

문항묶음: 원리의 이해와 적용*

이 지 현

김 수 영†

이화여자대학교

구조방정식 모형의 틀에서 관심 있는 구인(construct)을 분석하는 방법이 널리 쓰이게 되면서, 모형을 어떻게 더 정확하게 설정하고 모형적합도를 높일 수 있을지에 대한 관심이 많아졌다. 이를 위한 방법 중 하나가 문항묶음(item parcel)을 구인의 지표로 사용하는 것이다. 문항묶음은 경로모형에서도 사용될 수 있으나, 주로 측정모형을 설정할 때 사용할 수 있는 기법으로, 두 개 이상의 개별문항 점수를 합산하거나 평균을 내어 만든 묶음을 구인의 지표로 사용하는 것이다. 문항묶음 사용에 대한 논쟁은 여전히 진행 중이며, 여러 방법론적 연구가 진행되고 있다. 최근 국내의 다양한 분야에서 문항묶음을 사용한 실질연구가 발표되고 있으나, 사용을 위한 중요한 가정을 확인하지 않거나 필요한 정보를 제공하지 않은 논문들이 많았다. 따라서 본 연구는 지난 20여 년간 이루어진 문항묶음 연구들을 통합하고, 이에 기반하여 일반 연구자들에게 적절한 절차를 제안하는 것을 목적으로 한다. 먼저, 문항묶음을 사용하는 이유를 측정학적 관점과 구조방정식모형 관점으로 나누어 설명한다. 다음으로는 문항묶음 사용을 둘러싼 여러 논쟁들과 제안점에 관해 논의하고, 실질적으로 문항묶음을 어떻게 만들 것인지를 다룬다. 마지막으로, 문항묶음을 사용하려는 연구자들이 참고할 수 있는 적절한 단계를 제안하고, 사용 시 주의점에 대해서 논의한다.

주요어 : 문항묶음, 묶음화, 구조방정식 모형, 측정모형

* 본 연구는 이지현(2015)의 이화여자대학교 석사학위 논문을 수정 보완한 것임.

* 이 논문은 2014년 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (NRF- 2014S1A3A2038236).

† 교신저자: 김수영, 이화여자대학교 심리학과, 서울시 서대문구 이화여대길 52, Tel. 02-3277-3792, E-mail: suyong.kim@ewha.ac.kr

구조방정식 모형과 같이 발전된 자료 분석 방법이 일반 연구자들에게 널리 쓰이게 되면서, 잠재변수를 다루기가 이전보다 용이해졌다. 이에 따라 관심 있는 구인(construct)을 측정하고, 그 측정구조나 여러 구인 간의 관계 등을 확인하려는 연구 또한 많아지고 있다. 일반적으로 구조방정식 모형에서 구조관계를 확인하기 이전에 측정모형의 적합성을 확인하는데(Anderson & Gerbing, 1988), 이때 각 구인을 측정하는 지표(indicators) 또는 지표변수(indicator variables)가 지니는 측정학적 특성이 매우 중요하다. 즉, 구인과 상응하는 지표와의 관계가 잘 정의되어야 그 이후 구인 간의 구조를 검증하는 것이 의미 있게 된다.

측정모형이나 구조방정식 모형에서 표준으로 쓰이는 최대우도(maximum likelihood, ML) 추정방법은 지표변수의 정규성 및 연속성 등을 가정한다. 만약 이러한 기본 가정을 만족하지 못한 경우 편향된 추정치를 주는 것으로 알려져 있다(Curran, West, & Finch, 1996; Muthén, 1989). 정규성이나 연속성을 잘 만족하지 않는 순위형 변수(예, 리커트 척도)를 사용할 때, 제한정보추정 또는 완전정보추정 방법을 이용한 비선형(nonlinear) 또는 범주형(categorical) 요인분석을 실시할 수 있다. 또는 순위형 변수의 범주 개수가 충분히 많다면, 순위형 변수를 마치 연속형 변수인 것처럼 가정하고 일반적인 선형(linear) 요인분석을 사용하기도 한다.

이와 같이 정규성 등을 잘 만족하지 못하는 지표변수를 이용할 때, 문항묶음을 사용하는 방법도 광범위하다. 문항묶음은 두 개 이상의 개별문항 점수를 합산하거나 평균을 내어 만든 것으로, 경로모형의 변수로 이용하거나, 요인모형에서 측정하고자 하는 구인의 지표로

사용한다(Kishton & Widaman, 1994). 이때, 합산 점수 혹은 평균점수 모두 문항묶음을 형성하기 위해 사용되지만, 되도록 평균점수를 사용할 것이 권장된다(Little, Rhemtulla, Gibson, & Schoemann, 2013). 이는 묶음에 포함된 문항 수가 다를 경우 묶음의 평균과 분산을 비슷한 수준으로 맞출 수 있고, 실제 척도를 반영하여 해석하고 비교하기 좋기 때문이다. 잠재변수를 측정하기 위해 개별문항을 그대로 사용하는 것보다 여러 문항들을 통합하여 사용하게 되면, 지표변수의 정규성 및 연속성 가정을 더 잘 만족시킬 수 있기 때문에 문항묶음화(item parceling)¹⁾는 자료를 다루는 중요한 기술로 자리 잡고 있다. 또한 지표변수로 문항묶음을 사용하게 되면, 연구자들에게 어려울 수 있는 비선형 요인분석을 실시할 필요 없이, 선형 요인분석을 사용할 수 있다는 것도 실질적인 큰 장점이다.

지난 20여 년간 방법론적 연구가 활발하게 진행되어 여러 문항묶음의 장단점들이 논의되었다. 예를 들어, 문항묶음을 사용하면 분포의 비정규성 문제가 완화될 수 있으며(Bandalos, 2002; Hau & Marsh, 2004), 더 안정적이고 정확한 모수 추정이 가능하고(Coffman & MacCallum, 2005; Nasser & Wisenbaker, 2006; Rogers & Schmitt, 2004), 모형적합도가 향상되는(Bandalos, 2002; Nasser & Wisenbaker, 2003; Rogers & Schmitt, 2004) 통계적 이점을 보이는 것으로 알려져 있다. 그러나 적절하지 않은 방식으로 사용된다면 오히려 연구의 타당성을 저해할 수 있다. 예를 들어, 모형이 잘못 설정된 경우에도 양호한 적합도 지수를 보임으로써 오히

¹⁾ 본 논문에서는 item parcel을 문항묶음이라고 명명하고, 맥락에 따라 parceling은 묶음화(문항묶음을 만드는 작업)로 명명한다.

러 모형의 설정 오류를 발견하지 못하게 할 수 있다(Bandalos, 2002; Marsh, Lüdtke, Nagengast, Morin, & Von Davier, 2013).

이러한 국제적인 방법론적 연구와 더불어 최근 국내에서도 심리학뿐 아니라, 경영학, 소비자학, 사회복지학 등 다양한 분야에서 문항 묶음을 이용한 연구가 발표되고 있다. 그러나 아직 ‘문항묶음’, ‘문항꾸러미’, ‘항목묶음’, ‘묶음지표’, ‘항목합산’ 등으로 각기 다르게 명명되고 있어 그 용어조차 통일되지 않은 상태이다. 또한 최근 국내에서 문항묶음을 사용했다고 밝힌 50여개의 논문을 확인한 결과, 문항 묶음을 만들기 전에 꼭 확인해야 할 사항인 구인의 차원성(dimensionality)에 대한 정보를 제공하지 않은 경우가 약 54%, 어떤 문항묶음화 방법을 사용했는지 밝히지 않은 경우는 약 26%, 어떤 문항을 묶었는지 구체적인 내용을 제시하지 않은 경우는 약 72%에 이르렀다. 또한 문항묶음의 개념을 혼동하거나 정확하지 않은 개념적 진술을 한 경우도 있었다. 이와 같은 일이 발생하는 원인 중 하나는 아마도, 연구자들이 참고할 수 있는 문항묶음에 관한 방법론적 논문이 국내에 부족하기 때문일 것이다.

국내에 발표된 논문 중에서는 조현철과 강석후(2007)와 우종필(2015)이 문항묶음에 대한 개념적 논의와 더불어 실증적 자료 분석을 통해 문항묶음의 효과를 다루고 있다. 조현철과 강석후(2007)는 두 개의 문항묶음을 사용했을 때 경로모수의 유의성과 모형적합도의 변화를 확인하였다. 이 연구는 국내에서 문항묶음화에 대한 개념적·실증적 연구를 시작했다는 것에 의의가 있다. 그러나 측정모형의 판별을 위한 지표변수의 최소 개수인 두 개의 문항묶음에 의한 효과만을 확인하였다는 한계가 있

다. 실제로 구인 측정을 위한 두 개의 지표변수는 여러 추정의 문제를 일으킨다는 것이 알려진 만큼(Kenny, 1979; Kline, 2011), 연구결과를 일반화하기는 어려울 것이다. 특히 이 논문에서는 다차원적 구인에 대한 묶음화를 다루고 있지 않은데, 다차원적 구인을 적절치 않은 방식으로 묶음화를 하는 경우 여러 통계적 및 해석적 문제가 발생할 수 있기 때문에 주의를 기울여야 하므로 중요한 내용이다.

우종필(2015)의 연구는 조현철과 강석후(2007)의 연구와 유사한 흐름을 따르지만, 다차원적 구인의 문항묶음 방법을 포함했다는 점에서 한 발 더 나아갔다고 할 수 있다. 구인의 차원성에 대한 고찰이 이전 논문보다 넓어졌으며, 차원성과 문항묶음 방법에 따라 모형 추정의 결과가 달라질 수 있다는 것을 확인하였고, 묶음화 이전 절차로 차원성 검정을 강조하고 있다. 그러나 문항묶음화 방법에 따라 추정의 정확성이 달라지는 것은 자료의 특성에 따를 수 있다는 쟁점이 있는데도 불구하고, 부정적인 영향만을 부각시키고 있어 독자의 혼란을 초래할 가능성이 있다. 또한 소개한 문항묶음 형성방법이 제한적이어서 상관 알고리즘(correlational algorithm)이나 방사 알고리즘(radial algorithm) 등과 같은 주요한 묶음화 방법을 포함하지 않고 있다는 약점도 존재한다.

두 논문은 공통적으로 문항묶음의 개념을 정리하고 실증적 자료를 통한 묶음화의 효과를 확인하는 국제 논문들과 그 흐름을 같이 하고 있다. 그러나 국제적으로 중요한 최근 리뷰논문인 Little 등(2013)의 논문을 참고하지 않았고, Little, Cunningham, Shahar와 Widaman (2002)과 Matsunaga(2008)의 연구에 크게 의존하여 문항묶음에 대한 통합적인 이해를 이끌

기에는 무리가 있다고 판단된다.

본 연구의 목적은 지난 20여 년간 다각도로 이루어진 문항묶음화에 대한 여러 방법론적 연구들을 통합하고, 이에 기반하여 일반 연구자들이 주어진 조건에서 어떤 방식으로 문항 묶음을 형성해 사용할지에 대한 적절하고 효과적인 방법을 제안하는 것이다. 문항묶음 사용은 빠른 속도로 증가하고 있는데 반해, 아직 국내에서 문항묶음의 전반적인 원리에 대하여 설명하고, 다양한 방법을 분류 및 정리하여, 연구자들에게 도움이 될 만한 절차를 제공함으로써 문항묶음에 대한 전반적인 이해를 높일 수 있는 방법론적 논문은 찾기 어렵다. 이에 본 연구는 문항묶음에 대한 측정학적 이해를 제공하고, 문항묶음 사용의 배경과 원리 및 쟁점에 대한 종합적인 논의를 하고자 한다. 이와 더불어, 이미 출판된 논문들에 존재하는 모호성(예, 선명하지 않은 문항묶음 알고리즘과 차원성 가정에 대한 논란 등)에 대하여도 구체적으로 논의하고자 한다.

본 연구에서는 가장 먼저, 주어진 연구 상황에서 문항묶음을 사용해야 하는가라는 기본적인 질문에 대한 답을 찾는 과정부터 시작한다. 이런 맥락에서 문항묶음 사용의 이유를 측정학적 관점과 구조방정식모형 관점으로 나누어 설명한다. 다음으로 문항묶음을 둘러싼 여러 쟁점들과 제안점, 즉 어떤 상황과 조건 하에서 어떻게 문항묶음을 사용하는 것이 적절한지를 종합적으로 논의한다. 다양한 조건에서 문항묶음 사용에 대한 지지와 반대가 모두 설득력 있게 맞서고 있으므로, 한두 개 논문의 주장만 받아들이면 편협한 결정을 내릴 가능성이 있다. 또한 문항묶음의 사용은 결국 연구자의 목적과 판단에 따르기 때문에 묶음화 원리에 대한 정확한 이해가 중요하다. 문

항묶음을 둘러싼 여러 쟁점들에 대한 이해를 바탕으로 문항묶음을 사용하기로 결정하였다면, 어떤 방식으로 묶음을 만들 것인가 또한 매우 중요하다. 실제로 어떤 방법이 가장 효과적인지를 확인하는 것이 여러 논문들의 연구주제일 정도로 문항묶음화 방법 선택은 핵심적인 사항이다. 본 연구에서는 선별적으로 몇 가지 주요한 묶음화 방법을 보인다. 마지막으로 연구자들이 실제로 문항묶음을 어떻게 적절하고 유용하게 사용할 수 있을 것인지 그 판단 기준과 사용 단계를 제안하고, 주의점에 대해서도 논의한다.

문항묶음의 사용 이유

개별문항을 사용하는 것에 비해 문항묶음을 사용하게 되면 여러 가지 얻을 수 있는 이점이 있지만, 동시에 주의해야 할 부분도 존재한다. 여기에서는 묶음화를 하는 이유와 그에 따른 영향을 크게 측정학적 관점과 구조방정식모형 관점에서 논의한다.

측정학적 관점

지표의 정규성과 연속성

문항묶음을 사용하는 이유 중 하나이기도 하면서 가장 대표적인 장점은 개별문항을 지표로 사용할 때 나타날 수 있는 측정학적 한계를 극복할 수 있다는 것이다. 구조방정식모형에 대한 표준 추정방법인 ML은 자료의 정규성과 연속성을 가정하는데, 이를 만족하지 못할 경우에 크게 두 가지 문제가 발생하는 것으로 알려져 있다(Curran, West, & Finch, 1996; Muthén & Kaplan, 1992). 첫째는 χ^2 검정

통계량을 과대추정하여 모형이 자료에 부합한다는 영가설을 과도하게 기각하는 것이고, 둘째는 각 개별추정치의 표준오차를 과소추정하여 모수검정의 영가설을 과도하게 기각하는 것이다. 연속형이면서 정규성을 약간 만족하지 못하는 경우에는 영향이 크지 않다는 연구도 있으나(Muthén & Kaplan, 1992), 비연속형 변수가 정규성까지 만족하지 못하는 경우는 심각한 문제가 발생할 수도 있다.

문항목록을 이용하게 되면 이와 같은 문제를 어느 정도 해결할 수 있다. 개별 점수를 더하거나 평균을 내어 사용하면, 범주형 척도이지만 연속형인 것처럼 다룰 수 있는 원리와 같다. 예를 들어, 사회과학에서 많이 쓰이는 5점 리커트형 척도(Likert-type scale)로 만들어진 문항은 일반적으로 1점 간격으로 1점부터 5점 사이의 범위를 지닌다. 평균으로 문항목록을 만든다면, 점수의 범위는 1점부터 5점으로 같지만, 얻을 수 있는 점수의 간격은 좁아진다. 따라서 하나의 문항목록은 하나의 개별문항보다 상대적으로 연속적인 점수로서의 특성을 더 나타낼 수 있게 된다. 또한 문항 하나의 응답 분포는 편향되기 쉬운 반면, 여러 개의 문항이 합쳐진 문항목록은 그 안에서 개별문항 서로가 분포의 비정규성을 상쇄할 수 있기 때문에 전체적으로 정규분포에 가까워지는 경향이 있다(Bandalos, 2002; Bandalos & Finney, 2001; Hau & Marsh, 2004; Matsunaga, 2008).

지표의 대표성

하나의 문항으로 어떤 구인의 특성을 대표하는 것은 한계가 있으므로, 보통 여러 문항으로 구인을 측정하여 지표들이 구인에 대한 대표성을 확보할 수 있도록 한다. 하나의 문항 점수보다 여러 개를 더하거나 평균을 내어

만든 점수가 더 신뢰롭고, 더 안정적인 점수일 수 있다(Cohen, Cohen, Teresi, Marchi, & Velez, 1990; Rushton, Brainerd, & Pressley, 1983). 많은 문항을 사용할수록 전체적으로 구인을 잘 대표하는 척도가 될 수 있겠지만, 하나의 구인을 측정하기 위해서 개별문항 전집(universe)을 모두 사용하는 것은 불가능하다. 게다가, 개별문항 각각이 측정오차(measurement error)를 가지게 되므로 문항이 많을수록 모형 내 전체적인 오차의 크기가 증가하게 된다. 하지만 그렇다고 해서 임의의 개별문항 몇 개를 무작위로 선택하여 지표로 사용하는 방식은 척도의 신뢰도에 좋지 않은 영향을 미칠 가능성이 크다.

그림 1을 통하여 구인과 그 구인을 대표하는 개별문항의 관계를 기하학적으로 살펴볼 수 있다. 그림 1에서 가운데 큰 점(●)은 해당 구인이 의미하는 개념의 중심점(centroid)이다. 작은 점(●)으로 표현된 개별문항들은 그 중심점을 기준으로 퍼져 있는데, 중심점에 가까이 있을수록 구인의 개념을 잘 대표하는 문항이라고 할 수 있다. 이때 중심점과 개별문항 간의 거리는 측정오차라고 볼 수 있다(Matsunaga, 2008). 그림 1에서 보듯이 많은 문항을 사용할수록 분명히 중심점의 다양한 측면을 더 잘 드러낼 수 있을 것이지만, 앞서 말한 대로 모형의 전체적인 측정오차가 증가한다. 측정오차가 커지면, 측정모형 지표의 오차 간 상관 혹은 오차와 요인 간 상관이 발생할 가능성이 증가하므로 모형적합도는 나빠진다(MacCallum, Widaman, Zhang, & Hong, 1999). 따라서 연구자는 많은 문항을 사용하여 구인을 더 정확하게 측정하는 것과 간명한 모형을 만들어 측정오차를 줄이는 것 사이에서 균형을 맞추어 측정모형을 설정해야 한다(Little, Lindenberger, &

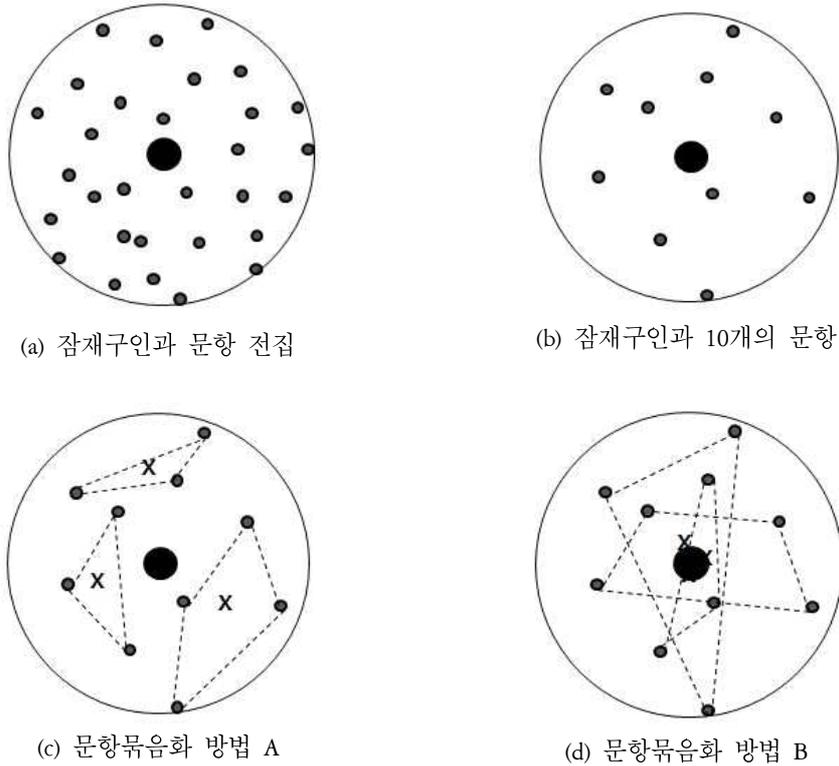


그림 1. 구인에 대한 지표의 대표성

잠재구인과 문항의 관계를 기하학적으로 나타낸다. 큰 점(●)은 해당 구인의 중심점, 작은 점(●)은 개별문항, X는 새로 만들어진 문항묶음의 중심점을 나타낸다.

Nesselroade, 1999).

이때 문항묶음을 사용하여 지표의 개수를 줄이면 이러한 문제를 어느 정도 해결할 수 있다. 각 문항묶음이 구인을 대표하는 정도는 하나의 개별문항보다 높아지는 동시에, 구인 측정을 위한 지표의 수는 감소하기 때문에 전체적인 측정오차는 줄어든다. 즉, 구인에 대한 전체 지표의 대표성과 정확성을 유지하면서 측정오차는 감소하게 되어, 개별문항을 사용할 때보다 좋은 측정학적 특성과 모형적합도를 모두 확보할 수 있게 된다. 그림 1(c)와 그림 1(d)는 모두 10개의 문항으로 세 개의 문항

묶음을 만드는 경우를 나타낸 것이고, X는 각 문항묶음의 중심점을 의미한다. 그림 1(c)와 그림 1(d)는 문항묶음 형성방법에 따라 다른 특성의 묶음이 만들어질 수 있다는 것을 보인다. 어떤 경우든 모두 개별문항을 사용했을 때와 비교하여 지표(X)가 구인의 중심점에 더 가까워지며 지표의 수도 적어져, 전체적인 측정오차가 감소한 것을 시각적으로 확인할 수 있다.

공통분산 증가, 고유분산 감소

문항묶음을 지표로 사용하면 개별문항을

사용할 때보다 상대적으로 지표의 공통분산(communality 또는 common variance)은 커지는 반면, 상대적으로 고유분산(오차분산, uniqueness 또는 unique variance)은 작아지게 되어 결과적으로 지표의 신뢰도와 모형적합도를 향상시킬 수 있다. 이러한 이점이 나타나는 원리는 진점수이론의 틀에서 분산과 신뢰도의 개념으로 이해할 수 있다. 식 (1)과 같이 진점수이론 모형에서는 관찰된 검사점수(x)가 진점수(τ)와 오차점수(e)의 합으로 결정된다.

$$x = \tau + e \quad (1)$$

문항의 특성을 더 잘 파악하기 위해 진점수를 식 (2)처럼 세분화할 수 있다.

$$\tau = \lambda\xi + s \quad (2)$$

먼저 진점수는 측정의 대상이 되는 구인의 특성(ξ)과 문항 고유의 특성(s)으로 이루어진다. 간명한 이해를 위해 요인부하 λ 를 1로 고정하면(Bollen, 1989), $\tau = \xi + s$ 가 된다. 이를 식 (1)에 대입하면 x 는 식 (3)과 같이 고쳐 쓸 수 있다.

$$x = \xi + s + e \quad (3)$$

즉, 문항의 관찰점수(x)는 구인의 특성(ξ), 문항 고유의 특성(s), 그리고 오차(e)로 이루어진다(Bollen, 1989; Little et al., 2002; Little et al., 2013; Matsunaga, 2008).²⁾ 여기서 $\delta = s + e$ ³⁾라

고 정의한다면, 최종적으로 식 (4)와 같이 관찰점수를 나타낼 수 있다.

$$x = \xi + \delta \quad (4)$$

ξ 는 구인의 특성을 반영하기 때문에 문항에 따라 변하지 않는 공통적인 것이며, δ 는 문항마다 다를 수 있다. 진점수이론에서는 모집단 수준에서 식 (4)의 각 요소들이 독립적이라고 가정하기 때문에, 문항묶음을 만들 경우 문항 특수적인 요소(δ)는 서로 상쇄되는 것이 일반적이다(Matsunaga, 2008). 이 가정에 기반하여 검사점수에 대한 분산 식을 쓰면 식 (5)와 같다.

$$Var(x) = Var(\xi) + Var(\delta) \quad (5)$$

여러 개의 문항이 있는 경우, 개념적으로 $Var(\xi)$ 는 공통분산을 의미하며, $Var(\delta)$ 는 고유분산을 의미한다.

개별문항이 아닌 문항묶음을 지표로 사용하게 되면, 지표의 개수는 줄어들고 묶음 내 문항들의 고유분산 부분은 상쇄되어, 묶음의 고유분산은 문항에 비해 작아지게 된다. 또한 개념적으로 하나의 문항묶음에는 여러 문항들의 속성이 포함되므로 문항묶음들 간 공통분산은 커진다. 즉, 각 지표(묶음)의 고유분산에 대한 공통분산의 비가 더 커진다. 같은 원리로 관찰분산에 대한 공통분산의 비율 ($\frac{Var(\xi)}{Var(x)}$)이라고 할 수 있는 지표의 신뢰도가

다.

²⁾ Little과 그 동료들 및 Matsunaga(2008)는 $x = t + s + e$ 로 단순화하여 표현하였는데, 이는 $t(=\tau)$ 와 ξ 를 구분하지 않았기에 혼란의 여지가 있어, 본 논문에서는 Bollen(1989)의 방식을 따른다.

³⁾ δ 는 ξ 를 측정할 때 x 의 무선오차를 의미한다. 문항 특수적 요소인 s 는 파악하기가 매우 어려우며 모호한 개념이기 때문에 보통 $Var(s) = 0$ 으로 가정하기도 한다(Bollen, 1989).

증가하고⁴⁾, 신뢰도가 증가하면 모형의 적합도 역시 좋아지는 경향이 있다.

구조방정식모형 관점

모형적합도

문항묶음을 이용하면 χ^2 값이나 근사 적합도지수(예, RMSEA, CFI 등)의 측면에서 모형적합도(model fit)가 좋아지게 된다(Bandalos, 2002; Bandalos & Finney, 2001; Landis, Beal, & Tesluk, 2000; Nasser & Wisenbaker, 2003). 앞서 설명한 대로 관찰분산에 대한 공통분산의 비율이 상대적으로 증가하기 때문이다. 이는 오차의 상관 및 오차 상관행렬의 요인부하량(unique factor loading) 등을 이용하여 좀 더 체계적으로 설명 가능하다(Bandalos, 2002; MacCallum et al., 1999; Matsunaga, 2008). 먼저 구조방정식 분야의 대표적인 모형이라고 할 수 있는 요인 모형은 표본수준에서 식 (6)과 같이 표현된다.

$$C_{yy} = AC_{\alpha\alpha}A' + AC_{cu}\theta' + \theta C_{uc}A' + \theta C_{uu}\theta' \quad (6)$$

위에서 C_{yy} 는 관찰변수 y 의 공분산 행렬, $C_{\alpha\alpha}$ 는 공통요인, 즉 구인 c 의 공분산 행렬, C_{cu} 와 C_{uc} 는 구인과 오차요인의 공분산 행렬, C_{uu} 는 오차요인 간의 공분산 행렬이다. A 와 θ 는 부하량, 즉 행렬에 대한 가중치를 의미하는데, 이 중 특히 적합도에 직접적인 영향을 미치는 가중치는 θ 로서, 이는 측정오차의 분산 행렬이다.

⁴⁾ 진점수이론의 신뢰도는

$$\frac{Var(\tau)}{Var(x)} = \frac{Var(\xi) + Var(s)}{Var(x)} \rightarrow \frac{Var(\xi)}{Var(x)} \text{로 표현될 수 있다.}$$

모형적합도는 기본적으로 C_{yy} 와 모집단에서의 y 공분산행렬, 즉 Σ_{yy} 의 차이에 의해 결정된다고 할 수 있기 때문에 표본오차를 반영하는 식 (6)에서 $AC_{cu}\theta' + \theta C_{uc}A' + \theta C_{uu}\theta'$ 부분이 작아질수록 적합도는 상승한다. 다시 말해, 구조방정식 모형에서 한 지표의 오차가 다른 지표의 오차와 상관이 없고, 다른 요인들과도 상관이 없게 되면 이 부분이 작아지게 되어 모형적합도가 향상된다. 개별문항을 이용하는 대신에 각 문항묶음의 특성을 비슷하게 만들면, 전체분산에서 오차분산의 비율이 줄어들게 된다. 결국 오차 간, 또는 오차와 다른 요인들 간의 공분산 행렬(C_{uu} , C_{cu} , C_{uc})의 크기를 감소시킨다. 또한 측정오차(θ)를 반영하는 고유분산의 크기도 문항묶음을 만들으로써 감소한다(Little et al., 2013). 상대적으로 높은 공통분산과 낮은 고유분산은 C_{yy} 가 다른 오차가 아닌 공통요인에 의해 더 많은 영향을 받는다는 의미이고, 이러한 원리로 문항묶음을 사용하면 개별문항을 사용할 때보다 모형적합도가 향상될 수 있다.

이렇듯 문항묶음의 사용은 모형적합도를 향상시키는 것이 일반적인데, 결론과 해석에 있어서 매우 주의해야 할 부분이 있다. 적합도가 향상되었다는 것이 연구자의 모형이 자료를 완전하게 반영하는 진정한 모형(true model)에 가까워졌다는 것을 항상 의미하지는 않는다는 것이다. 보통 개별문항을 이용하여 모형을 설정하면, 문항이 실제로는 여러 구인에 이중으로 부하되어 있거나 설정한 모형이 존재하는 오차 간 상관 등을 제대로 반영하지 못할 때, 자연스럽게 낮은 모형적합도가 산출됨으로써 그 문제를 암시하게 된다. 하지만 문항묶음을 이용하게 되면 개별문항의 특이성이 완화되고 상쇄되어 문제가 되는 부분이 가

려질 수 있고, 그 결과 모형의 설정 오류가 존재함에도 불구하고 겉으로 드러나는 적합도는 매우 좋을 수 있다(Bandalos, 2002, 2008; Bandalos & Finney, 2001; Little et al., 2013).

문항묶음을 사용하면 많은 경우에 모형적합도가 상승하지만, 결과를 제대로 해석하기 위해서는 개별문항에 대한 이해가 선행되어야 한다. 따라서 문항묶음 사용을 결정하기 전에 연구자는 문항 수준의 분석을 하고 선행 연구들을 확인하여 개별문항의 특성을 파악해야 한다. 문항묶음으로 인해 개별문항의 중요한 실질적 의미나 특성이 올바르게 드러나지 못하게 된다면, 좋은 모형적합도란 것은 단지 공허한 메아리일 뿐이다.

추정의 안정성

만약 지표로 사용하는 관찰변수가 바람직하지 않은 측정학적 특성(예, 큰 측정오차)을 보이거나 모형 내 설정되지 않은 교차부하(cross loadings)가 존재하는 경우, 모형을 분석하는 과정에서 여러 문제가 나타날 수 있다. 예를 들어, 추정치의 표준오차가 커지고, 추정과정에서 수렴하기 위해 더 많은 시간과 반복 계산 횟수를 보이며, 작은 차이(예, 다른 표본자료 사용, 시작값의 차이, 다른 추정방법 이용 등)에도 모수 추정치의 변화폭이 커지는 등이다. 이런 경우에는 같은 모형으로 분석을 할 때마다 다른 결론을 도출할 수 있기 때문에 결과를 일반화하는 데 한계를 지니게 된다. 또한 추정된 값들을 신뢰할 수 없고, 모형의 적합도 역시 좋지 않은 경우가 많다.

추정과정에서 발생하는 이런 문제들을 제어하고, 안정적인 추정치를 얻기 위해 연구자가 할 수 있는 일들 중 하나는 모형을 간명하게 만드는 것이다. 구조방정식모형을 간단하게

만드는 여러 방법 중 하나는 사례수를 늘리거나 추정하고자 하는 모수의 개수를 줄이는 방법이 일반적이다. 즉, 모수의 개수에 대한 사례수의 비율을 증가시키는 것이라고 할 수 있다. 이런 맥락에서 문항묶음을 만들면 개별문항을 사용할 때보다 모형에서 추정해야 하는 모수(요인부하, 측정오차의 분산 등)의 수가 적어지고, 더욱 간명한 모형을 만들 수 있다. 결국 문항묶음을 사용하면 더 안정적인 추정 결과를 얻을 수 있게 되는 것이다(Bagozzi & Edwards, 1998; Bandalos & Finney, 2001; MacCallum et al., 1999; Marsh, Hau, Bella, & Grayson, 1998; Matsunaga, 2008). 그리고 이러한 문항묶음의 효과는 표본 크기가 작은 경우처럼 여러 추정 조건이 이상적이지 않을 때 더욱 강력하게 나타나게 된다(Bandalos, 2002; MacCallum et al., 1999).

문항묶음에 대한 쟁점

연구자는 자신이 처한 상황과 조건에 따라 문항묶음의 사용여부를 결정해야 한다. 이때 개별문항을 사용해야 하는지, 문항묶음을 사용해야 하는지, 또한 문항묶음을 사용한다면 어떤 방법을 적용해야 하는지 등이 모두 합리적으로 결정되어야 한다. 이러한 관점에서 문항묶음을 둘러싼 여러 쟁점에 대해 인지하고 이해하는 것은 매우 중요하다.

연구자의 목적

연구자의 연구 주제와 목적은 문항묶음 사용을 결정하는 데 있어 중요한 기준이다. 만약 연구의 목적이 검사 개발이나 타당화와 같

이 개별문항의 특성과 수행정도를 파악해야 하는 것이라면, 문항묶음의 사용은 바람직하지 않다. 이때는 각 문항이 해당 구인을 잘 반영하고 있는지, 문항이 다른 구인에 이중으로 부과된 것은 아닌지, 문항들의 공통분산과 고유분산이 상대적으로 어떤 크기인지 등 개별문항과 관련된 정보가 중요하다. 이와 같은 연구에서 문항묶음을 사용하면, 각 문항의 고유한 특성이 사라지게 되고 검사 문항에 따른 척도의 신뢰도와 타당도는 파악하기 어렵게 된다. 따라서 이러한 연구를 수행한다면 문항묶음 사용을 권하지 않는다. 반면에 연구자가 구인 간의 구조적 관계, 즉 구조방정식 모형을 검정하고 추정하는 데 관심이 있다면, 문항묶음 수준에서만 모형이 성립되어도 큰 문제는 없다(Little et al., 2013; Matsunaga, 2008).

차원성

차원성(dimensionality)은 척도가 측정하는 구인이 어떤 구조로 이루어져 있는지를 의미한다. 일차원성(unidimensionality)이란 문항 또는 지표변수들이 동일한 구인을 나타낸다(indicate)는 것을 의미하며, 다차원성(multidimensionality)이란 문항(또는 지표변수)들이 하나 이상의 구인을 다루고 있음을 의미한다. 여러 연구자들이 문항묶음을 사용하기 위한 전제조건으로서 구인의 일차원성 가정을 언급한다(Bandalos, 2002; Bandalos & Finney, 2001; Kim, 2000; Marsh et al., 2013). 구인이 정확하게 일차원적이라면 어떤 방법으로 문항묶음을 만들어도 추정 결과에 차이가 없을 것이다. 그러나 다차원성을 나타내는 문항들을 임의의 방법으로 묶을 경우, 모형적합도는 향상될 수도 있지만 구인의 실제 구조를 정확히 반영하지 못하는

문제가 발생할 수도 있다(Bandalos, 2008, Marsh et al., 2013).

이후 문항묶음을 만드는 방법에 대해 논의할 때 더 자세히 다루겠지만, 다차원적 구인을 ‘고유분산 분배전략(unicqueness distributed strategy)’을 적용하여 각 문항묶음이 서로 비슷한 특성을 지니도록 묶음화하면, 구인의 구조가 명확히 드러나지 못하고 실제보다 모형적합도 지수가 과대추정되어 산출되는 문제가 발생할 수 있다. 그렇다고 ‘고유분산 고립전략(unicqueness isolated strategy)’을 적용하여 비슷한 문항끼리 묶는다고 해도, 구인 구조를 정확히 파악하고 있어야 그에 맞게 묶을 수 있기 때문에 간단한 문제가 아니다. 따라서 문항묶음을 사용하기 전에 개별문항들이 실제로 어떤 구조를 형성하고 있는지를 파악해야 한다.

이처럼 문항묶음화 과정에서 차원성 가정을 중요하게 여겼기 때문에 차원성을 확인하는 방법도 여러 가지가 제안되었는데, 크게 이론적 관점과 경험적 관점으로 나눌 수 있다. 먼저 이론적인 관점에서 당연하지만 중요한 방법은 요인 구조가 이미 잘 알려진 검증된 척도를 사용하느냐는 것으로서, 자료를 수집하기 이전에 고려해야 할 사항이다(Bandalos, 2002; Matsunaga, 2008). 좋은 척도를 사용하는 것은 문항묶음을 사용하기 위한 조건이며, 연구 전체의 신뢰성과 타당성을 위해서도 중요하다.

경험적인 관점에서 일차원성을 확인하는 방법은 탐색적 요인분석(EFA) 또는 확인적 요인분석(CFA)을 실시하는 것이다. 이론적으로 구조가 정해지지 않았다면, EFA를 통해 차원성을 설정하는 근거를 얻을 수 있다(Hall, Snell, & Foust, 1999). 모든 문항들이 하나의 요인에 대해서만 상대적으로 높은 부하량을 보이면

일차원성이 있다고 판단하는 것이다. 그러나 실제로 일차원이더라도 EFA에서 명확한 일 요인을 나타내기는 어려우며, 다차원이어도 분석 결과 일요인 구조가 적절하다는 결과가 나올 수도 있으므로 이 방법에는 한계가 있다(Gerbing & Anderson, 1988; Hattie, 1985; Matsunaga, 2008). CFA를 이용하는 방법은 일요인 모형을 추정하여 모형적합도와 수정지수를 확인하는 것이다. 적합도가 양호하고, 수정지수들이 크지 않으면 구인이 일차원성을 만족한다고 결론을 내릴 수 있다. 하지만 적합도라는 것이 차원성 이외에도 표집오차나 표본 크기의 영향을 받기 때문에, 이 역시 완전한 차원성 확인방법은 아니라는 한계점이 존재한다.

위에서 언급한 그 어떤 한 가지 방법만으로는 차원성을 확인할 수 없으며, 사실 100% 일차원성을 확인할 수 있는 그 어떤 방법도 없다. 그럼에도 불구하고, 문항묶음을 이용하기 전에 문항 및 척도에 대한 충분한 이론적 탐구와 경험적 자료분석을 통해 척도의 차원성에 대한 근거를 확보하는 것은 중요하다(Kim & Hagtvet, 2003; Little et al., 2002). 하지만 일차원성은 상황에 따라 판단하기가 상당히 모호할 수 있다. 예를 들어, Spearman(1927)은 인간의 지능이 일반지능(general intelligence)으로 대표된다고 본 반면, Thurstone이나 Gardner 등은 인간의 지능이 독립적인 몇 개의 특수지능(specific intelligence)으로 구성되어 있다고 보았다(Gardner & Hatch, 1989; Guilford, 1972). 이처럼 같은 구인에 대해서도 연구자마다 다른 차원성을 설정할 수 있으며, 다른 지표를 사용하여 다른 결론을 내릴 수도 있다. 이와 같은 이유로 Little 등(2002)은 구인의 일차원성이 어느 정도 자의적 판단에 따른다고 하였다.

만약 척도를 구성하는 문항들이 다차원을 형성하고 있다고 가정한다면 문항묶음은 사용할 수 없는 것인가? 이에 대해 Little과 그의 동료들(Little et al., 2002; Little et al., 2013)은 느슨한 일차원성 가정 또는 다차원성을 가정하는 문항묶음을 제안한다. 다차원적 구인을 다룰 때, 알맞은 방법으로 문항묶음을 만들면 구인의 특성을 잘 반영하면서도 적절한 모형을 설정할 수 있다는 것이다. 이는 다차원성을 가정하는 문항묶음 방법 자체에 대한 새로운 제안을 한 것은 아니지만, 문항묶음 사용의 폭을 넓혔다는 점에서 의의가 있다고 할 수 있다.

다집단 분석

구조방정식 모형을 이용한 다집단 분석을 통해 서로 다른 집단 간 구조관계를 비교하는 것이 광범위하게 사용되고 있다. 또는 측정도구의 개발 과정에서, 문항들이 서로 다른 집단에서 동일한 기능을 하는지를 파악하기 위한 목적으로 다집단 분석이 이용되기도 한다. 이런 경우, 일반적으로 집단 간 측정불변성(measurement invariance)이 먼저 확보되어야 한다. 이는 구조관계를 확인하기 이전에, 구조관계의 대상이 되는 잠재변수들이 상응하는 지표에 의해 집단 간 동일한 의미로 정의되는지 확인하는 절차이다. 일반적으로 형태동일성(configural invariance) 모형을 첫 단계로 검정하는데, 요인과 지표변수의 관계가 집단 간 동일하고 모수 추정치는 다를 수 있다고 가정한다. 만약 형태동일성 모형의 적합도가 양호하면, 연구자가 원하는 다음 단계의 더 강한 측정동일성을 확인할 수 있다.

이처럼 형태동일성 모형의 적합도가 상당히

중요하기 때문에, 이를 높이기 위한 목적으로 문항묶음을 사용하는 경우가 있다. 하지만 측정불변성 확인과정에서의 문항묶음은 각 개별 문항의 집단 간 차이를 확인하지 못하게 할 수 있다. 즉, 문항 수준에서 실제로 존재하는 집단 간 차이를 문항묶음의 수준에서는 파악할 수 없게 되는 것이다(Marsh et al., 2013; Meade & Kroustalis, 2006). 이러한 문제는 척도 개발 시 검사의 타당성 등에 매우 심각한 영향을 끼칠 수 있다. Meade와 Kroustalis(2006)는 개별문항을 사용해야 집단 간 차이를 보이는 문항끼리의 관계 및 문항과 구인과의 관계에서 많은 정보를 얻을 수 있기 때문에 개별문항에 기반한 다집단 분석을 강조한다.

반면, Little 등(2013)은 구조모형 맥락에서 구인 간 관계의 집단차이를 확인하고자 할 때는 문항묶음화가 정당화될 수 있다고 주장한다. 구조모형 맥락에서는 측정오차를 줄이고 구인에 대한 간명함 모형을 설정하는 것이 목적이기 때문이다. 이러한 주장은 구조관계의 집단 간 차이를 분석하는 연구에서는 굳이 모든 문항의 완전한 측정동일성이 필요하지 않다는 여러 연구(Byrne, Shavelson, & Muthén, 1989; Millsap & Kwok, 2004; Steenkamp & Baumgartner, 1998)와 이론적으로 맞닿아 있다. 집단 간 측정불변성 확인 시 개별문항의 사용을 강조한 Meade와 Kroustalis(2006)도, 연구자가 구조모형 부분에 더 관심이 있다면 측정학적 이점이 있는 문항묶음을 사용할 수도 있다고 논의하였다. 이때 중요한 것은, 집단 간 동일한 방식으로 문항묶음을 만들어야 하고 집단 간 동일성을 의미 있게 해석하기 위해 내용기반 접근법(content-based approach)을 이용해야 한다는 것이다. 마지막으로 주의할 점은, 문항 묶음을 지표로 사용했을 때 측정불변성은 문

항 수준이 아닌 묶음 수준에서 이루어진 것이라고 명확히 해석해야 한다. 결론적으로, 측정 불변성을 검정할 때 문항묶음을 사용할 지에 대한 결정권은 연구자에게 있으며, 논리적으로 이를 뒷받침할 수 있다면 사용이 가능하다.

추정의 정확성

묶음화에 따른 모수 추정의 정확성 혹은 편향의 정도는 연구마다 상당히 다른 결과를 보인다. Bandalos(2002)는 시뮬레이션 연구를 통해 개별문항보다 문항묶음을 사용할 때 경로계수 추정 편향의 정도가 더 작으며, 하나의 묶음에 포함된 문항 수가 많을수록 더 정확한 추정치를 산출한다는 것을 확인했다. 그러나 몇몇 다른 시뮬레이션 연구는 문항묶음을 사용한 모형의 모수 추정치가 과소 혹은 과대추정될 수 있다는 것을 보였다(Hall et al., 1999; Stephenson & Holbert, 2003).

이렇게 상반된 결과는 혼란스러울 수 있다. 여러 연구의 시뮬레이션을 자세히 확인해본 결과, 자료 생성의 조건(예, 자료 분포의 형태, 비정규성의 정도, 표본크기, 요인 간 상관 등)이 상당히 달랐고, 그에 따라 결과도 달랐다. 예를 들어, 생성된 개별문항의 측정학적 특성이 이상적인 상황에서의 시뮬레이션 결과를 보면, 문항묶음화가 추정의 정확성에 대해 거의 효과를 미치지 않거나, 심지어 추정 결과를 왜곡하기도 하였다(Hall et al., 1999; Marsh et al., 1998; Nasser & Wisenbaker, 2006; Stephenson & Holbert, 2003). 반면에 이상적이지 않은 특성을 가진 문항들을 이용한 연구들을 보면 문항묶음화가 추정의 정확성을 증가시키는 방향으로 작용하였다. 즉, 개별문항이 여러 측정학적 성질을 만족하지 못할 때 문항묶음

을 사용하면, 측정학적 성질의 교정효과를 크게 얻을 수 있다(Little et al., 2002; Matsunaga, 2008).

추정의 정확성 측면에서 문항묶음을 이용해야 하는가 또는 개별문항을 이용해야 하는가라는 논쟁에 대한 답은 결국 연구자들이 가지고 있는 자료의 특성에 기반한다. 그러므로 문항묶음을 고려하는 연구자들은 문항 수준에서 그 측정학적 성질이나 분포의 정규성 등을 미리 확인해 보는 것이 필요하다. 그런데 아마도 현실 속에서 연구자가 수집한 자료가 매우 이상적인 특성을 지니기는 어려울 것이다. 이런 관점에서 보면, 위에서 보인 추정의 정확성에 대한 상반된 연구결과는 오히려 문항묶음의 사용 이유가 될 수도 있다. 연구자가 문항묶음을 고려한다는 것이 이미 개별문항이 잘 작동하고 있지 않은 경우가 대부분이기 때문이다.

범주형 변수의 문항묶음

앞서 언급했듯이 문항묶음을 사용하면 지표의 정규성과 연속성을 개별문항에 비해 더 잘 만족시키게 되므로 리커트형 척도를 이용해서도 선형 요인분석을 실시할 수 있다. 심지어 이분형 문항의 경우에도 묶음들이 서로 비슷한 평균과 분산을 지니도록 평행(parallel) 묶음을 만들어서 선형 요인분석을 실시하기도 한다. 이때 개별문항들은 일차원성 가정을 만족해야 하며, 문항 난이도를 기준으로 각 묶음에 동일하게 분배하는 것이 일반적이다(Cook, Dorans, & Eignor, 1988).

그렇다면 편의성의 측면에서 이분형이나 순위형 문항을 묶음으로 만들고 선형 요인분석을 실시하는 것은 아무런 문제가 없는 것일

까? Bandalos(2008)의 시뮬레이션 연구를 보면, 문항들이 일차원적이지 않을 때 문항묶음의 사용이 경로계수에 심각한 편향을 야기했다고 밝혔다. 특히 이분형 문항인 경우에 묶음화에 의해 문항 고유의 특성이 왜곡되는 결과를 얻을 수도 있다. 따라서 범주형 문항을 분석할 때는 문항묶음보다 가중최소제곱(weighted least squares, WLS) 등의 제한정보 추정방법을 이용한 범주형 요인분석을 실시하는 것이 더 적절할 수도 있다(Bandalos, 2008; Marsh et al., 2013). 또는 일차원성을 만족하는 범주형 문항들은 문항반응이론(item response theory, IRT)의 전통에서 2모수 로지스틱 모형이나 Samejima (1970)의 등급반응모형(graded response model, GRM) 등을 완전정보 추정방법을 이용하여 분석할 수 있다. 언급한 IRT 모형들은 구조방정식 전통에서 일반화 요인분석 모형과 같으므로 다른 모수화(parameterization)를 통하여 추정하는 것도 가능하다.

정리하여 말하자면, 경우에 따라 범주형 자료를 분석할 때 문항묶음을 만들어 분석하는 것이 유일한 방법이거나 가장 좋은 방법은 아닐 수 있다는 것을 기억할 필요가 있다. 자료의 특성에 맞는 추정방법을 사용하거나 다른 모형을 사용하여 분석할 수도 있으므로, 문항묶음을 사용할지 그렇지 않을지는 연구의 방향과 연구자의 능력에 따라 결정하면 될 것이다.

문항묶음 형성방법

문항묶음의 개수

여러 쟁점들을 인지하고도 문항묶음을 사용

하는 것이 적절하다고 판단했다면, 그 다음으로 요인 당 몇 개의 문항묶음(즉, 지표변수)을 생성할지 결정해야 한다. 이 질문은 CFA에서 지표의 개수가 몇 개인 것이 가장 효율적인가라는 질문과 같은 맥락이다. 이는 문항묶음도 측정모형에서 사용되는 지표이기 때문이다. 요인이 두 개 이상 있을 때, 모형판별을 위한 최소 지표변수의 개수는 두 개지만, 추정과정에서 오류를 발생시키기 쉬운 것으로 알려져 있다(Kline, 2011). 반면 지표변수의 개수가 너무 많으면, 측정오차가 커지거나 모수 간 거짓 상관(spurious correlation)이 발생하기 쉬우며, 결과적으로 모형적합도가 나빠지고 추정의 안정성을 해칠 수 있다. 또한 여러 연구들은 문항묶음의 개수가 적을수록 모형적합도는 높았지만, 동시에 추정치는 더 편향되는 것을 확인하였다(Bandalos, 2002; Rogers & Schmitt, 2004).

이와 같은 여러 가지 상황을 고려하여 Little 등(2013)과 Matsunaga(2008)는 요인 당 세 개의 문항묶음을 사용하는 것을 추천하였다. 요인 당 세 개의 지표변수를 사용하게 되면, 측정모형에서 각 잠재변수의 일요인 모형 부분이 완전판별(just-identified)이 된다. 물론 이때 요인이 여러 개라면 전체 구조방정식 모형은 과대판별(over-identified)이 되지만, 측정모형의 자유도는 최소화 되어 0에 가깝게 된다. 이렇게 되면 구조방정식 모형에서 산출된 모형적합도를 측정모형의 적합도 부분과 분리할 수 있는 장점이 생긴다. 이는 Anderson과 Gerbing(1988)이 제안한 것처럼 구조방정식 모형의 2단계 접근법을 실행해야 하는 이유와 상통한다.

사실 적절한 문항묶음의 개수는 연구모형, 표본크기, 자료의 분포 특성 등에 따라 다르기 때문에 이를 무조건 세 개로 단정 짓기는 어렵다. 만약 묶음을 형성하기 위한 개별문항

의 개수가 충분한 상황이라면, 일반적으로 요인 당 세 개 또는 네 개의 문항묶음이 추천된다. 측정모형의 간명성과 모형적합도 및 추정치의 정확성 등을 모두 고려했을 때 이것이 적절한 개수일 것이다. 만약 개별문항의 개수가 충분하지 못한 상황이라면, 요인 당 두 개의 문항묶음도 가능하다. 앞서 설명한대로 만약 세 개의 문항묶음을 만든다면 적합도의 해석에 있어서 이점을 얻을 수 있겠지만, 최종 결정은 연구자의 판단이다.

문항묶음 형성방법

몇 개의 문항묶음을 만들 것인지 결정하였다면, 이제 어떤 방법을 적용할지 결정해야 한다. 이때 연구자들은 자신의 연구에서 사용한 문항묶음 방법에 대해 정직하고 명확하게 기술해야 한다(Little et al., 2013). 모형적합도가 가장 좋은 모형을 찾거나 유의한 경로계수를 얻기 위해 같은 자료에 여러 문항묶음 방법을 적용해 보는 것은 옳지 않다. 즉, 미리 논리적인 절차에 따라 어떤 방식을 통해 문항묶음을 만들 것인지 결정해야 한다. 이런 이유로 각 방법이 어떤 목적과 원리를 따르는지 이해해야 할 것이다. 사실 모든 개별문항들이 완전하게 일차원을 형성하고 있다면 어떤 묶음방법을 사용해도 비슷한 결과가 나올 것이라는 것을 앞에서도 언급하였다. 그러나 문항 수가 많아질수록 그 문항들이 순수하게 하나의 구인만을 나타내기는 어려우며, 연구자도 파악하지 못한 숨겨진 구인이 있을 수도 있다. 이런 경우, 묶음을 만드는 방법에 따라 모형적합도나 모수추정 결과, 해석 등이 영향을 받을 수 있으므로 적합한 전략을 선택해야 한다(Landis et al., 2000; Rogers & Schmitt, 2004).

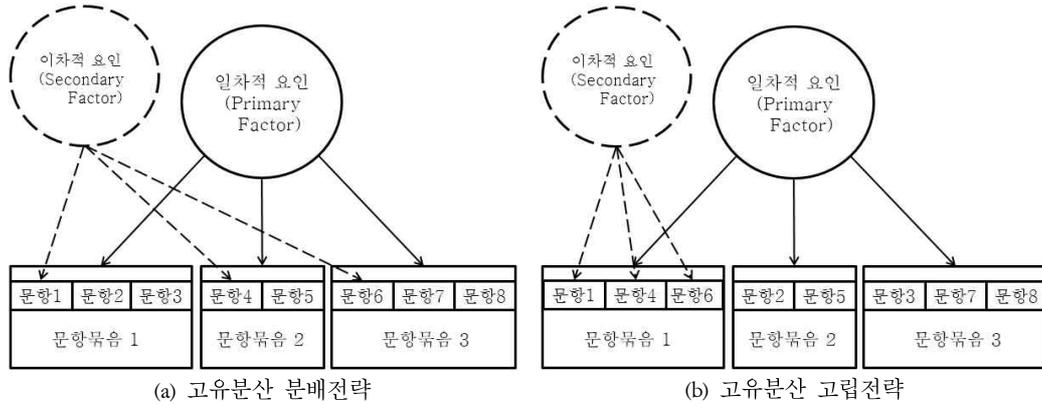


그림 2. 고유분산 분배전략 및 고립전략

묶음화 전략은 크게 고유분산 분배전략 (uniqueness distributed strategy)과 고유분산 고립 전략(uniqueness isolated strategy)으로 나눌 수 있다(Bandalos, 2002; Bandalos & Finney, 2001; Hagtvet & Nasser, 2004; Hall et al.,1999)⁵⁾. 이 두 전략은 개별문항의 고유분산을 다루는 방식에 차이를 보인다. 달리 말하면, 모형에 포함되지 않은 이차적 요인(secondary factor)의 영향을 처리하는 방식이 다르다. 이차적 요인이란 그림 2에 점선으로 표현된 것처럼 모형에 포함되지는 않았지만 몇몇 개별문항이 특성을 공유하고 있는 구인이다. 이차적 요인은 이론적으로 중요하지 않으므로 연구자가 의도적으로 포함하지 않았거나, 중요함에도 불구하고 파악하지 못한 것일 수 있다. 이는 종종 문항의 표현 방식이나 단어 선택, 부정문 및 긍정

문의 사용 등과 같은 방법효과에 의해 생길 수도 있다.

고유분산 분배전략은 그림 2(a)처럼 이차적 요인을 형성하는 문항들을 여러 묶음으로 분산시켜서 문항들의 고유분산을 각 묶음에 고르게 분배하는 전략이다. 이렇게 하면 문항묶음들은 비슷한 특성을 지니게 되고, 이차적 요인의 영향이 각 묶음에 고르게 분배되어 전체 모형에서 지표의 특성이 잘 드러나지 않게 된다. 이 전략은 결국 상대적으로 문항묶음의 고유분산은 작게 만들고, 공통분산은 크게 만드는 효과가 있다. 이에 반해 그림 2(b)의 고유분산 고립전략은 하나의 이차적 요인에 의해 영향 받는 문항들, 즉 비슷한 특성을 지닌 문항들을 하나의 묶음으로 만드는 것이다. 이렇게 만들어진 묶음들은 분배전략에 의하여 만들어진 묶음들에 비하여 각 특성들이 더욱 두드러지고, 상대적으로 큰 고유분산의 비율을 갖게 되며, 해당 묶음이 이차적 요인의 특성을 드러낼 수 있다.

실제 자료와 시뮬레이션 자료를 통해 두 전략의 차이를 확인한 많은 연구들이 있다. 고유분산 분배전략을 사용하면, 이차적 요인이

5) 문항묶음 방법을 분류하는 방식은 학자들마다 조금씩 다르게 명명하고 있다. 그러나 그 원리는 동일하다. 고유분산 분배전략은 영역-대표 접근법(domain-representative approach), 분배-대표 묶음화(distributed-representative parceling)와 같은 의미이다. 고유분산 고립전략은 국면-대표 접근법(facet-representative approach), 동질적 묶음화(homogeneous parceling)와 같은 의미이다.

나 오차의 영향이 감소하여 모형적합도가 향상되는 경향이 있다. 그러나 모수의 추정은 더 편향되는 경향이 있고, 모형이 잘못 설정되었을 경우에도 적합도가 높을 수 있는 문제를 보이기도 한다. 이에 반해, 고유분산 고립 전략을 사용하면 모형적합도는 낮은 경향이 있지만, 모수를 더 정확히 추정하고 모형 설정이 잘못된 경우에 낮은 적합도를 보임으로써 문제점을 암시하는 장점을 보인다(Hall et al., 1999; Landis et al., 2000; Rogers & Schmitt, 2004).

연구자는 설정된 모형과 구인에 대한 자신의 관점과 논리에 따라 문항묶음 방법을 선택하면 된다. 이때 가장 중요하게 고려해야 할 것은 구인의 차원성과 이차적 요인의 실재유무로, 특히 자료가 다차원성을 보이는 경우에는 고유분산 고립전략을 채택하는 것이 적절하다(Hall et al., 1999; Landis et al., 2000; Little et al., 2013; Rogers & Schmitt, 2004). 물론 분배 전략을 사용하면 높은 모형적합도를 얻을 수 있지만, 정확한 모형을 설정하고 검증하는 것이 구조방정식 모형을 사용하는 목적이기 때문에 상황에 따라 낮은 추정 편향 정도를 보이고 다차원적 구인의 특성을 잘 드러내는 고유분산 고립방법이 더 적절할 것이다.

이제 문항묶음화의 두 가지 전략에 따른 문항묶음 알고리즘을 소개한다. 고유분산 분배 전략에 속하는 무선 및 요인 알고리즘과, 고유분산 고립전략에 속하는 상관 알고리즘 및 방사 알고리즘과 내용기반 접근법, 이렇게 다섯 가지가 주요한 문항묶음화 방법이다. 이는 여러 출판된 논문들에서 다루고 있는 것들과 저자들이 실질적으로 중요하다고 판단한 알고리즘을 종합하여 선정한 것이다(Matsunaga, 2008; Little et al., 2013; Rogers & Schmitt,

2004). 요인 알고리즘, 상관 알고리즘, 그리고 방사 알고리즘을 적용하여 실제로 문항묶음을 만드는 방법은 부록에 예를 들어 제시하였다. 참고로 하나의 전략을 따르는 여러 알고리즘은 비슷한 작용을 하므로 같은 전략 내에서는 연구자가 임의로 선택하여도 큰 문제는 없다.

무선 알고리즘

무선 알고리즘(random algorithm)은 문항들 각각 문항묶음에 무작위로 배정하는 것으로, 가장 간단한 방법이다. 이때 문항들은 하나의 구인을 측정하는 동일한 문항 전집에서 표집된 문항 세트이며 서로 교환 가능한 비슷한 특성을 지닌다고 가정한다. 이렇게 완전한 일차원성을 만족한다는 가정 하에 임의로 만들어진 문항묶음들은 서로 비슷한 특성을 지니게 되고, 대략적으로 동일한 공통분산 및 오차분산을 지닐 것으로 기대된다(Little et al., 2013; Matsunaga, 2008). 그러나 이런 엄격한 일차원성 가정을 만족시키지 못한다면, 오히려 서로 특성이 다른 문항묶음이 형성되는 경향이 있다. 일반적으로 완전한 일차원성 가정은 현실에서 만족하기 매우 어려우므로 무선 알고리즘의 사용은 정당화되기 어렵다(Hall et al., 1999; Little et al., 2013).

요인 알고리즘

요인 알고리즘(factorial algorithm) 또는 단일 요인(single-factor) 방법(Landis et al., 2000)은 문항의 요인부하 크기를 기준으로 문항을 묶는 것이다. 요인 알고리즘의 구체적인 실행 방법은 여러 연구에서 조금씩 차이를 보이기도 하지만 그 원리는 크게 다르지 않다(Landis et al., 2000; Little et al., 2013; Matsunaga, 2008; Rogers & Schmitt, 2004). 이 알고리즘은 일요인 확인

적(또는 탐색적) 요인분석을 실시하고, 요인부하⁶⁾가 큰 문항들과 작은 문항들이 각 묶음에 배정되는 방식이다. 요인부하가 큰 문항은 상대적으로 오차분산이 작고, 요인부하가 작은 문항은 오차분산이 클 것이기 때문에 이 방법은 문항 특수적 요소인 오차를 묶음 간에 균등하게 나누려는 방법이다. 요인 알고리즘은 연구자가 정교하게 고안된 방법을 통하여 비슷한 묶음을 만들 수 있도록 하므로 무선 알고리즘에 비해 고유분산을 분배하는 목적을 더 정확히 달성할 수 있다.

상관 알고리즘

상관 알고리즘(correlational algorithm)은 상관 계수를 이용하는 단계적인 과정이다. 첫 단계에서 모든 문항 간 이변량 상관계수를 계산하여, 가장 높은 상관을 보이는 문항 쌍을 첫 번째 묶음에 배정하고, 두 번째로 높은 상관을 보이는 쌍을 두 번째 묶음에 놓는다. 이런 순서대로 연구자가 계획한 개수의 묶음에 문항을 한 쌍씩 배정한다. 두 번째 단계에서는 만들어진 문항묶음과 남은 문항들 간 이변량 상관계수를 다시 계산하여 각 묶음과 가장 높은 상관을 보이는 문항을 상응하는 묶음에 배정한다. 즉, 모든 묶음이 하나의 추가적인 문항을 할당받게 된다. 이런 방식으로 모든 문항이 소진될 때까지 같은 과정을 반복한다. 이렇게 하면 결과적으로 서로 비슷한 문항끼리 묶이게 되어 각 묶음의 특성이 드러나게 된다.

방사 알고리즘

방사 알고리즘(radial algorithm; Cattell, 1956; Cattell & Burdsal, 1975)은 개념적으로 요인 알고리즘과 상관 알고리즘 방식을 결합한 것이라고 할 수 있다. 요인분석을 통해 문항 간 혹은 문항과 묶음의 관계를 파악하는 것은 요인 알고리즘의 논리이며, 비슷한 문항끼리 단계적으로 묶음을 만드는 것은 상관 알고리즘의 논리이다. 먼저 일요인 분석을 하여 요인 부하의 크기 차이가 가장 작은 문항 쌍을 첫 번째 묶음에, 두 번째로 작은 차이를 보이는 문항 쌍을 두 번째 묶음에 배정한다. 이러한 방식으로 계획한 개수의 묶음에 두 문항씩 배정하면, 남은 문항들과 묶음들을 사용하여 다시 일요인 분석을 실시한다. 이번에는 하나의 요인에 대해 묶음의 요인부하와 차이가 가장 작은 문항을 상응하는 묶음에 배정한다. 앞선 방법들처럼 모든 문항이 소진될 때까지 같은 과정을 반복한다.

상관 알고리즘처럼 방사 알고리즘도 고유분산 고립전략의 한 방법으로, 비슷한 특성의 문항들을 묶으려는 목적으로 이루어진다. 다만 상관 알고리즘이 문항을 서로 직접적으로 비교하는 것에 반해, 방사 알고리즘은 전체 문항에 대해 하나의 잠재 요인을 가정하여 그 요인을 통해 문항과 묶음을 비교한다. 참고로 방사 알고리즘은 요인 알고리즘과 상관 알고리즘을 결합한 개념이지만, 시기적으로는 오히려 더 앞선 방법이다.

내용기반 접근법

내용기반 접근법(content-based approach)은 구인에 대한 이론이나 내용적 근거에 기반하여 문항묶음을 구성하는 방법이다. 다시 말해, 하나의 구인을 이루는 여러 개의 문항들 중에서

⁶⁾ 비표준화 혹은 표준화 요인부하에 따라 결과가 다를 수 있지만, 구체적인 제약은 없다. 연구자가 임의로 선택할 수 있는 부분이다.

내용적으로 비슷한 문항들끼리 문항묶음을 형성하는 방법이다. 만약 큰 개념으로서 하나의 구인을 측정하는 척도가 여러 개의 하위척도로 이루어져 있다면 하위척도에 속한 문항들로 각각 문항묶음을 형성할 수 있다. 이때 각 하위척도는 완전히 구별되는 다차원을 나타낸 다기보다, 넓은 의미로서 하나의 개념을 형성하는 구인에 대한 세부적인 하위차원이라고 할 수 있다.

앞에서 다룬 네 개의 알고리즘은 기본적으로 문항 간 관계를 통계적으로 확인하여 문항을 묶는 방식이었다. 이 방법들은 지표의 특성을 직접 통제하여 측정모형을 설정할 수 있다는 통계적 이점을 갖고 있지만, 개별묶음에 대한 의미는 모호할 수 있다. 이에 반해, 내용기반 접근법을 적용하면 각 묶음이 하위구인의 평균치라는 의미를 갖게 되므로 묶음 각각에 대한 의미 있는 해석을 할 수 있다. 이런 이유로, 앞에서 언급한 것과 같이 문항묶음을 이용한 다집단 분석을 실시하는 경우에는 내용기반 접근법을 적용하는 것이 적절하다 (Little et al., 2013).

제안 및 논의

주어진 상황에서 문항묶음을 사용해야 하는지, 그리고 사용한다면 어떤 방법을 적용해야 하는지 등에 대해서는 학자들의 의견 차이가 있지만, 올바르게 사용한다면 여러 이점을 얻을 수 있는 것이 문항묶음 기법이다. 그러나 국내에서는 전반적으로 그 개념과 활용에 대한 연구가 아직 활발히 이루어지지 않고 있으며, 문항묶음 사용을 위한 적절한 절차를 보고하지 않은 연구들이 많은 것이 사실이다.

이에 본 연구는 문항묶음 원리에 대한 연구자들의 이해를 높이고, 실질적으로 사용하는데 도움이 주기 위한 목적에 따라 작성되었다. 이를 위해 문항묶음의 개념을 정리하고, 이 기법을 사용하는 이유와 원리를 확인하였다. 또한 이를 둘러싼 여러 쟁점의 각 입장을 이해하고, 실제로 사용할 수 있는 몇 가지 유용한 문항묶음 형성방법을 설명하였다.

연구자들에게 실질적인 도움을 주고자 하는 본 연구의 맥락에서 그림 3에는 문항묶음 사용을 고려할 때 따를 수 있는 단계를 간략하게 나타낸 차트가 제공되어 있다. 이 차트에 따라 자신의 연구에서 문항묶음을 사용하는 것이 적절한지, 그리고 어떤 방법으로 문항묶음을 만드는 것이 좋을지 도움을 얻을 수 있을 것이다. 문항묶음의 사용을 고려할 때 가장 먼저 확인해야 할 것은 연구자의 목적이다. 만약 척도개발과 같이 문항 고유의 특성과 기능을 파악하는 것이 중요할 때 문항묶음을 사용하면, 개별문항 자체가 사라지고 그 특성 파악이 어렵기 때문에 적절하지 않다. 반면 구조방정식 모형을 추정하고 검증하는 것이 연구의 목적이라면, 문항묶음을 사용함으로써 지표의 연속성과 정규성을 확보하고, 측정오차를 줄이는 등 여러 이점을 얻을 수 있다. 따라서 자신의 연구목적이 무엇인지를 정확히 파악하는 것이 문항묶음을 사용하기 위한 첫 단계이자 가장 중요한 단계이다.

위와 같이 연구목적을 기준으로 문항묶음의 사용이 적절하다고 판단했다면, 다음 단계에서는 구인의 차원성을 파악해야 한다. 만약 구인의 차원성에 대한 이론이 잘 정립되어 있거나 연구자가 설정한 합리적인 이론이 있다면 그대로 적용할 수도 있겠지만, 문항 수준에서 그 구조를 확인해 보는 것이 정확한

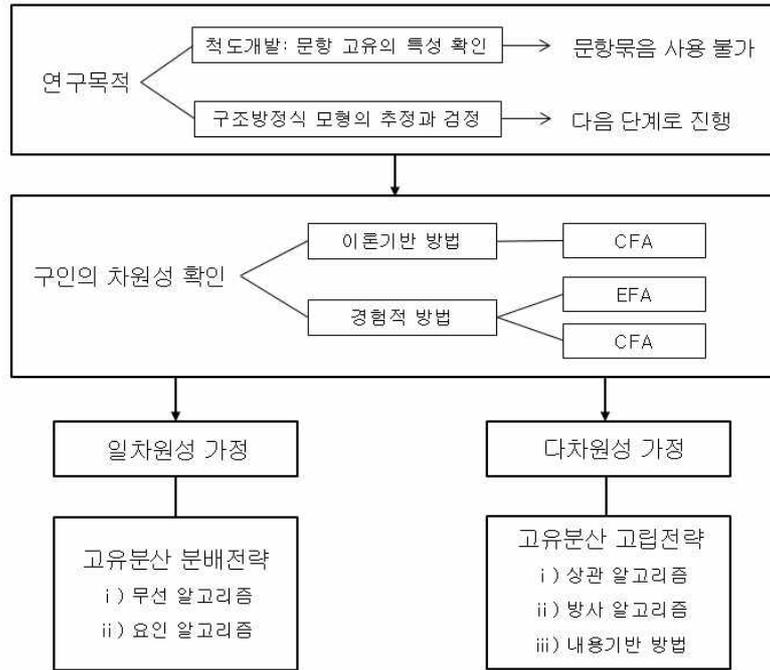


그림 3. 문항묶음 사용 단계

구조방정식 모형의 설정과 추정에 도움이 될 것이다. 이를 위해서는 문항 수준에서 CFA를 실시하여 모형적합도와 수정지수를 살펴보는 것이 좋은 방법이다. 모형적합도가 적절하며 큰 수정지수가 보이지 않는다면, 이론에서 설정된 차원성을 기반으로 문항묶음을 만드는 것이 타당하다고 볼 수 있다. 그러나 모형적합도가 좋지 않다면 이론을 다시 탐색하여 구인의 구조를 재설정하고, 큰 수정지수가 나타났다면 어느 부분에서 큰 오차 공분산이 발생했는지, 그리고 그 원인이 무엇인지 파악해야 한다.

반면 구인의 차원성에 대한 이론이 잘 정립되어있지 않다면, 경험적인 방법으로 문항 수준에서 EFA를 실시하여 구조를 파악할 수 있다. EFA 분석 이후, 설정한 요인구조로 CFA를 통해 적합도를 확인하는 절차를 거치면 차원

성에 대한 더 명확한 근거를 얻을 수 있다. 혹은 일요인 CFA를 실시하여 바로 적합도와 수정지수를 통해 일차원인지 또는 설정한 다차원이 적절한지 등을 확인할 수도 있다. 차원성은 모형적합도의 해석과 묶음방법 선택에 밀접한 관련이 있기 때문에 반드시 확인해야 할 사항이다.

마지막으로, 연구자가 설정한 차원성을 바탕으로 문항묶음 형성방법을 결정한다. 일차원성을 가정했다면, 묶음의 오차분산을 줄일 수 있으며 모형적합도를 향상시킬 수 있는 고유분산 분배전략을 채택하는 것이 여러 이점을 얻을 수 있다. 이와 다르게, 구인이 다차원성을 지닌다고 가정했다면, 모수를 더 정확히 추정할 수 있고 요인구조를 잘 드러내는 고유분산 고립전략을 적용하는 것이 적절하다.

문항묶음을 사용하는 연구자들이 지켜야 할

것이 있다. 앞서 언급한 대로 연구자들은 문항묶음을 사용하는 이유를 선명하게 밝혀야 하며, 방법 선택 과정에 대해서도 명확하고 정직해야 한다. 또한 어떤 방법을 이용하여 문항들을 묶었으며, 각 묶음에는 어떤 문항들이 포함되었는지와 같은 자세한 정보를 제공해야 한다. 이러한 정보가 쌓이면 같은 구인이 묶음화에 따라 신뢰도와 타당도 등에서 어떻게 차이가 났는지, 혹은 다른 표본에서도 결과가 일관적인지 등을 통해 해당 구인의 특성을 파악하고 후속 연구를 이끄는 데 도움이 될 것이다.

문항묶음은 어떻게 사용하는지에 따라 연구에 이점이 될 수도, 단점이 될 수도 있기 때문에 개념과 작용에 대한 정확한 이해가 필요하다. 앞에서도 강조했던 것과 같이 문항묶음 사용에 대한 결정은 상당한 자유도가 있기 때문에 연구자 자신의 판단이 매우 중요하다. 연구자는 자신이 가진 자료와 본문에서 논의된 많은 논쟁적인 내용을 동시에 고려하여 문항묶음을 최대한 적절하게 사용해야 할 것이다. 본 논문이 구조방정식 모형을 포함한 양적 연구의 질을 높이는 데 기여할 수 있을 것으로 기대한다.

참고문헌

우종필 (2015). 구조방정식모델에서 다차원성 개념의 항목묶음 편향에 대한 연구. *경영학연구*, 44(4), 1131-1147.

조현철, 강석후 (2007). 구조방정식 모델에서 항목묶음이 인과 모수의 검정과 적합도 평가에 미치는 영향. *Journal of Global Academy of Marketing*, 17(3), 133-153.

Anderson, J. C., & Gerbing, D. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin*, 103(3), 411-423.

Bagozzi, R. P., & Edwards, J. R. (1998). A general approach for representing constructs in organizational research. *Organizational Research Methods*, 1(1), 45-87.

Bandalos, D. L. (2002). The effects of item parceling on goodness-of-fit and parameter estimate bias in structural equation modeling. *Structural Equation Modeling*, 9(1), 78-102.

Bandalos, D. L. (2008). Is parceling really necessary? A comparison of results from item parceling and categorical variable methodology. *Structural Equation Modeling*, 15(2), 211-240.

Bandalos, D. L., & Finney, S. J. (2001). Item parceling issues in structural equation modeling. In G. A. Marcoulides & R. E. Schumacker (Eds.), *New developments and techniques in structural equation modeling* (pp. 269-296). Mahwah, NJ: LEA.

Bollen, K. A. (1989). *Structural equations with latent variables*. New York: Wiley.

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

Cattell, R. B. (1956). Second-order personality factors in the questionnaire realm. *Journal of Consulting Psychology*, 20(6), 411-418.

Cattell, R. B., & Burdsal Jr, C. A. (1975). The radial parcel double factoring design: A solution to the item-vs-parcel controversy.

- Multivariate Behavioral Research*, 10(2), 165-179.
- Coffman, D. L., & MacCallum, R. C. (2005). Using parcels to convert path analysis models into latent variable models. *Multivariate Behavioral Research*, 40(2), 235-259.
- Cohen, P., Cohen, J., Teresi, J., Marchi, M., & Velez, C. N. (1990). Problems in the measurement of latent variables in structural equations causal models. *Applied Psychological Measurement*, 14, 183-196.
- Cook, L. L., Dorans, N. J., & Eignor, D. R. (1988). An assessment of the dimensionality of three SAT-verbal test editions. *Journal of Educational and Behavioral Statistics*, 13(1), 19-43.
- Curran, P. J., West, S. G., & Finch, J. F. (1996). The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychological Methods*, 1(1), 16 - 29.
- Gardner, H., & Hatch, T. (1989). Educational implications of the theory of multiple intelligences. *Educational Researcher*, 18(8), 4-10.
- Gerbing, D. W., & Anderson, J. C. (1988). An updated paradigm for scale development incorporating unidimensionality and its assessment. *Journal of Marketing Research*, 25, 186-192.
- Guilford, J. P. (1972). Thurstone's primary mental abilities and structure-of-intellect abilities. *Psychological Bulletin*, 77(2), 129-143.
- Hagtvet, K. A., & Nasser, F. M. (2004). How well do item parcels represent conceptually defined latent constructs? A two-facet approach. *Structural Equation Modeling*, 11(2), 168-193.
- Hall, R. J., Snell, A. F., & Foust, M. S. (1999). Item parceling strategies in SEM: Investigating the subtle effects of unmodeled secondary constructs. *Organizational Research Methods*, 2(3), 233-256.
- Hattie, J. (1985). Methodology review: assessing unidimensionality of tests and items. *Applied Psychological Measurement*, 9(2), 139-164.
- Hau, K. T., & Marsh, H. W. (2004). The use of item parcels in structural equation modelling: Non normal data and small sample sizes. *British Journal of Mathematical and Statistical Psychology*, 57(2), 327-351.
- Kenny, D. A. (1979). *Correlational and causality*. New York: Wiley.
- Kim, S. (2000). Assessment of item parcels in representing latent variables. *Unpublished Doctoral Dissertation, University of Georgia, Athens*.
- Kim, S., & Hagtvet, K. A. (2003). The impact of misspecified item parceling on representing latent variables in covariance structure modeling: A simulation study. *Structural Equation Modeling*, 10(1), 101-127.
- Kishton, J. M., & Widaman, K. F. (1994). Unidimensional versus domain representative parceling of questionnaire items: An empirical example. *Educational and Psychological Measurement*, 54(3), 757-765.
- Kline, R. B. (2011). *Principles and practice of structural equation modeling*. New York, NY: Guilford.
- Landis, R. S., Beal, D. J., & Tesluk, P. E. (2000). A comparison of approaches to forming

- composite measures in structural equation models. *Organizational Research Methods*, 3(2), 186-207.
- Little, T. D., Cunningham, W. A., Shahar, G., & Widaman, K. F. (2002). To parcel or not to parcel: Exploring the question, weighing the merits. *Structural Equation Modeling*, 9(2), 151-173.
- Little, T. D., Lindenberger, U., & Nesselroade, J. R. (1999). On selecting indicators for multivariate measurement and modeling with latent variables: When "good" indicators are bad and "bad" indicators are good. *Psychological Methods*, 4(2), 192-211.
- Little, T. D., Rhemtulla, M., Gibson, K., & Schoemann, A. M. (2013). Why the items versus parcels controversy needn't be one. *Psychological Methods*, 18(3), 285-300.
- MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods*, 4(1), 84-99.
- Marsh, H. W., Hau, K., Balla, J. R., & Grayson, D. (1998). Is more ever too much? The number of indicators per factor in confirmatory factor analysis. *Multivariate Behavioral Research*, 33(2), 181-220.
- Marsh, H. W., Lüdtke, O., Nagengast, B., Morin, A. J., & Von Davier, M. (2013). Why item parcels are (almost) never appropriate: Two wrongs do not make a right—Camouflaging misspecification with item parcels in CFA models. *Psychological Methods*, 18(3), 257-284.
- Matsunaga, M. (2008). Item parceling in structural equation modeling: A primer. *Communication Methods and Measures*, 2(4), 260-293.
- Meade, A. W., & Kroustalis, C. M. (2006). Problems with item parceling for confirmatory factor analytic tests of measurement invariance. *Organizational Research Methods*, 9(3), 369-403.
- Millsap, R. E., & Kwok, O. M. (2004). Evaluating the impact of partial factorial invariance on selection in two populations. *Psychological methods*, 9(1), 93-115.
- Muthén, B. (1989). Multiple-group structural modeling with non-normal continuous variables. *British Journal of Mathematical and Statistical Psychology*, 42, 55-62.
- Muthén, 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.
- Nasser, F., & Wisenbaker, J. (2003). A Monte Carlo study investigating the impact of item parceling on measures of fit in confirmatory factor analysis. *Educational and Psychological Measurement*, 63(5), 729-757.
- Nasser-Abu Alhija, F., & Wisenbaker, J. (2006). A monte carlo study investigating the impact of item parceling strategies on parameter estimates and their standard errors in CFA. *Structural Equation Modeling*, 13(2), 204-228.
- Rogers, W. M., & Schmitt, N. (2004). Parameter recovery and model fit using multidimensional composites: A comparison of four empirical parceling algorithms. *Multivariate Behavioral Research*, 39(3), 379-412.
- Rushton, J. P., Brainerd, C. J., & Pressley, M. (1983). Behavioral development and construct

- validity: The principle of aggregation. *Psychological Bulletin*, 94(1), 18-38.
- Samejima, F. (1970). Estimation of latent ability using a response pattern of graded scores. *Psychometrika*, 35(1), 139-139.
- Spearman, C. (1927). *The abilities of man*. Oxford, England: Macmillan. [Abstracts]
- Steenkamp, J. B. E., & Baumgartner, H. (1998). Assessing measurement invariance in cross-national consumer research. *Journal of Consumer Research*, 25(1), 78-107.
- Stephenson, M. T., & Holbert, R. L. (2003). A Monte Carlo simulation of observable versus latent variable structural equation modeling techniques. *Communication Research*, 30(3), 332-354.
- 1차원고접수 : 2016. 02. 11.
수정원고접수 : 2016. 05. 16.
최종게재결정 : 2016. 06. 16.

Item Parceling: Understanding and Applying the Principles

Jihyun Lee

Su-Young Kim

Ewha Womans University

Item parcels, defined as the sum or mean scores of individual item scores, have been widely used in covariance structure models, such as structural equation models, because of their statistical advantages. The aims of the present study are to integrate studies on item parceling in the past 20 years and to propose appropriate methodological strategies with which researchers can apply when utilizing item parceling technology. This study first outlines the reasons why item parcels are used in light of psychometric properties and a structural equation model, and then it discusses controversies surrounding item parceling. After explaining and comparing the strategies for creating parcels, the study gives some appropriate suggestions. Finally, the study proposes steps researchers can follow and emphasizes precautions that should be considered. This study highlights the fact that using parcels is not a technique employed to resolve any problems in a model, but is a useful method to enhance the quality of research if it is used in an appropriate manner. The present study is expected to contribute to the field where item parcels are used and to enhance the quality of quantitative research.

Key words : item parcels, parceling, structural equation modeling, measurement model

부록 I

문항묶음 형성방법 예시

본문에서 소개한 다섯 가지 방법 중 무선 알고리즘과 내용기반 접근법은 추가적인 설명이 필요 없을 만큼 간단하므로 요인 알고리즘, 상관 알고리즘, 방사 알고리즘을 적용하는 방법에 대해서만 실제 자료를 통해 예시를 보이려고 한다. 사용한 자료는 임의로 생성된 4점 리커트형 척도의 8문항에 대한 300명의 응답이고, 세 개의 문항묶음을 만들기 위해 각 알고리즘을 적용한다.

요인 알고리즘

확인적 요인분석을 실시하여 얻은 비표준화 요인부하 결과가 표 I-1에 제시되어 있다. 본문의 각주에서 언급한 것처럼, 표준화 요인부하를 기준으로 하여도 문제가 되지는 않는다. 이 예에서는 4번과 5번 문항의 요인부하가 같은 값으로 추정되었다. 이런 경우에는 연구자가 임의로 배정하면 된다.

표 I-1. 일요인 확인적 요인분석 결과

문항	2번	1번	4번	5번	7번	3번	8번	6번
요인부하	1.002	1.000	.999	.999	.971	.964	.811	.793

크기 순서에 따라 첫 번째 묶음에는 2번과 6번, 두 번째 묶음에는 1번과 8번, 세 번째 묶음에는 4번과 3번을 배정한다. 남은 문항은 5번과 7번 문항이고, 7번 문항을 세 번째 묶음에, 그리고 5번 문항을 두 번째 묶음에 배치하면서 단계가 끝난다. 최종적으로 첫 번째 묶음에 2번과 6번 문항이, 두 번째 묶음에 1번, 5번, 8번 문항이, 세 번째 묶음에 3번, 4번, 7번이 속하게 된다. 설명한 방법 외에도 요인부하 크기에 따라 가장 큰 값부터 묶음에 차례대로 배정하고, 각 묶음에 하나씩 문항을 배정한 다음에는 역순으로 다시 문항을 배정하고, 또 다시 차례대로 배정하는 방식으로 문항을 할당할 수 있다. 예를 들어, 2번을 첫 번째 묶음에, 1번을 두 번째 묶음에, 4번을 세 번째 묶음에 할당하고, 다음에는 역순으로 5번은 세 번째, 7번은 두 번째, 3번은 첫 번째 묶음에 할당하는 식이다.

상관 알고리즘

상관 알고리즘을 사용하기 위해서는 가장 먼저 개별문항들 간 이변량 상관계수를 계산해야 한다. 그 결과는 표 I-2에 제시되어 있다. 가장 큰 상관을 보이는 문항 쌍은 7번과 8번(.642)이므로 이 두 문항을 첫 번째 묶음에 배정한다. 그 다음으로 큰 상관을 보이는 문항 쌍은 3번과 4

번(.622)이므로 두 번째 묶음에 배정한다. 세 번째에는 1번과 2번(.549) 문항 쌍이 배정된다. 각 문항묶음은 두 문항의 평균 값을 갖는다.

두 번째 단계로, 만들어진 문항묶음과 남은 개별문항 간의 이변량 상관계수를 다시 계산한다. 그 결과는 표 I-3에 제시되어 있다. 5번 문항은 묶음 2와 가장 큰 상관을 보이고(.589), 6번 문항은 묶음 1과 가장 큰 상관을 보이므로(.537) 각각 해당 묶음에 배정한다. 최종적으로 첫 번째 묶음에는 6번, 7번, 8번 문항이, 두 번째 묶음에는 3번, 4번, 5번 문항이, 세 번째 묶음에는 1번과 2번 문항이 포함된다.

표 I-2. 개별문항 간 상관계수

	1번	2번	3번	4번	5번	6번	7번	8번
1번	1.000							
2번	.549	1.000						
3번	.486	.523	1.000					
4번	.463	.464	.622	1.000				
5번	.511	.516	.532	.529	1.000			
6번	.373	.350	.372	.310	.368	1.000		
7번	.467	.467	.432	.418	.407	.548	1.000	
8번	.415	.407	.382	.427	.454	.417	.642	1.000

표 I-3. 문항묶음과 개별문항 간 상관계수

	5번	6번	묶음1	묶음2	묶음3
5번	1.000				
6번	.368	1.000			
묶음1	.473	.537	1.000		
묶음2	.589	.378	.509	1.000	
묶음3	.583	.411	.552	.609	1.000

방사 알고리즘

먼저 요인분석을 실시하여 개별문항의 요인부하를 표 I-1과 같이 추정해야 한다. 추가적으로 각 문항의 요인부하 추정치의 차이를 표 I-4와 같이 계산한다. 가장 작은 차이(0)를 보이는 4번과 5번 문항을 첫 번째 묶음에, 남은 문항들 중 그 다음으로 작은 차이(.002)를 보이는 1번과 2번을 두 번째 묶음에, 그리고 3번과 7번을 세 번째 묶음에 배정하고 평균값으로 문항묶음을 만든다.

표 I-4. 개별문항의 요인부하 차이

문항	2번-1번	1번-4번	4번-5번	5번-7번	7번-3번	3번-8번	8번-6번
요인부하 차이	.002	.001	0	.028	.007	.153	.018

이렇게 만들어진 세 개의 문항묶음과 남은 6번, 8번 문항으로 요인분석을 다시 실시하여 요인부하 추정치를 확인한다. 결과는 표 I-5에 제공되어 있다.

표 I-5. 묶음과 개별문항의 요인부하

	묶음3	8번	묶음2	6번	묶음1
요인부하	1.152	1.116	1.013	1.000	.979

묶음2와 6번 문항이 가장 작은 요인부하 크기 차이(.013)를 보이므로 6번 문항을 묶음2에 배정한다. 다음으로 묶음3과 8번이 .036의 크기 차이를 보이므로 8번을 묶음 3에 배정한다. 최종적으로 첫 번째 문항묶음에는 4번과 5번 문항, 두 번째 문항묶음에는 1번, 2번, 6번 문항, 세 번째 문항묶음에는 3번, 7번, 8번 문항이 포함된다.