A thesis submitted to the Committee of the Graduate School of Education, Cheju National University in partial fulfillment of the requirements for the degree of master of Education in August, 2000.

Second, according to the 3-parameter logistic model of the item response theory, because of the 75.2 % of the test setting in the adequate discriminate index, it distinguishes the upper group from the lower group, and also the degree of difficulty in the items is normally distributed from easy to hard items, and the pseudo-guessing parameter shows that over 0.2 and under 0.2 test have been setting at the rate of 41.0, 59.0% each.

Third, the basic unidimensionality assumption in the item response theory on the test is nearly satisfied.

Fourth, between the test grades by the classical test theory and the ability parameter on the interrelation coefficient show very high relations.

Finally, the characteristic of test items between academic and vocational high schools shows some meaningful difference except the peusedo-guessing parameter.

附 錄

부록 1) 검사별 · 문항별 특성(난이도, 변별도, 추측도)

검사	문항 번호	CTT		IRT			검사	문항 번호	CTT		IRT		
		b	a	a	b	c			b	a	a	b	c
jal	1	0.64	0.51	0.67	-0.21	0.15	ja2	1	0.42	0.34	0.66	0.42	0.24
	2	-	-	-	-	-		2	0.58	0.33	0.61	0.86	0.35
	3	0.19	0.32	0.76	2.17	0.15		3	0.57	0.41	0.66	0.55	0.28
	4	0.40	0.50	0.74	1.03	0.16		4	0.56	0.53	0.77	0.30	0.22
	5	0.57	0.57	0.69	0.11	0.15		5	0.23	0.26	0.74	2.35	0.16
	6	0.78	0.35	0.61	-1.05	0.13		6	0.87	0.20	0.61	-1.86	0.24
	7	0.23	0.30	0.71	2.21	0.16		7	0.56	0.53	0.70	0.22	0.20
	8	0.79	0.40	0.69	-1.16	0.12		8	0.56	0.71	0.93	0.10	0.17
	9	0.31	0.45	0.74	1.51	0.15		9	0.65	0.49	0.68	-0.19	0.24
	10	0.71	0.46	0.70	-0.64	0.13		10	0.46	0.38	0.69	1.17	0.26
	11	0.75	0.20	0.44	-0.86	0.16		11	0.34	0.31	0.66	2.02	0.23
	12	0.61	0.45	0.62	-0.08	0.16		12	0.66	0.55	0.77	-0.34	0.20
	13	0.57	0.34	0.53	0.75	0.20		13	0.53	0.54	0.72	0.38	0.20
	14	0.67	0.39	0.58	-0.33	0.16		14	0.61	0.71	0.95	-0.17	0.16
	15	0.64	0.44	0.64	-0.15	0.16		15	0.62	0.54	0.75	-0.10	0.20
	16	0.58	0.40	0.58	0.45	0.19		16	0.74	0.39	0.66	-0.79	0.23
	17	0.57	0.54	0.66	0.21	0.16		17	0.51	0.36	0.56	1.20	0.30
	18	0.59	0.37	0.57	0.41	0.19		18	0.82	0.41	0.74	-1.31	0.21
	19	0.60	0.35	0.52	0.36	0.20		19	0.85	0.31	0.68	-1.54	0.24
	20	0.80	0.34	0.61	-1.28	0.12		20	0.63	0.34	0.59	0.17	0.30
ja3	1	0.91	0.20	0.64	-2.24	0.22	ja4	1	0.81	0.12	0.44	-1.25	0.30
	2	0.68	0.19	0.46	0.27	0.38		2	0.66	0.67	0.87	-0.34	0.18
	3	0.16	0.18	0.70	3.00	0.13		3	0.84	0.41	0.77	-1.38	0.26
	4	0.75	0.54	0.78	-0.86	0.19		4	0.68	0.59	0.77	-0.41	0.21
	5	0.68	0.39	0.54	-0.35	0.25		5	0.63	0.56	0.70	0.01	0.25
	6	0.68	0.52	0.72	-0.44	0.20		6	0.84	0.35	0.69	-1.48	0.24
	7	0.78	0.43	0.72	-1.06	0.21		7	0.86	0.42	0.99	-1.50	0.21
	8	0.49	0.53	0.66	0.56	0.18		8	0.63	0.45	0.68	0.37	0.34
	9	0.46	0.57	0.77	0.60	0.16		9	0.74	0.24	0.57	0.07	0.38
	10	0.54	0.45	0.63	0.44	0.22		10	0.69	0.42	0.60	-0.31	0.30
	11	0.62	0.41	0.57	0.12	0.26		11	0.86	0.37	0.74	-1.53	0.27
	12	0.72	0.59	0.87	-0.70	0.17		12	0.85	0.38	0.77	-1.47	0.24
	13	0.49	0.59	0.78	0.50	0.17		13	0.82	0.36	0.62	-1.32	0.28
	14	0.69	0.49	0.70	-0.46	0.23		14	0.80	0.50	0.87	-1.09	0.21
	15	0.61	0.58	0.79	-0.13	0.17		15	0.61	0.28	0.64	1.06	0.38
	16	0.76	0.36	0.61	-1.00	0.21		16	0.79	0.46	0.80	-1.07	0.22
	17	0.66	0.27	0.50	-0.04	0.29		17	0.82	0.41	0.75	-1.28	0.22
	18	0.72	0.36	0.57	-0.67	0.24		18	0.79	0.44	0.73	-1.01	0.25
	19	0.57	0.60	0.73	0.09	0.18		19	0.64	0.56	0.68	-0.23	0.21
	20	0.87	0.24	0.64	-1.83	0.22		20	0.68	0.52	0.71	-0.27	0.26
	21	0.66	0.39	0.58	-0.18	0.28		-	-	-	-	-	-

* CTT : 고전검사이론, IRT : 문항반응이론(3-모수로지스틱 모형)

* a : 난이도, b : 변별도, c : 추측도

검사	문항 번호	CTT		IRT		
		b	a	a	b	c
ji1	1	0.66	0.40	0.58	-0.31	0.23
	2	0.48	0.56	0.77	0.66	0.20
	3	0.30	0.30	0.74	1.90	0.19
	4	0.58	0.48	0.68	0.24	0.24
	5	0.37	0.47	0.79	1.12	0.17
	6	0.76	0.26	0.53	-0.81	0.28
	7	0.78	0.49	0.84	-1.01	0.18
	8	0.71	0.26	0.51	-0.69	0.24
	9	0.65	0.38	0.56	-0.12	0.25
	10	0.51	0.50	0.63	0.49	0.18
	11	0.84	0.35	0.73	-1.45	0.21
	12	0.35	0.24	0.65	2.46	0.27
	13	0.52	0.36	0.60	0.81	0.26
	14	0.68	0.40	0.61	-0.38	0.23
	15	0.65	0.36	0.62	0.08	0.31
	16	0.45	0.49	0.72	0.75	0.18
	17	0.52	0.69	0.90	0.24	0.16
	18	0.41	0.29	0.63	1.64	0.26
	19	0.79	0.39	0.75	-1.17	0.19
	20	0.88	0.28	0.83	-1.72	0.19

검사	문항 번호	CTT		IRT		
		b	a	a	b	c
ji2	1	0.30	0.34	0.73	1.93	0.19
	2	0.56	0.46	0.67	0.42	0.24
	3	0.48	0.50	0.73	0.69	0.21
	4	0.66	0.46	0.60	-0.17	0.25
	5	0.14	0.10	0.79	3.00	0.12
	6	0.64	0.44	0.61	-0.11	0.23
	7	0.78	0.51	0.80	-1.10	0.17
	8	0.21	0.23	0.70	2.93	0.16
	9	0.38	0.18	0.72	2.54	0.31
	10	0.55	0.51	0.63	0.32	0.20
	11	0.80	0.38	0.71	-1.22	0.20
	12	0.35	0.63	0.87	1.04	0.13
	13	0.55	0.73	0.84	0.06	0.14
	14	0.64	0.57	0.74	-0.26	0.19
	15	0.79	0.45	0.76	-1.11	0.19
	16	0.72	0.47	0.67	-0.72	0.20
	17	0.71	0.47	0.64	-0.63	0.22
	18	0.58	0.55	0.61	0.05	0.17
	19	0.84	0.35	0.67	-1.52	0.21
	20	0.69	0.40	0.58	-0.59	0.20

검사	문항 번호	CTT		IRT		
		b	a	a	b	c
ji3	1	0.83	0.34	0.73	-1.32	0.24
	2	0.50	0.36	0.79	1.26	0.34
	3	0.57	0.56	0.82	0.36	0.26
	4	0.73	0.41	0.69	-0.44	0.31
	5	0.35	0.38	0.75	1.58	0.21
	6	0.80	0.32	0.62	-1.12	0.28
	7	0.92	0.15	0.65	-2.42	0.27
	8	0.45	0.50	0.72	0.83	0.21
	9	0.79	0.40	0.75	-0.99	0.25
	10	0.59	0.58	0.88	0.06	0.21
	11	0.71	0.44	0.70	-0.26	0.32
	12	0.92	0.22	0.87	-2.11	0.24
	13	0.86	0.34	0.87	-1.45	0.24
	14	0.94	0.16	0.90	-2.32	0.24
	15	0.72	0.51	0.78	-0.56	0.24
	16	0.57	0.51	0.74	0.29	0.24
	17	0.58	0.29	0.67	0.93	0.38
	18	0.84	0.28	0.68	-1.44	0.25
	19	0.84	0.32	0.72	-1.45	0.25
	20	0.56	0.34	0.67	0.95	0.35

검사	문항 번호	CTT		IRT		
		b	a	a	b	c
ji4	1	0.94	0.03	0.52	-3.00	0.21
	2	0.96	0.11	0.76	-3.00	0.19
	3	0.27	0.38	0.68	1.92	0.15
	4	0.94	0.08	0.61	-3.00	0.20
	5	0.54	0.15	0.68	2.93	0.15
	6	0.89	0.31	0.78	-1.89	0.20
	7	0.95	0.08	0.63	-3.00	0.20
	8	0.75	0.49	0.68	-0.92	0.19
	9	0.50	0.42	0.60	0.72	0.22
	10	0.67	0.56	0.69	-0.42	0.19
	11	0.73	0.43	0.62	-0.81	0.20
	12	0.58	0.49	0.64	0.30	0.25
	13	0.38	0.33	0.67	1.92	0.26
	14	0.89	0.18	0.63	-2.13	0.20
	15	0.80	0.45	0.78	-1.22	0.18
	16	0.29	0.43	0.79	1.65	0.15
	17	0.72	0.52	0.65	-0.75	0.19
	18	0.71	0.49	0.66	-0.69	0.19
	19	0.65	0.44	0.60	-0.15	0.25
	20	0.79	0.39	0.63	-1.16	0.22

검사	문항번호	CTT		IRT		
		b	a	a	b	c
nj1	1	0.73	0.49	0.79	-0.63	0.23
	2	0.66	0.44	0.66	-0.29	0.22
	3	0.60	0.58	0.78	0.02	0.21
	4	0.55	0.51	0.79	0.73	0.31
	5	0.62	0.58	0.85	-0.06	0.22
	6	0.67	0.61	0.96	-0.46	0.17
	7	0.55	0.63	0.83	0.23	0.19
	8	0.72	0.41	0.76	-0.70	0.20
	9	0.78	0.38	0.71	-1.08	0.20
	10	0.36	0.46	0.78	1.22	0.18
	11	0.30	0.28	0.80	2.07	0.21
	12	0.28	0.41	0.82	1.79	0.17
	13	0.67	0.36	0.62	-0.18	0.26
	14	0.30	0.58	0.95	1.22	0.14
	15	0.23	0.41	0.97	1.67	0.12
	16	0.52	0.37	0.63	0.79	0.26
	17	0.28	0.40	0.85	1.66	0.16
	18	0.58	0.57	0.76	0.14	0.22
	19	0.44	0.50	0.87	0.96	0.23
	20	0.46	0.42	0.76	1.10	0.26
	21	0.70	0.49	0.78	-0.54	0.21
	22	0.87	0.14	0.61	-1.89	0.22
	23	0.72	0.45	0.76	-0.63	0.21
	24	0.60	0.41	0.66	0.04	0.22

검사	문항번호	CTT		IRT		
		b	a	a	b	c
nj2	1	0.88	0.18	0.71	-1.88	0.18
	2	0.46	0.69	0.93	0.43	0.13
	3	0.57	0.65	0.93	-0.00	0.15
	4	0.59	0.75	1.15	-0.13	0.14
	5	0.55	0.73	1.02	0.07	0.14
	6	0.94	0.15	0.93	-2.25	0.17
	7	0.56	0.75	1.09	-0.00	0.14
	8	0.90	0.16	0.74	-2.07	0.18
	9	0.48	0.82	1.26	0.31	0.13
	10	0.54	0.58	0.86	0.22	0.17
	11	0.84	0.28	0.74	-1.46	0.18
	12	0.64	0.67	0.95	-0.26	0.17
	13	0.59	0.66	0.94	-0.10	0.16
	14	0.66	0.47	0.72	-0.36	0.19
	15	0.43	0.62	0.91	0.70	0.16
	16	0.72	0.25	0.61	-0.73	0.19
	17	0.40	0.54	0.88	0.91	0.16
	18	0.48	0.14	0.82	2.09	0.38
	19	0.42	0.69	0.92	0.58	0.12
	20	0.66	0.47	0.77	-0.34	0.19
	21	0.78	0.29	0.67	-1.13	0.18
	22	0.80	0.23	0.60	-1.31	0.20
	23	0.51	0.56	0.85	0.49	0.20
	-	-	-	-	-	-

검사	문항번호	CTT		IRT		
		b	a	a	b	c
nn1	1	0.82	0.36	0.60	-1.42	0.22
	2	0.41	0.46	0.66	1.14	0.19
	3	0.62	0.54	0.63	-0.13	0.20
	4	0.92	0.16	0.60	-2.58	0.21
	5	0.29	0.24	0.63	2.36	0.20
	6	0.78	0.37	0.57	-1.10	0.24
	7	0.15	0.11	0.69	3.00	0.13
	8	0.71	0.53	0.75	-0.67	0.19
	9	0.52	0.51	0.70	0.52	0.21
	10	0.58	0.74	0.86	-0.09	0.14
	11	0.68	0.13	0.41	0.23	0.37
	12	0.39	0.66	0.82	0.71	0.11
	13	0.87	0.26	0.63	-1.84	0.21
	14	0.68	0.45	0.62	-0.48	0.20
	15	0.48	0.35	0.55	1.09	0.25
	16	0.72	0.30	0.50	-0.73	0.25
	17	0.72	0.28	0.49	-0.66	0.27
	18	0.80	0.40	0.67	-1.26	0.20
	19	0.43	0.43	0.65	1.14	0.22
	20	0.21	0.34	0.71	2.17	0.12
	21	0.39	0.51	0.73	0.96	0.14
	22	0.67	0.29	0.49	-0.23	0.26
	23	0.67	0.49	0.60	-0.39	0.21
	24	0.66	0.36	0.54	-0.20	0.25
	25	0.31	0.17	0.64	2.70	0.24
	26	0.60	0.47	0.55	0.02	0.21
	27	0.41	0.34	0.54	1.58	0.23
	28	0.75	0.32	0.52	-0.94	0.24

검사	문항번호	CTT		IRT		
		b	a	a	b	c
nn2	1	0.39	0.68	0.57	-0.24	0.37
	2	0.64	0.61	0.75	-0.25	0.22
	3	0.56	0.49	0.72	0.49	0.22
	4	0.62	0.49	0.78	0.31	0.18
	5	0.39	0.80	0.62	-1.62	0.26
	6	0.30	0.25	0.74	2.07	0.18
	7	0.67	0.61	0.80	-0.21	0.22
	8	0.27	0.87	0.61	-2.60	0.25
	9	0.28	0.91	0.75	-2.90	0.25
	10	0.53	0.72	0.69	-0.95	0.22
	11	0.27	0.59	0.57	0.99	0.37
	12	0.24	0.42	0.65	2.37	0.36
	13	0.57	0.64	0.70	-0.41	0.22
	14	0.42	0.54	0.58	0.54	0.29
	15	0.43	0.71	0.55	-0.90	0.26
	16	0.58	0.70	0.70	-0.79	0.22
	17	0.38	0.71	0.54	-0.69	0.31
	18	0.44	0.66	0.59	-0.26	0.30
	19	0.31	0.28	0.67	2.17	0.21
	20	0.53	0.58	0.64	0.06	0.25
	21	0.39	0.39	0.66	1.47	0.25
	22	0.34	0.82	0.52	-1.86	0.28
	23	0.36	0.85	0.65	-2.15	0.25
	24	0.55	0.41	0.66	0.96	0.19
	25	0.58	0.55	0.66	0.24	0.24
	26	0.53	0.71	0.63	-0.84	0.24
	27	0.41	0.80	0.64	-1.54	0.25
	28	0.59	0.58	0.63	-0.07	0.20

검사	문항번호	CTT		IRT		
		b	a	a	b	c
sa1	1	0.46	0.47	0.85	1.03	0.26
	2	0.71	0.43	0.71	-0.28	0.31
	3	0.68	0.43	0.69	-0.34	0.25
	4	0.49	0.36	0.73	1.16	0.31
	5	0.64	0.58	0.84	-0.18	0.23
	6	0.60	0.59	0.79	0.11	0.25
	7	0.85	0.32	0.79	-1.46	0.26
	8	0.42	0.52	0.77	1.07	0.22
	9	0.42	0.52	0.87	1.11	0.23
	10	0.29	0.43	0.90	1.58	0.18
	11	0.38	0.54	0.87	1.13	0.19
	12	0.59	0.53	0.79	0.23	0.27
	13	0.18	0.23	0.79	3.00	0.18
	14	0.29	0.24	0.77	2.78	0.25
	15	0.61	0.53	0.81	0.06	0.25
	16	0.35	0.36	0.80	1.69	0.24
	17	0.60	0.14	0.78	2.65	0.58
	18	0.75	0.33	0.67	-0.48	0.41
	19	0.55	0.36	0.73	1.94	0.49
-	-	-	-	-	-	
-	-	-	-	-	-	

검사	문항번호	CTT		IRT		
		b	a	a	b	c
sa2	1	0.84	0.45	0.93	-1.35	0.21
	2	0.76	0.50	0.87	-0.86	0.21
	3	0.54	0.43	0.74	0.63	0.27
	4	0.37	0.40	0.79	1.65	0.25
	5	0.38	0.50	0.90	1.19	0.21
	6	0.78	0.52	0.91	-0.89	0.22
	7	0.65	0.76	1.07	-0.30	0.18
	8	0.70	0.52	0.89	-0.44	0.23
	9	0.66	0.65	0.87	-0.29	0.20
	10	0.57	0.75	0.94	0.07	0.18
	11	0.24	0.27	0.83	2.34	0.18
	12	0.60	0.64	0.85	0.08	0.23
	13	0.71	0.54	0.78	-0.50	0.23
	14	0.32	0.38	0.81	1.62	0.20
	15	0.38	0.24	0.82	1.89	0.28
	16	0.84	0.30	0.74	-1.37	0.25
	17	0.66	0.33	0.69	0.38	0.40
	18	0.82	0.28	0.71	-1.15	0.30
	19	0.57	0.51	0.71	0.27	0.22
	20	0.31	0.17	0.82	2.52	0.26
	21	0.46	0.48	0.86	0.70	0.24

검사	문항번호	CTT		IRT		
		b	a	a	b	c
sa3	1	0.72	0.56	0.97	-0.67	0.16
	2	0.32	0.32	0.77	1.58	0.17
	3	0.71	0.47	0.79	-0.74	0.18
	4	0.49	0.48	0.78	0.49	0.17
	5	0.12	-0.04	0.77	0.72	0.17
	6	0.54	0.40	0.67	0.42	0.21
	7	0.13	0.08	0.87	0.85	0.14
	8	0.54	0.65	0.87	0.12	0.14
	9	0.43	0.48	0.89	0.68	0.15
	10	0.34	0.33	0.79	1.59	0.15
	11	0.30	0.30	0.80	1.89	0.19
	12	0.08	-0.03	0.80	0.47	0.19
	13	0.34	0.32	0.79	1.88	0.22
	14	0.16	0.17	0.82	2.95	0.14
	15	0.44	0.54	0.81	0.72	0.16
	16	0.25	0.35	0.85	1.84	0.14
	17	0.26	0.30	0.79	2.07	0.17
	18	0.15	-0.01	0.79	2.51	0.19
	19	0.19	0.13	0.82	3.00	0.17
	20	0.24	0.36	0.86	1.70	0.12
	21	0.54	0.36	0.70	0.54	0.24
	22	0.42	0.38	0.74	1.35	0.24
	23	0.25	0.19	0.79	3.00	0.22
	24	0.48	0.46	0.77	0.70	0.22

검사	문항번호	CTT		IRT		
		b	a	a	b	c
sa4	1	0.52	0.47	0.71	0.56	0.22
	2	0.49	0.55	0.73	0.52	0.18
	3	0.26	0.19	0.77	2.42	0.19
	4	0.53	0.40	0.69	0.55	0.23
	5	0.47	0.63	0.79	0.52	0.16
	6	0.62	0.37	0.65	0.02	0.22
	7	0.76	0.55	0.79	-0.93	0.18
	8	0.57	0.66	0.75	0.02	0.17
	9	0.56	0.37	0.62	0.68	0.28
	10	0.41	0.60	0.79	0.80	0.15
	11	0.47	0.44	0.76	0.72	0.19
	12	0.57	0.25	0.65	0.72	0.32
	13	0.26	0.20	0.74	2.42	0.19
	14	0.48	0.68	0.85	0.42	0.15
	15	0.55	0.64	0.87	0.20	0.17
	16	0.60	0.48	0.71	0.18	0.24
	17	0.43	0.54	0.80	0.81	0.17
	18	0.70	0.52	0.77	-0.48	0.21
	19	0.76	0.46	0.74	-0.87	0.20
	20	0.39	0.27	0.72	1.63	0.24
	21	0.50	0.55	0.86	0.43	0.18
	22	0.49	0.50	0.81	0.56	0.20
	23	0.26	0.40	0.77	1.61	0.13
	24	0.38	0.54	0.80	0.92	0.15

검사	문항 번호	CTT		IRT		
		b	a	a	b	c
jml	1	0.82	0.44	1.08	-1.19	0.13
	2	0.74	0.77	1.60	-0.69	0.09
	3	0.82	0.54	1.31	-1.00	0.17
	4	0.78	0.62	1.25	-0.86	0.13
	5	0.80	0.56	1.34	-1.00	0.12
	6	0.74	0.68	1.37	-0.59	0.17
	7	0.81	0.55	1.33	-1.07	0.13
	8	0.71	0.77	1.45	-0.57	0.11
	9	0.67	0.69	1.24	-0.29	0.16
	10	0.85	0.42	1.16	-1.35	0.13
	11	0.73	0.66	1.20	-0.62	0.13
	12	0.75	0.63	1.21	-0.75	0.12
	13	0.84	0.40	1.00	-1.30	0.15
	14	0.88	0.39	1.22	-1.48	0.14
	15	0.73	0.77	1.58	-0.61	0.12
	16	0.73	0.60	1.15	-0.46	0.21
	17	0.63	0.79	1.37	-0.18	0.12
	18	0.55	0.66	0.97	0.17	0.15
	19	0.51	0.70	1.06	0.34	0.15
	20	0.49	0.73	1.15	0.36	0.13
	21	0.63	0.81	1.44	-0.21	0.11
	22	0.49	0.40	1.01	0.84	0.26
	23	0.74	0.67	1.24	-0.71	0.11
	24	0.69	0.70	1.26	-0.48	0.12
-	-	-	-	-	-	

검사	문항 번호	CTT		IRT		
		b	a	a	b	c
sbl	1	0.54	0.42	0.64	0.43	0.22
	2	0.70	0.51	0.78	-0.56	0.18
	3	0.49	0.53	0.84	0.65	0.21
	4	0.74	0.49	0.72	-0.79	0.20
	5	0.74	0.48	0.77	-0.73	0.21
	6	0.32	0.47	0.82	1.49	0.16
	7	0.88	0.34	0.93	-1.65	0.19
	8	0.54	0.53	0.70	0.49	0.24
	9	0.89	0.28	0.77	-1.85	0.21
	10	0.32	0.43	0.75	1.56	0.18
	11	0.88	0.26	0.74	-1.78	0.20
	12	0.82	0.30	0.66	-1.37	0.21
	13	0.87	0.28	0.72	-1.73	0.20
	14	0.21	0.26	0.79	2.58	0.16
	15	0.64	0.51	0.74	-0.09	0.23
	16	0.34	0.24	0.75	2.06	0.25
	17	0.39	0.26	0.74	1.95	0.28
	18	0.30	0.38	0.82	1.46	0.15
	19	0.61	0.60	0.87	-0.08	0.18
	20	0.73	0.56	0.96	-0.65	0.19
	21	0.55	0.62	0.78	0.27	0.20
	22	0.25	0.35	0.84	1.78	0.14
	23	0.31	0.37	0.86	1.59	0.18
	24	0.27	0.10	0.61	1.14	0.28
	25	0.70	0.55	0.91	-0.57	0.18

부록 2) 검사의 단일차원성의 검증 결과

(1) 검사 j11 의 요인분석 결과

공통성			설명된 총분산						
문항	초기	추출	성분	초기 고유값			추출 제곱합 적재값		
				합계	% 분산	% 누적	합계	% 분산	% 누적
1	1.000	.465	1	3.013	15.064	15.064	3.013	15.064	15.064
2	1.000	.467	2	1.387	6.936	21.999	1.387	6.936	21.999
3	1.000	.615	3	1.209	6.044	28.043	1.209	6.044	28.043
4	1.000	.692	4	1.163	5.815	33.858	1.163	5.815	33.858
5	1.000	.526	5	1.132	5.658	39.517	1.132	5.658	39.517
6	1.000	.673	6	1.078	5.391	44.908	1.078	5.391	44.908
7	1.000	.472	7	1.037	5.186	50.094	1.037	5.186	50.094
8	1.000	.548	8	1.007	5.036	55.130	1.007	5.036	55.130
9	1.000	.687	9	.941	4.703	59.833			
10	1.000	.356	10	.920	4.601	64.434			
11	1.000	.550	11	.885	4.423	68.857			
12	1.000	.523	12	.845	4.226	73.084			
13	1.000	.388	13	.801	4.003	77.087			
14	1.000	.573	14	.776	3.881	80.967			
15	1.000	.677	15	.733	3.665	84.632			
16	1.000	.436	16	.714	3.569	88.201			
17	1.000	.480	17	.625	3.126	91.327			
18	1.000	.708	18	.614	3.068	94.395			
19	1.000	.590	19	.589	2.946	97.341			
20	1.000	.601	20	.532	2.659	100.000			

추출 방법: 주성분 분석

(2) 검사 j12 의 요인분석 결과

공통성			설명된 총분산						
문항	초기	추출	성분	초기 고유값			추출 제곱합 적재값		
				합계	% 분산	% 누적	합계	% 분산	% 누적
1	1.000	.665	1	3.479	17.396	17.396	3.479	17.396	17.396
2	1.000	.433	2	1.337	6.683	24.079	1.337	6.683	24.079
3	1.000	.386	3	1.205	6.027	30.107	1.205	6.027	30.107
4	1.000	.390	4	1.127	5.633	35.740	1.127	5.633	35.740
5	1.000	.489	5	1.103	5.517	41.257	1.103	5.517	41.257
6	1.000	.507	6	1.089	5.443	46.699	1.089	5.443	46.699
7	1.000	.595	7	1.033	5.164	51.863	1.033	5.164	51.863
8	1.000	.607	8	.999	4.995	56.858			
9	1.000	.641	9	.941	4.706	61.564			
10	1.000	.561	10	.894	4.469	66.033			
11	1.000	.503	11	.873	4.364	70.397			
12	1.000	.586	12	.820	4.099	74.497			
13	1.000	.539	13	.801	4.005	78.502			
14	1.000	.593	14	.729	3.644	82.146			
15	1.000	.488	15	.683	3.417	85.563			
16	1.000	.476	16	.650	3.249	88.812			
17	1.000	.539	17	.633	3.163	91.976			
18	1.000	.596	18	.591	2.954	94.929			
19	1.000	.388	19	.529	2.643	97.572			
20	1.000	.392	20	.486	2.428	100.000			

추출 방법: 주성분 분석

(3) 검사 ji3 의 요인분석 결과

공통성			설명된 총분산						
문항	초기	추출	성분	초기 고유값			추출 제곱합 적재값		
				합계	% 분산	% 누적	합계	% 분산	% 누적
1	1.000	.468	1	3.610	18.050	18.050	3.610	18.050	18.050
2	1.000	.327	2	1.491	7.456	25.506	1.491	7.456	25.506
3	1.000	.451	3	1.268	6.340	31.846	1.268	6.340	31.846
4	1.000	.560	4	1.168	5.839	37.685	1.168	5.839	37.685
5	1.000	.606	5	1.117	5.583	43.268	1.117	5.583	43.268
6	1.000	.628	6	1.066	5.328	48.596	1.066	5.328	48.596
7	1.000	.479	7	.995	4.973	53.569			
8	1.000	.378	8	.957	4.785	58.354			
9	1.000	.457	9	.908	4.540	62.894			
10	1.000	.440	10	.869	4.346	67.240			
11	1.000	.415	11	.838	4.190	71.430			
12	1.000	.513	12	.815	4.075	75.504			
13	1.000	.488	13	.730	3.649	79.154			
14	1.000	.492	14	.683	3.416	82.570			
15	1.000	.472	15	.668	3.339	85.909			
16	1.000	.519	16	.656	3.282	89.191			
17	1.000	.525	17	.617	3.086	92.277			
18	1.000	.595	18	.550	2.748	95.025			
19	1.000	.576	19	.527	2.634	97.658			
20	1.000	.330	20	.468	2.342	100.000			

추출 방법: 주성분 분석

(4) 검사 ji4 의 요인분석결과

공통성			설명된 총분산						
문항	초기	추출	성분	초기 고유값			추출 제곱합 적재값		
				합계	% 분산	% 누적	합계	% 분산	% 누적
1	1.000	.377	1	3.086	15.432	15.432	3.086	15.432	15.432
2	1.000	.558	2	1.576	7.879	23.311	1.576	7.879	23.311
3	1.000	.559	3	1.188	5.941	29.252	1.188	5.941	29.252
4	1.000	.563	4	1.157	5.787	35.040	1.157	5.787	35.040
5	1.000	.722	5	1.132	5.660	40.700	1.132	5.660	40.700
6	1.000	.554	6	1.097	5.485	46.184	1.097	5.485	46.184
7	1.000	.438	7	1.053	5.266	51.451	1.053	5.266	51.451
8	1.000	.580	8	.993	4.964	56.414			
9	1.000	.482	9	.951	4.753	61.168			
10	1.000	.437	10	.928	4.642	65.809			
11	1.000	.385	11	.870	4.352	70.162			
12	1.000	.616	12	.829	4.146	74.307			
13	1.000	.658	13	.801	4.005	78.312			
14	1.000	.421	14	.725	3.625	81.937			
15	1.000	.429	15	.704	3.520	85.456			
16	1.000	.574	16	.646	3.231	88.688			
17	1.000	.553	17	.608	3.040	91.728			
18	1.000	.496	18	.598	2.989	94.717			
19	1.000	.425	19	.545	2.724	97.441			
20	1.000	.464	20	.512	2.559	100.000			

추출 방법: 주성분 분석.

부록 3) 문제 내용과 문항 특성의 예시

Item(sal)	CTT		IRT																													
	b	a	a	b	c																											
<p>1. 다음의 표는 서울역을 출발하여 부산역에 도착할 때까지의 새마을호 열차의 시각표이다. 이 표에서 서울-천안, 천안-대전, 대전-동대구, 동대구-부산 사이의 평균 시속을 구했을 때 어느 구간에서 가장 빠른가?</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th>서울역에서의 주행거리(km)</th> <th>역이름</th> <th>시각(시:분)</th> </tr> </thead> <tbody> <tr> <td>0.0</td> <td>서울</td> <td>발 7:00</td> </tr> <tr> <td>97.1</td> <td>천안</td> <td>착 7:55</td> </tr> <tr> <td>166.8</td> <td>발</td> <td>7:56</td> </tr> <tr> <td></td> <td>대전</td> <td>착 8:56</td> </tr> <tr> <td>327.1</td> <td>발</td> <td>8:38</td> </tr> <tr> <td></td> <td>동대구</td> <td>착 10:08</td> </tr> <tr> <td>444.5</td> <td>발</td> <td>10:09</td> </tr> <tr> <td></td> <td>부산</td> <td>착 11:20</td> </tr> </tbody> </table> <p>① 서울-천안 ② 천안-대전 ③ 대전-동대구 ④ 동대구-부산 ⑤ 모든구간에서 시속이 같다</p>	서울역에서의 주행거리(km)	역이름	시각(시:분)	0.0	서울	발 7:00	97.1	천안	착 7:55	166.8	발	7:56		대전	착 8:56	327.1	발	8:38		동대구	착 10:08	444.5	발	10:09		부산	착 11:20	0.46	0.47	0.85	1.03	0.26
서울역에서의 주행거리(km)	역이름	시각(시:분)																														
0.0	서울	발 7:00																														
97.1	천안	착 7:55																														
166.8	발	7:56																														
	대전	착 8:56																														
327.1	발	8:38																														
	동대구	착 10:08																														
444.5	발	10:09																														
	부산	착 11:20																														
<p>2. 다음 중 속력의 단위는 ?</p> <p>① m/s ② m/s² ③ s/m ④ kg ⑤ m</p>	0.71	0.43	0.71	-	0.28																											
<p>3. 다음은 운동하는 모형자동차의 운동을 시간기록계를 이용하여 종이테이프에 기록한 것이다. 평균속력이 가장 빠른 구간은 ?</p> <p style="text-align: center;">.....</p> <p style="text-align: center;">A<-2cm->B<-----6cm----->C<--4cm----->D</p> <p>① A-B ② B-C ③ C-D ④ A-C ⑤ B-D</p>	0.68	0.43	0.69	-	0.25																											
<p>4. 위 문제에서 사용한 시간기록계의 타점주기가 1/50초이면 B-C 구간에서의 평균속력은 몇 cm/s 인가 ?</p> <p>① 20 ② 40 ③ 60 ④ 80 ⑤ 100</p>	0.49	0.36	0.73	1.16	0.31																											
<p>5. 고속버스가 서울-광주 사이의 거리 320 km를 4시간에 달린 경우의 속력은 얼마인가 ?</p> <p>① 320 km/h ② 80 km/h ③ 1280 km/h ④ 324 km/h ⑤ 316 km/h</p>	0.64	0.58	0.84	-	0.23																											
<p>6. 평균속력과 순간속력에 대한 설명이다 맞는 것은 ?</p> <p>① 평균속력은 중간간의 운동상태를 고려한 속력이다. ② 자동차가 달릴 때 운전석 앞의 속력계의 눈금은 평균속력을 나타내는 값이다. ③ 순간속력은 매우 짧은 시간 동안에 이동한 거리를 말한다. ④ 일반적으로 속력이라고 할 때는 평균속력을 의미하는 경우가 많다.</p>	0.60	0.59	0.79	0.11	0.24																											
<p>6. 평균속력과 순간속력에 대한 설명이다 맞는 것은 ?</p> <p>① 평균속력은 중간간의 운동상태를 고려한 속력이다. ② 자동차가 달릴 때 운전석 앞의 속력계의 눈금은 평균속력을 나타내는 값이다. ③ 순간속력은 매우 짧은 시간 동안에 이동한 거리를 말한다. ④ 일반적으로 속력이라고 할 때는 평균속력을 의미하는 경우가 많다.</p>	0.60	0.59	0.79	0.11	0.24																											
<p>* CTT : 고전검사이론, IRT : 문항반응이론(3-모수 로지스틱 모형) * a : 난이도, b : 변별도, c : 추측도</p>																																

<p>10. 직선 위를 운동하는 물체의 시간 t_1에서의 속도를 v_1, 시간 t_2에서의 속도를 v_2라 할 때 t_1과 t_2사이의 가속도를 바르게 나타낸 식은 ?</p> <p>① $\frac{v_2 - v_1}{t_2 - t_1}$ ② $\frac{v_1 - v_2}{t_1 - t_2}$ ③ $\frac{v_2 - v_1}{t_1 + t_2}$</p> <p>④ $\frac{v_1 + v_2}{t_1 + t_2}$ ⑤ $\frac{v_2 + v_1}{t_2 - t_1}$</p>	0.29	0.43	0.90	1.58	0.18
<p>11. 정지하고 있던 자동차가 출발하여 10초후에 20 m/s로 되었다. 이 자동차의 가속도는 얼마인가 ?</p> <p>① 30% ② 10% ③ 200% ④ 0.5% ⑤ 2%</p>	0.38	0.54	0.87	1.13	0.19
<p>12. 정지한 물체가 10%의 등가속도 직선 운동을 1분 동안 한다면 물체의 속력은 몇 m/s로 되겠는가 ?</p> <p>① 10 ② 0.1 ③ 600 ④ 60 ⑤ $\frac{1}{6}$</p>	0.59	0.53	0.79	0.23	0.27
<p>14. 높이 19.6m인 곳에서 돌을 자유 낙하시켰다. 지면에 도달하는 시간과 지면에 도달하는 순간의 속력은 얼마인가 ?</p> <p>-----</p> <p>지면에 도달하는 시간 지면에 도달하는 순간의 속력</p> <p>-----</p> <p>① 1 초 9.8 m/s</p> <p>② 1 초 19.6 m/s</p> <p>③ 2 초 9.8 m/s</p> <p>④ 2 초 19.6 m/s</p> <p>⑤ 3 초 19.6 m/s</p> <p>-----</p>	0.29	0.24	0.77	2.78	0.25
<p>15. 다음의 물리적 현상들 중에서 뉴턴의 운동 제1법칙을 나타내는 예가 아닌 것은 ?</p> <p>① 정지한 버스가 급히 출발하면 타고 있던 사람은 뒤로 넘어지려 한다.</p> <p>② 달리던 버스가 급히 정지하면 타고 있던 사람은 앞으로 넘어지려 한다.</p> <p>③ 컵 위에 카드를 올려놓고 그 위에 동전을 올려놓은 다음 카드를 갑자기 빼면 동전은 컵 속으로 떨어진다.</p> <p>④ 나무도막을 여러 개 수직으로 쌓아 놓고 중간의 나무도막을 갑자기 빼내면 위의 나무도막은 아래로 내려간다.</p> <p>⑤ 물체에 힘을 가할 때 큰 힘을 가할수록 물체는 빨리 움직인다.</p>	0.61	0.53	0.81	0.06	0.25
<p>16. 수평면에 정지하고 있는 자동차를 뒤에서 힘을 가하여 밀면 자동차는 앞으로 나아간다. 그러나 밀던 손을 떼면 운동하던 자동차는 잠시 후 멈추게 된다. 이러한 이유를 가장 바르게 설명한 것은 ?</p> <p>① 지면과 자동차바퀴사이에 운동방향과 같은 방향으로 마찰력이 작용하기 때문이다.</p> <p>② 손과 자동차바퀴사이에 운동방향과 같은 방향으로 마찰력이 작용하기 때문이다.</p> <p>③ 지면과 자동차바퀴사이에 운동방향과 반대 방향으로 마찰력이 작용하기 때문이다.</p> <p>④ 손과 자동차바퀴사이에 운동방향과 반대 방향으로 마찰력이 작용하기 때문이다.</p> <p>⑤ 손으로 미는 힘이 작용하지 않았기 때문이다.</p>	0.35	0.36	0.8	1.70	0.24

Item(sa3)	CTT		IRT		
	b	a	a	b	c
7. 높이 19.6m인 곳에서 돌을 자유낙하 시켰다. 지상 9.8m인 지점에 도달하는 시간은 얼마인가?(단, $g=9.8\text{m/s}^2$ 이다.) ① 1.4초 ② 2초 ③ 9.8초 ④ 19.6초 ⑤ 20초	0.13	0.08	0.87	0.85	0.14
8. 매끄러운 수평면 위에 정지하고 있던 질량 1kg인 물체를 일정한 힘으로 끌었더니 10초후에 속도가 10%였다. 작용한 힘의 크기는 얼마인가? ① 1N ② 2N ③ 3N ④ 4N ⑤ 10N	0.54	0.65	0.87	0.12	0.14
10. 다음 중에서 장소에 관계없이 일정한 값을 갖는 것은? ① 질량 ② 무게 ③ 중력가속도 ④ 중력 ⑤ 만유인력	0.34	0.33	0.79	1.59	0.15
11. 무게와 질량에 대한 다음의 설명 중 잘못 된 것은? ① 무게는 장소에 따라 일정한 값을 갖는다. ② 무게는 적도에서 작고, 극지방으로 갈수록 커진다. ③ 질량은 물체의 고유한 양으로서 장소에 관계없이 일정한 값을 갖는다. ④ 질량은 양팔저울로 측정한다. ⑤ 무게는 용수철 저울로 측정한다.	0.30	0.30	0.80	1.89	0.19
12. 다음 중 관성과 관계가 없는 것은? ① 정지하고 있던 버스가 갑자기 출발하면 서 있던 사람은 뒤로 넘어진다. ② 옷의 먼지를 털거나 담배재를 손가락으로 두드려 털다. ③ 만치를 잡고 손잡이 부분을 바닥에 내려치면 망치의 쇠부분이 나무 손잡이에 더욱 깊게 박힌다. ④ 뛰어가던 사람이 발이 돌부리에 걸리면 넘어진다. ⑤ 노를 저으면 배가 앞으로 나간다.	0.08	0.03	0.80	0.47	0.19
13. 다음은 대전열을 나열한 것이다. (+)털가죽-상아-유리-명주-나무-고무-셀룰로이드-에보나이트(-) 보기와 같이 짝지은 물질들끼리 마찰을 시켰다. 앞에 있는 물질이 양(+) 전기를 띠는 것들로 짝지어진 것은? -----<보기>----- a.털가죽-상아 b.고무-나무 c.명주-털가죽 d.유리-상아 e.명주-셀룰로이드 f.유리-털가죽 g.셀룰로이드-상아 -----	0.34	0.32	0.79	1.88	0.22
① a, b, c, d, e, f ② c, d, f, g ③ a, e ④ a, b ⑤ d, g					
15. 마찰력에 대한 설명으로 맞는 것은? ① 마찰력의 방향은 운동의 방향과 항상 반대방향이다. ② 마찰력은 방향은 물체의 운동 방향과 항상 같은 방향이다. ③ 마찰력이 없으면 육성선수의 기록은 더욱 향상된다. ④ 운동 마찰력은 정지 마찰력보다 항상 크다. ⑤ 마찰력은 수직 항력에 반비례한다.	0.44	0.54	0.81	0.72	0.16
16. 화석준금 1도 간격은 1기압일 때 물이 어는 점과 끓는점 사이의 1/()의 간격과 같다. ()속에 알맞은 것은? ① 50 ② 100 ③ 150 ④ 180 ⑤ 200	0.25	0.34	0.85	1.84	0.14
17. 질량 m, 가한 열량 Q, 물체의 온도를 t라 할 때, 비열을 구하는 식으로 맞는 것은? ① $\frac{m \cdot c}{Q}$ ② $\frac{Q}{m \cdot c}$ ③ $\frac{Q}{m \cdot t}$ ④ $\frac{m \cdot t}{Q}$ ⑤ $Q \cdot m \cdot t$	0.42	0.38	0.74	1.35	0.24

19. 비열과 열용량에 대한 설명으로 잘못 된 것은 ? ① 어떤 물질의 온도를 1K만큼 올리는데 필요한 열량을 그 물질의 열용량이라 한다. ② 어떤 물질의 온도를 1K 올리는데 필요한 열량을 그 물질의 비열이라 한다. ③ 어떤 물질 1kg의 온도를 1K만큼 올리는데 필요한 열량을 비열이라 한다. ④ 어떤 물질의 온도를 1℃만큼 올리는데 필요한 열량을 그 물질의 열용량이라 한다. ⑤ 어떤 물질 1kg의 온도를 1℃ 올리는데 필요한 열량을 비열이라 한다.	0.19	0.13	0.82	3.00	0.17
21. 에어컨을 집안에 설치할 때 위쪽에 설치하는 이유로 가장 적당한 것은 ? ① 에어컨의 무게가 크기 때문 ② 벽에 가까이 해야 하기 때문 ③ 공기의 대류현상 때문에 ④ 열복사 현상을 최대한 이용하려는 것 ⑤ 공기에 의한 열전도 현상을 적게 하기 위함이다.	0.54	0.36	0.70	0.54	0.24
24. 보온병을 보면 안쪽 벽은 거울처럼 되어 있고 용기 둘레는 진공으로 되어 있다. 그 이유는 무엇일까 ? ① 전도열만을 막기 위해서 ② 대류에 의한 열만을 막기 위해서 ③ 복사열만을 막기 위해서 ④ 상품가치를 높이기 위해서 ⑤ 전도, 대류 및 복사에 의한 열의 출입을 막기 위해서	0.48	0.46	0.77	0.70	0.22

Item(sa4)	CTT		IRT		
	b	a	a	b	c
1. 다음 중 화력 발전소의 장점은 ? ① 소음이 없다. ② 폐수로 인한 공해가 없다. ③ 대도시 근처에 세울 수 있다. ④ 수자원관리가 가능하다. ⑤ 연료공급을 무한정하게 받을 수 있다.	0.52	0.47	0.71	0.56	0.22
2. 다음 중 원자력 발전의 장점과 관계가 없는 것은 ? ① 핵연료는 화석연료보다 오래 사용할 수 있다. ② 발전 단위 용량이 크다. ③ 화석연료에 비하여 환경오염이 비교적 크지 않다. ④ 방사능오염이나 누출에 대비하여야 한다.	0.49	0.55	0.73	0.52	0.18
3. 다음 중 송전 전압과 손실전력과 관계를 바르게 설명한 것은 ? ① 손실전력은 송전전압에 비례한다. ② 손실전력은 송전전압에 반비례한다. ③ 손실전력은 송전전압의 제곱에 비례한다. ④ 손실전력은 송전전압의 제곱에 반비례한다. ⑤ 손실전력의 제곱은 송전전압에 비례한다.	0.26	0.19	0.77	2.42	0.19
4. 송전시에 전선을 통해 손실되는 전력에 영향을 미치는 요인이 아닌 것은 ? ① 전선의 길이 ② 전선의 무게 ③ 전선의 굵기 ④ 전선의 전압 ⑤ 전선의 종류	0.53	0.40	0.69	0.55	0.23
5. 전압을 높여서 송전하는 이유는 무엇일까 ? ① 전력손실을 줄이기 위하여 ② 전력손실을 높이기 위하여 ③ 위험을 방지하기 위하여 ④ 발전소의 연료소모를 줄이기 위하여 ⑤ 화력발전의 열효율을 높이기 위하여	0.47	0.63	0.79	0.52	0.16

6. 발전소에서 가정까지 전력을 수송하는 과정에서 발전소에서 변전소까지는 송전전압을 높여주지만, 변전소에서 가정으로 송전할 때에는 송전전압을 낮추어 송전하는 까닭은 ? ① 전력손실을 줄이기 위하여 ② 전력손실을 높이기 위하여 ③ 위험을 방지하기 위하여 ④ 발전소의 연료소모를 줄이기 위하여 ⑤ 화력발전의 열효율을 높이기 위하여	0.62	0.37	0.65	0.02	0.22
7. 감전과 누전을 예방할 수 있는 방법과 관계없는 것은 ? ① 도선의 피복이 벗겨지지 않도록 한다. ② 물기가 많거나 금속과 자주접촉하는 곳에서는 콘센트의 두껍을 덮도록 한다. ③ 고압선에 도체가 가까이 가지 않도록 한다. ④ 도선을 단단히 연결한 다음 절연테이프를 잘 감아둔다. ⑤ 전압을 높여준다.	0.76	0.55	0.79	0.93	0.18
12. 다음 중 굳기(경도)가 가장 큰 광물은 어느 것인가 ? ① 석영 ② 활석 ③ 금강석 ④ 장석 ⑤ 강옥	0.57	0.25	0.65	0.72	0.32
8. 광물의 굳기를 조사하기 위하여 <보기>와 같은 방법으로 실험을 하였다. 실험 결과 굳기가 가장 큰 것은 ? -----<보기>----- a. 방해석으로 광물 A를 긁어 보았더니 긁히지 않았다. b. 광물 A와 B를 서로 문질렀더니 광물 A의 가루가 생겼다. c. 유리로 긁었더니 광물 A는 긁혔으나 광물 B는 아무런지도 않았다. d. 광물 B를 유리로 긁었더니 유리가 긁혔다. -----	0.57	0.66	0.75	0.02	0.17
① 광물A ② 광물B ③ 방해석 ④ 유리 ⑤ 굳기가 서로 같다.					
9. 암석을 이루는 기본 단위물질을 누엇이라고 하는가 ? ① 광물 ② 원소 ③ 산소 ④ 흙 ⑤ 원자	0.56	0.37	0.62	0.68	0.28
10. 황동석, 금, 활철석의 색은 모두 황색이다. 이들 광물을 구별하려면 어떤 방법을 이용하는 것이 좋은가 ? ① 조흔판을 이용하여 광물의 조흔색을 관찰한다. ② 돋보기를 이용하여 색을 관찰한다. ③ 쪼개짐과 깨짐의 차이를 이용하여 구별한다. ④ 모스경도계를 이용하여 굳기를 알아본다. ⑤ 전기전도도를 이용하여 구별한다.	0.41	0.60	0.79	0.80	0.15
11. 광물의 조흔색이란 ? ① 표면색을 말한다. ② 액체상태의 색을 말한다. ③ 덩어리 상태의 색을 말한다. ④ 기체상태의 색을 말한다. ⑤ 가루상태의 색을 말한다.	0.47	0.44	0.76	0.72	0.19
13. 공기 중에서 500g중인 물체가 물 속에서 측정한 결과 450g중이었다. 이 물체의 물 속에서 가벼워진 무게는 50g중이다. 이 50g중에 대한 다음 설명 중 맞는 것은? ① 이 물체의 무게이다. ② 이 물체의 부피이다. ③ 이 물체의 비중이다. ④ 이 물체의 밀도이다. ⑤ 이 물체의 부피에 해당하는 물의 무게이다.	0.26	0.20	0.74	2.42	0.19
14. 지각의 여러 개의 판으로 구성하여 지구 표면을 덮고있다는 학설은 다음 중 어느 것인가? ① 맨틀 대류설 ② 판 구조론 ③ 판입의 법칙 ④ 지층 누층의 법칙 ⑤ 부정합의 법칙	0.48	0.68	0.85	0.42	0.15

15. 다음 중 부정합의 생성 순서로 맞는 것은? ① 퇴적→용기→침식→침강→퇴적 ② 용기→퇴적→침식→용기→침강 ③ 침식→용기→퇴적→침강→퇴적 ④ 침강→용기→퇴적→침식→퇴적 ⑤ 퇴적→용기→침강→침식→퇴적	0.55	0.64	0.87	0.20	0.17
16. 지각 변동이 없었다고 했을 때 아래 지층이 위 지층보다 오래된 지층임을 말해주는 법칙은? ① 동일 과정의 법칙 ② 지층 누층의 법칙 ③ 부정합의 법칙 ④ 관입의 법칙 ⑤ 동물군 천이의 법칙	0.60	0.48	0.71	0.18	0.24
17. 다음 지사 연구 법칙 중 화성 활동과 가장 관계가 깊은 것은? ① 누층의 법칙 ② 관입의 법칙 ③ 부정합의 법칙 ④ 동일 과정의 법칙 ⑤ 동물군 천이의 법칙	0.43	0.54	0.80	0.81	0.17
18. 가장 오래된 지층에서 발견되는 화석은? ① 공룡 ② 시조새 ③ 삼엽충 ④ 매머드 ⑤ 포유류	0.70	0.52	0.77	- 0.48	0.21
19. 중생대 초에 나타나 중생대에 번성하다가 중생대 말에 멸종한 거대한 동물은? ① 잠자리 ② 오징어류 ③ 양서류 ④ 공룡 ⑤ 매머드	0.76	0.46	0.74	- 0.87	0.20
20. 다음 중 표준 화석으로 적당한 것은? ① A ② B ③ C ④ D ⑤ E	0.39	0.27	0.72	1.63	0.24
21. 다음 중 고생대의 표준 화석으로 맞는 것은? ① 조류(새) ② 양서류 ③ 파충류 ④ 어류 ⑤ 삼엽충	0.50	0.55	0.86	0.43	0.18
22. 생물 진화에 대한 설명 중 잘못된 것은? ① 대체로 작은 생물에서 큰 생물로 진화한다. ② 대체로 모든 생물의 종들은 영원히 계속 진화를 한다. ③ 대체로 원시 생물로 진화한다. ④ 대체로 단순한 생물에서 복잡한 생물로 진화한다. ⑤ 대체로 생물은 환경에 적응하기 위하여 진화한다.	0.49	0.50	0.81	0.56	0.20
23. 전 해수의 평균 염분은 다음 중 어느 것인가? ① 0.08% ② 35% ③ 1.66% ④ 1.66‰ ⑤ 35‰	0.26	0.40	0.77	1.61	0.13
24. 해수의 염분에 대한 다음 설명 중 맞는 것은? ① 해수 1kg중에 녹아있는 염류의 g수를 말한다. ② 해수 100g중에 녹아있는 염류의 g수를 말한다. ③ 해수 1m ³ 중에 녹아있는 염류의 g수를 말한다. ④ 해수 1cm ³ 중에 녹아있는 염류의 g수를 말한다. ⑤ 해수 1kl중에 녹아있는 염류의 g수를 말한다.	0.38	0.54	0.80	0.92	0.15