



저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

강 태 훈 교수지도
석사학위 청구논문

선다형문항에서 차등배점을 고려한 IRT
능력모수 추정을 위한 탐색적 연구

- 등급반응모형 활용을 중심으로 -

2015

성신여자대학교 대학원

교육학과

김윤주

선다형문항에서 차등배점을 고려한 IRT
능력모수 추정을 위한 탐색적 연구

- 등급반응모형 활용을 중심으로 -

강 태 훈 교수지도

이 논문을 석사학위논문으로 제출함.

2014년 11월

성신여자대학교 대학원

교육학과

김운주

인 준 서

김윤주의 석사학위 논문으로 인준함.

2014년 11월

심사위원장 _____ 인

심사위원 _____ 인

심사위원 _____ 인

성신여자대학교 대학원

논문개요

본 연구의 목적은, 문항반응이론의 분석 맥락 하에서 차등배점 부여방식을 고려한 능력 추정방법을 탐색 및 개발하는 데에 있다. 이러한 목적을 이루기 위하여, 등급반응모형(Graded Response Model: GRM)을 활용하였으며, 이를 통하여 선다형 문항들에 차등배점 부여방식이 부여되었을 때에 이를 반영하여 능력 추정치를 구하는 방법을 제시하고자 하였다.

우리나라 학교현장에서 사용되는 학생들의 성취도를 평가하기 위한 검사 점수는 일반적으로 문항에 전문가의 주관적 판단에 의해 부여된 차등배점을 적용하여 이를 합산하는 방식으로 산출된다. 이러한 차등배점 부여 방식 하에서는 두 개인의 정답 문항 수가 같더라도 서로 다른 검사 점수가 부여될 수 있기 때문에 동등배점방식으로 점수를 산출한 결과와는 상이한 영향을 미칠 수 있다. 선행연구에 의하면, 대다수의 연구에서 문항 점수의 단순 합산으로 피험자의 능력을 나타내는 경우에 동등배점에 비하여 차등배점을 활용하는 것이 능력 추정의 정확도 측면에서 별다른 차이를 가져오지 않으며, 검사의 신뢰도나 타당도 역시 유의미하게 개선하지 못하는 것으로 보고하고 있다. 하지만, 위와 같은 선행연구에도 불구하고 현재 주요 학업성취도 시험(예를 들어, 국가수준학업성취도검사, 대학수학능력시험 등)에서는 학생들의 점수의 폭을 넓히기 위함뿐만 아니라 검사의 특정한 목적에 따라 전문가에 의해 부여된 차등배점 방식을 사용하고 있다.

이와 같은 상황에서 피험자의 능력을 추정하고 검사 결과를 분석함에 있어서 현대검사이론이라고 불리는 문항반응이론에서는 전문가의 주관적 판단에 의해 부여된 차등배점을 고려한 자료 분석 방법론이 거의 존재하지 않은 실정이다. 따라서 본 연구는 전문가의 주관적 판단에 의해 부여된 차등배점이 적용되는 시험 결과에 대하여 이러한 차등배점을 문항반응이론의 틀 안

에서 다를 수 있는 방법론이 요청된다는 문제의식에서 시작되었다.

본 연구에서 제시하는 전문가 판단에 따라 선다형 문항에 사전에 부여된 배점 혹은 가중치를 등급반응모형을 활용하여 고려하는 방법은 다음과 같다. 우선, 이분 문항반응모형을 활용하여 전통적 방법으로 문항의 변별도와 곤란도를 추정한다. 다음, 각 선다형 문항에 부여된 배점을 마치 다분형 문항에서 산출된 문항의 점수인 것처럼 다룬다. 이는, 한 선다형 문항의 배점이 3점이라면 이를 마치 0, 1, 2, 3점으로 채점된 주관식 혹은 구성형 문항에서 도출된 점수로 본다는 의미이다. 이 때 물론 실제 문항 자료에서는 0점과 3점만 존재하며 1점과 2점은 산출될 수 없을 것이다. 이러한 상황을 본 연구에서는, 등급반응모형에서 활용하게 되는 경계특성 곡선의 각 위치모수를 하나의 위치모수에 근접시키는 방식으로 다루어 해결하였다. 이렇게 산출된 각 문항의 문항 모수 추정치와 피험자의 문항 반응 자료를 바탕으로, 최종적으로, 각 피험자의 능력 모수를 추정하였다.

본 연구에서는 위의 방법을 적용하여 산출된 능력추정치와 전통적인 IRT 방식으로 산출된 능력추정치를 비교하고자, 두 추정치 간의 상관을 산출하였다. 그 결과, 두 추정치 간의 매우 높은 상관이 나타나 피험자의 순위를 부여하는 측면에서 두 추정 방법 간의 차이는 그리 크지 않음을 알 수 있었다. 그러나 각 원점수 구간별 RMSD를 통해 원점수 범위의 양 극단으로 갈수록 두 방법에 의한 능력추정치의 차이가 다소 증가함을 볼 수 있었다. 또한 두 가지 방법으로 산출된 능력모수의 순위와 원점수(차등배점을 고려한 원점수와 고려하지 않은 원점수 포함) 순위의 관련성을 보기 위해 두 가지 추정 방법으로 산출된 능력 추정치 값과 두 종류의 원점수 간의 순위상관계수를 산출한 결과, 원점수 서열에서 차등배점을 고려하였을 때에 추정된 능력모수의 서열이 차등배점을 고려하지 않은 능력모수의 서열에서보다 순위의 역전이 더 적게 일어나는 것으로 나타났다.

위의 연구 결과를 토대로 본 연구에서의 결론은 다음과 같다. 차등배점을 고려한 피험자의 능력추정치와 고려하지 않은 능력추정치간의 차이가 그리 크지 않았으며, 차등배점을 고려한 능력추정 방식이 원점수의 순위를 더 잘 반영해줄 수 있는 것으로 나타났다. 따라서 검사의 목적에 따라 사전에 차등배점이 부여된 상황에서 문항반응이론을 통하여 피험자의 능력추정이 요구될 때에 본 연구에서 제안하는 차등배점을 고려한 능력추정방식은 기존의 IRT방식보다 원점수의 순위를 보다 더 잘 반영해 줄 수 있다는 측면에서 효과적 대안으로 고려될 수 있을 것이다.

목 차

논문개요

I. 서론	1
1. 연구의 필요성 및 목적	1
2. 연구 문제	5
II. 이론적 배경	6
1. 문항배점방식	6
2. 등급반응모형	11
3. 차등배점 고려를 위한 위치모수 고정법	20
4. 관련 선행연구	24
III. 연구 방법	28
1. 분석대상	28
2. 연구절차	30
IV. 연구 결과	34
1. 검사의 양호도 및 일차원성 가정 검증	34
2. 문항 모수 산출 결과	37
3. 두 능력 추정치 간 차이 비교	39
4. 원점수 순위와의 관련성	43

V. 결론 및 논의 47

참 고 문 헌

ABSTRACT

표 목 차

<표 III-1> 검사 문항의 배점 구성	29
<표 III-2> 검사 문항의 평가 내용 및 행동 영역	29
<표 IV-1> 검사의 평균, 표준편차, 최소값, 최대값	34
<표 IV-2> 요인 설명량	35
<표 IV-3> 검사문항의 난이도, 변별도 모수	38
<표 IV-4> 능력추정치 기술통계량	39
<표 IV-5> 능력추정치와 원점수의 상관	39
<표 IV-6> 원점수 구간 별 능력추정치 간 RMSD	41
<표 IV-7> 피험자 응답패턴	42
<표 IV-8> 순위상관계수	43
<표 IV-9> 차등배점을 고려하지 않은 능력추정치에 대한 피험자 순위 변화정도	44
<표 IV-10> 차등배점을 고려한 능력추정치에 대한 피험자 순위 변화정도	44
<표 IV-11> 원점수 구간 별 피험자 순위 변화정도	46

그림 목 차

[그림 II-1] 등급반응모형에서 문항반응범주특성곡선의 기본원리	13
[그림 II-2] 문항반응범주특성 곡선	14
[그림 II-3] 경계특성 곡선	15
[그림 II-4] 경계특성곡선에서 위치모수 조정 원리	22
[그림 IV-1] 전체 피험자의 검사점수 분포	35
[그림 IV-2] 스크리 도표	36
[그림 IV-3] 차등배점 고려 여부에 따른 θ 값 산포도	40
[그림 IV-4] 원점수 구간별 RMSD	41

I. 서론

1. 연구의 필요성 및 목적

가정, 학교, 사회 전반에서 이루어지는 교육은 그 목적에 부합하게 인간으로 하여금 행동의 변화를 가져왔는지 판단하는 과정이 필요하다. 이러한 과정을 교육평가라 정의할 수 있다. 교육평가는 교육과 관련된 모든 것의 양, 정도, 질, 가치, 장점 등을 체계적으로 측정하여 판단하는 주관적 행위로서 교육목적에 대한 가치를 판단하는 행위를 의미한다(성태제, 2010). 교육평가를 통해 우리는 교육과 관련된 정보를 수집하고 이를 통해 교육적인 의사결정을 내릴 수 있다. 교육평가는 그 방법에 따라 양적인 평가와 질적인 평가로 나뉘며, 양적인 평가방법에서 자료를 수집하는 여러가지의 방식 중에서 교육상황에서 대표적으로 사용되는 방식의 하나가 검사이다(양길석, 2007).

검사는 교육현장뿐만 아니라 피험자의 특성에 대한 평가가 요구되는 다양한 상황에서 피험자의 보이지 않는 잠재능력을 추정하는데 활용된다. 우리나라에서 실시되는 대표적인 검사의 예로는 학생들의 학업성취를 진단하고자 실시하는 국가수준학업성취도 평가, 대학 입시에서 학생을 선발하기 위한 목적으로 실시되는 대학수학능력시험이 있으며, 이외에도 성격 검사, 적성검사, 흥미검사 등 피험자의 잠재적 특성을 측정하는 다양한 검사가 존재한다. 교육현장에서 실시되는 검사는 학생들의 학업성취 정도를 진단하는 목적으로 쓰이지만, 이와 더불어 선발, 배치, 자격부여 등 중요한 의사결정을 하는데 사용되기 때문에 검사를 실시하는데 있어서 검사는 피험자의 능력을 제대로 추정할 수 있는 적절한 타당도와 신뢰도를 갖추어

야 함이 강조되며, 검사점수의 부여방식에 있어서도 공정성을 유지해야 할 필요가 있다(김신영, 노국향, 1999).

검사에 있어서 검사점수란 검사가 재고자 하는 구인 혹은 구인들에 대한 피험자의 반응의 결과물을 요약한 것이라 할 수 있다(Thissen & Wainer, 2001). 고전검사이론에서의 검사 점수는 각 문항의 점수를 합한 총점으로 이루어지는데, 이 총점은 피험자의 능력에 대한 추정치로 활용된다(Crocker & Algina, 1986; 박찬호, 강태훈, 2011). 검사점수를 부여하는 배점 방식은 일반적으로 모든 문항에 대해 동일한 가중치를 주거나, 문항에 따라 가중치를 다르게 부여하는 방식에 따라 동등배점방식과 차등배점방식으로 구분된다. 동등배점방식으로 실시되는 검사의 예로는 의·치 학교육입문검사, 법학적성시험, 공직적격성평가, 전문연구요원 선발시험 등이 있으며(김주훈 외, 2010), 차등배점방식 예로는 대표적으로 대학수학능력시험과 국가수준학업성취도 평가 등이 있다. 차등배점이 부여되는 방법은 전문가의 주관적인 판단에 의한 방법, 합리적 방법(예: 난이도에 의해 부여하는 방법, 문항점수와 총점과의 상관에 의해 부여하는 방법), 경험·통계적 방법(예: 신뢰도를 최대화하는 방법) 등이 있다(Wang & Stanly, 1970; 김지현, 2000).

우리나라 학교현장에서 사용되는 학생들의 성취도를 평가하기 위한 검사점수의 종합방식은 합리적 또는 경험·통계적 방법 보다는 전문가의 주관적 판단에 의한 방법이 주로 사용되었다(김신영, 노국향, 1999). 사전에 전문가에 주관적 판단에 따라 문항의 길이, 푸는속도, 난이도, 중요도 등을 고려하여 문항에 차등한 배점을 부여하고, 각 개인의 원점수는 맞춘 문항에 부여된 배점을 합하여 산출된다. 이와 같은 차등배점방식 하에서는 두 개인의 정답 문항 수가 같더라도 서로 다른 검사 점수를 부여받게 될 수 있다. 이는 전문가의 판단에 의한 차등배점 부여 방식은 검사결과에 대해

여 동등배점방식과는 상이한 영향을 미칠 수 있음을 의미한다.

현재까지 차등배점 부여방식의 효용성에 대한 연구는 활발히 이루어져 왔다. 선행연구에 의하면 일반적으로 문항의 총합으로 피험자의 능력을 추정하는 경우에는 차등배점을 활용하는 것이 의미 없고, 검사의 신뢰도나 타당도를 개선하지 못하는 것으로 알려져 있다(Guilford, 1954; Gulliksen, 1950; Lord & Novick, 1968; Wang & Stanley 1970; 박찬호, 강태훈, 2011, 재인용). 성태제(1994)는 대학수학능력시험에서의 차등배점 부여 근거가 미약하다고 지적하고 있으며, 김신영·노국향(1999)연구에서도 역시 차등배점의 논리의 일관성을 파악하기 어려우며, 특별한 이유와 근거가 없다면 동일배점을 사용하는 것이 더 적합하다고 제안하였다. 이 외 차등배점과 관련된 연구에서도 차등배점 부여방식은 동등배점에 비해 더 많은 가짓수의 점수를 생성할 수 있다는 장점이 있을 뿐 다른 긍정적인 효과는 없으므로 피험자를 선발이나 자격부여를 목적으로 등수를 매겨야 하는 경우가 아니라면 동등배점부여방식이 더 적합하다고 제시하고 있다(양길석, 2007; 김주훈 외, 2010). 하지만 위와 같은 선행연구에도 불구하고 현재 대다수의 시험에서는 점수의 폭을 넓히기 위한 목적뿐만 아니라, 검사문항의 속성과 특정한 검사의 목적을 반영하여 피험자를 변별하기 위해 차등배점 방식을 사용하고 있다.

일반적으로 검사를 통해 얻게 되는 개인의 원점수는 각 문항의 배점을 합산하여 산출된다. 따라서 문항에 차등배점이 주어지게 되는 경우에 개인의 원점수는 이를 반영하여 총점으로 산출된다. 그러나 현대검사이론으로 불리우는 문항반응이론 하에서는 검사총점에 의하여 분석하는 것이 아니라 문항마다 갖고 있는 문항 고유의 문항특성곡선에 의해 피험자의 능력과 문항 모수를 추정한다. 따라서 전통적인 문항반응이론의 관점에서 볼 때, 각 선다형 문항에 대한 채점은 맞고 틀리고의 이분적 결정에 따라

1점과 0점이 주어질 뿐이며 문항에 대한 가중치는 문항변별도 모수에 의하여 통계적으로 고려될 뿐, 사전에 부여된 가중치는 일반적으로 고려하지 않는다. 따라서 아직까지 문항반응이론을 통한 검사 자료 분석에 있어서 전문가의 주관적 판단에 의한 문항의 차등배점을 고려하는 분석 방법론은 거의 없다고 볼 수 있다.

국내외에서 자격부여, 선발, 진단 등의 목적으로 실시되는 대다수의 검사가 전문가의 주관적 판단에 의한 차등배점 부여방식으로 배점이 부여됨에도 불구하고 문항반응이론 하에서는 이를 고려하는 분석방법이 없는 상황에서, 본 연구는 문항반응이론에서 차등배점부여방식을 고려한 능력 추정방법을 탐색해야할 필요하다는 문제의식에서 시작되었다. 만일 실시하는 검사의 목적에 따라 합리적으로 문항에 차등배점이 부여되었다면, 이러한 차등배점을 검사 자료 분석 과정에서 고려할 때에 분석결과는 검사의 목적을 보다 더 잘 반영할 수 있을 것이다. 문항반응이론은 피험자 집단에 따라 문항의 모수와 피험자의 능력추정치가 변하지 않는다는 점, 즉 문항모수 불변성과 능력모수 불변성이라는 점에서 고전검사이론과 구분되는 장점을 지니고 있다. 따라서 사전에 부여된 문항의 가중치를 문항반응이론 하에서 고려하는 방법을 개발함을 통하여 문항반응이론의 장점을 취하면서 동시에 문항에 부여된 차등배점을 고려하여 피험자의 능력을 추정함을 통해 검사의 목적을 보다 더 잘 반영할 수 있는 효과적인 대안으로 고려될 수 있을 것이라 기대한다.

따라서 본 연구의 목적은 문항반응이론의 분석 맥락 하에서 차등배점 부여방식을 고려한 능력 추정방법을 탐색 및 개발하고자 한다. 그리고 전문가 판단에 따른 차등배점을 고려하지 않는 기존의 방법론과의 비교를 통하여 본 연구에서 제안하는 차등배점을 고려한 능력추정방식의 적용가능성을 모색하고자 한다.

2. 연구문제

본 연구는 문항반응이론을 통해 차등배점이 부여된 검사를 분석할 수 있도록 Samejima(1969)의 등급반응모형을 활용하여 새로운 분석방법을 탐색하고 개발하고자 하는데 그 목적이 있다. 더 나아가 이를 실제 자료에 적용해봄으로써 본 연구에서 제시하는 새로운 능력모수 추정방법을 통해 산출된 피험자의 능력 추정치와 피험자의 기존 모형에서의 피험자 능력 추정치의 비교를 통해 새로운 추정방법의 적용가능성을 보고자 하였다. 이를 위해 먼저 차등배점을 고려하였을 때와 고려하지 않았을 때의 능력추정치에 어느 정도의 차이가 있는지 살펴보고, 원점수 순위와의 관련성과 순위변화 정도를 통하여 본 연구에서 제안하는 새로운 방법을 통한 능력추정치가 원점수의 순위를 얼마나 잘 반영하는지를 보고자 하였다. 이하 본 연구에서는 편의상 차등배점을 고려한 능력추정치는 “ID”(IRT의 I와 Differential의 D를 합한 약어)로, 차등배점을 고려하지 않은 능력추정치는 “IE”(IRT의 I와 Equal의 E를 합한 약어)로 지칭한다.

본 연구의 연구문제는 다음과 같다.

첫째, 차등배점이 부여된 검사 자료를 분석함에 있어서 ID와 IE간의 차이가 있는가?

둘째, ID와 IE의 원점수 순위와의 관련성은 어떠한가?

셋째, 피험자 능력모수 추정에 있어서 차등배점 고려여부에 따라 피험자의 순위 변화 정도는 어떠한가?

II. 이론적 배경

1. 문항배점 방식

문항의 부여된 차등배점을 고려하여 피험자의 능력추정 방법을 제안하기 전에, 본 연구에서 다루고자 하는 차등배점의 의미를 명확히 하기 위해 문항의 배점부여 방식에 대해 살펴보려고 한다. 일반적으로 검사점수는 검사를 구성하는 여러개의 문항 또는 하위 검사에서 얻어진 점수를 합산하는 방식으로 산출된다(김주훈 외, 2010). 이 때에 문항에 배점을 부여되는 방식은 각 문항에 동일한 배점을 부여하느냐 혹은 서로 다른 배점을 부여하느냐에 따라 동등배점방식과 차등배점방식으로 나뉘게 되며, 차등배점방식이 경우 그 방법에 따라 다시 세분화된다.

가. 동등배점방식

동등배점방식은 모든 문항에 동일한 배점을 부여하는 방법이다. n 개 문항에 대한 동등배점 공식은 아래 식 (1)과 같다.

$$X = I_1 + I_2 + I_3 + \dots + I_n \quad (1)$$

X 는 검사 총점을 의미하며 I_1 부터 I_n 은 n 개의 문항 점수를 의미한다. 동등배점 방식에서 각 문항의 점수는 총 점수를 문항의 수로 나누어 산출하여 각 문항의 점수는 검사 점수에 대해 동일한 비율을 갖는다. 동등배점 방식

은 채점절차와 검사점수의 해석이 용이하다(김신영, 노국향, 1999)는 장점이 있다. 그러나 배점을 통해 문항의 차별적인 속성(난이도, 문제풀이시간, 중요도 등)을 반영할 수 없고, n개의 문항일 때 점수의 가짓수가 n+1개로 생성되어 상대적으로 점수의 폭이 좁아 동점자가 많이 발생할 수 있다는 제한점이 있다.

나. 차등배점방식

1) 차등배점의 정의

차등배점이란 검사를 이루고 있는 문항 혹은 하위검사에 대하여 서로 다른 배점을 부여하는 방식을 의미한다. 차등배점의 기본가정은 검사를 구성하는 문항의 측정내용, 측정수준이 동일하지 않다는 것이다(양길석, 2007). 차등배점은 난이도, 중요도, 풀이시간 등 각 문항의 속성을 고려하여 배점을 부여할 수 있으며, 검사를 실시하는 목적에 따라서 적합한 방법으로 부여할 수 있다. 차등배점은 문항 별로 점수가 다르기 때문에 같은 문항을 맞춘 피험자라 할지라도 서로 다른 점수를 얻을 수 있다. 차등배점 부여시 n개의 문항에 대한 검사점수는 식 (2)와 같이 산출된다.

$$X = w_1I_1 + w_2I_2 + w_3I_3 + \dots + w_nI_n \quad (2)$$

차등배점은 부분점수 부여 방식과 그 의미를 구분할 필요가 있다. 차등배점 방식은 부분점수 부여 방식과는 달리 각 문항에 대하여 고정된 점수를 부여한다. 부분점수 부여 방식은 한 문항에 대하여 피험자에 응답에 따라 정해진 점수 범위 내에서 서로 다른 점수를 부여하기 때문에 같은 문항이라

할지라도 피험자에 따라서 서로 다른 점수를 받게 될 수도 있다. 그러나 차등배점 방식은 정답과 오답의 경우만 존재하는 이분형 문항에 대해 부여하게 되며, 문항에 따라 다른 배점을 부여하게 된다. 따라서 해당 문항에 대해 정답을 맞춘 피험자는 서로 동일한 점수를 부여받게 된다.

2) 차등배점을 부여하는 방법

가) 전문가 판단 방법

전문가 판단 방법은 전문가가 문항의 중요도, 난이도, 문항의 형태 등을 고려하여 주관적으로 문항에 배점을 부여하는 방법이다. 현재 대학수학능력시험, 국가수준학업성취도 평가, 학교 현장에서 실시되는 각종 시험 등에서 사용되는 가장 보편적인 차등배점 부여 방법이다. 전문가 판단 방법은 검사의 활용목적에 따라 전문가의 판단에 따라 임의로 점수를 부여하기 때문에 피험자들을 검사의 목적에 맞게 변별을 할 수 있다는 장점을 지닌다. 그러나 전문가의 개인적이고 주관적인 판단에 의존하기 때문에 판단의 준거가 명료하지 않은 경우에 배점에 대한 타당성을 증명하기 힘들고, 객관성을 잃기 쉽다는 단점을 지닌다. 따라서 전문가 판단 방법은 검사의 활용목적과 명료한 준거에 의해 문항에 적절한 배점을 부여하는 전문가적인 판단이 무엇보다 절실히 요구된다. 전문가 판단 방법은 검사 시행 이전에 문항에 대한 가중치가 부여되기 때문에 검사 이후 즉시 검사 점수를 산출할 수 있고, 피험자가 문항에 대한 가중치를 미리 알 수 있다는 점에서 다른 방법과 차이가 있다.

나) 합리적인 방법

합리적인 방법은 외부의 기준이 되는 변인이 없는 경우에 합리적인 근거

에 의해 문항 점수를 합산하여 검사점수를 산출하는 방법이다(김신영, 노국향, 1999). 합리적인 방법에는 최대검사신뢰도에 따른 차등배점부여 방법, 검사점수와 문항점수의 상관에 따른 차등방법, 문항 난이도에 따른 차등방법 등이 있다.

① 최대 검사 신뢰도에 따른 차등배점 부여 방법

최대검사신뢰도에 따른 배점부여방법은 문항의 배점을 검사의 신뢰도가 최대가 되도록 부여하는 방법이다. 고전검사이론에서 검사의 신뢰도는 관찰 점수의 분산에 대한 진점수 분산의 비로 정의한다. 이에 근거하여 가중치가 부여되는 검사의 신뢰도를 아래 식 (3)과 같이 나타내고, 이 신뢰도 값이 최대가 되도록 하는 가중치 값을 산출하는 방식이다(Wang & Stanly, 1970).

$$\rho_{xx'} = \frac{\sum w_i^2 \rho_{ii'} + \sum \sum w_i w_j \rho_{ij}}{\sum w_i^2 + \sum \sum w_i w_j \rho_{ij}} \quad (3)$$

여기에서 $\rho_{xx'}$ 는 검사 신뢰도, w_i 는 문항 i 에 대한 가중치, $\rho_{ii'}$ 는 문항의 신뢰도, ρ_{ij} 는 문항간의 상호상관을 의미한다.

② 검사점수와 문항점수의 상관에 따른 차등배점 부여 방법

검사점수와 문항점수 간의 상관에 따라 차등배점을 부여하는 방법은 문항이 검사점수에 미치는 영향이 동일하다는 가정을 바탕으로 검사를 이루는 각 문항과 검사점수 간의 상관이 동일하도록 가중치를 부여하는 방법이다.

단일 영역 혹은 사실 내용에 대한 문항이 아닌 여러 부분에 대한 복합적인 문항으로 이루어지는 검사에 적합한 방법이다(김지현, 2000).

검사점수와 문항점수 간의 상관은 아래 식 (4)를 통해 구할 수 있으며, 이

를 통해 Kaiser(1967)에 의해 유도된 식 (5)를 이용하여 문항간 상관을 동일하게 하는 가중치를 산출할 수 있다.

$$C_i = Cov(X_i, X_1) + Cov(X_i, X_2) + \dots + Cov(X_i, X_{i-1}) + VarX_i + Cov(X_i, X_{i+1}) + \dots + Cov(X_i, X_n) \quad (4)$$

$$\rho_i = \frac{w_i + \sum w_j \rho_{ij}}{\sqrt{w_i^2 + \sum \sum w_i w_j \rho_{ij}}} \quad (5)$$

w_i = 각 문항의 가중치, ρ_{ij} = 문항 i 와 j 의 상관계수

이 방법은 문항 간의 상관이 정적일 경우에만 적용할 수 있음을 유의해야 한다.

③ 문항 난이도에 따른 차등배점 부여 방법

난이도에 따른 차등배점 부여 방법은 난이도가 높은 문항을 맞춘 피험자가 쉬운 문항을 맞춘 피험자보다 능력이 더 뛰어나다는 가정에 근거하여 문항에 가중치를 부여하는 방법이다. 이 방법에 의하면 난이도가 높은 문항에 높은 가중치가 부여되고, 상대적으로 난이도가 낮은 문항에 대해서는 낮은 가중치가 부여된다. 따라서 문항에 대한 배점은 문항을 맞추지 못한 피험자의 비율로도 나타낼 수 있다. 그런데 이 방법은 어려운 문항을 맞출 경우 더 높은 점수를 받지만, 어려운 문항을 틀린 학생은 쉬운 문항을 틀린 학생보다 더 많은 점수의 삭감이 있으므로 난이도에 의해 점수부여 시 이 부분에 대해서 신중히 고려해야 할 필요가 있다(김신영, 노국향, 1999).

다) 경험·통계적인 방법

경험·통계적인 방법은 외부에 준거가 있는 경우에 검사가 최대의 타당도를 갖도록 가중치를 부여하는 방법이다. 주로 중다회귀분석, 요인분석, 중다판별분석등의 통계적인 방법을 통해서 가중치를 부여한다(임인재, 1993). 경험통계적인 방법은 가중치를 각 문항에 적용하기보다는 하위검사에 가중치를 부여할 때 주로 쓰이며, 중다회귀분석 방법이 경험·통계적 방법으로 가장 흔히 쓰이는 방법이다.

위와 같이 차등배점을 부여하는 방법은 다양하게 존재하나, 본 연구에서는 차등배점을 부여하는 방식 가운데서도 보편적인 검사상황에서 흔히 쓰이게 되는 전문가의 주관적 판단에 의한 차등배점방식에 초점을 맞추어 연구를 실시하였다.

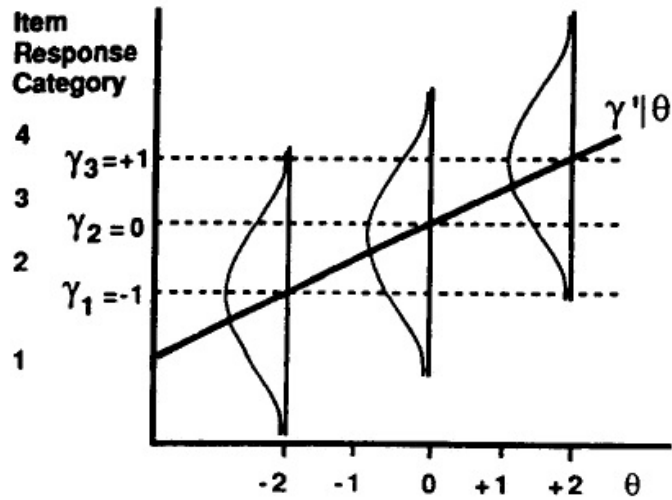
2. 등급반응모형

가. 등급반응모형의 개념

이분문항에서는 피험자의 반응이 인지검사에서는 ‘정답’ 혹은 ‘오답’, 태도나 정의적인 특성을 측정하는 검사에서는 ‘선호’, 혹은 ‘비선호’ 등 두가지 경우로 이분화 되어 나타나게 된다. 이분문항에서는 문제를 푸는 과정에서 피험자의 응답을 두 가지의 경우로 구분하기 때문에 문제를 푸는 과정에서 피험자의 능력이나 특성을 세분화할 수 없다. 예를 들어 특정 문항에 대해 정답을 맞추지 못한 피험자라 할지라도 정답에 근접하게 사고한 피험자와 아예 정답을 맞추지 못하는 피험자의 능력은 차이가 있지만 이분문항에서는 이들의 능력을 구분할 수 없다. 다분 문항은 이러한 관점에서 문항이

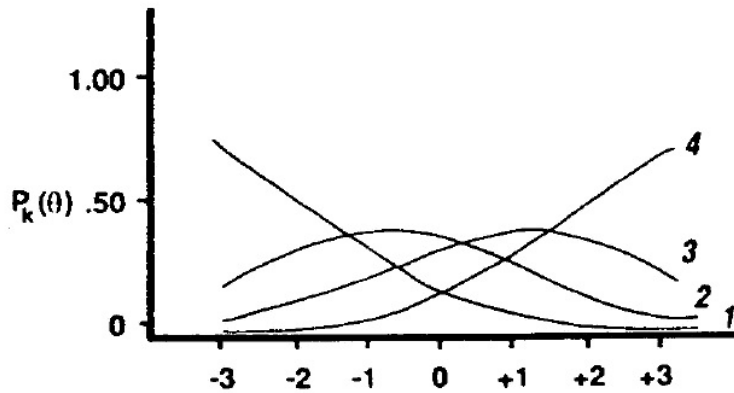
3개 이상의 점수범주를 가질 수 있도록 고안된 문항이라 할 수 있다. 예를 들어, 4점만점의 문항의 경우 0, 1, 2, 3, 4점의 5개의 범주로 나눌 수 있다. 다분문항을 분석할 수 있는 다분문항반응모형에는 등급반응모형(Graded Response Model: GRM), 부분점수모형(Partial Credit Model: PCM), 평정척도모형(Rating Scale Model: RSM), 일반화 부분점수 모형(Generalized Partial Credit Model: GPCM) 등이 있다.

본 연구에서 활용할 등급반응모형(Graded Response Model)은 검사를 이루고 있는 문항이 다분 문항인 경우에 적용할 수 있는 다분 반응모형 중 하나로, 문항에서 피험자가 받을 수 있는 문항점수의 범주가 3개 이상인 경우에 각 범주별 피험자의 정답을 맞출 확률을 산출할 수 있는 모형이다. 등급반응모형은 1969년에 Samejima에 의해 제안되었으며, 2모수 문항반응모형을 다분 문항으로 확장된 모형이다. 따라서 점수범주가 2개인 이분문항에 대한 2모수 문항반응모형은 등급반응모형의 특수한 예라고 볼 수 있다. 등급반응모형에서는 각 문항에 대한 피험자의 반응을 m 개의 범주로 구분하여 첫 번째 범주는 가장 낮은 단계이고, 마지막 범주는 가장 높은 단계의 반응으로 분류한다. 어떤 수행 과제에 대해서 0, 1, 2, 3점으로 부분점수를 인정하는 채점방식을 더한다든지, 심리검사에서 Likert 척도에 대해 ‘매우불만족’, ‘불만족’, ‘보통’, ‘만족’, ‘매우만족’으로 응답을 척도화하면 피험자의 반응은 등급반응이 된다(성태제, 1998). 등급반응모형에서는 [그림 II-1]과 같이 문항변수 r_i 를 피험자의 반응경향을 나타내는 연속변수로 보고 피험자의 각 능력수준에 따라 조건분포를 가정한다(Baker, 1992; 송미영, 1993).



[그림 II-1] 등급반응모형에서 문항반응범주특성곡선의 기본원리

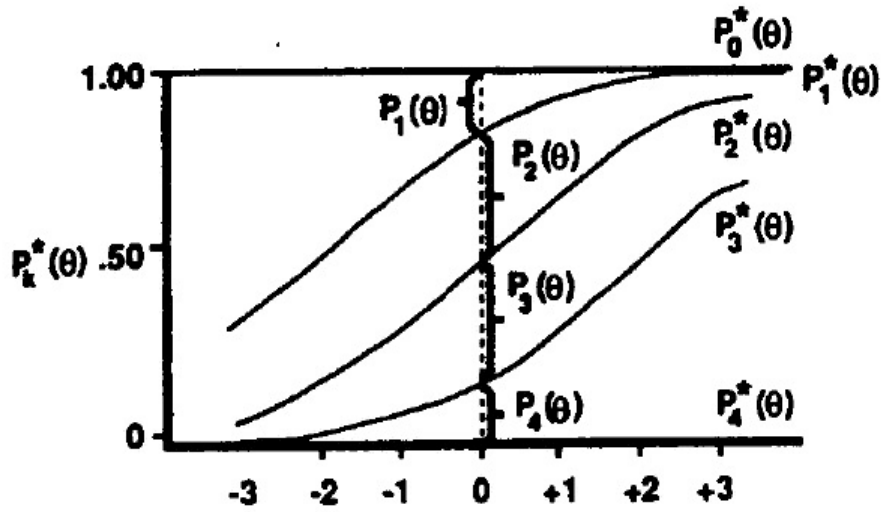
위의 [그림 II-1]에서 r_k 는 문항반응범주 k 와 $k+1$ 를 구분하는 기준치이다. 각 θ 값의 조건분포에서 r_k 와 r_{k+1} 사이에 해당하는 확률은 k 범주의 점수를 받을 확률을 뜻하며 이를 $P_k(\theta_j)$ 로 나타낼 수 있다. 즉, $P_k(\theta_j)$ 는 θ_j 의 능력을 가진 피험자가 k 범주의 점수를 맞을 확률을 의미한다. $P_k(\theta_j)$ 의 값을 각 범주별로 연결하게 되면 [그림 II-2]와 같이 문항반응범주특성곡선(item response category characteristic curve)을 그릴 수 있다.



[그림 II-2] 문항반응범주특성곡선

문항반응범주특성곡선은 피험자의 각 능력 수준에 따라 각 범주에 속할 확률을 나타낸다. 낮은 수준의 능력을 가진 피험자는 낮은 범주에 속할 확률은 높고, 높은 범주에 속할 확률은 낮다. 반면 높은 수준의 능력을 가진 피험자는 낮은 범주에 속할 확률은 낮은 반면, 높은 범주에 속할 확률이 높게 나타남을 볼 수 있다. 임의의 θ 값에서 각 범주에 속할 확률을 모두 합하면 1이 된다.

$P_k(\theta_j)$ 가 θ_j 의 능력을 가진 피험자가 k 범주에 속할 확률을 나타낼 때, $P_k^*(\theta_j)$ 는 θ_j 의 능력을 가진 피험자가 $k+1$ 범주 이상의 범주에 속할 확률로 정의할 수 있다. $P_k^*(\theta_j)$ 는 k 와 $k+1$ 범주의 경계가 되어, $1-P_k^*(\theta_j)$ 는 k 이하의 점수에 속할 확률이 된다. $P_k^*(\theta_j)$ 를 연결하면 [그림 II-3]과 같은 경계특성곡선으로 나타낼 수 있다.



[그림 II-3] 경계특성곡선

총 4개의 범주가 있을 때 $P_1^*(\theta)$ 는 2점 이상의 범주에 속할 확률이므로 $P_1^*(\theta) = P_2(\theta) + P_3(\theta) + P_4(\theta)$ 로 나타낼 수 있다. n개의 범주가 있다고 가정할 때, $P_k^*(\theta)$ 는 식 (6)과 같이 정의할 수 있다.

$$P_k^*(\theta) = P_{k+1}(\theta) + P_{k+2}(\theta) + \dots + P_n(\theta) \quad (6)$$

$$P_0^*(\theta) = \sum_{k=1}^n P_k(\theta) = 1$$

각 경계특성곡선은 $k+1$ 이상의 범주에 속하는 경우를 정답으로 보고 k 이하 범주에 속할 확률을 오답으로 보았을 때의 이분문항반응이론에서 2모수 문항특성곡선과 일치하게 된다. 따라서 문항 i 에 대해서 $P_k^*(\theta_j)$ 는 아래 식 (7)과 같이 정의할 수 있다.

$$P_{ik}^*(\theta_j) = \frac{\exp[1.7a_i(\theta_j - b_{ik})]}{1 + \exp[1.7a_i(\theta_j - b_{ik})]} \quad (7)$$

$P_{ik}^*(\theta_j)$: 문항 i 에서 피험자 j 가 $k+1$ 이상의 범주에 속할 확률

a_i : 문항 i 의 변별도 모수

θ_j : 피험자의 능력 추정치

b_{ik} : 문항 i 의 k 범주에 대한 난이도 모수

k 범주에 속할 확률인 $P_k(\theta)$ 는 다음 식 (8)과 같이 나타낼 수 있다.

$$P_k(\theta) = P_{k-1}^*(\theta) - P_k^*(\theta) \quad (8)$$

각 경계특성곡선은 변별도 모수와 난이도 모수(위치모수)로 결정된다. 각 범주에 해당하는 경계특성곡선은 같은 변별도 모수를 갖고, n 개의 범주에 대해서 $n-1$ 개의 위치모수를 갖게 된다. 경계특성곡선에서 위치모수 b_{ik} 는 $P_k^*(\theta)$ 를 경계특성곡선에서 k 범주보다 높은 범주에 속할 확률이 0.5에 해당하는 피험자의 능력을 의미한다. 문항반응범주특성곡선에서의 위치모수를 b' , 경계특성곡선에서의 위치모수를 b 라 할 때 그 관계는 아래와 같다(성태제, 1998).

범주	위치모수
1	$b'_1 = b_1$
$k (1 < k < n)$	$b'_k = \frac{b_k + b_{k-1}}{2}$
n	$b'_n = b_{n-1}$

n개의 범주가 있는 경우에 첫 번째 범주에 해당하는 b'_1 은 경계특성곡선에서의 첫 번째 범주보다 높은 범주에 속할 확률이 0.5에 해당하는 능력추정치인 b_1 과 같다. 그리고 중간범주에 해당하는 k 범주에 대한 문항범주특성곡선의 위치 모수 b'_k 는 인접한 k 와 $k-1$ 범주의 경계특성곡선의 위치모수의 중간지점이 된다. 그리고 마지막 범주인 n 번째 범주의 문항범주특성곡선의 위치모수 b'_n 은 $n-1$ 범주에 대한 경계특성곡선에서의 위치모수와 동일하다.

등급반응모형은 각 범주에 대해서 서열적인 관계를 가정하기 때문에 경계특성곡선에서 각 위치모수는 서열적이다. n개의 범주에 대한 각 경계특성곡선의 위치모수의 관계는 다음과 같다.

$$b_{n-1} > b_{n-2} > \dots > b_2 > b_1$$

가장 높은 범주에 대한 위치모수가 가장 큰 값을 갖고, 가장 낮은 범주에 대한 위치모수는 가장 작은 값을 갖게 된다.

나. 문항 및 능력모수 추정 방법

1) 최대우도추정법

문항모수와 능력모수를 추정하기 위한 IRT의 접근법은 크게 최대우도 추정법(Maximal Likelyhood Estimation: MLE)과 베이저안 추정법(bayesian estimation)로 나눌 수 있다.

등급반응모형에서 최대우도추정방법에 기반한 문항모수를 추정하기 위해서는 피험자의 능력 수준에 따라 J개의 집단으로 나눈다. 그리고 피험자의 반응을 n개의 범주로 나눌 때 j집단의 피험자 n_j 명 중에서 k번째 문항 범주에 속한 피험자가 r_k 명이라면 능력수준 θ_j 의 피험자가 k번째 문항반응 범주에 속한 관찰 비율은 $p_k = \frac{r_k}{n_j}$ 이고, 이론적인 확률은 $P_k(\theta_j)$ 이다(송미영, 1993). 각 능력집단에서 각 문항 반응 범주에 속할 확률에 대한 우도함수는 아래 식과 같다.

$$\text{Prob}(r_1, r_2, \dots, r_n) = \prod_{j=1}^J \frac{n_j!}{r_1! r_2! \dots r_n!} P_1^{r_1} P_2^{r_2} \dots P_n^{r_n} \quad (9)$$

위의 우도함수 값을 최대로 하는 문항모수 값이 관찰된 자료에 가장 적합한 문항모수값이 되고 이러한 방식으로 문항모수를 산출하는 방법이 최대우도추정방법이다. 최대우도추정법에 의해 문항모수와 피험자의 능력을 추정하는 방법은 결합최대우도추정법(joint maximum likelyhood estimation), 주변최대우도추정법(marginal maximum likelyhood estimation) 등으로 나뉜다(Swaminathan, 1983).

2) 사후기대 추정법

베이지안 추정법은 피험자의 능력분포에 대해 사전정보를 사용하는 방법이다. 베이지안 추정법의 핵심이 되는 식은 다음과 같다(Hambleton & Swaminathan, 1985).

$$P(B|A) = [P(A|B)P(B)]/P(A) \quad (10)$$

능력추정에 있어서는 위의 식에서 A 값이 θ 로, B 값은 u 로 관찰된 문항 반응 벡터를 의미한다. 따라서 (식 11)과 같이 나타낼 수 있다.

$$g(\theta|u) = \frac{L(u|\theta)g(\theta)}{P(u)} = \frac{L(u|\theta)g(\theta)}{\int_{-\infty}^{+\infty} L(u|\theta)g(\theta)d\theta} \quad (11)$$

$g(\theta|u)$ 는 사후분포가 되고, $g(\theta)$ 는 θ 의 사전분포이다. $L(u|\theta)$ 는 관찰된 반응의 우도함수가 된다. 위의 관계에 의해 $g(\theta|u)$ 는 $L(u|\theta)$ 와 $g(\theta)$ 의 곱에 비례한다. 사후기대 추정법(Expected a Priori)은 베이지안 추정 방법 중 하나로 능력모수 θ 의 사후기대(EAP) 추정량 $\hat{\theta}_{EAP}$ 는 사후분포의 평균으로 아래 식과 같이 정의된다.

$$\hat{\theta}_{EAP} = \int_{-\infty}^{+\infty} \theta g(\theta|u)d\theta = \frac{\int_{-\infty}^{+\infty} \theta L(u|\theta)g(\theta)d\theta}{\int_{-\infty}^{+\infty} L(u|\theta)g(\theta)d\theta} \quad (12)$$

3. 차등배점 고려를 위한 문항모수 고정 방법

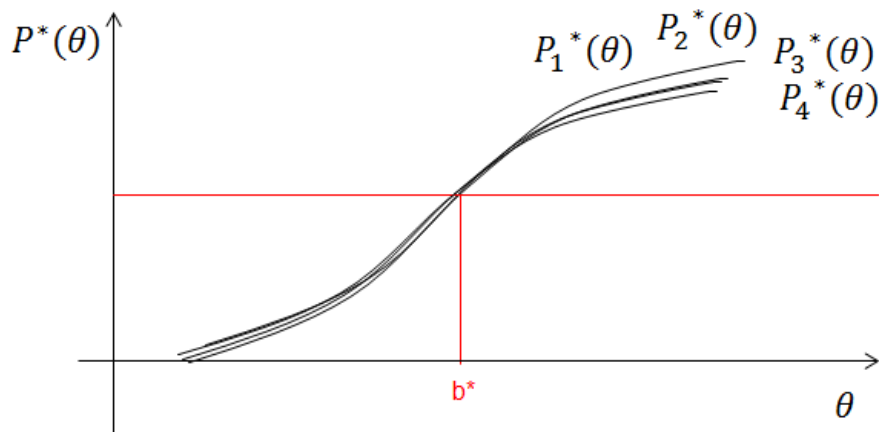
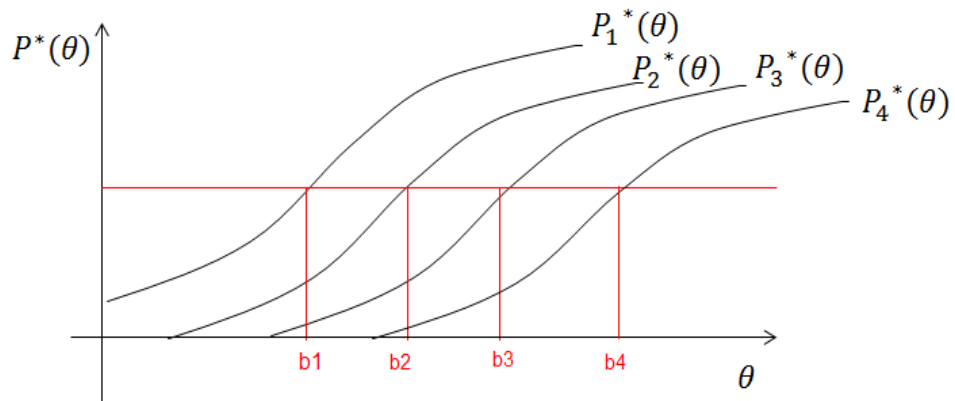
이분형 검사자료에서 문항반응이론을 통해 모수를 추정하는 경우, 문항을 맞추었을 경우를 1, 틀렸을 경우를 0으로 보고 피험자의 능력과 문항모수를 산출한다. 따라서 사전에 문항에 각기 다른 배점이 부여되었을 지라도 문항 반응이론 하에서 자료를 분석 할 때에는 이러한 차등배점을 고려하지 않으며, 2모수나 3모수 로지스틱 모형 하에서는 변별도가 문항의 부여되는 가중치 역할을 한다. 즉 같은 개수의 문항을 맞춘 피험자라 할지라도 변별도가 높은 문항을 맞추었을 때 더 높은 능력추정치로 얻게 된다.

본 연구에서는 특정한 목적에 의해 차등배점과 같이 검사를 시행하기 전에 부여된 가중치가 있는 경우 기존 문항반응이론에서는 고려하지 않은 사전 가중치를 반영한 능력추정치를 산출하기 위해 등급반응모형을 활용하여 차등배점을 고려한 문항모수 고정 방법을 도출하고자 하며 그 적용가능성을 탐색하고자 하였다. 문항모수 고정 방법은 Samejima(1969)의 등급반응모형을 활용한 방법으로, 문항의 배점을 다분반응문항의 부분점수와 같이 간주하고 각 범주에 해당하는 경계특성곡선의 모수를 조정하여 피험자의 능력 모수를 추정하는 방법이다. 예를 들어 배점이 4점인 문항이라면, 실제 피험자가 응답한 문항반응 자료에는 0점과 4점만 있지만 0, 1, 2, 3, 4 총 5개의 부분점수가 있다고 가정한다. 그리고 각 범주에 해당하는 경계특성곡선의 변별도와 난이도 모수를 2모수 로지스틱 모형 하에서 산출된 변별도와 난이도에 거의 근접하도록 조정하여 하나의 2모수 로지스틱 곡선처럼 나타나도록 조정한다. 즉, 4점이 부여된 문항인 경우 0, 1, 2, 3, 4의 5개의 범주가 있다고 가정하고, 각 범주에 대한 경계특성곡선의 b_1, b_2, b_3, b_4 의 값을 하나의 새로운 b^* 값에 근접하도록 각 위치모수의 값을 아래 식과 같이 조정해준다.

$$b_k = b^* + \epsilon_k \quad (13)$$

여기에서 b^* 값은 차등배점을 고려하지 않은 이분형 자료에서 산출된 난이도 모수를 의미한다. ϵ_k 는 각 범주에 대해 b^* 값에 더해지는 매우 작은 값을 의미한다. 위와 같이 경계특성곡선의 각 범주에 해당하는 위치모수를 조정하면 하나의 곡선에 거의 근접하게 되고 결국 이분형 자료에서 산출된 곡선과 변별도와 난이도 모수가 동일한 곡선이 된다. 5점문항인 경우에는 0, 1, 2, 3, 4, 5 이렇게 6개의 범주를 가정하여, 위와 동일한 방식으로 조정해준다. 경계특성곡선에서 위치모수 조정을 도식화하면 [그림 II-4]와 같다.

위와 같은 방법은 검사 동등화 방법 중 고정문항모수 추정(Fixed Item Parameter Calibration; FIPC) 방법과 유사하다. FIPC 방법은 두 검사를 치른 피험자들을 같은 척도에 두기 위하여 가교문항의 모수를 기준검사의 가교문항 모수로 고정하는 방법으로, 두 검사의 임의적인 선형변환이 되는 것을 막아 척도의 비결정성 문제를 해결하고, 피험자 집단의 능력모수를 같은 척도 상에서 추정될 수 있도록 하는 방법이다(박산하, 2014). 본 연구에서 적용하는 방법은 가교문항이 아닌 전체 문항에 대한 모수를 고정한다는 점에서 FIPC 방법과 차이가 있다.



[그림 II-4] 경계특성곡선에서 위치모수 조정 원리

차등배점을 고려하기 위한 문항모수 고정 방법은 PARSCALE4.1 프로그램 (Muraki & Bock, 2003)을 사용할 수 있다. PARSCALE에서 각 경계특성곡선은 아래 식과 같은 형태로 주어진다. 즉 각 경계특성곡선의 난이도 모수는 τ 값으로 조정되며, n 개의 범주가 주어질 때 식 (15)와 같이 $n-1$ 개의 τ 값의 합은 항상 0이 된다.

$$P_{ijx}^*(x | \theta_j, \alpha_i, \beta_{ij}) = \frac{\exp[\alpha_i(\theta_j - (\beta_i - \tau_{xi}))]}{1 + \exp[\alpha_i(\theta_j - \beta_i - \tau_{xi})]} \quad (14)$$

$$\tau_1 + \tau_2 + \dots + \tau_{n-1} = 0 \quad (15)$$

등급반응모형에서의 문항모수를 2모수 로지스틱모형으로 산출된 문항모수 (변별도, 난이도)의 값으로 대체하고, 각 범주에 대한 τ 값을 매우 작은 값(ϵ)으로 조정하여 'prm'파일을 생성한 뒤 'prm' 파일 상의 문항모수를 고정하여 피험자의 능력 모수를 추정하면 각 문항의 차등배점을 고려한 피험자 능력을 추정할 수 있다.

한가지 유의해야 할 점은 PARSCALE 상에서 문항모수를 고정하여 추정할 때에 따로 명령어를 입력하지 않으면 디폴트로 능력모수의 재척도화가 실행된다. 따라서 사전 능력모수를 계속 갱신하고, 추정 후 능력모수의 재척도화가 일어나지 않도록 이를 방지하기 위해 POSTERIOR 명령어와 NOADJUST 명령어를 입력하여 능력모수를 추정해야 한다(Kang & Peterson, 2009).

차등배점을 고려한 피험자 능력 추정 과정을 요약하면 다음과 같다.

1. 이분문항반응모형으로 문항모수를 산출한다.
2. 문항에 부여된 배점에 따라 각 문항의 범주 수를 생성한다.
3. 각 범주에 해당하는 경계특성곡선의 모수를 1에서 산출된 문항모수로 고정한다.
4. 고정된 문항모수를 통해 등급반응모형 하에서 피험자의 능력 모수를 추정한다.

4. 관련 선행연구

차등배점은 1920년대에 처음 도입되어 그 이후에 차등배점과 관련된 연구들이 등장하기 시작하였다. 특히 많은 선행연구에서 동등배점 방식과의 비교를 통해 차등배점의 도입의 타당성에 대한 연구가 이루어졌다. Douglass Spancer(1923)의 연구에서는 문항의 길이가 충분히 길다면 차등배점과 동등배점의 결과의 큰 차이가 없다고 하였으며, Wilk(1938), Stalnaker(1939)의 연구에서도 임의로 부여된 차등배점에 대해서 검사결과에 무의미한 영향을 미친다는 결론을 내렸다. Ewen(1967), Hulin(1968)의 연구에서도 문항수의 합산으로 이루어진 결과가 차등배점을 부여하는 결과보다 타당한 결과를 가져왔다는 결론을 맺었다(Waters, 1969; 김지현 2000). 위와같이 차등배점에 타당성에 대한 선행연구는 차등배점을 통해 얻는 이익이 차등배점을 고려하여 각 가중치를 산출하여 검사점수를 종합하려는 노력에 비해 별로 의미가 없다고 보고하였다(Guilford, 1954; Gulliksen, 1950).

국내에서 이루어진 차등배점관련 연구를 살펴보면, 김신영·노국향(1999)

연구에서는 대학수학능력시험 모의고사 자료를 통해 차등배점이 부여된 문항 특성, 동일배점간의 상호상관, 전문가의 평정 일치도 등을 비교 분석한 결과 차등배점이 부여되어야 할 특별한 이유와 근거가 없다면 동일배점을 부여하는 것이 타당하다는 결론을 맺었다. 양길석(2007)의 연구에서는 한국어능력시험을 대상으로 차등배점과 동일배점 방식에 따른 검사점수간 상관 및 각 방식별 검사신뢰도, 합격여부의 일치도 등을 분석한 결과 검사 양호도를 확보하면서 점수 가짓수를 늘린다는 점에서 차등배점의 일반적인 장점은 인정하나, 반대로 수험자를 줄세울 필요가 없는 자격부여 시험에서는 굳이 차등배점을 적용할 필요가 없음을 시사하였다. 김주훈 외(2010)의 연구에서는 의학교육입문검사를 대상으로 문헌 연구 및 전문가 협의회 등의 연구방법을 활용하여 차등 배점 방식의 타당성을 분석한 결과, 차등배점 부여시 그 기준이 선형적 추측에 기반하기 때문에 차등배점 자체에 임의성이 개입될 수 있으므로 차등배점 부여에 신중해야하며, 충분히 사회적 합의가 이루어질 수 있는지 검토해야 함을 제안하였다. 또한 고사별 성격과 목적에 맞게 배점방식을 결정해야 할 것을 언급하였다. 국내 선행연구의 결과를 종합하면, 차등배점은 주관적 판단에 의해 부여되기 때문에 부여의 타당성을 위해서는 배점에 대한 논리적인 근거에 의해 신중하게 부여되어야 함을 강조하고 있다.

위의 선행연구에서 지적하듯이 차등배점 부여는 일관된 논리로 부여되기 힘들며, 그 결과가 동등배점을 부여했을 때와 큰 차이가 없다고 보고하기도 하였다. 그러나 현재 대다수의 검사는 선발의 목적으로 피험자의 점수의 폭을 넓히기 위해서 혹은 특정한 검사의 목적을 반영하기 위해 차등배점이 보편적으로 활용되고 있다. 대표적으로 국내에서는 대학수학능력시험, 국가수준학업성취도평가 등을 예로 들 수 있다. 이러한 상황에서 현재 차등배점이 부여된 검사의 결과를 분석하는 방법은 고전검사이론에서만 다루어질 뿐,

문항반응이론 하에서는 거의 다루어지지 않는다고 볼 수 있다. 전문가의 주관적 판단에 의한 차등배점이 부여된 검사자료 분석에 대한 방법론적인 측면에서 이제까지의 선행연구에서는 차등배점을 고점검사이론 하에서 문항의 배점을 합산하는 방식으로 원점수를 산출하여 분석이 이루어져왔다. 그러나 현재 널리 활용되고 있는 검사이론인 문항반응이론 하에서는 검사자료를 동등배점으로 간주하고 문항반응모형으로 산출되는 자체 문항모수를 통하여 가중치가 통계적으로 부여될 뿐, 사전에 전문가에 의해 부여된 차등배점을 고려한 방법론은 거의 존재하지 않는다.

문항에 부여된 차등배점에 대한 직접적인 접근은 아니지만 문항에 사전에 부여한 가중치를 고려하여 문항반응이론을 통해 분석한 연구를 살펴보면, 국내에서는 박찬호·강태훈(2011)연구에서 전문가가 판정한 문항의 가중치를 문항의 변별도모수 대신 활용하여 일반화 부분점수 모형(GPCM)에 전문가 판정 문항의 가중치를 부여한 제한적 일반화 부분점수 모형(rGPCM)을 제시하여 이를 통해 사전에 문항에 전문가가 판정한 가중치를 피험자의 능력 추정 과정에 반영하였다. 이 연구에서는 문항에 부여된 배점에 대한 고려는 아니었지만, 각 문항에 통계적으로 추정되는 변별도 모수대신 전문가가 판정한 가중치를 부여한 rGPCM 모형도 GPCM과 유사한 수준의 적합도를 보였다는 점에서 충분한 활용가치를 지님을 보여주었다.

국외에서는 Sykes & Hou(2003) 연구에서 문항반응이론을 통해 점수를 추정하기 이전에 하위검사에 대해 사전에 부여된 가중치에 대해서 검사 특성 곡선(TCC)의 비율을 증가시켜 수정된 TCC를 통해 점수를 산출하는 방식을 도입하여 하위검사의 가중치를 고려한 문항반응이론의 방법을 제안하였다. 이 연구에서는 구성형 문항과 선다형 문항이 혼합되어 있는 검사상황에서 구성형 문항에 더 큰 가중치를 부여하기 위해 고안된 방법으로 구성형 문항의 TCC의 비율을 증가시켜 문항반응이론 모형을 통해 모수를 추정하는 방

법이다. 이 역시 문항에 부여된 차등배점을 고려한 방법은 아니지만 각 하위 검사에 해당하는 문항의 가중치를 다르게 부여하는 방법을 제시함으로써 문항반응이론 하에서 문항에 주어진 가중치를 고려하는 보다 직접적인 방법을 제안한 연구라 볼 수 있다.

위의 선행연구에서 살펴본 바와 같이 아직까지 문항반응이론 하에서 각 문항에 사전에 부여된 차등배점을 고려한 능력추정방법은 존재하지 않는다고 볼 수 있다. 따라서 본 연구에서는 차등배점을 부여하는 것이 불가피한 검사상황에서 문항반응이론을 통하여 차등배점을 문항에 가중치로 부여하여 피험자의 능력을 추정할 수 있는 방법을 제안하고 이를 실제검사 자료에 적용하여 기존의 문항반응이론을 통한 분석과 비교함을 통해 그 적용가능성을 탐색하고자 한다.

Ⅲ. 연구 방법

본 연구의 목적은 사전에 문항에 가중치가 부여된 검사 자료에서 문항반응이론을 통하여 이를 고려하는 방안을 모색하고자, 차등배점을 고려한 IRT 능력모수를 산출하는 방법을 제안하고, 두 방법에 의해 산출된 능력추정치(ID, IE)와 원점수간의 비교를 통하여 차등배점을 고려하는 새로운 방법에 대한 적용가능성을 검토하고자 하는 것이다. 이와 같은 목적을 위하여 본 연구에서는 경기도 지역에서 실시된 학생들의 학업성취도 검사 자료를 연구자료로 사용하여, 본 연구에서 제시하는 차등배점을 고려하는 방법으로 산출된 ID와 기존 IRT 방식으로 산출된 IE간의 상관 및 RMSD를 산출하고, 피험자의 반응패턴을 비교하고자 한다. 그리고 능력추정치와 원점수간 순위의 관련성을 보기 위해 두 능력추정치 ID, IE와 원점수간의 순위 상관과 순위변화정도를 보고자 하였다.

1. 분석 대상

가. 자료의 구성

본 연구에서 차등배점을 고려한 IRT 능력모수 추정을 위해 사용한 자료는 2014년도에 실시된 경기도 지역 중학생의 학업성취도 관련 자료이다. 대상은 전체 경기도 지역에서 총화표집된 중학교 3학년 학생 965명이며, 수학과목의 성취도 평가 자료를 사용하였다. 연구 자료에서 문항 수는 총 20개의 선택형 문항으로 구성되어 있다. 검사의 모든 문항은 5지선다형 문항이며, 문항 배점은 문항마다 차등배점이 부여되었으며, 문항의 배점 구성은

<표 III-1>과 같다.

<표 III-1> 검사 문항의 배점 구성

배점	문항 번호	문항 수
4점	1, 2, 7, 11, 18	5개
5점	4, 5, 6, 8, 9, 10, 14, 16, 19, 20	10개
6점	3, 12, 13, 15, 17	5개
합계		20개

검사 자료에서는 세 종류의 배점을 부여하였고, 문항의 배점은 사전에 전문가의 주관적 판단에 의해 난이도에 따라 배점이 부여되었다. 상대적으로 어렵다고 판단되는 문항에서는 6점, 중간수준 난이도의 문항에는 5점, 쉽고 판단되는 문항에는 4점이 부여되었다. 검사의 총점은 100점으로 1점 간격으로 차등배점이 부여되었는데, 4점 문항은 5개, 5점 문항은 10개, 6점 문항은 5개로 5점 문항이 가장 많았다. 검사에서 측정하는 평가내용과 행동영역은 <표 III-2>와 같다.

<표 III-2> 검사 문항의 평가 내용 및 행동 영역

	영역	문항번호	문항 수
평가내용	수와연산	1,2,3,4,5	5
	문자와 식	6,7,8,9,10,11,12,13	8
	확률과 통계	14,15	2
	기하	16,17,18,19,20	5
행동영역	개념 및 절차	4,5,7,9,11,18	6
	문제해결	2,6,8,12,14,19	6
	의사소통	1,10,15,16	4
	추론	3,13,17,20	4

평가내용은 중학교 3학년 수준의 수와 연산 영역, 기하 영역, 문자와 식 영역, 확률과 통계 영역 등 수학의 전 영역을 아우르는 다양한 영역의 문항으로 구성되었으며, 문자와 식 영역에 해당하는 문항이 총 8문항으로 가장 많이 출제되었다. 행동영역은 개념 및 절차 영역, 문제해결 영역, 의사소통 영역, 추론 영역, 총 4가지 영역으로 이루어졌다.

2. 연구 절차

가. 검사 양호도 및 일차원성 가정 검증

본 연구에서는 차등배점을 고려한 IRT 추정 방법을 실제 자료에 적용하기 위하여 다음과 같은 분석을 하였다. 우선 문항모수와 피험자 모수를 추정하기 전에 SPSS(Statistical Package for Social Science) 19.0 ver.을 활용하여 검사의 양호도 분석을 실시하였다. 검사도구의 신뢰도를 검증하기 위하여 문항 간의 내적 일관성 신뢰도 계수인 cronbach α 계수를 산출하였다. 또한 문항반응이론을 통해 문항의 자료를 분석하기 위해 문항반응이론의 기본 가정인 일차원성 가정을 충족하는지에 대한 자료의 적합도 검증을 위하여 주성분 분석(Principal component analysis)을 통해 이를 확인하고자 하였다.

나. 자료의 문항모수 및 피험자 모수 추정

차등배점을 고려하여 모수를 추정하기 전에, 전통적인 IRT 방식으로 문항 모수와 피험자 능력 모수를 추정하였다. 검사 자료의 문항 변별도와 난이도

를 추정하기 위하여 PARSCALE 프로그램을 사용하였다¹⁾. PARSCALE은 다분문항반응모형 하에서 문항모수와 능력모수를 추정할 수 있도록 Muraki & Bock(1993)에 의해 개발된 프로그램이다. 등급반응모형은 2모수 로지스틱모형에서 확장된 모형이기 때문에 이분반응모형인 2모수 모형은 등급반응모형에서 범주의 개수가 2개인 모형이라고 볼 수 있다. 따라서 명령문에서 범주의 수를 2개로 지정하여 2모수 로지스틱 모형 하에서 연구 자료의 문항모수를 추정하였다. PARSCALE에서는 피험자의 능력 모수를 추정하기 위해 최대우도추정법(Maximum Likelihood Estimation:MLE), 사후기대추정법(Expected a Posteriori: EAP)등의 방법을 선택할 수 있다. 본 연구에서는 EAP 방법을 통해 피험자 능력모수를 산출하였다. 이 과정에서 산출된 문항모수는 이 후의 과정에서 각 경계특성곡선의 모수를 고정하는데 활용되었으며, 피험자 능력 모수 즉 전통적 방법을 통해 산출된 피험자 능력 모수(IE)는 차등배점을 고려한 능력모수(ID) 값과 비교하기 위해 산출하였다.

다. 경계특성곡선의 문항모수 및 τ 값 조정

차등배점을 고려한 피험자 능력모수 추정을 위하여 본 연구에서는 문항에 부여된 배점을 다분반응문항에 부여된 부분점수의 개념으로 간주하여, 등급반응모형을 통해 분석하고자 하였다. 만약 배점이 4점인 경우에 실제 자료에는 0점과 4점만 존재하지만 0, 1, 2, 3, 4점 총 5개의 점수가 존재하는 것으로 가정하였다. 그리고 각 범주에 대한 경계특성곡선의 변별도와 난이도모수를 2모수 로지스틱 모형으로 산출한 변별도와 난이도 모수 값과 매우 근접한 값으로 고정하여, 결국 각 경계특성곡선이 하나의 곡선과 동일하게 나타나도록 조정해주었다. 연구 자료에서는 총 세 종류의 배점이 부여되었

1) 실행 명령어 파일 [부록 1] 참조.

으므로 4점의 범주의 개수는 5개, 5점인 경우는 6개, 6점인 경우는 7개의 범주가 되도록 하여 각 경계특성곡선의 모수를 고정하였다. 각 경계특성곡선이 하나의 곡선에 거의 근접해야 하므로 각 범주의 τ 값을 0.0009 미만으로 조정하였다.

위와 같이 고정된 변별도와 난이도 모수값과 τ 값이 반영된 'prm 파일'을 생성하고, 생성된 'prm 파일'을 문항모수로 하여 배점이 부여된 자료에서 PARSCALE 프로그램을 통해 피험자 능력을 추정하였다.

라. 차등배점 고려 여부에 따른 IRT 능력추정치 비교

본 연구에서 차등배점을 고려하여 분석한 결과와 기존 IRT 방식으로 산출한 ID와 IE 값을 비교하고 RD, RE와의 관련성을 보았다.

차등배점을 고려여부에 따라 산출된 값에 차이가 있는지 보기 위하여 상관계수를 이용하여 두 추정치의 관계를 보았다. 그리고 두 추정치간의 차이를 보기위하여 각 원점수(RD) 구간별 RMSD(Root Mean Standard Difference)를 산출하였다. RMSD 산출 식은 아래와 같다.

$$\sqrt{\frac{\sum(\text{표준화된 } X - \text{표준화된 } Y)^2}{n}} \quad (16)$$

이후 두 능력추정치 차이를 보다 구체적으로 보기 위하여, 각 점수대에 해당하는 같은 점수의 피험자 ID와 IE를 비교하고자 문항 반응패턴을 분석하였다.

그리고 능력추정치와 원점수와의 관련성은 산출된 능력추정치와 원점수간의 관계는 비선형적인 관계이기 때문에 상관인 아닌 순위의 관련성을 보

고자 하였다. 서열 측면에서 차등배점 고려 여부에 따라 두 가지 방법으로 산출된 ID와 IE 그리고 원점수 순위관련성을 보기 위하여 Kendall의 τ 분석을 실시하였다. Kendall의 τ 분석을 통해 ID, IE, 원점수 간의 등위상관을 산출하였다. 차등배점을 Kendall의 τ 분석은 두 변수가 비연속 서열척도일 때 두 변수간의 순위의 일치도를 보기위한 방법이다. 일반적으로 순위의 일치도를 보기위한 분석 방법에는 Spearman의 등위상관계수(ρ)나 Kendall의 타우(τ)분석이 있다. 두 방법의 차이점은 등위상관계수의 경우 순위를 점수처럼 다루어 상관관계를 도출하는 반면에 Kendall의 τ 는 순위의 역전에 기초하여 상관을 다룬다는 점에서 차이가 있다(박산하, 2014). 본 연구에서는 차등배점의 고려여부에 따른 피험자 능력모수 서열이 기존 원점수 서열에서 어느 정도의 순위변동이 일어났는지에 대해 관심 있기 때문에 Kendall의 타우(τ)분석을 실시하였다. Kendall의 τ 를 산출하는 식은 아래와 같다.

$$\tau = 1 - \frac{2(\text{역전횟수})}{\text{총순서쌍}} \quad (17)$$

차등배점을 반영한 원점수(RD)의 서열과 차등배점을 고려한 능력추정방식을 통해 산출된 서열의 관련성을 보다 구체적으로 보기 위해 원점수 서열을 기준으로 두 능력추정치(ID, IE)의 순위변화정도를 비교하고자 하였다. 피험자의 상대적 위치의 변화정도는 다음과 같이 정의하였다.

$$\text{피험자의 상대적 위치 변화정도} = \text{원점수 순위} - \text{능력추정치 순위} \quad (18)$$

피험자 상대적 위치의 변화정도를 50등 간격으로 구분하여 각 구간에 속하는 피험자의 빈도를 비교하였다. 또한 원점수 범위를 상(70점-100점), 중(40점-70점), 하(0점-40점)로 나누어 각 범위에서의 상대적 위치변화 정도를 비교하였다.

IV. 연구 결과

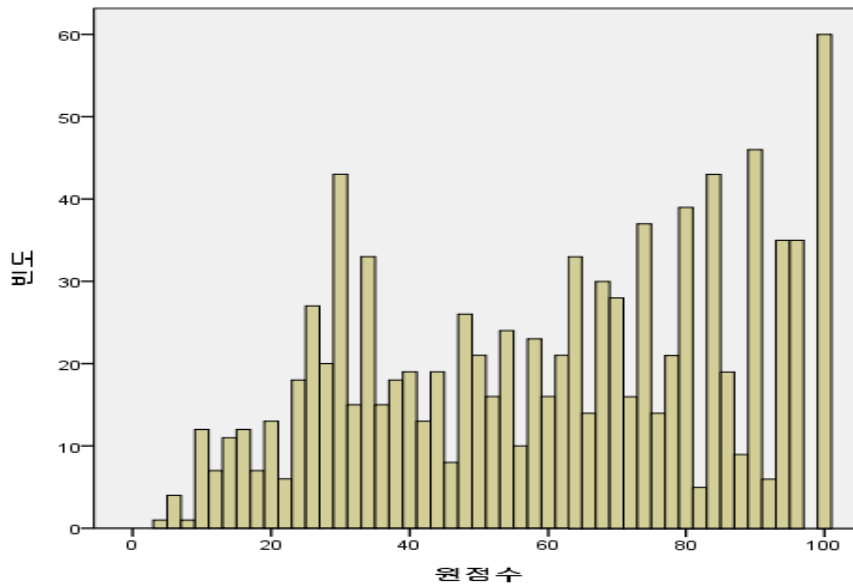
1. 검사의 양호도 및 일차원성 가정 검증

검사점수의 기술 통계치는 아래 <표 IV-1>과 같다.

<표 IV-1>검사의 평균, 표준편차, 최소값, 최대값

	N	평균	표준편차	최소값	최대값
검사점수	965	59.99	26.22	4	100

검사에 참여한 피험자 수는 총 965명이고, 검사점수의 평균은 100점 만점의 59.99점이다. 검사점수의 최소값은 4점, 최대값은 100점이며 표준편차는 26.22로 나타났다. 전체 피험자의 검사점수 분포는 [그림 IV-1]과 같다. 피험자의 검사 점수는 대체로 각 점수대에 고루 퍼져있으며, 50점을 기준으로 50점 이상의 점수대에 속하는 피험자가 50점 미만의 피험자보다 다소 많음을 알 수 있다.



[그림 IV-1] 전체 피험자의 검사점수 분포

검사의 양호도를 보기 위하여 검사 자료의 신뢰도를 산출한 결과 cronbach- α 계수가 .879로 양호한 편이이었다.

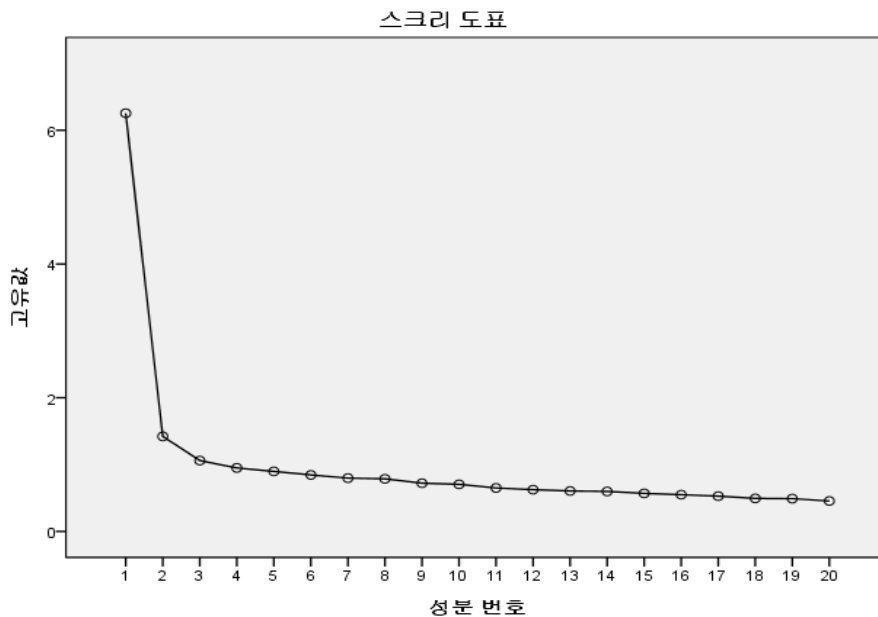
검사 자료가 문항반응이론의 일차원성 가정을 충족하는지에 대한 여부를 알아보기 위해 주성분 분석(principle component analysis)을 실행한 결과는 <표 IV-2>와 같다.

<표 IV-2>요인 설명량

요인	고유값	설명량(%)	누적 설명량(%)
1	6.255	31.27	31.27
2	1.421	7.11	38.38
3	1.060	5.30	43.68

주성분분석 결과 3개의 요인이 추출되었으며, 요인 1의 고유값은 6.255로 전체 변량 중에서 31.27%를 설명하고 있다. 나머지 요인의 고유값은 1.421, 1.060, 설명량은 각각 7.11%, 5.30%로 급격히 감소했음을 확인할 수 있다. [그림 III-1]에서 제시한 스크리 도표(scree plot)를 통해 요인 1의 고유값이 나머지 요인에 비해 상대적으로 매우 크다는 것을 확인할 수 있다.

Reckase(1979)에 의하면 현실적으로 검사자료가 완벽하게 단일차원을 측정하기는 어렵기 때문에 제 1요인이 전체 검사 점수 분산의 20%이상을 설명할 수 있으면 자료가 문항반응이론에서 일차원성가정을 만족한다고 볼 수 있다고 하였다. 따라서 본 연구 분석 자료에서 1요인의 설명량이 31.27%로 20%이상으로 나타나 일차원성 가정을 충족하였다고 볼 수 있다.



[그림 IV-2] 스크리 도표

2. 문항 모수 산출 결과

2모수 로지스틱 모형을 통해 자료의 문항모수를 산출한 결과는 <표IV-3>와 같다. 본 연구에서 활용하는 검사자료는 난이도에 의해 차등배점이 부여되어 전문가의 판단에 의해 난이도가 높다고 판단되는 문항에 더 높은 배점이 부여되었다. 차등배점을 부여한 기준에 부합하게 실제로 배점이 부여되었는지 살펴보기 위해 문항반응이론을 통해 산출한 난이도모수를 배점별로 살펴보았다. 몇 가지 문항의 경우 예외적인 경우가 있었다. 예를 들어 3번문항의 경우 배점은 6점이 부여되었으나, 난이도는 -0.832로 높은 배점이 부여되었음에도 불구하고 난이도는 쉬운 것으로 나타났고, 18번 문항의 경우에는 배점은 4점이 부여되었으나 난이도는 -0.079로 배점에 비해 비교적 어려운 문항이라고 볼 수 있다. 하지만 전체적으로 보았을 때 4점 문항의 난이도모수 평균이 -0.7604, 5점문항은 -0.4027, 6점문항은 -0.0846으로 배점이 높아질수록 난이도모수 평균값이 높아짐을 알 수 있으므로 전반적으로 난이도가 높은 문항에 높은 배점을 부여되었다고 볼 수 있다.

<표Ⅳ-3> 검사문항의 난이도, 변별도 모수

배점	문항번호	난이도(b모수)	변별도(a모수)
4점	1	-1.035	1.112
	2	-1.042	1.092
	7	-0.699	1.498
	11	-0.947	1.294
	18	-0.079	0.786
	평균	-0.7604	1.1564
5점	4	-0.68	1.001
	5	-0.72	1.443
	6	-0.243	1.17
	8	-0.406	1.233
	9	-0.811	0.61
	10	-0.462	0.944
	14	-0.108	1.091
	16	0.043	0.688
	19	0.134	0.362
	20	-0.774	0.833
	평균	-0.4027	0.9375
6점	3	-0.832	0.57
	12	-0.119	0.967
	13	-0.142	1.155
	15	0.715	0.411
	17	-0.045	0.701
	평균	-0.0846	0.7608

3. 두 능력추정치 간 차이 비교

ID 추정치와 IE추정치의 기술통계량은 <표 IV-4>과 같다.

<표 IV-4> 능력추정치의 기술통계량

	N	최소값	최대값	평균	표준편차
ID	965	-2.0573	1.7064	.0065	.9304
IE	965	-1.7427	1.9741	-.0895	.7888

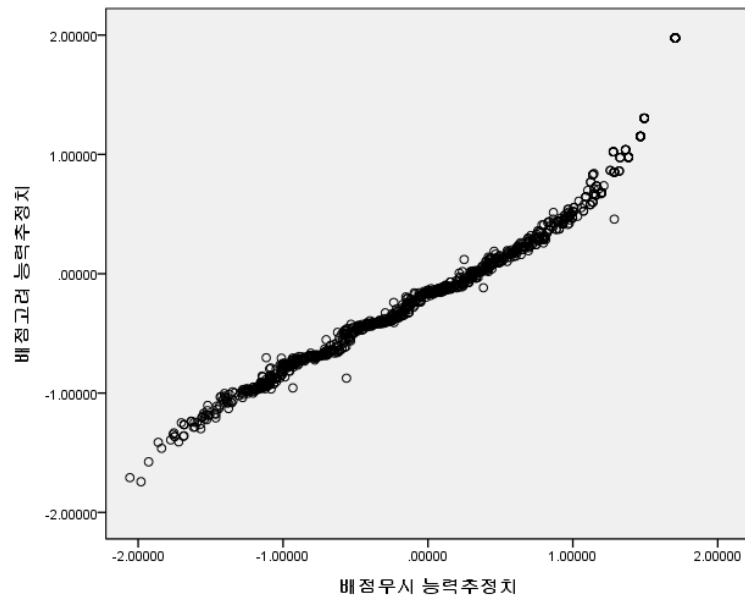
차등배점을 고려한 방법으로 산출된 능력추정치와 차등배점을 고려하지 않은 기존의 IRT 방식으로 산출된 능력추정치간의 결과를 비교하기 위하여 ID 능력추정치, IE 능력추정치, 원점수 간의 상관은 <표 IV-5>과 같다.

<표 IV-5> 능력추정치와 원점수의 상관

	IE	ID	원점수
IE	1.000	0.962 ^{***}	0.981 ^{***}
ID		1.000	0.920 ^{***}
RD			1.000

*** $p < .001$

배점을 고려한 능력추정치와 배점을 고려하지 않은 능력추정치의 상관은 유의수준 .001에서 .962의 매우 높은 상관을 보여, 두 추정치는 거의 유사하게 나타남을 알 수 있다. 두 가지 방법으로 산출한 θ 값의 관계를 도식화하면 [그림 IV-3]과 같다.

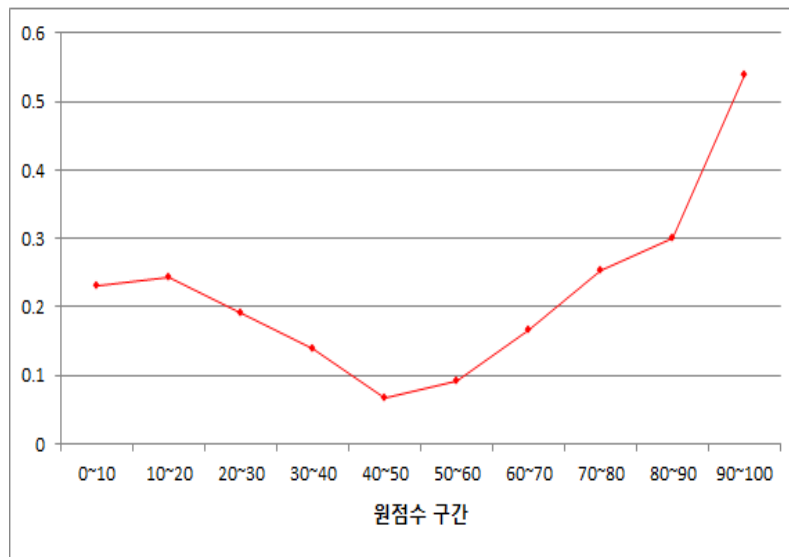


[그림 IV-3] 차등배점 고려 여부에 따른 θ 값 산포도

원점수 구간 별 ID와 IE의 차이를 보기 위해 평균자승편차(Root Mean Square Difference; RMSD)를 산출한 결과는 <표 IV-6>과 같다. RMSD 산출 결과 중간범위(40점-50점)의 원점수에 해당하는 피험자의 능력추정치에 대한 RMSD값이 가장 낮게 나타났으며, 90~100점의 범위에 해당하는 원점수 구간에서 가장 크게 나타났다. 각 원점수 구간별 능력추정치 간의 RMSD값을 도식화하면 [그림 IV-4]와 같다. [그림 IV-4]에서 나타난 바와 같이 원점수의 양 극단으로 갈수록 능력추정치 간 RMSD값이 커짐을 확인할 수 있었다.

<표 IV-6> 원점수 구간 별 능력추정치 간 RMSD

원점수 범위	RMSD
0이상 ~ 10미만	0.231226
10 ~ 20	0.244286
20 ~ 30	0.191523
30 ~ 40	0.140145
40 ~ 50	0.066432
50 ~ 60	0.09126
60 ~ 70	0.165338
70 ~ 80	0.252255
80 ~ 90	0.301493
90 ~ 100	0.53953
전체	0.275717



[그림 IV-4] 원점수 구간별 RMSD

차등배점 고려여부에 따라 능력추정치 값의 차이가 상대적으로 큰 경우에 해당하는 피험자의 반응패턴에는 어떤 차이가 있는지 보기 위해 전체 피험자 965명 중 가장 차이가 크게 나타난 피험자 6명을 선별하여 그들의 문항 반응패턴을 분석한 결과는 <표 IV-7>과 같다.

<표 IV-7> 원점수 구간 별 능력추정치 간 응답패턴

피험자	응답패턴	원점수	표준화 ID	표준화 IE
1	11111111111111111111	100	2.61620	1.82710
2	00100000000010000000	12	-1.47005	-1.83700
3	001000000000100001001	23	-1.15920	-1.51692
4	11111111111101011111	88	.63634	1.03011
5	11111111111011010111	82	.44047	.85008
6	1001111011111010101	65	-.03532	.40413

먼저 피험자 1의 경우 모든 문항에 대해 정답을 맞춘 경우로 이 때 표준화한 능력추정치의 차이가 가장 크게 나타났는데, 차등배점을 고려한 경우에 틀린 개수가 동일하더라도 문항의 배점이 다양해지기 때문에 점수의 폭이 넓어지면서 모든 문항을 맞춘 경우에 표준화된 점수가 차등배점을 고려하기 이전보다 높아진 것으로 해석될 수 있다. 피험자 2의 경우는 6점의 배점이 부여된 3번문항과 13번문항만을 맞춘 경우로 전체 문항 중 2문항만 맞추었지만, 해당 문항의 배점이 가장 높은 문항을 맞추어 ID가 IE에 비해 값이 상당히 커지게 됨을 볼 수 있다. 피험자 3의 경우도 마찬가지로 전체에서 4문항을 맞췄지만 배점이 6점인 문항을 3문항, 배점이 5점인 문항 1문항을 맞추어 전체적으로 배점이 높은 문항만 맞추어 ID가 IE보다 더 높은 능력추정치를 얻게 됨을 확인 할 수 있다. 피험자 4, 5, 6의 경우는 반대의 경우로, 대부분의 문항을 맞추었으나 틀린 문항의 배점이 높아 ID가 IE보다 낮아짐을 나타내는 예이다. 피험자 6의 경우를 대표적으로 보면 전체 틀린

문항이 6개이며, 그 중 6점 문항 5개, 5점 문항 1개로 배점이 높은 문항 위주로 틀리게 되어 ID가 IE에 비해 현저하게 낮게 나타났다.

4. 원점수 순위와의 관련성

차등배점에 대한 고려여부에 따라 능력추정치 간의 상관성이 높더라도 피험자 개개인이 얻는 능력추정치의 상대적 위치는 추정방식에 따라 변하게 된다. 차등배점 고려 여부에 따라 두 가지 상황에서 산출된 능력 모수(IE, ID)의 서열이 원점수 서열을 기준으로 어느 정도의 순위변화가 있는지 보기 위하여 Kendall의 τ 분석을 실시하였다. 결과는 <표 IV-8>과 같다.

<표 IV-8> 순위상관계수

	IE	ID	원점수
IE	1.000	.972 ^{***}	.906 ^{***}
ID		1.000	.922 ^{***}
RD			1.000

*** $p < .001$

순위상관계수를 산출한 결과 원점수에 대해서 IE의 서열과의 상관보다 ID의 서열이 .922로 상관이 더 높게 나타났다. 따라서 학생의 원점수를 통해 매겨진 서열은 ID 서열과 더 유사하다는 것을 알 수 있다.

검사 문항에 차등배점이 부여된 경우, 이를 반영한 원점수의 서열과 차등배점을 고려한 능력추정방식을 통해 산출된 서열의 관련성을 보다 구체적으로 보기 위해 원점수 서열에 대하여 두 능력추정치(IE, ID)의 순위변화정도

를 보고자 하였다. 차등배점이 반영된 원점수 서열을 기준으로 차등배점을 고려한 능력추정방식을 통해 산출된 피험자 순위와 차등배점을 고려하지 않은 방식으로 산출된 피험자 순위의 변화정도를 본 결과는 <표Ⅳ-9>와 같다. 피험자의 상대적 위치의 변화정도(원점수 순위-능력추정치(IE, ID) 순위)는 50등 간격으로 분류하였다.

<표 Ⅳ-9> 차등배점을 고려하지 않은 능력추정에 대한 피험자 순위 변화정도

순위 변화 정도*	빈도(명)	백분율(%)
변화없음	73	7.56
1-10	170	17.62
10-20	146	15.13
20-50	320	33.16
50-100	231	23.94
100-150	23	2.38
150-200	2	0.21
전체	965	100

*순위변화정도=원점수 순위-IE 순위

원점수 순위를 기준으로 차등배점을 고려하지 않은 능력추정방법을 통해 산출한 피험자 순위의 변화 정도는 전체의 7.56%가 순위의 변동이 없었으며, 전체 피험자의 17.62%가 1-10등내에서의 순위변동이 있었다. 가장 많은 피험자가 속한 순위 변화 정도 구간은 50-100등 구간으로 전체 피험자 중 33.16%에 해당하는 피험자가 50등에서 100등 사이로 순위가 변동되었다. 100등 이상 순위의 차이를 보인 피험자는 전체의 2.59%였다.

<표 IV-10> 차등배점을 고려한 능력추정에 대한 피험자 순위 변화정도

순위 변화 정도*	빈도(명)	백분율(%)
변화없음	73	7.56
1-10	181	18.76
10-20	188	19.48
20-50	329	34.09
50-100	185	19.17
100-150	9	0.93
250-200	0	0.00
전체	965	100

*순위변화정도=원점수 순위-ID 순위

차등배점을 고려한 능력추정방법으로 산출한 피험자 순위 결과는 <표 IV-10>과 같다. 순위의 변화가 없는 피험자는 73명으로 전체의 7.56%에 해당하였으며, 이는 앞서 차등배점을 고려하지 않은 결과와 동일한 빈도로 나타났다. 1-10등의 순위변화정도에 속한 피험자는 18.76%였으며, 20-50등 구간에 속한 피험자가 전체의 34.09%로 가장 많은 피험자가 속해있었다. 100등 이상 순위의 차이를 보인 피험자는 전체의 0.93%였다. 앞서 차등배점을 고려하지 않은 결과와 비교해볼 때 20등내에서 순위의 변동을 보인 피험자는 ID에서 45.80%, IE에서 40.31%로 나타났으며, 100등 이상으로 순위의 변동이 크게 나타난 피험자의 빈도는 ID에서 0.93%, IE 2.59%로 IE에서 더 많이 나타났다. 즉, ID를 산출하였을 때 피험자의 순위가 원점수의 순위와 더 근접하게 산출됨을 볼 수 있다.

<표 IV-11>는 각 점수대별 순위변화정도를 보기위해, 원점수를 70점에서 100점은 '상', 40점에서 70점은 '중', 0점에서 40점은 '하' 이렇게 세 개의 구간으로 나누어 각 점수대에서의 두 능력추정치(ID, IE)의 순위변화정도를 나타낸 결과이다.

<표 IV-11> 원점수 구간 별 피험자 순위 변화정도

(단위: 명(%))

원점수 구간		상(70점-100점)		중(40점-70점)		하(0점-40점)	
능력추정치		ID	IE	ID	IE	ID	IE
순위 변화 정도	변화없음	63(16.4)	63(16.4)	1(0.3)	4(1.3)	9(3.2)	6(2.1)
	1-10	59(15.3)	49(12.7)	60(20.1)	52(17.4)	62(22.1)	69(24.6)
	10-20	64(16.6)	47(12.2)	61(20.4)	52(17.4)	63(22.4)	47(16.7)
	20-50	135(35.1)	136(35.3)	107(35.8)	104(34.8)	87(31.0)	80(28.5)
	50-100	59(15.3)	78(20.3)	67(22.4)	77(25.8)	59(21.0)	76(27.1)
	100-150	5(1.3)	11(2.9)	3(1.0)	10(3.3)	1(0.4)	2(0.7)
	250-200	0(0.0)	1(0.3)	0(0.0)	0(0.0)	0(0.0)	1(0.4)
전체		385(100)	385(100)	299(100)	299(100)	281(100)	281(100)

원점수 서열과 두 능력추정방식을 통해 산출된 순위가 일치하는 비율은 각각 '상'에 해당하는 점수대에서 가장 높았으며, 모든 점수대에서 순위변화 정도가 20-50등 사이인 피험자의 비율이 가장 높았다. ID와 IE의 순위변동 결과를 비교해보면 IE의 경우 모든 점수대에서 50등 이상의 순위변화 빈도가 더 높게 나타났다. 즉 모든 점수대에서 ID가 IE에 비해 원점수와의 순위변동의 정도가 더 작게 나타났다.

V. 결론 및 논의

본 연구에서는 문항반응이론에서 차등배점부여방식을 고려한 능력 추정방법을 탐색해야할 필요성을 인지하고 등급반응모형(Graded Response Model: GRM)을 활용하여 각 경계특성곡선의 문항모수를 고정하여 차등배점을 반영한 피험자 능력추정 방법을 제시하였다. 그리고 이 방법을 실제 검사자료에 적용해봄으로써 그 적용가능성을 탐색하고자 차등배점부여방식을 고려하였을 때의 문항반응이론 모형과 기존 동등배점 하에서의 문항반응이론 모형을 비교하였다.

능력모수를 추정하는데 있어서 사전에 부여된 문항의 배점을 고려하여 추정한 능력모수 추정치(ID)와 차등배점을 고려하지 않았을 때의 능력모수 추정치(IE)와의 상관관계를 산출한 결과 두 추정치간의 매우 높은 상관관계를 보여 두 추정치 간의 큰 차이는 나지 않음을 알 수 있었다. 이는 차등배점을 고려한 능력추정방식이 검사자료를 분석하는데 있어서 기존의 IRT방식과 크게 다르지 않은 추정치를 산출한다는 것을 의미한다. 두 추정치의 차이를 세부적으로 확인하기 위하여 원점수의 구간별로 RMSD를 산출한 결과 중간범위(40점-50점)의 해당하는 원점수를 가진 피험자의 능력추정치와 차등배점을 고려하였을 때와 고려하지 않았을 때의 차이가 가장 낮게 산출되었고, 원점수 범위의 양 극단으로 갈수록 두 방법에 의한 능력추정치의 차이가 증가함을 볼 수 있었다. ID와 IE의 차이가 큰 피험자의 문항반응패턴을 분석한 결과, 0점에 가까운 피험자와 만점에 가까운 피험자 집단에서 ID와 IE의 차이가 크게 나타났다. 0점에 가까운 피험자는 맞춘 문항이 높은 배점의 문항인 경우에 같은 문항수의 낮은 배점 문항을 맞춘 다른 피험자보다 더 높은 능력을 지녔다고 간주하여 ID가 IE보다 상대적으로 높게 산출되면서 차이가 크게 나타났으며, 100점에 가까운 피험자의 경우에는 틀린 문항의 배점이

높은 경우에 차등배점을 고려한 능력모수의 값이 고려하지 않은 값에 비해 현저히 낮아짐을 볼 수 있었다. 이를 통해 차등배점을 고려했을 경우에 양극단 점수대에 속하는 피 시험자는 그들이 맞춘 문항 혹은 틀린 문항의 배점으로 인하여 능력모수 값에 영향을 많이 받는 것을 확인할 수 있었다.

산출된 능력추정치와 원점수와의 순위의 관련성을 보기 위하여 ID, IE, 원점수, 간의 등위상관을 산출한 결과 IE에 비해 ID 순위가 원점수 순위와 더 높은 상관을 보였다. ID와 IE의 순위변화정도를 분석한 결과에서도 마찬가지로 IE에 비해 ID에서 순위변화정도가 더 적게 나타났다. 이를 통해 차등배점을 고려하였을 때에 추정된 능력모수의 서열이 차등배점을 고려하지 않은 능력모수의 서열에서보다 순위의 역전이 더 적게 일어나는 것을 알 수 있었다. 위의 결과를 토대로 본 논문의 결론은 다음과 같다.

차등배점을 고려한 피 시험자의 능력추정치와 고려하지 않은 능력추정치 간의 상관 분석 결과 두 추정치 간 상관이 매우 높게 나타나, 차등배점을 고려하였을 때와 고려하지 않았을 때 능력추정치 간에는 큰 차이가 나지 않음을 알 수 있었다. 반면 원점수 구간별 두 능력추정치 간 RMSD와 응답패턴 분석 결과 원점수의 양 극단에서 능력추정치간 차이가 다소 커지는 것으로 나타났는데, 이는 다분문항반응의 모형의 복잡성으로 인하여 능력추정치 분포의 퍼짐정도가 이분모형에 비해 커지게 되어 피 시험자의 능력추정치 범위가 넓어지게 되는 현상에서 그 원인을 추측할 수 있다.

또한 원점수 순위와의 관련성에 관해 분석한 결과 ID와 IE 모두 원점수와의 등위상관 분석 결과 매우 높은 상관을 보였으며, ID와 원점수 순위와의 상관이 IE와 원점수 순위와의 상관보다 다소 높게 산출되었다. 순위변화정도 분석 결과에서도 마찬가지로 ID에서 순위의 역전이 IE보다 덜 일어난 것으로 나타났다. 따라서 원점수 순위와의 관련성은 차등배점을 고려한 능력추정 방식이 차등배점이 부여된 원점수의 순위를 보다 더 잘 반영해줄 수

있음을 시사한다.

이를 통해 검사의 목적에 따라 사전에 문항 배점이 부여된 상황에서 본 연구에서 제안하는 차등배점을 고려한 능력추정방식은 원점수와 높은 관련성과 원점수의 서열을 보다 더 잘 반영해줄 수 있다는 측면에서 문항반응 이론을 통한 차등배점과 고려한 능력모수 추정이 요구되는 상황에서 효과적인 대안으로 고려될 수 있을 것이다.

등급반응모형을 활용하여 차등배점을 고려한 능력모수 추정법을 제시하고 실제 검사자료에 활용될 수 있는 가능성을 보였으나 연구에서 제안한 방법은 특정 상황에서 그 활용에 제한점을 지니고 있다.

본 연구에서 쓰인 연구자료는 차등배점이 4점, 5점, 6점으로 점수의 폭이 넓지 않고 배점이 자연수로 부여되었으나 만일 배점이 소수단위로 부여되는 경우나 점수의 배점이 매우 큰 폭으로 부여되는 경우에 점수 범주의 수를 지정하기에 어려움이 있어 그 활용이 제한될 수 있다. 또 다른 제한점은 본 연구에서 제시한 차등배점을 고려한 능력 추정방법의 검사에 부여된 차등배점이 검사의 목적에 부합하게 제대로 부여되었다고 기본적으로 가정하기 때문에 만일 차등배점을 부여하는 논리가 일관적이지 못하거나 목적과 어긋나게 배점이 부여되었다면 이를 고려한 능력추정치도 잘못 추정될 수 있다. 따라서 능력 추정방법을 적용하기 이전에 차등배점이 본래의 목적에 맞게 제대로 부여되었는지 확인하는 절차가 필요하다.

본 연구방법의 제한점을 해결하고 결과를 뒷받침하기 위해 다음과 같은 후속연구가 필요하다. 먼저, 피험자 수, 검사의 길이, 배점의 종류 등의 조건이 다르게 주어진 상황에서 본 연구에서 제시하는 방법의 적용 결과를 비교할 필요가 있다. 또한 본 연구에서는 한가지 조건내에서 결과를 비교하였으나 다양한 검사 조건하에서 본연구에서 제시하는 방법의 적용가능성을 탐색할 필요가 있다. 뿐만 아니라 일반적으로 차등배점은 본 연구에서 활용한

검사자료와 같이 난이도에 따라 어렵다고 판단되는 문항에 더 높은 배점이 부여되고, 상대적으로 쉬운 문항에는 낮은 배점이 부여되지만, 이 외 검사내용의 중요도나 문항을 푸는데 걸리는 시간 등 다양한 기준을 통해 부여될 수 있는 차등배점을 고려하여 본 연구방법에서 제시하는 능력추정방법의 활용가능성과 결과 비교가 필요하다. 이 외에 문항반응이론 하에 차등배점이 부여된 자료를 차등배점을 고려하여 분석할 수 있는 방법에 대한 보다 더 심층적인 연구가 이어져야 할 것이다.

본 연구는 아직까지 문항반응이론 하에서 문항에 사전에 주어진 차등배점을 고려한 피험자의 능력추정방법이 없는 상황에서 새로운 분석 방법을 제시하고 그 적용가능성을 보였다는 점에서 의의를 지닌다. 기존에 고전검사이론 하에서 검사점수의 총점 즉 원점수에서만 차등배점을 고려하였다면, 본 연구에서 제안하는 방법을 통해 문항반응이론의 틀 내에서 모수를 추정하는 과정에서 문항에 주어진 차등배점을 가중치로서 적용하여 피험자의 능력모수를 추정할 수 있게 되었고, 이는 검사의 성격에 따라 특정 검사의 목적을 반영할 수 있다는 차등배점의 이점과 문항반응이론의 강점을 동시에 살릴 수 있는 분석방법으로 활용될 수 있을 것이다. 따라서 선발이나 혹은 배치, 자격부여 등과 같은 목적으로 실시되는 검사에서 차등배점이 부여된 상황에서 본 연구는 평가가 본래의 취지에 부합한 방향으로 제대로 이루어질 수 있도록 도울 수 있는 효과적인 대안으로 활용되리라 기대하는 바이다.

참 고 문 헌

- 김신영, 노국향(1999). 선택형 검사문항에 대한 차등 배점의 타당성에 관한 연구. **교육평가연구**, 12(2), 235-250.
- 김주훈, 동효관, 송미영, 남민우, 김미영, 최원호, 이재봉, 주지은, 이은경, 김지운(2010). 문항의 배점 결정 요인 및 타당성 분석: 2009학년도 의학교육입문검사를 중심으로. **교육과정평가연구**, 13(2), 197-218.
- 김지현(2000). 대학수학능력시험 차등배점과 문항난이도모수와의 관계 분석. 이화여자대학교 석사학위논문.
- 남명호, 김신영, 남현우, 권순달, 박정, 조지민(2002). 2005학년도 대학수학능력시험 점수 체제 연구. 한국교육과정평가원 연구보고 CAT 2002-32.
- 박산하(2014). 교정문항모수 추정방법에 의한 검사동등화가 피험자 순위변호에 미치는 영향. 이화여자대학교 석사학위 논문.
- 박정(2001). **다분 문항반응이론 모형**. 서울: 교육과학사
- 박찬호, 강태훈(2011). 전문가 판정에 의한 차등 배점을 활용한 제한적 일반화부분점수 모형의 적용. **교육평가연구**, 24(3), 781-797.
- 성태제(1998). **문항반응이론의 개념과 컴퓨터 프로그램의 활용. 교육측정·평가의 새지평**. 서울: 교육과학사
- 성태제(2010). **현대교육평가**. 서울:학지사.
- 송미영(1994). 이분반응모형과 등급반응모형에 의한 문항특성과 피험자 능력 추정의 정확성. 이화여자대학교 석사학위 논문.
- 양길석(2007). 전문가 판단방법에 따른 문항 간 차등 배점의 타당성 연구. **한국교육학연구**, 13(1). 197-219.
- 임인재(1993). **심리측정의 원리**. 서울:교육과학사.

- Baker, F. B. & Kim, S. H. (2004). *Item Response Theory: Parameter Estimation Techniques*. New York, NY: Marcel Dekker, Inc.
- Birnbaum, A. (1968). Some latent trait models and their use in inferring an examinee's ability. In F. M. Lord & M. R. Novick (Eds.), *Statistical theories of mental test scores*(pp. 397-479). Reading MA: MIT press.
- Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. Fort Worth, TX: Holt, Rinehart and Winston.
- Cronbach, L. J.(1977). *Educational psychology*. (3rd ed.). New York: Harcourt Brace Jovanovich.
- Decker, R. L.(1955). A study of three specific problem in the measurement and interpretation of employee attributes. *Psychological Monographs*, 69.
- Ewen, R. B.(1967). Weighting components of job satisfaction. *Journal of Applied Psychology*, 51, 68-73
- Findleym W. G.(1963). Purpose of school testing programs and their efficient development. In W. G. Findeley(Ed.), *Sixty-second yearbook of the National Society for the study of education*, Part II. Chicago, IL:University Chicago Press, 1-27.
- Fitzpatrick, J.L., Sanders, J.R., & Worthen, B. R.(2004). *Program evaluation: Alternative approaches and practical guidelines(3rd Ed.)*. Pearson Education, Inc.
- Fruchter, B.(1962), *Prediction of airman success from responses to items of the kelly Activity Reference Report*. Texas: Lackland Air force Base.
- Guilford, J.P.(1954). *Psychometric Methods*. New York: McGraw-Hill.
- Gulliken, H.(1950). *Theory of Mental Tests*. New York:Wiley

- Lawshe, C. H. & Schucker, R. E.(1959). The relative efficiency of four test weighting methods in multiple prediction. *Educational and Psychological Measurement*, 19, 103-114.
- Lord, F. M.,& Novick, M. R.(1968). *Statistical Theories of Mental Test scores*. Reading, Mass.: Addison-Wesley.
- Reckase, M. D. (1979). Unifactor latent trait models applied to multifactor tests: Results and implications. *Journal of Educational Statistics*, 4, 207-230.
- Samejima, F. (1969). *Estimation of ability using a response pattern of graded scores*. Psychometrika Monograph, No. 17.
- Stalnaker, J. M.(1939). Weighting test items and test-item option : An overview of the analytical and empirical literature. *Educational and Psychological Measurement*, 30, 21-35.
- Thissen, D., & Steinberg, L. (1986). A taxonomy of item response models. *Psychometrika*, 51, 567-577.
- Tyler, R. W.(1949). *Basic Principles of curriculum and instruction*. Chicago IL: The University of Chicago Press.
- Wang, M. W.,& Stanley, J.C.(1970). Differential weighting: A review of methods and empirical studies. *Review of Educational Research*, 4, 663-705
- Wilk, S. S.(1938). Weighting systems for linear of correlated variables when there is no dependent variable. *Psychometrika*, 3, 23-40.

ABSTRACT

An exploratory study on IRT ability parameter estimation considering the differential item weighting in multiple choice items

Kim, Yoon-Ju

Dept. of Education

The graduate school of Sungshin Women's University

This study investigated a method of ability parameter estimation considering Differential Item Weighting(DIW) in multiple choice items. In this study, DIW refers to assigning different points on items based on expert judgement.

The purpose of this study is to develop the ability estimation method considering DIW and to compare with original IRT estimation.

In order to achieve the purpose, this study suggested the new estimation method by using the Graded Response Model. Then compared estimated θ by the new method with estimated θ by original IRT method.

The results of the data analysis is that the degree of difference between θ considering DIW and θ not considering DIW was low and ranking of θ considering DIW is more similar to ranking of the raw score. Therefore the method of IRT ability parameter estimation

considering DIW could be used as an effective alternative in the situation required analysis the test data with differential weighted item.

[부록 1] PARSCALE 위치모수 고정 코드

```
fake.PSL
Modified_GRM
>COMMENTS
  TOTAL ITEM NUMBER 20
  5 four point items, 15 five point items, AND 5 six point items
  GRM
>FILE  DFNAME='Mat_3_score_change1_add.dat', IFNAME='mod.prm',
SAVE ;
>SAVE  PARM='mod_para1.PAR', SCORE='mod_score';
>INPUT  NIDCH=8, NTOTAL=20, NTEST=1, LENGTH=20 ;
(8A1,1x,5A1,T15,10A1,T25,5A1)
>TEST1  TNAME=MATH_FAKE, ITEM=(1(1)20), NBLOCK=20,
INAMES=(IT01, IT02, IT07, IT11, IT18, IT04, IT05, IT06, IT08, IT09,
        IT10, IT14, IT16, IT19, IT20, IT03, IT12, IT13, IT15, IT17);
>BLOCK1  BNAME=1, NITEMS=1, NCAT=5, ORI=(0,1,2,3,4),
MOD=(1,2,3,4,5), REP=5, SKIP;
>BLOCK2  BNAME=2, NITEMS=1, NCAT=6, ORI=(0,1,2,3,4,5),
MOD=(1,2,3,4,5,6), REP=10, SKIP;
>BLOCK3  BNAME=3, NITEMS=1, NCAT=7, ORI=(0,1,2,3,4,5,6),
MOD=(1,2,3,4,5,6,7), REP=5, SKIP;
>CALIB  PARTIAL, LOGISTIC, SCALE=1.7, NQPTS=101,
CYCLE=(500,1,2,2,1,1), FREE=(NOADJUST, NOADJUST),
POSTERIOR,NEWTON=0,
        CRIT=0.001, SPRIOR;
>SCORE  EAP;
```

[부록 2] PRM 파일 예시

```

Modified.PSL

Modified_GRM

MATH_FAK  20  20   7   0   1
   1   1   1   1   1   1   1   1   1   1   1   1   1   1
   1   1   1   1   1   1

GROUP 01
BLOCK      5IT01  1.11150  0.08408 -1.03488  0.05726  0.00000
0.00000
   0.00000 -0.00002 -0.00001  0.00003  0.00000
   0.00000  1.25844  0.81067  0.95293  0.94281
BLOCK      5IT02  1.09166  0.08295 -1.04240  0.05839  0.00000
0.00000
   0.00000 -0.00002 -0.00001  0.00003  0.00000
   0.00000  1.25844  0.81067  0.95293  0.94281
BLOCK      5IT07  1.49780  0.10086 -0.69947  0.03839  0.00000
0.00000
   0.00000 -0.00002 -0.00001  0.00003  0.00000
   0.00000  1.25844  0.81067  0.95293  0.94281
BLOCK      5IT11  1.29422  0.09406 -0.94727  0.04803  0.00000
0.00000
   0.00000 -0.00002 -0.00001  0.00003  0.00000
   0.00000  1.25844  0.81067  0.95293  0.94281
BLOCK      5IT18  0.78627  0.05608 -0.07925  0.05527  0.00000
0.00000
   0.00000 -0.00002 -0.00001  0.00003  0.00000
   0.00000  1.25844  0.81067  0.95293  0.94281
BLOCK      6IT04  1.00064  0.07047 -0.68048  0.05179  0.00000
0.00000
   0.00000 -0.00004 -0.00002  0.00001  0.00005  0.00000
   0.00000  1.24336  0.52868  0.56772  0.56731  0.54729
BLOCK      6IT05  1.44328  0.09775 -0.71973  0.03971  0.00000

```

0.00000						
0.00000	-0.00004	-0.00002	0.00001	0.00005	0.00000	
0.00000	1.24336	0.52868	0.56772	0.56731	0.54729	
BLOCK	6IT06	1.17004	0.07490	-0.24348	0.04152	0.00000
0.00000						
0.00000	-0.00004	-0.00002	0.00001	0.00005	0.00000	
0.00000	1.24336	0.52868	0.56772	0.56731	0.54729	
BLOCK	6IT08	1.23263	0.07965	-0.40604	0.04086	0.00000
0.00000						
0.00000	-0.00004	-0.00002	0.00001	0.00005	0.00000	
0.00000	1.24336	0.52868	0.56772	0.56731	0.54729	
BLOCK	6IT09	0.61018	0.05287	-0.81130	0.08427	0.00000
0.00000						
0.00000	-0.00004	-0.00002	0.00001	0.00005	0.00000	
0.00000	1.24336	0.52868	0.56772	0.56731	0.54729	
BLOCK	6IT10	0.94358	0.06512	-0.46170	0.05046	0.00000
0.00000						
0.00000	-0.00004	-0.00002	0.00001	0.00005	0.00000	
0.00000	1.24336	0.52868	0.56772	0.56731	0.54729	
BLOCK	6IT14	1.09099	0.07018	-0.10767	0.04337	0.00000
0.00000						
0.00000	-0.00004	-0.00002	0.00001	0.00005	0.00000	
0.00000	1.24336	0.52868	0.56772	0.56731	0.54729	
BLOCK	6IT16	0.68773	0.05199	0.04317	0.06173	0.00000
0.00000						
0.00000	-0.00004	-0.00002	0.00001	0.00005	0.00000	
0.00000	1.24336	0.52868	0.56772	0.56731	0.54729	
BLOCK	6IT19	0.36155	0.04187	0.13405	0.10889	0.00000
0.00000						
0.00000	-0.00004	-0.00002	0.00001	0.00005	0.00000	
0.00000	1.24336	0.52868	0.56772	0.56731	0.54729	
BLOCK	6IT20	0.83307	0.06296	-0.77405	0.06290	0.00000
0.00000						
0.00000	-0.00004	-0.00002	0.00001	0.00005	0.00000	

0.00000	1.24336	0.52868	0.56772	0.56731	0.54729	
BLOCK	7IT03	0.57037	0.05133	-0.83210	0.09057	0.00000
0.00000						
0.00000	-0.00003	-0.00002	-0.00001	0.00001	0.00005	0.00000
0.00000	1.50979	1.19927	1.11268	1.10056	1.09806	1.09081
BLOCK	7IT12	0.96677	0.06416	-0.11947	0.04730	0.00000
0.00000						
0.00000	-0.00003	-0.00002	-0.00001	0.00001	0.00005	0.00000
0.00000	1.50979	1.19927	1.11268	1.10056	1.09806	
1.09081						
BLOCK	7IT13	1.15464	0.07358	-0.14166	0.04170	0.00000
0.00000						
0.00000	-0.00003	-0.00002	-0.00001	0.00001	0.00005	0.00000
0.00000	1.50979	1.19927	1.11268	1.10056	1.09806	1.09081
BLOCK	7IT15	0.41081	0.04345	0.71483	0.11694	0.00000
0.00000						
0.00000	-0.00003	-0.00002	-0.00001	0.00001	0.00005	0.00000
0.00000	1.50979	1.19927	1.11268	1.10056	1.09806	1.09081
BLOCK	7IT17	0.70088	0.05263	-0.04545	0.06056	0.00000
0.00000						
0.00000	-0.00003	-0.00002	-0.00001	0.00001	0.00005	0.00000
0.00000	1.50979	1.19927	1.11268	1.10056	1.09806	1.09081