



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

정의적 특성 검사 자료 분석을 위한  
문항반응모형 탐색

2013

성신여자대학교 대학원  
교육학과  
오민아

정의적 특성 검사 자료 분석을 위한  
문항반응모형 탐색

강 태 훈 교수지도

이 논문을 석사학위논문으로 제출함.

2013년 5월

성신여자대학교 대학원

교육학과

오민아

# 인 준 서

오민아의 석사학위 논문으로 인준함.

심사위원 \_\_\_\_\_ 인

심사위원 \_\_\_\_\_ 인

심사위원 \_\_\_\_\_ 인

성신여자대학교 대학원

## 논문개요

본 논문에서는, 다분문항반응 자료 분석을 위하여 흔히 활용되는 일반화 부분점수모형과 정의적 특성 검사 자료 분석에 대한 효용성의 측면에서 최근 부각되고 있는 일반화등급전개모형의 적용가능성을 상호 비교하고자 하였다. 이를 위한 연구문제는 다음과 같다.

첫째, 실제 정의적 특성 검사 자료를 분석함에 있어서 각각 전개기제와 누적기제를 따르는 두 모형을 적용하여 문항 분석을 시도할 경우, 어떠한 해석상의 차이가 존재하는가?

둘째, 두 모형은 전개기제를 따르거나 누적기제를 따르는 다분문항자료를 분석함에 있어서 자료에 대한 효율적 설명력과 능력모수에 대한 복원력에 있어서 분명한 차별성을 가진다고 볼 수 있는가?

연구는 실제자료에 대한 분석과 모의실험자료에 대한 분석의 절차로 수행되었다. Likert 4점 척도에 기반하여 수집된 다분문항으로 이루어진 청소년의 인터넷 게임중독 진단 검사 자료를 실제 데이터로 활용한 실제자료 분석 결과, DIC에 의하면 일반화등급전개모형이 일반화부분점수모형보다 효율적으로 자료를 잘 설명하고 있었다. 일반화등급전개모형에 의하여 문항해석을 시도한 결과, 피험자의 특성 수준이 일정 수준 이상일 경우 긍정적인 응답 확률이 낮아지는 상황을 고려할 수 있었고, 몇몇의 문항은 누적기제를 보이기도 하였다.

두 모형을 생성모수로 한 모의 실험 데이터 역시, 4개의 범주로 이루어진 다분문항반응자료였으며, 두 모형을 추정모형으로 적용하여 피험자 특성 복원연구를 수행하였다.

생성모형이 추정모형과 일치할 때 피험자 특성 복원이 가장 잘 이루어졌

고, 이는 실제 데이터가 따르는 기제를 고려하여 문항분석 모형을 선택해야 함을 의미하였다. 하지만 DIC값은 일반화등급전개모형이 생성모형과 관계없이 모든 데이터를 효율적으로 설명한다고 나타나, 또 다른 모형선택지수에 관한 후속연구의 필요성이 제기되었다.

# 목 차

|                                 |    |
|---------------------------------|----|
| <b>I. 서론</b> .....              | 1  |
| 1. 연구의 필요성 및 목적 .....           | 1  |
| 2. 연구문제 .....                   | 5  |
| <b>II. 이론적 배경</b> .....         | 6  |
| 1. 고전검사이론에 의한 정의적 특성 검사 .....   | 8  |
| 1) Thurstone 척도 .....           | 8  |
| 2) Likert 척도 .....              | 15 |
| 3) Guttman 척도 .....             | 19 |
| 2. 다분문항반응이론에 의한 정의적 특성 검사 ..... | 23 |
| 1) 일반화부분점수모형 .....              | 27 |
| 2) 일반화등급전개모형 .....              | 31 |
| <b>III. 연구방법</b> .....          | 36 |
| 1. 실제 자료 분석 .....               | 36 |
| 1) 자료의 구성 .....                 | 36 |
| 2) 자료의 분석 .....                 | 41 |
| 2. 모의실험을 통한 자료 분석 .....         | 43 |
| 1) 모의 실험 설계 .....               | 43 |
| 2) 피험자 모수 추정 및 모형선택지수 .....     | 45 |

|   |    |
|---|----|
| <b>IV. 연구결과</b> .....                           | 46 |
| 1. 실제 자료 분석 결과 .....                            | 46 |
| 1) 검사도구의 타당도와 신뢰도 분석 및 기술통계 .....               | 46 |
| 2) 일반화등급전개모형과 일반화부분점수모형을 통한 문항 분석 및 문항 해석 ..... | 51 |
| 2. 모의실험에 의한 자료 분석 결과 .....                      | 56 |
| 1) 피험자 모수 추정치 .....                             | 56 |
| 2) 모형선택지수 .....                                 | 58 |
| <br>  |    |
| <b>V. 논의 및 결론</b> .....                         | 61 |
| 1. 전개기체에 기반한 정의적 특성 검사 문항의 해석 .....             | 61 |
| 2. 일반화등급전개모형의 정의적 특성 검사의 다분문항자료에 대한 설명력 .....   | 62 |

참고문헌

Abstract

부    록

## 표 목 차

|   |    |
|---|----|
| <표1> 사형제도에 관한 Thurstone척도 검사 도구 .....                 | 10 |
| <표2> 두 가지 진술문에 대한 평정자들의 표준점수행렬표 .....                 | 12 |
| <표3> 청소년의 부모관련 심리건강상태에 관한 Likert척도 검사 도구 .....        | 17 |
| <표4> Maslow의 욕구위계설 검증을 위한 Guttman척도 .....             | 21 |
| <표5> 이분문항에 관한 Guttman척도의 이상적 반응유형 .....               | 22 |
| <표6> 게임중독 진단 검사 응답자의 일반적 특성 .....                     | 38 |
| <표7> 온라인게임 및 가족여가활동실태조사의 구성 내용 .....                  | 39 |
| <표8> 청소년용 인터넷 게임중독 진단 Likert척도 검사 도구 .....            | 40 |
| <표9> 모의실험 조건 .....                                    | 44 |
| <표10> 게임중독 진단 응답자료의 주성분분석에 따른 요인별 누적분산비율 .....        | 47 |
| <표11> 게임중독 진단 검사의 요인간 상관행렬 .....                      | 49 |
| <표12> 게임중독 진단 검사의 요인분석 결과 .....                       | 49 |
| <표13> 게임중독 진단 검사 응답자의 성별 및 학교수준에 따른 기술통계 .....        | 50 |
| <표14> 일반화등급전개모형과 일반화부분점수모형에 따른 문항모수치 및 DIC .....      | 49 |
| <표15> 일반화등급전개모형에 의한 자료 생성시 피험자 특성 생성모수에 대한 상관계수 ..... | 56 |
| <표16> 일반화등급전개모형에 의한 자료 생성시 원점수에 대한 상관계수 .....         | 57 |
| <표17> 일반화부분점수모형에 의한 자료 생성시 피험자 특성 생성모수에 대한 상관계수 ..... | 49 |
| <표18> 일반화부분점수모형에 의한 자료 생성시 원점수에 대한 상관계수 .....         | 58 |
| <표19> N=500일 경우의 DIC .....                            | 59 |
| <표20> N=1000일 경우의 DIC .....                           | 60 |

## 그 립 목 차

|   |    |
|---|----|
| <그림1>전쟁과 평화에 관한 진술문 위치와 피험자 모집단 태도 분포 .....                           | 13 |
| <그림2> Thurstone척도 문항의 문항특성곡선 .....                                    | 14 |
| <그림3> Likert척도 문항의 문항특성곡선 .....                                       | 16 |
| <그림4> Guttman척도 문항의 문항특성곡선 .....                                      | 22 |
| <그림5> 전개기제와 누적기제 하에서의 문항특성곡선 .....                                    | 26 |
| <그림6> 일반화부분점수모형의 문항범주특성곡선과 기대반응곡선 .....                               | 30 |
| <그림7> 관찰가능범주와 주관적반응 범주 .....  | 32 |
| <그림8> 네 개 범주 문항의 주관적반응범주 반응곡선 .....                                   | 33 |
| <그림9> 네 개 범주 문항의 관찰가능범주 반응곡선 .....                                    | 35 |
| <그림10> 네 개 범주 문항의 반응기대값곡선 .....                                       | 35 |
| <그림11> 게임중독 진단 응답자료의 주성분 분석에 따른 스크리 도표 .....                          | 48 |
| <그림12> 문항8“갈수록 게임을 하는 시간이 길어진다”의 GGUM에 관한 ICCC와 REVC .....            | 53 |
| <그림13> 문항12“게임을 안하겠다고 마음먹고도 다시 게임을 하게 된다”의 GGUM에 관한 ICCC와 REVC .....  | 53 |
| <그림14> 문항10“게임을 그만두어야 하는 경우에도 그만두는 것이 어렵다”의 GGUM에 관한 ICCC와 REVC ..... | 54 |
| <그림15> 문항20“게임을 못하게 되면 화가 난다”의 GGUM에 관한 ICCC와 REVC .....              | 55 |
| <그림16> 문항3“게임 속의 내가 실제의 나보다 더 좋다”의 GPCM에 관한 ICCC와 REVC .....          | 55 |
| <그림17> 가상의 동일한 문항에 대한 GPCM과 GGUM의 REVC .....                          | 64 |

# I. 서 론

## 1. 연구의 필요성 및 목적

가족공동체와 지역공동체의 유기적 협업 관계에서 개인의 삶이 영위되었던 전통사회와 달리, 복잡성이 증가한 현대사회는 보다 전문화된 분업의 형태로 유지되고 있는 바, 양육과 교육, 질병관리, 여가 등 이전에는 가족 내에서 관리했던 삶의 제 부분을 각 영역을 담당하는 전문기관 혹은 전문가에게 의탁하고 있다. 이에 따라 분업화된 영역을 담당한 개인에게 사회는 보다 심화된 전문성을 요구하게 되었고, 개인의 적성과 흥미가 그에게 요구되는 전문성 신장에 부합하는지 여부도 중요한 관심 사안이 되었다. 이러한 사회적 변화 양상에 따라 직업세계와 같이, 개인이 담당할 구체적인 분야에 관한 전문성과 더불어 개인의 흥미와 적성과 같은 정의적 영역에 관한 평가가 요구되어 지는 바, 다양한 제 분야의 검사와 평가가 발전되어 왔다.

검사란 피험자의 보이지 않는 잠재적 특성에 관한 정보를 간접적으로 얻기 위한 도구로서, 잠재적 특성은 크게 지적능력과 같은 인지적 특성(cognitive domain)과 태도, 의견, 흥미와 같은 정의적 특성(affective domain)으로 나뉠 수 있다(성태제, 2001). 전자의 경우 학업성취도를 검사구인의 대표적인 예로 들 수 있고 후자의 경우 태도, 성격, 의견, 자아개념, 그 이외의 비인지적 특성들(non-cognitive traits)을 검사구인의 예로 들 수 있으며, 특히 태도와 같은 정의적 특성 평가의 경우, 사회심리학 분야에서 기인하였다(Anastasi & Urbina, 1997).

검사결과가 활용되는 대표적인 사례로서, 학교 현장에서 교사는 다양한

검사를 실시하고 그 결과를 해석하여 학교, 학생, 학부모 등에게 보고하며 (황정규 외, 2012) 이는 상급학교 진학의 주요한 근거자료로 작용한다. 국가 수준의 대규모 학업성취도와 같은 검사결과는 잠재적인 국가경쟁력을 대표하는 하나의 지표로 활용되며 국가 및 다양한 단체와 기관에서는 자체적으로 실시한 검사에 기반하여 사전에 설정한 준거점수에 부합한 피험자들에게 특정 자격을 부여하고, 기업체는 개인의 적성검사, 직업기초능력검사 등에 기반하여 인재를 선발한다. 이처럼 검사결과에 기반한 수많은 의사결정이 이루어지고 있는 가운데, 피험자의 잠재적 특성과 능력을 타당하고 신뢰롭게 구현할 수 있는 검사를 개발하고 관리하는 것의 중요성은 날로 더해지고 있다.

평가목적에 부합한 적절한 검사를 개발하고, 각 검사를 구성하는 문항을 분석하는 데에 활용되는 검사이론은 크게 고전검사이론(classical test theory)과 문항반응이론(item response theory)으로 나뉜다. 고전검사이론이 검사를 분석의 기본단위로 상정하는 것과 달리 문항반응이론은 검사를 구성하는 문항을 분석의 기본단위로 하며, 피험자 특성 수준과 해당 문항에 대하여 특정 점수를 획득할 확률 간의 관계를 나타내는 문항특성곡선(item characteristic curve)을 바탕으로 검사에 관한 연구를 풍부하게 확장한 이점이 있다. 한 가지 예로서, 단순한 검사 총점으로 피험자의 특성 수준을 나타내는 고전검사이론과 달리 문항반응이론은 각 문항의 특성과 이에 대한 피험자의 반응행태를 반영한 피험자 특성 추정치를 산출해 낼 수 있다 (Bryce, 2002). 동일한 검사에 대한 피험자의 반응 행태가 피험자 특성 추정치로 나타날 경우, 고전검사이론에서는 단순 총점의 방식으로 나타나지만, 문항반응이론에 의하면 해당 검사의 특성과 검사가 수행된 배경 및 피험자의 반응행태에 부합한 구체화된 모형에 의하여 각기 다른 특성 추정치를 산

출해낼 수 있다. 따라서 피험자의 반응을 적절하고 정확하게 분석하는 것이 문항반응이론의 주요한 연구 분야이기도 하다(임미경, 2001).

문항반응이론이 활용하는 피험자 반응에 관한 가정 중 한 가지는, 피험자의 특성수준이 높아질수록 검사 문항을 맞히거나 이에 대하여 긍정적으로 응답할 확률이 높아진다는 우세반응모형(dominant response model)이다. 이때 인지적 특성 검사의 경우, 피험자의 지적 능력이 높을수록 문항을 맞힐 확률이 높아진다는 해석이 가능하며, 정의적 특성 검사의 경우, 검사가 측정하고자 하는 정의적 구인의 수준이 높은 피험자일수록 문항의 진술문에 긍정적으로 응답할 확률이 높아진다는 해석이 가능해진다.

하지만 최근 정의적 특성 검사의 경우, 피험자는 문항이 측정하고자 하는 구인에 대한 자신의 수준과, 문항의 진술문이 나타내는 구인의 수준을 비교하여 두 수준이 부합할수록 긍정적으로 응답하되, 자신의 의견 및 태도와 진술문이 나타내는 의견 및 태도가 다를수록 이에 긍정하지 않는다는 이상점모형(ideal point model)이 대안적 가정이 될 수 있다는 주장이 제기되고, 실제로 이상점모형에 근거한 피험자 특성을 추정한 분석적 연구가 수행되었다(Andrich, 1988; Andrich, 1989; Andrich & Luo, 1993; Andrich, 1995; Andrich, 1996; Carter, 2010; Chernyshenko, Stark, Drasgow & Williams, 2007; Drasgow, Chernyshenko & Stark, 2010; Javaras, 2004; Kantrowitz & Tusinski, 2010; Roberts, 1995; Roberts, Laughlin & Donoghue, 1998; Roberts, Donoghue & Laughlin, 2000; Roberts, Lin & Laughlin, 2001; Sgammato, 2009; Stark, 2002; Stark, Chernyshenko, Drasgow & Williams, 2006, 강태훈, 김동일, 2010; 권재웅, 2012; 이선희, 2008; 박동건, 최명옥, 이원선, 이혜민, 이준희, 2012). 이상점 반응 과정이란 피험자의 특성수준이 높아짐에 따라 긍정적으로 응답할 확률이 함께 높아지는 누적기제가 아닌, 피

험자의 특성수준과 문항의 위치가 가까워질수록 긍정적으로 응답할 확률이 높아지는, 꺾임현상이 존재하는 전개기제를 가정한다.

강태훈, 김동일(2010)과 Drasgow et al.(2010)에 의하면 대부분의 문항반응이론가들은 검사가 측정하고자 하는 피험자의 특성 수준이 높을수록 해당 구인에 관하여 긍정적으로 진술된 문항에 높은 점수의 범주를 선택한다는 우세 반응 과정을 가정한 누적기제의 문항반응모형을 활용하였으나 이러한 가정은 인지적 특성 검사의 분석에 적합할 수 있으며, 비인지적 특성을 측정하는 검사의 경우 다른 기제를 가정한 문항반응모형을 적용하는 것이 바람직할 수 있다. Roberts et al.(1998)은 각 문항의 문항특성곡선(item characteristic curve; ICC)을 통하여 실제로 검사 문항이 누적기제와 전개기제 중 어떠한 심리적 기제를 따르는지 명확히 구분하여 정의적 특성에 관한 연구를 수행해야 한다고 하였으며, Thurstone척도나 이에 대응하는 전개기제의 문항반응모형을 활용할 것을 제안하였다.

이와 관련한 분석적 연구결과를 살펴보면, 우세반응 모형보다 전개 기제를 따르는 이상점 모형이 보다 타당한 검사를 개발하고 피험자의 특성을 추정한다는 연구와(Andrich, 1988; Roberts, 1995; Roberts et al., 2000) 실제로 우세반응 모형과 이상점 모형의 적용 가능성을 탐색한 국내의 연구들(강태훈, 김동일, 2010; 권재웅, 2012; 박동건 외, 2012; 이선희, 2008)이 존재하고 있다.

Drasgow et al.(2010)이 앞으로 30년 이내에는 이상점 반응에 기반한 문항반응모형이 정의적 특성 검사를 개발하는데 활용될 것이라고 주장하는 가운데, 보다 적절한 측정모형을 통하여 다양한 연구에 활용되는 검사 도구의 질적 수준을 향상시키고, 자료 해석의 타당성을 높일 수 있다고 예상되는 바, 이상점 반응에 기반한 전개기제 문항반응모형의 활용 가능성을 탐색하

고 적용할 필요성이 제기된다. 특히 이상점모형이 문항모수와 피험자 특성을 추정할 때에, 우세반응모형 만큼의 우수한 추정치를 보여줄 수 있는지에 관한 모의실험 연구를 통하여, 우세반응모형에 대한 대안으로 이상점모형이 활용될 수 있는지 탐색해야 할 것이다(Brown, Maydeu-Olivares, 2010).

## 2. 연구문제

본 연구는 다분문항반응의 누적기제 모형으로서 일반적으로 활용되는 일반화부분점수모형과 전개기제 모형의 일반화등급전개모형이 갖는 정의적 특성 검사 데이터에 관한 설명력을 비교함으로써 정의적 특성 검사 데이터 분석에 전개기제의 활용가능성을 탐색해 보고자 하였다.

연구는 크게 두 가지 측면으로 구성되어 수행되었다. 첫째, Likert척도에 기반하여 수집된 피험자의 정의적 영역 특성 검사 자료를 기반으로 두 모형을 적용하여 문항을 분석하였다. 둘째, 보다 다양한 형태의 정의적 특성 검사 데이터에 전개기제를 적용할 수 있는지 탐색하기 위하여 실제 데이터 및 랜덤 샘플링의 문항모수를 생성 모수로 설정한 모의실험 연구를 수행하였다. 이를 위한 연구문제는 다음과 같다.

첫째, 실제 정의적 특성 검사 자료를 분석함에 있어서 각각 전개기제와 누적기제를 따르는 두 모형을 적용하여 문항 분석을 시도할 경우, 어떠한 해석상의 차이가 존재하는가?

둘째, 두 모형은 전개기제를 따르거나 누적기제를 따르는 다분문항자료를 분석함에 있어서 자료에 대한 효율적 설명력과 능력모수에 대한 복원력에 있어서 분명한 차별성을 가진다고 볼 수 있는가?

## Ⅱ. 이론적 배경

모든 학문 영역에서의 ‘측정’은 해당 분야의 기초로서 측정을 통하여 연구자의 관심 대상에 대한 객관적인 측정이 이루어질 때 보다 합리적인 연구가 수행될 수 있다(백순근, 2010). 교육측정에서는 개인의 지능, 학업성취도와 같은 인지적 특성과 태도, 의견, 만족도 등과 같은 정의적 특성을 수량화하는 것이 주요 연구 분야이며, 수량화된 연구 결과를 토대로 각종 교육학의 실제적인 타 분야의 이론적 적합성이 가시적 수준에서 논의될 수 있을 것으로 사료된다. 이와 같은 교육학의 기초 분야인 교육측정에서 피험자의 특성을 측정하는 도구가 검사이며, 도구의 종류는 측정구인의 특성이나 유형에 따라 달라진다(백순근, 김미림, 2010).

정의적 특성과 관련한 검사이론은 고전검사이론과 문항반응이론으로 크게 나뉘는 가운데, 정의적 특성 검사의 시작은 동일한 특성 수준을 가진 피험자에게 동일한 수치를 부여하는 고전검사이론의 척도화 방법에서 기인하였다(김명소, 1992). 그 중 비교적 수월한 개발과 활용이 가능한 Likert척도(1932)가 광범위하게 사용되고 있으며, 그 예로 ‘매우 아니다’, ‘아니다’, ‘보통이다’, ‘그렇다’, ‘매우 그렇다’와 같은 5점 척도, 혹은 이보다 더 적거나 많은 수의 범주로 이루어진 설문조사 문항의 응답 형식을 고려해볼 수 있다. 설문조사는 각종 이슈에 관한 대중의견을 조사하거나, 정책수립의 근거 자료 확보를 위한 전국민 실태조사 혹은 표본조사 등을 위하여 공공부문은 물론, 각종 기업, 단체, 개인 연구자에 의하여 활용되고 있다. 이처럼 빈번하게 활용되는 정의적 특성 검사에 따른 결과가 보다 타당하고 신뢰롭게 피험자들의 진(true) 특성을 반영할 때, 이를 기반으로 형성된 정책이나 연구결과

등의 의미 역시 타당하고 신뢰롭다고 볼 수 있을 것이다. 이러한 관점에서 피험자들의 특성을 타당하게 구현하고자, 문항반응이론은 Likert 척도에 따른 피험자 응답 데이터와 정의적 특성 간의 관계를 고려하여 정의적 특성 검사에 대한 피험자들의 응답을 분석한다(Javaras, 2004).

정의적 특성에 관한 검사이론이 비록 고전검사이론과 문항반응이론으로 나뉘지만, 고전검사이론에서 발달해 온 척도화 방법에 기반한 피험자 반응을 토대로 문항반응이론은 문항분석을 수행하며, 누적기제와 전개기제와 같은 문항특성곡선의 개념 역시 척도화 방법에서 고려되었던 개념으로, 두 검사이론의 맥락을 완벽히 구분하여 살펴보는 것은 무의미하다. 또한 Fan(1998)의 고전검사이론과 문항반응이론의 문항 분석 결과가 매우 유사하다는 연구결과와 Stage(2003)의 문항반응모형과 비교하여 고전검사 모형이 보다 적절히 검사 데이터를 분석하였다는 연구결과가 제시하는 바와 같이, 문항반응이론이 반드시 고전검사이론보다 우수하다는 일률적 평가를 내릴 수는 없다. 하지만 문항반응이론은 문항반응 함수의 시각화를 통하여 고전검사이론보다 문항에 관한 세분화된 정보를 제공하고 대량의 데이터와 컴퓨터에 기반한 문항분석을 토대로 보다 현실적인 이론모형을 구축하였으며, 고전검사이론의 Likert척도가 Stevens(1946)의 척도구분에 따라 엄밀히 서열척도에 지나지 않지만, 문항반응이론의 동간척도를 지닌 피험자 특성 추정치를 산출하는 이점이 있다.

이에 따라 본 연구의 목적이 서로 다른 문항반응모형에 의한 정의적 특성 검사의 가능성을 탐색하는 것이지만, 고전검사이론의 척도화방법의 논리와 형식을 우선적으로 살펴보고, 이에 근거한 문항반응이론의 논리를 살펴보고자 한다.

## 1. 고전검사이론에 의한 정의적 특성 검사

20세기 초, Spearman이 제안한 기본가정인 ‘관찰점수=진점수+오차’의 기본 전제를 따르는 고전검사이론은 진점수를 추정하기 위하여 현실에서 불가능한 반복 측정을 가정한다. 분석의 기본단위를 ‘검사’로 상정하는 고전검사이론은, 피험자의 검사총점을 기반으로 검사분석을 수행하기 때문에 피험자의 능력 수준에 따라 검사총점이 달라짐에 따라 검사 특성 또한 변동되는 단점을 가지고 있다. 또한 같은 수준의 특성을 지닌 피험자라 하더라도 곤란도가 높은 검사에 대하여 낮은 검사총점을 갖고, 곤란도가 낮은 검사에 대하여 높은 검사총점을 가짐으로써 결과적으로 피험자 특성이 검사 특성에 종속되게 된다. 측정의 오차와 검사의 신뢰도가 모든 수준의 피험자에게 동일하다는 비현실적인 가정 또한 고전검사이론의 제한점이다. 그럼에도 불구하고 고전검사이론은 간단한 문항분석 절차와 검사분석 등의 장점을 가지고 있기 때문에 현재까지 국내의 측정분야에서 활발히 사용되고 있다(성태제, 2001). 정의적 특성을 측정하는 대표적인 척도로서, 높은 신뢰도와 타당도를 지닌 Thurstone척도, Likert척도, Guttman척도가 존재한다(황은희, 2010).

### 1) Thurstone 척도

#### ① 척도의 구성

Thurstone척도는 물리적 자극에 대한 감각반응을 측정하는 것과 같이, 인간의 정의적 특성 역시 수량적으로 측정이 가능하다는 가정 하에 개발된 척도방법이다. 즉 감각반응을 일으키는 물리적 자극에 강도가 있듯이 피험자

의 호의적, 중립적, 비호의적 등의 태도나 의견을 불러일으키는 언어적 자극에도 강도가 있다는 가정 하에, 언어적 자극의 강도를 측정하여 진술문과 각 진술문에 대한 자극의 크기로 이루어진 검사도구를 마련하여 피험자의 태도를 측정하고자 하였다(Thurstone, 1927). 태도측정에 있어서 자극의 역할을 하는 것은 측정하고자 하는 하나의 속성에 대한 문항의 진술문으로, 척도는 일차원성이 가정된 문항들로 구성되며, 반응에 해당하는 것은 문항에 대한 응답자의 동의여부이다. 피험자는 자신의 태도와 유사한 진술문에 동의를 표하게 된다.

<표1>은 사형제도에 대한 찬반여부를 묻는 Thurstone 척도로서, 피험자는 자신이 찬성하는 문항에 표시하고 표시한 문항의 척도점수의 평균을 자신의 태도값으로 얻게 된다(Peterson & Thurstone, 1933)

<표1> 사형제도에 관한 Thurstone척도 검사 도구

| 문항<br>번호 | 진술문   | 척도<br>점수 | 찬성<br>여부                 |
|----------|---|----------|--------------------------|
| 1        | 사형제도가 옳지 않을 수도 있지만 범죄를 예방하는데 가장 좋은 방법이다.      | 7.2      | <input type="checkbox"/> |
| 2        | 사형제도는 결코 정당화 될 수 없다.                          | 0.0      | <input type="checkbox"/> |
| 3        | 사형제도는 필요한 제도이긴 하나, 발생하지 않기를 바란다.              | 6.2      | <input type="checkbox"/> |
| 4        | 성별, 연령 등을 떠나 누구든 살인을 했다면, 그도 사형되어야 한다.        | 10.4     | <input type="checkbox"/> |
| 5        | 사형제도는 범죄를 다루는 정상적인 방법으로 볼 수 없다.               | 2.4      | <input type="checkbox"/> |
| 6        | 사형제도는 옳지 않지만, 우리의 불완전한 사회에 필요한 제도이다.          | 6.2      | <input type="checkbox"/> |
| 7        | 모든 범죄자는 사형되어야 한다.                             | 11.0     | <input type="checkbox"/> |
| 8        | 사형제도는 범죄 예방에 결코 효과적이지 않았다.                    | 2.7      | <input type="checkbox"/> |
| 9        | 사형제도가 옳다고 생각하지는 않지만, 필요없다고 생각하지도 않는다.         | 5.4      | <input type="checkbox"/> |
| 10       | 몇몇의 범죄자에 대하여 사형제도가 반드시 필요한 경우가 있다.            | 8.5      | <input type="checkbox"/> |
| 11       | 사형제도보다 태형을 부활시키는 것이 더 효과적이라고 생각한다.            | 3.9      | <input type="checkbox"/> |
| 12       | 어떤 상황에서도 사형제도는 옳지 않다고 생각한다.                   | 0.1      | <input type="checkbox"/> |
| 13       | 현대 사회에서 사형제도는 필요하지 않다.                        | 3.0      | <input type="checkbox"/> |
| 14       | 사형제도가 존재하는 한, 우리 사회는 문명화되었다고 볼 수 없다.          | 1.5      | <input type="checkbox"/> |
| 15       | 중신형이 사형제도보다 더 효과적이다.                          | 3.4      | <input type="checkbox"/> |
| 16       | 범죄자를 사형하는 것은 문명화된 사회의 수치다.                    | 0.9      | <input type="checkbox"/> |
| 17       | 사형제도는 정당하고 반드시 필요한 제도이다.                      | 9.6      | <input type="checkbox"/> |
| 18       | 사형제도가 옳다고 생각하진 않지만, 실제로 그것을 폐지하는 것이 바람직하진 않다. | 5.8      | <input type="checkbox"/> |
| 19       | 사형제도는 우리 시대의 가장 치욕적인 관습이다.                    | 0.6      | <input type="checkbox"/> |
| 20       | 사형은 범죄자에게 그의 행동의 대가를 지불하는 제도이다.               | 9.4      | <input type="checkbox"/> |
| 21       | 사형제도를 집행하며, 생명의 신성불가침을 가르칠 수는 없다.             | 2.0      | <input type="checkbox"/> |
| 22       | 사형제도의 집행 여부에 대하여 큰 차이를 못 느낀다.                 | 5.5      | <input type="checkbox"/> |
| 23       | 사형제도는 사전에 계획하여 살인을 저지른 범죄자에게만 정당화 될 수 있다.     | 7.9      | <input type="checkbox"/> |
| 24       | 사형제도는 보다 더 자주 집행될 필요가 있다.                     | 9.1      | <input type="checkbox"/> |

출처: Peterson & Thurstone(1933)

## ② 척도의 제작

Thurstone 척도제작의 절차는 일반적으로 예비 진술문을 수집·작성하고 진술문을 분류(평정)하며 각 진술문의 척도치를 계산한 후, 최종 진술문을 선택하는 절차로 요약할 수 있다. Thurstone 척도는 자극에 대한 응답치의 차이는 근본적으로 피험자의 차이가 아닌, 자극간의 차이에서 나타난다고 보는 자극중심방법의 척도화 방법(박도순, 2006)이기 때문에, 응답자의 태도 척도치를 산출하기에 앞서 문항에 대한 척도치를 설정하는 작업이 선행되어야 한다.

개인의 진술문에 대한 변별반응은 일차원의 특성 연속성 상에서 하나의 값을 가지게 되고, 동일한 평정자에게 동일한 자극, 즉 문항을 반복하여 제시하면 그 자극에 대한 많은 수의 변별반응이 생긴다(Thurstone, 1928). 결과적으로 변별반응의 빈도수는 정규분포를 이루며, 해당 분포의 평균이 문항의 최종 척도치로 산출된다. 하지만 실제로 한 평정자가 동일한 문항을 연속적으로 판단할 경우, 피로나 연습효과가 생길 수 있기 때문에 여러 명의 평정자가 동일한 문항을 판단한 결과도 역시 정규분포를 이룬다는 가정 하에 척도값을 도출한다.

Thurstone은 적정 수준의 평정자수로 300명을 추천하였으나 일반적으로 100명 이상의 평정자면 충분하며(김명소, 1992) 각 진술문을 비교·판단하는 방식에 따라 3가지 척도화 방법이 존재한다.

첫 번째 도구는 Thurstone(1927)이 제시한 상대적 비교법(law of comparative judgment)으로 응답자들로 하여금 두 개씩 짝을 지은 문항을 제시하고 검사가 측정하고자 하는 속성에 대하여 어느 문항의 속성이 더 큰지 판단하도록 한다. 응답자들은 가능한 모든 짝진 문항의 경우의 수만큼

짜진 문항의 비교를 수행해야 하기 때문에 전체 문항의 수가  $n$ 이라면 총  $n(n-1)/2$ 번의 판단을 수행해야 한다. 예를 들어, 진술문  $i$ 와  $j$ 의 비교시,  $i$ 를 더 긍정적이라고 답한 응답자의 빈도를  $f_{ij}$ 라고 할 때, 전체 평정자 집단에서  $f_{ij}$ 가 차지하는 비율인  $p_{ij}$ 를 표준정규분포의  $Z$ 값인  $Z_{ij}$ 로 대체할 수 있고, 모든 가능한 짜진 진술문의 비교·판단에 의하여 도출된  $Z_{ij}$  값으로 표준점수 행렬표를 작성하면 <표2>와 같다.

<표2> 두 가지 진술문에 대한 평정자들의 표준점수행렬표  
(열의 진술문을 행의 진술문보다 더 긍정적이라고 응답한 평정자들의 표준점수행렬표)

| 진술문 | 1              | 2              | · | i              | · | n              |
|-----|----------------|----------------|---|----------------|---|----------------|
| 1   | $z_{11}$       | $z_{21}$       | · | $z_{i1}$       | · | $z_{n1}$       |
| 2   | $z_{12}$       | $z_{22}$       | · | $z_{i2}$       | · | $z_{n2}$       |
| ·   | ·              | ·              | · | ·              | · | ·              |
| j   | $z_{1j}$       | $z_{2j}$       | · | $z_{ij}$       | · | $z_{nj}$       |
| ·   | ·              | ·              | · | ·              | · | ·              |
| n   | $z_{1n}$       | $z_{2n}$       | · | $z_{in}$       | · | $z_{nn}$       |
| 총합  | $\sum z_{1.}$  | $\sum z_{2.}$  | · | $\sum z_{i.}$  | · | $\sum z_{n.}$  |
| 평균  | $\bar{z}_{1.}$ | $\bar{z}_{2.}$ | · | $\bar{z}_{3.}$ | · | $\bar{z}_{n.}$ |

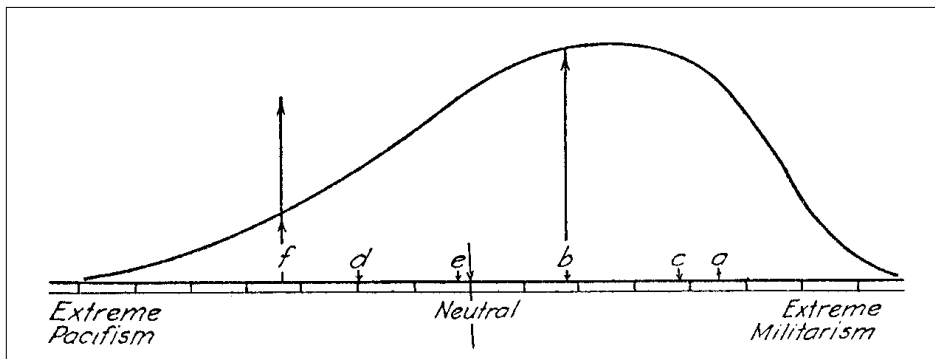
출처: 박도순(2006)

<표2>의 마지막 행에 나타난 각 진술문의  $Z_n$ 이 해당 문항의 척도값이 된다. 응답자에게 요구되는 판단 수를 줄이기 위하여 개발된 연속적 동간법(the method of successive intervals)은 상대적 비교법과 달리 평정자들로 하여금 각 진술문을  $m$ 개의 서열화된 범주로 나눌 것을 요구한다(Safir, 1937).  $m$ 개의 범주는 심리적 연속선상을 따라 연구자들에 의하여 선행적으로 정의된 연속적 구간이며, 각 응답자가 판단한 결과로부터 도출된 진술문의 범주별 누적비율이 자극의 척도값과 표준편차를 구하는데 활용된다.

Thurstone & Chave(1929)는 연속적 동간법과 유사한, 유사동간법(the

method of equal-appearing intervals)을 개발하는데, 이것은 척도화된 결과가 동간성을 이루고 있다고 가정한다. 유사동간법에서도 역시 평정자들은 연구자들에 의하여 정의된 m개의 범주로 언어적 자극, 즉 문항을 나누는데 이 때 각 자극에 할당된 범주 점수의 평균 혹은 중앙치가 척도값으로 활용된다.

Thurstone(1928)은 전쟁과 평화에 대한 여섯 개의 태도 진술문을 예시로 제시하며, 아래와 같은 척도화 방법에 의하여 산출된 문항의 위치와 모집단의 태도분포를 제시하였다. 결과적으로 피험자의 태도 위치와 문항의 위치가 같은 연속선상에 놓여지게 됨을 알 수 있다.



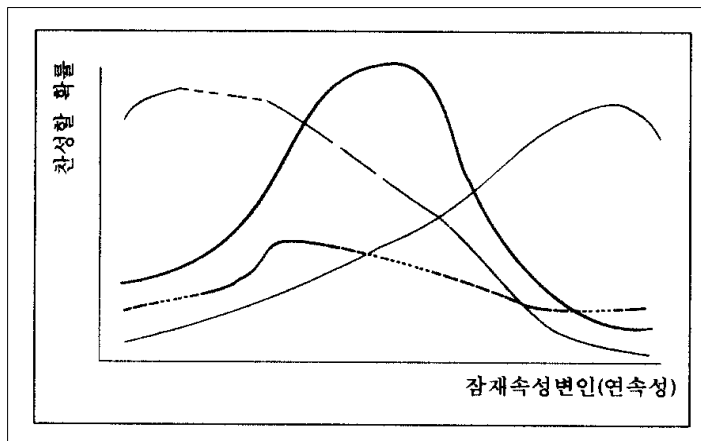
<그림1> 전쟁과 평화에 관한 진술문 위치와 피험자 모집단 태도 분포(Thurstone, 1928)

Thurstone에 의하면 피험자들은 반전주의에 대한 자신의 태도 위치와 근접한 문항에 긍정적으로 응답할 확률이 크기 때문에, 평화를 희망하는 피험자는 비교적 d, f문항에 긍정적으로 응답할 확률이 높으며, 중립적 태도의 피험자는 b와 e문항에, 반대의 태도를 지닌 피험자는 a와 c문항에 찬성할 확률이 높다.

Thurstone은 이와 같이 태도에 관한 진술문을 척도화하는 공식을 개발하였으나 피험자를 태도의 연속선상에 위치시키는 모형을 제시하지 않고 대신

피험자들의 태도를 측정하기 위한 직관적인 논리를 활용하였다(Robert, 1995). 진술문의 척도값이 산출된 후, 응답자들은 각각의 진술문에 대한 동의 혹은 반대 여부를 나타낼 것을 요구받는데, 이때 피험자는 자신의 태도와 유사한 특성을 가진 진술문에 동의할 것이기 때문에, 피험자가 동의한 태도 진술문의 척도값들의 중앙값 혹은 평균값으로 피험자의 최종 태도 점수는 확정된다.

Thurstone척도에 있어서 중요한 것은, 측정하고자 하는 태도의 연속선상의 전 범위를 포괄하는 문항들로 척도가 구성되어야 한다는 사실이다. 바로 이러한 점이 Likert 척도와 구별되는 점으로, 속성에 관한 선호도에 있어서 문항의 특성수준이 다르기 때문에 척도가 측정하는 속성이 강한 피험자라 하더라도 모든 문항에 동의하지 않을 것이 가정된다. 이러한 특성은 비단조적 문항특성곡선(non-monotonic item characteristic curve)의 형태로 구체화되며 단조적 증가형태의 누적곡선 형태만을 가정하는 리커트척도의 문항 특성곡선과 구별된다.



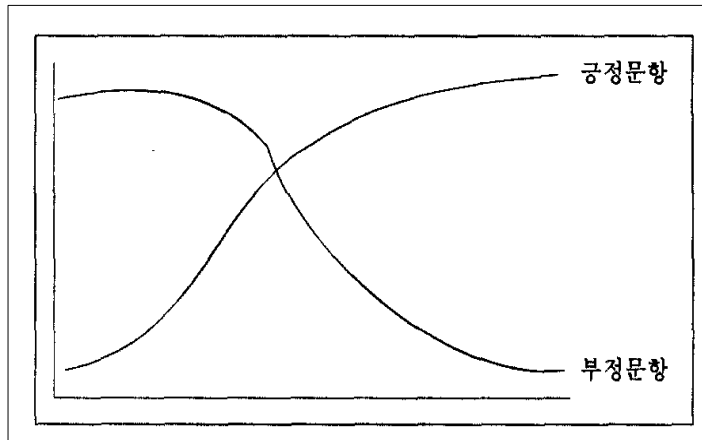
<그림2> Thurstone척도 문항의 문항특성곡선(김명소, 1997)

## 2) Likert 척도

### ① 척도의 구성

Likert척도 제작법(Likert, 1932)은 Thurstone 척도가 척도구성에 있어 수많은 응답자와 그들의 판단 과정을 요구하는 등의 문제를 고려하여 개발된 방법으로, 척도제작 절차가 수월하고 척도점수 산출과정이 직관적으로 이해되는 측면이 있기 때문에 현재 가장 넓게 활용되는 척도법이다(권재웅, 2012). Likert 척도 제작법은 Thurstone 척도 제작법에서 수행된 평정자에 의한 문항척도화 과정을 거치지 않고, 직접 응답자들을 척도화하는 응답자 중심방법으로서, 척도화의 목적은 태도연속선상에서 응답자의 태도 위치를 도출하는 것이다. 각 문항은 측정하고자 하는 속성에 대한 강한 긍정이나 강한 부정의 태도를 나타내는 문항들로 구성되고, 응답자들은 각 문항이 나타내는 특정 태도에 관한 진술문에 동의 혹은 비동의를 표시한다. 응답자가 각 문항에 동의하는 정도에 따라 문항별 척도점수가 산출되며 전체 문항의 척도점수의 총합이 응답자의 최종 척도치로 작용한다.

측정하고자 하는 태도의 속성이 강한 응답자일수록, 긍정적인 진술문을 담고 있는 문항에 강한 긍정을 나타낼 것이고, 해당 태도에 관한 부정적인 문항에 동의하지 않을 것이기 때문에, Likert 문항특성곡선은 누적적(cumulative)인 형태로 도출된다. 따라서 Thurstone척도와 같이 중립적 태도의 문항이 없고 대체로 강한 긍정 혹은 강한 부정을 나타내는 문항들로 검사가 구성되는데(김명소, 1992), 한 가지 형식의 진술문만을 사용할 경우에 나타날 수 있는 반응경향을 감소시키기 위함이다(임인재, 김신영, 박현정, 2006).



<그림3> Likert척도 문항의 문항특성곡선(김명소, 1997)

각 문항에 대한 선택지는 ㉠대단히 찬성한다, ㉡찬성한다, ㉢보통이다 혹은 잘 모르겠다, ㉣반대한다, ㉤적극 반대한다로 이루어진 5개의 유목으로 나뉠 수 있고, 진술문의 특성에 따라 3점 척도, 4점 척도, 7점 척도 등이 일반적으로 사용된다(박도순, 2006). 다음은 한국청소년정책연구원에서 실시한 청소년 건강 실태조사에서, 부모의 태도와 관련한 영역을 측정하기 위한 Likert 척도의 예(최인재, 2010)이다. 왼쪽의 문항을 읽고 응답자는 자신에게 해당하는 오른쪽의 범주에 표시하게 된다.

<표3> 청소년의 부모관련 심리건강상태에 관한 Likert척도 검사 도구

| 문항<br>번호 | 문항                           | 전혀<br>그렇지<br>않다 | 그렇지<br>않다 | 그렇다 | 매우<br>그렇다 |
|----------|------------------------------|-----------------|-----------|-----|-----------|
| 1        | 정신적으로 편안하게 해 주신다             | ①               | ②         | ③   | ④         |
| 2        | 나를 잘 알고 이해해 주신다              | ①               | ②         | ③   | ④         |
| 3        | 따뜻하게 대해 주신다                  | ①               | ②         | ③   | ④         |
| 4        | 고민을 들어 주신다                   | ①               | ②         | ③   | ④         |
| 5        | 힘들고 어려울 때 도와 주신다             | ①               | ②         | ③   | ④         |
| 6        | 실망하거나 좌절할 때 위로해 주신다          | ①               | ②         | ③   | ④         |
| 7        | 진학이나 진로문제에 대해 조언해 주신다        | ①               | ②         | ③   | ④         |
| 8        | 공부에 도움이 되는 말씀을 해 주신다         | ①               | ②         | ③   | ④         |
| 9        | 좋은 책이나 필요한 정보를 알려 주신다        | ①               | ②         | ③   | ④         |
| 10       | 올바른 공부 태도와 사는 방식에 대해 가르쳐 주신다 | ①               | ②         | ③   | ④         |
| 11       | 용돈을 주신다                      | ①               | ②         | ③   | ④         |
| 12       | 공부에 필요한 것들을 사 주신다            | ①               | ②         | ③   | ④         |
| 13       | 준비물이 필요할 때 돈을 주신다            | ①               | ②         | ③   | ④         |
| 14       | 경제적인 걱정 없이 생활하게 해 주신다        | ①               | ②         | ③   | ④         |

출처: 최인재(2010)

## ② 척도의 제작

Likert척도 제작의 절차는 Thurstone 척도 제작의 절차와 비슷하지만, 태도 점수를 결정하는 방법이 훨씬 간단하다. 박도순(2006)이 제시한 Likert 제작 절차를 토대로 Likert척도 제작 원리를 살펴보면 다음과 같다. 우선 진술문을 수집·작성하여 20개 이상의 진술문을 만든 후 진술문의 성질을 고려하여 응답의 범주를 나눈 후, 측정하고자 하는 태도나 속성에 다양성을 나

타넬 사람들(보통 50~100명)에게 척도검사를 실시한다(김명소, 1997). 각 진술문에 반응한 피험자의 사례수와 백분율을 산출하여, 선택지에 해당하는 점수비중을 결정한다. 선택지의 점수비중(scoring weight)을 결정하는 방법으로, 임의적 비중 방법(arbitrary weighting method), 표준편차 비중방법(sigma-deviate weighting method), 표준점수 비중방법(standard-score weighting method)의 세 가지 결정방법이 있다. 임의적 비중방법은 각 선택지의 순서대로 1, 2, 3, 4, 5와 같이 임의적으로 점수비중을 두는 것이고, 표준편차 비중방법은 하나의 태도가 정상분포를 이룬다는 가정 하에 점수 비중을 결정하는 방법이다. 응답자가 선택한 각 문항의 선택지의 빈도를 비율화하여, 누가비율의 중앙치에 해당하는 z점수를 산출하고, 산출된 z점수를 양수로 바꿀 수 있는 가장 작은 수를 더하여 모두 양수화 하면 최종 선택지의 점수가 결정된다. 표준점수 비중방법은 임의적 비중방법에 의한 점수를 응답자들의 선택에 기초하여 표준화하는 방법으로, 임의적 비중방법에 표준점수 개념을 추가한 것이다. Likert 척도의 이론적 가정에 가장 적절히 부합하는 것은 표준편차 비중방법이지만, Likert에 의하면 임의적 비중방법과 표준편차 비중방법의 상관인 .99에 도달하고, 임의적 비중방법의 사용이 훨씬 간단하기 때문에, 일반적으로 임의적 비중방법이 가장 많이 활용된다.

각 선택지의 점수가 결정되면, 문항분석을 통하여 최종 진술문을 선택하는데, 보통 진술문에 대한 반응점수와 총점간의 상관계수와 척도의 내적일관성을 측정하기 위한 Cronbach- $\alpha$  계수에 기여하는 문항들로 최종 진술문을 선택한다.

Likert 척도제작법은 제작 과정이 수월하고 신뢰도가 높은 장점이 있으나(Roberts et al, 1995), 타당도는 제작된 척도에 따라 다르며 척도의 동간성을 보장할 수는 없다(임인재 외, 2006). 특히 측정하고자 하는 속성에 대한

긍정 혹은 부정적인 진술문만을 척도에 포함함으로써 총점과 문항간 상관계수가 높아질 수도 있으나 중간수준의 진술문을 포함하지 않아 중간수준의 태도를 지닌 응답자의 타당한 척도치를 도출할 수 없다는 주장이 제기된다(Drasgow et al., 2010; Roberts, 1995; Roberts et al., 2000). Roberts et al.(1995)에 의하면, Likert 척도를 개발하는 과정에서 적절한 문항을 선택할 때 중간수준의 문항은 제외되는 경향이 있으며, 실제로 Likert(1932) 또한 중간수준의 문항을 제외하는 것이 바람직하다고 보았다.

또한 Likert 척도의 문항을 통해 얻어진 중간수준의 척도점수는 모든 문항에 대하여 ‘보통이다’ 혹은 ‘별로 의견이 없다’의 반응을 통하여 얻어진 점수일 수도 있고 전 문항의 반에 ‘대단히 찬성한다’를 표시하고 나머지 반에 ‘대단히 반대한다’라는 극단적인 선택지에 반응을 하여 얻어질 수도 있는 문제점이 있다(임인재 외, 2006).

### 3) Guttman 척도

#### ① 척도의 구성

Guttman척도는 응답자의 반응에 변산이 생기는 이유는 Thurstone척도와 같이 자극에 의해서만 발생하거나, Likert척도의 경우처럼 응답자 자체의 다양성에서 오는 것이 아니라, 자극을 판단하는 응답자와 자극 간의 차이에 근거하여 나타난다고 보는, 반응중심방법의 척도화방법이다(Guttman.1950; Guttman, 1953) Guttman은 Thurstone척도와 Likert척도의 문항이 일차원 속성을 충족하지 못함을 문제로 지적하며 일차원성을 보이는 척도제작법을 개

발하였는데, 각 문항이 완벽한 일차원성을 보장할 수 있는 증거는 전체점수를 토대로 각각의 문항에 대한 반응을 모두 예측할 수 있을 때 제공된다(박도순, 2006). 즉 일차원성이 확보된 Guttman 척도에서 모든 문항은 같은 구인에 관하여 진술하고 있기 때문에, 다분문항의 경우, 총점이 높은 응답자는 상대적으로 낮은 총점을 받은 응답자보다 모든 문항에 대하여 더 높은 점수를 받아야 하고, 이분문항의 경우에는 총점이 더 높은 응답자가 총점이 낮은 피험자보다 긍정적으로 답한 문항의 수가 많아야 한다.

Thurstone척도제작에서는 각 진술문이 가진 특성의 정도를 판단하는 것이 응답자들의 주요 임무였던 반면, Guttman척도제작의 경우, 응답자는 객관적인 입장에서 진술문을 판정하는 것이 아니라, 응답자의 주관적 입장에서 판정할 것을 요구받는다. 판정 형식은 특정 진술문에 대한 찬성여부나 진술문을 자극의 특징에 따라 분류하는 것과 같은 유목적 반응이 될 수 있고, 서로 다른 자극과 비교하는 상대적인 반응일 수도 있다(임인재 외, 2006).

<표4>는 Porat(1977)이 Maslow(1943)의 욕구위계설을 검증하기 위하여 활용한 Guttman척도의 예이다. Maslow(1943)는 인간의 욕구는 생리적 욕구, 안전의 욕구, 애정의 욕구, 존경의 욕구, 자아실현의 욕구의 5단계로 이루어져 있으며 이전 단계의 욕구가 충족되어야 다음 단계의 욕구가 동기화된다고 주장하였다. Porat(1977)은 각 단계에 해당하는 문항을, 직장 내 맥락에 따라 수정하여 최종적으로 5개의 문항을 101명의 산업조직의 관리자 및 전문가에게 실시하였다.

<표4> Maslow(1943)의 욕구위계설 검증을 위한 Guttman 척도

| 척도 수준 | 척도 항목                | 매우 불만 | 불만 | 약간 불만 | 보통 | 약간 만족 | 만족 | 매우 만족 |
|-------|----------------------|-------|----|-------|----|-------|----|-------|
| I     | 물질적 욕구(봉급)           | ①     | ②  | ③     | ④  | ⑤     | ⑥  | ⑦     |
| II    | 안전 욕구(직장 고용기간)       | ①     | ②  | ③     | ④  | ⑤     | ⑥  | ⑦     |
| III   | 존경 욕구(동료와 상사로부터의 존중) | ①     | ②  | ③     | ④  | ⑤     | ⑥  | ⑦     |
| IV    | 자율성 욕구(선택의 자율성)      | ①     | ②  | ③     | ④  | ⑤     | ⑥  | ⑦     |
| V     | 자아실현 욕구(전문성 개발 기회)   | ①     | ②  | ③     | ④  | ⑤     | ⑥  | ⑦     |

출처: Porat(1977)

② 척도의 제작

Guttman척도는 각 진술문이 완벽한 누적척도를 이루며, 검사의 구인을 이루는 특성의 전 범위를 포괄하는 진술문들의 위계적 서열이 응답자의 반응에 의하여 확인되었을 때 비로소 척도가 될 만하다고 간주한다. 따라서 일련의 문항에 대한 응답자의 반응이 척도를 형성하고 있는지의 여부를 판정하는 과정이 Guttman척도 제작과정이라고 할 수 있고, 엄밀하게 말하면 일련의 문항들이 제대로 구성된 것인지 평가하는 절차라고도 볼 수 있다(박도순, 2006). 이 때 특정 척도가 응답자의 반응에 의하여 척도를 구성하고 있는지 여부를 분석하는 과정을 스칼로그램 분석(scalogram analysis)이라고 한다(Guttman, 1950).

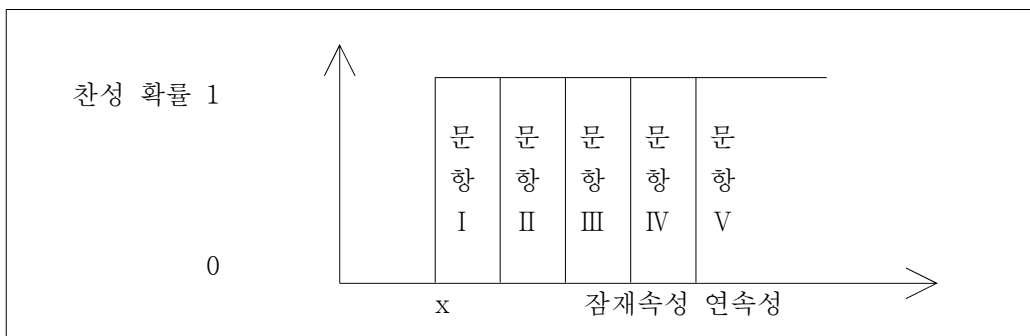
척도제작자는 측정하고자 하는 속성을 적절히 대표하는 문항을 만들고 조사대상이 되는 모집단의 표본을 추출하여 응답자들로 하여금 자신과 일치한다고 생각되는 각각의 진술문에 찬반 여부 혹은 찬반의 정도를 표시하는 범주에 표시하도록 한다. 다음으로 해당 척도가 일차원적일 것이라는 가설 하

에 표본의 응답자료를 바탕으로 척도구성가능성(scalability)을 검증하는데, 이 때 각각 12개의 문항과 100명 내외의 응답자가 필요하다. 위에서 예로 든 Maslow(1943)의 욕구위계설을 검증하기 위한 Guttman 척도는 1-4점 척도를 0으로, 5-7점 척도를 1로 재코딩하여 이분문항의 형태로 변환한 뒤, 응답자들의 분포를 고려하였다. 이 경우 욕구위계설이 완벽한 이론이라면 나타날 수 있는 Guttman척도에 대한 반응유형 역시, Guttman의 완벽한 누적척도를 따를 것이며 다음과 같이 여섯 개의 유형만이 가능하다.

<표5> 이분문항에 관한 Guttman척도의 이상적 반응유형

| I | II | III | IV | V | Score |
|---|----|-----|----|---|-------|
| 1 | 1  | 1   | 1  | 1 | 5     |
| 1 | 1  | 1   | 1  | 0 | 4     |
| 1 | 1  | 1   | 0  | 0 | 3     |
| 1 | 1  | 0   | 0  | 0 | 2     |
| 1 | 0  | 0   | 0  | 0 | 1     |
| 0 | 0  | 0   | 0  | 0 | 0     |

출처: Porat(1977)



<그림4> Guttman척도 문항의 문항특성곡선(김명소, 1997)

위에 나타난 반응유형을 제외한 나머지 유형은 비척도 반응유형으로서, 모든 Guttman척도의 제작 절차에 있어서 완벽한 누적척도만 실제로 나타나지 않기 때문에 비척도 반응유형, 즉 오차의 정도에 대한 기준이 존재한다. Guttman은 재생가능성 비율(reproducibility)의 개념을 사용하여, 비척도 반응유형의 피험자를 그 반응형태에 가장 가까운 완벽한 척도반응 유형을 갖도록 수정하였을 때 생기는 차이를 오차의 정도로 삼고, 다음과 같은 재생가능성지수(coefficient of reproducibility) 공식을 제안하였다.

$$CR = 1 - \frac{\text{오차반응의총수}}{\text{반응의총수}} = 1 - \frac{\text{오차반응의총수}}{\text{문항의수} \times \text{피험자수}}$$

Guttman은 구성된 척도의 재생가능성지수가 .90이어야 척도가 될만 (scalable)하다고 하였으며, Guttman척도는 Thurstone척도와 같이 경험적 관찰과 모형의 적합도를 검토할 수 있기 때문에 Likert척도보다 우세함을 보인다(김명소, 1992).

## 2. 다분문항반응이론에 의한 정의적 특성 검사

검사를 구성하는 각 문항을 분석단위로 하는 문항반응이론은 Binet와 Simon이 나이에 따라 문항의 정답을 맞힐 확률을 그린 문항특성곡선에 근거하여 발달하였는데, 고전검사이론과 같이 비현실적인 진점수 추론 과정을 가정하지 않고 현실적이고 정교한 검사 모형을 제공한다. 문항반응이론은 고전검사이론에 비하여 보다 정확하고 정교한 피험자 특성 추정치, 피험자 능력 수준에 따라 각기 다른 능력 추정의 오차, 표본에 종속되지 않는 문항

과 피험자 특성 추정치 등을 제공함으로써 보다 타당한 문항분석을 수행한다(Sgammato, 2009).

이처럼 문항반응이론이 발전함에 따라 문항과 피험자 모수가 동일한 척도상의 값으로 나타내짐으로써 정의적 특성에 관한 개개의 진술문에 관한 피험자의 문항반응특성곡선을 나타낼 수 있게 되었고, 보다 정교한 문항분석이 가능해지게 되었다. 실제로 피험자가 어떠한 심리적 기제를 따라 자신의 정의적 특성을 나타내는지에 관한 연구가 진행되었고(Lin, Carter, Zickar, Dalal, Adams & Wolford, 2009; Lake & Zickar, 2010), 누적기제와 전개기제의 문항반응모형에 대한 정의적 특성 검사 데이터의 적합도를 살펴보는 연구와(Chernyshenko, Stark, Chan, Drasgow & Williams 2001; 강태훈, 김동일, 2010; 박동건 외, 2012; 이선희, 2008) 보다 타당한 피험자의 특성을 추정하고자 응답 자료의 기제에 적합한 문항반응모형이 개발되기도 하였다(Andrich, 1988; Roberts, 1995; Roberts et al., 2000). 이때 활용되는 피험자의 응답 자료는 Likert 척도의 형태이며 피험자의 응답 데이터와 정의적 특성 간의 관계는 이상점 반응 과정(ideal point response process)을 가정한 전개기제(unfolding mechanism)와 우세 반응 과정(dominance response process)을 가정한 누적기제(cumulative mechanism)로 나눌 수 있다(Coombs, 1964).

Thurstone과 Likert가 척도를 개발할 당시, 그들은 피험자들이 자신이 개발한 척도의 진술문에 응답할 때 어떠한 심리적 응답과정을 거치는지 체계적인 이론을 정립하지는 않았으나, 이후 Coombs(1964)에 의하여 정의된 이상점 과정(ideal point process)과 우세반응 과정(domain response process)은 각각 Thurstone척도와 Likert척도에 대한 피험자들의 응답과정을 설명할 수 있었다(Roberts, 1995; Roberts et al., 1999). 피험자들이 자신의 태도와

유사한 문항의 진술문에 동의하며, 태도와 문항의 위치가 가장 근접한 문항에 높은 확률로 동의하는 것을 이상점 과정이라고 하며, 이에 의한 피험자의 응답 자료는 전개기제를 통하여 가장 잘 분석된다.

Likert 척도에 응답하는 피험자들은 우세반응 과정에 근거하여, 문항에 관하여 자신의 태도 위치가 우세한 정도만큼 긍정적으로 작성된 문항에 동의하며, 이들의 응답 자료는 누적기제에 의하여 분석된다.

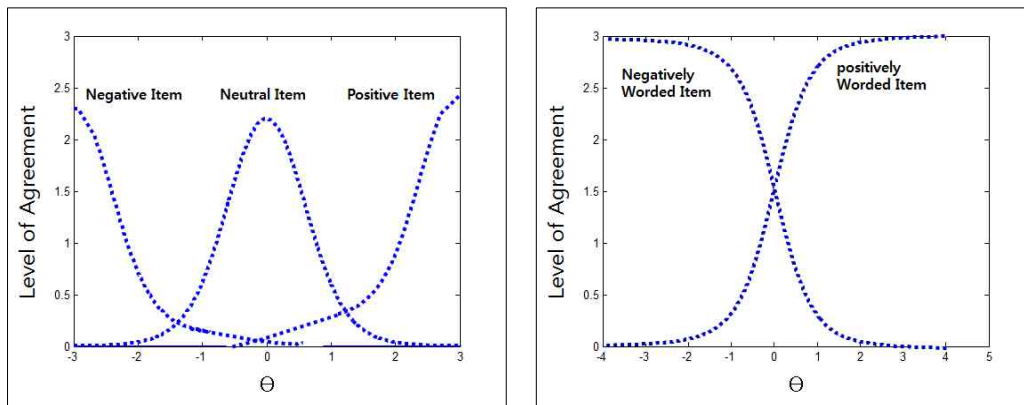
Thurstone척도와 Likert척도의 문항 간의 가장 큰 차이점은 중립적 수준의 진술문의 유무로서, 즉 전개기제 형태의 곡선의 유무로 볼 수도 있다. <그림6>에서 보는 바와 같이 모든 Thurstone 척도 문항이 전개기제를 나타내는 것은 아니다(Green, 1954). 부정적이거나 긍정적인 방향으로 치우친 문항의 특성곡선은 문항의 위치보다 더 극단적인 태도를 지닌 피험자가 거의 없기 때문에 꺾임현상이 거의 나타나지 않지만, 중립적 수준의 문항의 경우, 양 극단의 피험자가 다수 존재하기 때문에 중립적 수준에서 멀어질수록 동의를의 수준은 낮아진다.

이상점과정과 우세반응과정에 의한 전개기제와 누적기제의 문항특성곡선은 중립적 수준(intermediate level)의 문항에서 가장 잘 나타나지만, 중립적 수준의 문항을 검사도구에 포함하는 것에 대한 반대의견도 존재한다. 중립적 수준의 문항은 일차원적 수준을 넘는 복잡한 다차원성을 가지기 때문에, 일차원적인 Likert척도를 활용하여 중립적 수준의 문항에 대한 피험자의 특성을 추정하는 것은 적합하지 않다는 입장이다(Oswald & Schell, 2010). 가령, 낙태에 대한 찬반여부를 묻는 중립적 수준의 진술문으로서 ‘낙태는 여성의 선택사항이지만, 단지 낙태의 편이성으로 인하여 낙태가 허용되어서는 안된다’라는 문항을 고려해 보자. 해당 문항에 반대한 피험자의 경우, ‘낙태는 여성의 선택사항’이라는 의견에 반대를 하였는지, 혹은 태아의 생명을 존

중하여 ‘단지 낙태의 편이성으로 인하여 낙태가 허용되어서는 안된다’라는 의견에 반대를 하였는지의 여부를 알 수 없다. 즉 해당 문항은 낙태에 대한 여성의 권리와 태아의 인권에 관한 두 가지 구인을 내포하는 문항으로서 중첩된 문항(double-barrel item)으로 여겨질 수 있다. Likert척도가 중립적 수준의 문항을 내포하지 않는 가장 큰 이유는 이처럼 중첩된 문항의 경우, 일차원성을 보장할 수 없기 때문이다(Likert, 1932).

이 외에도 이상점 모형에 적합한 중립적 문항을 제작하는 것이 힘들기 때문에, 실제로 척도를 활용하는 실무자들에게 중립적 수준의 문항을 제작하는 명확한 방안을 제공할 필요성도 제기되며, 중간수준의 문항은 Thurstone 척도의 독특한 특성으로서, 모든 정의적 특성에 관한 척도에서 중간수준의 문항을 고려하는 것은 옳지 않다는 입장도 존재한다(Brown & Maydeu-Olivares, 2010; Dalal, Withrow, Gibby & Zickar, 2010; Waples, Weyhrauch, Connell & Culbertson, 2010).

하지만 Roberts et al.(1995)이 수행한 모의실험 연구 결과에 따르면, 전통적으로 Likert 척도를 개발하는 과정을 따른 가상의 Likert 척도 문항에서도 꺾임현상이 나타나 전개기제를 이루는 모습을 볼 수 있었다.



<그림5> 전개기제와 누적기제 하에서의 문항특성곡선 (Roberts et al., 1998)

지난 100년간 심리측정학자들은 피험자의 특성 수준이 높아질수록 반응 확률이 높아진다는 우세모형에 근거한 검사이론을 활발히 연구한 반면 최근에 와서야 몇몇의 심리측정이론가들에 의하여 이상점과정에 근거한 비인지적 특성 분야의 검사이론이 연구되기 시작하였다(Drasgow et al., 2010; Roberts et al., 2000). 문항반응이론의 맥락에서 이상점모형에 근거한 전개기제는, 비인지적 특성을 측정하는 검사의 구인 연속선상에서 피험자의 위치와 문항의 위치가 근접할수록 피험자가 해당 문항에 동의할 확률이 크다는 전제를 갖는다. 전개기제모형의 문항반응모형은 이분 문항에 적합한 모형(Andrich, 1988; Andrich & Luo, 1993; Desarbo & Hoffman, 1986; Hoijsink, 1990, 1991)과 다분 문항에 적합한 모형(Andrich, 1996; Roberts, 1995, Roberts et al., 2000)이 개발되었다.

본 연구는 정의적 특성 검사에 일반적으로 활용되는, 다분문항의 Likert 척도에 기반한 검사 데이터를 토대로, 대표적인 전개기제의 일반화등급전개 모형과 일반적으로 활용되는 누적기제인 일반화부분점수모형을 활용한다.

### (1) 일반화부분점수모형(Generalized Partial Credit Model)

Muraki(1992)의 일반화부분점수모형(*generalized partial credit model*)은 Masters(1982)의 부분점수모형(*partial credit model*)을, 문항 변별도 지수를 고려하는 보다 일반화된 모형으로 확장시킨 것으로, 기본적인 모형의 형성 과정은 부분점수모형과 동일하다.

일반화부분점수모형은 다분 문항에서 특정한 범주를 택할 확률을 두 지점에서의 누적 확률의 차이로 계산하는, Samejima(1969)의 등급반응모형으로 대표되는 ‘차이 모형(*difference model*)’과는 달리, 각 범주에 대한 반응 확률

을 구하기 위하여 범주에 대한 반응의 합으로 나누어주는 ‘총합으로 나누기 (divide-by-total model)’ 모형에 해당한다.

일반화부분점수모형은 식(1)과 같이  $\theta$ 의 능력을 가진 피험자가 각 범주에 반응할 확률을, 이분모형을 적용하여 인접한 범주  $k-1$ 과  $k$ 의 두 선택지에서  $k$ 를 선택할 조건부확률로 모두 표현한 후, 그 확률값들의 합이 1임을 조건으로 하여 각 범주의 개별 확률을 구하게 된다.

$$C_{ik} = P_{ik|k-1,k}(\theta) = \frac{P_{ik}(\theta)}{P_{i,k-1}(\theta) + P_{ik}(\theta)} = \frac{\exp[a_i(\theta - b_{ik})]}{1 + \exp[a_i(\theta - b_{ik})]}, \quad (k=1,2,\dots,C) \quad (1)$$

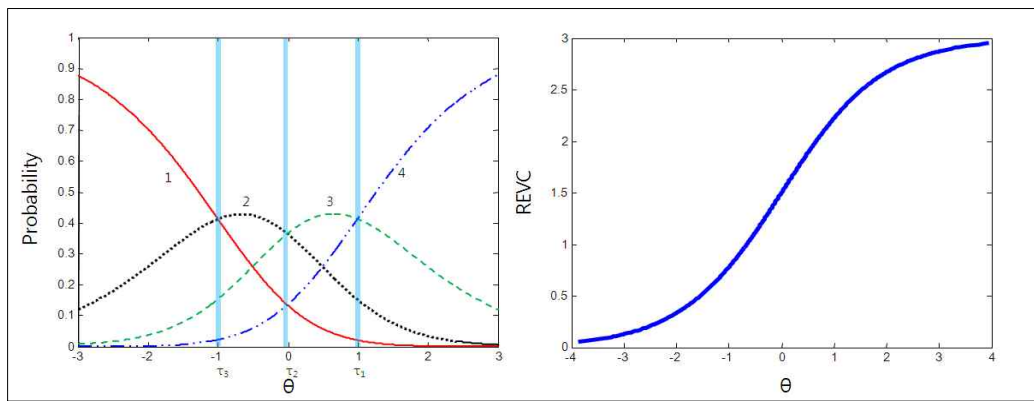
$\theta$ 는 피험자의 특성 수준을 나타내며  $a_i$ 는  $i$ 번째 문항의 변별도를,  $b_{ik}$ 는  $i$ 번째 문항의  $k$ 번째 범주의 단계모수를 의미한다. 단계모수란, Masters(1982)가 부분점수모형을 제시하며 활용한 개념으로서,  $b_{ik}$ 는  $P_{j,k-1}(\theta)$ 과  $P_{jk}(\theta)$ 을 나타내는 문항범주특성곡선(item category characteristic curve)이 교차하는 검사 특성 연속선상의 위치를 의미하며, 즉 인접한 범주에 대한 상대적 난이도를 나타낸다. 이로써 Masters는 Samejima(1969)의 등급반응모형보다 유연한 다분문항반응이론을 구축할 수 있었다.

각 개별 범주에 반응할 확률을 토대로  $\theta$ 의 능력수준을 가진 피험자  $j$ 가  $\alpha_i$ 의 변별도를 가진 다분문항에 대하여 각 범주( $y=0,1, \dots, C$ )에 반응할 확률을 수식으로 표현하면 다음과 같다.

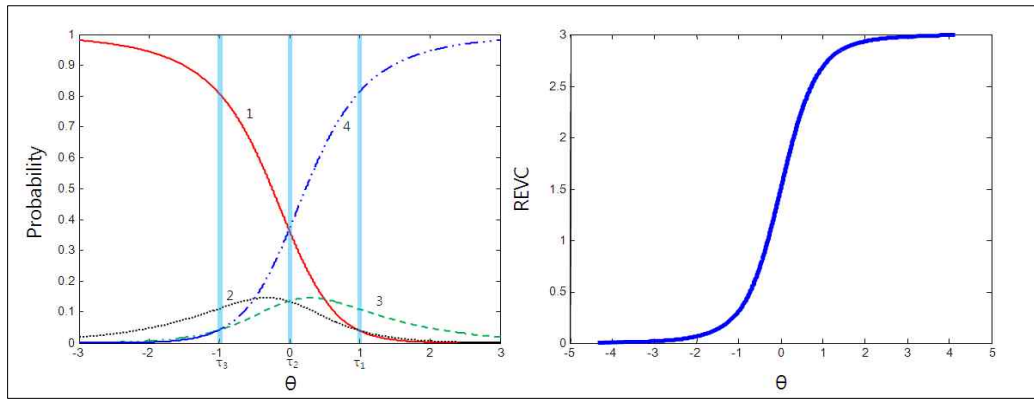
$$P(Y_i = y | \theta_j, \alpha_i, \delta_i, \tau_{ki}) = \frac{\exp\left\{\sum_{k=0}^y \alpha_i [\theta_j - (\delta_i - \tau_{ki})]\right\}}{\sum_{w=0}^C \exp\left\{\sum_{k=0}^w \alpha_i [\theta_j - (\delta_i - \tau_{ki})]\right\}} \quad (2)$$

부분점수모형에서 변별도 모수  $\alpha$ 가 추가됨에 따라 모형 식별을 위하여 보통  $\tau_{i0} = 0, \sum_{k=0}^c \tau_{ik} = 0$ , 그리고  $\exp \sum_{k=0}^c \alpha_i [\theta_j - (\delta_i - \tau_{ik})] = 1$ 이 가정된다(강태훈 외, 2012).  $(\delta_i - \tau_{ik})$ 는 k-1번째 범주에 대한 k번째 범주의 단계모수  $b_{ik}$ 를 문항난이도  $\delta_i$ 와 검사 문항 내의 단계난이도  $\tau_{ik}$ 의 차이로 표현한 형태이다.

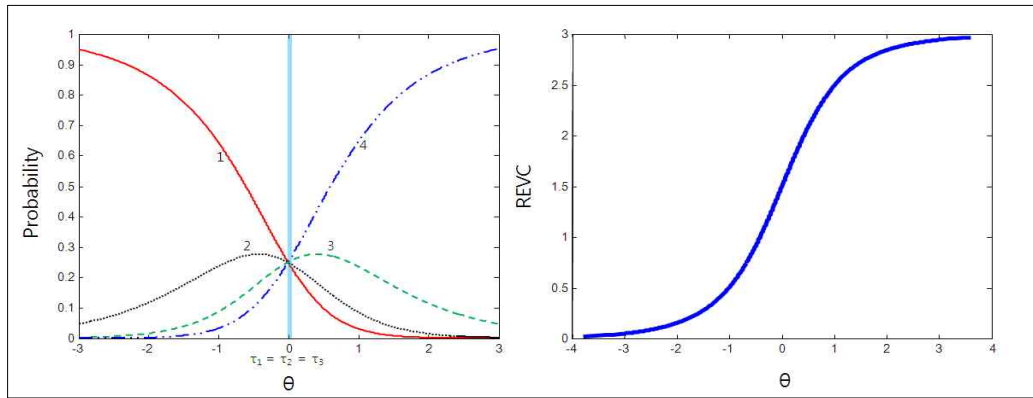
식 (2.1)과 (2.2)를 통하여, 피험자 특성 수준에 따른 문항범주특성곡선 (item category characteristic curve: ICC)과 반응기대값곡선(response expected value curve: REVC)을 각각 나타내면 아래 그림과 같다.



a.  $\alpha_i = 1, \delta_i = 0, \tau_1 = -1, \tau_2 = 0, \tau_3 = 1$



b.  $\alpha_i = 1, \delta_i = 0, \tau_1 = 1, \tau_2 = 0, \tau_3 = -1$



c.  $\alpha_i = 1, \delta_i = 0, \tau_1 = 0, \tau_2 = 0, \tau_3 = 0$

<그림6> 일반화부분점수모형의 문항범주특성곡선과 기대반응곡선

문항범주특성곡선에서 보듯이 단계모수가 인접한 두 곡선은 단 한 번만 교차하게 되고, 교차지점은 검사 특성의 척도 연속선상의 어디에서든지 나타날 수 있기 때문에, 한 문항에 대한  $\tau_{jk}$ 가 반드시 서열적이지 않으며, 모두 같은 값을 가질 수도 있다. <그림7-a>는 단계모수가 서열성을 가질 경우를 나타낸 경우이며, <그림7-b>는 단계모수 크기의 순서가  $\tau_1 > \tau_2 > \tau_3$  이며, <그림7-c>는 단계모수의 크기가 모두 동일한 경우이다. 세 경우 모두 단계모수의 서열성의 여부나 크기에 관계없이 기대반응곡선에서 누적기제를 따를 수 있다. 이는 특정한 범주의 점수를 받는 것으로 나타나는, 한 단계의 완성은 이전 단계를 완성하였다는 가정하에 가능하기 때문에, 비록 단계 난이도의 서열성은 변화할 수 있으나, 문항 범주 자체의 서열성은 변화하지 않기 때문이다(박정, 2001).

피험자 특성 수준이 증가할수록 해당 문항에 대한 반응 기댓값이 점차 높아지는 단조증가형태의 반응기댓값곡선은 일반화부분점수모형이 대표적인 우세반응모형임을 나타낸다.

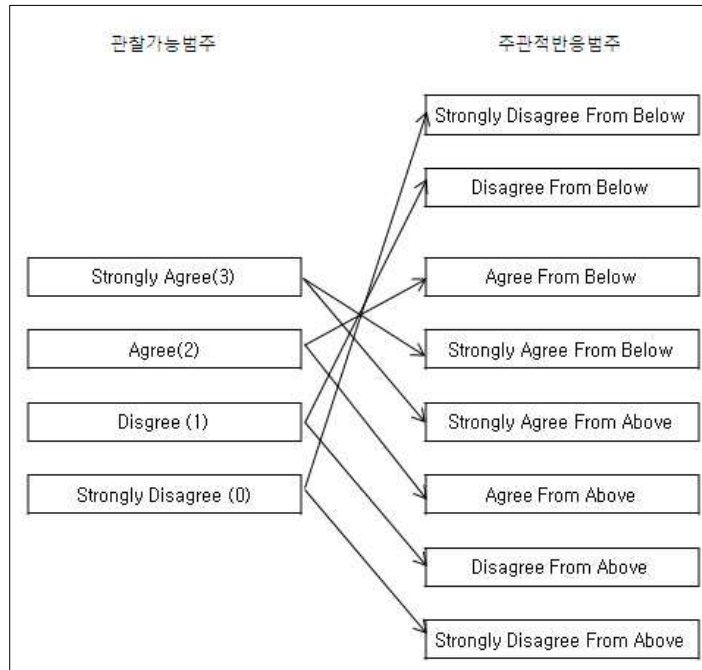
일반화부분점수모형에서  $\alpha_i$ 가 모든 문항에서 1이라고 가정하면 Masters의

부분점수모형이 되며  $\tau_{ik}$ 가 모든 문항에서 같다고 제약될 경우, Andrich의 평정척도모형을 나타내게 된다. 즉 일반화부분점수모형은 부분점수모형과 평정척도모형을 위계적으로 내포하는 보다 일반화된 모형이라고 볼 수 있다(강태훈, 김동일, 2010). 또한 문항의 범주가 총2개(C=2)이고, k=1,2일 때 일반화부분점수모형은 이분문항의 2모수로지스틱모형의 형태와 동일해지는데, 즉 문항반응이론을 적용하여 검사결과를 분석할 때 널리 활용되어지는 Birnbaum(1968)의 2모수로지스틱모형은 일반화부분점수모형의 특수한 형태라고도 볼 수 있다(Muraki, 1992).

## (2) 일반화등급전개모형(Generalized Graded Unfolding Model)

일반화등급전개모형은 이상점 반응에 기반한 전개기제 모형이지만, 그 구성 과정에서 누적기제 모형을 활용하며, 특히 Muraki(1992)의 일반화부분점수모형이 가진 일반성에 기인하여 해당 누적기제를 활용한다(Roberts et al, 2000).

일반화등급전개모형은 전개기제를 나타내기 위하여 문항에 대한 피험자들의 응답을, 진술문에 대한 표면적인 응답형식인 관찰가능범주(observable response categories: ORCs)와 피험자가 심리적으로 나타낸 내면적 응답형식인 주관적반응범주(subjective response categories: SRCs)로 구별한다. <그림8>은 반응범주의 구분의 구체적인 예를 들기 위하여, Roberts et al(2010)의 낙태에 관한 태도검사 문항 중, 가장 중립적 위치에 나타난 “낙태에 대한 나의 감정은 매우 복잡적이다”라는 문항을 가장 반대하는 0점부터, 가장 찬성하는 3점까지에 이르는 4가지 관찰반응범주와 이에 대한 주관적반응범주를, Javars(2004)가 제시한 mapping의 형식으로 나타낸 것이다.



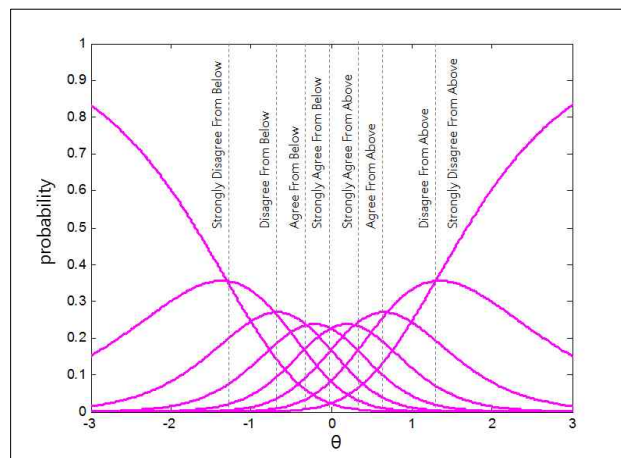
<그림7> “낙태에 대한 나의 감정은 매우 복잡적이다”에 관한 관찰가능 범주와 주관적반응 범주

문항에 대한 실제 피험자들의 응답범주는 위와 같이 매우 찬성, 찬성, 반대, 매우 반대의 네 가지로 나뉘지만, 각각의 관찰가능범주는 두 개의 주관적반응범주로부터 기인함을 알 수 있다. 예를 들어, 위의 그림에서 “strongly disagree”를 표시한 피험자의 경우, 낙태 자체를 무조건 반대하기 때문일 수도 있고(strongly disagree from below), 낙태에 대하여 무조건적인 찬성입장을 가졌기 때문일 수도 있다(Strongly Disagree From Above).

공식(2)는 Roberts et al.(2010)이 Muraki의 누적기제를 활용하여 나타낸 피험자의 주관적반응범주를 보여준다.

$$P(Y_i = y|\theta_j) = \frac{\exp\left\{\alpha_i \left[ y(\theta_j - \delta_i) - \sum_{k=0}^y \tau_{ik} \right]\right\}}{\sum_{w=0}^M \left\{ \exp\left\{\alpha_i \left[ w(\theta_j - \delta_i) - \sum_{k=0}^w \tau_{ik} \right]\right\} \right\}}, \quad \sum_{k=0}^M \tau_{ik} = 0 \quad (2)$$

$Y_i$ 는 문항  $i$ 에 대한 주관적 반응을 나타내고,  $0, 1, 2, \dots, M$ 값을 가지며, 0일 때에는 피험자가 문항의 진술문에 가장 강력하게 반대함을,  $M$ (=주관적 반응 범주의 수-1)일 경우에는 피험자가 문항의 진술문에 가장 높은 정도로 찬성함을 의미한다. 일반화부분점수모형에서와 마찬가지로,  $\alpha_i$ 는 문항  $i$ 의 변별도를,  $\tau_{ik}$ 는 문항  $i$ 의 위치 연속선상에서의  $k$ 번째 주관적 반응범주의 경계 위치모수를 나타내며,  $\tau_{i0}$ 는 임의적으로 0으로 정의되지만, 해당 값은 어떤 상수를 취하더라도 문항범주반응확률에 영향을 미치지 않는다(Muraki, 1992). <그림9>에서 알 수 있듯이, 일반화등급전개모형의 주관적반응범주는 누적기제를 따르며, 각각의 주관적반응범주가 가장 높은 확률로 나타나는 지점에 해당 범주곡선의 명칭이 표시되어 있다.



<그림8> 네 개 범주 문항의 주관적반응범주 반응곡선 (Roberts et al., 2000)

위의 mapping에서 보았듯이, 각 관찰가능범주는 2개의 주관적반응범주에 기인하고, Roberts et al.은 각 관찰가능범주에 대한 확률은 두 개의 해당 주관적반응범주의 확률의 단순합으로 나타난다. 이는 공식(3)으로 구체화 되었다.

$$P(Z_i = z|\theta_j) = P(Y_i = z|\theta_j) + P(Y_i = (M-z)|\theta_j) \quad (3)$$

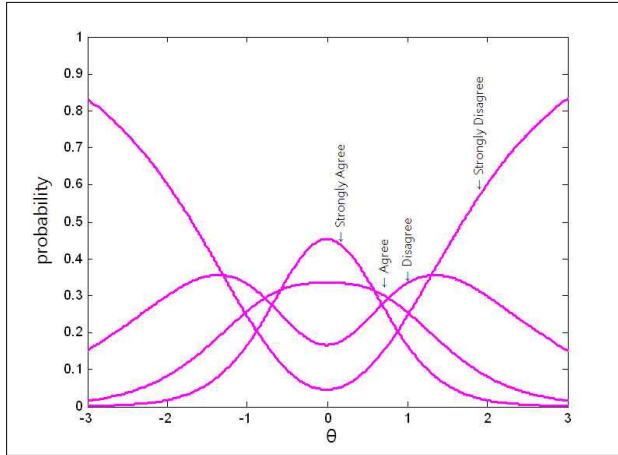
$Z_i$ 는 문항 $i$ 에 대한 관찰가능한 반응이며,  $0, 1, 2, \dots, C$ 의 값을 갖는 가운데,  $0$ 은 가장 강력한 반대를,  $C$ 는 가장 강력한 찬성의 수준을 의미하며, 구체적으로 ‘관찰가능범주-1’의 값을 갖는다. 마지막으로  $M=2C+1$ 을 의미한다.

일반화등급전개모형은  $\tau_{ik}$  값들이  $(\theta_j - \delta_i) = 0$ 인 지점을 중심으로 대칭을 이룬다고 가정하며, 이에 따라  $z_i \neq 0$ 이 아닌 경우,  $\tau_{i(C+1)} = 0$ ,  $\tau_{iz} = -\tau_{i(M-z+1)}$ 이 성립하는 가운데, 경계위치모수의 대칭성을 나타내는 공식(4)가 도출된다. 식(4)를 식(2)에 대입함으로써 아래와 같은 최종적으로 일반화등급전개모형의 공식이 완성된다.

$$\sum_{k=0}^z \tau_{iz} = \sum_{k=0}^{M-z} \tau_{ik} \quad (4)$$

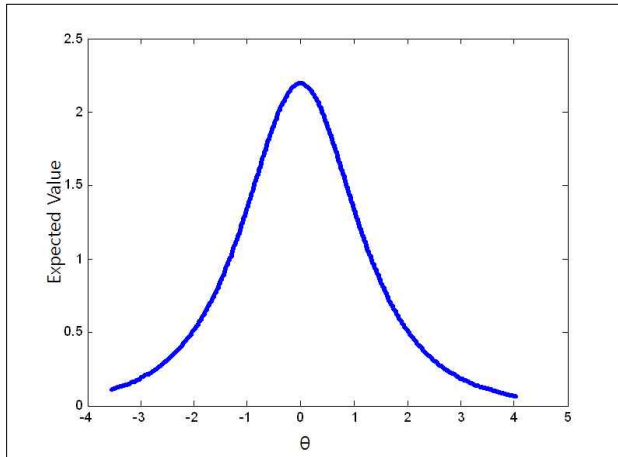
$$P(Z_i = z|\theta_j) = \frac{\exp\left\{\alpha_i \left[ z(\theta_j - \delta_i) - \sum_{k=0}^z \tau_{ik} \right]\right\} + \exp\left\{\alpha_i \left[ (M-z)(\theta_j - \delta_i) - \sum_{k=0}^z \tau_{ik} \right]\right\}}{\sum_{w=0}^C \left\{ \exp\left\{\alpha_i \left[ w(\theta_j - \delta_i) - \sum_{k=0}^w \tau_{ik} \right]\right\} + \exp\left\{\alpha_i \left[ (M-w)(\theta_j - \delta_i) - \sum_{k=0}^w \tau_{ik} \right]\right\} \right\}} \quad (5)$$

최종적으로 일반화등급전개모형은 <그림10>과 같은 관찰가능범주응답의 확률을 나타낸다.



<그림9> 네 개 범주 문항의 관찰가능범주 반응곡선(Roberts et al., 2000)

매우 반대, 반대, 찬성, 매우 찬성에 각각 0, 1, 2, 3의 범주점수를 부여한 후, 피험자 특성에 따른 기대점수를 반응기대값곡선(response expected value curve: REVC)으로 나타내면 다음과 같은 단봉분포 형태의 전개기제 곡선이 도출된다.



<그림10> 네 개 범주 문항의 반응기대값곡선 (Roberts et al., 2000)

### Ⅲ. 연구방법

#### 1. 실제 자료 분석

##### 1) 자료의 구성

###### ① 자료의 출처 및 수집 목적

본 연구에서 정의적 특성 검사 데이터로 활용한 실제 자료는 한국청소년정책연구원의 2009 한국 아동·청소년 온라인게임 및 가족여가활동실태조사를 통하여 수집된 것으로, 보다 구체적으로 설문조사의 목적을 살펴보면 첫째, 청소년의 게임사용 및 중독 수준의 파악, 둘째, 청소년의 여가활동 실태 파악, 셋째, 청소년의 가족요인 현황 파악을 들 수 있다. 이기봉 외(2009)에 의하면, 청소년의 인터넷 게임중독의 원인은 게임 요인 이외에도 개인의 심리적 특성, 사회·환경적 요인에 의해 영향을 받는다고 알려져 있다(김교정 외, 2006; 이수진, 2004; 최명선 외, 2007). 한국청소년정책연구원은 이러한 청소년 게임중독 실태를 파악하고 가족과 함께하는 건전한 여가활동을 통하여 사회·환경적 요인에서 해결을 도모하기 위한 연구를 수행하였다. 이와 관련한 보고서, 설문지 및 코드북 등의 데이터는 한국청소년정책연구원 홈페이지 데이터 아카이브(<http://archive.hypi.re.kr>)에서 다운받을 수 있다.

## ② 표본의 구성

한국청소년정책연구원은 보다 효율적이고 신뢰로우며 타당한 한국청소년들의 실태파악을 위한 기초 통계자료의 구축을 위하여 해당 연구기관에서 추진한 조사 연구 과제들에 대하여 “2009 아동·청소년 통합조사”를 실시하였다. 그 결과 한국의 아동·청소년의 전반적인 실태를 파악할 수 있는 (제주도를 제외한) 전국 단위의 데이터를 구축할 수 있었다. 본 연구에서 활용한 데이터 역시 통합조사의 일환으로 수집된 자료로서, 제주도를 제외한 전국 단위의 초등학교 4학년부터 고등학교 3학년까지의 학생을 조사대상으로 하고 있다.

설문조사를 위하여 2008 교육통계연보를 표집틀로 한 확률표집 방법을 활용하였고, 12개(서울시/인천시/대전시/광주시/대구시/울산시/부산시/경기도/충청도/전라도/강원도/경상도)의 층화지역에 학생수에 비례한 목표표본수를 할당하여 1개 학교에서 1개 학년을 추출한 후, 1개 학급을 전원 조사하였다. 설문조사결과의 최종 분석에 포함된 조사대상자는 9,180명이었으나, 공개된 데이터는 중고등학생 대상의 6,499명의 데이터이며, 본 연구에서 직접적으로 활용한 게임중독 진단 문항에 대한 답변에 결측치가 있는 120명의 데이터를 제외한 총6,379명의 데이터를 활용하였다. 응답자의 구체적인 일반적 특성은 다음과 같다.

<표6> 게임중독 진단 검사 응답자의 일반적 특성

| 구분  |      | N     | %     |      |
|-----|------|-------|-------|------|
| 전체  |      | 6,379 | 100.0 |      |
| 성별  | 남자   | 3,413 | 53.5  |      |
|     | 여자   | 2,966 | 46.5  |      |
| 학교별 | 중학교  | 1학년   | 990   | 15.5 |
|     |      | 2학년   | 1,120 | 17.6 |
|     |      | 3학년   | 1,028 | 16.1 |
|     | 고등학교 | 1학년   | 978   | 15.3 |
|     |      | 2학년   | 1,121 | 17.6 |
|     |      | 3학년   | 1,142 | 17.9 |
| 권역별 | 특별시  | 1,021 | 16.0  |      |
|     | 광역시  | 2,173 | 34.1  |      |
|     | 시군   | 3,185 | 49.9  |      |

③ 검사도구의 구성

2009 한국 아동·청소년 온라인게임 및 가족여가활동실태조사에서 중등학생을 대상으로 실시한 설문조사 내용은 8개의 영역으로 구성되어 있다. 게임관련 배경질문, 게임중독 진단, 여가활동 배경질문, 여가 장애, 여가 만족, 가족 기능, 부모와 자녀 간 의사소통 및 일반적 특성으로 구체적인 내용은 다음의 표와 같다(이기봉 외, 2009).

<표7> 온라인게임 및 가족여가활동실태조사의 구성 내용

| 대분류        | 조사영역                             | 문항수 |
|------------|----------------------------------|-----|
| 게임관련 배경질문  | 일 평균 게임 시간, 게임 빈도, 게임 장소 등       | 8   |
| 게임중독 진단    | 게임지향적 생활                         | 5   |
|            | 내성과 통제력 상실                       | 8   |
|            | 금단과 정서 경험                        | 7   |
| 여가활동 배경질문  | 여가 유형, 희망 여가활동, 여가활동 장소 및 동행인물 등 | 9   |
| 여가 장애      | 주관적 장애                           | 6   |
|            | 여가 선호와 참여 관계에 따른 장애              | 7   |
| 여가만족       | 심리적 만족                           | 4   |
|            | 교육적 만족                           | 4   |
|            | 사회적 만족                           | 3   |
|            | 휴식적 만족                           | 4   |
|            | 생리적 만족                           | 3   |
|            | 환경적 만족                           | 2   |
| 가족 기능      | 가족 응집력                           | 10  |
|            | 가족 적응력                           | 10  |
| 부모-자녀 의사소통 | 개방형 의사소통                         | 10  |
|            | 문제형 의사소통                         | 10  |
| 일반적 특성     | 성별, 학교급, 부모학력, 맞벌이여부 등           | 13  |

출처: 이기봉 외(2009)

이 중 본 연구에서 활용한 데이터는 게임중독을 진단한 20개의 문항에 관한 응답자료로서 검사도구는 한국정보화진흥원(2006)에서 개발하여 한국정보화진흥원의 인터넷중독대응센터의 홈페이지에서 활용하는 청소년용 인터넷 게임중독 척도로서, ‘전혀 그렇지 않다’, ‘때때로 그렇다’, ‘자주 그렇다’,

‘항상 그렇다’의 4점 리커트척도로 구성되어 있다.

한국정보화진흥원은 게임중독 척도의 하위요인을 ‘게임지향적 생활’, ‘내성과 통제력 상실’, ‘금단과 정서 경험’으로 구분하였다. ‘게임지향적 생활’은 실제생활보다 게임을 하는 가상공간에 집착하며 의존하는 정도를 측정하며, ‘내성과 통제력 상실’은 온라인 게임에 할애하는 시간과 관련한 영역을 측정하고, ‘금단과 정서 경험’은 게임행위를 하지 못하게 된 경우 나타나는 부정적 증상의 정도를 측정한다. 한국정보화진흥원(2006)이 제시한 각 하위영역별 구체적인 문항내용은 다음과 같다.

<표8> 청소년용 인터넷 게임중독 진단 Likert척도 검사 도구

| 하위영역       | 문항내용                                   | 전혀 그렇지 않다 | 때때로 그렇다 | 자주 그렇다 | 항상 그렇다 |
|------------|--|-----------|---------|--------|--------|
| 게임 지향적 생활  | 1. 게임을 하는 것이 친한 친구들과 어울리는 것보다 더 좋다.    | 1         | 2       | 3      | 4      |
|            | 2. 게임공간에서의 생활이 실제생활보다 더 좋다.            | 1         | 2       | 3      | 4      |
|            | 3. 게임 속의 내가 실제의 나보다 더 좋다.              | 1         | 2       | 3      | 4      |
|            | 4. 게임에서 사귀 친구들이 실제 친구들과보다 나를 더 알아준다.   | 1         | 2       | 3      | 4      |
|            | 5. 게임에서 사람을 사귀는 것이 더 편하고 자신 있다.        | 1         | 2       | 3      | 4      |
| 내성과 통제력 상실 | 6. 밤 늦게까지 게임을 하느라 시간 가는 줄 모른다.         | 1         | 2       | 3      | 4      |
|            | 7. 게임을 하느라 해야 할 일을 못한다.                | 1         | 2       | 3      | 4      |
|            | 8. 갈수록 게임을 하는 시간이 길어진다.                | 1         | 2       | 3      | 4      |
|            | 9. 점점 더 오랜 시간 게임을 해야 만족하게 된다.          | 1         | 2       | 3      | 4      |
|            | 10. 게임을 그만두어야 하는 경우에도 게임을 그만두는 것이 어렵다. | 1         | 2       | 3      | 4      |
|            | 11. 게임 하는 시간을 줄이려고 노력하지만 실패한다.         | 1         | 2       | 3      | 4      |
|            | 12. 게임을 안 하겠다고 마음먹고도 다시 게임을 하게 된다.     | 1         | 2       | 3      | 4      |
|            | 13. 게임 생각 때문에 공부에 집중하기 어렵다.            | 1         | 2       | 3      | 4      |
| 금단과 정서 경험  | 14. 게임을 못한다는 것은 견디기 힘든 일이다.            | 1         | 2       | 3      | 4      |
|            | 15. 게임을 하지 않을 때에도 게임 생각을 하게 된다.        | 1         | 2       | 3      | 4      |
|            | 16. 게임으로 인해 생활에 문제가 생기더라도 게임을 해야 한다.   | 1         | 2       | 3      | 4      |
|            | 17. 게임을 하지 못하면 불안하고 초조하다.              | 1         | 2       | 3      | 4      |
|            | 18. 다른 일 때문에 게임을 못하게 될까봐 걱정된다.         | 1         | 2       | 3      | 4      |
|            | 19. 누가 게임을 못 하게하면 신경질이 난다.             | 1         | 2       | 3      | 4      |
|            | 20. 게임을 못하게 되면 화가 난다.                  | 1         | 2       | 3      | 4      |

출처: 한국정보화진흥원(2006)

## 2) 자료의 분석

### ① 검사도구의 타당도와 신뢰도 분석 및 기술통계

게임중독 척도의 양호도를 분석하고자 SPSS 18.0 ver.을 활용하여 검사도구의 타당도와 신뢰도를 분석하고 기술통계치를 산출하였다. 먼저 타당도를 분석하기 위하여, 응답자들의 문항반응 데이터가 검사의 세 가지 하위척도인 ‘게임지향적 생활’, ‘내성과 통제력 상실’, ‘금단과 정서 경험’을 나타내는지 확인하고자 주성분분석과 요인분석을 수행하였다. 구체적인 절차를 살펴보면, 해당 데이터가 요인분석에 적합한지의 여부를 확인하기 위하여 Kaiser-Mayer-Olkin(KMO) 및 Bartlett 구형성 검정을 실시하였고, 요인의 수를 결정하기 위하여 주성분분석을 실시하였으며, 요인 간 상관관계 여부를 살펴본 후 요인의 회전방식을 결정하여 요인분석을 수행하였다.

게임중독 척도의 신뢰도는 Cronbach- $\alpha$ 를 통하여 분석되었으며, 게임중독 척도의 총점을 기준으로 판별되는 인터넷게임 사용자 유형에 대한 성별, 학교급별 기술통계치를 산출하였고, 기술통계치에 대한 t-검증을 실시하였다.

### ② 일반화등급전개모형과 일반화부분점수모형을 통한 문항 분석 및 해석

전개기제 모형의 일반적 형태를 갖춘 일반화등급전개모형과 누적기제 모형으로서 대표적으로 활용되는 일반화부분점수모형을 통하여 20개 문항에 대한 응답자들의 데이터를 분석하여 문항모수 및 피험자모수를 추정하였다. 일반적으로 일반화부분점수모형을 통한 문항응답 데이터는 주변최대우도 (marginal maximum likelihood; MML)법을 사용하는 GGUM2004(Roberts

& Cui, 2004) 컴퓨터 프로그램이 적용될 수 있고, 일반화부분점수모형은 MML 방법에 EM알고리즘을 더한 MMLE 추정법을 통하여 PARSCALE(프로그램을 적용하여 문항모수 및 피험자모수를 추정한다. 하지만 두 모형의 비교의 공정성을 위하여 동일한 추정법을 활용할 필요성이 제기되었고, 복잡한 다분문항모형에 따른 안정적인 모수추정의 필요성까지 고려하여 마르코프 연쇄 몬테 칼로 방법(Markov chain Monte Carlo; MCMC)을 통한 문항모수 및 피험자모수 추정을 수행하였다. MCMC 추정 방법은 베이저안 통계학에 기반한 추정법으로 Patz et al.(1999)에 의하여 다층수준, 다중수준 구조를 포함한 복잡한 데이터 구성을 탄력적으로 수용한다는 것이 입증되었고, Torre et al.(2006)에 의하여 일반화등급전개모형에 기반한 MCMC 추정법과 MML 추정법을 비교연구하여 MCMC 추정법이 전개기제모형에서도 타당하게 활용될 수 있음이 입증되기도 하였다. 국내에서는 강태훈, 김동일(2010)에 의하여 일반화등급전개모형과 일반화부분점수모형의 모수추정이 타당하게 이루어질 수 있음이 밝혀졌다.

MCMC 추정을 위해 사용된 컴퓨터 프로그램은 WinBUGS(Spiegelhalter et al., 2003) 1.4.3 ver.로서 MCMC 방법으로 베이저안 모형을 적합시키는 통계 패키지 프로그램이다. 최초로 캠브리지 대학에서 BUGS(Bayesian Inference Using Gibbs Sampling)로 개발된 소프트웨어를 Windows 환경에서 사용할 수 있도록 개발된 프로그램이다.

주어진 데이터에 관한 두 문항반응이론의 적합성을 비교하기 위하여, 통계적 모형 중 간명하면서도 자료의 적합도가 뛰어난 모형에 관한 정보를 제공하는 deviance information criterion(이하 DIC; Spiegelhalter et al., 2002)를 고려하였고, DIC는 다음과 같은 공식을 활용하여 계산된다(Kang et al., 2009).

$$DIC = \overline{D(\theta)} + pD \quad (6)$$

DIC는 모형이 데이터에 얼마나 잘 적합하는지에 대한 측정치인  $\overline{D(\theta)}$ 와 모형이 복잡해짐에 따라 모형선택지수에 패널티를 가하도록 구성된, 모형의 자유도에 기반한  $pD$ 로 구성되어 있다.  $pD = \overline{D(\theta)} - D(\bar{\theta})$ 로서,  $D(\bar{\theta})$ 는 모수의 사후 추정치에서의 편차이며,  $\bar{\theta}$ 는  $\theta$ 의 기댓값이다.

DIC 지수는 Winbugs에 의하여 제공되며, 그 수치가 작을수록 자료를 보다 효율적으로 설명함을 의미한다.  $\overline{D(\theta)}$ 는 모형의 모수가 늘어날수록 모형의 설명력이 늘어나기 때문에 작아지지만,  $pD$ 는 모형의 모수가 커질수록 함께 커지기 때문에 DIC는 적합도를 모형의 모수에 따라 효율적으로 나타내는 지수라고 볼 수 있다.

## 2. 모의실험을 통한 자료 분석

### 1) 모의실험 설계

일반화등급진개모형을 정의적 특성 검사에 적용하기 위한 추가적 검증을 실시하기 위하여, 게임중독 척도의 실제 데이터를 통하여 추출한 변별도, 경계모수 및  $N(0,1)$ 의 분포에서 표집한 문항 난이도모수를 생성 문항 모수로 활용하여 모의실험을 수행하였다. 실제 데이터를 통해 추출한 문항의 난이도가 다소 높은 수준에 편포되어 있어, 고른 난이도를 가진 검사도구를 가정하기 위하여 별도의 난이도 모수를 표집하였다.

모의실험 조건은 생성모형과 피험자 수에 따라 각각 2가지 경우를 상정하여 총 4개의 조건으로 구성되었고, 피험자 특성의 확률분포는  $N(0,1)$ 을 따

른다는 가정하에 표집되었다. 각 조건마다 10개의 모의실험 자료를 생성한 뒤 두 모형 모두를 통하여 분석하였으며, 보다 구체적으로 모의실험 조건을 살펴보면 다음과 같다.

첫째, 일반화등급전개모형과 일반화부분점수모형에 의하여 생성된 Likert 척도 형식의 자료를 각각의 대응하는 모형이 적절히 추정하는지 알아보기 위하여 두 가지의 생성모형을 모의실험 조건으로 구성하였다. 즉 일반화등급전개모형과 일반화부분점수모형에 의하여 각각의 자료는 생성되었다.

다음으로, 생성모형에 따른 능력 추정의 정확성과 함께 피험자 수가 이에 미치는 영향을 고려하기 위하여 500명, 1000명의 피험자 수 조건이 반영되었다. Roberts et al.(2000)에 따르면 일반화등급전개모형의 모수 복원 연구를 수행한 결과, 750명의 표본 크기에서 정확한 모수 추정 결과를 얻었다고 보고 하였기 때문에, 부족한 표본 크기 500명과, 충분한 표본 크기 1000명을 고려하였다. 이와 같은 모의실험 조건은 <표9>와 같이 요약될 수 있다.

<표9> 모의실험 조건

|         | 생성모형   |        |        |        |
|---------|--------|--------|--------|--------|
|         | GGUM   |        | GPCM   |        |
| 사례수     | 500    | 1000   | 500    | 1000   |
| 문항수     | 20     | 20     | 20     | 20     |
| 피험자특성분포 | N(0,1) | N(0,1) | N(0,1) | N(0,1) |
| 추정모형    | GGUM   | GPCM   | GGUM   | GPCM   |

## 2) 피험자 모수 추정 및 모형선택지수

일반화등급전개모형과 일반화부분점수모형을 적용하여 문항 및 피험자 모수와 모형선택지수를 추정하기 위하여, 실제 데이터의 분석과 마찬가지로 Winbugs프로그램 1.4ver.(Spiegelhalter, Thomas, Best, & Lunn, 2003)을 통하여 MCMC방법을 활용하였다. 피험자 모수에 대한 복원 연구를 수행하기 위하여 자료 분석 결과 도출된 두 모형에 의한 피험자 모수 추정치와, 진(true)능력 모수치, 모의실험 자료의 원점수 간의 상관계수를 산출하였고, 복원 연구에 부합한 모형선택지수가 도출되었는지 부가적으로 살펴보았다.

## IV. 연구결과

### 1. 실제 자료 분석 결과

#### 1) 검사도구의 타당도와 신뢰도 분석 및 기술통계

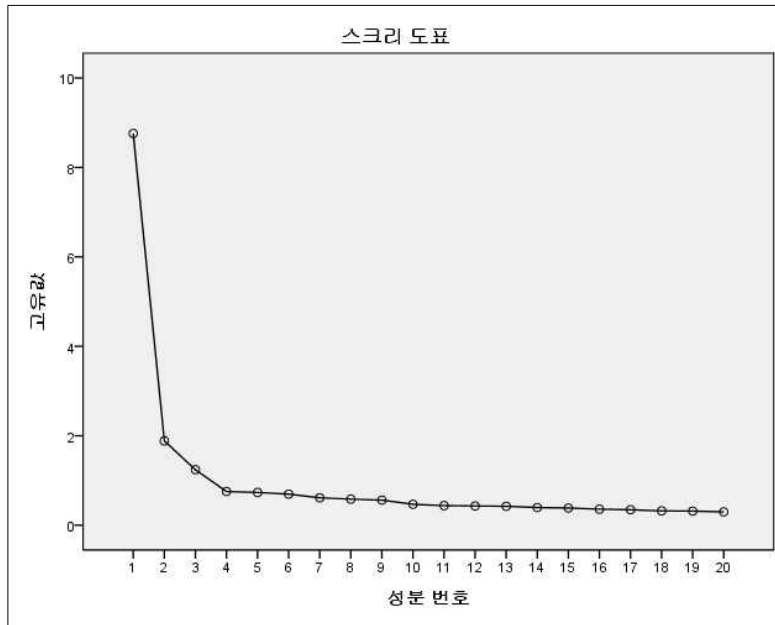
게임중독 척도의 타당성을 분석하기 위하여 요인분석을 수행하기 위하여 해당 데이터가 요인분석에 적합한지 KMO 검정을 수행한 결과, 그 값이 .950으로 도출되었다. Kaiser(1974)에 의하면 KMO값이 1에 가까울수록 표본의 상관이 요인분석을 수행하기에 적합하다는 것으로, .9 이상이면 좋고, .8정도이면 양호한 수준이 되, .6이나 .7은 보통, 마지막으로 .5이하이면 부적합하다고 판정한다. 따라서 본 데이터의 KMO 검정결과는 매우 양호하다고 볼 수 있다. 다음으로 Bartlett 구형성 검정치를 고려한 결과,  $\chi^2 = 65684.3$ ,  $df=190$ ,  $p<.001$ 로 산출되었다. 이 값은 주어진 상관행렬자료가 단위행렬이라는 영가설을 기각하여, 요인분석을 수행하기에 적절하다고 판단할 수 있는 결과라고 볼 수 있다.

게임중독 척도가 담고 있는 요인의 수를 결정하기 위하여 주성분분석을 실시하여 계산된 누적분산비율은 다음과 같다.

<표10> 게임중독 진단 응답자료의 주성분분석에 따른 요인별 누적분산비율

| 요인 | 합계    | % 분산   | % 누적    |
|----|-------|--------|---------|
| 1  | 8.759 | 43.794 | 43.794  |
| 2  | 1.886 | 9.431  | 53.225  |
| 3  | 1.241 | 6.204  | 59.429  |
| 4  | .753  | 3.765  | 63.194  |
| 5  | .732  | 3.660  | 66.854  |
| 6  | .696  | 3.480  | 70.335  |
| 7  | .613  | 3.064  | 73.399  |
| 8  | .584  | 2.918  | 76.316  |
| 9  | .560  | 2.800  | 79.116  |
| 10 | .468  | 2.340  | 81.456  |
| 11 | .438  | 2.190  | 83.646  |
| 12 | .430  | 2.152  | 85.799  |
| 13 | .423  | 2.115  | 87.913  |
| 14 | .396  | 1.978  | 89.891  |
| 15 | .385  | 1.924  | 91.816  |
| 16 | .358  | 1.792  | 93.607  |
| 17 | .346  | 1.729  | 95.336  |
| 18 | .320  | 1.602  | 96.938  |
| 19 | .315  | 1.577  | 98.515  |
| 20 | .297  | 1.485  | 100.000 |

요인은 측정변수의 수만큼 계산되기 때문에 문항수인 20개의 수만큼 계산되었다. 각 요인의 분산비율을 보면 첫 번째 요인은 43.8%를, 두 번째 요인은 9.4%를, 세 번째 요인은 6.2%를 설명함을 알 수 있고, 네 번째 요인부터 3.7%의 분산을 설명하며 이하의 요인으로 갈수록 비슷한 분산의 크기만큼 설명함을 알 수 있다. 또한 주성분 분석 결과 도출된 스크리 도표는 다음과 같다.



<그림11> 게임중독 진단 응답자료의 주성분 분석에 따른 스크리 도표

스크리 도표에 따르면, 네 번째 요인부터 그 이하의 요인까지 비슷한 고유값을 지님을 볼 수 있다. 요인의 수를 결정할 때에는 스크리도표, 누적분산비율, 해석가능성을 기준으로 삼을 수 있는데(이순목, 2010), 해석가능성은 자료분석 이전에 연구자가 몇 개의 요인이 “해석가능한”지 고려할 수 있음을 의미한다. 본 게임중독 척도가 개발될 때, ‘게임지향적 생활’, ‘내성과 통제력 상실’, ‘금단과 정서경험’이라는 세가지 하위척도로 개발되었음을 고려하여 세 가지 요인이 해석가능하고, 누적분산비율과 스크리 도표를 고려하였을 때에도 총3가지 요인이 존재한다고 볼 수 있겠다. 요인 수를 3개로 결정한 후, 요인간 상관여부를 고려하기 위하여 요인의 회전방식을 promax로 하여 요인분석을 한 결과 다음과 같은 요인간 상관계수와 사교회전에 따른 요인분석 결과가 도출되었다.

<표11> 게임중독 진단 검사의 요인간 상관행렬

|     | 요인1   | 요인2   | 요인3   |
|-----|-------|-------|-------|
| 요인1 | 1.000 |       |       |
| 요인2 | .666  | 1.000 |       |
| 요인3 | .444  | .525  | 1.000 |

<표12> 게임중독 진단 검사의 요인분석 결과

|      | 요인    |       |       | 공통분  |
|------|-------|-------|-------|------|
|      | 요인1   | 요인2   | 요인3   |      |
| 문항1  | .170  | -.111 | .628  | .432 |
| 문항2  | -.002 | .046  | .760  | .615 |
| 문항3  | -.131 | .109  | .788  | .629 |
| 문항4  | -.013 | -.023 | .815  | .636 |
| 문항5  | .041  | -.022 | .774  | .610 |
| 문항6  | .779  | -.095 | .046  | .546 |
| 문항7  | .878  | -.092 | -.016 | .661 |
| 문항8  | .799  | -.049 | .035  | .613 |
| 문항9  | .565  | .190  | .067  | .549 |
| 문항10 | .762  | .075  | -.057 | .624 |
| 문항11 | .802  | .012  | -.030 | .635 |
| 문항12 | .761  | .040  | .004  | .624 |
| 문항13 | .517  | .252  | .024  | .522 |
| 문항14 | .142  | .674  | .002  | .603 |
| 문항15 | .260  | .519  | .039  | .549 |
| 문항16 | -.040 | .772  | .054  | .602 |
| 문항17 | -.165 | .888  | .038  | .651 |
| 문항18 | -.008 | .804  | -.021 | .620 |
| 문항19 | .211  | .614  | -.051 | .554 |
| 문항20 | .007  | .803  | -.050 | .612 |
| 고유치  | 7.370 | 7.220 | 5.085 |      |

요인간 상관을 고려하여, 게임중독척도를 사교회전 한 결과, 요인1은 문항 6 - 문항13을, 요인2는 문항14 - 문항20을, 요인3은 문항1 - 문항5를 포함하고 있음이 나타났다. 게임중독 척도 개발시, 고려하였던 하위척도를 반영하여 요인을 해석하여 보면, 요인1은 ‘내성과 통제력 상실’을, 요인2는 ‘금단과 정서경험’을, 요인3은 ‘게임지향적 생활’을 나타낸다고 볼 수 있다. 보통 관찰변수를 일정한 요인으로 분류할 때, .3 이상이 일반적인 기준으로 받아들여짐(강태훈 외, 2013)을 고려한다면 해당 게임중독 척도는 .517 - .888의 높은 요인값을 보임에 따라 명확한 하위척도를 구성하며 척도의 타당성을 확보한다고 사료된다. 해당 척도의 신뢰도를 분석하기 위하여 Cronbach- $\alpha$  값을 산출한 결과, .929의 높은 값을 보였다.

게임 중독을 진단하기 위한 네 개의 범주, “전혀 그렇지 않다”, “때때로 그렇다”, “자주 그렇다”, “항상 그렇다”에 각각 1, 2, 3, 4의 점수를 부여한 후 산출한 기술통계치는 다음과 같다.

<표13> 게임중독 진단 검사 응답자의 성별 및 학교수준에 따른 기술통계

|    | 구분   | 사례수   | 평균<br>(표준편차) | t/유의수준     |
|----|------|-------|--------------|------------|
| 전체 | -    | 6,379 | 26.57(8.11)  | -          |
| 성별 | 남자   | 3,413 | 29.13(8.64)  | 29.37/0.00 |
|    | 여자   | 2,966 | 23.62(6.27)  |            |
| 학교 | 중학교  | 3,138 | 26.95(8.13)  | 3.71/0.00  |
|    | 고등학교 | 3,241 | 26.20(8.08)  |            |

총 6379명의 응답자의 최소점수는 20점이었고 최대점수는 80점인 가운데, 평균점수는 26.57로 비교적 낮은 점수를 보였다. 남학생 3,413명과 여학생

2,966명의 평균은 각각 29.13과 23.62였으며 남학생의 표준편차(8.64)는 여학생의 표준편차(6.27)보다 더 높은 경향을 나타내었고, 원점수 역시 남학생이 여학생에 비하여 통계적으로 유의미하게 높았다. 학교수준에 따른 평균점수 차이를 본 결과, 중학교 학생들이 평균 26.95를 보였고 고등학교 학생들은 평균 26.20을 보였고, 중학교 학생들의 원점수가 통계적으로 유의미하게 높았다.

## 2) 일반화등급전개모형과 일반화부분점수모형을 통한 문항 분석 및 문항 해석

일반화등급전개모형과 일반화부분점수모형의 적용가능성을 본격적으로 비교하기 위하여, 실제 게임중독 데이터에 두 모형을 적용하여 문항모수와 모형 선택 지수를 도출한 결과는 <표13>과 같다.

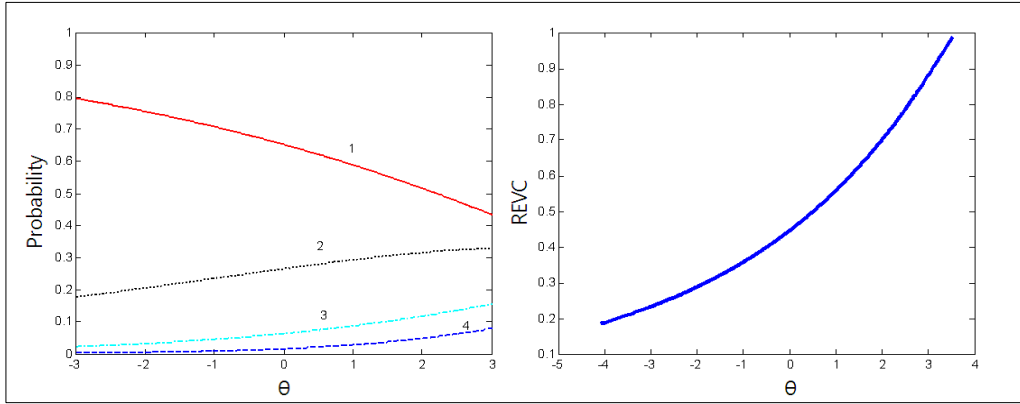
일반화등급전개모형과 일반화부분점수모형의 모형선택지수, DIC를 살펴본 결과, 일반화등급전개모형의 DIC는 139251.00으로, 일반화부분점수모형의 DIC는 130604.00으로 나타남으로써, 해당 데이터는 일반화부분점수모형에 의하여 보다 더 적합하게 설명됨을 알 수 있었다.

일반화등급전개모형으로 추정된 피험자 특성 추정치의 평균은 -1.20으로 일반화부분점수모형에 의한 피험자 특성 추정치의 평균, -0.12보다 높은 수준이었으며, 두 모형에 의한 피험자 특성 추정치 간의 적률상관계수의 절대값은 0.951로, 높은 상관의 정도를 보였다.

<표14> 일반화등급전개모형과 일반화부분점수모형에 따른 문항모수치 및 DIC

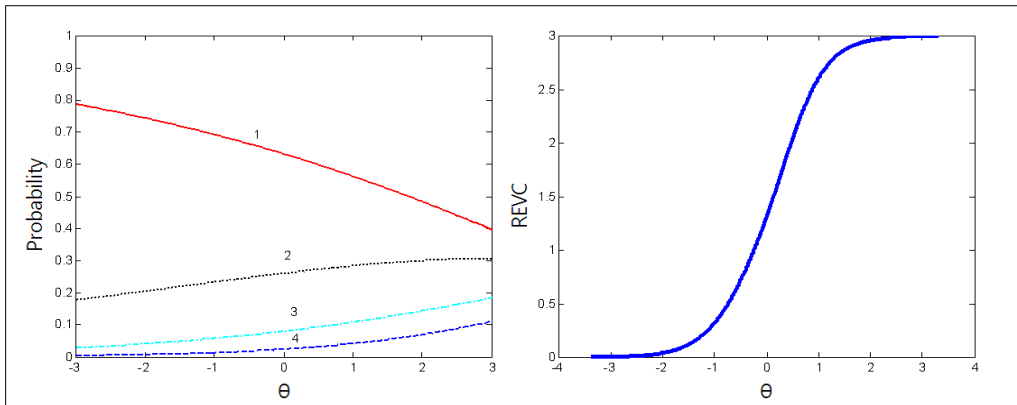
| 문항  | 일반화등급전개모형 |       |       |       |       | 일반화부분점수모형 |      |      |       |       |
|-----|-----------|-------|-------|-------|-------|-----------|------|------|-------|-------|
|     | ai        | δi    | τi1   | τi2   | τi3   | ai        | δi   | τi1  | τi2   | τi3   |
| 1   | 1.21      | -1.51 | -0.53 | 0.86  | 0.19  | 1.05      | 0.71 | 0.81 | -0.82 | 0.01  |
| 2   | 0.20      | 6.13  | 1.43  | 2.07  | -0.46 | 1.28      | 0.68 | 0.41 | -0.32 | -0.10 |
| 3   | 1.39      | -1.59 | 0.12  | 0.42  | 0.25  | 1.24      | 0.90 | 0.18 | -0.18 | 0.01  |
| 4   | 0.20      | 6.09  | 3.63  | 2.44  | -0.43 | 1.32      | 0.86 | 0.21 | -0.24 | 0.03  |
| 5   | 1.41      | -1.54 | -0.19 | 0.29  | 0.39  | 1.24      | 0.74 | 0.42 | -0.16 | -0.26 |
| 6   | 2.23      | -1.36 | -0.79 | -0.09 | 0.25  | 1.63      | 0.27 | 0.74 | -0.16 | -0.58 |
| 7   | 2.97      | -1.32 | -1.05 | -0.13 | 0.23  | 2.18      | 0.12 | 0.98 | -0.23 | -0.75 |
| 8   | 0.20      | 6.47  | -1.97 | 0.72  | 1.30  | 2.17      | 0.23 | 0.79 | -0.19 | -0.60 |
| 9   | 2.95      | -1.45 | -0.70 | -0.08 | 0.12  | 2.50      | 0.30 | 0.61 | -0.16 | -0.45 |
| 10  | 2.97      | -1.38 | -0.88 | -0.14 | 0.11  | 2.31      | 0.16 | 0.75 | -0.20 | -0.55 |
| 11  | 2.97      | -1.36 | -0.71 | -0.12 | 0.10  | 2.44      | 0.21 | 0.61 | -0.15 | -0.47 |
| 12  | 0.20      | 6.40  | -1.96 | -0.43 | 0.56  | 2.34      | 0.09 | 0.71 | -0.14 | -0.57 |
| 13  | 2.92      | -1.44 | -0.75 | -0.04 | 0.14  | 2.28      | 0.33 | 0.69 | -0.21 | -0.48 |
| 14  | 2.98      | -1.62 | -0.67 | -0.15 | 0.09  | 2.74      | 0.37 | 0.50 | -0.10 | -0.41 |
| 15  | 2.95      | -1.55 | -0.85 | -0.05 | 0.15  | 2.42      | 0.35 | 0.74 | -0.23 | -0.50 |
| 16  | 2.99      | -1.60 | -0.26 | 0.05  | 0.29  | 3.48      | 0.65 | 0.45 | 0.00  | -0.46 |
| 17  | 2.99      | -1.65 | -0.24 | 0.17  | 0.11  | 3.41      | 0.68 | 0.41 | -0.12 | -0.30 |
| 18  | 2.98      | -1.63 | -0.47 | 0.01  | 0.15  | 2.91      | 0.54 | 0.47 | -0.10 | -0.36 |
| 19  | 2.67      | -1.67 | -1.06 | -0.18 | -0.06 | 2.02      | 0.22 | 0.70 | -0.33 | -0.38 |
| 20  | 2.93      | -1.69 | -0.69 | -0.06 | -0.04 | 2.23      | 0.44 | 0.45 | -0.24 | -0.21 |
| DIC | 139251.00 |       |       |       |       | 130604.00 |      |      |       |       |

일반화등급전개모형에서 가장 높은 위치모수(6.47)를 보인 8번 문항은 “갈수록 게임을 하는 시간이 길어진다”는 진술문을 가진 문항으로, 위치모수가 매우 높음에 따라 일반화등급전개모형으로 추정하였음에도 누적기제를 보였다. 즉 피험자 특성 6.47의 수준까지 단조증가의 기댓값 곡선을 보여주기 때문에, 게임중독의 수준이 극단적으로 높은 수준을 제외하고 게임중독의 수준이 높아질수록 긍정적으로 응답할 확률이 높아지는 누적기제를 따른다고 볼 수 있다. 이러한 문항분석 결과에 비추어, 실제로 기본 의식주생활 및 학교수업시간을 제외한 모든 시간을 이미 게임에 소비하고 있기 때문에 더 이상 증가시킬 게임 시간이 없는 극단적인 경우가 아니라면, 게임 중독이 심할수록 게임하는 시간은 점차 늘어날 것이라고 예상해 볼 수 있다.



<그림12> 문항8. “갈수록 게임을 하는 시간이 길어진다”의 GGUM에 관한 ICCC와 REVC

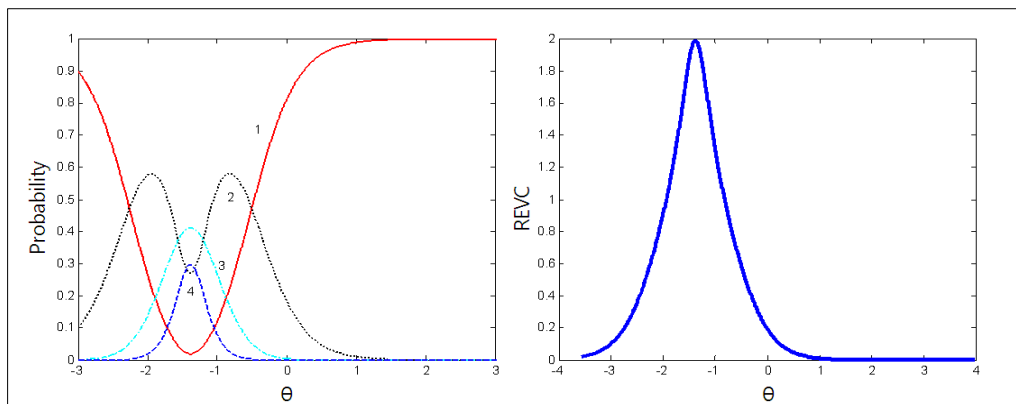
“게임을 안 하겠다고 마음먹고도 다시 게임을 하게 된다.”는 문항 12번 역시 대표 위치모수가 6.40으로 누적기제의 기댓값곡선을 보여주고 있다. 게임에 대한 중독 정도가 병리적일 만큼 높아서 게임을 안 하겠다고 마음먹지 않는 경우를 제외한, 대부분의 피험자들이 게임중독 수준이 높을수록 게임을 안 하겠다고 마음먹고도, 생각과는 다르게 다시 게임을 하게 된다고 응답하였음을 알 수 있다. 게임을 안 하겠다는 생각까지 못할 만큼 게임에 심각하게 중독된 피험자는 극히 드물기 때문에, 일반적인 피험자 특성 수준에서 해당 문항은 누적기제를 따르는 것으로 보인다.



<그림13> 문항12. “게임을 안 하겠다고 마음먹고도 다시 게임을 하게 된다”의 GGUM에 관한 ICCC와 REVC

마찬가지로 각각 6.13, 6.09의 위치모수를 보인 문항2, 4 역시 누적기제를 따르는 반응기대값곡선을 나타냈음을 알 수 있다(부록 참조).

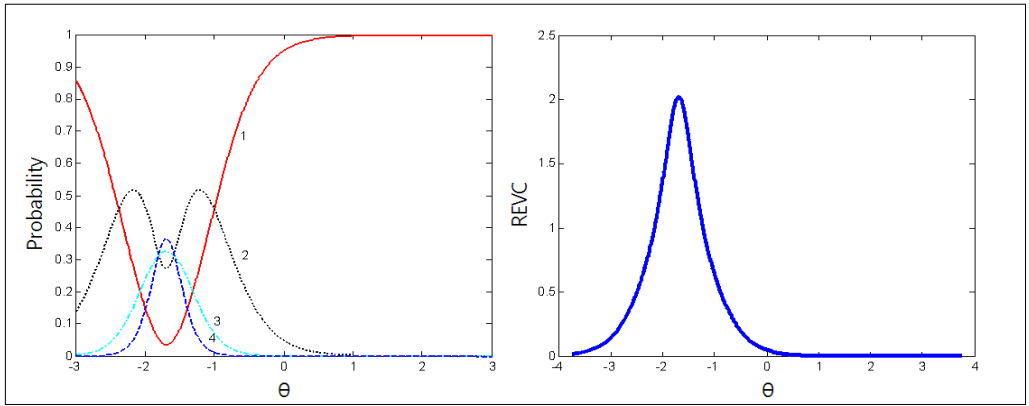
위치모수가 -1.38인 문항10은 “게임을 그만두어야 하는 경우에도 게임을 그만두는 것이 어렵다”는 진술문을 가졌으며, -1.38 이상의 특성 수준을 가진 피험자는 게임 중독의 수준이 높을수록 긍정적 수준으로 해당 문항에 응답할 확률이 점차 낮아진다. 일정수준까지의 게임중독의 경우에는 게임을 그만두어야 할 경우, 게임을 그만두는 것이 어렵게 느껴지지만, 특정 수준을 넘은 정도로 게임중독이 심각할 경우, 반드시 게임을 그만두어야 하는 피치 못할 상황이 아니라면, 그만둘 생각조차 하지 않기 때문에 오히려 게임을 그만두어야 하는 상황에 쉽게 게임을 그만둘 수 있을 것으로 예상해 볼 수 있다.



<그림14> 문항10. “게임을 그만두어야 하는 경우에도 게임을 그만두는 것이 어렵다”의 GGUM에 관한 ICC와 REVC

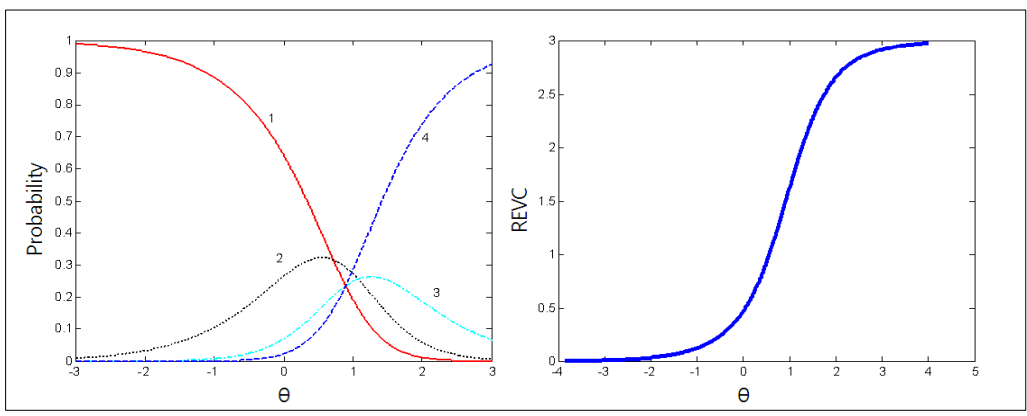
일반화등급전개모형을 적용하였을 때 가장 낮은 위치모수를 보인 문항은 20번, “게임을 못하게 되면 화가 난다”는 진술문으로서 -1.69의 곤란도 모수를 나타낸다. 이 경우 역시 게임중독 수준이 일정 수준을 넘어 극단적인

수준으로 높을 경우, 게임을 못하게 되었을 때 화가 나는 정도가 아니라 기운이 없거나 삶의 의욕을 느끼지 못하는 경우 등과 같이 보다 더 심각한 증상을 나타낼 수도 있다는 예상을 가능하게 한다.



<그림15> 문항20. “게임을 못하게 되면 화가 난다”의 GGUM에 관한 ICCC와 REVC

일반화부분점수모형에서 가장 높은 위치모수(0.90)를 보인 문항은 “게임 속의 내가 실제의 나보다 더 좋다”는 진술문을 가진 3번 문항으로, 게임중독의 수준이 높은 학생이어야 긍정적으로 응답할 수 있는 문항이었음이 밝혀졌다. 반면 문항 12번은 일반화부분점수모형에 대하여 0.09의 가장 낮은 위치모수를 보인, “게임을 안 하겠다고 마음먹고도 다시 게임을 하게 된다.”는 진술문의 문항이었다.



<그림16> 문항3. “게임 속의 내가 실제의 나보다 더 좋다”의 GPCM에 관한 ICCC와 REVC

## 2. 모의실험에 의한 자료 분석 결과

### 1) 피험자 모수 추정치

<표15>와 <표16>은 생성모형이 일반화등급전개모형일 경우, 표본수가 500명과 1000명일 경우에 대한 생성 피험자 모수, 원점수, 일반화등급전개모형에 의한 추정치( $\hat{\theta}_{GGUM}$ ), 일반화부분점수모형에 의한 추정치( $\hat{\theta}_{GPCM}$ ) 간의 상관계수의 평균 및 표준편차를 나타낸다.

먼저 <표15>를 통하여 피험자 생성 모수를 원점수와 두 모형에 의한 추정치가 각각 잘 대표하는지 살펴보면,  $\hat{\theta}_{GGUM}$ ,  $\hat{\theta}_{GPCM}$ , 원점수 순으로 생성 모수를 잘 복원하였으나, 일반화부분점수모형에 의한 피험자 특성 추정치와 원점수는 생성모수를 적절히 나타내고 있지 못하였다. 만약 피험자가 이상 점 과정에 기반하여 문항에 응답하여 전개기제를 따르는 문항응답 데이터가 존재할 경우, 원점수와 일반화부분점수모형에 의한 추정치는 피험자의 특성을 타당하게 나타낼 수 없음을 알 수 있다.

<표15> 일반화등급전개모형에 의한 자료 생성시 피험자 특성 생성모수에 대한 상관계수

| 표본크기<br>(N) | 피험자 특성 생성모수와의 상관계수 평균 및 표준편차( $\bar{r}(SD)$ ) |                       |                       |
|-------------|---|-----------------------|-----------------------|
|             | 원점수   | $\hat{\theta}_{GGUM}$ | $\hat{\theta}_{GPCM}$ |
| 500         | 0.09(0.04)                                    | 0.94(0.04)            | 0.434(0.07)           |
| 1000        | 0.05(0.04)                                    | 0.87(0.12)            | 0.348(0.09)           |

다음으로 <표16>을 통하여 일반화부분점수모형에 의하여 생성된 모의 자료의 원점수와  $\hat{\theta}_{GGUM}$ ,  $\hat{\theta}_{GPCM}$  간의 상관계수를 보면, 일반적으로 누적기제를

가정한 Likert 척도의 점수 산출방법과 동일한 방식으로 나타난 원점수와 누적기제의 대표적 다분문항반응모형인 일반화부분점수모형에 의한 추정치 간에 높은 상관관계를 볼 수 있었으나, 원점수와 일반화등급전개모형에 의한 피험자 특성 추정치 간의 상관관계는 상대적으로 매우 낮은 수준이었다. 흔히 Likert 척도를 이용한 정의적 특성 검사를 수행하고 난 후, 원점수에 기반하여 피험자의 특성 수준을 나타내지만, 만약 피험자가 이상점 과정에 기반하여 정의적 특성 검사에 대한 응답을 하였다면 원점수는 타당한 피험자의 특성 추정치가 될 수 없음을 보여준다.

<표16> 일반화등급전개모형에 의한 자료 생성시 원점수에 대한 상관계수

| 표본 크기(N) | 원점수와의 상관계수 평균 및 표준편차( $\bar{r}$ (SD)) |                       |
|----------|---------------------------------------|-----------------------|
|          | $\hat{\theta}_{GGUM}$                 | $\hat{\theta}_{GPCM}$ |
| 500      | 0.09(0.04)                            | 0.75(0.03)            |
| 1000     | 0.06(0.03)                            | 0.76(0.03)            |

<표17>과 <표18>은 일반화부분점수모형을 생성모형으로 하여 모의실험 자료를 생성 및 분석한 후 도출된 생성 피험자 모수, 원점수,  $\hat{\theta}_{GGUM}$ ,  $\hat{\theta}_{GPCM}$  간의 상관계수를 나타낸다. 일반화부분점수모형에 의한 피험자 생성 모수를 가장 대표할 수 있는 피험자 특성의 순서는  $\hat{\theta}_{GPCM}$ , 원점수,  $\hat{\theta}_{GGUM}$ 에 의한 피험자 특성 추정치였다.  $\hat{\theta}_{GPCM}$ 에 의한 피험자 특성 추정치는 0.98의 높은 수준의 상관계수를 보였으며, 원점수 역시 0.97이상의 높은 상관의 정도를 보였다. 하지만 전개기제를 따르는 일반화등급전개모형에 의한 특성 추정치는 생성 모수와 0.12, 0.30의 낮은 상관을 보였다.

<표17> 일반화부분점수모형에 의한 자료 생성시 피험자 특성 생성모수에 대한 상관계수

| 표본크기<br>(N) | 피험자 특성 생성모수와의 상관계수 평균 및 표준편차( $\bar{r}(SD)$ ) |                       |                       |
|-------------|---|-----------------------|-----------------------|
|             | 원점수   | $\hat{\theta}_{GGUM}$ | $\hat{\theta}_{GPCM}$ |
| 500         | 0.97(0.00)                                    | 0.12(0.06)            | 0.98(0.00)            |
| 1000        | 0.97(0.00)                                    | 0.30(0.20)            | 0.98(0.01)            |

누적기체에 의하여 생성된 자료의 원점수와 피험자 모수 추정치 간의 상관계수를 보면, 전개기체에 의한 피험자 모수 추정치와 원점수는 낮은 상관계수의 정도를 보이며, 누적기체에 의한 피험자 모수 추정치와 원점수는 높은 상관계수의 정도를 보임을 알 수 있다. 만약 정의적 특성에 대한 검사 문항에 피험자가 우세 반응 과정을 통하여 응답하였다면, 원점수는 피험자 특성 추정치로서 타당하게 기능할 수 있음을 보여준다.

<표18> 일반화부분점수모형에 의한 자료 생성시 원점수에 대한 상관계수

| 표본 크기(N) | 원점수와의 상관계수 평균 및 표준편차( $\bar{r}(SD)$ ) |                       |
|----------|---------------------------------------|-----------------------|
|          | $\hat{\theta}_{GGUM}$                 | $\hat{\theta}_{GPCM}$ |
| 500      | 0.12(0.07)                            | 0.99(0.00)            |
| 1000     | 0.30(0.21)                            | 0.99(0.01)            |

## 2) 모형선택지수

일반화등급전개모형과 일반화부분점수모형에 대하여 Winbugs를 통하여 MCMC방법에 따른 피험자특성 복원연구를 수행한 결과, 피험자가 완벽히 이상점 과정에 기반하여 문항에 응답하였을 경우, 일반화등급전개모형에 의한 피험자 특성 추정이 적절하고 우세 과정에 기반한 피험자의 응답 데이터

는 일반화부분점수모형에 의하여 분석되어야 적절하다는 것을 알 수 있었다. 이러한 복원 결과를 통계적 모형선택지수가 잘 나타내 주는지 살펴보기 위하여 DIC 지수를 산출한 결과는 <표19>와 <표20>과 같다. 이 경우 표본의 크기가 클수록 DIC 값 역시 커지는 것을 알 수 있는데, DIC의 경우, 표본의 크기가 증가함에 따라 그 값이 커질 수밖에 없기 때문이다.

생성모형과 관계없이 추정모형이 일반화등급전개모형일 경우 DIC값이 일반화부분점수모형의 것보다 더 작게 산출되었고, 이러한 모형선택지수의 도출결과는 피험자 특성 복원 연구에서 도출된 결과와 달랐다. 복원 연구의 경우, 생성모형과 추정모형이 부합할 때 가장 우수한 피험자 특성 복원이 이루어졌으나, DIC에 의하면 생성모형과 관계없이 일반화등급전개모형이 응답 데이터를 보다 효율적으로 설명한다는 추론이 가능하기 때문이다.

<표19> N=500일 경우의 DIC

| DIC                   |                  |                  |                  |
|-----------------------|------------------|------------------|------------------|
| 생성모형                  |                  |                  |                  |
| GGUM                  |                  | GPCM             |                  |
| 추정모형                  |                  | 추정모형             |                  |
| GGUM                  | GPCM             | GGUM             | GPCM             |
| 16177.0               | 19223.0          | 10470.7          | 14034.5          |
| 16734.3               | 19256.9          | 11751.3          | 13831.3          |
| 16427.9               | 19397.1          | 11216.8          | 13995.1          |
| 16035.1               | 18985.5          | 12041.1          | 14128.1          |
| 16299.2               | 19001.4          | 11147.8          | 13784.5          |
| 16283.6               | 18828.7          | 11524.4          | 14177.3          |
| 16121.0               | 18969.4          | 11340.7          | 14216.5          |
| 16296.6               | 19116.9          | 11922.7          | 14388.1          |
| 15221.4               | 19287.6          | 11478.9          | 14508.0          |
| 16085.6               | 18974.4          | 11574.2          | 14398.1          |
| $\overline{DIC}$ (SD) |                  |                  |                  |
| 16168.17(388.47)      | 19104.09(180.31) | 11446.86(447.46) | 14146.15(241.39) |

<표20> N=1000일 경우의 DIC

| DIC                   |                   |                  |                   |
|-----------------------|-------------------|------------------|-------------------|
| 생성모형                  |                   |                  |                   |
| GGUM                  |                   | GPCM             |                   |
| 추정모형                  |                   | 추정모형             |                   |
| GGUM                  | GPCM              | GGUM             | GPCM              |
| 32423.1               | 38092.0           | 23352.5          | 28655.4           |
| 33104.6               | 38200.5           | 23562.4          | 28485.1           |
| 33209.0               | 38314.0           | 23291.0          | 27890.5           |
| 33406.5               | 38751.8           | 23177.0          | 28166.2           |
| 34390.0               | 38063.4           | 25249.1          | 28092.8           |
| 32682.3               | 38408.6           | 24857.3          | 28623.8           |
| 32242.6               | 38002.9           | 23735.5          | 28419.7           |
| 32471.0               | 38350.4           | 22067.3          | 28578.8           |
| 32597.1               | 38298.1           | 23082.0          | 28509.8           |
| 32833.6               | 38018.5           | 23584.2          | 28542.2           |
| $\overline{DIC}$ (SD) |                   |                  |                   |
| 32935.98(631.42 )     | 38250.02(228.60 ) | 23595.83(898.97) | 28396.43(257.27 ) |

## V. 논의 및 결론

### 1. 전개기제에 기반한 정의적 특성 검사 문항의 해석

게임중독 진단 검사의 총 20개의 문항 중, 위치모수를 중심으로 몇몇의 문항에 대하여 전개기제를 적용한 결과를 바탕으로 피험자 특성과 관련한 문항의 해석을 시도한 결과, 누적기제에 적용한 문항 해석과 다른 차이점을 발견할 수 있었다.

첫째, 누적기제에 기반하여 단순히 피험자의 특성 수준이 높아질수록 해당 문항에 대한 긍정적인 응답 확률이 높아진다는 해석을 하기보다, 피험자의 특성 수준이 일정 수준 이상일 경우 오히려 긍정적인 응답 확률이 낮아질 수 있는 상황을 고려해 볼 수 있었다. 동일한 문항임에도 이는 보다 다양한 정보를 산출할 수 있는 근거를 제공할 수 있는 가능성을 보여주었다.

둘째, 누적기제에 기반한 일반화부분점수모형에 의하여 문항을 해석할 경우, 모든 문항특성곡선이 누적기제를 따르게 되지만, 전개기제에 기반한 일반화등급전개모형에 의하여 문항을 해석할 때에는 문항의 위치모수에 따라 각각의 문항이 전개기제를 따르기도 하고, 누적기제를 따르기도 하였다. 즉, 누적기제에 의한 문항해석은 일률적인 누적기제에 기반한 문항해석을 낳지만, 전개기제에 의한 문항해석은 전개기제 뿐만 아니라 누적기제 또한 나타냄으로써 상대적으로 풍부한 문항에 관한 정보를 제공하였다.

일반화등급전개모형의 차별성은 풍부한 정보를 제공할 수 있는 검사구성의 가능성을 도모하는 가운데, 해석의 의미가 타당도를 갖추기 위하여 피험자의 응답과정이 실제로 이상점 반응 과정에 기반하는지에 관한 연구가 필

요하다고 보인다. 일부 연구자들에 의하여 피험자들이 실제로 문항의 진술문과 자신의 태도 수준을 비교하여 정의적 특성 검사에 응답하는지, 응답한다면 어떠한 조건에서 그러한 응답 기제를 나타내는지에 관한 연구가 수행되었는데 Carter & Dalal(2010)는 문항의 진술문에 대하여 극단적으로 자세히 해석하거나 간단하게 해석하는 피험자의 경우 누적기제에 가까운 반응곡선을 보였고, 중간 수준의 해석을 수행하는 피험자는 보다 전개기제에 가까운 반응곡선을 나타내었다고 주장하였다. 같은 연구에 의하면 문항의 내용과 검사의 실시 형식, 검사가 측정하고자 하는 구인의 성질에 따라 피험자들의 응답 기제가 달라지기도 하였는데, 다른 개념에 비하여 보다 추상적인 구인이 유의미한 수의 전개기제 문항을 나타내었고, Carter, Dalal & Zickler(2010)에 의한 다른 연구에 의하면 수량형용사의 사용이 응답기제에 영향을 미치기도 하였다. 이와 같이 피험자들의 응답과정에 나타나는 이상점 기제에 관한 선행연구가 일부 존재하기는 하나, 심리학적 전개 기제에 관한 일반화된 연구결과는 미미한 실정으로(Carter et al., 2010), 전개기제를 결정하는 요인과 조건을 보다 일반화할 수 있는 포괄적인 후속 연구가 요청되는 바이다.

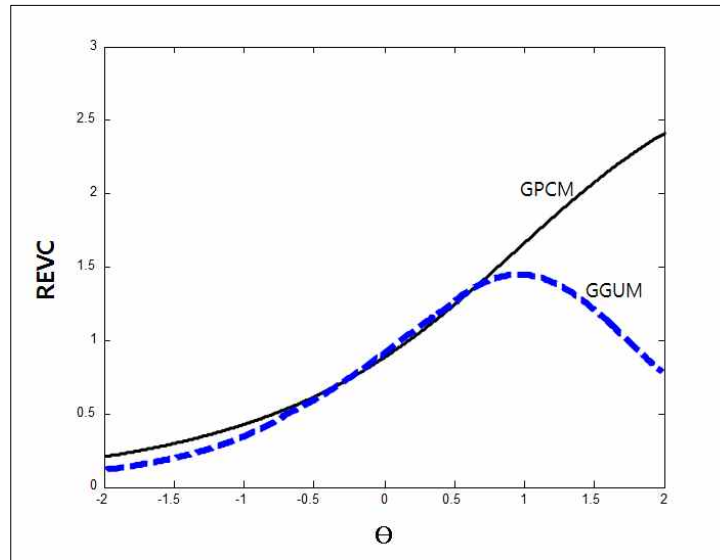
## 2. 일반화등급전개모형의 정의적 특성 검사의 다분문항자료에 대한 설명력

모의실험을 통하여 피험자 특성의 복원 연구를 수행한 결과, 생성모형이 일반화등급전개모형일 경우, 일반화등급전개모형에 의한 피험자 특성 추정치가 진능력모수를 가장 잘 복원하였고, 원점수는 피험자 특성을 타당하게

판별하지 못하였으며, 문항특성곡선이 누적기제를 따른다고 가정하는 일반화부분점수모형 또한 적합한 추정 모형이 될 수 없었다. 반면 일반화부분점수모형에 의하여 생성된 응답 데이터의 경우, 일반화부분점수모형에 의한 피험자 특성 추정치가 원점수와 거의 유사한 정도로 가장 좋은 복원력을 나타내었고, 일반화등급전개모형에 의한 피험자 특성 추정치는 피험자 특성 생성모수와 0.3 이하의 낮은 상관을 나타내었다. 즉 생성모형과 추정모형이 일치할 때 가장 우수한 피험자 특성 복원이 이루어진 것으로, 일반화등급전개모형이 모든 정의적 특성 검사에 대한 분석 모형의 대안이 될 수 없음을 알 수 있다. 따라서 피험자가 이상점과정에 기반하여, 문항 진술문의 내용과 자신이 가진 검사가 측정하고자 하는 구인수준을 비교하여 척도에 응답하였을 경우, 일반화등급전개모형은 적합한 문항분석 모형이 될 수 있지만, 우세 과정에 근거하여 응답한 피험자의 응답 자료는 일반화부분점수모형이 보다 타당한 분석 모형이 될 수 있을 것이다.

이처럼 복원 연구 결과가 실제 정의적 특성 검사 데이터를 분석하는 연구자나 실무자들이 자료에 적합한 모형을 선택할 필요성이 있음을 나타내기 때문에, 모형선택지수 DIC가 선택을 위한 지표역할을 해낼 수 있는지 추가적으로 살펴보았다. 하지만 복원 연구와 다르게, DIC는 생성모형에 관계없이 모든 모의 응답 자료를 일반화등급전개모형이 효율적으로 설명한다고 나타냄으로써, 모형선택에 대한 적합한 판단기준이 될 수 없다고 판단되었다.

가상의 동일한 다분문항에 대하여 일반화등급전개모형과 일반화부분점수모형을 적용하여 문항분석을 하였을 때, 다음과 같은 기대반응곡선이 나타날 수 있다.



<그림17> 가상의 동일한 문항에 관한 GPCM과 GGUM의 REVC

만약 대부분의 피험자 특성이 일반화등급전개모형의 대표 문항곤란도보다 작을 경우, 두 모형의 문항범주특성곡선은 해당 표본의 피험자 모수 범위에서 크게 다르지 않기 때문에, 두 모형의 모형 선택 지수가 비슷하게 나올 가능성이 제기될 수 있다. 하지만 본 연구의 경우, 일반화등급전개모형을 통하여 피험자 특성을 추정하였을 때의 DIC가 일반화부분점수모형을 적용하였을 때보다 일괄적으로 더 작게 나왔고, 이는 피험자들의 정의적 특성에 관한 응답 데이터가 일반화부분점수모형과 일반화등급전개모형 중 어느 모형에 의하여 보다 효율적으로 설명되는지에 관한 지표로서 DIC의 기능에 한계가 있음을 나타낸다. 연구자들의 모형 선택을 도울 수 있는 모형선택지수에 대한 후속 연구의 필요성이 제기되는 부분이다.

정의적 특성 검사에 관한 피험자들의 응답 데이터를 이상점 기제의 모형으로 분석할지, 우세반응 기제에 의한 모형으로 분석할지 선택할 수 있는 대안적 방법으로서 연구자의 전문성에 기인한 판단을 고려해볼 수 있다. 하

지만 Thurstone척도 이후 이상점 기제가 다시금 주목받은 것은 비교적 최근의 일이며, 이와 관련한 선행연구가 우세반응과정과 비교하여 부족한 상황 속에서 피험자들의 응답기제에 관한 선택을 연구자 개인의 판단에 맡기기보다, 이들의 판단을 도울 수 있는 명확한 지침을 제공하는 것이 선행되어야 할 것이다. 이를 위하여 우선적으로 해결해야 할 점은 Dalal et al.(2010)이 제시한, 연구자나 실무자들이 이상점 기제에 근거한 정의적 특성 검사도구를 활용할 경우 가질 수 있는 질문들로 구체화될 수 있다고 보여진다. Dalal et al.(2010)은 피험자의 수가 적기 때문에 복잡한 문항반응이론모형을 적용하지 않고 이상점 전개 기제에 근거하여 피험자들의 태도 점수를 산출하는 방법, 타당한 이상점 전개 과정을 나타낼 수 있는 문항작성 방법, 중립적 수준의 문항을 고려하여 검사도구를 구성할 경우 짧은 길이의 검사도구를 지향하는 최근의 흐름에 위배되는 것에 대한 해결책 등이 이상점 기제의 실제 적용에 있어 해결되어야 할 점이라고 지적하였다. 즉 누적기제를 따르는 문항반응이론 모형을 적용하지 않고, 간단히 Likrt척도에 기반한 원점수로 태도 점수를 산출하는 것과 같은 수준으로 전개기제를 실용화할 수 있도록 보다 다양한 연구가 수행되어야 할 것이다.

## 참 고 문 헌

- 강태훈, 김동일(2010). 정의적 특성 검사 자료 분석을 위한 문항반응모형 탐색. **교육평가연구**, 23(1), 149-170.
- 강태훈, 김명연(2012). 모의실험 연구를 통한 등급반응모형과 일반화부분점수모형 비교. **교육평가연구**, 25(3), 479-496.
- 구슬기(2011). 일반화부분점수모형에 의한 대인관계능력 진단도구의 문항특성분석. 이화여자대학교 일반대학원 석사학위 논문.
- 권재용(2012). 성격 검사 문항반응과정에 대한 고찰. 고려대학교 일반대학원 석사학위 논문.
- 김명소(1992). 심리검사 제작의 이론과 실제: 척도화 방법. **심리검사 제작의 이론과 실제**, -( ), 197-227.
- 김명소(1997). 척도개발의 일차원 접근법. **精神病理學**, 6(2), 183-201
- 김병구, 이경옥, 김민화, 김승옥, 김혜수(2006). **유·아동 및 청소년의 인터넷 게임중독 척도 개발 연구** (연구보고 06-15), 서울: 한국정보문화진흥원.
- 박도순(2006). **척도법**. 서울: 교육과학사.
- 박동건, 최명옥, 이원선, 이혜민, 이준희(2012). 성격 측정에서 우세반응 모형과 이상점반응 모형의 비교. **한국심리학회지: 산업 및 조직**, 25(2), 421-452.
- 박정(2001). **다분 문항반응이론 모형**. 서울: 교육과학사.
- 백순근(2010). 교육측정의 기초. 백순근(편), **교육측정의 이론과 실제**, 1-10. 서울: 교육과학사.
- 백순근, 김미림(2010). Likert형 태도 척도의 문항진술방식에 따른 척도 및 문항 특성 비교. **교육평가연구**, 23(4), 869-886

- 성태제(2001). **문항반응이론의 이해와 적용**. 서울: 교육과학사.
- 이기봉, 설수영, 원형중, 설민신(2009). **청소년의 게임중독 예방을 위한 가족단위 여가프로그램 활성화 방안** (연구보고 09-R11). 서울: 한국청소년정책연구원
- 이기봉(2010). **한국 아동·청소년 온라인 게임 및 가족여가활동실태조사**. 서울: 한국청소년정책연구원.
- 이선희(2008). 성격검사문항에 대한 이상점 반응 모형의 적용. **한국심리학회지: 일반**, 27(3), 839-857.
- 이순목(2010). **요인분석의 기초**. 서울: 교육과학사.
- 임미경(2001). 등급반응모형, 평정척도모형, 부분점수모형의 문항모수와 피험자모수 추정치 비교분석. 이화여자대학교 일반대학원 석사학위 논문.
- 임인재, 김신영, 박현정(2006). **심리측정의 원리**. 서울: 학연사.
- 최인재, 모상현, 이선영(2012). **아동·청소년 정신건강 증진을 위한 지원방안 연구 III : 총괄보고서** (연구보고 12-R15). 서울: 한국청소년정책연구원.
- 황은희(2010). 태도 및 흥미검사의 특성과 활용방안. 백순근(편), **교육측정의 이론과 실제**, 351-365. 서울: 교육과학사.
- 황정규, 서민원, 최종근, 김민성, 양명희, 김재철, 강태훈, 이대식, 김준엽, 신종호, 박인우, 김동일(2012). **교육평가의 이해**. 서울: 학지사.
- Anastasi, A., & Urbina, S.(1997). *Psychological Testing*(7th ed.), New Jersey: Prentice Hall.
- Andrich, D.(1988). The application of an unfolding model of the PIRT type to the measurement of attitude. *Applied Psychological Measurement*, 12, 33-51.

- Andrich, D.(1989). A probabilistic IRT model for unfolding preference data. *Applied Psychological Measurement, 13*(2), 247-264.
- Andrich, D., & Luo, G.(1993). A hyperbolic cosine *latent trait model for unfolding dichotomous single-stimulus responses*. *Applied Psychological Measurement, 17*(3), 253-276.
- Andrich, D.(1995). Hyperbolic cosine latent trait models for unfolding direct responses and pairwise preferences. *Applied Psychological Measurement, 19*(3), 269-290.
- Andrich, D.(1996). A hyperbolic cosine latent trait model for unfolding polytomous responses: Reconciling Thurstone and Likert methodologies. *British Journal of Mathematical and Statistical Psychology, 49*, 347-365.
- 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.
- Brown, A., & Maydeu-Olivares, A.(2010). Issues that should not be overlooked in the dominance versus ideal point controversy. *Industrial and Organizational Psychology, 3*, 489-493.
- Carter, N. T., & Dalal, D. K.(2010). An ideal point account of the JDI Work satisfaction scale. *Personality and Individual Differences, 49*, 743-748.
- Carter, N.T., Lake, C. J., & Zickar, M., J.(2010). Toward understanding the psychology of unfolding. *Industrial and Organizational Psychology, 3*, 511-514.

- Chernyshenko, O. S., Stark, S., Chan, K.-Y., Drasgow, F., & Williams, B. (2001). Fitting item response theory models to two personality inventories: Issues and insights. *Multivariate Behavioral Research*, 36, 523 - 562.
- Chernyshenko, O. S., Stark, S., Drasgow, F., & Williams, B. A. (2007). Constructing personality scales under the assumptions of an ideal point response process: toward increasing the flexibility of personality measures. *Psychological Assessment*, 19(1), 88-106.
- Dalal, D. K., Withrow, S., Gibby, R. E., & Zickar, M. J.(2010). Six questions that practitioners (might) have about ideal point response process items. *Industrial and Organizational Psychology*, 3, 498-501.
- DeSarbo, W. S., & Hoffman, D. L.(1986). Simple and weighted unfolding threshold models for the spatial representation of binary choice data. *Applied Psychological Measurement*, 10, 247-264.
- Drasgow, F., Chernyshenko, O. S., & Strak, S.(2010). 75 years after Likert: Thurstone was right!. *Industrial and Organizational Psychology*, 3, 465-476.
- Fan, X.(1998). Item response theory and classical test theory: An empirical comparison of their item/person statistics. *Educational and Psychological Measurement*, 58, 357-381.
- Green, B. F.(1954). Attitude measurement. In G. Lindzey (Ed.), *Handbook of social psychology* (1st ed., vol. 1, pp. 335-369). Cambridge. MA: Addison-Wesley.
- Guttman, L.(1950). The basis for scalogram analysis. In S. A. Stouffer et

- al. (Eds.), *Measurement and Prediction*. Princeton, NJ: Princeton University Press.
- Guttman, L.(1953). Image theory for the structure of quantitative variates. *Psychometrika*, 18, 277-296.
- Hojtink, H.(1990). A latent trait model for dichotomous choice data. *Psychometrika*, 55, 641-656.
- Hojtink, H.(1991). The measurement of latent traits by proximity items. *Applied Psychological Measurement*, 15, 153-169.
- Javaras, K. N.(2004). Statistical analysis of Likert data on attitudes. Doctoral dissertation, University of Oxford.
- Kaiser, H. F.(1974). An index of factorial simplicity. *Psychometrika*, 39(1), 31-36.
- Kang, T., Cohen, A. S., & Sung, H.-J.(2009). Model selection indices for polytomous items. *Applied Psychological Measurement*, 33, 499-518.
- Kantrowitz, T. M., & Tuzinski, K. A.(2010). The ideal point model in action: how the use of computer adaptive personality scales benefits organizations. *Industrial and Organizational Psychology*, 3, 507-510.
- Lake, C. J., & Zickar, M. J. (2010). Respondent and test delivery characteristics that induce item unfolding. Unpublished manuscript.
- Lin, B. C., Carter, N. T., Zickar, M. J., Dalal, D. K., Adams, J. E., & Wolford, K. A.(2009). Unfolding item response theory: Analysis of employment drug testing attitudes. Paper presented at the 24th Annual Meeting of the Society for Industrial and Organizational Psychology, New Orleans, LA.

- Masters, G. N.(1982). A Rasch model for partial credit scoring. *Psychometrika*, 47, 149-174.
- Muraki, E.(1992). A generalized partial credit model: Application of an EM algorithm. *Applied Psychological Measurement*, 16, 159-176.
- Oswald, F. L., & Schell, K. L.(2010). Developing and scaling personality measures: Thurstone was right-but so far, Likert was not wrong. *Industrial and Organizational Psychology*, 3, 481-484.
- Peterson, R. C., & Thurstone, L. L.(1933). *Motion Pictures and the Social Attitudes of Children: A Payne Fund Study*. New York: Macmillan & Company.
- ([http://www.brocku.ca/MeadProject/Thurstone/1933/1933\\_04.html](http://www.brocku.ca/MeadProject/Thurstone/1933/1933_04.html))
- Porat, A. B.(1977). Guttman scale test for Maslow Need Hierarchy. *The Journal of Psychology*, 97, 85-92.
- Reeve, B. B.(2002). *An introduction to modern measurement theory*. Bethesda: National Cancer Inst, 2002.
- Roberts, J. S.(1995). Item response theory approaches to attitude measurement. Doctoral dissertation, University of South Carolina, Columbia.
- Roberts, J. S., Laughlin, J. E., & Donoghue, J. R.(1998). Validity issues in the Likert and Thurstone approaches to attitude measurement, *Educational and Psychological Measurement*, 59(2), 211-233.
- Roberts, J. S., Donoghue, J. R., & Laughlin, J. E.(2000). A general item response theory model for unfolding unidimensional polytomous responses. *Applied Psychological Measurement*, 24, 3-32.

- Roberts, J. S., Lin, Y., & Laughlin, J. E.(2001). Computerized adaptive testing with the generalized graded unfolding model. *Applied Psychological Measurement, 25*(2), 177–196.
- Robert, F. D.(2011). *Scale Development: Theory and Applications (Applied Social Research Methods)*. SAGE Publications.
- Safir, M. A.(1937). A comparative study of scales constructed by three psychophysical methods. *Psychometrika, 2*, 179–198.
- Sgammato, A. N.(2009). An application of unfolding and cumulative item response theory models for noncognitive scaling: examining the assumptions and applicability of the generalized graded unfolding model. Doctoral dissertation, The University of North Carolina at Chapel Hill.
- Stage, C.(2003). *Classical test theory or item response theory: The Swedish Experience*(EM No. 42). Umea, Sweden: Umea Universitet Department of Educational Measurement.
- Stark, S.(2002). A new IRT approach to test construction and scoring designed to reduce the effects of faking in personality assessment: the generalized graded unfolding model for multi-unidimensional paired comparison responses. Doctoral dissertation, University of Illinois at Urbana-Champaign.
- Stark, S., Chernyshenko, O. S., Drasgow, F., & Williams, B. A.(2006). Examining assumption about item responding in personality assessment: should ideal point methods be considered for scale development and scoring?

- Torre, J., Stark, S., & Chernyshenko, O. S.(2006). Markov Chain Monte Carlo Estimation of item parameters for the generalized graded unfolding model. *Applied Psychological Measurement, 30*(3), 216-232.
- Thurstone, L. L.(1928). Attitudes can be measured. *The American Journal of Sociology, 33*, 529-554.
- Thurstone, L. L.(1929). Theory of attitude measurement. *Psychological Review, 36*, 222-241.
- Thurstone, L. L., & Chave, E. J.(1929). *The measurement of attitude: A psychophysical method and some experiments with a scale for measuring attitude toward the church*. Chicago: University of Chicago Press.
- Waples, C. J., Weyhrauch, W. S., Connell, A. R., & Culbertson S. S.(2010). Questionable defeats and discounted victories for Likert rating scales. *Industrial and Organizational Psychology, 3*, 477-480.

# ABSTRACT

## Item response models for analyzing affective testing data

Min Ah, Oh

Department of Education

The graduate school of Sungshin Women's University

This study compared generalized partial credit model that used commonly for analyzing polytomous-item dataset and generalized graded unfolding model assuming ideal point process in analyzing affective polytomous-item testing data. In order to explore two model's practicability, two study problems was set up as following. First, is it possible that generalized graded unfolding model is used to interpret item's statements that measure affective construct? Second, could generalized graded unfolding model be applied to analyze affective polytomous-item testing dataset?

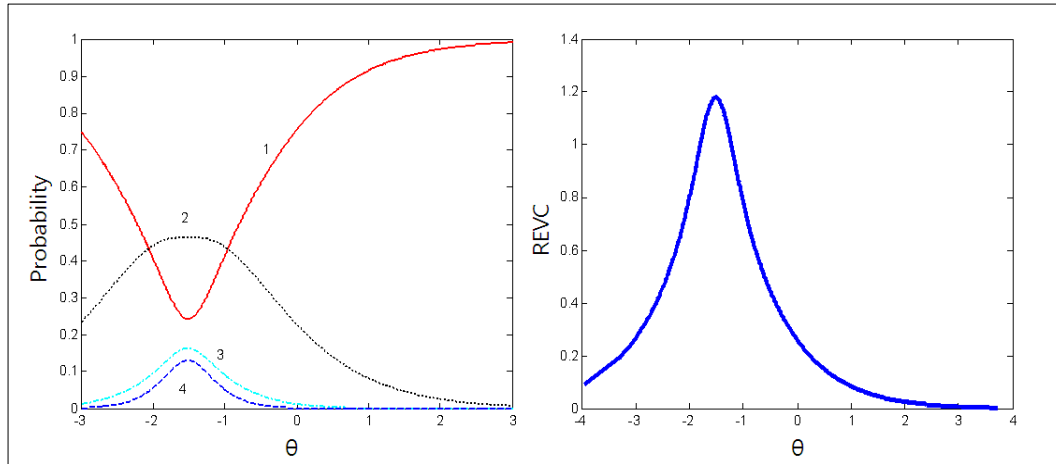
For these questions, this study conducted analyzing real data of internet game addiction for adolescents and simulated data that were generated by using generalized graded unfolding model and generalized partial credit model. Real data and simulated data were Likert-type scale composed of four response categories-strongly disagree, disagree, agree, and strongly agree.

The result of analyzing real data was that generalized partial credit model appeared a little better than generalized graded unfolding model in DIC, a model selection index. And generalized graded unfolding model was more informative method in interpreting item than generalized partial credit model.

Simulated data were generated by using generalized graded unfolding model and generalized partial credit model. The former was used to generate affective testing data in unfolding mechanism and the latter cumulative mechanism. The item response theory models for calibrating were also generalized graded unfolding model and generalized partial credit model. In this simulation study, the correlation coefficients among true simulees's true theta, raw scores and estimated trait parameters were computed by two models to investigate accuracy in recovering the true parameters. Simulee's true parameters were best recovered when the generating model was equal to calibrating model. However DIC showed that the test data analysis could be more efficient when generating model was used as calibrating model. It means that DIC could not be the model selection index when researcher or practitioner need to choose among generalized graded unfolding model and generalized partial credit model and other statistical model selection indices are required to help their model selection.

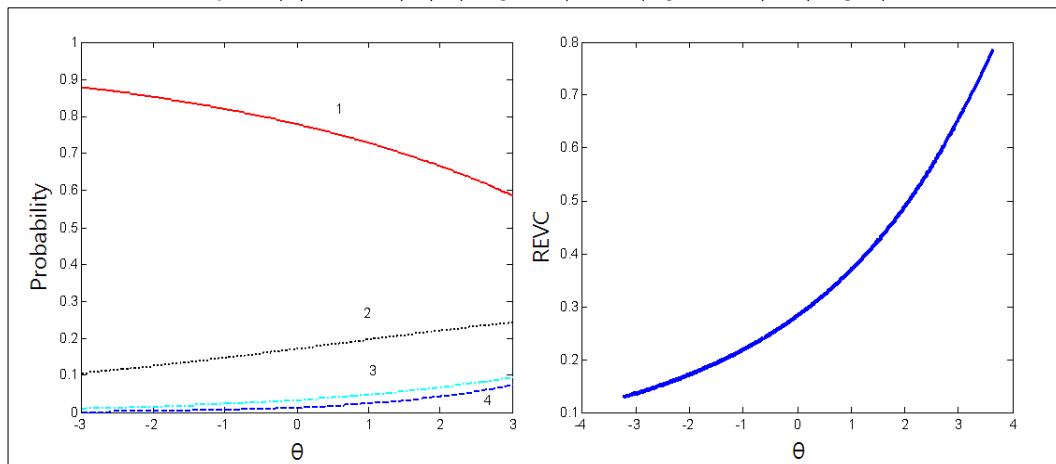
## 부록 1 : 일반화등급전개모형에 따른 ICCC와 REVC

문항1. 게임을 하는 것이 친한 친구들과 어울리는 것보다 더 좋다.



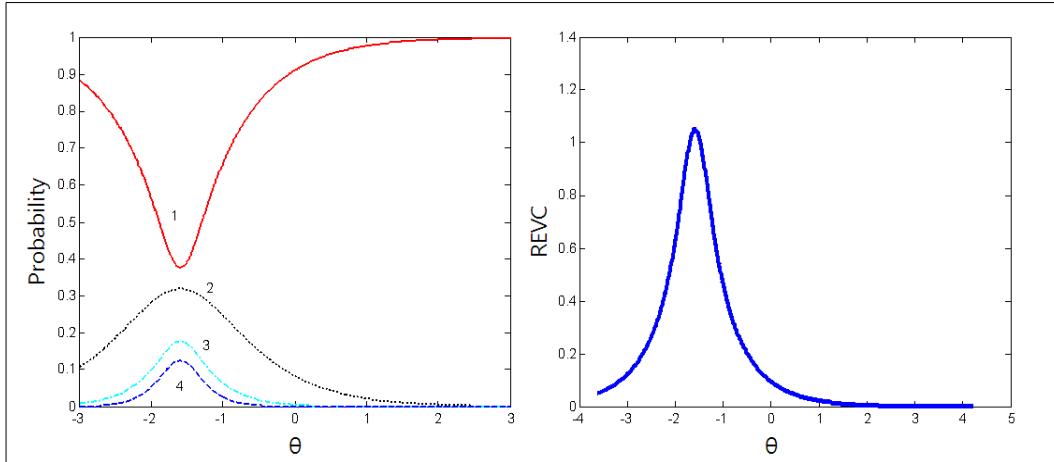
| $\alpha_i$ | $\delta_i$ | $\tau_{i1}$ | $\tau_{i2}$ | $\tau_{i3}$ |
|------------|------------|-------------|-------------|-------------|
| 1.21       | -1.51      | -0.53       | 0.86        | 0.19        |

문항2. 게임공간에서의 생활이 실제생활보다 더 좋다.



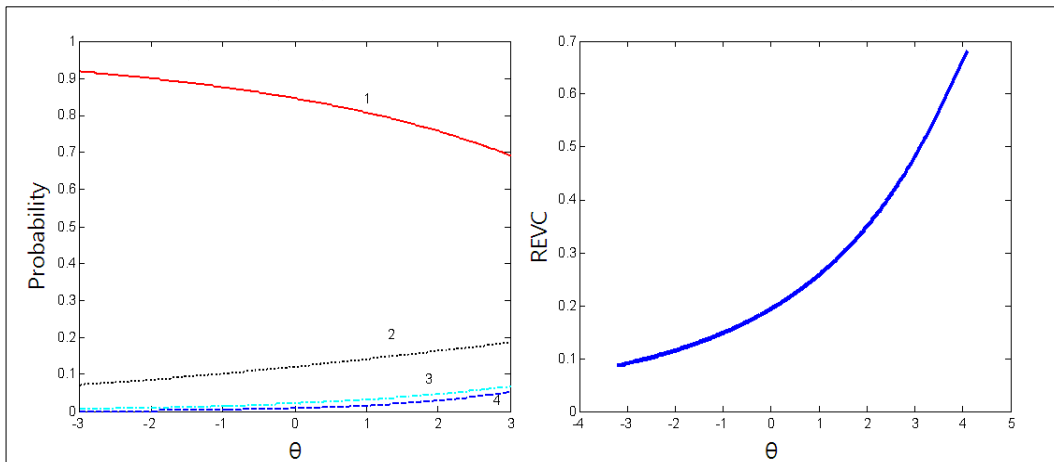
| $\alpha_i$ | $\delta_i$ | $\tau_{i1}$ | $\tau_{i2}$ | $\tau_{i3}$ |
|------------|------------|-------------|-------------|-------------|
| 0.20       | 6.13       | 1.43        | 2.07        | -0.46       |

문항3. 게임 속의 내가 실제의 나보다 더 좋다.



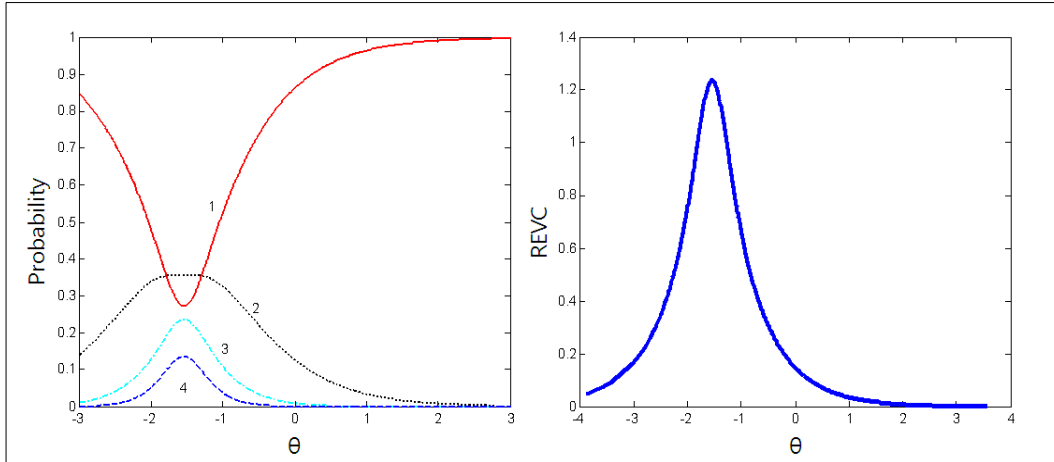
| $\alpha_i$ | $\delta_i$ | $\tau_{i1}$ | $\tau_{i2}$ | $\tau_{i3}$ |
|------------|------------|-------------|-------------|-------------|
| 1.39       | -1.59      | 0.12        | 0.42        | 0.25        |

문항4. 게임에서 사귀 친구들이 실제 친구들보다 나를 더 알아준다.



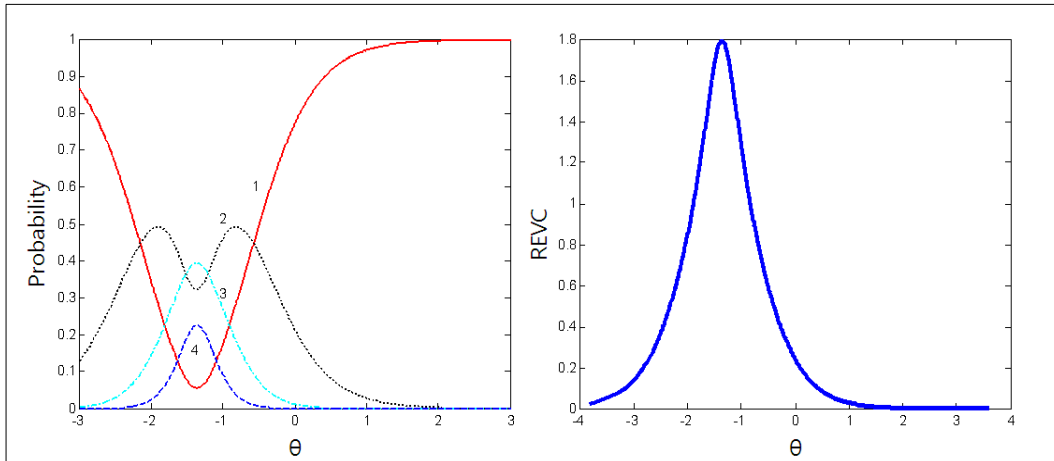
| $\alpha_i$ | $\delta_i$ | $\tau_{i1}$ | $\tau_{i2}$ | $\tau_{i3}$ |
|------------|------------|-------------|-------------|-------------|
| 0.20       | 6.09       | 3.63        | 2.44        | -0.43       |

문항5. 게임에서 사람을 사귀는 것이 더 편하고 자신 있다.



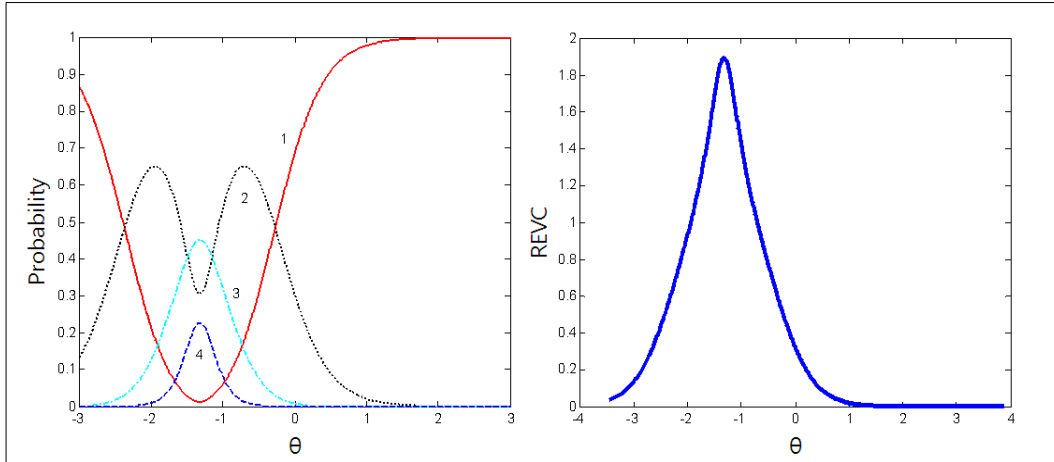
| $\alpha_i$ | $\delta_i$ | $\tau_{i1}$ | $\tau_{i2}$ | $\tau_{i3}$ |
|------------|------------|-------------|-------------|-------------|
| 1.41       | -1.54      | -0.19       | 0.29        | 0.39        |

문항6. 밤 늦게까지 게임을 하느라 시간 가는 줄 모른다.



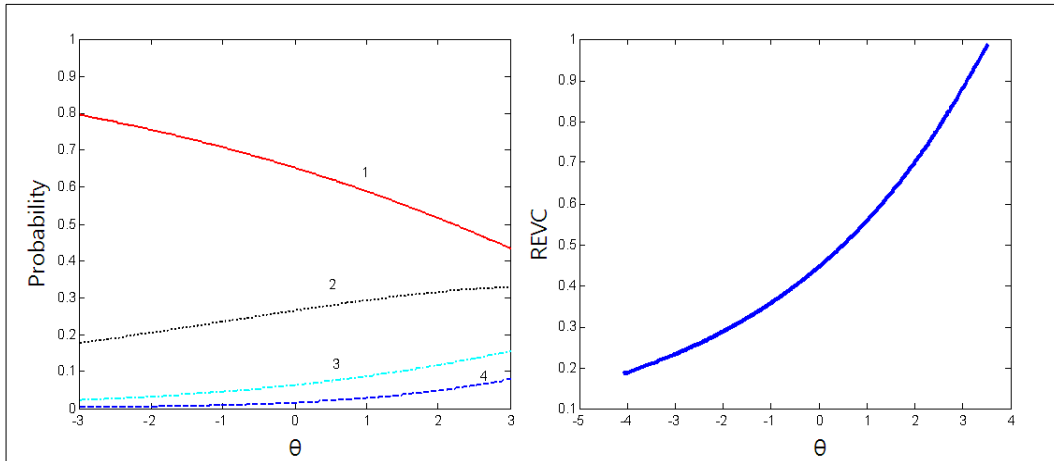
| $\alpha_i$ | $\delta_i$ | $\tau_{i1}$ | $\tau_{i2}$ | $\tau_{i3}$ |
|------------|------------|-------------|-------------|-------------|
| 2.23       | -1.36      | -0.79       | -0.09       | 0.25        |

문항7. 게임을 하느라 해야 할 일을 못한다.



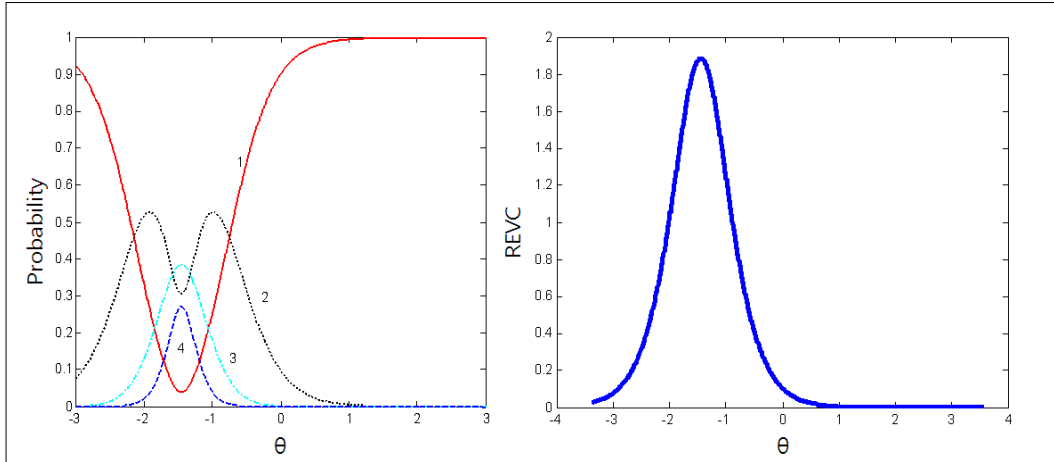
| $\alpha_i$ | $\delta_i$ | $\tau_{i1}$ | $\tau_{i2}$ | $\tau_{i3}$ |
|------------|------------|-------------|-------------|-------------|
| 2.97       | -1.32      | -1.05       | -0.13       | 0.23        |

문항8. 갈수록 게임을 하는 시간이 길어진다.



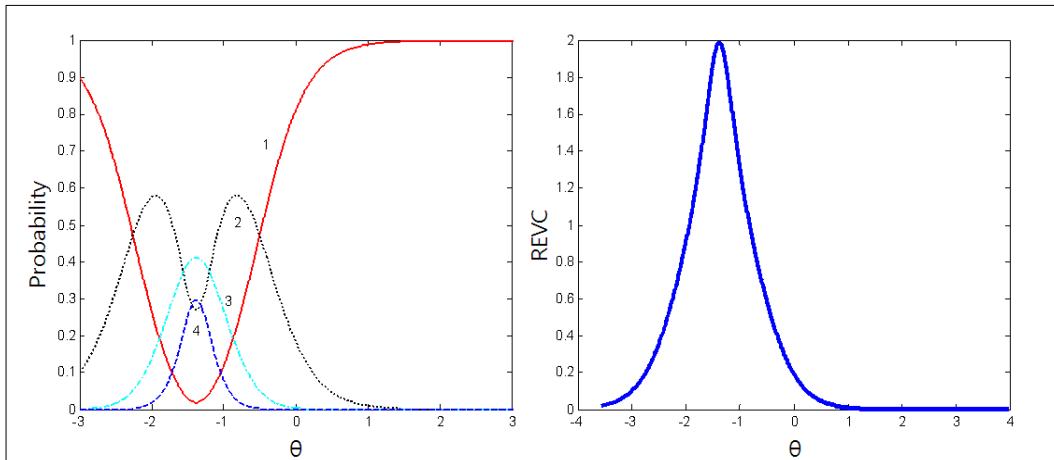
| $\alpha_i$ | $\delta_i$ | $\tau_{i1}$ | $\tau_{i2}$ | $\tau_{i3}$ |
|------------|------------|-------------|-------------|-------------|
| 0.20       | 6.47       | -1.97       | 0.72        | 1.30        |

문항9. 점점 더 오랜 시간 게임을 해야 만족하게 된다.



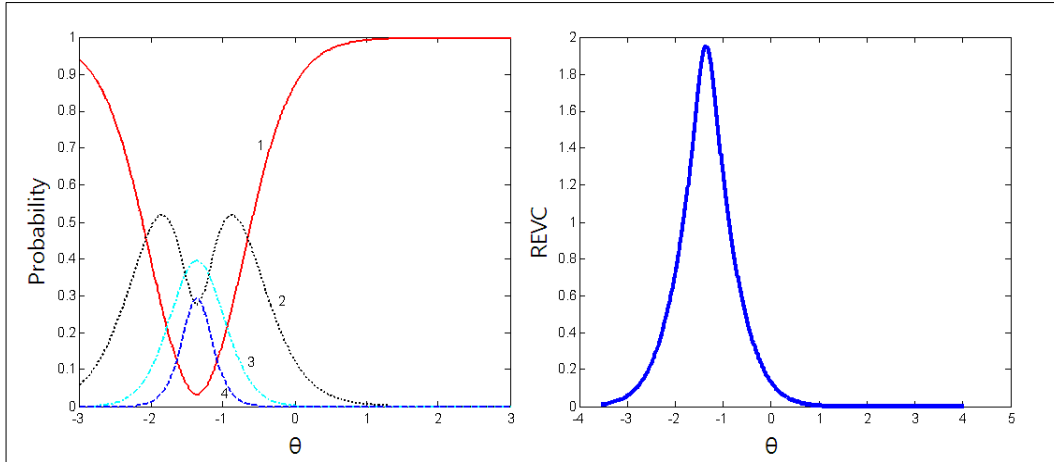
| $\alpha_i$ | $\delta_i$ | $\tau_{i1}$ | $\tau_{i2}$ | $\tau_{i3}$ |
|------------|------------|-------------|-------------|-------------|
| 2.95       | -1.45      | -0.70       | -0.08       | 0.12        |

문항10. 게임을 그만두어야 하는 경우에도 게임을 그만두는 것이 어렵다.



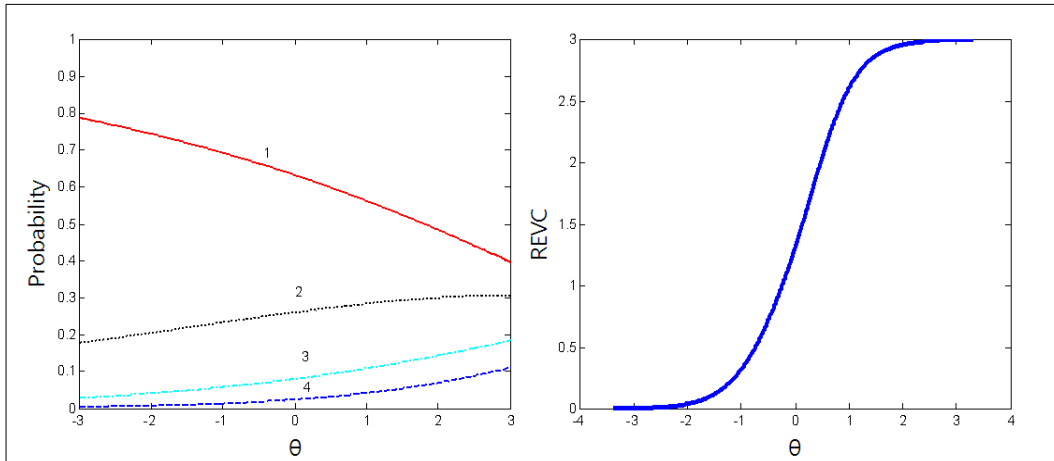
| $\alpha_i$ | $\delta_i$ | $\tau_{i1}$ | $\tau_{i2}$ | $\tau_{i3}$ |
|------------|------------|-------------|-------------|-------------|
| 2.97       | -1.38      | -0.88       | -0.14       | 0.11        |

문항11. 게임 하는 시간을 줄이려고 노력하지만 실패한다.



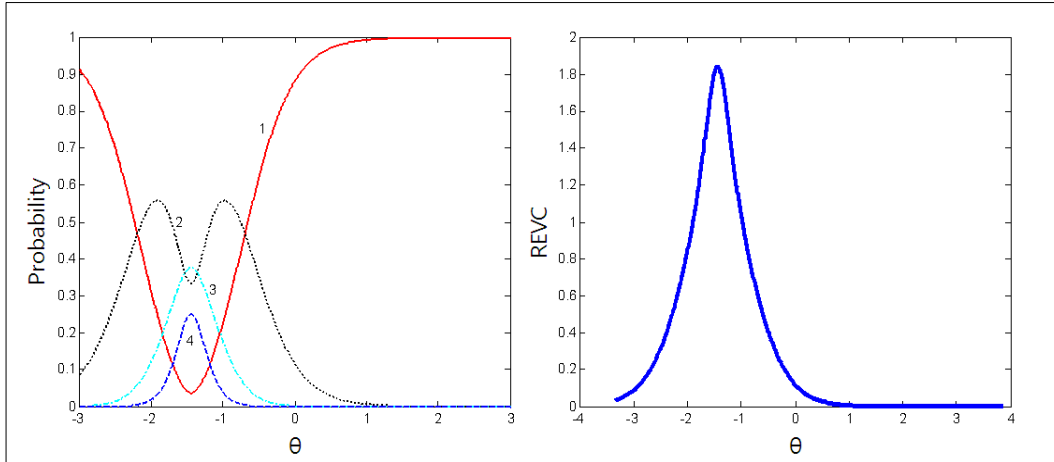
| $\alpha_i$ | $\delta_i$ | $\tau_{i1}$ | $\tau_{i2}$ | $\tau_{i3}$ |
|------------|------------|-------------|-------------|-------------|
| 2.97       | -1.36      | -0.71       | -0.12       | 0.10        |

문항12. 게임을 안 하겠다고 마음먹고도 다시 게임을 하게 된다.



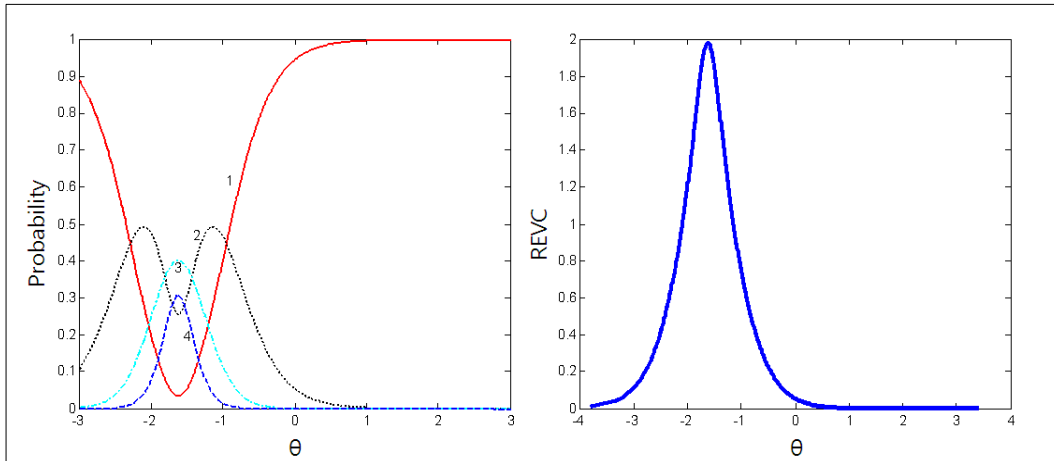
| $\alpha_i$ | $\delta_i$ | $\tau_{i1}$ | $\tau_{i2}$ | $\tau_{i3}$ |
|------------|------------|-------------|-------------|-------------|
| 0.20       | 6.40       | -1.96       | -0.43       | 0.56        |

문항13. 게임 생각 때문에 공부에 집중하기 어렵다.



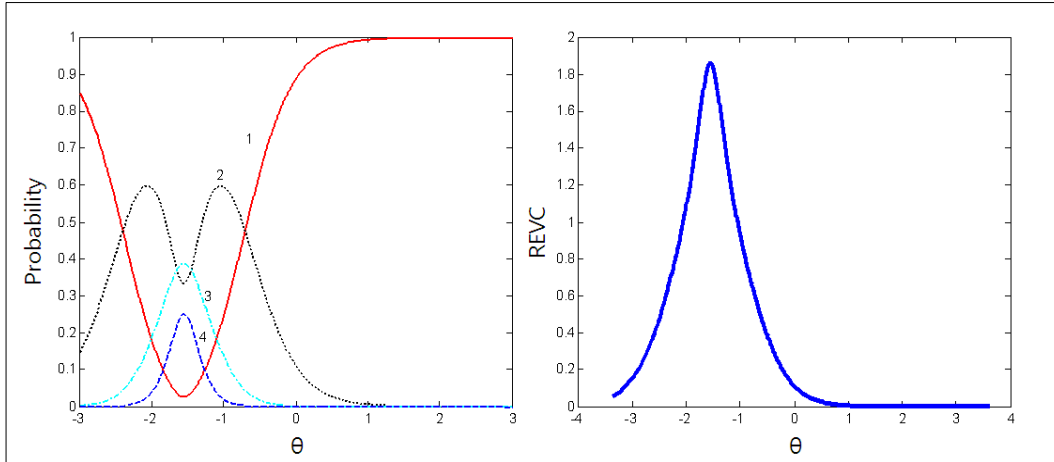
| $\alpha_i$ | $\delta_i$ | $\tau_{i1}$ | $\tau_{i2}$ | $\tau_{i3}$ |
|------------|------------|-------------|-------------|-------------|
| 2.92       | -1.44      | -0.75       | -0.04       | 0.14        |

문항14. 게임을 못한다는 것은 견디기 힘든 일이다.



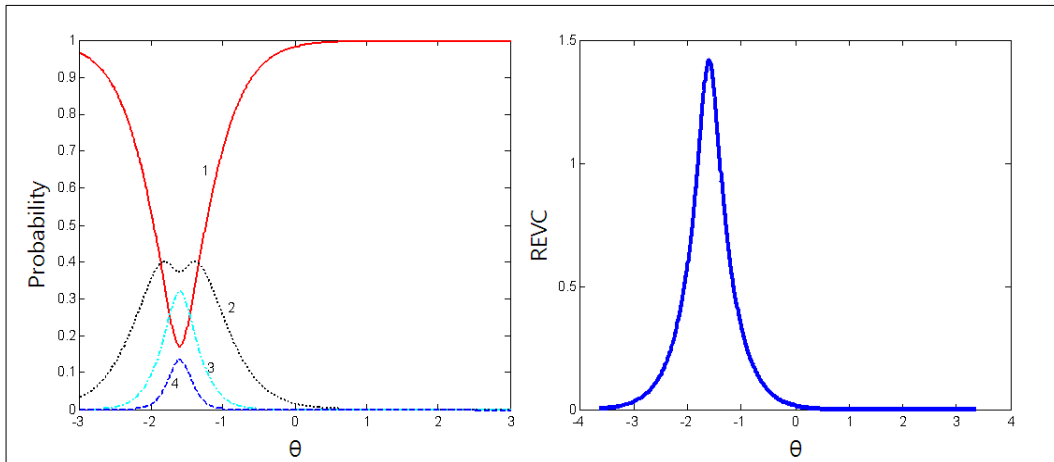
| $\alpha_i$ | $\delta_i$ | $\tau_{i1}$ | $\tau_{i2}$ | $\tau_{i3}$ |
|------------|------------|-------------|-------------|-------------|
| 2.98       | -1.62      | -0.67       | -0.15       | 0.09        |

문항15 게임을 하지 않을 때에도 게임 생각을 하게 된다.



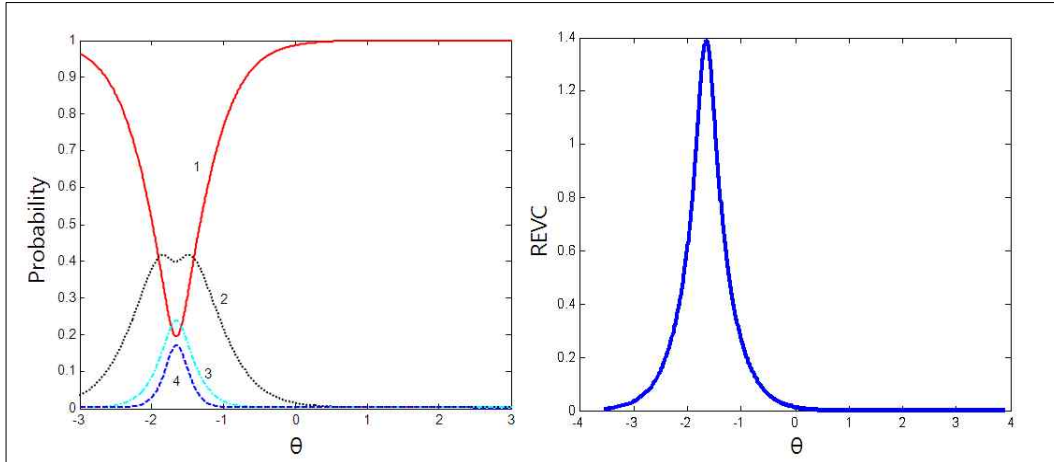
| $\alpha_i$ | $\delta_i$ | $\tau_{i1}$ | $\tau_{i2}$ | $\tau_{i3}$ |
|------------|------------|-------------|-------------|-------------|
| 2.95       | -1.55      | -0.85       | -0.05       | 0.15        |

문항16. 게임으로 인해 생활에 문제가 생기더라도 게임을 해야 한다.



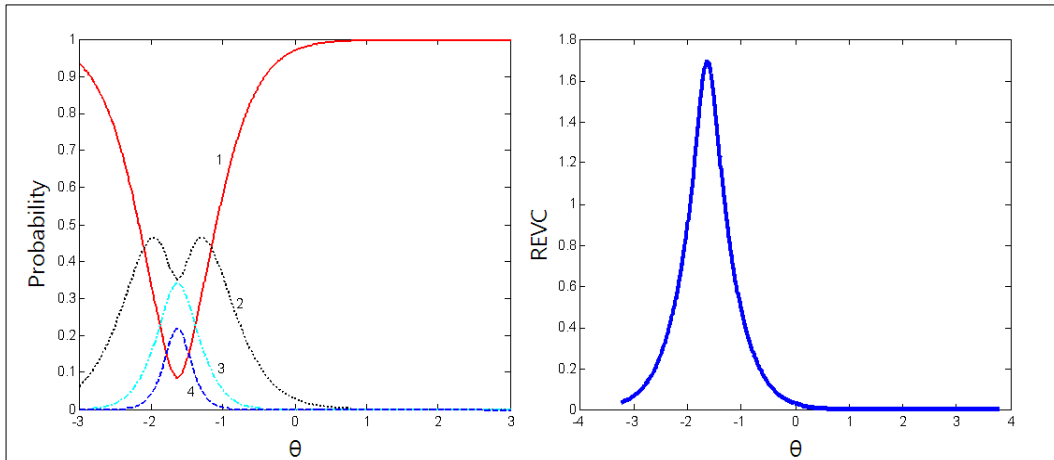
| $\alpha_i$ | $\delta_i$ | $\tau_{i1}$ | $\tau_{i2}$ | $\tau_{i3}$ |
|------------|------------|-------------|-------------|-------------|
| 2.99       | -1.60      | -0.26       | 0.05        | 0.29        |

문항17. 게임을 하지 못하면 불안하고 초조하다.



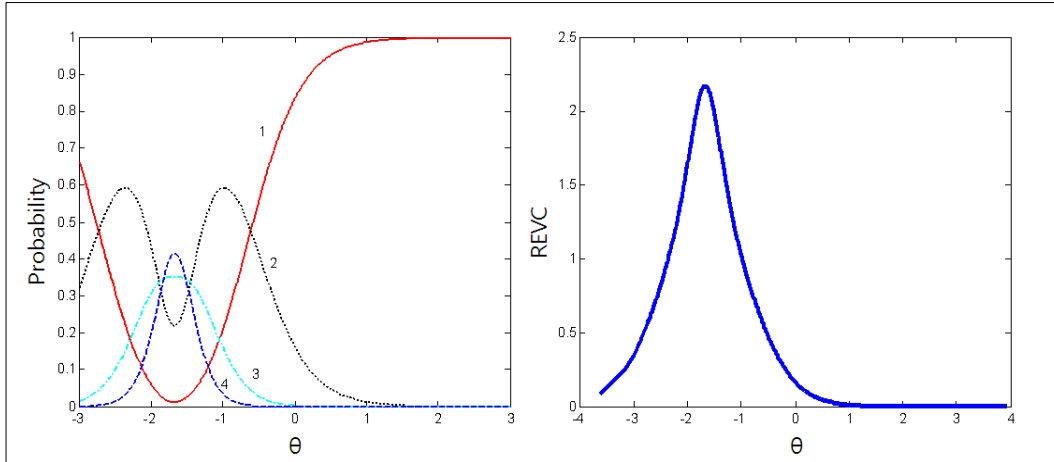
| $\alpha_i$ | $\delta_i$ | $\tau_{i1}$ | $\tau_{i2}$ | $\tau_{i3}$ |
|------------|------------|-------------|-------------|-------------|
| 2.99       | -1.65      | -0.24       | 0.17        | 0.11        |

문항18. 다른 일 때문에 게임을 못하게 될까봐 걱정된다.



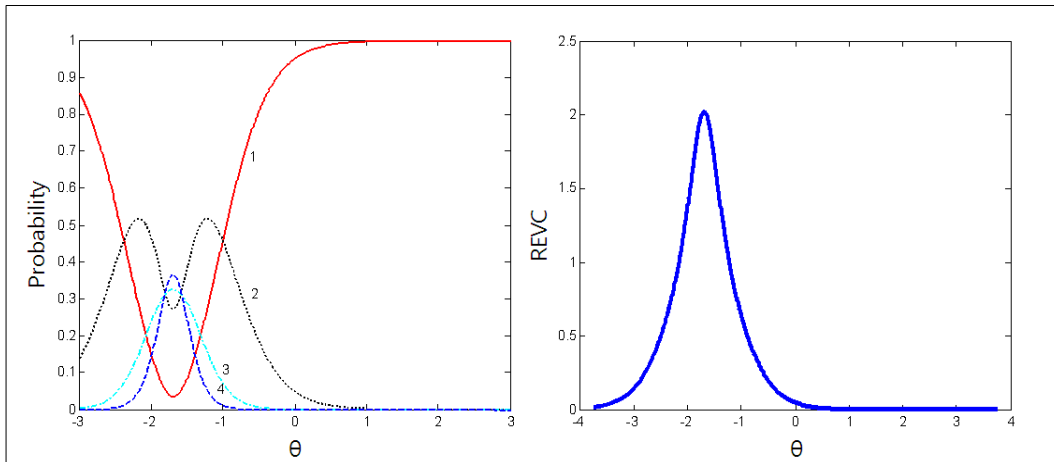
| $\alpha_i$ | $\delta_i$ | $\tau_{i1}$ | $\tau_{i2}$ | $\tau_{i3}$ |
|------------|------------|-------------|-------------|-------------|
| 2.98       | -1.63      | -0.47       | 0.01        | 0.15        |

문항19. 누가 게임을 못 하게 하면 신경질이 난다.



| $\alpha_i$ | $\delta_i$ | $\tau_{i1}$ | $\tau_{i2}$ | $\tau_{i3}$ |
|------------|------------|-------------|-------------|-------------|
| 2.67       | -1.67      | -1.06       | -0.18       | -0.06       |

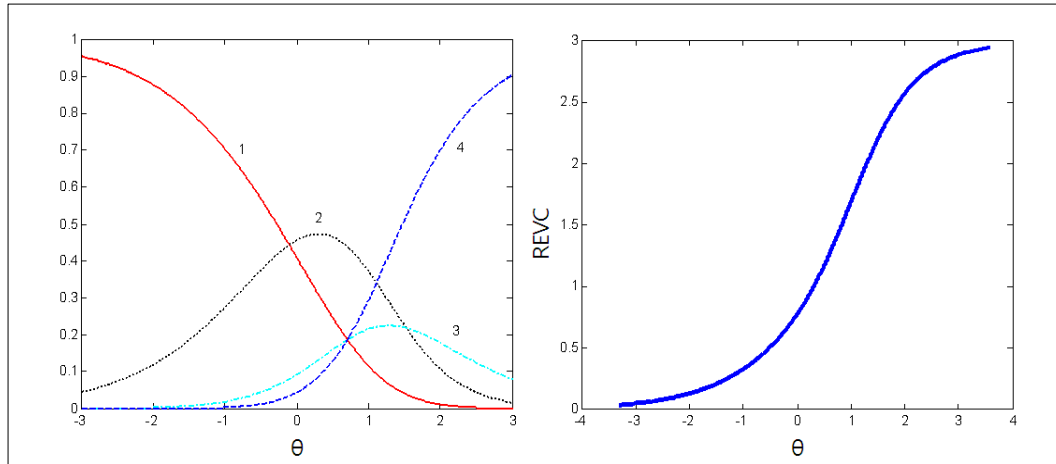
문항20. 게임을 못하게 되면 화가 난다.



| $\alpha_i$ | $\delta_i$ | $\tau_{i1}$ | $\tau_{i2}$ | $\tau_{i3}$ |
|------------|------------|-------------|-------------|-------------|
| 2.93       | -1.69      | -0.69       | -0.06       | -0.04       |

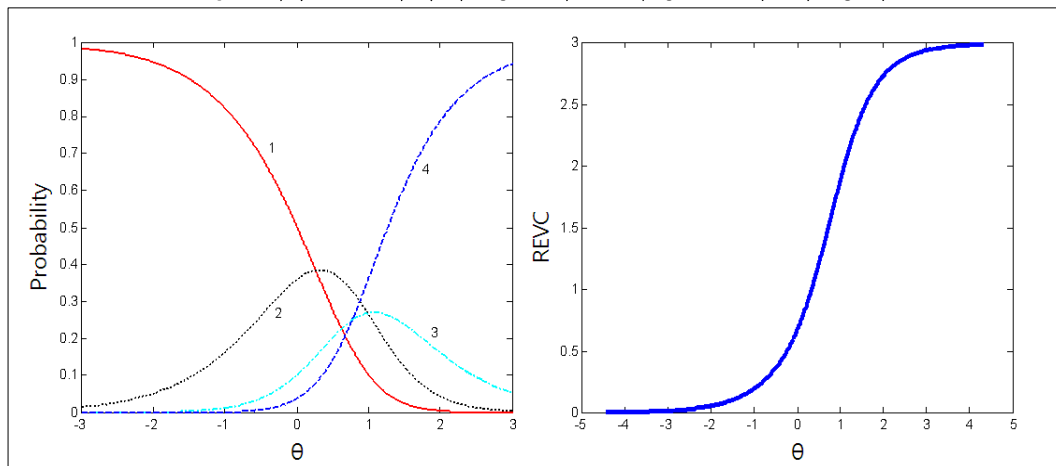
## 부록2 : 일반화부분점수모형에 따른 ICCC와 REVC

문항1. 게임을 하는 것이 친한 친구들과 어울리는 것보다 더 좋다.



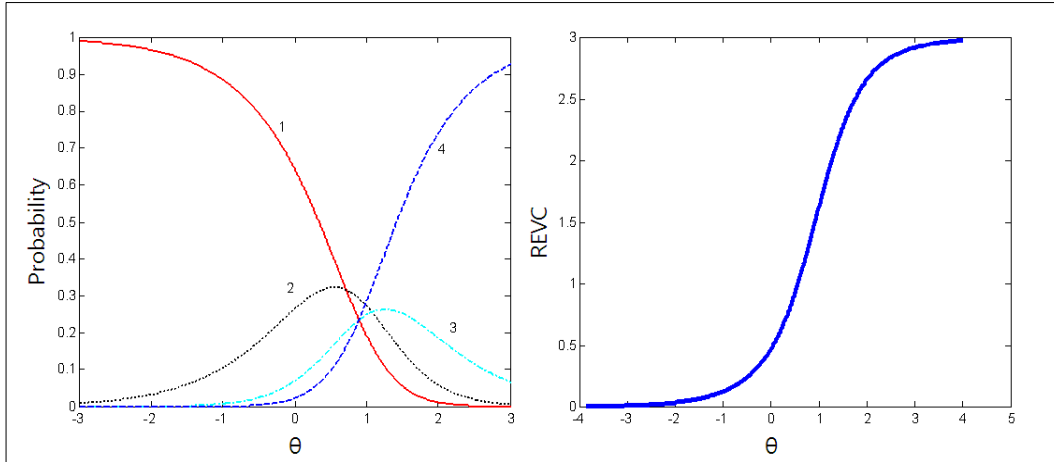
| $\alpha_i$ | $\delta_i$ | $\tau_{i1}$ | $\tau_{i2}$ | $\tau_{i3}$ |
|------------|------------|-------------|-------------|-------------|
| 1.05       | 0.71       | 0.81        | -0.82       | 0.01        |

문항2. 게임공간에서의 생활이 실제생활보다 더 좋다.



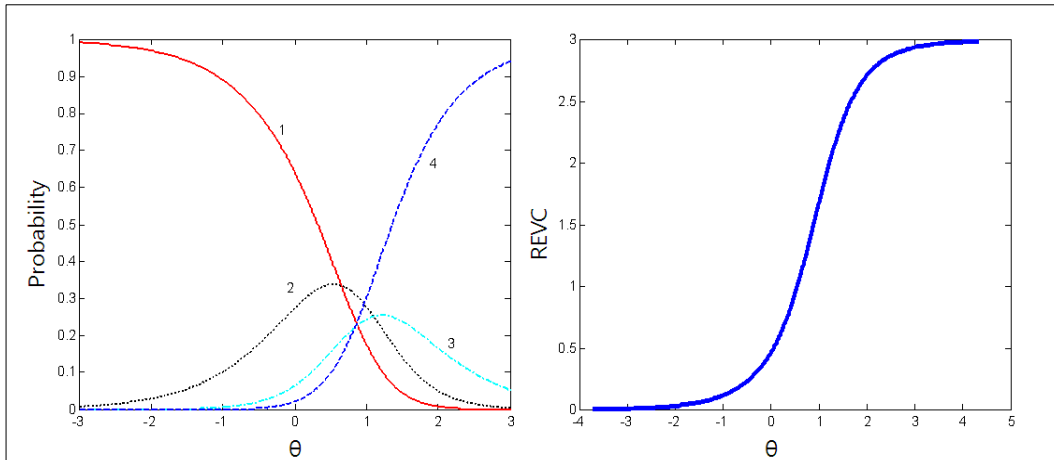
| $\alpha_i$ | $\delta_i$ | $\tau_{i1}$ | $\tau_{i2}$ | $\tau_{i3}$ |
|------------|------------|-------------|-------------|-------------|
| 1.28       | 0.68       | 0.41        | -0.32       | -0.10       |

문항3. 게임 속의 내가 실제의 나보다 더 좋다.



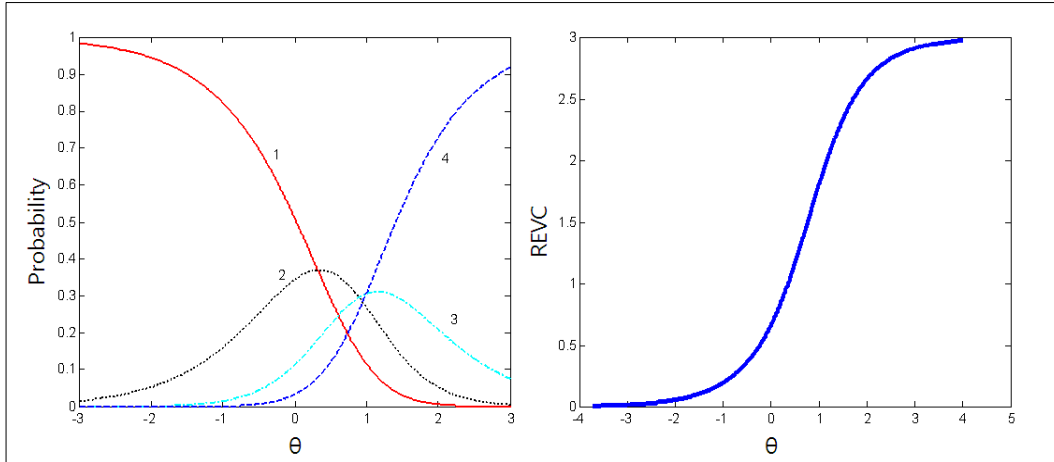
| $\alpha_i$ | $\delta_i$ | $\tau_{i1}$ | $\tau_{i2}$ | $\tau_{i3}$ |
|------------|------------|-------------|-------------|-------------|
| 1.24       | 0.90       | 0.18        | -0.18       | 0.01        |

문항4. 게임에서 사귄 친구들이 실제 친구들보다 나를 더 알아준다.



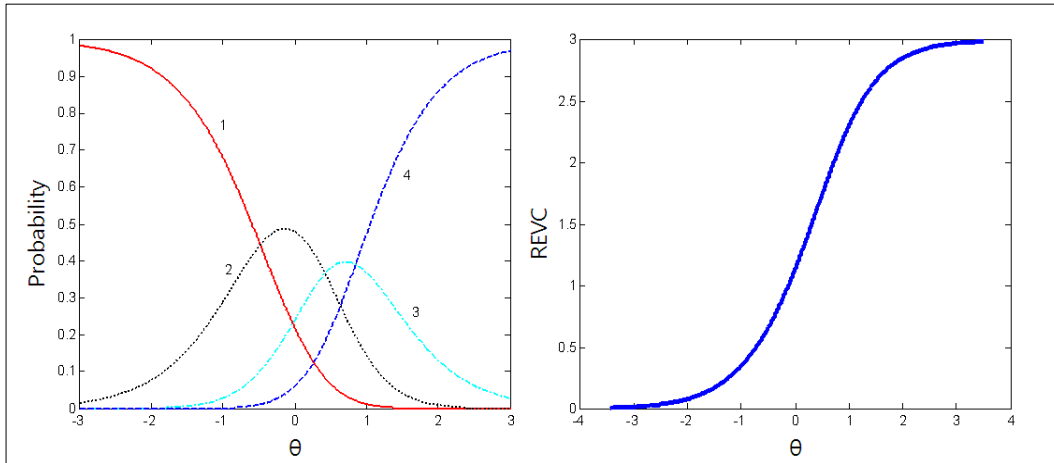
| $\alpha_i$ | $\delta_i$ | $\tau_{i1}$ | $\tau_{i2}$ | $\tau_{i3}$ |
|------------|------------|-------------|-------------|-------------|
| 1.32       | 0.86       | 0.21        | -0.24       | 0.03        |

문항5. 게임에서 사람을 사귀는 것이 더 편하고 자신 있다.



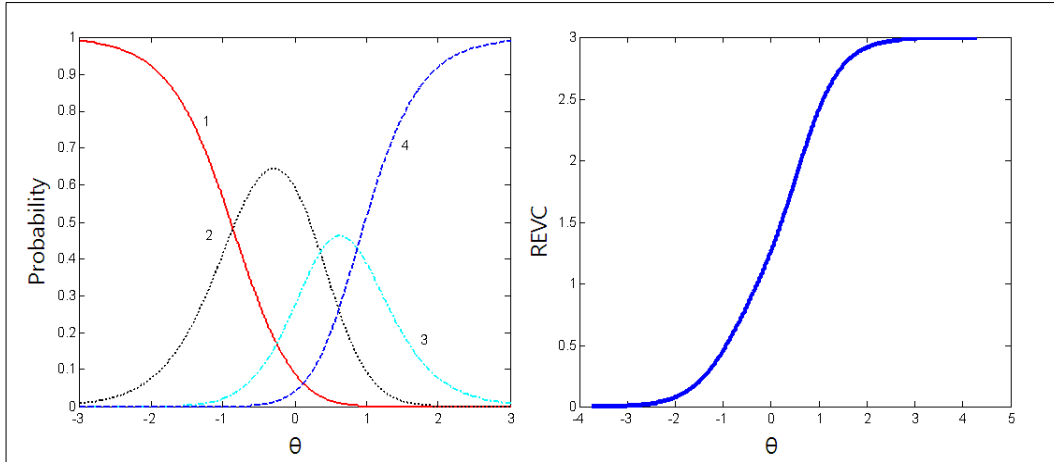
| $\alpha_i$ | $\delta_i$ | $\tau_{i1}$ | $\tau_{i2}$ | $\tau_{i3}$ |
|------------|------------|-------------|-------------|-------------|
| 1.24       | 0.74       | 0.42        | -0.16       | -0.26       |

문항6. 밤 늦게까지 게임을 하느라 시간 가는 줄 모른다.



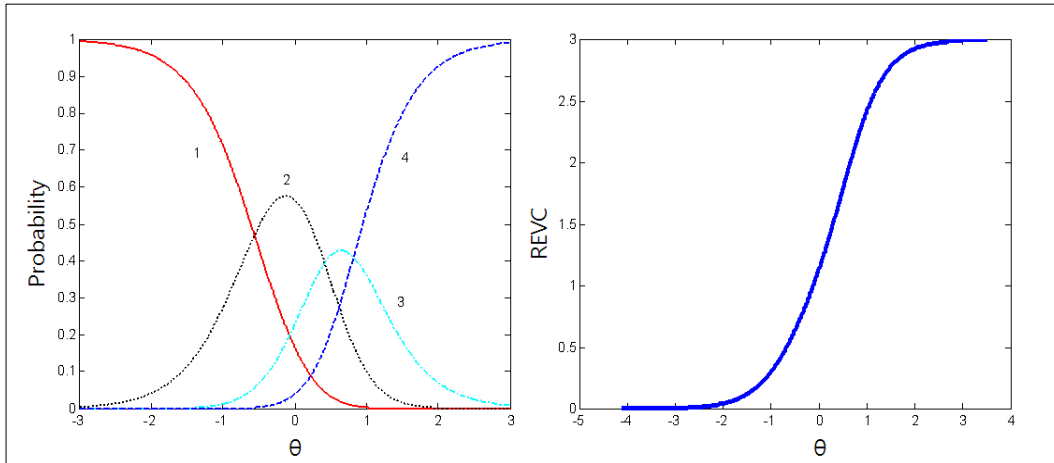
| $\alpha_i$ | $\delta_i$ | $\tau_{i1}$ | $\tau_{i2}$ | $\tau_{i3}$ |
|------------|------------|-------------|-------------|-------------|
| 1.63       | 0.27       | 0.74        | -0.16       | -0.58       |

문항7. 게임을 하느라 해야 할 일을 못한다.



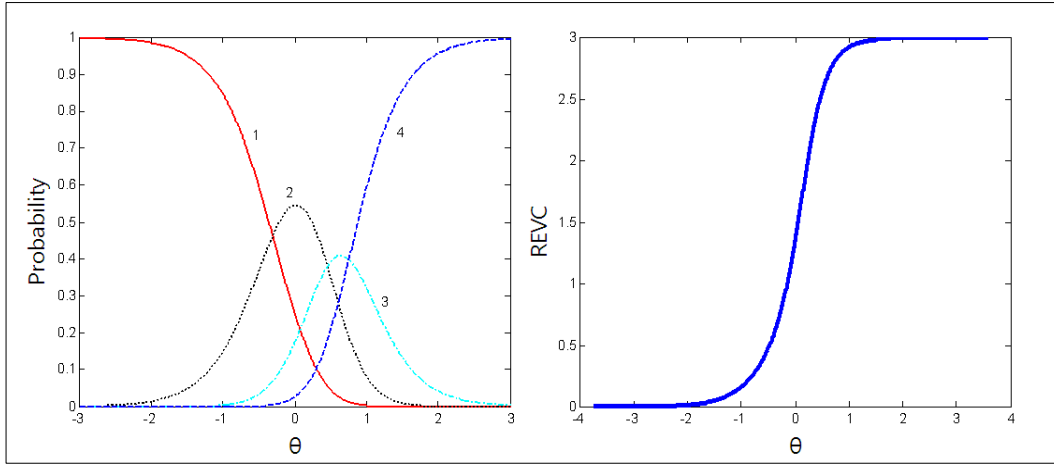
| $\alpha_i$ | $\delta_i$ | $\tau_{i1}$ | $\tau_{i2}$ | $\tau_{i3}$ |
|------------|------------|-------------|-------------|-------------|
| 2.18       | 0.12       | 0.98        | -0.23       | -0.75       |

문항8. 갈수록 게임을 하는 시간이 길어진다.



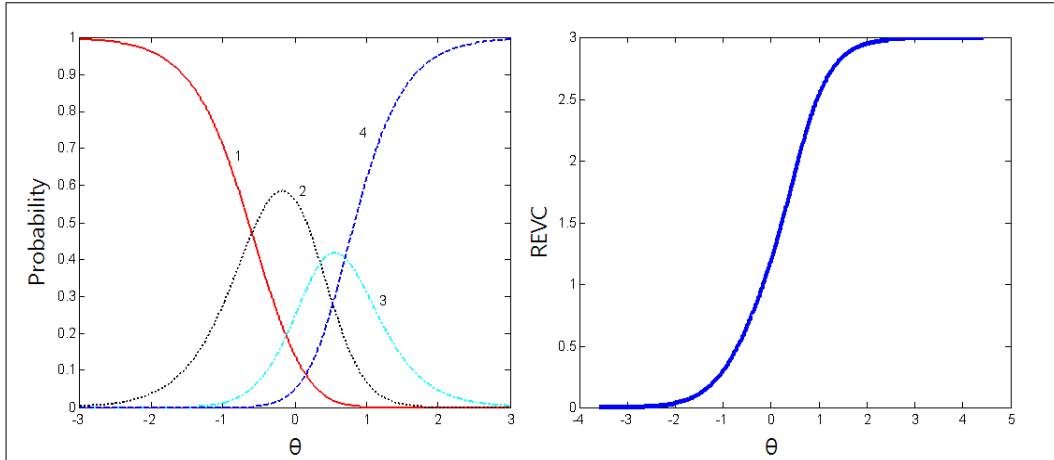
| $\alpha_i$ | $\delta_i$ | $\tau_{i1}$ | $\tau_{i2}$ | $\tau_{i3}$ |
|------------|------------|-------------|-------------|-------------|
| 2.17       | 0.23       | 0.79        | -0.19       | -0.60       |

문항9. 점점 더 오랜 시간 게임을 해야 만족하게 된다.



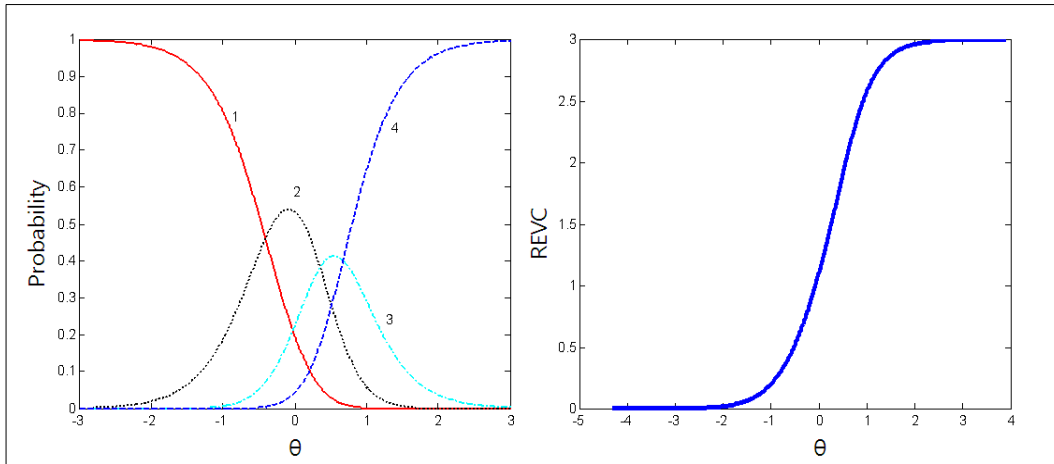
| $\alpha_i$ | $\delta_i$ | $\tau_{i1}$ | $\tau_{i2}$ | $\tau_{i3}$ |
|------------|------------|-------------|-------------|-------------|
| 2.50       | 0.30       | 0.61        | -0.16       | -0.45       |

문항10. 게임을 그만두어야 하는 경우에도 게임을 그만두는 것이 어렵다.



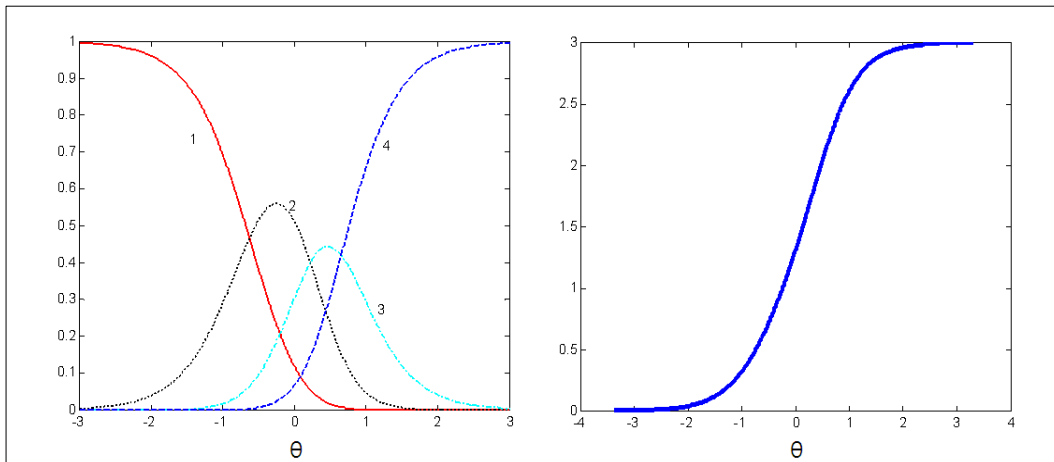
| $\alpha_i$ | $\delta_i$ | $\tau_{i1}$ | $\tau_{i2}$ | $\tau_{i3}$ |
|------------|------------|-------------|-------------|-------------|
| 2.31       | 0.16       | 0.75        | -0.20       | -0.55       |

문항11. 게임 하는 시간을 줄이려고 노력하지만 실패한다.



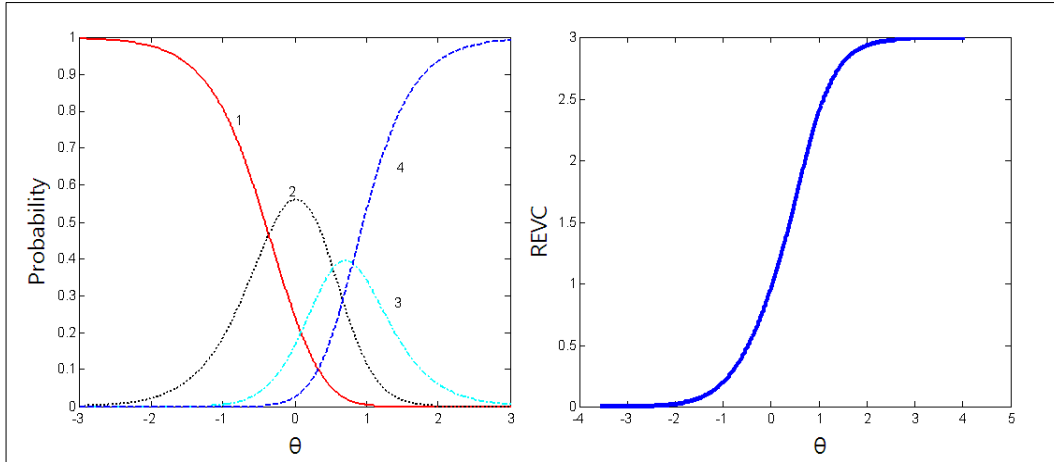
| $\alpha_i$ | $\delta_i$ | $\tau_{i1}$ | $\tau_{i2}$ | $\tau_{i3}$ |
|------------|------------|-------------|-------------|-------------|
| 2.44       | 0.21       | 0.61        | -0.15       | -0.47       |

문항12. 게임을 안 하겠다고 마음먹고도 다시 게임을 하게 된다.



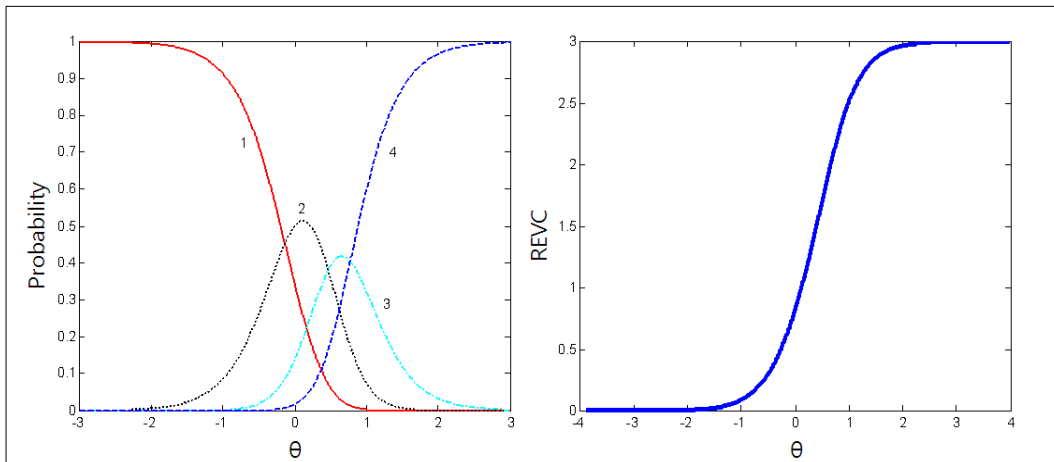
| $\alpha_i$ | $\delta_i$ | $\tau_{i1}$ | $\tau_{i2}$ | $\tau_{i3}$ |
|------------|------------|-------------|-------------|-------------|
| 2.34       | 0.09       | 0.71        | -0.14       | -0.57       |

문항13. 게임 생각 때문에 공부에 집중하기 어렵다.



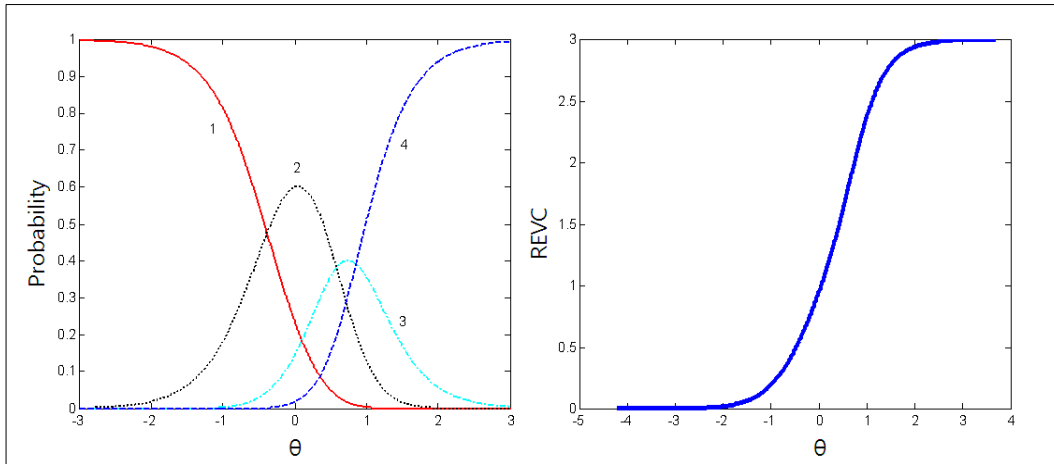
| $\alpha_i$ | $\delta_i$ | $\tau_{i1}$ | $\tau_{i2}$ | $\tau_{i3}$ |
|------------|------------|-------------|-------------|-------------|
| 2.28       | 0.33       | 0.69        | -0.21       | -0.48       |

문항14. 게임을 못한다는 것은 견디기 힘든 일이다.



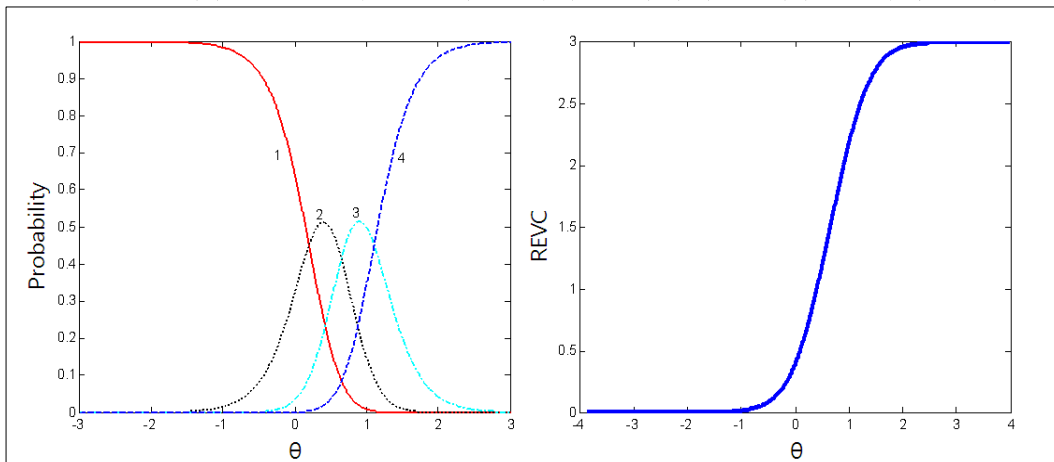
| $\alpha_i$ | $\delta_i$ | $\tau_{i1}$ | $\tau_{i2}$ | $\tau_{i3}$ |
|------------|------------|-------------|-------------|-------------|
| 2.74       | 0.37       | 0.50        | -0.10       | -0.41       |

문항15. 게임을 하지 않을 때에도 게임 생각을 하게 된다.



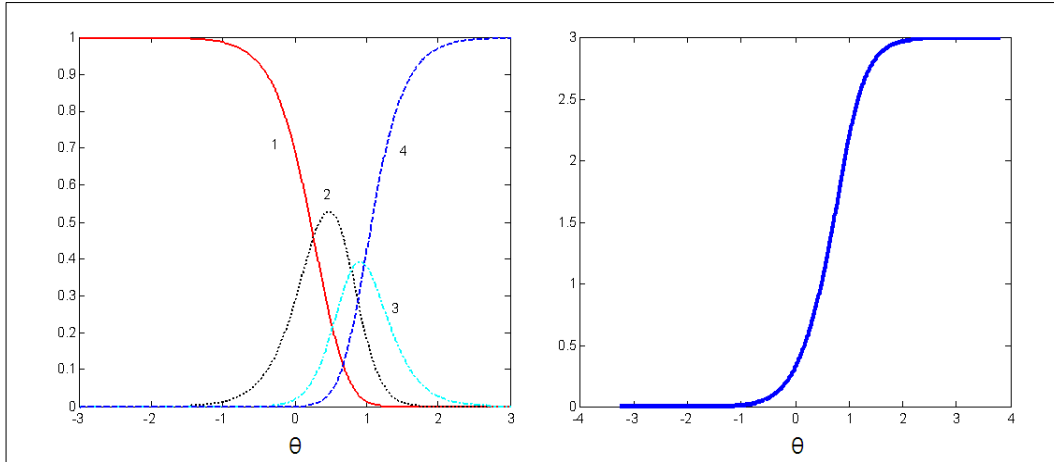
| $\alpha_i$ | $\delta_i$ | $\tau_{i1}$ | $\tau_{i2}$ | $\tau_{i3}$ |
|------------|------------|-------------|-------------|-------------|
| 2.42       | 0.35       | 0.74        | -0.23       | -0.50       |

문항16. 게임으로 인해 생활에 문제가 생기더라도 게임을 해야 한다.



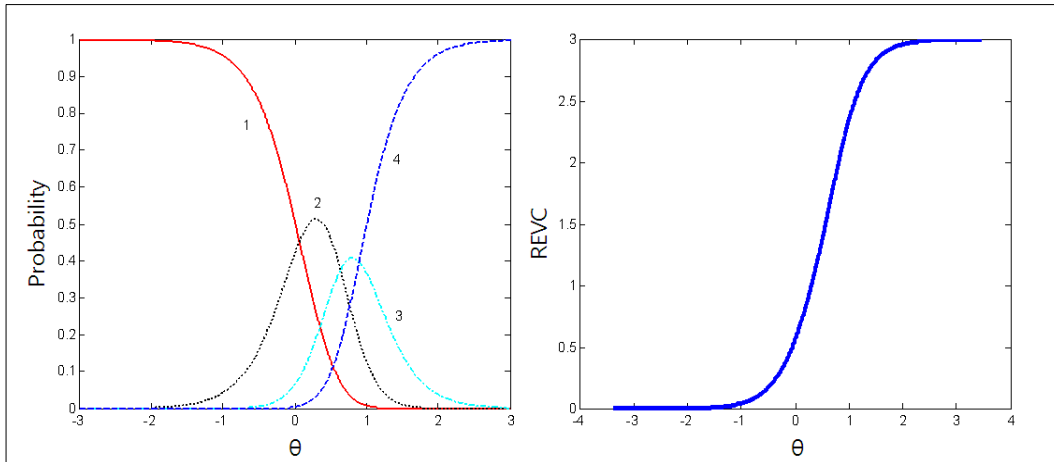
| $\alpha_i$ | $\delta_i$ | $\tau_{i1}$ | $\tau_{i2}$ | $\tau_{i3}$ |
|------------|------------|-------------|-------------|-------------|
| 3.48       | 0.65       | 0.45        | 0.00        | -0.46       |

문항17. 게임을 하지 못하면 불안하고 초조하다.



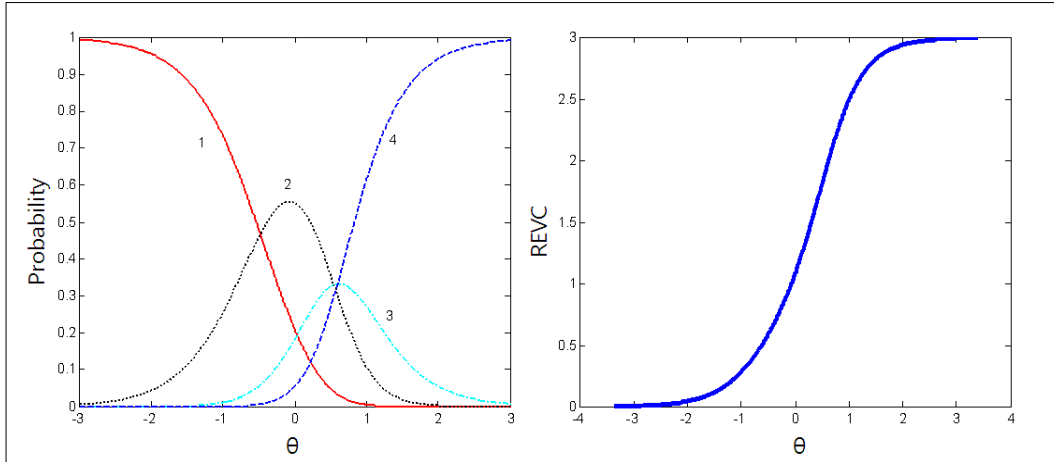
| $\alpha_i$ | $\delta_i$ | $\tau_{i1}$ | $\tau_{i2}$ | $\tau_{i3}$ |
|------------|------------|-------------|-------------|-------------|
| 3.41       | 0.68       | 0.41        | -0.12       | -0.30       |

문항18. 다른 일 때문에 게임을 못하게 될까봐 걱정된다.



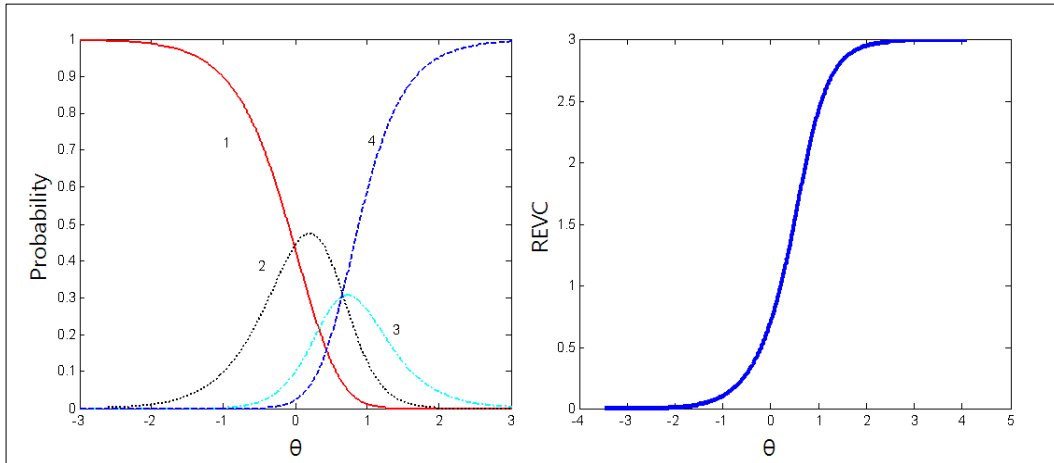
| $\alpha_i$ | $\delta_i$ | $\tau_{i1}$ | $\tau_{i2}$ | $\tau_{i3}$ |
|------------|------------|-------------|-------------|-------------|
| 2.91       | 0.54       | 0.47        | -0.10       | -0.36       |

문항19. 누가 게임을 못 하게 하면 신경질이 난다.



| $\alpha_i$ | $\delta_i$ | $\tau_{i1}$ | $\tau_{i2}$ | $\tau_{i3}$ |
|------------|------------|-------------|-------------|-------------|
| 2.02       | 0.22       | 0.70        | -0.33       | -0.38       |

문항20. 게임을 못하게 되면 화가 난다.



| $\alpha_i$ | $\delta_i$ | $\tau_{i1}$ | $\tau_{i2}$ | $\tau_{i3}$ |
|------------|------------|-------------|-------------|-------------|
| 2.23       | 0.44       | 0.45        | -0.24       | -0.21       |

### 부록 3 : 일반화등급전개모형과 일반화부분점수모형을 위한

#### WinBUGS 코드

```

# WinBUGS code for calibrating Generalized Graded Unfolding Model
model
{
for (i in 1:T) {
  cumtau[i,1] <- 0; M[i] <- 2*(mI[i]-1)+1;
  # tau1,2,3,4 were used instead of tau0,1,2,3. So, tau1=0 always.
  for (k in 1:mI[i]-1) {
    cumtau[i,k+1] <- tau[i,k+1] + cumtau[i,k];
  }
}

# GGUM
for (j in 1:N) {
  for (i in 1:T) {
    for (k in 1:mI[i]){
      numer[j,i,k] <- exp(a[i]*((k-1)*(theta[j]-d[i])-cumtau[i,k])) +
exp(a[i]*((M[i]-(k-1))*(theta[j]-d[i])-cumtau[i,k]))
    }
    denom[j,i] <- sum(numer[j,i,1:mI[i]]);
  }
}
for (j in 1:N) {
  for (i in 1:T) {
    for (k in 1:mI[i]){
      p[j,i,k] <- numer[j,i,k]/denom[j,i];
    }
    r[j,i] ~ dcat(p[j,i,1:mI[i]]);
  }
}
theta[j] ~ dnorm(mu,1);
}
mu ~ dnorm(0,1);

# Priors
# item discrimination
#for (i in 1:T) {
#  a[i] ~ dlnorm(0,1);    }
for (i in 1:T) {
  a[i] ~ dunif(0.2,3.0);    }

# item location
for (i in 1:T){
  d[i]~dnorm(0,1);    }
# for (i in 1:T){
#  d[i]~dunif(-2.0,2.0);    }

# The first Category of every item
for (i in 1:T) {
  tau[i,1] <- 0;    }
# The other Category of every item
for (i in 1:T) {
  for (k in 2:(mI[i])) {
    tau[i,k] ~ dnorm(0,5);    }}
}

```

```

# WinBUGS code for calibrating Generalized Partial Credit Model

model
{
for (j in 1:N){
  for (i in 1:T) {
    denom[j,i,1] <- 1;
    numer[j,i,1] <- 0;
    enumer[j,i,1] <- 1;
  }

  # GPCM
  for (j in 1:N) {
    for (i in 1:T) {
      for (k in 2:mI[i]) {
        numer[j,i,k] <- a[i]*(theta[j] - b[i] + tau[i,k] ) + numer[j,i,k-1];
        enumer[j,i,k] <- exp(numer[j,i,k]);
        denom[j,i,k] <- enumer[j,i,k] + denom[j,i,k-1];
      }
      denom2[j,i,1] <- denom[j,i,mI[i]];
    }
  }

  for (j in 1:N) {
    for (i in 1:T) {
      for (k in 1:mI[i]){
        p[j,i,k] <- enumer[j,i,k]/denom2[j,i,1];
      }
      r[j,i] ~ dcat(p[j,i,1:mI[i]]);
    }
  }
  theta[j] ~ dnorm(mu,1);
}
mu ~ dnorm(0,1);

# Priors
# item discrimination
for (i in 1:T) {
  a[i] ~ dlnorm(0,1) ;
}

# item difficulty
# for (i in 1:T){
#   b.pre[i]~dnorm(0,1) ;
# }
# for (i in 1:T){
#   b[i] <- b.pre[i]-mean(b.pre[1:T]) ;
# }
for (i in 1:T){
  b[i]~dnorm(0,1);
}

# The first Category of every item
for (i in 1:T) {
  tau[i,1] <- 0;
}

# From the second to (mI-1)th Categories of each item
for (i in 1:T){
  for (k in 2:(mI[i]-1)) {
    tau[i,k]~dnorm(0,.5);
  }
}

# The last Category of every item: It makes the sum of all category parameters be
0
for (i in 1:T){
  tau[i,mI[i]] <- -sum(tau[i,2:(mI[i]-1)]);
}

```

```
# If you are interested in parameters about the categories without item difficulty
# for (i in 1:T) {
#   step[i,1] <- 0;           }
# for (i in 1:T){
#   for (k in 2:mI[i]){
#     step[i,k] <- b[i] - tau[i,k] ;   }}
# }
```