

다척도 심리검사 자료의 이질적 극단응답양식 탐색을 위한 혼합문항반응나무모형의 적용

이 주 연*

중앙대학교

오 정 수

Eastern Illinois University

장 윤 선†

대구교육대학교

문항반응나무모형(Item Response Tree: IRTree)은 응답자의 설문 응답 시 발생할 수 있는 인지적 응답 반응 과정을 나무구조로 반영하여 문항반응이론에 접목시킨 모형으로, 최근 응답 편향의 일종인 극단응답반응양식(Extreme Response Style: ERS)과 관련한 연구에서 활발하게 활용되고 있다. 나아가 ERS로 인해 발생할 수 있는 자료의 이질성을 탐색하는 방법으로 IRTree 모형에 잠재집단모형을 추가한 혼합문항반응나무모형(Mixture Item Response Tree: MixIRTTree)을 제안되고 있다. 본 연구에서는 MixIRTTree 모형의 적용 가능성을 탐색하기 위하여 한국청소년정책연구원의 한국 아동·청소년 패널조사의 부적정서 문제 중 신체적 증상과 우울 척도를 활용하였다. 연구 결과에 따르면 서로 다른 수준의 ERS를 보이는 세 개의 잠재집단을 확인할 수 있었다. 첫 번째로, 잠재집단 3은 잠재집단 1과 2에 비하여 적은 ERS 경향을 보였고, 잠재집단 1이 모든 척도에서 가장 강한 ERS 경향을 보였다. 두 번째로, ERS 경향을 보인 잠재집단 1과 2에서는 척도에 따라서 ERS 경향이 다르게 나타났다. 잠재집단 1은 우울 척도에서 상대적으로 더 강한 ERS 경향을 보였으며, 잠재집단 2는 신체적 증상 척도에 상대적으로 더 강한 ERS를 보이는 것으로 나타났다. 이와 같은 결과를 바탕으로, 본 연구에서 제안된 MixIRTTree와 같은 심리 측정모형을 활용하여 잠재목표특성치와 ERS를 동시에 측정하면서 ERS로 발생할 수 있는 잠재 집단 간 차이를 탐색하는 것이 필요함을 경험적 자료 분석을 통하여 시사하였다.

주요어 : 문항반응나무모형, 혼합문항반응이론, 극단반응양식, 한국 아동·청소년 패널조사

* 주저자: 이주연, 중앙대학교 인지자원개발전략 연구소 연구교수, 서울특별시 동작구 흑석로 84 303관 1419호, JyeonLeeKang@gmail.com

† 교신저자: 장윤선, 대구교육대학교 교육학과 조교수, 대구광역시 남구 중앙대로 219 인문사회관 516호, ysj@dnue.ac.kr

자기 보고식 설문 조사는 성격이나 태도 같은 잠재적 심리 특성을 측정하기 위해서 일반적으로 널리 사용되고 있다(Bolt & Johnson, 2009; Jin & Wang, 2014). 자기 보고식 설문의 결과를 활용할 때는 보통 관측된 점수(observed score)를 사용하게 되는데, 이러한 관측점수는 해당 설문이 측정하고자 하는 잠재 특성과 함께 응답자의 응답 편향(response bias)에 영향을 받게 된다(Jin & Wang, 2014). 응답 편향이란 측정하고자 하는 특성이 아닌 다른 요인으로 인하여 발생하는 응답의 체계적 경향성을 의미한다. 예를 들면 사회적으로 바람직하게 응답하는 사회적 바람직성 편향(social desirability bias)이나, 문항의 내용과 관계없이 응답자가 문항 응답에 대한 동기 부족 등의 이유로 발생하는 무작위 응답(random guessing)이나 부주의 응답(careless response)이 대표적인 예이다.

이 중 응답반응양식(Response Styles: 이하 RS)은 일반적으로 문항의 내용과는 관계없이 모든 문항에 걸쳐 체계적(systematic)으로 발생하는 응답 편향의 일종으로, 특정 응답지에 대한 응답자의 개인적인 선호양식을 뜻한다(Cronbach, 1946, 1950; Messick, 1991). 예를 들면, ‘매우 그렇다’ 또는 ‘매우 그렇지 않다’ 같은 극단반응을 선호하는 극단응답반응양식(Extreme Response Style: 이하 ERS), 5단계 리커트 척도에서 중간응답인 ‘보통이다’를 선호하는 중간응답반응양식(Middle Response Style: 이하 MRS) 같은 것이 대표적이다.

많은 선행연구에서는 실제 자료 분석을 통해 RS의 존재를 증명하였고, 나아가 RS로 인한 구인 간 상관관계수, 척도의 신뢰도와 요인 구조의 왜곡 같은 문제가 발생할 수 있음을 지적하였다(Johnson, Kulesa, Cho, & Shavitt,

2005; Plieninger & Meiser, 2014; 2016; Weijters, Geuens, & Schillewaert, 2010; Wetzel, Carstensen, & Böhnke, 2013). Cronbach(1946)는 응답자들이 동일한 지식 수준과 태도를 가지고 있음에도 불구하고 RS로 인해 관측점수가 서로 달라질 수 있다고 보고하였다. 만약 의료적 진단이나 학생의 선발 및 분류에 자기 보고식으로 수집된 관측점수를 사용할 경우, RS로 인한 결과의 왜곡은 더 심각한 문제가 될 수 있다(Gollwitzer, Eid, & Jürgensen, 2005). 따라서 자기 보고식 설문자료를 분석하기에 앞서 응답자들의 RS가 어느 정도인지, RS의 정도가 최종 결과에 어떤 영향을 주는지 확인해 볼 필요가 있다.

RS의 영향을 탐색하는 방법으로는 별도의 통계적 모형을 사용하지 않고 응답의 빈도를 바탕으로 분석하는 방법(model-free approach)과 RS에 대한 통계적 가정을 반영한 모형 기반 방법(model-based approach)이 있다(Böckenholt, 2017). 모형에 기반하지 않고 RS를 측정하는 방법은 주로 해당 RS에 응답한 빈도를 중심으로 계산된다(De Beuckelaer, Ruten, & Weijters, 2009). 예를 들어, 5점 리커트 척도에서 ERS를 측정하는 경우는 $[f(1)+f(5)]/K$ 로, K 는 전체 문항 수를 뜻하며, $f(1)$ 은 전체 문항 수 중 1(매우 그렇지 않다)에 응답한 문항의 수를 뜻한다.

그러나 이 방법은 RS를 체계적인 잠재특성치로 고려하지 않기 때문에, 여전히 측정하려고 했던 잠재특성치와 RS를 따로 구분하여 측정하는 것이 어렵다(Böckenholt, 2017). 이 경우 RS 측정 후 사후 처리에 어려움이 있을 수 있다. 예를 들어, 응답자가 무작위로 응답지를 선택해서 발생하는 부주의한 응답의 경우 해당 응답을 이상치(outlier)로 간주하여 전체

설문 결과에서 제거하면 된다(Meade & Craig, 2012). 그러나 RS를 개인의 체계적인 잠재특성으로 고려하면 응답자에 따라 RS의 정도에 차이가 있으므로, 어느 정도의 RS를 이상치로 분류할지 그 기준값을 설정하는 것이 어려울 수 있다. 즉, 통계적 모형에 기반하지 않고 RS의 정도를 임의의 기준에 따라 처리한 후 분석을 실시할 경우, 그 결과의 타당성에 대한 의문이 발생할 수 있다. 따라서 문항의 내용과 관련된 잠재특성치와 RS에 대한 잠재특성치를 동시에 측정하는 심리측정모형을 사용할 필요가 있다(Sohn, 2016).

RS 측정을 위한 IRT 심리측정 모형

RS를 잠재특성치로 가정하고, 하나의 심리측정 모형에서 측정하고자 했던 목표잠재특성치(target latent trait)와 함께 추정함으로써, 목표잠재특성치 측정의 정확도를 증가시킬 수 있다(Bolt & Johnson, 2009; Bolt & Newton, 2011; Jin & Wang, 2014). 이를 위해 많은 선행 연구에서는 문항반응이론(Item Response Theory: 이하 IRT)을 RS에 대한 잠재특성치를 측정하기 위해 활용해왔다(Bolt & Johnson, 2009; Bolt & Newton, 2011; Jin & Wang, 2014; Jin, Chen, & Wang, 2018). 예를 들면 Bolt와 Johnson(2009)은 RS와 목표잠재특성치를 동시에 측정할 수 있는 다차원 명명 반응 모형(multidimensional nominal response model)을 제안하고, 이 모형을 통해 RS가 교정된 목표잠재특성치를 측정하는 것이 자료 해석의 왜곡을 줄일 수 있음을 보였다. Jin, Chen과 Wang(2018)의 연구에서는 혼합다분문항반응모형(mixture polytomous IRT)을 활용하여 응답자를 정상반응, ERS, MRS를 보이는 세 개의 잠재집

단으로 구분하는 것을 제안하였다.

최근 IRT를 활용하여 RS를 측정하는 새로운 모형으로 문항반응나무(Item Response Tree: 이하 IRTree) 모형이 제안되었다. IRTree 모형은 응답자가 최종적으로 특정 응답지를 선택할 때까지, 다수의 심리적 결정과정(psychological decision process)을 거친다고 가정하고 이러한 과정을 나무구조로 구성하여 문항반응이론과 혼합한 모형이다(De Boeck & Partchev, 2012; Jeon & De Boeck, 2016). IRTree 모형에서 응답자의 최종 선택은 연구자가 사전에 가정한 나무구조를 따르게 되며, 이 때 사용되는 나무구조는 연구자가 가지고 있는 인지과정에 대한 가설이나 이론적 근거에 따라 다양하게 구성될 수 있다. 예를 들어 4점 리커트 척도를 가정하면, 제시된 문항에 대해 긍정 또는 부정의 응답방향(그렇다 또는 그렇지 않다)을 먼저 결정하고 이어서 의견의 강도(매우 그렇다 또는 그렇다)를 결정하는 응답자의 인지과정을 단계적인 나무구조로 구성할 수 있다(Tourangeau, Rips, & Rasinski, 2000). 이는 설문 응답 시 발생할 수 있는 다양한 인지과정을 반영할 수 있는 IRTree 모형의 장점으로, 최근 IRTree 모형은 RS를 측정하는 많은 연구에 활용되고 있다(Lee & Jang, 2021; Jang, Kim, & Lee, 2019; Böckenholt, 2012, 2014, 2017, 2019; Böckenholt & Meiser, 2017; Jeon & De Boeck, 2016; Jeon & De Boeck, 2019; Pleninger & Meiser, 2014).

혼합문항반응나무모형의 적용

최근 선행연구에서는 IRTree 모형을 혼합문항반응모형으로 확장한 혼합문항반응나무모형(Mixture item response tree model: 이하

MixIRTtree model)이 제안되었다(Khorramdel, von Davier, & Pokropek, 2019; Kim & Bolt, 2021; Tijmstra, Bolsinova, & Jeon, 2018). MixIRTtree 모형은 IRTree 모형에 혼합모형을 적용함으로써, RS의 차이에 따른 응답자 간 이질성(heterogeneity)을 잠재집단으로 탐색하는 것이 가능하다. 여기서 RS로 인한 응답자 간 이질성이란, RS의 특성 때문에 응답자들 사이에 척도에 응답하는 인지과정에서 차이가 발생하는 것을 뜻한다. 예를 들면, ‘잘 모르겠다’같은 중립적 성향을 물어보는 경우, 어떤 응답자들은 이 응답지를 ‘동의한다’와 ‘동의하지 않는다’의 중간으로 인식하지만, 어떤 응답자들은 자신의 의견을 솔직하게 표현하기를 꺼리거나 해당 척도가 묻는 내용 자체에 관심이 없기 때문에 이 중립적 선택지를 선택할 수 있다(Baumgartner & Steenkamp, 2001). 이렇게 ‘보통이다’나 ‘잘 모르겠다’같은 중간 응답지를 선택할 때, 한 자료에 응답자들 간의 서로 다른 인지과정이 혼합되어 있을 경우, 단일 IRTree 모형을 가지고 MRS를 측정하게 되면 MRS 측정의 안정성과 정확성을 보장하기 어렵다. Tijmstra et al.(2018)은 이러한 이론적 가설을 바탕으로 중립적 선택지에 대한 전혀 다른 인지과정을 측정하기 위해서 두 개의 잠재집단을 가정하고, 하나의 잠재집단은 중립적 선택지에 대한 인지과정을 나무구조로 분리한 IRTree 모형을 적용하고, 다른 잠재집단에는 일반화부분점수모형(generalized partial credit model)을 적용하였다. 이를 통해 Tijmstra et al. (2018)는 잠재집단에 서로 다른 측정모형을 적용하여, MRS와 관련하여 응답자 간 서로 다른 인지과정을 사용할 수 있음을 경험적 자료 분석을 통해서 보여줬다.

반면 ERS와 관련해서는, ERS가 잠재목표특

성치에 영향을 받는 특성인지 아닌지에 대한 가정에 따라 응답자들 간의 차이가 발생할 수 있다. 전통적으로 ERS는 잠재목표특성치, 즉 척도의 내용과는 관계없이(content-irrelevant) 체계적으로 발생하는 특성으로 정의되었지만(Paulhus, 1991), 최근 연구에서는 일부 응답자에게는 ERS가 척도의 내용에 영향을 받는 특성(content-relevant)으로 보고되고 있다(Kim & Bolt, 2021; Meiser, Plieninger, & Henninger, 2019). 예를 들면, 어떤 응답자들은 높은 잠재목표특성치 때문에 ‘매우 그렇다’를 선택할 수 있다(Kim & Bolt, 2021). 반면 어떤 응답자들이 문항에 빠르게 응답하거나 응답 시 단순히 사고하는 특성(Naemi, Beal, & Payne, 2009) 때문에 ERS가 높다면, 이 경우의 응답자들의 ERS는 잠재목표특성치에 영향을 받지 않게 된다. Kim과 Bolt(2021)의 연구에서는 두 개의 잠재집단을 사전에 정하고, 잠재집단 1에는 잠재목표특성치와 ERS가 독립된 잠재특성으로 측정하는 IRTree 모형을 적용하고, 잠재집단 2에는 ERS가 잠재목표특성치에 영향을 받는 것을 가정한 IRTree 모형을 적용하였다. 이러한 MixIRTtree 모형 적용을 통하여 ERS가 잠재목표특성치에 영향을 받는 특성에 관한 경험적 근거를 더하였다.

측정 구인 특성에 따른 RS의 이질성

이러한 ERS나 MRS로 인한 이질성에 관한 논의 외에도, 선행연구에서는 RS가 다수의 척도에서 일관적이고 체계적으로 발생하는 잠재적 특성치인가에 대한 논의가 진행중에 있다. 전통적으로 RS는 척도의 종류와 관계없는 일관적인 특성으로 여겨져 왔다(Bolt & Newton, 2011; Jackson & Messick, 1958; Paulhus, 1991;

Wetzel et al., 2013). 예를 들면, 강한 ERS 경향성을 가진 응답자는 척도의 종류에 관계없이 일관되게 강한 ERS를 보이는 것으로 가정되어 왔다. 그러나 최근 연구들은 RS가 척도에 따라 다르게 발생할 수 있음을 지적하였다 (Ames, 2022; Weijters et al., 2010). Ames(2022)의 연구에서는 서로 다른 세 개의 척도를 활용하여, 첫 번째 IRTree 모형은 각 척도의 잠재목표특성치, ERS, MRS를 모두 구분하여 추정하는 다차원IRTree를 가정하고, 두 번째 IRTree 모형은 각 척도의 잠재목표특성치는 구분하여 측정하지만 ERS와 MRS는 모든 척도에 대해 동일하다고 가정하고 측정하였다. 두 가지 IRTree 모형의 비교를 통해서 ERS와 MRS는 서로 다른 세 개의 척도에서 안정적이고 체계적으로 발생하는 특성이 아니고, 척도에 따라 다른 경향성을 보이는 것으로 나타났다.

이렇게 척도의 구인 특성에 따라 RS가 다르게 나타날 수 있음에도 불구하고, 이러한 이질성을 IRTree를 활용하여 분석한 경험적 연구는 아직 충분하지 않다. 척도의 구인의 특성에 따른 RS의 이질성을 IRTree 모형에 반영하는 방법은 크게 두 가지가 가능하다. 하나는 Ames(2022)의 연구와 같이 RS를 각 척도마다 다른 차원(dimensionality)으로 가정하여 다차원 모형을 활용하는 것이다. 이 경우, RS는 측정하는 구인마다 다른 차원의 잠재특성으로 추정된다.

다른 방법으로는 척도의 구인에 따른 RS의 이질성으로 인하여 모집단이 두 개 이상의 잠재집단으로 구성되어 있다고 가정하는 것으로, 이 경우는 모집단 분포를 혼합분포로 보는 MixIRTree 모형을 활용할 수 있다. 예를 들어, ERS를 반영하는 IRTree 모형을 모든 잠재

집단에 동일하게 적용하고 최적의 잠재집단 수를 선택한 후, 특정 잠재집단에서 척도별로 ERS의 정도가 다른지, 아니면 체계적으로 발생하는지를 탐색적으로 살펴볼 수 있다. 이러한 접근은 잠재집단의 특성을 RS에 관한 연구자의 가설에 따라 사전에 정의하는 기존의 확인적 접근과는 달리(Kim & Bolt, 2021; Tijmstra et al. 2018), 척도 구인에 따른 RS의 이질성에 대한 다양한 양상을 탐색적으로 살펴볼 수 있다는 장점이 있다.

따라서 본 연구는 MixIRTree모형을 통해 심리척도에서 척도의 구인 특성으로 인한 응답자 간 ERS 이질성에 관한 특성을 탐색하고 이에 대한 경험적 증거를 더하고자 하였다. 이를 위해 본 연구에서는 (1) ERS를 보이는 응답 과정은 잠재집단별로 동일하고, (2) 부정정서와 관련된 두 잠재특성치를 측정하는 척도를 사용하여 잠재집단은 RS에 의해서만 존재하고, 부정정서에 의한 잠재집단은 존재하지 않는 것으로 설정하여 ERS에서 발생하는 이질성의 특징에 대해서 집중적으로 조사하고자 하였다. Khorramdel et al.(2019)의 연구에서도 (1) 단일 IRTree 측정 모형을 모든 잠재집단에 가정하고 (2) 탐색적으로 잠재집단의 수를 결정하는 MixIRTree 모형을 활용하였으나, 잠재집단 별로 RS의 수준 차이를 살펴보고 잠재집단 별로 성별, 교육 수준 같은 특성이 어떻게 다른지를 살펴보는데 집중하고 있다. 또한 분석자료가 정서, 성격, 태도 같은 심리적으로 중요한 업무에 필요한 기술을 얼마나 자주 습득하고 노력했느냐에 따른 빈도의 정도를 물어본 자료에 국한되어있다. 만약 척도에 따라 ERS가 발생하는 이유와 양상이 응답자 간에 차이가 있다면, 부정 정서같은 척도의 경우 이러한 응답자 간 이질성은 더욱 두드러

질 수 있다. 이는 부적 정서를 측정하는 척도에서는 응답자가 극단응답을 다른 응답 범주보다 더 빈번하게 선택할 수 있기 때문에 (Deng et al., 2018), 척도에 따른 ERS간의 차이가 존재한다면 이러한 차이가 더 드러날 수 있다.

본 연구의 연구목적은 구체적으로 다음과 같다.

연구 문제 1. 부적정서를 측정하는 척도에서 ERS의 양상이 서로 다른 잠재집단이 존재하는가?

연구 문제 2. ERS의 이질성으로 인한 잠재집단이 존재할 경우, 그 특성은 어떠한가?

연구 문제 3. 동일 척도 내 문항에서 ERS 정도가 다른 문항이 존재하는가?

IRTree 모형

IRTree 모형의 정의

IRTree 모형은 Böckenholt(2012)와 De Boeck과 Partchev(2012)가 제안한 모형으로, IRT 모형에 응답자가 설문에 응답할 때 사용하는 인지과정(cognitive process)을 접목한 모형이다. 이때 응답자의 최종문항반응에 도달하기까지의 인지과정은 일련의 순차적인 의사결정과정으로 이루어지고, 이를 나무구조(tree structure or tree diagram)로 표현하게 된다.

나무구조는 응답자의 인지과정에 대한 연구자의 경험적·이론적 가설에 따라 다양하게 구성될 수 있다(Böckenholt, 2017; Jeon & De Boeck, 2016). 그림 1에서 각각의 인지과정은 결절(node)로 정의되며, 화살표는 각 인지과정

에서의 응답자의 판단 결과와 다음 인지과정의 방향을 나타낸다. 이러한 인지과정을 바탕으로 관측치인 다분 문항반응은 각 의사결정 단계에 따라 이분 문항반응(0 혹은 1)으로 바뀌게 된다. 예를 들면, 응답자가 어떤 문항에 ‘매우 그렇다’(4점)로 응답했을 경우, 그림 1의 왼쪽의 나무구조를 반영하면, 응답자는 첫 번째 결절(인지 과정1)에서 1이라는 가상의 이분 문항반응을 얻게 된다. 다음으로 두 번째 결절(인지과정2)에서는 ‘매우 그렇다’를 선택했으므로 1이라는 가상의 이분 문항반응을 갖는다. 즉, 4점이라는 최초의 관측치인 다분 문항반응은 선행적으로 결정된 나무구조의 인지과정의 수에 따라 1, 1이라는 이분 문항반응으로 구성된다. 이렇게 다분 문항반응을 나무구조를 반영하는 가상의 이분 문항반응으로 변환하는 것을 가상 문항이라고 한다.

그림 1의 왼쪽에 제시된 나무구조는 설문에 응답할 때 일반적으로 발생할 수 있는 네 가지 단계의 인지과정 이론을 바탕으로 한다 (Tourangeau et al., 2000). 첫 번째로, 응답자는 우선 문항을 이해하고 해석하려고 하며 (comprehension), 다음으로 문항과 관련된 자신의 기억(generic memories)이나 지식을 회상하게 된다(retrieval). 이 두 단계에서 응답자는 자신의 정보를 바탕으로 반응 문항의 방향성에 대한 판단을 내리게 된다(judgment). 즉, 이 단계에서 응답자는 먼저 문항의 내용에 동의하거나, 동의하지 않거나 혹은 그렇거나 그렇지 않거나 하는 전반적인 방향을 결정한다. 다음으로 응답자는 문항에 대한 강도를 결정하게 된다.

Böckenholt(2012, 2014)는 Tourangeau의 이론을 활용하여 4점 리커트 척도의 경우, 그림 1의 왼쪽 나무구조를 반영하여 ERS를 측정할

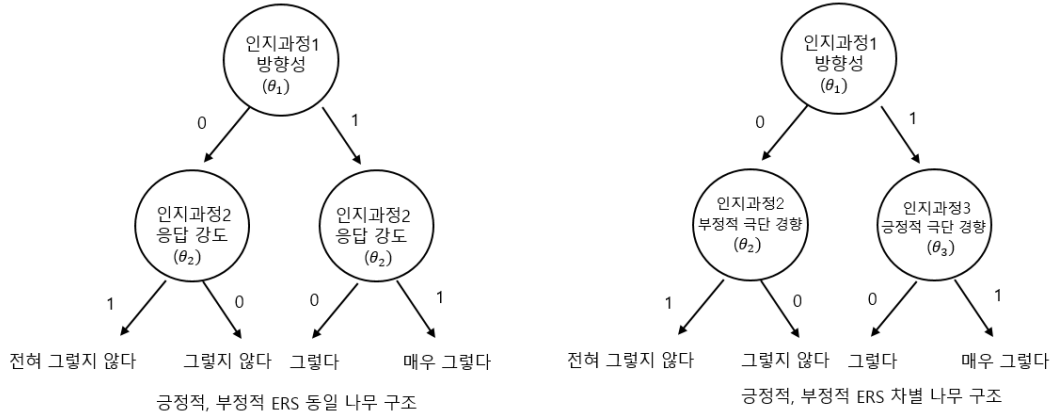


그림 1. ERS 측정을 위한 내재적 나무 구조

것을 제안하였다. 그림 1의 오른쪽 나무구조는 긍정적 방향의 ERS와 부정적 방향의 ERS가 다르게 나타날 수 있는 것을 반영하고 있다. Jeon과 De Boeck(2019)의 연구에서는 일부 문항은 부정방향의 ERS보다 긍정적 방향의 ERS에서 더 낮은 문항곤란도 추정치를 보였다. 즉, 일부 문항에서는 응답자가 긍정적 방향의 극단반응 응답지를 선택하는 것이 부정적 방향의 극단반응 응답지를 선택하는 것보다 쉬울 수 있다.

본 연구에서는 Böckenholt(2012, 2014)의 연구를 바탕으로, 그림 1의 왼쪽의 나무구조를 모든 잠재집단에 동일하게 적용하였다. 표 1은 원래의 관측치인 다분 문항반응이 나무구조를 바탕으로 가상 이분 문항반응으로 변환되는 예시를 보여주고 있다.

IRTree 모형에서 모수추정

IRTree 모형은 가상의 이분 문항반응자료를

표 1. 그림 1의 나무구조를 바탕으로 한 가상 이분 문항반응

다분 문항반응	가상 이분 문항반응				
	긍정적, 부정적 ERS 동일 나무구조		긍정적, 부정적 ERS 차별 나무구조		
	인지과정1: 방향성	인지과정2: 응답강도	인지과정1: 방향성	인지과정2: 부정적 응답강도	인지과정3: 긍정적 응답강도
1	0	1	0	1	NA
2	0	0	0	0	NA
3	1	0	1	NA	0
4	1	1	1	NA	1

참조. 1=매우 그렇지않다; 2=그렇지않다; 3=그렇다; 4=매우 그렇다; NA=결측치

사용하여 인지과정마다 문항 모수와 잠재특성치(이하 능력모수라 명명)를 측정하게 된다. 그림 1에서 왼쪽 나무구조의 θ_1 은 방향성에 대한 능력모수를 뜻하며, θ_2 은 극단반응양식에 대한 능력모수로 정의된다. 각 인지과정은 Rasch나 2모수 로지스틱모형(two-parameter logistic model)같은 이분 IRT모형이 적용될 수 있다. 각 인지과정에 2모수 로지스틱모형을 적용할 때의 문항반응 확률은 다음과 같다.

$$P(Y_{ijp} = 1) = \frac{\exp(\alpha_{ip}(\theta_{jp} - \beta_{ip}))}{1 + \exp(\alpha_{ip}(\theta_{jp} - \beta_{ip}))} \quad (1)$$

Y_{ijp} 는 p 인지과정(방향성 혹은 응답 강도)의 i 번째 문항에서 j 응답자의 가상 이분 문항반응을 뜻하고, θ_{jp} 는 p 번째 인지과정에서 j 응답자의 능력모수를 뜻한다. α_{ip} 는 p 번째 인지과정에서 i 번째 문항의 문항 변별도를 뜻하며, β_{ip} 는 p 번째 인지과정에서 i 번째 문항의 문항 난이도를 뜻한다. 위의 식(1)에서 문항 변별도를 1로 가정하면 Rasch 모형의 문항반응 확률이 된다.

각 인지과정은 독립적으로 발생한다고 가정하므로, 원래의 다분 문항반응에 응답할 확률은 다음 식 (2)와 같다(Jeon & De Boeck, 2019).

$$P(Y_{ij} = y) = \prod_{p=1}^P P(Y_{ijp} = m_p)^{T_{ip}} \quad (2)$$

$$= P(Y_{ji1} = m_1)^{T_{j1}} \times \dots \times P(Y_{jip} = m_p)^{T_{jp}}$$

Y_{ij} 는 i 번째 문항에서 j 응답자의 다분 문항반응을 뜻하고, m_p 는 p 번째 인지과정에서

가상 이분 문항반응을 뜻하며, T_{ijp} 는 0 혹은 1의 값을 갖는 변수로 다분 문항반응이 해당 인지과정에 포함이 되면 1의 값을 가지고, 아니면 0의 값을 가지게 된다. 예를 들면, 그림 1의 부정적 ERS와 긍정적 ERS에 서로 다른 인지과정을 가정할 때는, 부정적 ERS의 T_{ij3} 는 0의 값을 가지게 되고, 반대로 긍정적 ERS의 경우 T_{ij2} 가 0의 값을 가지게 된다.

MixIRTree 모형

MixIRT에서 잠재집단의 이질성은 잠재집단마다 서로 다른 문항 모수 추정치로 표현된다. 따라서 IRTree 모형에 혼합분포를 적용하게 되면, 우선 최적의 잠재집단 수를 결정한 후, 잠재집단 별로 각 인지과정에 대해 서로 다른 문항 모수를 추정하게 된다. 본 연구는 각 인지과정에 Rasch 모형을 적용하였으므로 문항의 변별도는 모두 1로 고정되고 문항 난이도는 잠재집단마다 다르게 추정된다. 즉, Rasch 모형을 적용한 MixIRTree 모형에 따르면 각 인지과정에서 가상 이분 문항반응에 1이라고 응답할 확률은 다음과 같이 정의된다.

$$P(Y_{ijpg} = 1) = \frac{\exp(\theta_{jpg} - \beta_{ipg})}{1 + \exp(\theta_{jpg} - \beta_{ipg})} \quad (3)$$

θ_{jpg} 는 p 번째 인지과정에서 g 번째 잠재집단의 j 응답자의 능력모수를 뜻하며, β_{ipg} 는 p 번째 인지과정에서 g 번째 잠재집단의 i 번째 문항의 문항 난이도를 뜻한다.

전술한 바와 같이 본 연구는 ERS에만 혼합분포가 존재하고 있다고 가정하므로, 각 다분

문항반응에 대한 확률은 그림 1에 제시된 왼쪽의 나무구조에 따라 다음과 같이 정의한다.

$$P(Y_{ij} = y) = P(d_{ij1} = m_1 | \theta_{j1}, \beta_{i1}) \times P(d_{ij2} = m_2 | \theta_{j2g}, \beta_{i2g}) \quad (4)$$

d_{ij1} 는 1번째 인지과정(방향 과정)에서 j 번째 사람의 i 문항에 대한 가상의 이분 문항반응을 뜻하며, m_1 은 0 혹은 1의 응답을 뜻한다. d_{ij2} 는 2번째 인지과정(극단응답)에서 j 번째 사람의 i 문항에 대한 가상의 이분 문항반응을 뜻하며, m_2 은 0 혹은 1의 응답을 뜻한다.

일반적으로 MixIRT에서는 척도의 비결정성 문제(scale indeterminacy)와 잠재집단의 비교를 위해 첫 번째 잠재집단의 능력모수 평균과 표준편차를 각각 0과 1로 고정하고, 다른 잠재집단의 능력모수 분포를 자유롭게 추정한다. 본 연구에서는 모형추정의 안정성을 위하여 모든 잠재집단의 평균을 0, 분산을 1로 고정하였다. 따라서 식 (5)에 제시된 바와 같이 모든 잠재집단의 능력모수는 분산-공분산 행렬이 $p \times p$ 단위행렬인 다변량 정규분포를 따른다.

$$\theta_{jpg} | \Sigma_{pg} \sim MVN(0, \Sigma_p) \quad (5)$$

$$p = 1, \dots, P$$

MVN은 다변량 정규분포를 뜻하며, Σ_{pg} 는 IRTree 모형의 모든 인지과정의 능력모수들의 분산-공분산 행렬을 뜻한다. 이 분산-공분산 행렬은 분산은 1로 고정되고, 공분산은 모든 잠재집단에서 같은 값을 가지는 것으로 가정하였다(Σ_p).

방 법

연구 대상

본 연구는 한국청소년정책연구원에서 제공하는 한국 아동·청소년 패널조사(KCYPS) 중, 2018년 당시 중학교 1학년 대상으로 구축된 3차년도 조사(2020년 실시) 자료를 사용하였다. 패널조사에 응답한 2,590명의 표본 중, 분석에 활용된 신체적 증상과 우울 척도에서의 결측치를 제외한 2,359명의 표본을 분석에 활용하였다. 전체 표본 중 남학생은 1,262명(53%), 여학생은 1,097명(47%) 이었다.

측정도구

본 연구의 목적인 MixIRTree를 적용한 ERS의 이질성탐색을 위해서 KCYPS의 개인발달 척도 중 정서문제를 다루는 구인이 활용되었다. KCYPS의 정서문제는 주의집중, 공격성, 사회적 위축, 신체적 증상과 우울로 구성되어 있다. 본 연구는 척도의 내용과 상관없이 응답자가 동일한 ERS를 보일 것이라고 가정하였으므로, 정서 문제 척도 중 두 개의 척도를 활용하였다. 정서문제의 다른 척도에 비해서 신체적 증상(8문항)과 우울(10문항) 척도는 상대적으로 문항 수가 많아 문항 난이도와 능력모수 추정의 안정성도 보장될 수 있으므로, 신체적 증상과 우울 척도가 사용되었다. 신체적 증상과 우울은 모두 4점 척도로 측정되었으며, 높은 점수는 해당 특성의 정도가 높은 것을 의미한다. 두 척도의 구체적인 측정 문항 및 신뢰도는 표 2와 같다.

표 2. 신체적 증상과 우울 척도 문항 및 신뢰도

요인	측정문항	신뢰도
신체적 증상	(1) 깊이 잠들지 못하고 자다가 잠에서 깨곤 한다	.92
	(2) 머리가 자주 아프다	
	(3) 속이 자주 메스거린다	
	(4) 공부를 하면 배가 아플 때가 있다	
	(5) 입맛이 없을 때가 있다	
	(6) 자주 피곤하다	
	(7) 숨쉬기 어려울 때가 있다	
	(8) 자주 열이 나는 것처럼 느껴진다	
우울	(9) 기운이 별로 없다	.86
	(10) 불행하다고 생각하거나 슬퍼하고 우울해한다	
	(11) 걱정이 많다	
	(12) 죽고 싶은 생각이 든다	
	(13) 울기를 잘한다	
	(14) 어떤 일이 잘못 되었을 때 나 때문이라는 생각을 자주 한다	
	(15) 외롭다	
	(16) 모든 일에 관심과 흥미가 없다	
	(17) 장래가 희망적이지 않은 것 같다	
	(18) 모든 일이 힘들다	

분석 방법

본 연구에서는 그림 1의 왼쪽 나무구조와 같이 긍정적 ERS와 부정적 ERS가 같을 것으로 가정하였다. 따라서 신체적 증상과 우울 척도 각각에서 관측치인 다분 문항반응이 응

답시 발생 할 수 있는 인지과정인 방향과정과 극단응답과정에 대한 가상의 이분 문항반응으로 변환되었다. 분석에 활용된 가상 이분 문항반응의 구체적 예시는 다음 표 3과 같다.

또한, ERS와 관련된 능력모수는 신체적 증상과 우울 척도의 모든 문항을 활용하여 측정

표 3. 가상 이분문항반응

관측된 다분 문항반응		방향과정		극단반응양식	
신체적 증상	우울	신체적 증상	우울	신체적 증상	우울
33323411	3231321142	11101100	1010100010	00000111	0001001110

하고, 여기에 잠재집단을 분석하여 ERS가 척도에 따라 다른 양상을 보이는지를 탐색하였다. 즉, 분석에 측정된 능력모수의 수는 총 3개로, 신체적 증상과 우울 척도의 방향 과정에서의 능력모수와 두 척도의 극단응답과정의 18개 가상 이분문항을 통해서 측정된 ERS 능력모수로 구성되었다.

분석을 위하여 *Mplus* 8.7(Muthén & Muthén, 2017)이 사용되었다. 모수추정을 위해서는 일반적으로 정규성 가정 위반에도 강건(robust)하다고 알려진 MLR(Maximum likelihood with Robust standard errors) 방법을 활용하였다(Muthén, 2004). 본 연구에서는 정보지수를 모형적합도 평가 지수로 활용하여 최적의 잠재집단 수를 탐색적으로 결정하였다. 정보지수로 AIC(Akaike Information Criterion), BIC(Bayesian Information Criterion), SABIC(Sample Adjusted BIC)가 사용되었고, 해당 정보지수의 값이 작을수록 모형적합도가 좋은 것을 의미한다. 또한 LRT(Likelihood Ratio Test)를 사용하여 최종 잠재집단 수를 결정하였다. 나무구조에 따른 가상이분반응(pseudo items)을 생성하는 R코드와 *MplusAutomation R package*(Hallquist & Wiley, 2018)를 사용하였다(부록 참조).

결 과

기술통계

분석에 활용된 신체적 증상과 우울 척도의 전체 문항과 각 문항에 대한 기술통계는 다음 표 4와 같다. 각 척도의 모든 문항에 대한 기술통계를 산출하기 위하여 사용된 값은 신체적 증상 8문항과 우울 10문항의 원점수의

평균값을 사용하였다. 전체 문항에 대한 평균은 신체적 증상이 1.83이고 우울척도 1.79로 나타났다. 표준편차는 각각 .58과 .61로 두 척도의 평균과 표준편차에 큰 차이가 없는 것으로 나타났다. 왜도의 경우 두 척도 모두 정규분포를 의미하는 -1과 1사이였지만, 첨도는 다소 높은 것으로 나타났다.

잠재집단 수 결정

잠재집단의 수를 최종적으로 결정하기 위해서 AIC, BIC, SABIC가 사용되었다. 그림 2에서 나타난 것처럼, 잠재집단의 수가 늘어남에 따라 모든 정보지수가 작아지는 경향을 보였다.

그러나 잠재집단의 수가 증가, 즉 모형이 더 복잡해질수록, 정보지수의 차이의 값이 점점 작아지는 것으로 나타났다. 예를 들면, 잠재집단의 수가 1에서 2로 갈 때 모든 정보지수의 차이의 값이 가장 컸지만, 2에서 4로 갈수록 모든 정보지수의 변화 정도는 점차 작아졌다.

No와 Hong(2012)의 연구에서는 표본 크기가 매우 큰 경우에는 요인분석과 마찬가지로 잠재프로파일 분석에서도 정보지수들의 기울기가 완만해지는 구간에서 최적의 잠재프로파일 수를 결정하는 것이 가능하다고 하였다. MixIRT의 경우에도 Li, Cohen, Kim과 Cho(2009)는 시뮬레이션 연구를 통해 AIC 같은 정보지수는 복잡한 모형을 선호하는 경향을 보였다고 보고하였다. 또한, 잠재집단의 수가 4의 경우에는 일부 문항에서 문항 난이도 모수 추정치가 지나치게 높거나, 표준오차가 측정되지 않았다. 따라서 LRT를 통해 2개의 잠재집단을 가진 모형과 3개의 잠재집단을 가진 모형의

표 4. 신체적 증상과 우울 척도의 기술통계

척도	문항	평균	표준편차	왜도	첨도
신체적 증상	1	1.75	.80	.75	2.71
	2	1.90	.86	.57	2.43
	3	1.74	.78	.73	2.75
	4	1.71	.76	.78	2.96
	5	1.97	.88	.42	2.18
	6	2.32	.93	-.01	2.02
	7	1.64	.75	1.02	3.52
	8	1.66	.77	.99	3.41
	전체	1.83	.58	.46	2.88
우울	9	1.84	.78	.62	2.80
	10	1.74	.77	.81	3.13
	11	2.19	.95	.22	2.01
	12	1.52	.69	1.22	4.09
	13	1.77	.84	.83	2.87
	14	1.85	.84	.66	2.63
	15	1.77	.82	.84	3.01
	16	1.70	.75	.94	3.65
	17	1.77	.81	.80	2.98
	18	1.77	.77	.73	3.00
전체	1.79	.61	.68	3.27	

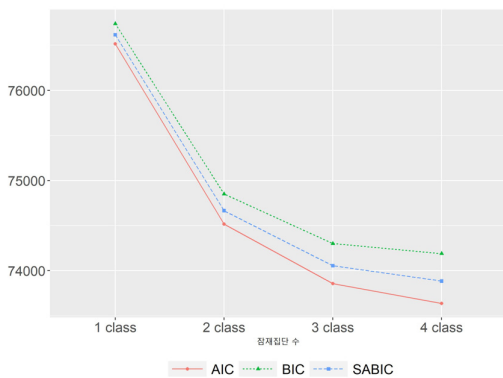


그림 2. 잠재집단 수 별 정보지수

적합도를 비교하여 최적의 모형을 선택하였다. LRT 결과($\chi^2=697.108$, $df=19$, $p < .001$), 잠재집단의 수가 3개인 모형이 최종적으로 선택되었다. 잠재집단으로 분류된 응답자의 수는 잠재집단 1 940명(39.8%), 잠재집단 2 653명(27.7%), 잠재집단 3 766명(32.5%)으로 잠재집단 1의 크기가 다른 잠재집단에 비해 다소 컸다.

나무구조에 따른 문항 난이도 추정치와 잠재 집단 문항 난이도

앞서 기술한 바와 같이 본 연구에서는 응답자가 응답에 반응하는 인지과정을 목표잠재 특성치를 나타내는 방향 과정과 ERS를 나타내는 극단응답과정으로 구분하고, 혼합분포가 ERS에서만 발생할 것으로 가정하였다. 따라서 추정된 문항 난이도는 신체적 증상과 우울 척도 각각의 방향 과정과 두 척도를 합친 총 18 문항에서 발생한 극단응답과정의 두 잠재집단의 문항 난이도가 추정되었다. 표 5에는 신체적 증상과 우울 척도의 방향 과정에서 각 문항에 대한 문항 난이도 추정치가 정리되어있고, 그림 3에는 두 척도의 극단응답과정에 대한 세 개의 잠재집단에서의 문항 난이도 추정치가 정리되어 있다.

방향 과정에서 문항 난이도 추정치가 높다는 것은 응답자가 해당 문항에 ‘그렇다’라고

응답하기 어려움을 뜻한다. 표 5의 신체적 증상 척도를 보면 6번 문항을 제외하고는 문항 난이도 추정치가 다소 높은 것으로 나타났다. 특히 3, 4, 7, 8번 문항의 경우 모두 문항 난이도 추정치가 2를 넘어, 해당 문항의 경우 응답자들이 ‘그렇다’쪽으로 응답하지 않는 것으로 보인다. 6번 문항은 ‘자주 피곤하다($\hat{\beta}= .17$)’로, ‘숨쉬기 어려울 때가 있다(7번 문항, $\hat{\beta}=2.55$)’나 ‘자주 열이 나는 것처럼 느껴진다(8번 문항, $\hat{\beta}=2.44$)’에 비해 응답자들이 더 ‘그렇다’로 응답하기 쉬운 문항으로 해석된다.

표 5의 우울 척도를 보면 11번 문항을 제외하고는 문항 난이도 추정치가 다소 높은 것으로 나타났다. 특히 10, 12, 16, 17, 18번 문항의 경우 모두 문항 난이도 추정치가 2를 넘어, 해당 문항의 경우 응답자들이 ‘그렇다’쪽으로 응답하기 어려운 것으로 보인다. 11번 문항은 ‘걱정이 많다($\hat{\beta}= .55$)’로, ‘죽고 싶은 생각이 든다(12번 문항, $\hat{\beta}=2.97$)’나 ‘모든 일

표 5. 신체적 증상과 우울 척도의 방향과정에서 문항난이도모수 추정치

문항	신체적 증상		문항	우울	
	추정치	SE		추정치	SE
1	1.97	.07	9	1.90	.07
2	1.49	.06	10	2.22	.07
3	2.11	.07	11	.55	.06
4	2.28	.07	12	2.97	.08
5	1.17	.06	13	1.86	.07
6	.17	.06	14	1.66	.06
7	2.55	.08	15	1.99	.07
8	2.44	.07	16	2.55	.08
-	-	-	17	2.01	.07
-	-	-	18	2.15	.07

에 관심과 흥미가 없다(16번 문항, $\hat{\beta}=2.55$)’에 비해 응답자들이 더 ‘그렇다’로 응답하기 쉬운 문항으로 해석된다.

잠재집단별 극단응답반응 경향

그림 3에는 세 잠재집단의 문항 난이도 추정치가 정리되어 있다. 잠재집단별로 다른 문항 난이도 추정치를 해석함으로써 각 잠재집단의 특성을 정의할 수 있다. 극단응답과정에서 문항 난이도 추정치가 높다는 것은 응답자가 해당 문항에 ‘매우 그렇다’나 ‘매우 그렇지 않다’를 응답하기 어려움을 뜻한다. 그림 3을 보면, 잠재집단 3(N=766)은 신체적 증상과 우울 척도의 대부분의 문항에서 문항 난이도 추정치가 2 이상으로 높아(최솟값 문항12: 1.36, 최댓값 문항2: 3.26), 응답자가 극단응답과정에서 극단반응 응답지를 선택하기 어려운 것으

로 보인다. 반면 잠재집단 1(N=940)과 잠재집단 2(N=653)는 극단응답과정에서 문항 난이도 추정치가 잠재집단 3에 비교해서 상대적으로 낮았다. 즉, 잠재집단 3의 응답자보다 상대적으로 잠재집단 1과 2의 응답자는 극단반응 응답지를 쉽게 선택하는 것으로 해석된다.

잠재집단별 극단응답과정에 대한 문항 난이도 추정치의 프로파일에서 주목할만한 점은 신체적 증상과 우울 척도별로 잠재집단별 ERS의 강도가 다르다는 점에 있다. 잠재집단 1과 잠재집단 2의 문항 난이도 추정치는 신체적 증상 척도(문항1-8)에서는 큰 차이가 없지만, 우울 척도(문항9-18)에서는 큰 차이가 나타났다. 나아가 잠재집단별 각 문항의 응답지 선택비율을 비교한 그림 4에 나타난 바와 같이, 잠재집단 1의 경우 두 척도 모두에서 부정적 ERS의 경향이 강하고, 이러한 ERS 경향성은 신체적 증상보다 우울 척도에서 더 강한

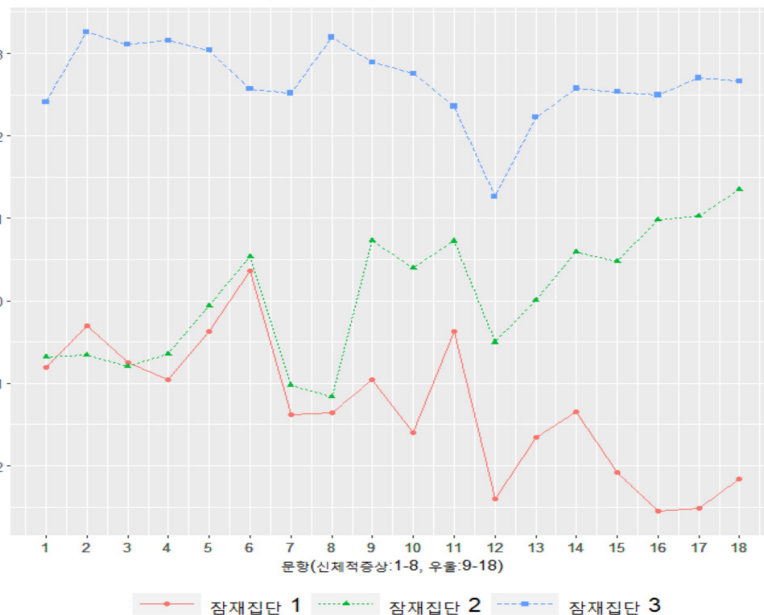


그림 3. 극단응답과정에서 잠재집단별 문항 난이도 추정치 프로파일

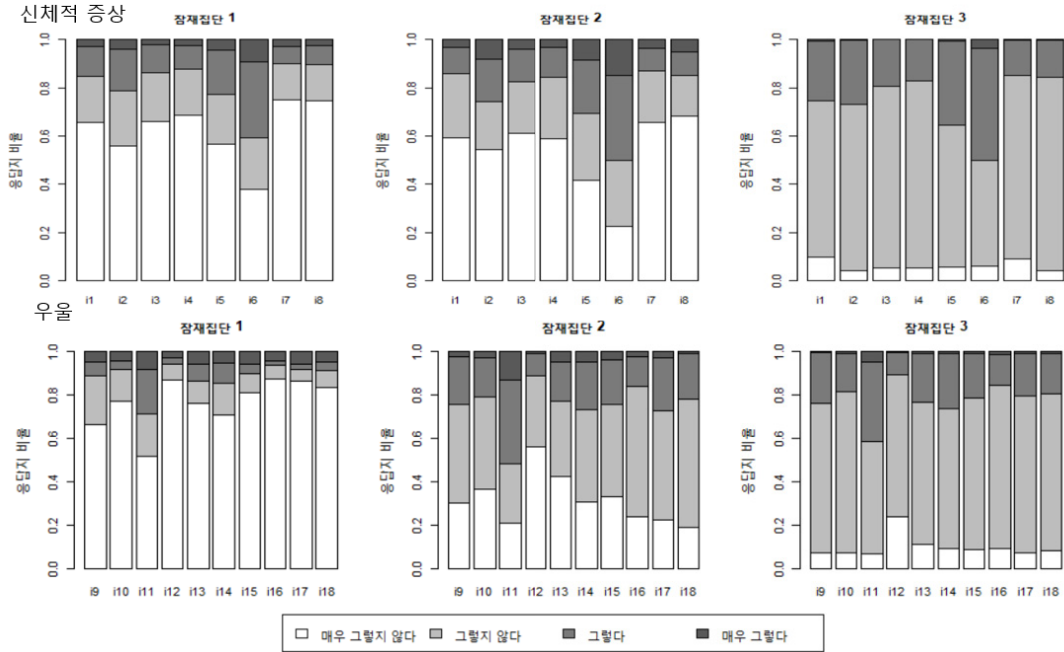


그림 4. 척도와 문항별 4점 리커트 척도 응답 범주 사용 비율

것으로 보인다. 반면에 잠재집단 2의 경우 부정적 ERS의 경향성이 우울 척도보다 신체적 척도에서 더 강한 것으로 보였다.

잠재집단 간 ERS 정도의 차이가 나는 문항은 우울 척도의 문항 12, 16, 17, 18로 보인다. 그림 3을 보면 ERS 경향을 거의 보이지 않는 잠재집단 3에서 다른 문항들에 비해 문항 12의 문항 난이도 추정치가 낮았다. 즉, 잠재집단 3의 응답자들은 다른 문항에 비해서 상대적으로 문항 12 ‘죽고 싶은 생각이 든다’에 대해서는 극단반응 응답지를 다른 문항에 비해서는 쉽게 선택하는 것으로 보인다. 문항 16, 17, 18의 경우 다른 문항들에 비해서 세 개의 잠재집단의 문항 난이도 추정치의 차이가 두드러지는 것으로 나타났다. 같은 우울척도에서도 문항 11 ‘걱정이 많다’에 비해서 문항 16 ‘모든 일에 관심과 흥미가 없다’, 문항 17

‘장래가 희망적이지 않은 것 같다’ 문항 18 ‘모든 일이 힘들다’ 같은 문항에, ERS 경향이 강한 잠재집단 1은 극단반응 응답지를 더 쉽게 선택하고, ERS 경향이 상대적으로 낮은 잠재집단 2는 극단반응 응답지를 오히려 쉽게 선택하지 못하는 것으로 해석된다.

결론 및 논의

본 연구는 최근 RS와 관련하여 다양한 연구가 활발하게 진행되어온 IRTree 모형에 잠재집단이 존재할 것을 가정하는 MixIRTtree 모형을 소개하고 실제 자료를 사용하여 MixIRTtree 모형의 활용 가능성을 보이고자 하였다. 이를 위해 본 연구에서는 혼합분포가 ERS를 보이는 극단응답과정에서만 발생할 것

으로 가정하고 부정정서를 측정하는 두 개의 척도를 활용하여 MixIRTree 모형을 적용하였다. 이러한 가정과 실제 자료 분석을 통해 본 연구에서는 부정정서에서 (1) ERS 경향에 따른 응답자들 간의 이질성을 확인하였고, (2) 척도에 따라 ERS를 보이는 경향이 잠재집단별로 달랐으며, (3) 이러한 차이는 특정 문항에서 두드러지게 나타날 수 있음을 확인할 수 있었다. 연구목적에 따른 본 연구의 구체적인 결과는 다음과 같다.

첫 번째로, 부정정서를 측정하는 척도의 극단응답과정에서 서로 다른 수준의 ERS 경향을 보이는 세 개의 잠재집단을 도출하였다. 예를 들면, 잠재집단 3은 1과 2에 비해 상대적으로 ERS 정도가 적은 집단으로 나타났다. 잠재집단 1의 ERS 정도가 가장 높았으며, 상대적으로 잠재집단 2는 중간 수준의 ERS를 보였다. 이를 통해, 부정정서를 측정하는 척도에서 ERS에 관련하여 응답자 간의 이질성이 ERS 경향의 수준에 따라 존재할 수 있음을 실제 자료를 통하여 확인할 수 있었다.

두 번째로, ERS를 보이는 잠재집단이라도 척도에 따라 다른 ERS 수준을 보이는 것을 확인할 수 있었다. ERS 경향이 강한 잠재집단 1에서는 신체적 증상보다는 상대적으로 우울 척도에서 부정적 ERS가 더 강하게 나타나지만, ERS 경향이 중간 수준인 잠재집단 2의 경우 우울 보다는 상대적으로 신체적 증상 척도에서 부정적 ERS를 활용하는 경향이 강하게 나타났다. 이 결과는 RS가 문항의 내용과 관계없이 모든 문항에서 체계적으로 나타나는 개인의 일관적인 경향성이라는 선행연구(Bolt & Newton, 2011; Jackson & Messick, 1958; Paulhus, 1991; Wetzal et al., 2013)의 견해와 반대된다. 이는 RS가 문항의 내용과 관계없이

발생하는 것이 아니라, 척도가 측정하고자 하는 내용에 따라 다를 수 있음을 지적한 선행 연구들을 지지한다(Ames, 2022; Meiser et al., 2019; Weijters et al., 2010). 즉, 척도의 성격에 따라 응답자의 극단응답반응에 대한 이질성의 정도가 더 두드러지게 나타날 수 있으며, 척도에 따라 ERS에 반응하는 정도가 잠재집단별로 다르게 나타날 수 있음을 의미한다. 예를 들면, 신체적 증상 척도가 응답자의 정서 문제와 관련된 신체적 증상의 빈도를 물어보는 반면, 우울 척도는 응답자의 정서 문제의 심리적 정도를 물어보기 때문에, 이러한 응답 범주에 대한 ERS를 응답자별로 다르게 해석하거나, 척도의 내용에 따라서 응답자들이 극단응답에 다른 반응을 보인 것으로 추측된다.

세 번째로, 척도의 특성뿐만 아니라 문항의 특정한 내용으로도 잠재집단별 ERS의 차이가 있는 것으로 나타났다. 예를 들면, 우울 척도의 문항 12 ‘죽고 싶은 생각이 든다’의 경우 잠재집단 1과 3에서는 같은 우울 척도의 다른 문항에 비하여 문항 난이도 추정치가 낮았다. 이는 ERS가 척도뿐만 아니라 같은 척도의 문항의 내용에 따라서도 영향을 받을 수 있음을 보여준다.

본 연구의 제한점과 이를 바탕으로 한 후속 연구를 위한 제언은 다음과 같다. 첫 번째로, 본 연구에서는 MixIRTree 분석을 위한 여러 가지 강한 가정을 사용하였기 때문에, 본 연구의 결과를 다른 심리척도까지 일반화시키는 것에 신중할 필요가 있다. 본 연구에 사용된 주요 가정으로는 (1) 부정정서 척도에서는 극단반응 응답지를 사용할 확률이 높고 이에 따라 ERS에서만 잠재집단이 발생할 것으로 가정하고, (2) 방향 과정과 극단응답과정(긍정적 ERS와 부정적 ERS가 동일)을 가지는 단일

나무구조를 모든 잠재집단에 적용하였다. 그러나 일부 선행연구에서 긍정적 ERS와 부정적 ERS가 다른 인지과정을 가질 수 있다고 주장한 것을 고려하면(Kim & Bolt, 2021; Jeon & De Beock, 2019), 후속 연구에서는 긍정적 ERS와 부정적 ERS를 분리하여 고려하는 나무구조를 가정하고 MixIRTtree를 시행할 필요가 있다. 실제로 잠재집단별 응답 범주의 사용 비율에서 ERS 경향이 강한 잠재집단 1의 경우 부정적 ERS를 긍정적 ERS보다 많이 사용하는 것으로 나타났다. 다만, 연구자가 관심이 있는 심리척도가 편향된 응답 자료가 많이 관측되는, 즉 정적 혹은 부적 편포가 심한 자료인 경우에는 나무구조에 따른 가상 이분 반응 자료에 결측치(표 1의 긍정적ERS와 부정적 ERS 차별 나무구조의 NA의 경우)가 많이 발생할 수 있으므로 모형 추정의 안정성을 고려하여 MixIRTtree 모형을 적용할 필요가 있다.

이상의 논의를 종합해보면, MixIRTtree 연구를 수행할 때는 연구자가 관심 있는 자료의 특성과 그 자료에서 발생할 수 있는 RS의 특성에 대해서 먼저 고려할 필요가 있다. 이러한 고려를 바탕으로 어떠한 나무구조를 선정할지, 잠재집단이 어떻게 발생할 수 있을지에 대한 가설을 사전에 설계할 필요가 있다. 예를 들면, 어떤 응답자는 긍정적 ERS와 부정적 ERS를 동일한 것으로 인식할 수 있지만, 어떤 응답자는 긍정적, 부정적 ERS를 구별하여 인식할 수 있다. 이 경우 잠재집단별로 단일나무구조를 가정하는 것보다는 확인적 MixIRTtree 선행연구들에서 제안된 것처럼(Kim & Bolt, 2021; Tijmstra et al., 2018) 서로 다른 나무구조가 잠재집단별로 가정될 필요가 있을 것이다.

나아가 본 연구에서는 RS를 잠재특성치로

가정하였지만, 일부 응답자의 경우 무선적으로(random) ‘매우 그렇지 않다’나 ‘매우 그렇다’라는 응답지를 선택할 수 있다. 실제로 본 연구의 잠재집단 1(N=940)에는 모든 문항에 1(매우 그렇지 않다)로 응답한 응답자가 121명에 해당했다. 이러한 응답자들의 ERS가 개인 또는 측정도구 고유의 잠재특성치인지, 아니면 무선으로 발생하는 것인지에 따라 향후 자료의 활용이 달라질 수 있다. 만약 모든 문항에 똑같은 응답 범주를 고른 응답자를 무선적으로 응답한 경우로 고려한다면, 이 경우에는 RS의 잠재특성치를 측정하는 것보다, 자료의 신뢰도와 타당도를 저해하는 이상치로 고려하고 자료에서 삭제하는 것이 추천될 수 있다.

두 번째로, 다양한 RS에 잠재집단이 어떻게 발생하는지에 대한 추가적인 연구가 필요하다. 본 연구는 4점 리커트 척도를 사용하였기 때문에 ERS만이 고려되었다. 그러나 ‘보통이다’같은 3점을 포함한 5점 리커트 척도 자료가 사용될 경우 MRS와 ERS같은 서로 다른 RS가 동시에 발생할 수 있다. 특히, 부적정서를 측정하는 척도에 ‘잘 모르겠다’ 같은 중립적 응답지가 3점으로 사용될 경우, 자신의 부적 감정 상태를 노출하는 것을 회피하는 경향의 응답자는 ‘잘 모르겠다’라는 응답지를 사용하는 것을 다른 응답지보다 더 선호할 수 있다. 이렇게 MRS를 강하게 보이는 응답자가 설문지에 응답할 때 사용하는 인지과정은 ERS를 강하게 보이는 응답자의 인지과정과 다를 수 있다. 나아가 특정 문항에 대해서만 ‘잘 모르겠다’같은 응답지를 강하게 사용할 수도 있으므로, 이러한 문항은 차별기능문항(differential item functioning)으로 작용할 수도 있다. 본 연구에서는 상호 연관된 잠재특성치를 측정하는 다차원 심리검사의 자료에서 발

생활 수 있는 이질적 ERS를 탐색하고 측정하는 것에 있어서 MixIRTree 모형의 적용 가능성 탐색에 중점을 두었다. 하지만 MixIRTree 모형은 단일척도 검사도구를 통해 수집된 자료에서 발생할 수 있는 여러 유형의 RS를 종합적으로 탐색하는 목적으로도 활용할 수 있다. 그러므로 단일 척도에서 다양한 RS가 발생할 수 있는 경우를 고려하여 척도의 성격에 따라 MixIRTree를 적용하는 연구가 필요하다.

마지막으로 심리척도에서 발생할 수 있는 이상점 반응에서 RS가 어떻게 발생하는지에 대하여 MixIRTree를 확장할 수 있다. 심리학 연구에서 활용되는 리커트 척도를 사용하는 자기보고식 설문자료는 문항반응의 기저로 대부분 응답자가 해당 척도의 잠재목표특성치를 강하게 가질수록 해당 문항에 동의할 확률이 단조적으로 증가하는 우세 과정(dominance process)을 가정하여 분석된다(Park, Choi, Lee, Lee, & Lee, 2012; Lee, 2008). 그러나 태도나 성격검사에서는 응답자가 해당 문항에 동의할 확률이 단조적으로 증가하다가 특정 지점 이후 감소하는 ‘이상점 반응 과정(ideal point response process)’이 발생할 수 있다(Part et al., 2012; Lee, 2008). 최근 Jin, Wu와 Chen(2021)의 연구에서는 리커트 척도에서 이러한 이상점 반응과 우세 과정을 동시에 측정할 수 있는 IRTree 모형을 제안하였다. 이 모형에서 이상점 반응에 해당하는 나무구조는 잠재목표특성치를 측정하고 우세 과정을 가정하는 나무구조에서는 ERS를 측정하는 것으로 가정되었다. 후속연구에서는 이러한 IRTree 모형을 MixIRTree로 확장하여, 심리검사에서 우세과정이나 이상점 반응 과정이 복합적으로 발생하는 응답자들의 인지과정을 잠재집단을 통해 탐색해보거나, 이상점 반응 과정에서 RS가 잠

재집단별로 어떻게 다른 양상을 보이는지 탐색하는 것에 적용할 수 있을 것이다.

참고문헌

- Ames, A. J. (2022). Measuring Response Style Stability Across Constructs With Item Response Trees. *Educational Psychological Measurement, 82*(2), 281-306.
<https://doi.org/10.1177/00131644211020103>
- Baumgartner, H., & Steenkamp, J. E. M. (2001). Response styles in marketing research: a cross national investigation. *Journal of Marketing Research, 38*(2), 143-156.
<https://doi.org/10.1509/jmkr.38.2.143.18840>
- Böckenholt, U. (2012). Modeling multiple response processes in judgment and choice. *Psychological Methods, 17*(4), 665-678.
<https://doi.org/10.1037/a0028111>
- Böckenholt, U. (2014). Modeling motivated misreports to sensitive survey questions. *Psychometrika, 79*(3), 515-537.
<https://doi.org/10.1007/s11336-013-9390-9>
- Böckenholt, U. (2017). Measuring response styles in Likert items. *Psychological Methods, 22*(1), 69-83.
<https://doi.org/10.1037/met0000106>
- Böckenholt, U. (2019). Assessing item feature effects with item response tree models. *British Journal of Mathematical and Statistical Psychology, 72*(3), 486-500.
<https://doi.org/10.1111/bmsp.12163>
- Böckenholt, U., & Meiser, T. (2017). Response style analysis with threshold and multi process

- IRT models: A review and tutorial. *British Journal of Mathematical and Statistical Psychology*, 70(1), 159-181.
<https://doi.org/10.1111/bmsp.12086>
- Bolt, D. M., & Johnson, T. R. (2009). Addressing score bias and differential item functioning due to individual differences in response style. *Applied Psychological Measurement*, 33(5), 335-352.
<https://doi.org/10.1177/0146621608329891>
- Bolt, D. M., & Newton, J. R. (2011). Multiscale measurement of extreme response style. *Educational and Psychological Measurement*, 71(5), 814-833.
<https://doi.org/10.1177/0013164410388411>
- Cronbach, L. J. (1946). Response sets and test validity. *Educational and Psychological Measurement*, 6(4), 475-494.
<https://doi.org/10.1177/001316444600600405>
- Cronbach, L. J. (1950). Further evidence on response sets and test design. *Educational and Psychological Measurement*, 10(1), 3-31.
<https://journals.sagepub.com/doi/pdf/10.1177/001316445001000101>
- De Beuckelaer, A., Rutten, A., & Weijters, B. (2009). Using ad hoc measures for response styles: A cautionary note. *Quality & Quantity*, 44(4), 761-775.
<http://doi.org/10.1007/s11135-009-9225-z>
- De Boeck, P., & Partchev, I. (2012). IRTrees: Tree-based item response models of the GLMM family. *Journal of Statistical Software*, 48(1), 1-28.
<https://doi.org/10.18637/jss.v048.c01>
- Deng, S., E. McCarthy, D., E. Piper, M., B. Baker, T., & Bolt, D. M. (2018). Extreme response style and the measurement of intra-individual variability in affect. *Multivariate Behavioral Research*, 53(2), 199-218.
<https://doi.org/10.1080/00273171.2017.1413636>
- Gollwitzer, M., Eid, M., & Jürgensen, R. (2005). Response styles in the assessment of anger expression. *Psychological Assessment*, 17(1), 56-69.
<https://doi.org/10.1037/1040-3590.17.1.56>
- Hallquist, M. N., & Wiley, J. F. (2018). MplusAutomation: an R package for facilitating large-scale latent variable analyses in Mplus. *Structural Equation Modeling: A Multidisciplinary Journal*, 25(4), 621-638.
<https://doi.org/10.1080/10705511.2017.1402334>
- Jackson, D. N., & Messick, S. (1958). Content and style in personality assessment. *Psychological Bulletin*, 55(4), 243-252.
<https://doi.org/10.1037/h0045996>
- Jang, Y., Kim, M., & Lee, Y. (2019). IRTree Model: An Alternative Approach for Self-Reported Ordinal Data Analysis. *Journal of Educational Evaluation*, 32(2), 303-323.
<http://dx.doi.org/10.31158/JEEV.2019.32.2.303>
- Jeon, M., & De Boeck, P. (2016). A generalized item response tree model for psychological assessments. *Behavior Research Methods*, 48(3), 1070-1085.
<https://doi.org/10.3758/s13428-015-0631-y>
- Jeon, M., & De Boeck, P. (2019). Evaluation on types of invariance in studying extreme response bias with an IRTree approach. *British Journal of Mathematical and Statistical Psychology*, 72(3), 517-537.
<http://doi.org/10.1111/bmsp.12182>

- Jin, K. Y., Chen, H. F., & Wang, W. C. (2018). Mixture item response models for inattentive responding behavior. *Organizational Research Methods, 21*(1), 197-225.
<https://doi.org/10.1177/1094428117725792>
- Jin, K. Y., & Wang, W. C. (2014). Generalized IRT models for extreme response style. *Educational and Psychological Measurement, 74*(1), 116-138.
<https://doi.org/10.1177/0013164413498876>
- Jin, K.-Y., Wu, Y.-J., & Chen, H.-F. (2022). A new multiprocess IRT Model with ideal points for likert - type items. *Journal of Educational and Behavioral Statistics. Journal of Educational and Behavioral Statistics, 47*(3), 297-321.
<https://doi.org/10.3102/10769986211057160>
- Johnson, T., & Bolt, D. M. (2010). On the use of factor-analytic multinomial logit item response models to account for individual differences in response style. *Journal of Educational and Behavioral Statistics, 35*(1), 92-114.
<https://doi.org/10.3102/1076998609340529>
- Johnson, T., Kulesa, P., Cho, Y. I., & Shavitt, S. (2005). The relation between culture and response styles: Evidence from 19 countries. *Journal of Cross-cultural Psychology, 36*(2), 264-277.
<https://doi.org/10.1177/0022022104272905>
- Khorramdel, L., von Davier, M., & Pokropek, A. (2019). Combining mixture distribution and multidimensional IRTree models for the measurement of extreme response styles. *British Journal of Mathematical and Statistical Psychology, 72*(3), 538-559.
<https://doi.org/10.1111/bmsp.12179>
- Kim, N., & Bolt, D. M. (2021). A Mixture IRTree Model for Extreme Response Style: Accounting for Response Process Uncertainty. *Educational and Psychological Measurement, 81*(1), 131-154.
<https://doi.org/10.1177/0013164420913915>
- Lee, J., & Jang, Y. (2021). A multilevel item response tree model for measuring a cross-national variability in response styles. *Journal of Educational Evaluation, 34*(1), 131-153.
<http://dx.doi.org/10.31158/JEEV.2021.34.1.131>
- Lee, S. (2008). Application of ideal point response model to personality scale. *Korean Journal of Psychology: General, 27*(3), 839-857.
<https://www.dbpia.co.kr/journal/articleDetail?nodeId=NODE06368101>
- Li, F., Cohen, A. S., Kim, S.-H., & Cho, S.-J. (2009). Model selection methods for dichotomous IRT models. *Applied Psychological Measurement, 33*(5), 353-373.
<http://doi:10.1177/0146621608326422>
- Meade, A. W., & Craig, S. B. (2012). Identifying careless responses in survey data. *Psychological Methods, 17*(3), 437-455.
<https://doi.org/10.1037/a0028085>
- Meiser, T., Plieninger, H., & Henninger, M. (2019). IRTree models with ordinal and multidimensional decision nodes for response styles and trait based rating responses. *British Journal of Mathematical and Statistical Psychology, 72*(3), 501-516.
<https://doi.org/10.1111/bmsp.12158>
- Messick, S. (1991). Psychology and methodology of response styles. In R. E. Snow & D. E.

- Wiley (Eds.), *Improving inquiry in social science: A volume in honor of Lee J. Cronbach* (pp. 161-200). Hillsdale, NJ: Lawrence Erlbaum.
- Muthén, B. O. (2004). Latent variable analysis. In D. Kaplan (Ed.), *Handbook of quantitative methodology for the social science* (pp. 345-368). Thousand Oaks, CA: Sage.
- Muthén, L. K. & Muthén, B. O. (1998-2017). *Mplus user's guide*(7th ed.). Los Angeles, CA: Muthén & Muthén.
- Naemi, B. D., Beal, D. J., & Payne, S. C. (2009). Personality predictors of extreme response style. *Journal of personality*, 77(1), 261-286.
<https://doi.org/10.1111/j.1467-6494.2008.00545.x>
- No, U., & Hong, S. (2012). Identifying latent profiles in the motivations of adolescents' use of computers and testing the effects of eco-system variables. *Studies on Korean Youth*, 23(3), 51-76.
<https://www.dbpia.co.kr/Journal/articleDetail?noDeId=NODE01937598>
- Park, D. -K., Choi, M., Lee, W., Lee, H., & Lee, C. (2012). Comparison between dominance process model and ideal point response model for personality assesment. *Korean Journal of Industrial and Organizational Psychology*, 25(2), 421-452.
<https://www.dbpia.co.kr/journal/articleDetail?noDeId=NODE06370437>
- Paulhus, D. L. (1991). Measurement and control of response bias. In J. P. Robinson, P. R. Shaver, & L. S. Wrightsman (Eds.), *Measures of personality and social psychological attitudes* (Vol. 1, pp. 17-59). Academic Press.
<https://doi.org/10.1016/B978-012-590241-0.50006-X>
- Plieninger, H., & Meiser, T. (2014). Validity of multiprocess IRT models for separating content and response styles. *Educational and Psychological Measurement*, 74(5), 875-899.
<https://doi.org/10.1177/0013164413514998>
- Sohn, W. (2016). Testing the Longitudinal Stability of Extreme and Middle Response Style. *Journal of Educational Evaluation*, 29(2), 279-299.
<https://scholar.kyobobook.co.kr/article/detail/4010024903560>
- Tijmstra, J., Bolsinova, M., & Jeon, M. (2018). General mixture item response models with different item response structures: Exposition with an application to Likert scales. *Behavior Research Methods*, 50(6), 2325-2344.
<https://doi.org/10.3758/s13428-017-0997-0>
- Tourangeau, R., Rips, L. J., & Rasinski, K. (2000). *The psychology of survey response*. New York, NY: Cambridge University Press.
- Weijters, B., Geuens, M., & Schillewaert, N. (2010). The stability of individual response styles. *Psychological Methods*, 15(1), 96-110.
<https://doi.org/10.1037/a0018721>
- Wetzel, E., Carstensen, C. H., & Böhnke, J. R. (2013). Consistency of extreme response style and non-extreme response style across traits. *Journal of Research in Personality*, 47(2), 178-189.
<https://doi.org/10.1016/j.jrp.2012.10.010>

1차원고접수 : 2022. 04. 01.

2차원고접수 : 2022. 05. 14.

최종게재결정 : 2022. 06. 12.

Mixture item response tree model for measuring extreme response style in multiple psychological scales

Juyeon Lee	Jungsu Oh	Yoonsun Jang
Chung-Ang University	Eastern Illinois University	Daegu National University of Education

Item response tree model (IRTtree) is a sort of a multidimensional item response theory that incorporates a tree structure depicting a respondent's cognitive response process into item response theory. IRTtree has recently received considerable attention from researchers for measuring extreme response style (ERS) that can threaten reliability and factor structure in self-report survey data. In addition to the work of IRTtree, we extend IRTtree as the mixture item response tree model (MixIRTtree), which combines latent class analysis with IRTtree for measuring and investigating the possible heterogeneity of data caused by ERS. For analyses, this study used the body symptom and depression scales of the emotional problems in the third wave of KCYPS 2018. The results of the study are as follows. First, the three latent classes were captured, showing the different levels of ERS. The latent class 3 did not show the ERS tendency than the latent classes 1 and 2. The latent class 1 showed the largest ERS tendency. Second, the ERS tendency across the latent classes was different between the negative affect scales. For example, the ERS tendency of the latent class 1 was larger in the depression scale than body symptom scale, while the ERS tendency of the latent class 2 was larger in the body symptom scale than the depression scale. Based on these findings, we provide practical implications of the MixIRTtree for measuring the heterogeneity of ERS and discuss the future direction for further study.

Key words : Item response tree model, Mixture item response tree model, Extreme response style, KCYPS

부록 1. IRTree 모형에 따른 “가상 이분 문항반응(pseudo item)” 생성 R code

```

# 신체적 증상, body symptom, 8 items
# node 1 = agree direction
# BS 척도 원래의 다분 문항 자료
BS.irtree_node1 = BS
  BS.irtree_node1[BS == 1] = 0
  BS.irtree_node1[BS == 2] = 0
  BS.irtree_node1[BS == 3] = 1
  BS.irtree_node1[BS == 4] = 1
  BS.irtree_node1[BS == NA] = NA

names(BS.irtree_node1) = c("BS1", "BS2",
"BS3", "BS4", "BS5", "BS6", "BS7", "BS8")

# node 2= ERS 과정
BS.irtree_node2 = BS
  BS.irtree_node2[BS == 1] = 1
  BS.irtree_node2[BS == 2] = 0
  BS.irtree_node2[BS == 3] = 0
  BS.irtree_node2[BS == 4] = 1
  BS.irtree_node2[BS == NA] = NA

names(BS.irtree_node2) = c("ext1", "ext2",
"ext3", "ext4", "ext5", "ext6", "ext7", "ext8")

# 우울, Depression, 10 items
# node 1 = agree direction
# DP 척도 원래의 다분 문항 자료
DP.irtree_node1 = DP
  DP.irtree_node1[DP == 1] = 0
  DP.irtree_node1[DP == 2] = 0

  DP.irtree_node1[DP == 3] = 1
  DP.irtree_node1[DP == 4] = 1
  DP.irtree_node1[DP == NA] = NA

names(DP.irtree_node1) = c("DP1", "DP2",
"DP3", "DP4", "DP5", "DP6", "DP7", "DP8",
"DP9", "DP10")

# node 2= ERS
DP.irtree_node2 = DP
  DP.irtree_node2[DP == 1] = 1
  DP.irtree_node2[DP == 2] = 0
  DP.irtree_node2[DP == 3] = 0
  DP.irtree_node2[DP == 4] = 1
  DP.irtree_node2[DP == NA] = NA

names(DP.irtree_node2) = c("ext9", "ext10",
"ext11", "ext12", "ext13", "ext14", "ext15", "ext16",
"ext17", "ext18")

# Combine
MULTI.tree = cbind(BS.irtree_node1,
DP.irtree_node1, BS.irtree_node2, DP.irtree_node2)

# NA as -99

MULTI.mplus = MULTI.tree
MULTI.mplus[is.na(MULTI.mplus)] = -99

```

부록 2 *Mplus* automation package를 활용한 MixIRTree 2 class 분석 R code

<pre> library(MplusAutomation) library(rhdf5) library(here) # mixirtree2 = Mplus code mixirtree2 <- mplusObject(TITLE = "MixIRTree:Rasch", VARIABLE = "usevar = BS1-BS8 DP1-DP10 ext1-ext18; categorical = BS1-BS8 DP1-DP10 ext1-ext18; MISSING ARE ALL (-99); CLASSES = c(2);", ANALYSIS = "ALGORITHM = INTEGRATION; TYPE=MIXTURE; STARTS = 500 10; processors = 2;", MODEL = " %OVERALL% BSf by BS1-BS8@1; DPf by DP1-DP10@1; fext by ext1-ext18@1; %c#1% ! Latent group 1 [BS1\$1-BS8\$1] (BSB1-BSB8); [DP1\$1-DP10\$1] (DPB1-DPB10); [ext1\$1-ext18\$1] (extC1B1-extC1B18); </pre>	<pre> [BSf@0]; BSf@1; [DPf@0]; DPf@1; [fext@0]; fext@1; %c#2% ! Latent group 2 [BS1\$1-BS8\$1] (BSB1-BSB8); [DP1\$1-DP10\$1] (DPB1-DPB10); [ext1\$1-ext18\$1] (extC2B1-extC2B18); [BSf@0]; BSf@1; [DPf@0]; DPf@1; [fext@0]; fext@1; ", OUTPUT = "TECH1 TECH8; STANDARDIZED;", SAVEDATA = "FILE= MixIRTREE2_FSCORES.dat; SAVE = cprob fscores;", usevariables = colnames(MULTI.mplus), rdata = MULTI.mplus) # end of Mplus code (t1.MixIRtree2.Mplus <- Sys.time()) set.seed(1987) MixIRtree2.fit = mplusModeler(mixirtree2 , dataout=here("2.mix2.dat"), modelout=here("2.mix2.inp"), check=TRUE, run = TRUE, hashfilename = FALSE) t2.MixIRtree2.Mplus <- Sys.time() t2.MixIRtree2.Mplus - t1.MixIRtree2.Mplus </pre>
--	--