한국심리학회지: 일반
Korean Journal of Psychology: General
2020, Vol. 39, No. 2, 175-204
http://dx.doi.org/10.22257/kjp.2020.6.39.2.175

범주형 확인적 요인분석 모형의 다집단 확장*

신 미 미[1]　　　　　김 수 영†

범주형 확인적 요인분석은 심리학을 비롯한 사회과학 분야에서 빈번히 활용되는 범주형 변수(예, 리커트 척도 등)의 특성을 분석에 반영한 측정모형이다. 해당 모형의 다집단 확장에서 측정불변성을 확인하기 위해서는 단일 집단의 경우보다 모형 판별을 위하여 더 세분화된 척도화 및 제약이 수반되어야 하며, 측정불변성의 확인 단계 및 중요도 역시 연속형 지표변수를 상정한 경우와는 달라진다. 본 연구는 범주형 확인적 요인분석의 측정불변성과 관련하여 지금까지 이루어져 온 전체적인 논의를 탐색, 취합하여 연구자들이 자신의 연구 방향에 부합하는 판별 방식 및 측정불변성 확인 단계를 결정할 수 있도록 적절한 합의점을 도출하는 것을 목적으로 한다. 이를 위하여 두 가지 측면에서 논의가 이루어진다. 첫째, 모형의 판별 부분에서는 제한정보추정에 근거하여 두 종류의 잠재변수에 대한 척도화 및 다집단 분석을 위한 추가제약 방식을 다룬다. 둘째, 측정불변성 확인 단계에서는 범주형 지표변수를 상정하며 나타난 경계 모수의 동일성 확보를 중심으로 현재까지 제시되어 있는 측정불변성 확인 단계를 정리한다. 또한 앞서 제시된 판별 방식 및 측정불변성 확인 단계를 통계 프로그램을 이용하여 실제 자료 분석에 적용해 봄으로써 실질적인 활용을 도모한다. 마지막으로 판별 방식 및 측정불변성 확인 단계와 관련한 쟁점을 통합하여 각각의 상황에 부합하는 접근 방식의 특성 및 제한점을 논의한다.

주요어 : 범주형 자료, 다집단 분석, 요인분석, 제한정보추정, 측정불변성

---

* 이 논문은 2019년 대한민국 교육부와 한국연구재단의 지원을 받아 수행된 연구임 (NRF-2019S1A5A2A03041362).

1) 이화여자대학교 심리학과 학생, E-mail: mimiworld212@gmail.com
† 교신저자: 김수영, 이화여자대학교 심리학과, 서울시 서대문구 이화여대길 52
Tel: 02-3277-3792, E-mail: suyoung.kim@ewha.ac.kr

심리학을 비롯한 사회과학 분야에서는 변수들의 관계에 집단이 미치는 영향을 살피는 다집단 분석을 빈번하게 시행한다. 구조방정식의 맥락에서 다집단 분석 시, 집단 간 측정불변성(measurement invariance)을 확보하는 것은 필수적인 전제조건이다. 이는 비교하고자 하는 집단이 동일한 요인구조를 형성하고 있다는 점을 확인해야 이후 진행하는 집단 간 평균 비교, 경로계수 비교의 타당성을 확보할 수 있기 때문이다. 집단 간 측정불변성이 확보되었다는 것은 요인이 주어진 상태에서 나타날 수 있는 관찰된 지표변수의 조건부 확률이 집단 사이에 같다는 것을 의미한다(Mellenbergh, 1989; Meredith, 1993). 지표변수와 요인의 관계를 결정짓는 모수들이 집단 간 동등해야 조건부 확률이 같아지기 때문에 측정불변성 확인을 위해서는 이 모수들이 집단 간 동등한지를 살피면 된다(Meredith, 1964; Elosua, 2011). 일반적으로 각 집단의 요인분석 모형이 같은 형태를 지니고 있는가를 살피는 형태동일성, 요인부하 모수가 동일한 값을 지니고 있는가를 살피는 요인부하동일성, 절편 모수가 동일한 값을 지니고 있는가를 살피는 절편동일성, 지표변수의 오차분산 및 공분산이 동일한 값을 지니고 있는가를 살피는 분산(공분산)동일성 네 가지 또는 그 일부를 순차적으로 확인한다.

측정불변성 확인을 위한 도구로는 확인적 요인분석(confirmatory factor analysis, CFA)이 탐색적 요인분석(exploaratory factor analysis, EFA)에 비해 더 빈번하게 채택된다(Flora & Curren, 2004; Schreiber, Nora, Stage, Barlow & King, 2006). CFA는 지표변수와 잠재변수의 선형관계를 가정하기 때문에 지표변수로서 관계의 선형성 확보가 용이한 연속형 변수를 선호한다(Lee, Poon, & Bentler, 1992). 그런데 CFA의 지표변수로 자주 활용되는 리커트 척도는 그 본질이 순위범주형이기 때문에 이를 연속형으로 취급하는 것은 문제를 일으킬 수 있다. 특히 범주의 수가 적어질수록 문제가 발생할 가능성이 높아지며(Johnson & Creech, 1983), 이때 발생할 수 있는 문제로는 편향된 모수 추정치 및 부정확한 표준오차와 검정통계량 등이 있다(Bernstein & Teng, 1989; Rhemtulla, Brosseau-Liard, & Savalei, 2012). 만약 측정모형에서 이와 같은 문제가 발생하면 이후 요인들 간 구조 관계를 살피고 집단 간 비교를 진행하는 것은 무리가 있다. 이러한 문제점을 보완하기 위하여 범주형 지표변수의 특성을 보다 정확하게 반영하는 범주형 확인적 요인분석(categorical CFA)을 적용할 필요가 있으며(Beauducel & Herzberg, 2006), 이는 다집단 분석으로의 확장에서도 마찬가지이다.

다집단 분석 시 측정모형으로 범주형 CFA를 선택하는 것은 단순히 연속형 CFA에서 추정 방식만 바꾼 것이 아니다. 모형의 판별을 위해 잠재변수들에 단위를 부여하는 척도화(scaling) 방식부터 측정불변성을 확보하기 위한 동일성 검정 단계까지 모두 변화하므로 다각적인 측면을 고려해야 한다. 이러한 척도화 방식과 동일성 검정 모두 연구자의 주관이나 연구 목적에 따라 다양한 이론과 시행 방식이 존재한다. 그런데 이와 관련한 논의들은 그동안 각종 문헌에서 산발적으로 조금씩 드러나 있었을 뿐 통합적인 시각에서 다루어진 바 없다. Yu, Millsap, West, Tien, Tanaka와 Grimm(2017)이 범주형 CFA를 사용한 종단 자료에서의 측정불변성 검정 과정에 대해 논의한 바 있으나, 종단 자료의 특성상 다층모형(Multilevel modeling)의 관점하에서 활용해야 한

다. 이에 본 연구에서는 범주형 CFA의 다집단 확장과 관련하여 지금까지 발표된 문헌을 탐색, 정리하여 각각의 제약방식 및 측정불변성 확인 단계들이 지니는 특성과 사용 시 유의점을 통합한다. 여기에서 더 나아가 통계 프로그램을 이용하여 실제 자료를 분석하는데 적용해 봄으로써 연구자들에게 실질적으로 유용한 정보를 제공하고자 한다.

본격적인 논의에 앞서, 본 연구에서는 다분상관 혹은 사분상관과 같은 자료의 요약치를 이용하는 제한정보추정(limited information estimation) 방식으로 범주형 CFA의 다집단 확장을 다룬다는 점을 밝힌다. 제한정보추정은 자료에 나타난 개별 반응값을 모형의 추정과정에 이용하는 완전정보추정(full information estimation)에 비해 수리적으로 더 편리한 방식이며, 범주형 CFA와 관련한 시뮬레이션 연구들에서도 제한정보추정이 완전정보추정과 유사하거나 더 나은 수행 능력을 보였다는 연구도 있기 때문이다(Forero & Maydeu-Olivares, 2009; Rhemtulla et al., 2012). 게다가 범주형 CFA의 완전정보추정은 구조방정식의 맥락에서 사용되는 여러 모형적합도 지수에 대한 제대로 된 합의가 이루어지지 않아 실질적으로 통계 프로그램을 이용하는 데 있어 어려움을 야기할 수 있다(Muthén, 2006)는 점도 고려하였다.

범주형 CFA의 다집단 확장은 두 가지 측면에서 살펴보아야 한다. 첫 번째 측면은 모수의 추정과 관련한 것으로 모형의 판별을 위한 제약 방식과 관련한 논의이다. 모형의 판별을 위해서는 잠재변수들에 단위를 부여하는 필수 제약인 척도화 방식에 대한 고찰이 필요하며, 이와 더불어 다집단 확장 시 추정 자체를 가능하게 만들기 위한 추가제약과 관련한 명확한 논의도 이루어져야 한다. 두 번째 측면은 측정불변성 확인 단계로, 지표변수와 요인의 관계를 결정짓는 모수들이 집단 간 동등한지를 살피는 절차와 관련한 것이다. 이 모수들 중 어떤 종류의 모수들이 집단 간 동등해야 하는가는 측정불변성을 어떻게 정의하는지에 따라 달라진다(Byrne, Shavelson, & Muthén, 1989; Cheung & Rensvold, 2000; Little, 1997; Meredith, 1964; Vandenberg & Lance, 2000). 지표변수가 범주형인 경우 측정불변성 확인과 관련해서는 범주형 분석에서의 주요 모수 중 하나인 경계(threshold)의 동일성을 어떻게 다룰 것인지가 주요 논점이 된다. 또한 모형의 판별을 위하여 선택한 척도화 방식은 이후 측정불변성을 확인하는 단계에 영향을 미치기 때문에 이 둘은 함께 다루어질 필요가 있다.

국내에도 범주형 지표변수의 본질을 고려한 측정모형의 개념을 소개하고 그 적용법을 모색한 방법론 문헌들이 있다. Kim(2016)은 잠재반응변수(latent response variable)의 개념을 이용하는 범주형 CFA를 소개하며 제한정보추정의 일종인 평균 및 분산조정 가중 최소제곱(mean and variance adjusted weighted least squares, WLSMV) 추정법을 설명하였다. 하지만 김수영(2016)은 범주형 CFA의 다집단 분석 확장에 대한 논의는 제공하지 않고 있다. Lee, Youn, Lee와 Jung(2016)은 탐색적 요인분석의 맥락에서 WLSMV 추정법의 특성에 대하여 소개하고, 실제 자료를 가지고 분석을 실시하는 예시를 보였다. Seo와 Lee(2017)는 제한정보추정 방식으로 문항의 난이도 및 변별도를 추출하여 심리검사 제작을 위한 요인분석 상황에서 문항반응이론적인 접근이 가능함을 시사했다. 하지만 두 연구 역시 주요 관심사가 각각 탐색

적 요인분석 및 개별 문항 수준에 있었기 때문에 척도를 개발하고자 하는 연구자들에게는 도움이 될 수 있지만, 요인 사이의 구조 관계를 파악하고 특히 이를 다집단 상황으로 확장하고자 하는 연구자들에게 도움을 주기엔 부족하다.

국내 내용영역 연구 중 범주형 CFA를 통해 집단 간 측정불변성을 확인한 연구로는 Lee, Chun과 Lee(2014)가 있다. Lee 등(2014)은 범주가 3개인 지표변수를 이용하는 확인적 요인분석 모형을 WLSMV 방법으로 추정하여 한국과 미국 두 집단 사이의 측정불변성을 확인하였다. 이를 분석하는 과정에서 모형의 판별을 위해 사용한 척도화 방식을 밝히고 있으며, 측정불변성 확인 단계에서 경계 모수 동일성이 확보되지 않자 집단 간 차이가 발생하게 된 원인에 대한 설명을 덧붙였다. 그러나 범주형 CFA의 기저에 있는 통계적인 특성(예, 다분상관 또는 사분상관, 잠재반응변수 등) 및 가능한 제약 방식에 대한 기술이 상세히 이루어지지 않아 연관된 개념을 이해하기에는 무리가 있다. 이외에 척도의 범주가 다섯 개 미만일 때 다집단 분석을 실시한 국내 내용영역 연구의 대다수는 지표변수의 정규성에 대한 철저한 검증이 이루어지지 않았음에도 불구하고 변수를 연속형으로 가정한 채 분석을 진행하였다(Lee, Lee, Lee, & Yeo, 2012; Lee & Yoo, 2015; Choi & Cho, 2013). 이러한 현상은 국내 연구자들이 범주형 지표변수의 특성을 반영하여 다집단 분석을 실시할 때 참고할 수 있는 방법론 연구가 절대적으로 부족했기 때문인 것으로 판단된다. 이미 해외 내용 영역 문헌에서는 범주형 CFA를 활용한 다집단 분석이 실시되고 있기 때문에 국내 연구를 위한 논의가 시급하다.

본 연구는 다집단 범주형 CFA의 측정불변성과 관련하여 지금까지 이루어져 온 전체적인 논의를 종합하여 연구자들이 자신의 연구 방향에 부합하는 판별 방식 및 측정불변성 확인 단계를 결정할 수 있도록 적절한 합의점을 도출하고자 한다. 범주형 지표변수의 본질을 고려하여 자료구조를 더욱 정확히 반영할 수 있는 다집단 분석 방법에 대한 관심이 높아지고 있는 상황에서, 국내의 연구자들이 참고할 수 있는 명확한 방법론 연구는 매우 부족하며, 해외의 논의도 상당히 산재되어 종합적으로 정리되지 못하고 있다. 따라서 본 연구는 방법론적 측면에서 관련 개념에 대해 상세히 기술하고 관련된 쟁점들을 통합한다. 이를 통해 연구자들은 연관된 원리에 대한 더욱 심화된 이해를 도모할 수 있을 것이다. 또한 통계 소프트웨어를 통해 실제 자료의 분석 예시를 보임으로써 다집단 범주형 CFA의 실질적인 활용을 모색한다.

이러한 목적을 달성하기 위해 먼저 범주형 CFA의 기저에 있는 잠재반응변수의 원리 및 범주형 CFA 모형의 판별과 추정 방법에 대해 소개한다. 이후 본격적으로 다집단 확장을 위한 여러 척도화 방식의 장단점 및 측정불변성 확인 단계에 미치게 될 영향을 상술하고 추가 제약 방식을 다분형과 이분형 지표변수 두 가지 측면에서 정리한다. 다음으로 측정불변성 확인 단계에서는 경계동일성을 중심으로 그간 학자들이 제시한 다양한 측정불변성 확인 단계가 어떠한 논리로 파생되었는지 기술한다. 마지막으로 실용적인 측면에서의 활용을 위하여 M*plus*(Muthén & Muthén, 1998-2019)로 실제 자료를 분석한 예시를 보인다.

범주형 확인적 요인분석

다집단 분석으로 논의를 확장하기에 앞서, 연속형 CFA와 달라지는 범주형 CFA의 주요 개념 및 모형의 판별, 그리고 추정에 이르는 과정 전반을 살피고자 한다. 이는 향후 다집단 모형의 판별을 위한 제약 방식을 논할 때 도움을 주기 위함이다.

잠재반응변수

잠재반응변수의 개념은 요인분석 시 범주형 지표변수를 다루기 위한 맥락에서 활용한다. 잠재반응변수는 범주형 지표변수 y_{ij}의 기저에 연속형 잠재 반응변수 y^*_{ij}가 있다고 가정했을 때 y^*_{ij}를 지칭한다(Mckelvey & Zavoina, 1975; Muthén, 1984). 여기서 i는 개인, j는 문항을 의미하며 y_{ij}는 개인 i가 문항 j에서 선택한 관찰값으로서 c라고 가정한다. 즉, c는 피험자가 실제로 선택한 범주를 나타낸다. 이때 범주형 지표변수와 잠재반응변수의 관계는 식 1과 같다.

$$y_{ij} = c,\ \text{if}\ \tau_{jc} \le y^*_{ij} < \tau_{j(c+1)} \qquad (1)$$

만약 범주의 개수가 m개라고 하면 관찰된 값 $c = 0, 1, ..., m-1$로 가정하게 된다. 이때 $\{\tau_{j0}, \tau_{j1}, ..., \tau_{jm}\}$은 문항 j의 경계 모수이며 $\tau_{j0} = -\infty$, $\tau_{jm} = \infty$로 가정하기 때문에 이 두 개의 모수를 제외한 $m-1$개의 경계 모수가 추정된다. 예를 들어, 그림 1과 같이 4개의 범주가 있는 지표변수의 경우 $c = 0, 1, 2, 3$이 되며 문항 j에 해당하는 경계 모수는 $\{\tau_{j0}, \tau_{j1}, \tau_{j2}, \tau_{j3}, \tau_{j4}\}$이다.

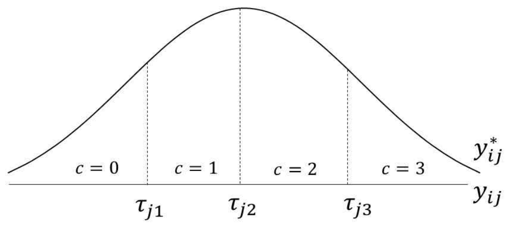

그림 1. 잠재반응변수와 범주형 변수

$\tau_{j0} = -\infty$, $\tau_{j4} = \infty$이므로 이 두 모수를 제외한 3개의 경계 모수가 추정된다. 잠재반응변수 y^*_{ij}는 식 2와 같이 정규분포를 따른다고 가정하는 것이 일반적이다.

$$y^*_{ij} \sim N(\mu_j, \sigma_j) \qquad (2)$$

위에서 μ_j는 y^*_{ij}의 평균, σ_j는 y^*_{ij}의 분산을 가리킨다. 범주형 CFA 모형의 설정은 이 잠재반응변수 y^*_{ij}와 요인을 선형적으로 연결하면서 이루어진다. 잠재반응변수 y^*_{ij}와 요인 η_i의 관계는 식 3과 같이 설정되며, 간명성을 위해 1요인 모형을 가정하였다.

$$y^*_{ij} = \nu_j + \lambda_j \eta_i + \epsilon_{ij} \qquad (3)$$

위에서 ν_j는 잠재절편, λ_j는 요인 η_i에 대한 y^*_{ij}의 요인부하, ϵ_{ij}는 개인이 각 문항에 대해 갖는 개인차, 즉 측정오차이다. 이때 요인 η_i는 식 4와 같이 정규분포를 따른다고 가정한다.

$$\eta_i \sim N(\kappa, \phi) \qquad (4)$$

위에서 κ는 요인의 평균, ϕ는 요인의 분산이다. 또한 오차 ϵ_{ij}도 식 5와 같이 정규분포를 따른다고 가정한다.

따른다고 가정한다.

$$\epsilon_{ij} \sim N(0, \theta_j) \tag{5}$$

위에서 잔차 ϵ_{ij}의 평균은 0으로 가정하며, 분산은 θ_j이다.

모형의 판별

모형의 추정을 위해서는 모형이 판별되어야(identified) 한다. 특히 잠재변수를 다루는 모형의 경우 단위가 존재하지 않는 잠재변수에 단위를 지정해 주는 척도화 과정이 반드시 수반된다. 연속형 확인적 요인분석의 경우에는 잠재변수가 요인 하나뿐이기 때문에 요인에 대한 척도화만 실시하면 된다. 그러나 범주형 확인적 요인분석의 경우에는 요인뿐만 아니라 잠재반응변수도 존재하기 때문에 두 가지 잠재변수에 대한 척도화를 모두 실시해야 한다. 잠재반응변수 y_{ij}^*의 단위를 지정하는 척도화를 진행하면 잠재반응변수의 평균 및 분산이 구체화 되어 y_{ij}^*의 정확한 분포가 결정되므로 모든 경계 모수를 추정하는 것이 가능해진다. 경계 모수가 추정되면 이를 기반으로 하여 지표변수들 사이의 상관이 구해진다. 이렇게 구해진 상관은 다분형 지표변수인 경우 다분상관, 이분형 지표변수인 경우 사분상관이 된다. 이에 더해 요인에도 단위를 지정하면 앞서 구한 지표변수들 사이의 상관을 이용하여 요인부하 및 요인분산과 같은 주요 모수들을 추정할 수 있게 된다.

잠재반응변수의 척도화

y_{ij}^*에 단위를 부여하는 방식은 Delta 방식과 Theta 방식 두 가지가 존재한다. Delta와 Theta는 학문적인 용어라기보다는 M*plus*에서 사용하는 명칭인데, 실제로 많은 학자들이 이 용어를 사용한다. Kamata와 Bauer(2008)는 Delta와 Theta를 각각 marginal과 conditional로 명명하기도 하였다. Delta 방식은 y_{ij}^*의 분산 σ_j를 전부 동일한 상수(보통, 1)로 고정한다. 잠재반응변수 자체에 단위를 부여해서 y_{ij}^*의 분포가 모든 지표변수에 대해 동일해지고, 이 분포를 표현할 때 요인 η_i와 상관없이 분포가 형성된다는 의미에서 marginal 방식으로 명명되었다. Theta 방식은 y_{ij}^*의 잔차분산 θ_j를 전부 동일한 상수(보통, 1)로 고정한다. 요인 η_i가 잠재반응변수 y_{ij}^*를 설명하고 있다는 조건하에 잔차분산을 통해 단위를 지정한다는 측면에서 conditional 방식이라 명명되었다. 단일집단 분석 시 잠재반응변수의 척도화로는 Delta 방식이 권장된다. 이는 Delta 방식이 Theta 방식에 비해 요인부하의 해석에 있어 가지는 장점이 상당히 명확하고, 실제로 Theta 방식이 Delta 방식을 보완하기 위한 측면에서 고안되었기 때문이다(Muthén, 2011). 단순 요인분석에서 더 나아가 확장된 분석을 실시하고자 할 때, Delta 방식을 적용하면 불가능한 분석들을 위해 나온 척도화 방법이 Theta 방식인 것이다.

요인의 척도화

또 다른 잠재변수인 요인의 단위를 지정하는 방식은 잠재반응변수의 척도화 방식보다 더욱 다양하다. 이 중 요인표준화 방식과 척도상수 방식이 대표적이며, 이외의 방식들은 두 방식을 혼용 또는 변형한 것(Reise, Widaman, & Pugh, 1993; Yoon & Millsap, 2007)이기 때문에 두 가지 방식을 중심으로 논의를

진행한다.

요인표준화 방식은 요인평균 κ와, 요인분산 ϕ를 모두 상수로 고정한다. 일반적으로 요인평균을 0, 분산을 1로 지정하여 요인의 단위를 표준화하기 때문에 표준화 방식이라 명명되었다. 이렇게 요인을 표준화하면 모든 요인부하를 추정하는 것이 가능해진다(Kim & Yoon, 2011). 또한 요인표준화 방식을 Delta 방식과 함께 사용하면 경계 모수들이 표준정규분포 상에 위치하게 되므로 추정치에 대한 해석이 상당히 용이하다. 경계 모수나 요인부하 같은 문항의 특성을 파악하기 위해서는 모든 요인부하 추정이 가능하고 추정치에 대한 해석이 용이한 요인표준화 방식을 선택하는 것이 좋다.

척도상수 방식은 요인마다 한 문항의 요인부하 λ_j와 경계 모수 τ_{jc} 하나씩을 상수로 고정한다. 보통 요인부하는 1, 경계 모수는 0으로 지정하는데, 이 방식을 사용하면 요인의 평균 및 분산 추정이 가능해진다(Kim & Yoon, 2011). 구조모형을 통한 요인들 간 관계 파악이 목적인 경우 요인의 평균과 분산을 추정하는 것이 가능한 척도상수 방식이 선호된다. 연속형 CFA의 경우 척도상수 방식을 사용하면 지정된 지표변수와 요인은 동일한 단위를 가지게 된다. 그러나 범주형 CFA의 경우 요인의 단위는 잠재반응변수의 척도화에 의해 영향받기 때문에 척도상수 방식의 사용이 지표변수의 단위와 요인의 단위를 동일하게 맞추어 주는 것은 아니라는 점에 유의해야 한다(Kamata & Bauer, 2008).

모형의 추정

범주형 지표변수를 사용하는 모형의 제한정보추정을 위해서는 최대우도(maximum likelihood, ML)와 가중최소제곱(weighted least squares, WLS)을 모두 이용할 수 있는데, 이 중 선호되는 것은 WLS 추정법이다(Koğar & Koğar, 2015). 하지만 전통적인 WLS 추정법은 추정에 소요되는 긴 시간(Katsikatsou, Moustaki, Yang-Wallentin, & Jöreskong, 2012), 안정적인 추정을 위해 요구되는 상당히 큰 표본크기(Muthén & Kaplan, 1992), χ^2검정통계량의 과대 추정(Dolan, 1994; Hutchinson & Olmos, 1998), 표준오차의 과소추정(Potthast, 1993) 등의 문제를 가지고 있다. 이후 WLS 추정법의 복잡한 가중치 행렬 문제를 보완하기 위해 대각 가중 최소 제곱(diagonally weighted least sauqres, DWLS) 방법이 고안되었다(Edwards, Wirth, Houts, & Xi, 2012). DWLS 추정법은 평균조정 가중최소제곱(mean adjusted weighted least squares, WLSM)으로도 불리며, 이는 또 다른 강건한 방식인 평균 및 분산조정 가중최소제곱(WLSMV, mean and variance adjusted weighted least squares; Muthén, du Toit, & Spisic, 1997)으로 확장되었다(Kim, 2016).

WLSMV 추정법은 추정된 표준오차의 편향이 거의 없고(Yang-Wallentin, Jöreskog, & Luo, 2010), 작은 크기의 표본에서도 잘 작동하는 것으로 알려져 있다(Maydeu-Olivares, 2001; Nussbeck, Eid, & Lischetzke, 2006). 특히 범주의 수가 2~3개로 적은 경우 연속형 변수를 가정하는 ML 추정법이 갖는 χ^2검정통계량의 과도한 기각, 요인부하량의 과소추정 문제 등을 보완할 수 있다(Beauducel & Herzberg, 2006)는 강점이 있다. 다만 이 추정 방식은 범주의 수가 적을수록 충분한 표본크기(예를 들어, 500 이상)를 필요로 한다는 점에 유의해야 하며(Yang & Green, 2010), 내재된 모형의 비교를

위해 M*plus*의 DIFFTEST 옵션과 함께 사용할 경우 제1종 오류가 ML이나 MLR[1]에 비해 상당히 크게 나타난다는 점이 보고된 바 있다 (Sass, Schmitt, & Marsh, 2014).

범주형 확인적 요인분석의 다집단 확장

범주형 CFA를 이용하여 다집단 분석을 실시하고자 하는 경우 관찰변수와 잠재반응변수 사이의 관계는 식 6과 같다.

$$y_{ijg} = c,\ \text{if}\ \tau_{jgc} \leq y^{*}_{ijg} < \tau_{jg(c+1)} \tag{6}$$

식 1과 달라진 점은 개인이 속한 집단을 나타내는 g가 추가되었다는 것이다. 이때 잠재반응변수와 요인 사이의 관계는 식 7과 같다.

$$y^{*}_{ijg} = \nu_{jg} + \lambda_{jg}\eta_{ig} + \epsilon_{ijg} \tag{7}$$

식 6과 마찬가지로 개인이 속한 집단을 나타내는 g가 추가되었다는 점이 기존의 범주형 CFA 모형의 식과 달라진 점이다. g가 추가되었다는 것은 추정되는 모수들이 집단에 따라 달라질 수 있음을 의미한다. 다시 말해, 집단의 개수만큼 CFA 모형이 존재하는 것이다.

필수제약

단일집단 분석의 경우 단위가 없는 두 잠재변수(잠재반응변수 및 요인)에 대한 척도화인 필수제약을 가하면 모형이 판별되며, 이때 선택할 수 있는 척도화 방식은 연구자의 관점 및 연구의 목적에 따라 달라진다. 다집단 분석의 경우에도 마찬가지로 다양한 방식을 자유롭게 선택할 수 있다. 하지만 판별을 위해 선택한 제약 방식은 추후 측정불변성 확인 단계에 영향을 미치기 때문에 각 척도화 방식의 특성 및 장단점을 고려한 제약을 실시해야 한다. 따라서 모형의 판별을 위한 필수제약에 대한 논의를 먼저 진행한다. 아래 논의에서는 집단이 두 개인 경우를 가정하며, 하나의 집단은 1, 나머지 집단은 0으로 설정한다.

잠재반응변수의 척도화

집단 1의 잠재반응변수 분산은 식 8과 같이 표현된다.

$$Var(y^{*}_{ij1}) = \sigma_{j1} = \lambda^{2}_{j1}\phi_{1} + \theta_{j1} \tag{8}$$

위 식에서 σ_{j1}은 집단 1의 잠재반응변수의 분산, λ_{j1}은 문항 j에 대한 집단 1의 요인부하, ϕ_1은 집단 1의 요인 분산, θ_{j1}은 문항 j에 대한 집단 1의 잠재반응변수 잔차분산이다. 척도화를 통해 y^{*}_{ij1}의 단위를 지정하면 집단 1의 경계 모수들이 추정되고, 경계 모수들을 통해 요인부하 λ_{j1}과 요인분산 ϕ_1이 추정된다.

Delta 방식은 집단 1의 잠재반응변수 y^{*}_{ij1}의 분산을 모두 동일한 상수로 고정하며 그 상수는 일반적으로 1이다. Delta 방식을 사용하여 잠재반응변수의 척도화를 실시하면 잠재반응변수의 분산은 식 9와 같아진다.

---
1) M*plus*의 추정 옵션명으로서 결측치가 있을 때 사용할 수 있는 강건한 최대우도 추정법이다. 다변량 정규성이 확보되지 않았을 때 χ^2검정통계량이 χ^2분포를 잘 따르지 않으므로, 이를 χ^2분포에 더 잘 근사하도록 만든 것이다.

$$Var(y^*_{ij1}) = 1 = \lambda^2_{j1}\phi_1 + \theta_{j1} \quad (9)$$

위에서 잔차분산 θ_{j1}은 $1-\lambda^2_1\phi_1$으로서 추정된 요인부하 λ_{j1}과 요인분산 ϕ_1을 대입하면 계산되어 나오는 값이다. 이는 θ_{j1}이 추정되는 모수가 아닌 얻어지는 값(remainder)이라는 것을 의미한다. 즉, θ_{j1}이 λ_{j1}과 ϕ_1에 의해 결정되므로 θ_{j1}은 이 두 모수로부터 분리될 수 없다. 따라서 Delta 방식을 이용하면 측정불변성 단계 중 분산(공분산)동일성(strict invariance)을 확인할 수 없다.

Theta 방식은 집단 1의 잠재반응변수 y^*_{ij1}의 잔차분산 θ_{j1}을 모두 동일한 상수로 고정하며 그 상수는 일반적으로 1이다. Theta 방식을 사용하여 잠재반응변수의 척도화를 실시하면 잠재반응변수의 분산은 식 10과 같아진다.

$$Var(y^*_{ij1}) = \sigma_{j1} = \lambda^2_{j1}\phi_1 + 1 \quad (10)$$

위의 식에 의하면 집단 1의 잠재반응변수 분산 σ_{j1}이 $\lambda^2_{j1}\phi_1+1$로서 요인부하 λ_{j1}과 요인분산 ϕ_1에 의해 결정된다. Theta 방식 사용 시 추정 없이 얻어지는 값은 σ_{j1}으로, 다른 집단의 잔차분산 θ_{j0}는 추정 가능하기 때문에 분산(공분산)동일성을 확인할 수 있다.

잠재반응변수의 척도화는 측정불변성을 어떻게 정의할 것인가에 따라 취사선택할 수 있다. Theta 방식을 사용하면 분산(공분산)동일성을 확인할 수 있다는 장점이 있으나, 그 외의 측면에서는 Delta 방식이 Theta 방식에 비해 가지는 장점이 상당히 명확하다. Delta 방식의 가장 큰 장점은 직관적인 이해가 쉽다는 점이다. Delta 방식은 잠재반응변수의 분산을 보통 1로 고정해 놓기 때문에 추정된 계수가 자연스럽게 표준화된 단위를 지니게 된다. 그에 반해, Theta 방식을 사용하면 잠재반응변수의 분산이 지표변수(문항)마다 달라지기 때문에 추정된 계수들을 표준화시키지 않는 이상 이들을 문항 및 집단 간에 비교하고 해석하는데 어려움이 생기게 된다. 실제로 Muthén(2011)은 Theta 방식이 비표준화 계수의 해석을 애매하게 만드는 것이 사실이라고 언급하는데, 특히 비표준화 계수들의 비교가 필수적인 다집단 분석의 경우에는 Delta 방식을 이용한 척도화가 더 편리하다. 또한 Theta 방식을 사용하면 요인이 갖는 설명력에 따라 잠재반응변수의 분산이 달라지는 문제가 발생한다. 식 10을 보면 잠재반응변수 분산 σ_{j1}은 $\lambda^2_{j1}\phi_1+1$로서 요인이 갖는 설명력인 요인부하 λ_{j1}의 크기에 의해 영향을 받는다. Theta 방식을 사용하면 일반적인 선형모형처럼 요인의 설명력이 증가함에 따라 잔차분산이 줄어드는 형태를 띠지 않기 때문에 전체 분산이 오히려 늘어나게 되며 그 크기도 지표변수 각각의 요인부하량에 따라 달라지는 것이다. 이로 인해 선형모형이 더 익숙한 연구자들에게 있어 Theta 방식을 사용하는 것이 추정된 계수에 대한 해석을 어렵게 만들 수 있다는 우려가 제기된다(Kamata & Bauer, 2008).

그간 범주형 확인적 요인분석 모형의 측정불변성 검정을 위해서 상당히 많은 연구자가 자연스럽게 Theta 방식을 채택해 왔다(Bass et al., 2011; Hempel & Bartkowski, 2008; Mathysseck et al., 2013; Moor Distel, Trull, & Boomsma, 2009; Xing & Hall, 2015; Zimprich, Allemand & Dellenbach, 2009). 이는 Theta 방식이 분산(공분산)동일성 검정을 가능하게 해주

기 때문이었다. 하지만 다집단 분석의 맥락에서 Theta 방식은 분산(공분산)동일성 확인을 위한 목적으로 제시되었을 뿐 그 이외의 이점은 존재하지 않는다(Muthén, 2011). 그런데 그간 연속형 요인분석의 경험에 비추어 보면 완벽한 분산(공분산)동일성을 확보하는 것은 상당히 어려운 일이며(Widaman & Reise, 1997), 사실상 측정불변성 검정의 주된 관심사가 아니기도 하다(Selig, Card, & Little, 2008). 그리고 이는 연속형 분석의 개념적, 수학적 확장이랄 수 있는 범주형 분석에서도 마찬가지일 것으로 판단된다. 즉, 측정불변성 확인 단계에서 분산(공분산)동일성이 반드시 확보되지 않더라도 집단 간 요인 구조가 동일하다고 볼 수 있다는 공감대가 형성되어 있는 것이다. 이러한 측면 및 Delta 방식이 지닌 이점을 고려하여 범주형 CFA의 다집단 확장을 위해서도 Delta 방식의 사용이 증가하는 추세이다(Christoph, Goldhammer, Zylka, & Harting, 2015; Forbush et al., 2013; Schlotz, Yim, Zoccola, & Schulz, 2011).

요인의 척도화

요인의 척도화는 잠재반응변수의 척도화에 비해 상대적으로 명확한 결론이 내려져 있다. 범주형 CFA의 다집단 확장을 위해서는 척도상수 방식을 사용해야 한다. 척도상수 방식을 이용하면 모든 집단에서 각 요인마다 임의의 요인부하 λ_{jg}를 전부 동일한 상수로 고정한다. 그 상수는 일반적으로 1이며 이는 식 11과 같다.

$$\lambda_{jg} = 1 \ \ \forall\, s, g \tag{11}$$

위에서 s는 각 요인마다 임의로 선택되어 상수로 고정되는 변수인 참조 문항(또는 참조 지표변수)을 지칭한다.

범주형 CFA 모형의 다집단 확장을 위해 척도상수 방식을 선택해야 하는 이유는 요인표준화 방식을 사용했을 때 두 집단의 요인분산이 동일하게 1로 고정되는 문제가 발생하기 때문이다. 서로 다른 집단의 비교는 집단의 특성이 각기 다르다는 것을 가정하고 그 양상을 살피는 것을 목적으로 한다. 집단의 특성이 서로 다를 것이라 가정하기 때문에 자연스럽게 집단 간 평균이나 분산도 다를 것이라 여겨진다. 따라서 요인부하도 각 집단에 따라 서로 다른 분산의 영향을 받는다고 보는 것이 마땅하다. 그래서 다집단 분석 상황에서 요인부하의 비교는 비표준화계수에 대해 이루어진다. 그런데 요인의 분산이 동일하게 고정되면 결과적으로 요인부하가 동일한 분산의 영향을 받게 된다. 집단 비교의 기본적인 가정이 위배되는 결과가 발생하게 되는 것이다. 따라서 다집단 분석을 목적으로 하고 있는 경우, 요인표준화 방식보다는 척도상수 방식을 선택하는 것이 타당하다. 이에 더해 집단 간 요인평균을 비교하고자 하는 목적이 있는 경우에도 요인평균의 추정이 가능한 척도상수 방식을 선택하는 것이 더 적절하다.

추가제약

범주형 CFA의 다집단 확장 시 필수제약만 가지고는 모형의 판별이 불가능하다. 문항이 p개, 경계 모수의 총 개수가 t개, 집단의 수가 2개일 때, 범주형 CFA의 다집단 확장 시 추정 가능한 최대 모수의 개수는 $2\left[\frac{p(p-1)}{2}+t\right]$이다. 판별 가능한 모수가 단일 집단에 비해 두 배로 늘어나긴 하지만, 집단의 수가 늘어난 만큼 추정해야 하는 모수의

개수도 더 많이 늘어나기 때문에 추가제약이 필수적이다.

그간 문헌에서 추가제약 방식에 대한 논의가 충분히 진행되지는 않았으나, 지금까지의 제한된 논의를 바탕으로 판별을 위한 제약을 다음과 같이 진행할 수 있다. 지표변수가 다분형인 경우 먼저 필수제약에 해당하는 잠재반응변수 및 요인에 대한 척도화를 실시한다. 잠재반응변수에 대한 척도화는 두 집단 중 하나의 집단에 진행된다. 한 집단의 잠재반응변수 평균을 0으로 고정한 뒤, Delta 방식의 경우 잠재반응변수의 분산을 1로, Theta 방식의 경우 잠재반응변수의 잔차분산을 1로 놓는다. 이렇게 하면 한 집단의 경계 모수가 추정된다. 요인에 대한 척도화는 일단 제약을 가한 집단의 요인평균을 0으로 고정한다. 이렇게 되면 나머지 집단의 요인평균을 추정할 수 있기 때문에 집단 간 요인의 평균 비교가 가능하다. 추가적으로 척도상수 방식을 사용하여 두 집단 모두에서 각 요인마다 임의로 선택한 하나씩의 요인부하를 1로 고정하면 필수제약이 완료된다. 이렇게 되면 한 집단의 모든 모수들이 판별된다. 이제 남은 집단의 모수들까지 판별 가능한 상태로 만들기 위한 추가제약을 실시한다. Millsap과 Yun-Tein(2004)의 추가제약 방식은 임의의 경계 모수를 집단 간 동일하게 놓는다. 모든 문항에서 임의로 선택한 하나의 경계 모수가 집단 간 동일하다는 제약이 가해지는 것이다. 또한 앞서 척도상수 방식에서 요인부하를 1로 지정한 문항들에 대해서는 또 다른 경계 모수를 하나 더 선택하여 집단 간 동일하다는 제약을 가해야 한다. 이러한 추가제약을 거치면 총 p(문항 수)$+r$(요인 수)개의 경계 모수를 제약하게 된다. 추가제약 시 다른 모수가 아닌 경계 모수를 제약한 이유는 다분형 지표변수의 경계 모수는 다른 모수들보다 상대적으로 그 수가 많고 특히 나머지 집단에서 r개의 요인평균을 추정하기 위해서이다.

다분형과는 다르게 지표변수가 이분형인 경우 경계 모수가 문항 당 하나밖에 없기 때문에 앞서 기술한 추가제약을 진행할 수 없다. Millsap과 Yun-Tein(2004)의 경우 모든 문항의 경계 모수를 동일하게 제약한 뒤 요인의 개수인 r개만큼 잠재반응변수의 분산이나 잔차분산을 집단 간 동일하게 제약한다. 하지만 이러한 추가제약을 가하면 경계 모수에 대한 동일성 검정을 실시할 수 없게 된다. 이를 보완하기 위해 Millsap과 Kim(2018)은 필수제약 이후 나머지 집단에서 모든 잠재반응변수의 분산이나 잔차분산을 1로 고정할 것을 제안한다. 잠재반응변수의 분산을 고정하는 것은 앞선 필수제약에서 Delta 방식을, 잠재반응변수의 잔차분산을 고정하는 것은 앞선 필수제약에서 Theta 방식을 사용한 경우에 해당한다. Millsap과 Kim(2018)의 방식을 이용한 추가제약을 실시하면 Theta 방식의 척도화를 진행했다고 할지라도 분산(공분산)동일성 검정을 실시할 수 없다. 추가제약으로 인하여 결국 모든 집단의 잔차분산이 동일하다는 제약이 가해지기 때문이다. 이후 앞서 요인부하를 1로 지정한 문항들의 경계 모수까지 집단 간 동일하다는 제약을 가하면 이분형 지표변수를 지니는 CFA 모형의 다집단 확장을 위한 판별이 완료된다.

측정불변성 확인

범주형 CFA의 측정불변성 확인 단계는 연속형 CFA와 마찬가지로 연구의 목적에 따라

다양할 수 있다(Widaman, 2000; Widaman & Reise, 1997). 연속형의 경우 측정불변성 확인 단계에 대해 많은 논의가 진행된 반면, 상대적으로 범주형 CFA의 측정불변성 확인 단계에 대한 논의는 많지 않다. 이는 범주형 분석을 연속형 분석의 단순 확장 형태로 보고 범주형 CFA의 측정불변성 확인 단계를 연속형의 경우와 유사하게 진행했기 때문인 것으로 판단된다. 그러나 연속형 CFA의 절편동일성은 생략하더라도 상관없으며, 잠재평균 비교 등의 목적이 있는 경우에만 확인하면 된다는 흐름이 존재(Bollen, 1989; Cooke, Kossen & Michie, 2001; Marsh, 2007; Kline, 2016)하는 반면, 이에 상응하는 범주형 CFA의 경계동일성을 확인하는 것은 생략할 수 없다고 여겨지는 것이 일반적이다(Glanville & Wildhagen, 2007; Looij-Jansen, Goedhart, Wilde, & Treffers, 2011; Meuleman & Billiet, 2012). 이는 경계 모수가 다분상관 혹은 사분상관을 구하는 데 핵심이 되기 때문이다. 다시 말해, 경계 모수가 요인부하 및 요인 분산 등 모형의 해석과 관련된 주요 모수를 추정하는 데 영향을 미치기 때문인 것이다. 따라서 범주형 CFA의 측정불변성 확인 단계는 경계 모수 동등성 확보를 중심으로 논의를 진행한다.

χ^2 차이검정

χ^2 차이검정의 보정

측정불변성 확인은 동일성 제약이 가해지지 않은 복잡한 모형과 동일성 제약이 가해진 단순한 모형의 χ^2값 차이를 통계적으로 검정함으로써 이루어진다. 이러한 검정은 두 내재된 모형의 χ^2차이값이 χ^2분포를 따르기 때문에 가능한 것이다. 그런데 범주형 자료를 위한 추정법을 사용하면 χ^2차이값이 χ^2분포를 따르지 않는다는 문제점이 발생한다. 만약 범주형 모형의 추정을 위해 MLM(mean adjusted maximum likelihood)[2)]이나 WLSM처럼 평균 조정만 이루어진 추정법을 사용하는 경우 각 모형의 결과에 제시된 χ^2값에 Satorra와 Bentler (2001)의 보정계수를 대입하여 보정된 값을 직접 계산하면 된다. 그러나 WLSMV, MLMV (mean and variance adjusted maximum likelihood)[3)], ULSMV(mean and variance adjusted unweighted least squares)[4)]처럼 평균에 더해 분산에 대한 조정까지 들어간 추정법을 사용하는 경우 분산(공분산) 행렬에 대한 보정까지 필요하다. 행렬 수준에서 이루어지는 계산은 상당히 복잡하기 때문에 M*plus*는 해당 추정법들을 사용했을 시 χ^2차이검정을 위한 DIFFTEST 옵션을 제공한다. DIFFTEST 옵션의 사용법은 이후 실제 데이터 분석 시 소개된다.

DIFFTEST 옵션 사용 시 χ^2차이검정은 비교하고자 하는 단순한 모형과 복잡한 모형의 합치함수 최소값들의 차이 T_{dif}에 대한 보정을 실시한 후 T_{dif}의 점근적 분포가 특정한 χ^2분포를 따른다는 가정 하에 실시된다(Asparouhov & Muthén, 2006). DIFFTEST 결과가 통계적으로 유의하지 않으면 보정된 차이값이 통계적으로 0과 다르지 않다는 것이며, 이는 곧 두 모형의 차이가 없다는 것을 의미한다.

2) 결측치가 없는 완전자료를 가정한 강건한 최대우도 추정법이다.

3) 평균 보정만 실시한 MLM 추정법에 분산에 대한 보정까지 실시한 추정법이다.

4) 비가중 최소제곱 추정(ULS)에 평균 및 분산 조정을 실시한 추정법으로서 변수의 분포에 대한 가정이 존재하지 않아 강건한 것으로 알려져 있다.

따라서 모형의 간명성 측면에서 제약이 더 가해진 단순한 모형을 선택하여 집단 간 측정불변성이 확보되었다는 결론을 내리게 된다. DIFFTEST 결과에 나타나는 검정통계량 차이가 비교하고자 하는 모형들의 χ^2검정통계량의 차이와 다른 이유는 T_{dif}를 가지고 χ^2값을 재구성하기 때문이다. χ^2값의 재구성은 Satterthwaite(1941)의 방식을 따라 이루어진다. 이 분포의 자유도 역시 동일한 방식을 따라 추정되기 때문에 일차적으로는 비교하고자 하는 두 모형의 자유도 차이와는 다른 값이다. 그런데 DIFFTEST 결과에 나타나는 자유도는 내재된 두 모형의 자유도 차이와 일치한다(Asparouhov & Muthén, 2010). 이는 비교하고자 하는 모형들의 자유도 차이에 대한 이차 보정을 실시한 것으로 연구자들의 혼란을 최소화하기 위해 M*plus* 6 이후에 도입된 것이다.

순차적 검정

순차적 검정(sequential test)은 각 모수의 종류별로 동일성 검정을 분리하여 실시한다. 순차적 검정 방식에는 순행절차와 역행절차가 있다. 형태동일성에 대한 검정 이후 동등성 제약을 순차적으로 가해주는 방식은 순행절차(자유모수기반 전략)이며 형태동일성 검정 이후 전체 모수에 대한 동등성 제약을 가해 준 상태에서 제약을 순차적으로 풀어주는 방식은 역행절차(제약모수기반 전략)이다. 순행절차와 역행절차를 비교했을 때에는 순행절차를 사용할 것이 권장되며(Lopez Rivas, Stark, & Chernyshenko, 2009; Wang & Yeh, 2003) 실제로도 순행절차를 더 많이 사용하는 경향이 있다(Dimitrov, 2010). 이러한 경향성의 원인은 개별 문항의 차별문항기능 여부를 밝혀내고자 하는 문항반응이론의 시각에서 찾아볼 수 있다. 역행절차는 일단 관심 문항을 제외한 다른 문항 모수들이 집단 간 동일하다는 가정 하에 관심 문항 모수의 측정불변성을 확인하지만, 순행절차는 관심 문항뿐만 아니라 이외의 문항 모수들의 측정불변성을 미리 가정하지 않는다. 따라서 순행절차를 사용했을 때 동등하지 않은 모수에 대한 동등성 제약으로 인한 추정의 오류를 방지할 수 있는 것이다.

순차적인 방식은 방법론 문헌(Maydeu-Olivares & Cai, 2006; Stark, Chernyshenko, Drasgow, & Williams, 2006) 뿐만 아니라 내용영역의 문헌(Bathinic, Wolff, & Huapt, 2007; Diniz, Pocinho, & Almeida, 2011; Lawrence, Rosenberg, Rimmer, Thombs, & Faurbach, 2010; Morin, Train, & Caci, 2016)에서도 빈번하게 적용되는 방식이다. 일반적으로 적용되는 절차는 형태동일성, 요인부하동일성, 경계동일성, 분산(공분산)동일성 순서대로 제약을 가해주는 방식이다. 이는 앞서 언급했듯이 범주형 분석을 연속형 분석의 단순 확장된 형태로 보고 연속형 분석에서 주로 선택하는 절차를 범주형 CFA에 그대로 적용한 결과로 볼 수 있다. Wu와 Estabrook(2016)은 순차적 검정에서 경계모수의 동등성 확보가 요인부하의 동일성에 선행하도록 할 것을 권장하기도 하였다. 이는 범주형 지표변수를 다루는 모형에서 경계 모수가 가지는 중요성을 고려한 것으로, 경계모수는 요인부하보다 먼저 추정되며 요인부하 추정을 위한 상관(사분상관, 다분상관)을 구하는 데 이용되기 때문이다.

동시 검정

동시 검정(omnibus test)은 여러 종류의 모수를 한 번에 검정하는 방식이다. 먼저 형태동일성을 확인한 뒤, 요인부하동일성과 경계동

일성 단계를 통합하여 두 가지 종류의 모수에 대한 검정을 동시에 실시한다(Glockner-Rist & Hoijtink, 2003; Sass & Schmitt, 2013; Temme, 2006). 이때 만약 동시 검정에서 동등성을 확보하지 못했을 경우 어떠한 모수에 대한 부분 불변성을 먼저 허용할 것인지에 대해 학자들마다 견해 차이가 존재한다.

경계동일성에 대한 제약을 먼저 풀어주어야 한다는 입장의 경우 요인부하의 불변성 확보를 우선으로 한다(Glockner-Rist & Hoijtink, 2003; Temme, 2006). 따라서 동시 검정이 기각되었을 때 경계 모수에 대한 동일성 제약을 먼저 풀어줌으로써 부분 경계동일성을 먼저 허용한다. 이러한 방식은 경계 모수의 차이 여부를 모형에 보다 정확히 반영함으로써 요인부하의 집단 간 차이에 대한 판단 오류를 최소화하고자 하는 것으로(Temme, 2006), 요인과 직접적인 관련성을 갖는 요인부하를 측정불변성을 위한 주요 모수로 취급한다. 반면 요인부하동일성에 대한 제약을 먼저 풀어주어야 한다는 입장의 경우 경계 모수의 불변성 확보를 우선으로 한다(Sass & Schmitt, 2013). 요인부하 및 경계동일성에 대한 동시 검정이 기각되었을 때 요인부하에 대한 동일성 제약을 먼저 풀어줌으로써 부분 요인부하동일성을 먼저 허용한다. 이는 요인과 관련한 주요 모수를 추정하는데 정보로 기능하는 경계 모수를 측정불변성을 위한 주요 모수로 취급한 것이다.

근사적합도 지수 차이

χ^2차이검정의 경우 χ^2 모형적합도 검정과 마찬가지로 표본의 크기가 커지면 영가설을 과도하게 기각하는 경향이 있다. 따라서 χ^2차이검정으로 측정불변성을 확인하면 실제 동일성이 존재함에도 불구하고, 그렇지 않은 것으로 결론을 내리게 되기도 한다. 그러므로 표본의 크기가 매우 커서 통계적인 검정을 통해 측정불변성을 확보하는 것이 적절치 않은 경우 실용적인 측면에서 근사적합도 지수의 차이를 살펴 측정불변성을 확인하기도 한다(Chen, 2007; Cheung & Rensvold, 2002; Meade, Johnson, & Braddy, 2008; Rutkowski & Svetina, 2014). 여기서는 동일성 제약을 가한 단순한 모형과 동일성 제약을 풀어준 복잡한 모형의 근사적합도 지수 차이가 특정 기준을 넘지 않으면 측정불변성이 확보되었다고 결론 내린다. 일반적인 연속형 CFA의 경우에 근사적합도 지수 차이에 대한 기준은 학자마다 다양하게 제시되어있지만 그 기준이 주관적이고 임의적이다. 관대한 기준을 중심으로 보면 ΔCFI의 경우 0.02(Rutkowski & Svetina, 2014)에서 0.01(Chen, 2007; Cheung & Rensvold, 2002; French & Finch, 2006)까지, ΔRMSEA의 경우 0.03(Rutkowski & Svetina, 2014)에서 0.015(Chen, 2007)까지로 제시되어 있다. 더 엄격한 기준이 제시되어야 한다는 의견도 있지만(Byrne, 2012; Meade et al., 2008), 현실적으로 측정불변성을 확보해야 하는 점을 고려하여 내용영역 연구의 측정불변성 확인 시 앞의 기준을 사용하는 것에 무리가 없다는 공감대가 형성되어 있으며, 이미 상당한 문헌에서 근사적합도 지수의 차이를 이용한 측정불변성 확인이 이루어지고 있다.

범주형 CFA의 경우에는 근사적합도 지수 차이와 관련한 연구가 드물다(Rutkowski & Svetina, 2017; Sass, Schmitt, & Marsh, 2014). 시뮬레이션을 통해 다양한 수준의 불변성을 설정한 범주형 확인적 요인분석 모형을 살펴본

결과, ΔCFI 및 ΔRMSEA의 경우 완벽한 측정 불변성이 확보된 모형에서는 Chen(2007)의 기준인 0.01과 0.015가 잘 작동하였는데(Sass et al., 2014), 불변성이 깨진 경우에는 그 차이를 감지하지 못하는 것으로 나타났다. 그럼에도 불구하고 χ^2차이검정의 과도한 기각 문제를 보완한다는 맥락에서 범주형의 경우에도 연속형의 경우처럼 CFI는 0.01, RMSEA는 0.015라는 기준을 사용하는 것이 크게 문제되지는 않을 것으로 판단된다. 이때 주의해야 할 점은 범주형 지표변수가 포함된 모형의 적합도지수인 WRMR의 차이를 이용하는 것은 WRMR의 표본크기 및 모형의 복잡성에 대한 민감성 때문에 권장되지 않는다는 것이다(Sass & Schmit, 2013).

실제자료의 적용

앞서 논의한 범주형 CFA의 측정불변성을 검정하는 방법을 실제 자료에 적용시켜 보고자 한다. 먼저 사용된 자료의 특성을 간략히 소개하고, 다집단 범주형 CFA를 실시한다.

분석 자료 및 분석 방법

분석에 사용된 자료는 한국복지패널의 13차년도 데이터이다. 한국복지패널 자료는 다양한 인구집단 별 생활실태와 복지욕구 등을 파악하기 위한 목적으로 2006년부터 연 단위로 총 열세 차례에 걸쳐 수집되었으며 설문 문항에 우울을 측정하는 축약형 CES-D(Center for Epidemiology Studies Depression) 척도를 포함하고 있다. 축약형 CES-D 척도는 11개 문항을 4점 리커트 척도(1=극히 드물다, 2=가끔 있었다, 3=종종 있었다, 4=대부분 그랬다)로 평정하며, 4개의 하위구인이 최종적으로 응답자의 우울을 측정하게 되는 구조이다. 본 연구에서는 2018년 전국에 거주하는 6,474가구의 가구원 12,469명을 대상으로 수집된 CES-D 척도 자료를 이용하였으며, 결측치는 일괄 삭제(listwise deletion)하여 최종 표본크기는 11,704였다. 축약형 CES-D 척도의 측정불변성은 남녀 집단에 대해 검정되었으며, 모든 분석은 M*plus* 7(Muthén & Muthén, 1998-2019)을 이용하였다. 분석에 사용된 M*plus* 명령어는 부록에 제공된다. 부록에 제공된 코드는 제약이 최소한으로 가해진 형태동일성 모형에 대한 것이다.

측정불변성 확인

자료에 대해 범주형 CFA를 이용하여 남녀 집단 간 측정불변성 확인 검정을 실시한다. 범주형 확인적 요인분석은 WLSMV를 통해 추정되었으며, 잠재반응변수의 척도화를 위해서는 Delta 방식이, 요인의 척도화를 위해서는 척도상수 방식이 선택되었다. Delta 방식을 통해 남자 집단의 잠재반응변수 분산이 전부 1로 고정되었으며, 척도상수 방식을 통해 남녀 집단 모두의 4개 요인 첫 번째 문항의 요인부하가 1로 고정되었다.

모형의 판별을 위한 추가제약은 Millsap과 Yun-Tein(2004)의 방식을 따랐다. M*plus*에서 Delta 방식을 지정하면 자동적으로 잠재반응변수의 평균이 0, 분산이 1로 고정된다. 이외의 제약은 연구자가 직접 지정하면 된다. 축약형 CES-D 척도를 위한 집단 간 경계 모수 동일 제약은 모든 문항의 세 번째 경계 모수와 앞서 요인부하를 1로 고정한 각 요인의 첫 번째 문항의 두 번째 경계 모수에 대해 실시하였다.

제약할 경계모수의 선택은 각 범주를 선택한 피험자의 누적 비율을 고려한 것으로, 누적 비율이 집단끼리 전반적으로 가장 비슷한 모수를 기준으로 하였다. 각 범주를 선택한 피험자의 누적 비율표는 부록에 제공된다. 이러한 선택은 측정불변성을 확보하는 데 있어 추가제약으로 인해 통계적 검정 없이 강제로 동등성 가정을 부여해야 하는 모수들에 대한 타당성 확보에 도움을 주기 위해서였다.

범주형 CFA의 측정불변성 확인을 위한 χ^2 차이검정은 M*plus*의 DIFFTEST 옵션을 사용했다. DIFFTEST 옵션을 이용하면 WLSMV나 MLMV, ULSMV 등의 추정법을 사용했을 때 보정된 χ^2차이검정을 실시하게 된다. DIFFTEST는 더 복잡한 모형의 SAVEDATA 커맨드에서 DIFFTEST를 위한 정보를 저장한 뒤 이후 더 단순한 모형의 ANALYSIS 커맨드에 DIFFTEST를 입력하면 실행된다. 측정불변성 확인 순서는 각 모수의 불변성을 차례대로 검정하는 순차적 검정 방식 중 형태동일성 다음에 경계동일성을 먼저 검정하는 경계동일성 선행 방식을 따랐다. Delta 방식의 척도화를 실시했으므로 분산(공분산)동일성은 확인하지 않는다.

형태동일성 모형의 경우 전반적인 적합도 지수($\chi^2(df=80)=1049.287$, $p<.001$, RMSEA$=0.046$, CFI$=0.992$)가 상당히 양호했기 때문에 두 집단 간 형태동일성이 성립하는 것으로 판단하였다. 형태동일성에 대한 검정 이후 경계 모수 선행 방식을 적용한 측정불변성 검정 결과는 표 1에서 확인할 수 있다. 경계동일성의 경우 χ^2차이검정이 통계적으로 유의하여($\Delta\chi^2(df=18)=65.123$, $p<.05$) 기각되었다. 결과가 통계적으로 유의하면 두 집단 간 측정불변성이 확보되지 않았다는 결론을 내려야 한다. 그런데 Byrne, Shavelson과 Muthén(1989)은 부분측정불변성 개념을 도입하여 측정불변성의 여러 단계에서 일부의 모수가 집단 간에 다름을 인정하는 보다 느슨한 측정불변성 확보 방법을 제안한 바 있다. 다만, 부분측정불변성이 성립하려면 일부 문항 모수에 발생하는 집단 차이가 이론적으로 충분한 합리성을 지녀야 하며, 모형의 적합도도 기준을 만족해야 한다(Byrne, Shavelson, & Muthén, 1989). 이에 따라 추정된 18개의 경계 모수 중 한 개의 경계 모수(문항 2의 두 번째 경계모수)에 대한 동등성 제약을 풀어주었다.

표 1. 범주형 확인적 요인분석을 적용한 남녀 집단 간 측정불변성 검정 결과($N=11{,}704$)

	모형적합도					χ^2차이검정^a			근사적합도 지수 차이	
측정불변성 단계	χ^2	df	p	RMSEA	CFI	$\Delta\chi^2$	p	df	ΔRMSEA	ΔCFI
형태동일성	1049.287	80	.000	0.046	0.992					
경계동일성	1086.463	98	.000	0.042	0.991	65.123	.000	18	-	-
부분경계동일성	1038.761	97	.000	0.041	0.992	20.134	.267	17	-	-
(부분)요인부하동일성	932.570	104	.000	0.037	0.993	39.581	.000	7		
부분요인부하동일성	933.334	101	.000	0.038	0.993	1.370	.849	4	-	-

a. M*plus* DIFFTEST 결과치이다.

동등성 제약을 풀어주는 기준은 형태동일성 모형에서 추정된 경계모수의 집단 간 차이가 가장 큰 순서였다. 부분경계동일성 모형은 형태동일성 모형과의 χ^2차이검정이 통계적으로 유의하지 않아 ($\Delta\chi^2(df=17)=20.134$, $p>.05$) 더 단순한 모형인 부분경계동일성이 성립하였다. 이후 부분경계동일성 모형에 모든 요인부하 간 동등성 제약을 가해 준 (부분)요인부하동일성의 경우 χ^2차이검정이 통계적으로 유의하여 ($\Delta\chi^2(df=7)=39.581$, $p<.05$) (부분)요인부하동일성은 성립하지 않았다. 따라서 3개의 요인부하(문항 3, 4, 6)에 대해서도 추가적으로 동등성 제약을 풀어주었다. 경계 모수와 마찬가지로 동등성 제약을 풀어주는 기준은 부분경계동일성 모형에서 추정된 요인부하의 집단 간 차이가 가장 큰 순서였다. 부분요인부하동일성 모형은 부분경계동일성 모형과의 χ^2차이검정이 통계적으로 유의하지 않아 ($\Delta\chi^2(df=4)=1.370$, $p>.05$) 더 단순한 모형인 부분요인부하동일성이 성립하였다. 따라서 남녀 집단 간 축약형 CES-D 척도에서 경계 모수와 요인부하 모수가 일부 다른 부분측정불변성이 확보되었다. 부분측정불변성이 성립되었다는 결론을 내릴 수 있는 기준은 하나로 정해진 바 없으며, 일부 학자들은 전체 중 두 개의 지표변수들만 동등성 제약을 만족하면 된다고 주장하기도 한다(Steenkamp & Baumgartner, 1998). 만약 부분요인부하동일성까지도 성립하지 않는다는 결론을 내렸다면 집단 간 동일한 척도를 사용한 측정이 이루어졌다고 보기 어렵다. 따라서 집단 간 비교보다는 집단 내에서 어떠한 결과가 발생했는지에 대한 해석이 각각 이루어져야 한다.

여기서 표 1의 부분경계동일성과 부분요인부하동일성의 χ^2와 df값을 보면, 부분경계동일성 모형의 df가 더 작음에도 불구하고 χ^2값은 더 큰 것을 확인할 수 있다. WLSMV를 사용하는 경우 이처럼 예상치 못한 현상이 발생할 수 있는데, 그 명확한 원인은 아직 연구된 바 없다. Muthén(2009)에 의하면 χ^2값이 예상하지 못한 방향으로 흘러가더라도 DIFFTEST 옵션을 통해 내재된 두 모형의 비교를 실시할 수 있다. 이는 DIFFTEST 절차의 내부적인 알고리즘이 결과에 나타난 χ^2값을 직접 보정하여 그 차이 값에 대한 통계적 검정을 실시하는 것이 아니라, 추정과정에서 얻어지는 각 모형의 합치함수 차이의 최소값에 강건한 보정을 실시하여 χ^2분포를 따르는 검정 통계량을 재구성하기 때문이다. 이러한 과정에서 얻어진 검정 통계량은 두 모형의 차이에 대한 통계적인 결정을 내릴 수 있도록 해 준다. 다만 χ^2값이 예상하지 못한 방향으로 흘러간 경우에는 두 모형 간 근사적합도 지수 차이를 살피는 방식은 사용할 수 없음에 유의해야 한다. 이는 근사적합도 지수가 보정 전 χ^2값을 기반으로 하여 구해지기 때문이다. 보정 전 χ^2값에 문제가 생긴 이상 두 모형을 보정 없이 비교할 수 없다.

이와 같이 자료를 범주형으로 취급하면 축약형 CES-D 척도의 남녀 집단 간 부분적인 측정불변성을 χ^2차이검정을 통해 확보할 수 있다. 축약형 CES-D 척도를 연속형으로 취급한 채 분석을 실시한 Hoe, Park과 Bae(2015)의 연구에서 χ^2차이검정을 통해 요인부하동일성과 절편동일성 모두를 확보할 수 없었던 것에 비하면 다른 결론이 도출된 것이다. Hoe 등(2015)은 χ^2차이검정이 아닌 근사적합도 지수

를 확인하는 방법으로 측정불변성을 확보하였는데, 근사적합도 지수 차이를 살피는 방법은 그 기준이 주관적이고 임의적이기 때문에 그 사용방법에 대한 완전한 공감대가 형성되지 않았으므로 주의해야 한다(Byrne, 2012). 또한 앞서 말한 대로 근사적합도 지수는 보정 전 χ^2값을 기반으로 하여 구해지므로 χ^2값이 예상치 못한 방식으로 나타날 때는 더욱더 주의해야 한다. 이러한 측면에서 자료의 범주형 구조를 분석에 반영하는 것이 타당한 측정불변성 확보에 일정 부분 기여할 수도 있음이 시사되었다.

제안 및 논의

본 연구는 범주형 확인적 요인분석의 다집단 확장과 관련하여 모형의 판별에서부터 측정불변성 확인 단계까지의 쟁점들을 유기적으로 통합하여 논의하는 것을 목적으로 하였다. 최근 리커트 척도의 범주형 본질을 더 정확하게 반영한 분석방법에 대한 관심이 높아지고 있는 추세(Beauducel & Herzberg, 2006; Rhemtulla & Brosseau-Liard, 2012)에서 본 연구는 범주형 확인적 요인분석의 다집단 확장 시 필요한 판별 및 측정불변성 검정에 대한 상세한 이해를 도모하고, 연구자들이 다집단 분석과 같이 단일 시점의 자료를 다루는 경우 측정불변성 확인 과정에서 조금 더 편리하게 사용할 수 있는 척도화 방식에 대한 유연한 논의가 이루어졌다. 또한 범주의 수가 적음에도 불구하고 그간 연속형으로 취급된 채 분석되었던 실제 자료를 범주형으로 분석하는 예시를 보임으로써 실질적인 활용을 모색하였다.

지금까지 이루어져 온 관련 논의들을 탐색한 결과 도출할 수 있었던 적절한 합의점은 다음과 같다. 첫째, 모형의 판별을 위한 척도화에서 잠재반응변수의 척도화는 Delta 방식을, 요인의 척도화는 척도상수 방식을 사용할 것을 권장한다. 이는 Delta 방식이 Theta 방식보다 추정된 계수의 해석을 용이하게 한다는 이점을 지니고 있기 때문이다. 잠재반응변수의 척도화와 관련된 논의에 비해 요인의 척도화는 척도상수 방식을 사용해야 한다는 상대적으로 명확한 결론이 내려져 있다. 이는 요인을 표준화하는 방법을 사용하면 집단 간 요인분산이 모두 동일해져 집단 간 비교에 문제가 발생하기 때문이다. 둘째, 모형의 판별을 위한 추가제약과 관련해서는 그동안 문헌에서 범주형 CFA의 다집단 확장을 위한 추가제약에 대한 논의가 명확히 이루어진 적이 많지 않다는 점을 짚어 보아야 한다. 사실상 추가제약 방식은 특정한 한 가지로 정해져 있는 것이 아니므로 현재 분석을 위해 가장 자주 활용되고 있는 Millsap과 Yun-Tein(2004)의 방식 이외에도 더욱 효율적인 추가제약을 실시할 방안이 있는가에 대해 향후 연구가 필요하다. 셋째, 측정불변성 확인 단계와 관련해서 범주형 CFA의 경우 경계 모수의 동등성 확보가 지니는 중요성에 주목해야 한다. 통상적으로 범주형 CFA의 측정불변성은 최소한 형태동일성, 요인부하동일성, 경계동일성까지 확보(strong invariance)해야 한다고 본다. 그러나 측정불변성을 확인하는 방식이나 단계는 연구자의 주관적인 판단이나 선택에 따라 다양해 질 수 있으며, 특히 경계동일성을 확인하는 순서는 요인부하동일성 확보의 전후 단계에 모두 위치할 수 있다. 마지막으로 위의 논의에서 도출된 합의점들을 실제 자료에 적용하여 본 결과, 자료의 범주형 취급이 χ^2차이검정의 과

도한 기각 문제를 보완하는 것이 어느 정도 가능할 수도 있음을 시사했다. 지표변수를 범주형으로 취급하고 분석한 결과 χ^2차이검정을 통해 부분적인 측정불변성 확보가 가능하였다.

본 연구는 범주형 확인적 요인분석의 다집단 확장을 위한 척도화 및 추가제약 방식, 측정불변성 확인 단계에 대한 논의를 통합하고 적절한 합의점을 도출하고자 하는 목적은 달성하였지만 여타 다른 모든 연구가 그러하듯 제한점을 지니고 있다. 먼저, 경로의 불변성에 대한 논의가 이루어지지 않았다. 다만 구조방정식 맥락에서 자료의 범주형 취급은 측정모형의 측정불변성 확인에 변화를 가져올 뿐, 측정불변성이 확보된 이후 실시하는 경로불변성과 관련된 분석은 자료를 연속형으로 취급한 경우와 다를 것이 없다. 다음으로, 집단이 두 개인 상황만을 가정하고 논의를 진행하였다. 집단이 세 개 이상인 경우 본 연구에서 논의된 척도화 방식을 그대로 적용하여 쌍별 비교를 실시할 수는 있지만, 일종오류의 증가로 인해 불변성을 잘못 검정할 확률이 높아진다(Rutkowski & Svetina, 2014). 또한 근사적합도지수를 이용한 검정을 실시하더라도 제시된 기준이 두 집단을 가정하였기 때문에 집단이 늘어날수록 그 기준을 충족하기는 어려워진다(Asparouhov & Muthén, 2014). 따라서 집단이 3개 이상인 경우에는 다층 요인혼합모형(multilevel factor mixture model)이나 베이지안 근사 척도불변성 검정(Bayesian approximate measurement invariance testing), 혹은 정렬 최적화(alignment optimization) 방식과 같은 대안적인 방법들을 고려해야 한다(Kim, Cao, Wang, & Nguyen, 2017).

언급된 제한점에도 불구하고, 본 연구는 다집단 범주형 확인적 요인분석 분야에 산재되어 있던 논의를 종합하여 방법론적 측면에서 심도 깊은 논의를 진행하였다는 데 그 의의가 있다. 구체적으로 모형의 판별과 제약 및 측정불변성 검정 단계에 내재된 원리를 유기적으로 살펴보고 더 나아가 통계 프로그램을 이용하여 실제 자료를 분석하는 방법을 소개함과 동시에 결과에 대한 해석을 보였다. 이를 통해 향후 연구자들이 자신의 연구 자료에 다집단 범주형 요인분석을 적용하는 것이 더욱 수월해질 수 있을 것이다.

참고문헌

Asparouhov, T., & Muthén, B., (2006). Robust chi square difference testing with mean and variance adjusted test statistics. *Matrix, 1*(5), 1-6.

Asparouhov, T., & Muthén, B. (2010). Simple second order chi-square correction. *Unpublished manuscript.*

Asparouhov, T., & Muthén, B. (2014). Auxiliary variables in mixture modeling: Three-step approaches using M plus. *Structural Equation Modeling, 21*(3), 329-341.

Baas, K. D., Cramer, A. O., Koeter, M. W., van de Lisdonk, E. H., van Weert, H. C., & Schene, A. H. (2011). Measurement invariance with respect to ethnicity of the Patient Health Questionnaire-9 (PHQ-9). *Journal of Affective Disorders, 129*(1-3), 229-235.

Batinic, B., Wolff, H. G., & Haupt, C. M. (2008). Construction and Factorial Structure of a Short Version of the Trendsetting

Questionnaire (TDS-K) A Cross-Validation Using Multigroup Confirmatory Factor Analyses. *European Journal of Psychological Assessment, 24*(2), 88-94.

Beauducel, A., & Herzberg, P. Y. (2006). On the performance of maximum likelihood versus means and variance adjusted weighted least squares estimation in CFA. *Structural Equation Modeling, 13*(2), 186-203.

Bernstein, I. H., & Teng, G. (1989). Factoring items and factoring scales are different: Spurious evidence for multidimensionality due to item categorization. *Psychological Bulletin, 105*(3), 467-477.

Bollen, K. A. (1989). A new incremental fit index for general structural equation models. *Sociological Methods & Research, 17*(3), 303-316.

Byrne, B. M. (2012). *Stuctural equation modeling with Mplus: Basic concepts, applications, and programming.* New York, NY: Taylor & Francis Group.

Byrne, B. M., Shavelson, R. J.. & Muthén, B. (1989). Testing for the equivalence of factor covariance and mean structures: The issue of partial meausrement invariance. *Psychological Bulletin, 105*(3), 456-466.

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement Invariance. *Structural Equation Modeling, 14*(3), 464-504.

Cheung, G. W., & Rensvold, R. B. (2000). Assessing extreme and acquiescence response sets in cross-cultural research using structural equations modeling. *Journal of Cross-Cultural Psychology, 31*(2), 187-212.

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling, 9*(2), 233-255.

Choi, S. & Cho, Y. (2013). Measurement invariance of self-esteem between male and female adolescents and method effects associated with negatively worded items. *Korean Journal of Psychology: General, 32*(3), 571-589.

Christoph, G., Goldhammer, F., Zylka, J., & Hartig, J. (2015). Adolescents' computer performance: The role of self-concept and motivational aspects. *Computers & Education, 81,* 1-12.

Cooke, D. J., Kosson, D. S., & Michie, C. (2001). Psychopathy and ethnicity: Structural, item, and test generalizability of the Psychopathy Checklist-Revised (PCL-R) in Caucasian and African American Participants, *Psychological Assessments, 13*(4), 531-542.

Dimitrov, D. M. (2010). Testing for factorial invariance in the context of construct validation. *Measurement and Evaluation in Counseling and Development, 43*(2), 121-149.

Diniz, A., Pocinho, M. D., & Almeida, L. S. (2011). Cognitive abilities, sociocultural background and academic achievement. *Psicothema,* 695-700.

Dolan, C. V. (1994). Factor analysis of variables with 2, 3, 5 and 7 response categories: A comparison of categorical variable estimators using simulated data. *British Journal of Mathematical and Statistical Psychology, 47*(2), 309-326.

Edwards, M. C., Wirth, R. J., Houts, C. R., &

Xi, N. (2012). Categorical data in the structural equation modeling framework. In R. Hoyle (Eds.), *Handbook of structural equation modeling* (pp. 195-208). New York, NY: The Guilford Press.

Elosua, P. (2011). Subjective values of quality of life dimensions in elderly people. A SEM preference model approach. *Social indicators research, 104*(3), 427-437.

Flora, D. B., & Curran, P. J. (2004). An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. *Psychological methods, 9*(4), 466-491.

Forbush, K. T., Wildes, J. E., Pollack, L. O., Dunbar, D., Luo, J., Patterson, K., ... & Watson, D. (2013). Development and validation of the Eating Pathology Symptoms Inventory (EPSI). *Psychological Assessment, 25*(3), 859-878.

Forero, C. G., Maydeu-Olivares, A., & Gallardo-Pujol, D. (2009). Factor analysis with ordinal indicators: A Monte Carlo study comparing DWLS and ULS estimation. *Structural Equation Modeling, 16*(4), 625-641.

French, B. F., & Finch, W. H. (2006). Confirmatory factor analytic procedures for the determination of measurement invariance. *Structural Equation Modeling, 13*(3), 378-402.

Glanville, J. L., & Wildhagen, T. (2007). The measurement of school engagement: Assessing dimensionality and measurement invariance across race and ethnicity. *Educational and Psychological Measurement, 67*(6), 1019-1041.

Glockner-Rist, A., & Hoijtink, H. (2003). The best of both worlds: Factor analysis of dichotomous data using item response theory and structural equation modeling. *Structural Equation Modeling, 10*(4), 544-565.

Hempel, L. M., & Bartkowski, J. P. (2008). Scripture, sin and salvation: Theological conservatism reconsidered. *Social Forces, 86*(4), 1647-1674.

Hoe, M, Park, B.-S., & Bae, S.-W. (2015). Testing measurement invariance of the 11-tiem Korean version CES-D scale. *Meantal Health and Social Work, 43*(2), 313-339.

Hutchinson, S. R., & Olmos, A. (1998). Behavior of descriptive fit indexes in confirmatory factor analysis using ordered categorical data. *Structural Equation Modeling, 5*(4), 344-364.

Johnson, D. R., & Creech, J. C. (1983). Ordinal measures in multiple indicator models: A simulation study of categorization error. *American Sociological Review, 48*(3) 398-407.

Kamata, A., & Bauer, D. J. (2008). A note on the relation between factor analytic and item response theory models. *Structural Equation Modeling, 15*(1), 136-153.

Katsikatsou, M., Moustaki, I., Yang-Wallentin, F., & Jöreskog, K. G. (2012). Pairwise likelihood estimation for factor analysis models with ordinal data. *Computational Statistics & Data Analysis, 56*(12), 4243-4258.

Kim, E. S., Cao, C., Wang, Y., & Nguyen, D. T. (2017). Measurement invariance testing with many groups: a comparison of five approaches. *Structural Equation Modeling, 24*(4), 524-544.

Kim, E. S., & Yoon, M. (2011). Testing measurement invariance: A comparison of multiple-group categorical CFA and IRT.

Structural Equation Modeling, 18(2), 212-228.

Kline, R. B. (2016). *Principles and practice of structural equation modeling.* New York, NY: Guilford publications.

Kim, S.-Y. (2016). *Fundamentals and extensions of structural equation modeling.* Seoul: Hakjisa.

Koğar, H., & Koğar, E. Y. (2015). Comparison of different estimation methods for categorical and ordinal data in confirmatory factor analysis. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi, 6*(2) 351-364.

Lawrence, J. W., Rosenberg, L., Rimmer, R. B., Thombs, B. D., & Fauerbach, J. A. (2010). Perceived stigmatization and social comfort: Validating the constructs and their measurement among pediatric burn survivors. *Rehabilitation Psychology, 55*(4), 360-371.

Lee, J. & Yoo, J. P. (2015). Influence of body shape perception on self-esteem among normal-weight middle school students in South Korea: The medicating effect of body shape satisfaction and gender differences. *Studies on Korean Youth, 26*(4), 267-297.

Lee, P., Chun, S., & Lee, S. (2014). Differential use of middle catetory "?" in job descriptive index between Korean and American samples: Application of mixed-model item response theory. *Korean Journal of Psychology: General, 33*(3), 647-669.

Lee, S., Lee, C., Lee, H., & Yeo, S. (2012). Invariance of conceptual structure and psychometric properties in Canadian problem gambling index. *Korean Journal of Psychology: General, 31*(1), 1-26.

Lee, S., Youn, C., Lee, M. M., & Jung, S. (2016). Exploratory factor analysis: How has it changed? *Korean Journal of Psychology: General, 35*(1), 217-255.

Lee, S. Y., Poon, W. Y., & Bentler, P. M. (1992). Structural equation models with continuous and polytomous variables. *Psychometrika, 57*(1), 89-105.

Looij-Jansen, P. M., Goedhart, A. W., de Wilde, E. J., & Treffers, P. D. (2011). Confirmatory factor analysis and factorial invariance analysis of the adolescent self report Strengths and Difficulties Questionnaire: How important are method effects and minor factors?. *British Journal of Clinical Psychology, 50*(2), 127-144.

Lopez Rivas, G. E., Stark, S., & Chernyshenko, O. S. (2009). The effects of referent item parameters on differential item functioning detection using the free baseline likelihood ratio test. *Applied Psychological Measurement, 33*(4), 251-265.

Little, T. D. (1997). Mean and covariance structures (MACS) analyses of cross-cultural data: Practical and theoretical issues. *Multivariate Behavioral Research, 32*(1), 53-76.

Marsh, H. W. (2007). Students' evaluations of university teaching: Dimensionality, reliability, validity, potential biases and usefulness. In Perry, R. P. and Smart, J. C. (Eds.), *The scholarship of teaching and learning in higher education: An evidence-based perspective* (pp. 319-383). Springer, Dordrecht.

Mathyssek, C. M., Olino, T. M., Hartman, C. A., Ormel, J., Verhulst, F. C., & Van Oort, F. V. (2013). Does the Revised Child Anxiety and Depression Scale (RCADS) measure anxiety symptoms consistently across adolescence? The TRAILS study. *International Journal of Methods in Psychiatric Research, 22*(1), 27-35.

Maydeu-Olivares, A. (2001). Multidimensional item response theory modeling of binary data: Large sample properties of NOHARM estimates. *Journal of Educational and Behavioral Statistics, 26*(1), 51-71.

Maydeu-Olivares, A., & Cai, L. (2006). A cautionary note on using G2(dif) to assess relative model fit in categorical data analysis. *Multivariate Behavioral Research, 41*(1), 55-64.

McKelvey, R. D., & Zavoina, W. (1975). A statistical model for the analysis of ordinal level dependent variables. *Journal of Mathematical Sociology, 4*(1), 103-120.

Meade, A. W., Johnson, E. C., & Braddy, P. W. (2008). Power and sensitivity of alternative fit indices in tests of measurement invariance. *Journal of Applied Psychology, 93*(3), 568-592.

Mellenbergh, G. J. (1989). Item bias and item response theory. *International journal of educational research, 13*(2), 127-143.

Meredith, W. (1964). Notes on factorial invariance. *Psychometrika, 29*(2), 177-185.

Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika, 58*(4), 525-543.

Meuleman, B., & Billiet, J. (2012). Measuring attitudes toward immigration in Europe: The cross-cultural validity of the ESS immigration scales. *Research & Methods, 21*(1), 5-29.

Millsap, R. E., & Kim, H. (2018). Factorial invariance across multiple populations in discrete and continuous data. In P. Irwing, T. Booth, D. J. Hughes (Eds.), *The Wiley Handbook of Psychometric Testing: A Multidisciplinary Reference on Survey, Scale and Test Development* (pp. 847-884), Hoboken, NJ, US: Wiley.

Millsap, R. E., & Yun-Tein, J. (2004). Assessing factorial invariance in ordered-categorical measures. *Multivariate Behavioral Research, 39*(3), 479-515.

Moor, M. H., Distel, M. A., Trull, T. J., & Boomsma, D. I. (2009). Assessment of borderline personality features in population samples: Is the Personality Assessment Inventory-Borderline Features scale measurement invariant across sex and age?. *Psychological Assessment, 21*(1), 125-130.

Morin, A. J., Tran, A., & Caci, H. (2016). Factorial validity of the ADHD Adult Symptom Rating Scale in a French community sample: results from the ChiP-ARD study. *Journal of Attention Disorders, 20*(6), 530-541.

Muthén, B. (1984). A general structural equation model with dichotomous, ordered categorical, and continuous latent variable indicators. *Psychometrika, 49*(1), 115-132.

Muthén, B. (2006). Model fit indices for two-part growth model. Retrived from http://www.statmodel.com/discussion/messages/14/1118.html?1352499941

Muthén, B. (2009). DIFFTEST related. Retrived from http://www.statmodel.com/discussion/messages/11/621.html?1540568737

Muthén, B. (2011). Multiple group analysis with categorical variables. Retrived from http://www.statmodel.com/discussion/messages/23/25.html?1553558399

Muthén, B., & Asparouhov, T. (2002). Latent variable analysis with categorical outcomes: Multiple-group and growth modeling in Mplus. *Mplus web notes, 4*(5), 1-22.

Muthén, B. O., du Toit, S. H. C., & Spisic, D. (1997). Robust interference using weighted least squares and quadratic estimating equations in the latent variable modeling with categorical and continuous outcomes. *Unpublished manuscript,* University of California, Los Angeles, USA.

Muthen, B., & Kaplan, D. (1992). A comparison of some methodologies for the factor analysis of non normal Likert variables: A note on the size of the model. *British Journal of Mathematical and Statistical Psychology, 45*(1), 19-30.

Muthén, L., & Muthén, B. (1998-2019). Mplus (Version 7). Los Angeles, CA: Muthén & Muthén.

Nussbeck, F. W., Eid, M., & Lischetzke, T. (2006). Analyzing multitrait-multimethod data with sturctural equation models for ordinal variables applying the WLSMV estimator: What sample size is needed for valid results? *British Journal of Mathematical and Statistical Psychology, 59*(1), 195-213.

Potthast, M. J. (1993). Confirmatory factor analysis of ordered categorical variables with large models. *British Journal of Mathematical and Statistical Psychology, 46*(2), 273-286.

Reise, S. P., Widaman, K. F., & Pugh, R. H. (1993). Confirmatory factor analysis and item response theory: Two approaches for exploring measurement invariance. *Psychological Bulletin, 114*(3), 552-566.

Rhemtulla, M., Brosseau-Liard, P. É., & Savalei, V. (2012). When can categorical variables be treated as continuous? A comparison of robust continuous and categorical SEM estimation methods under suboptimal conditions. *Psychological Methods, 17*(3), 354-373.

Rutkowski, L., & Svetina, D. (2014). Assessing the hypothesis of measurement invariance in the context of large-scale international surveys. *Educational and Psychological Measurement, 74*(1), 31-57.

Rutkowski, L., & Svetina, D. (2017). Measurement invariance in international surveys: Categorical indicators and fit measure performance. *Applied Measurement in Education, 30*(1), 39-51.

Sass, D. A., & Schmitt, T. A. (2013). Testing measurement and structural invariance: Implications for practice. In T. Teo (Eds.) *Handbook of quantitative methods for educational research* (pp. 315-345). Rotterdam: Sense Publishers.

Sass, D. A., Schmitt, T. A., & Marsh, H. W. (2014). Evaluating model fit with ordered categorical data within a measurement invariance framework: A comparison of estimators. *Structural Equation Modeling, 21*(2), 167-180.

Satorra, A., & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika, 66*(4), 507-514.

Satterthwaite, F. E. (1941). Synthesis of variance.

Psychometrika, 6(5), 309-316.

Schlotz, W., Yim, I. S., Zoccola, P. M., Jansen, L., & Schulz, P. (2011). The perceived stress reactivity scale: Measurement invariance, stability, and validity in three countries. *Psychological assessment, 23*(1), 80-94.

Schreiber, J. B., Nora, A., Stage, F. K., Barlow, E. A., & King, J. (2006). Reporting structural equation modeling and confirmatory factor analysis results: A review. *The Journal of educational research, 99*(6), 323-338.

Selig, J. P., Card, N. A., & Little, T. D. (2008). Latent variable structural equation modeling in cross-cultural research: Multigroup and multilevel approaches. In F. J. R. van de Vijver, & D. A. van Hemert, Y. H. Poortinga (Eds.), *Multilevel analysis of individuals and cultures*(pp. 93-119), New York, NY: Taylor & Francis.

Seo, D. & Lee, S. (2017). Item response theory approach using item factor anlaysis in the development of psycholgocial test. A packet for the Korean Psychological Association, 137.

Stark, S., Chernyshenko, O. S., Drasgow, F., & Williams, B. A. (2006). Examining assumptions about item responding in personality assessment: Should ideal point methods be considered for scale development and scoring?. *Journal of Applied Psychology, 91*(1), 25-39.

Steenkamp, J. B. E., & Baumgartner, H. (1998). Assessing measurement invariance in cross-national consumer research. *Journal of consumer research, 25*(1), 78-90.

Temme, D. (2006). Assessing measurement invariance of ordinal indicators in cross-national research. In S. Diehl, R. Terlutter (Eds.), *International advertising and communication: Current insights and empirical findings* (pp. 455-472), Springer Science & Business Media.

Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods, 3*(1), 4-70.

Wang, W. C., & Yeh, Y. L. (2003). Effects of anchor item methods on differential item functioning detection with the likelihood ratio test. *Applied Psychological Measurement, 27*(6), 479-498.

Widaman, K. F. (2000). Testing cross-group and cross-time constraints on parameters using the general linear model. In T. D. Little, K.U. Schnabel, & J. Baumert (Eds.), *Modeling longitudinal and multilevel data: Practical issues, applied approaches, and specific examples* (pp. 163-185, 269-281). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

Widaman, K. F., & Reise, S. P. (1997). Exploring the measurement invariance of psychological instruments: Applications in the substance use domain. In K. J. Bryant, M. Windle, & S. G. West (Eds.), *The science of prevention: Methodological advances from alchol and substance abuse research* (pp. 281-324), Washington, DC, US: American Psychologocal Association.

Wu, H., & Estabrook, R. (2016). Identification of confirmatory factor analysis models of different levels of invariance for ordered categorical outcomes. *Psychometrika, 81*(4), 1014-1045.

Xing, C., & Hall, J. A. (2015). Confirmatory factor analysis and measurement invariance testing with ordinal data: An application in revising the flirting styles inventory. *Communication Methods and Measures, 9*(3), 123-151.

Yang, Y., & Green, S. B. (2010). A note on structural equation modeling estimates of reliability. *Structural Equation Modeling, 17*(1), 66-81.

Yang-Wallentin, F., Jöreskog, K. G., & Luo, H. (2010). Confirmatory factor analysis of ordinal variables with misspecified models. *Structural Equation Modeling, 17*(3), 392-423.

Yoon, M., & Millsap, R. E. (2007). Detecting violations of factorial invariance using data-based specification searches: A Monte Carlo study. *Structural Equation Modeling, 14*(3), 435-463.

Liu, Y., Millsap, R. E., West, S. G., Tein, J.-Y., Tanaka, R., & Grimm, K. J. (2017). Testing measurement invariance in longitudinal data with ordered-categorical measures. *Psychological Methods, 22*(3), 486-506.

Zimprich, D., Allemand, M., & Dellenbach, M. (2009). Openness to experience, fluid intelligence, and crystallized intelligence in middle-aged and old adults. *Journal of Research in Personality, 43*(3), 444-454.

1차원고접수 : 2019. 11. 07.
2차원고접수 : 2020. 03. 07.
최종게재결정 : 2020. 04. 03.

한국심리학회지: 일반
Korean Journal of Psychology: General
2020, Vol. 39, No. 2, 175-204
http://dx.doi.org/10.22257/kjp.2020.6.39.2.175

A multi-group analysis using the categorical confirmatory factor analysis

Mimi Shin Su-Young Kim

Department of Psychology, Ewha Womans University

Categorical confirmatory factor analysis (CFA) is a measurement model that incorporates the categorical nature of the Likert scale, which is frequently used in the social sciences, including but not limited to psychology. The categorical CFA has quite different and complex model specification and identification as compared to the typical continuous CFA. A multi-group extension of the categorical CFA needs even more restrictions to be properly specified and identified. Moreover, the steps in checking the measurement invariance with the categorical indicator variables differ from those with the case of continuous indicator variables. The present study aims to help researchers choose proper methods and procedures for using a multi-group categorical CFA by investigating and integrating unorganized existing literature. To achieve this goal, we focus on two aspects. Based on the limited information estimation, we introduce scaling methods for the two types of latent variables, latent response variables and factors, and also address the method for parameter restrictions for multi-group analysis. Next, focusing on threshold parameters of the categorical CFA, we provide possible procedures for checking the measurement invariance. We then illustrate an application of the whole procedures using real data, and finally discuss the results.

Key words : categorical data, multi-group analysis, factor analysis, limited information estimation, measurement invariance

부 록

A. 다집단 범주형 확인적 요인분석 - 형태동일성 모형

```
TITLE: multigroup cat-CFA_형태동일성
DATA: FILE = CES-D.dat;
VARIABLE: NAMES ARE phy1-phy4 pos1-pos2 dep1-dep3 rel1-rel2 g;
          USEVARIABLES ARE phy1-phy4 pos1-pos2 dep1-dep3 rel1-rel2 g;
          MISSING ARE all(999);
          CATEGORICAL ARE phy1-phy4 pos1-pos2 dep1-dep3 rel1-rel2;
                                  !집단변수를 제외한 모든 변수를 범주형으로 지정
          GROUPING IS g (0=male 1=female);
ANALYSIS: ESTIMATOR = WLSMV;
                                  !WLSMV추정법 사용
          PARAMETERIZATION = Delta;
                                  !잠재반응변수의 척도화 방식으로 Delta방식 사용
MODEL: phy BY phy1@1 phy2-phy4; pos BY pos1@1 pos2;
       dep BY dep1@1 dep2-dep3; rel BY rel1@1 rel2;
       depress BY phy@1 pos dep rel;
       phy pos dep rel depress;
       [phy@0 pos@0 dep@0 rel@0 depress@0];
       [phy1$1 phy2$1 phy3$1 phy4$1 pos1$1 pos2$1
        dep1$1 dep2$1 dep3$1 rel1$1 rel2$1];
       [phy1$2 phy2$2 phy3$2 phy4$2 pos1$2 pos2$2
        dep1$2 dep2$2 dep3$2 rel1$2 rel2$2];
       [phy1$3 phy2$3 phy3$3 phy4$3 pos1$3 pos2$3
        dep1$3 dep2$3 dep3$3 rel1$3 rel2$3];
                                  !추정된 문항 별 경계 모수
        {phy1-phy4@1}; {pos1-pos2@1}; {dep1-dep3@1}; {rel1-rel2@1};
                                  !한 집단의 잠재반응변수 분산을 1로 고정
MODEL female:
      phy BY phy1@1 phy2-phy4; pos BY pos1@1 pos2;
      dep BY dep1@1 dep2-dep3; rel BY rel1@1 rel2;
                  !요인부하동일성 확인 시 위의 명령어 전부 삭제
                  !부분요인부하동일성 허용 시 동일성 제약을 풀어줄 문항 기재
```

```
    depress BY phy@1 pos dep rel;
    phy pos dep rel depress;
    [phy pos dep rel depress@0];
    [phy1$1 phy2$1 phy3$1 phy4$1 pos1$1 pos2$1
     dep1$1 dep2$1 dep3$1 rel1$1 rel2$1];
    [phy2$2 phy3$2 phy4$2 pos2$2
     dep2$2 dep3$2 rel2$2];
                !경계 모수에 대한 추가제약
                !경계동일성 확인 시 위의 명령어 전부 삭제
                !부분경계동일성 허용 시 동일성 제약을 풀어줄 경계 모수 기재
SAVEDATA: DIFFTEST IS dif.dat;
                !이후 내재된 모형의 ANALYSIS 명령어에 DIFFTEST=dif.dat 기재 시
                !내재된 모형의 결과에서 DIFFTEST실시결과 확인 가능
```

B. 범주 별 피험자 누적 선택 비율-남, 녀 집단 기준

	범주1 (남)	범주1 (여)	범주1 (집단차)	범주2 (남)	범주2 (여)	범주2 (집단차)	범주3 (남)	범주3 (여)	범주3 (집단차)
문항1	0.835	0.717	0.118	0.953	0.909	0.044	0.989	0.979	0.010
문항2	0.666	0.565	0.101	0.920	0.872	0.048	0.981	0.972	0.009
문항3	0.733	0.581	0.152	0.922	0.860	0.062	0.982	0.967	0.015
문항4	0.759	0.69	0.069	0.910	0.909	0.001	0.955	0.978	0.023
문항5	0.759	0.691	0.068	0.927	0.894	0.033	0.986	0.984	0.002
문항6	0.724	0.668	0.056	0.901	0.881	0.020	0.980	0.978	0.002
문항7	0.809	0.706	0.103	0.966	0.937	0.029	0.993	0.986	0.007
문항8	0.816	0.739	0.077	0.968	0.948	0.020	0.993	0.987	0.006
문항9	0.856	0.767	0.089	0.976	0.954	0.022	0.995	0.991	0.004
문항10	0.936	0.931	0.005	0.990	0.986	0.004	0.997	0.995	0.002
문항11	0.947	0.941	0.006	0.992	0.990	0.002	0.998	0.996	0.002

전반적으로 살펴볼 때 집단 간 누적비율 차이가 가장 없는 세 번째 경계모수를 최소제약모형의 동등성 제약 대상으로 지정함.