

## 다수준 분석 시에 동질상향적 개념의 측정을 위해 참조점 이동의 필요 여부에 대한 판단방법으로서 다집단 분석\*

안 정 원<sup>1)†</sup> 이 순 목<sup>2)</sup> 정 현 주<sup>3)</sup> 이 회 영<sup>4)</sup>

<sup>1)</sup>Fordham University

<sup>2)</sup>성균관대학교 심리학과/인재개발학과

<sup>3)</sup>University of Massachusetts Amherst

<sup>4)</sup>성균관대학교 인재개발학과

본 연구에서는 다수준분석 시 집단참조 일치모형에 의한 동질상향 개념이 개인 및 집단 수준 모두에서 사용될 때, 참조점이 다른 '나(self)' 및 '우리(we)' 자료가 모두 필요한지, 둘 중 하나로도 충분한지 여부를 판단하기 위한 방법으로서 다집단 분석을 소개하였다. 동질상향적(composition) 개념화를 위하여 점수들 간의 일치도를 기초로 하는 접근에는 Chan(1998)이 제안한 개인참조의 일치모형(direct consensus model)과 집단참조의 일치모형(referent-shift consensus model) 있다. 전자는 '자기' 참조 문항으로 개인수준 자료를 수집한 후 조사내용을 평균하여 집단 수준 변수를 구성하는 경우로서, 개인참조 반응의 내용(측정치간 공분산)이 집단 차이에서 발견되는 내용과 동질성이 있다는 이론적 논거가 필요하다. 후자는 참조점을 '개인'에서 '집단'으로 이동하여 측정하는 경우인데, 구성개념이보다 잘 표상될 수 있다는 장점이 있지만, 해당 개념이 개인 수준에서도 사용된다면 집단에 대한 측정과 달리 개인에 대한 측정도 실시해야 하는 불편이 있다. 실제에서는 많은 경우, '자기' 참조 문항으로 개인 수준 자료를 수집한 후, 조사내용을 통합(aggregate)하여 동일 개념에 대한 상위수준의 측정치로 사용한다. 이러한 방식은 편리하지만 여기에는 개인참조 자료와, 연산(평균 또는 합산)을 통해 상층(예: 집단, 조직)으로 참조점이 이동된 자료 간에 측정동일성이 성립한다는 가정을 검증없이 받아들이는 문제가 있다. 본 연구에서는 다집단 분석 방법을 활용하여 그 가정을 검증하는 예를 보였다. 어떤 개념에 대한 측정에서 참조점을 달리한 두 자료간에 측정동일성이 검증되면, 추후의 연구들은 그것을 근거로 '개인' 또는 '조직'을 참조로 하여 1회의 측정만으로 다수준 분석을 할 수 있다. 또한 이런 경우, '집단' 참조의 자료를 하층의 측정모형과 상층의 동질상향 모형에 사용하는 편이, 집단차이에 의한 효과를 보다 분명히 할 수 있다. 그러나 이러한 측정동일성이 성립하지 않는다면 연구자는 개인수준 모형과 집단수준 모형이 모두 분석될 경우 반드시 참조점을 바꾸어가며 '나' 자료와 '우리' 자료를 각각 수집해야 한다.

주요어 : 다수준분석, 동질상향, 참조점 이동, 다집단 분석, 측정동일성, 다수준구조방정식모형

\* 지난 31권 2호에 게재된 안정원, 이순목(2018)의 “조직몰입 구성개념의 개념화와 타당화: 은행의 정규직 종사자를 중심으로” 논문은 안정원의 성균관대학교 박사학위논문에서 근거하였다.

† 교신저자 : 안정원, Fordham university, 441 East Fordham Road, Bronx, NY 10458, [jahn20@fordham.edu](mailto:jahn20@fordham.edu)

조직은 근본적으로 다수준적 시스템의 특성을 가진다(Kozlowski & Klein, 2000). 즉 조직-부서/팀-개인의 각 층이 유기적 상호작용을 하며 통합적으로 운영되는 시스템이다. 다수준 분석은 조직의 다수준적 시스템 특성을 방법론적으로 구현할 수 있는 대안이 되면서 다양한 응용심리 분야에서의 활용이 증대되고 있다(이희영, 김종규, 이순목, 2013; 최은하, 2012; 최은하, 구병모, 2013; 한정원, 이경수, 박찬신, 손영우, 2009; 한태영, 김원형, 2006). 본 연구에서는 다수준분석 시 동질상향(composition)하여 2층에서 사용할 개념을 개인들에게서 측정할 때, 참조점을 반드시 집단수준으로 해야 하는가에 대한 판단방법으로서 다집단 분석 방법을 소개하였다. 다수준분석에서 동질상향적 개념이란 원래 개인수준에서 측정되는 개념이 상층에서 동일한 내용과 의미, 기능을 가지는 것으로서(Kozlowski & Klein, 2000), 조직 내 상이한 수준에 걸쳐서 구성개념이 동질적임을 의미한다. 예컨대, '팀 분위기'는 팀원들이 분위기에 대한 개념과 지각을 공유하는 정도가 높기 때문에(James, 1982), 동질상향적 개념으로 볼 수 있다(이순목, 2010).

이러한 동질상향적 개념을 측정하기 위한 모형에는, Chan(1998)이 제안한 개인참조[점수]의 일치 모형(direct consensus model)과 집단참조[점수]의 일치모형(referent-shift consensus model)이 있다. 개인참조 일치 모형은 '자기' 참조 문항으로 개인수준 자료를 수집한 후 조사내용을 평균하여 집단수준 변수를 구성하는 경우이다. 이 모형은 개인참조 반응의 내용(측정치간 공분산)이 집단 차이에서 발견되는 내용과 동질성이 있다는 이론적 논거가 뒷받침될 때 사용될 수 있다. 집단참조 일치모형은 참조점을 '개인'에서 '집단'으로 이동하여 측정

하는 것인데, 이때 응답자들은 자신들이 속한 조직 환경에 대해 객관적인 "평가자(raters)"로서의 기능을 하게 된다(Chan, 1998). 보다 구체적인 예시와 설명은 동질상향적 개념의 측정 부분에서 확인할 수 있다. 연구 구성개념타당도 측면에서 집단참조 일치모형의 사용이 이론적으로는 더 타당하지만, 해당 개념이 개인수준에서도 분석된다면(많은 경우 그러함) 동일 문항에 대해 참조점을 달리하여 이중으로 측정해야 하는 어려움이 있다. 특히 조직심리에서는 개인의 인식, 태도, 행동에 영향을 미치는 변수들을 개인수준에서 뿐 아니라 상위 수준의 개념으로 고려해야 하는 경우가 많은데(예: Griffin & Mason, 2002; Lee 2003), 이때 하층의 측정모형이 없는 상층 모형은 생각하기 어렵다. 따라서 하층과 상층에 걸쳐 이중 측정을 하는 것이 원칙이지만 참조점의 상이함을 제외하면 반복 질문이 되어 현실적으로 응답자들의 협조를 받기가 어려운 점이 있다. 그래서 이 경우에도 '자기' 참조 문항으로 개인수준 자료를 수집한 후 조사내용을 평균하여 집단수준 변수로 사용하는 경우가 더 많다(예: James, Demaree & Hater, 1980; Johns & James, 1979). 다수준분석 시 상층에서 사용될 동질상향적 개념을 위한 이와 같은 방식은 편리하긴 하나, 여기에는 개인참조 자료와 상층(예: 집단, 조직)으로 참조점이 이동된 자료 간에 측정동일성이 성립한다는 암묵적 가정이 내포되어 있다. 통계적으로는 두 자료에서 측정되는 개념 간 동등성을 검증없이 가정하는 것이다(Bliese, 2000).

이와 같은 이유로 다수준분석 시 연구자는 어떤 개념에 대하여 개인 참조 방식으로 수집된 개인수준의 통합치를 상층 개념의 측정치로 사용하기 전에, 참조점을 개인으로 한 경

우와 상층(예: 집단, 조직)으로 한 경우 사이에 측정동일성이 성립한다는 가정을 검증해야 한다. 또한 이 측정동일성의 검증은, 집단 참조 방식으로 수집된 자료를, 하층과 상층에 동시에 사용할 수 있는 근거가 될 수 있다. 모든 연구에서 동일 문항에 대해 참조점을 달리하여 두 번 측정하기는 어렵다. 그러므로 선행 연구에서 측정대상인 개념에 대해 참조점을 달리한 두 자료간에 측정동일성이 성립한다는 검증이 있으면, 추후의 연구들은 그것을 근거로 다수준분석 시 ‘개인’ 또는 ‘조직’을 참조로 하여 1회의 측정만으로 자료를 분석에 사용할 수 있다. 개인 참조 측정치와 집단 참조 측정치 간에 측정 구조의 동일함이 검증된 경우라면, 다수준 분석에서 상층개념의 측정치로서 ‘집단’ 참조의 자료를 사용하는 편이, 집단차이에 의한 효과를 보다 분명히 할 수 있다.

이에 본 연구는 다집단 측정동일성의 방식(Vandenberg & Lance, 2000)을 개인 참조 자료와 집단 참조 자료 간 측정동일성을 검토하기 위한 방법론으로 제시하였다. 측정동일성이 성립한다는 것은 구체적으로 상이한 두 자료간에 요인구조의 틀과 단위가 동일함을 의미하며, 단층 자료일 경우에는 일반적인 다집단 요인분석(multigroup factor analysis)을 통해 검증할 수 있다. 그러나 본 연구는 조직의 위계구조가 반영된 다수준적 속성의 자료를 분석 대상으로 하기 때문에, 자료의 집단 효과를 통제 후 일반적인 다집단 요인분석의 절차를 따랐다. 집단 효과는 다수준 구조방정식모형(Heck, 2001; Muthen, 1991, 1994) 분석을 통해 통제하였다. 이로써 개인수준에서 참조점을 달리하여 수집된 ‘나(self)와 ‘우리(we)’ 자료 간에 측정동일성 가정을 직접 검증할 수 있다.

이 동일성이 성립하면 연구자는 굳이 참조점을 바꾸어가며 두 번의 측정을 하지 않고도, ‘자기’ 또는 ‘집단’ 참조 문항으로 개인수준의 자료를 수집한 후, 이것을 하층의 측정모형과 집단별 통합치에 의한 상층 모형에 사용할 수 있는 정당성을 확보할 수 있다. 그러나 이 동일성이 성립하지 않는다면 연구자는 개인수준 모형과 집단수준 모형이 모두 분석될 경우 반드시 참조점을 바꾸어가며 ‘나(self) 자료와 ‘우리(we)’ 자료를 각각 수집해야 한다.

이에 대한 논의를 위해 다수준적 관점의 필요성, 다수준분석에서 동질상향의 개념 및 조건, 다수준 구조방정식모형, 그리고 다집단 분석에 대해 살펴본 후, 실제 사회복지 기관의 조직분배공정성 변수에 대한 경험자료를 사용하여 측정동일성 가정을 검증하는 예를 제시하였다.

#### 다수준적 관점의 필요

인간과 자연이 존재하는 이 세계는 기본적으로 시스템적이다. 시스템적이라는 것은 구성요소들이 공동의 목표를 가지고 유기적으로 연계·상호작용하며 유지되는 것을 의미하며 전체는 부분의 합보다 크다는 아리스토텔레스적 세계관을 전제한다. 위계적 구조를 가지는 조직은 근본적으로 다수준적 시스템의 특성을 가진다(Kozlowski & Klein, 2000). 다수준적 관점의 핵심은 하층의 현상은 상층의 맥락 안에 내재되며(embedded), 상층의 현상은 종종 하층 요소들 간에 상호작용과 역동을 통해 발현(emergence)된다는 것에 대한 인식이다. 그러므로 다수준적 자료의 연구자는 조직이나 집단 내 어느 한 단면이나 수준에서 나타나는 특정 현상을 파악하고 해석할 때, 개인수준의 하층

요소들과 조직(집단)의 상층 요소들이 어떤 관계를 가지는지 교차수준적 관점에서 논의할 필요가 있다. 예컨대, 조직맥락적 요소들이 구성원의 태도와 행동에 어떤 영향을 어떻게, 왜 미치는 지에 대해 이해하고 해석할 수 있어야 조직수준의 개입과 개인수준의 개입을 보다 차별화하여 제안할 수 있다.

그동안 인사·조직분야의 양적 연구들은 조직이 다수준적 구조라는 것을 인지하면서도 이러한 교차수준적 논의와 해석 없이, 개인수준에서 파악한 구성개념의 특성이나 관계를 조직 수준의 것으로 직접 해석하는 잘못된 분석단위(unit of analysis)의 오류를 범해왔다. 예컨대 '구성원에 대한 교육 및 훈련을 실시하면 조직효과성이 향상될 것이다' 라는 논리에 따라 조직은 직원교육을 실시한다. 그러나 여기에는 '개인 수준에서의 훈련 효과가 통합 상향되어 조직 수준에서 가치 있는 효과를 생성한다' 는 숨은 가정이 있다(Kozlowski & Klein, 2000). 교육훈련의 조직효과성에 대한 기여는 훈련 모델에서도 가정하고 있지만(Goldstein, 1993; McGehee & Thayer, 1961), 개인수준에서 구성원의 훈련이 상층수준에서의 효과성으로 나타나기 위해서, 많은 경우 암묵적으로만 가정되는 상향효과와 특성 및 개인수준 훈련에 영향을 미치는 조직맥락적 효과를 어떻게 개념화하고 검증할 것인지에 대한 이론적 안내나 노력은 거의 없다. 그러므로 개인수준의 KSA에 초점이 맞춰진 교육훈련의 요구분석과 과정설계 및 평가에 대한 이론적 모형은 다수준적 관점이 결여된 것이고, 실제로 조직수준에서의 효과성에 대한 설명 모형이 누락된 것이다(Kozlowski & Salas, 1997).

다수준적 관점의 결여는 방법론적으로도 오류를 수반한다. 다수준적 자료의 분산은 다수

준에 걸쳐 존재한다(Muthen & Satorra, 1995). 즉 조직에 내재된 개인들을 대상으로 수집된 자료에는 기본적으로 다수준적 속성이 내재되므로 응답치의 총분산은 집단 간과 집단 내 성분으로 구성되는데(Zyphur et al, 2008), 이러한 자료를 개인수준의 단일모형으로 분석할 경우 개인의 인식, 태도, 행동에 영향을 미치는 조직맥락효과가 표준오차로 처리됨으로써 구성개념 간 관계 검증에 오류가 발생하는 것이다(Kozlowski & Klein, 2000, 40쪽 참조). 이와 같은 오류를 원자적 오류(atomic fallacy)라고 한다. 반대로 상층의 통합측정치로 하층의 생태에 대한 이해 없이 하층 변수간 관계를 추론할 때 발생하는 생태적 오류(ecological fallacy)도 있다. 예컨대, '부서 수준의 직원훈련비 투자를 늘리면 개인성과가 증대될 것이다' 라는 논리에 따라 조직의 직원훈련비 투자를 늘리는 것과 같다. 즉 개인성과의 증대를 직원훈련비 증가의 함수로 잘못 생각하고 개인들이 어떻게 성과를 내는 지와 같은 생태에 대한 이해 없이, 개인성과를 기대한다. 이것은 직원들의 성과를 조직이 투자한 훈련비 총액이라는 변수에 동일시하여 추론하는 오류를 범하는 것이다.

한편 다수준적 관점의 분석을 위하여 상층개념을 정의하는 데는 두 가지 방식이 있다. 즉, 상층의 고유 개념(global property)을 사용하는 방식과 하층으로부터 상향 발현된 개념(emergent property)을 사용하는 방식이 있는데, 물론 두 가지를 다 사용할 수도 있다. 특히 상향 발현된 개념에는 동질상향(composition)과 이질상향(compilation)이 있는데, 이질상향은 아직 이론적으로 개발이 덜 되었기 때문에 본 논문의 초점은 동질상향 모형에만 국한한다. 본 논문에서는 이론적으로 상층과 하층간에

동일 개념이 가정될 때, 하층개념이 집단 내 개인 간에 합산 또는 평균으로 통합되어 상층에서의 측정치로 사용되는 동질상향적 개념화를 하는 경우를 중심으로 살펴보았다.

### 동질상향(composition)의 개념 및 조건

#### 동질상향의 개념

조직 내 현상들은 구성원들의 인지나 특성, 상호작용의 결과로 나타나는 경우가 많다. 예컨대, 조직분위기와 같은 구성개념은 개인들이 심리적으로 인식하는 분위기와 조직수준에서 나타나는 분위기의 내용과 의미가 같고, 동일한 논리적 개념 연결망(nomological network)을 가지므로(Jones & James, 1979; Kozlowski & Hults, 1987), 구성원들 응답의 통합치(평균 또는 합산 결과)를 상층개념의 측정치로 사용하게 되는 동질상향(composition)적 개념이다(James & Jones, 1974). 반면 조직/팀 수준의 학습, 수행, 갈등, 효과성 등과 같은 현상들은 조직 내 구성원들의 상이한 특성과 행동, 인식이 복합되어 나타나는 집단의 특성으로(Kozlowski, 1998, 1999), 동질상향보다는 좀 더 복잡한 과정에 의해서 상층으로 구체화되어 나타난다. 즉, 그 현상들은 구성원들의 응답의 통합치로 대표될 수 없고, 서로 다른 개별 구성원들의 이질적 특성이 독특한 형태(pattern), 변산 또는 네트워크 등을 이루면서 상향 발현되는 개념인 이질상향(compilation)적 개념이다. 현재 상향적 개념화의 대부분은 동질상향적이고 이질상향적 개념화의 시도는 거의 없다(예외: 최은하, 2012). 본 연구에서도 동질상향에서의 조건을 다루고 있다. 다수준적 이론 개발 시, 상층 구성개념을 구성하는 동질상향적 접근 방식은 연구자에게 연구의 편리함을 제공하므로

(Tay, Woo & Vermunt 2014) 중요한 개념적 도구가 된다. 예컨대, 개인수준의 심리적 구성개념이 집단 내 구성원들간 공유되어 집단수준의 특성으로 논의될 때(예. 팀 효능감), 과연 개인 수준의 실제적이고 심리적인 개념이 집단수준에서도 유지되어 집단 간 비교가 가능한 지에 대한 의문이 생긴다. 이 때 동질상향적 접근 방식은 상층개념의 의미와 속성이 하층의 상응하는 개념과 유사하다고 가정함으로써 상층개념에 대한 이론적 근거를 제공해주고, 집단수준에서의 수량화를 통해 변수 간 관계도 추론할 수 있게 해준다.

#### 동질상향의 조건

그렇다면 무엇을 동질상향적 개념으로 볼 것인가? 사실 어떤 구성개념이 동질상향적인지 아닌지는 구성개념의 속성이나 그것이 발현되어지는 과정 등에 관한 충분한 이론적 답론을 거친 후 정립되어야 할 부분이다. 본 연구에서는 다수준 연구 시 구성개념 타당화 관점에서, 어떤 심리측정적 조건들이 충족되어야 동질상향적 개념이라고 평가할 수 있는지에 대해 논의하였다. 이와 관련하여 Tay 등 (2014)은 상층과 하층 구성개념의 동질성은 집단 내 개별 점수간 유사성, 교차수준적 측정 동일성, 그리고 개념 간 관계모형의 교차수준적 유사성의 세 가지 측면을 모두 포함한다고 제시하였다. 그러나 마지막의 관계모형 동일성은 이론구조만의 동일성으로서 넓은 의미의 타당화에는 포함되지만, 심리측정적 동질성에는 해당되지 않으므로 본 연구에서는 앞의 두 가지 조건에 대해서만 제시하였다.

첫째, 집단 내 개별 점수 간 유사성은 상호 효과의 원재료가 되는 기본 내용의 형태와 양이 집단 내 모든 개인들에게 있어 본질적으로

동일함을 의미한다. 따라서 점수의 일치도나 일관성과 같이 집단 내 수렴을 나타내는 지수들(예:  $a_{WG}$ ,  $r_{WG}$ ,  $ICC(1)$ ,  $ICC(2)$ )에서 바람직한 값이 동질상향의 요건이 된다. 즉 집단내 견해가 수렴하면 개인의 점수들을 합산(평균)하여 집단의 점수로 사용할 수 있지만, 점수의 분산이 크면 개인 수준의 대표값에 의미가 없기 때문에 개인수준의 통합치를 집단수준 개념에 대한 측정치로 사용할 수 없다.  $a_{WG}$ 와  $r_{WG}$ 은 각각 .7 이상 일 때 집단내 일치도 및 일관성이 높다고 간주한다(James et al., 1984; Kozlowski & Hattrup, 1992). 계층 내 상관계수인  $ICC(1)$ 은 .05 이상일 때 집단 간 분산 크기를 고려할 필요가 있다(Dyer et al., 2005). 팀 평균의 신뢰도에 해당하는(집단의 평균이 얼마나 안정적인가)  $ICC(2)$ 는 일반적인 신뢰도와 마찬가지로 .7 이상을 양호하다고 간주하지만, 다른 다수준 통계지수와 함께 전체적으로 고려할 필요가 있다(Bliese, 2000).

둘째, 교차수준적 측정동일성(cross-level measurement invariance)은 수준 간 요인구조의 측정동일성에 의해 동질의 상층 구성개념을 정의하고 구체화한다(Chen, Bliese, & Mathieu, 2005). 앞서 동질상향적 개념화를 위한 다수준적 접근 방식에는 개인참조의 일치 모형과 집단참조의 일치모형이 있다고 언급하였다. 개인참조의 일치 모형은 '자기' 참조 문항의 점수 통합치를 집단수준 개념의 측정치로 사용하고, 집단참조의 일치모형은 '개인'에서 '집단'으로 참조점을 이동하여 측정한 점수를 통합하여 집단 수준의 측정치로 사용한다(Chan,

1998; Kozlowski & Klein, 2000). 어떤 모형을 사용할지는 해당 구성개념의 속성이나 그것이 발현되어지는 과정 등에 관한 이론에 근거하여 연구자가 택할 수 있지만, 어느 경우에서든 동질상향적 개념이라면 구성개념 타당화 관점에서 층간 요인구조의 측정동일성이 지지되어야 한다. 즉, 동질상향적 변수 및 변수 간 관계에 대한 추론이 타당하기 위해서는 교차수준적으로 동일한 구성개념이 측정되고 있는지에 대한 확인이 필요하다(Chen, Mathieu, & Bliese, 2004; Cronbach, 1976; Marsh et al., 2009; Raudenbush & Bryk, 2002; Schweig, 2013; Tay et al., 2014; Zyphur, Kaplan, & Christian, 2008). 그러기 위해서 우선, 다수준 분석 시 구성개념의 요인 수효가 개인과 집단 수준에서 동일하게 유지되어야 한다(Chen et al., 2004). 이러한 검증이 병행되지 않는 가운데, 개인수준의 점수 통합치를 기계적으로 집단수준 변수의 측정치로 사용하는 것은 잘못된 접근이다. 예를 들어, 세 개 하위 차원으로 구성된 조직공정성(분배, 절차, 상호작용) 변수를 측정한 개인 수준의 자료에서, 각 차원의 측정치들을 바로 평균하여 집단수준의 조직공정성(분배, 절차, 상호작용) 변수를 구성할 수는 없다. 이처럼 검증되지 않은 교차수준적 측정동일성 가정의 단점은 단순히 그 가정이 잘못된 것임에 그치지 않고, 다수준적 구성개념이 서로 다른 분석 수준에서 어느 정도로 상이한 의미나 요인구조를 가지는지, 나아가 구성개념들간 관계에서 어느 정도로 상이한 기능을 나타내는지 발견할 기회를 놓친다는 데 있다(참조: Cronbach, 1976; Sirotnik, 1980). 교차수준적 측정동일성은 엄격한 동일성 제약의 검증을 요구하므로 반드시 이론과 자료 검증을 통해 타당화 되어야 함에도 불구하고, 많은 다

1) 주: 조직과학에서 선호하는 용어인  $ICC(1)$ 과  $ICC(2)$ 를 문헌에 따라 달리 부르기도 한다. 즉,  $ICC(1)$ 을  $ICC$ ,  $ICC(2)$ 를 group mean reliability로 부르는 경우도 있다.

수준 연구자들이 이 부분의 고려 없이 동일성을 가정하고 모형 분석을 하는 데는, 많은 다수준분석 프로그램(예: WABA 및 HLM 소프트웨어가 그러함)에서 그런 모형을 기본으로 채택했던 것도 이유가 된다.

다수준분석 시 동질상향적 개념의 타당화를 위한 심리측정적 요건으로 집단 내 개인 점수 간 유사성과 교차수준적 측정동일성이 모두 충족되어야 하지만(Tay et al, 2014), 현재의 관행은 주로 앞서 언급한 집단내 점수의 유사성을 중심으로 평가하고 있다. 이것은 구성개념에 대한 측정모형보다는 측정치만을 중심으로 하는 것이다. 반면, 자료의 위계적 구조를 반영하여 보다 명시적으로 하층과 상층 자료 각각에 요인분석을 실시하고 측정모형의 층간 동일성을 직접 평가하는 연구는 거의 없다(Dyer, Hanges, & Hall, 2005). 이는 현재의 다수준분석 방법에서는 이 부분의 검증을 위한 분석 틀이 없기 때문이긴 하지만(Tay et al., 2014; Zyphur et al, 2008), 그렇다고 방법이 전혀 없는 것이 아니다. 두 가지 대안이 있다. 하나의 방법은, 원자료에서 다수준 구조방정식모형 분석을 통해 집단내 효과와 집단간 효과를 서로 오염되지 않게 분리한 각각의 공분산 행렬을 추출하여, 그것을 대상으로 다집단 측정동일성을 검증할 수 있다(참조: Schweig, 2013). 이 경우 상층의 집단 수가 충분히 많아야 한다. 또 하나의 방법은 하층의 개인응답 자료를 가지고 다수준 구조방정식모형을 분석하면서 층간 제약을 부과하여 측정동일성을 검증할 수 있다. 이 경우는 상층의 집단 수가 앞서의 대안에서 보다 적어도 될 것이다. 이러한 맥락에서 본 연구의 목적은 다음과 같다.

## 본 연구의 목적

본 연구의 목적은 다수준분석 시 집단참조 일치모형에 의한 동질상향 개념이 개인 및 집단 수준 모두에서 측정모형으로 사용될 때, 참조점이 다른 '나(self)' 및 '우리(we)' 자료가 모두 필요한지, 둘 중 하나로도 충분한지 여부를 판단하기 위한 방법으로서 다집단 분석을 소개하는 데 있다. 특히 응용심리에서는 개인의 인식, 태도, 행동에 영향을 미치는 변수들을 개인수준에서 뿐 아니라 상위수준의 개념으로 고려해야 하는 경우가 많다. 이때 측정대상 구성개념의 이론적 상향 모형이 집단참조 일치모형이라면, 연구자는 참조점을 바꾸어가며 '나(self)' 자료와 '우리(we)' 자료를 각각 수집해야 하는 어려움이 있다. 여기서 구성원들이 평가하는 상층의 속성은 개념적으로는 '자기(self)' 참조의 구성개념과 다르다. 그러나 응답자는 모두 1층에 소속되기도 하면서 해당 집단에 소속되기도 하여 '개인' 참조 응답이라고 해도 집단 맥락이 반영되는 것을 완벽하게 배제할 수 가 없고, '집단' 참조 응답에서도 개인 수준의 생태가 반영되는 것을 완벽하게 통제할 수는 없다. 즉, 개인 참조와 집단 참조 간에 측정의 유사성이 발생한다. 그 정도가 커서 참조점을 달리한 두 자료간에 측정동일성이 성립한다는 검증이 있으면, 추후의 연구들은 해당 구성개념에 대해 '자기' 또는 '집단' 참조 문항을 통한 한 번의 측정만으로 다수준분석을 할 수 있는 정당성이 확보된다. 개인 수준에서 수집된 '자기' 및 '집단' 참조의 두 구성개념이 같다고 하려면, 집단효과를 통제한 후 '자기' 참조 자료와 '집단' 참조 자료 간에 요인구조의 틀과 단위가 동일함(일반화 가능성)이 검증되어야 한다.

본 연구는 다수준 구조방정식모형을 통해 '자기' 및 '집단' 참조 자료 각각에서 집단효과가 통제된 공분산 행렬을 추출하여, 그것을 대상으로 다집단 분석의 측정동일성 검증 방식을 사용하여 두 자료의 일반화 가능성을 검토하였다. 이를 통해 참조점을 달리한 '나' 자료와 '우리' 자료 간에 측정동일성 가정을 수량적으로 검증할 수 있고, 다수준분석 시 집단참조 일치모형에 의한 동질상향 개념이 개인 및 집단 수준 모두에서 사용될 때, 참조점이 다른 '나' 및 '우리' 자료가 모두 필요하지, 둘 중 하나라도 충분한지 여부를 판단할 수 있다.

#### 다수준 구조방정식 모형 개요

측정동일성이 성립한다는 것은 구체적으로 상이한 두 자료 간에 요인구조의 틀과 단위가 동일함을 의미하며, 단층 자료일 경우에는 일반적인 다집단 요인분석(multigroup factor analysis)을 통해 검증할 수 있다. 요인분석이란 다수의 문항/변수들간의 상호 연관성(공분산구조, 상관관계)을 분석하여 이들간에 공통적으로 작용하고 있는 잠재 요인이나 구성개념을 추출하는 통계적 기법이다. 요인분석을 포함하여 보편적으로 많이 사용되는 통계적 방법들은 관찰치의 독립성을 가정한다. 그러나 본 연구에서와 같은 다수준적 속성의 자료는 자료의 위계적 구조로 인해 동일 집단내 응답치들은 동일 집단효과에 의해서 발생하는 상관을 포함할 수밖에 없는데 이것은 관찰치의 독립성 가정에 근본적으로 위배된다. 이로 인해 다수준적 속성의 자료에서 일반적인 요인분석을 하면, 요인계수나 요인간 상관과 같은 모수 추정치들이 왜곡될 수 있고 연구자

는 잘못된 결론을 내릴 수 있다. 이런 경우 자료의 위계적 구조를 반영할 수 있는 다수준 구조방정식모형(MSEM, Multilevel Structural Equation Modeling)을 사용하면 유용하다.

다수준 구조방정식모형은 구조방정식모형에서 자료의 위계적 구조를 반영하기 위해, 개인들의 자료를 집단내 및 집단간 성분으로 나누는 모형을 세운다(Muthen & Satorra, 1989, 1995). 즉, 측정대상 변수의 변산을 각 층과 관련된 분산성분으로 나누어주는 것인데(Muthen & Muthen, 1998), 개인의 응답 점수를 개인성분(집단평균으로부터의 편차, group-mean centered)과 집단성분(집단소속이 가지는 효과)로 나누어 준다. 따라서 개인수준 모형은 집단내에서 개인들 간(between individuals)의 차이에 의한 변산을, 집단수준 모형은 집단 간(between groups)의 차이에 의한 변산을 반영한다. 개인수준 모형은 각 집단 내에서 평균 중심화(편차점수화)된 자료를 전체 집단에 걸쳐 통합한 자료(pooled-within data)를 분석하는데, 이 때 각 조직 간에 동일모형이면서 측정오차만 다르다는 가정을 세운다. 다수준 공분산 구조분석이 어려운 이유는 각 집단별로 공분산 행렬을 계산하는 복잡성 때문인데, Muthen (1994)의 접근 방식(pooled-within data를 분석)은 이러한 문제를 간소화해주었다. 또한 다수준 구조방정식모형에서는 관찰치간 독립을 집단(cluster)간에만 가정하고, 집단들에 대해서는 무선 표집을 가정한다. 그에 따라 집단수준에서 산출되는 요인평균(factor mean)은 집단에 걸쳐 변하는 확률변수(random variable)가 된다.

본 연구에서는 개인수준에서 참조점을 달리하여 수집된 '나'와 '우리' 자료 간에 측정동일성 가정을 다집단 분석을 통해 직접 검증하기 전에, 자료의 다수준적 속성을 통제하기



위한 방안으로서 다수준 구조방정식모형을 사용하였다. 즉, Mplus 7.0의 다수준분석 기능을 사용하여, ‘나’ 자료와 ‘우리’ 자료 각각에서 집단효과를 통제하고서 산출되는 집단내 공분산 행렬의 통합(pooled-within level covariance matrix,  $\Sigma_{pw(I)}$ ,  $\Sigma_{pw(We)}$ )을 추출하여, 일반적인 다집단 분석을 위한 개인수준 자료로 사용하였다.

### 다집단 분석 개요

다집단 분석은 상이한 모집단을 대상으로 하는 경우를 말하는데, 여기서 모집단의 개념은 연구자가 설정하는, 이론적으로 상이한 상황에서의 자료를 중심으로 정의된다. 따라서 자료가 질적으로 구분되는 상황이나 측면을 상이한 모집단으로 간주한다(이순목, 2014). 즉, 다집단 분석은 구조방정식모형의 일반화 또는 차별화를 위한 방법론으로서, 상이한 모집단에 동일한 측정치를 사용하여 수집된 자료에 모형을 합치시켜 일반화 또는 차별화를 시도하는 것이다(이순목, 1993; 이순목 & 김한조, 2011; Joreskog, 1971; Meredith, 1993; Sorbom, 1974). 일반화 연구의 예로는 척도나 모형에 대하여 집단 간 일반화의 시도, 비교문화적 검토, 또는 기존의 검사 양식과 동일한 구성개념을 측정하는 대안적 검사 양식의 개발에서 두 양식이 평행한지의 검토 등을 들 수 있다. 또한 차별화를 중심으로 한 연구의 예는, 서로 다른 시점에서 모형의 변화를 파악하거나 시점에 따라 진행되는 조직의 변화를 연구하는 경우이다(Vandenberg & Lance, 2000).

일반화는 집단 간에 동일성(동질성)이 성립되는지를 보는 것이므로 다집단 분석은 일련의 순차적인 동일성 검증 단계를 포함한다.

이순목(2014), 이순목과 김한조(2011), 및 이순목, 윤수철, 차정은, 김종남, 여성철(2012)에 따르면, 첫 번째 단계는 각 요인들이 특정의 측정변수에 의해서 구체화되는 “측정의 틀(configuration)” 또는 “요인과 측정변수 간 관계”가 집단 간에 동일한지를 검증하는 측정틀 동일성(configuration invariance) 검증이다. 집단 간 측정틀의 동일성이 성립되고 나면 그 다음으로 측정을 할 때 사용되는 단위에 대한 고려가 필요하다. 상이한 상황(집단) 간에 개념의 해석 및 계량화 과정에서 동일한 단위(요인계수)를 써야 비교가 가능하기 때문이다. 이것을 측정단위 동일성(metric invariance) 검증이라고 하며, 전반적/부분적 측정단위 동일성의 성립은 다음 단계인 측정원점 동일성 검증으로 이어진다. 측정을 하는 원점, 즉 속성(이론변수)의 값이 0 일 때 부여되는 측정치(속성을 측정치로 변환하는 1차식에서의 절편, scale)가 동일해야 같은 단위를 적용할 때 동일하게 해석될 수 있는 수량 값이 산출된다. 상이한 집단 간에 측정 시 원점이 동일하면, 측정원점(scale) 동일성(scalar invariance)이 성립한다고 한다. 전반적/부분적 측정원점의 동일성이 성립하게 되면 다집단 분석의 궁극적인 목적이라고 할 수 있는 집단 간 요인평균 비교가 가능해진다. 그 외에도 요인분산/공분산 동일성 검증 및 측정오차분산 동일성 검증 등이 가능하다. 이러한 순차적인 검증단계를 통해, 선행단계의 검증이 통과된 모형은 다음 단계의 모형과 비교 시에 기저모형이 된다.

본 연구의 맥락에서 일반화를 보면, 개인수준에서 ‘나’를 참조점으로 한 자료와 ‘우리’를 참조점으로 한 자료는 상이한 모집단에서 구한 표본으로 간주되어 다집단 분석이 가능하다. 그러나 전형적인 다집단 분석과 차이가

있다면 앞서 언급한 바와 같이, 본 연구에 사용된 두 자료는 명시적으로 다수의 집단이 내재된 경우로서, 각 자료에서 다수 집단들의 효과를 통제된 결과를 통합하여 다집단 분석을 위한 '나' 자료와 '우리' 자료를 정의한 것이다.

### 동질상향적 개념의 측정

다수준분석에서 집단참조 일치모형을 따르는 동질상향 개념의 연구변수가, 개인과 집단 수준의 측정모형 모두에 사용된다면 연구자는 '자기' 참조 자료와 '집단' 참조 자료 각각을 준비해야 한다. 예컨대, 작업 팀 연구에서 팀 수준의 자기효능감은 '집단'을 참조점으로 한 자료에서 개별 점수의 일치도를 근거로 하여 상향적으로 정의된 개념이다(예. Guzzo, Yost, Campbell, & Shea, 1993; Kozlowski et al., 1994). 이 때 연구자는 측정을 위해 "나는 이 과업을 수행하는데 자신이 있다"와 같은 '자기' 참조 문항 세트와, "나는 우리 팀이 이 과업을 수행할 수 있다고 확신한다"와 같이 효능감 인식의 참조점이 나(self)에서 우리 팀(we)으로 이동한 '팀' 참조의 문항 세트를 각각 구성해야 한다. 이것은 동일한 응답자 집단에게 상이한 척도를 적용하여 질적으로 구분되는(모집단이 다른) 자료표본을 산출하는 것이다.

'팀(we)' 참조 문항들을 통해 새롭게 정의된 개념을 '팀 효능감(psychological team efficacy)'이라고 할 때, 이것은 구성원들이 팀 수준의 효능감에 대해 가지는 개인의 믿음과 확신으로 정의되며, 효능감에 대한 기본적인 내용은 본래의 구성개념(자기효능감)과 동일하게 유지되면서 내용의 참조점만 변화(self->team)한 경우이다. 이처럼 '팀 효능감(psychological team

efficacy)'은 본래 개인수준에서 개념화된 것(Guzzo et al., 1993)임에도 불구하고, 이것의 심리측정적 속성은 조직 수준에서만 나타난다. 즉, 개인들은 자신이 속한 팀에 대해 각자의 믿음과 확신을 가지지만, 공유된 믿음과 확신은 오직 조직 수준의 분석에서 나타나며 파악될 수 있다. 따라서 '팀(we)' 참조 문항의 응답치를 상층의 '조직 수준의 팀 효능감(organizational team efficacy)' 측정치로 사용하려면 이 문항들의 집단 내 일치도가 높아야 한다(Chan, 1998). 나아가 앞서 동질상향의 조건 부분에서 언급하였듯이 교차수준적으로 동일한 구성개념이 측정되고 있는지에 대한 확인도 필요하다.

### 측정동일성 검증의 단계

측정동일성 검증을 하기 전에, 집단간 공분산 자료동일성의 검증이 반드시 아니지만 실시되는 경우가 있다.

#### 집단간 공분산 자료동일성에 대한 검증

자료의 공분산 행렬에는 대각선에 분산이 있고 비대각선에 공분산이 있다. 이 때 두 집단의 공분산행렬을 각각  $S_1$  과  $S_2$  라고 할 때, 이들이 모집단에서는 같은 행렬이라는 가설이 검증될 수 있다. 즉, " $H_0: \Sigma_1 = \Sigma_2$ " 이 검증될 수 있다. 구조방정식모형 이론에서는 이 가설이 성립하면 자료는 통합되고 하나의 모집단에서 나온 자료로 취급한다. 즉, 다집단이 아닌 상황이 되어 '집단간' 모형의 일반화를 해야 하는 경우가 아니고, 하나의 집단에 잘 맞는 모형을 연구해야 하는 경우가 된다. 한편 이 검증의 논리와 필요성에 대해서는 부정적인 견해도 있다(Byrne, Shavelson & Muthen,

1989). 자료동일성 검증 시에는 비교대상이 되는 기저모형(baseline model)이 없기 때문에, 적절한 합치도 수준이 어느 정도인지에 대한 의사결정이 어렵고, 기저모형이 사용되는 그 이후 단계의 검증과 연관되지 않을 수 있으므로, 굳이 선수조건으로 검증될 이유가 없다는 것이다. 그러나 Vandenberg와 Lance(2000)가 검토한 67개 연구 중 15개에서는 이 검증을 실시하였다.

본 연구에서는 ‘나’ 및 ‘우리’ 참조의 자료동일성 검증을 통해 두 자료가 통합 가능한지, 즉 두 자료를 동일 모집단에서 수집된 표본들로 볼 수 있는지 여부를 검토하였다. 분석 대상 자료는 앞서 언급한 바와 같이, Mplus 7.0의 다수준분석 기능을 통해, ‘나’ 자료와 ‘우리’ 자료 각각에서 집단효과를 통제한 후 개인성분에 의한 효과를 통합한 두 개의 집단내 공분산 행렬(pooled-within level covariance matrix,  $\Sigma_{pw(I)}$ ,  $\Sigma_{pw(W_e)}$ )이 된다. 공분산 자료간 동일성이 기각되면, 측정동일성 검증 단계로 간다.

### 측정동일성 검증의 단계

측정동일성 검증은 다음의 절차를 따른다. **첫째**, 측정틀 동일성(configural invariance) 검증은 두 자료 간에 측정되고 있는 구성개념이 유사한 차원의 구조를 가질 경우에 실시하는 것으로, 각 측정변수가 어떤 요인을 나타내는 지(요인계수를 자유모수로 표시) 또는 아닌지(요인계수가 “0”)의 모양(configuration)이 집단 간에 동일한지를 검증한다. 본 연구에서는 분배공정성 5개 문항에 대한 1요인 모형이 측정의 틀이다. 측정틀 동일성이 성립되면, 참조점을 달리한 두 자료에서 분배공정성의 개념구조 또는 개념 이해의 틀이 동일하다고 해석할 수 있다. 이순목, 김한조 (2011)에 따르면 이

검증은 두 가지 순서로 이루어진다. 먼저, 각 자료 집단에 대해 개별적으로 합치시켰을 때 적절한 수준의 합치도를 보이는 모형이 집단 간에 상당한 정도로 유사해야 한다. 요인계수가 0이 아닌 것과 0인 것의 패턴이 집단 간에 전반적으로 같은 것이 바람직하다. 그렇지 않으면 집단 간에 부분적인 틀동일성을 검증하는 경우가 된다. 다음으로, 각 모형을 자료에 동시에 합치시키면서 틀동일성 검증을 할 때 적절한 수준의 합치도를 보여야 한다. 집단 간에 측정틀 동일성이 지지되는 모형은, 이후에 측정단위 동일성의 가설을 검증할 때 비교대상인 기저 모형(baseline model)이 된다. 간혹, 각 요인의 분산이 수렴하는지를 검토하지 않고 넘어가는 경우가 있는데, 수렴타당도를 위해 반드시 검토되어야 한다. 수렴이 안 되는 요인들에 대해 측정동일성 검증을 하는 것은 의미가 없기 때문이다.

**둘째**, 측정단위 동일성(metric invariance) 검증을 실시한다. 측정단위는 요인의 1단위 변화가 측정치에 미치는 효과, 즉 요인계수인데, 구성개념이 어떻게 양적으로 구체화되며 표준화될 경우 상대적으로 강조되는 측면이 무엇인지를 나타낸다. 측정단위 동일성은 구조방정식모형에서 집단 간 요인계수를 동일하게 제약함으로써 검증할 수 있다. 본 연구에서는 분배공정성(이론변수)이 5개 문항들(측정변수)로 구체화되는 방식이 문항 참조점(‘나’ vs ‘우리’)이 다른 자료 간에 동일한지에 대하여 측정단위 동일성 검증을 이용하여 확인하였다.

**셋째**, 측정원점 동일성 검증(scalar invariance)을 실시한다. 측정원점은 요인의 척도 상에서 ‘영점’에 해당되는 지점(Bollen, 1989)으로, 요인점수 ‘0’에 해당되는 측정치(측정변수 방정식에서의 절편)가 집단 간에 동일한지를 검증하

는 것이다. 측정틀, 측정단위 그리고 측정원점이 동일하게 되면, 원칙적으로 두 집단에서 동등한 점수는 측정변수인 척도점수에서도 동일한 수준의 값을 가진다고 볼 수 있다. 본 연구에서는 자료동일성과 측정틀 및 측정단위 동일성까지만 검토하였다. 공분산구조 검증에서 평균구조는(절편 포함) 관심의 대상이 아니므로 측정원점 동일성은 크게 필요하지 않다.

나아가 측정오차분산 동일성(uniqueness invariance) 검증이 있다. 측정오차가 집단 간에 동일한지를 검토하는 과정으로서, 측정도구 자체에 관심이 있을 때 실시한다. 즉, 측정변수 수준에서의 신뢰도가 연구의 중심일 때 검토하며, 요인평균비교가 목적이라면 이 검증은 무관하고 생략될 수 있다(Schmitt & Kuljanin, 2008, 예. 김아영 등, 2004; 홍세희, 황매향, 이은설, 2005). 또한 요인분산 동일성(factor variance invariance) 검증은 두 집단 간 평균비교 이전에 검토한다. 요인분산은 구성개념의 척도 상에서의 변산 정도를 나타내며, 분산이 클수록 집단 내 구성개념 척도에서 1 단위의 간격이 넓다는 것을 의미한다. 요인분산 동일성이 성립하게 되면, 두 집단 간 개념 구조에서 변산이 유사함을 의미한다. 끝으로 집단 간 측정체계 즉, 측정틀, 측정단위 그리고 측정원점이 동일하면 요인의 평균(factor mean)을 비교할 수 있다. 집단들 중 한 집단을 참조집단(요인평균=0)으로 고정하고 측정원점 동일화 제약을 하게 되면 나머지 집단들의 요인평균들이 참조집단에 대비되는 값으로 산출된다. 이들의 크기를 비교하기 위해서 Cohen(1988)의 효과크기( $d$ )를 계산하여 그 크기 정도를 해석한다. 측정체계가 동일하고 요인평균의 차이가 크지 않다면, 두 집단에서 구성개념에 대한 인식 정도가 유사함을 의미한다.

### 모형 선정을 위한 합치도

측정동일성 검증 시 기저모형에 비해서 한 단계 더 제약된 모형에서 과연 합치도에 차이가 없는지(동일성 지지), 차이가 충분히 있는지(동일성 기각)에 대한 결정과 모형 평가를 위해 모형 합치도 지수를 사용한다(이순목, 2014). 모형 합치도 지수는 자료와 모형 간 합치도를 수치로 나타낸 값으로서 단일 집단에서 주로 사용하는 합치도 지수들로는 표본 크기에 민감한 검증용 합치도인 카이제곱치, 그리고 이를 보완할 수 있는 판단적 합치도로서 표본 크기에 상대적으로 덜 민감한 TLI(터커-루이스 계수, 비표준합치도 NNFI, Bentler & Bonett, 1980), RMSEA(개략화 오차평균, Steiger & Lind, 1980), CFI(비교합치도, Bentler, 1990), SRMR(표준화 원소간 평균차이, Bentler, 1995) 등이 있다. 이 합치도 지수들은 다집단 분석에서 기저모형이 없는 자료동일성 및 측정틀 동일성 검증 시 모형의 합치도 판단에서도 사용된다. 이들 합치도 지수들에 대해 전통적인 기준은 CFI와 TLI는 보통 .90 이상일 때 적합하다고 보며, RMSEA는 .08이하, SRMR은 .10 이하일 경우 적절하다고 본다(Vandenberg & Lance, 2000). 본 연구에서는 이 전통적인 기준 값을 참고하였다.

측정단위 동일성 검증부터는 기저모형과 비교 검증되는 제약모형 간의 합치도 차이( $\Delta$ 합치도)를 참조하는데, 이들은 단일집단 분석에서의 합치도들과 독립적이면서 서로 중복이 되지 않아야 한다. 동일화 제약이 부과된 모수들에 대한 동일성 여부를 판단하기 위해 Chen(2007)이 추천한  $\Delta$ CFI,  $\Delta$ RMSEA를 참고하였다. Chen(2007)이 제시한 기준치에 따라  $\Delta$ CFI가 -.01이상(절대값으로 .01 이하),  $\Delta$ RMSEA가 .015이하일 경우 측정동일성이 유지되는

것으로 판단할 수 있다.

연구자의 관심은 다수준분석 시 구성개념이 집단참조 일치모형의 상향 모형을 따를 때, 참조점을 달리한(자기, 집단) 측정치들 간에 구성개념을 측정하는 체계가 동일한지에 대한 확인에 있다. 이러한 다집단 분석에서 공분산 구조가 분석되므로 동일성 검증 단계들 중 자료동일성, 측정틀 및 측정단위 동일성 검증의 실시만으로 확인이 가능하다. 따라서 본 논문은 경험자료를 대상으로 위 세 가지 동일성 검증 과정을 통해, 동질상향적 개념의 측정을 위해 참조점을 각각 '나'와 '우리'로 달리한 두 번의 측정이 필요한지 여부를 판단하는 방법을 제시하였다.

## 방 법

### 분석 절차

본 연구는 분배공정성 개념에 대하여 참조점을 '나' 및 '우리'로 측정된 313명(33개 집단)의 개인 응답자료를 사용하여, 일련의 측정 동일성을 검증하기 위해 다음의 분석 절차를 따랐다. 즉 1) '나' 및 '우리' 참조 분배공정성 변수의 집단수준 변수로의 상향가능성 검토, 2) '나' 및 '우리' 참조 자료에서 다수준 구조 방정식모형 분석을 통해, 집단효과를 통제한 통합 집단내 공분산 행렬(pooled-within level covariance matrix)  $\Sigma_{pw(I)}$ 와  $\Sigma_{pw(We)}$ 를 추출, 3)  $\Sigma_{pw(I)} = \Sigma_{pw(We)}$ 이 성립하는지에 대한 집단 간 자료동일성 검증, 4) 측정틀 동일성 검증, 5) 측정단위 동일성 검증을 하였다. 소프트웨어로는 Mplus 7 (Muthén & Muthén, 2010)을 분석도구로 사용하였다. 자료가 5점 척도로 수

집되어 정규분포 가정을 크게 위반하지 않을 것으로 보고(Bovaird & Koziol, 2012), 다집단 분석은 최대우도(maximum likelihood)법을, 다수준 분석은 Mplus에서 기본으로 하고 있는 MLR (maximum likelihood estimation with robust standard errors)법을 사용하여 해를 추정하였다.

### 연구대상 및 측정도구

본 연구는 이회영(2012)의 자료 가운데, Chan(1998)의 집단참조 일치모형에 근거하여 '나'와 '우리'를 참조로 한 분배공정성 문항에 대한 응답 자료를 분석에 사용하였다. 동일한 응답자에게 '개인의 지각된 분배공정성'과 '집단분배공정성' 모두를 측정하였다. 원자료의 응답자는 전국에 분포한 354명의 사회복지사였으며, 이 가운데 결측 및 예외값(outlier)를 제외하고 최종 85개 사회복지기관에 재직 중인 313명의 자료가 사용되었다(남자 110명, 35.1%; 여자 203명, 64.9%, 평균연령 남자: 33세, 여자: 30세). 예외값은 자료의 25분위와 75분위 사이를 사분위간 범위(Interquartile range, IQR)로 정의하고, 25분위 아래로  $1.5 \times IQR$ , 75분위 위로  $1.5 \times IQR$ 를 벗어나는 값으로 정의하였다(Moore, 2010). 문항은 Moorman(1991)과 Niehoff와 Moorman(1993)이 개발한 5점 척도의 조직공정성(organizational justice)을 사용하였다. '나'를 참조로 한 척도와 '우리'를 참조로 한 척도의 내적일관성 신뢰도는 각각 .88과 .90이다. 측정에 사용된 문항은 표 1과 같다.

자료 수집 결과, 개인수준에서 두 개의 표본자료가 구성되었다. 즉, '나' 참조 문항으로 측정된 '개인의 지각된 분배공정성'과 '우리' 참조 문항으로 측정된 '집단분배공정성' 자료이다. '집단분배공정성'은 구성원들이 조직의 분

표 1. 집단참조 일치모형에 근거한 조직분배공정성 문항

	'나(self)' 참조 문항	'우리(we)' 참조 문항
	문항	
조직 분배 공정성	1. 내가 맡은 책임의 양에 비추어 볼 때, 나는 공정한 보상을 받고 있다.	1. 내가 맡은 책임의 양에 비추어 볼 때, 우리 복지관은 공정한 보상을 해준다.
	2. 내가 쌓아온 경험의 양에 비추어 볼 때, 나는 공정한 보상을 받고 있다.	2. 내가 쌓아온 경험의 양에 비추어 볼 때, 우리 복지관은 공정한 보상을 해준다.
	3. 내가 업무에 기울인 노력의 양에 비추어 볼 때, 나는 공정한 보상을 받고 있다.	3. 내가 업무에 기울인 노력의 양에 비추어 볼 때, 우리 복지관은 공정한 보상을 해준다.
	4. 내가 달성한 업무의 성과에 비추어 볼 때, 나는 공정한 보상을 받고 있다.	4. 내가 달성한 업무의 성과에 비추어 볼 때, 우리 복지관은 공정한 보상을 해준다.
	5. 내가 맡은 업무에서 받은 스트레스나 긴장에 비추어 볼 때, 나는 공정한 보상을 받고 있다.	5. 내가 맡은 업무에서 받은 스트레스나 긴장에 비추어 볼 때, 우리 복지관은 공정한 보상을 해준다.

배공정성에 대해 가지는 인식을 나타내는 개념으로서, 본래의 구성개념(분배공정성)과 기본적인 내용은 동일하게 유지되면서 내용의 참조점이 변화(self->team)한 경우이다.

### 집단의 재구성

원자료는 85개 복지관에 재직 중인 313명의 응답자료로 집단의 크기가 1~10명으로 매우 다양하여 다수준분석 시 집단 내 응답의 유사성이 낮고, 또한 집단 간 변별도 어려웠다. 그래서 본 연구의 목적에 맞는 자료로 활용하기 위해 자료를 재구성하였다. 다수준분석을 위한 최소한의 집단 수는 대개 몇십개 이상은 되어야 모집단에 대한 충분한 정보를 포함할 수 있고, 집단의 크기 또한 10명 이상인 경우 다른 집단과 차별된다고 볼 수 있다(Snijders & Bosker, 1999). 만약, 예산과 시간 등 현실적인

제약으로 충분한 표본을 구하지 못한 경우 추정의 정확성과 검증력 확보를 위해서는 집단의 수가 집단의 크기보다 더 중요하다(Maas & Hox, 2005). 따라서 원자료에 대해 동일한 권역(경인, 강원, 경상, 전라, 충청, 제주권역)에 위치하면서, 기관유형(종합복지관, 노인복지관, 장애인복지관)과 운영체(사회복지법인, 비사회복지법인)가 가능하면 동일하고, 응답 평균 및 분산이 유사한 복지관들로 집단을 재구성하였다. 결과로 총 33개 집단, 집단별 크기는 7~18명으로 재구성된 자료를 분석에 사용하였다.

## 결 과

### 집단수준 변수로의 상향가능성 검토

개인 통합치를 상향적 개념에 대한 상층집

표 2. '나' 및 '우리' 참조 분배공정성 변수의 집단수준 변수로의 상향가능성 검토 결과

	$a_{WG}$	ICC(1)	ICC(2)
나(I)	.77 (.51~.93)	.16	.65
우리(we)	.79 (.55~.92)	.16	.65

주. N=313(33개 집단)

수로 구성할 수 있으려면, 집단 내 구성원 간 응답이 유사하고 집단 간 변별이 되어야 한다. 따라서 질적인 측면에서 집단의 동질성뿐 만 아니라 양적인 측면에서 집단 내 개별점수의 신뢰도(ICC(1)), 집단 평균점수의 신뢰도(ICC(2)), 집단 내 일치도( $a_{WG}$ ) 값을 비교할 수 있다. 신뢰도에 대한 선행연구를 살펴보면, ICC(1)은 .3이 넘는 경우가 거의 없으며(Bliese, 2000), ICC(2)의 경우 .7이상인 경우 수용 가능하다(Klein & Koslowski, 2000).  $a_{WG}$ 의 값은 .60일 경우 약한 일치도를, .70 이상일 경우 적절한 일치도 값이라고 본다(Brown & Hauenstein, 2005). '나' 및 '우리' 참조 분배공정성 변수의 집단수준 변수로의 상향가능성에 대한 검토 결과를 표 2에 제시하였다.

위 표와 같이 개인수준의 '나' 및 '우리' 참조 분배공정성에 대한  $a_{WG}$ 는 각각 평균 .77과 .79로 적절한 일치도 수준을 보였고, ICC(1)과 ICC(2)도 두 자료 모두에서 동일한 값으로 양호하게 나타났다. 따라서 '나' 및 '우리' 참조 분배공정성 응답결과는 두 자료 모두에서 33개 집단 간 변별이 되며, 집단 내 유사성이

있다고 볼 수 있다.

### 집단 간 공분산 동일성 검증

자료동일성은 분배공정성 5개 문항에 대해 참조점을 달리한 두 자료 간에 공분산동일성 검증을 통해 가능하다. 이를 위해 Mplus 7.0의 다수준분석 기능을 사용하여, 각 자료에서 집단효과를 통제한 통합 집단내 공분산행렬(pooled-within level covariance matrix)인  $\Sigma_{pw(I)}$ 와  $\Sigma_{pw(We)}$ 을 추출하여,  $\Sigma_{pw(I)} = \Sigma_{pw(We)}$ 이 성립하는지를 검토하였다. 이는 추출된 공분산 행렬( $\Sigma_{pw(I)}$ ,  $\Sigma_{pw(We)}$ ) 간에 요인분산 및 공분산이 동일한지에 대한 검증을 통해 가능하다. 구체적으로 다집단 분석의 틀에서, 각 집단에 대해 5개 측정치 각각을 요인처럼 취급하였다. 즉, 문항의 분산을 요인분산으로, 문항간 공분산을 그들 요인간 공분산으로 간주하여, 두 집단간에 요인분산·공분산 동일화 제약을 하여 모형의 합치도를 검토하였다(표 3 참조).

요인분산·공분산 동일화 제약을 한 결과, CFI와 TLI가 .90 이상, SRMR이 .10 이하로 만

표 3. 집단간 요인분산·공분산 제약 결과

$\chi^2$	df	p	CFI	TLI	RMSEA (90% C.I.)	SRMR
69.485 (나: 27.919, 우리: 41.565)	15	.000	.960	.947	.114 (.088~.142)	.073

주. 각 자료의 표본크기는 N:(집단수)= 280 (Stapleton, 2006, 356쪽)

축할 만한 합치도 지수를 보임에 따라, '나' 및 '우리' 참조의 자료동일성이 어느 정도 지지되었다. 이것을 근거로 자료의 통합이 고려될 수 있다. 그럴 경우, 다집단이 아닌 상황이 되어 두 자료가 동일개념을 측정한다고 보고, 하나의 모집단에 잘 맞는 모형을 연구하면 된다. 그러므로 참조점을 달리한 두 번의 측정 없이 '나' 또는 '우리'로 측정한 자료 중 하나를 다수준분석에 사용할 수 있다. 그러나 RMSEA가 .1 이상으로 높은 편이므로 좀 더 명확한 검토를 목적으로 본 연구에서는 측정틀 및 측정단위 동일성까지 검토하였다.

측정틀 동일성 검증

측정틀 동일성 검증은 두 모형 간 요인구조 동일성을 검증하는 것으로서, 측정틀 동일성이 성립되면 참조점을 달리한 두 공분산 행렬 자료( $\Sigma_{pw(I)}$ ,  $\Sigma_{pw(We)}$ )에서 분배공정성의 개념구조 또는 개념 이해의 틀이 동일하다고 해석할

수 있다. 이론변수의 척도(scale)를 제공하기 위해서 Vandenberg(2002)가 추천한 방식을 따랐다. 즉 각 모형에 대해 1요인 확인적 요인분석을 실시한 결과(표 4 참조), 전반적으로 요인계수 패턴이 유사한 가운데, 요인계수가 크고 집단 간 유사한 3번 문항을 참조지표로 하여 측정틀 동일성을 검증하였다.

표 5를 보면, 참조점별 1요인 모형은 '우리' 참조의 RMSEA를 제외하고, CFI와 TLI가 .90 이상, SRMR이 .10 이하로 매우 좋은 합치도 지수를 보였다. 또한 각 모형에서 요인분산이 0이 아니면서 유의하므로, 참조점이 다른 두 집단 모두에서 분배공정성 개념의 수렴타당도가 확인되었다. 그에 따라 참조점이 다른 두 집단 전체에 대한 측정틀 동일성 검증을 실시하였다.  $\chi^2$ 는 표본크기에 민감하므로 참고로만 제시하였다. 전체집단의 모형에 대한 전반적 합치도들을 볼 때 두 집단 간 측정틀 동일성이 성립한다고 판단되었다. 즉, '나'를 참조로 한 분배공정성의 개념과 '우리'를 참조로 한

표 4. 확인적 요인분석을 통한 참조점별 요인계수 및 요인분산

구분	DF1	DF2	DF3	DF4	DF5	요인분산
나	.721	.807	.819	.687	.733	.245 ( $p < .000$ )
우리	.732	.741	.828	.859	.771	.284 ( $p < .000$ )

주.  $N_f=280$

표 5. 참조점별 모형 합치도 및 측정틀 동일성 검증결과

참조점	$\chi^2$	df	p	CFI	TLI	RMSEA (90% C.I.)	SRMR	요인분산
나	4.184	5	.523	1.000	1.003	.000(.000~.076)	.014	.245 ( $p < .000$ )
우리	20.569	5	.001	.980	.959	.105(.061~.155)	.029	.284 ( $p < .000$ )
전체	24.753	10	.006	.989	.979	.073(.037~.109)	.023	.274 ( $p < .000$ ) .345 ( $p < .000$ )

주.  $N=280$



분배공정성 개념 간에 이해의 틀이 동일하다고 해석할 수 있다. 한편 RMSEA 값은 보통 모형이 작고 자유도가 낮을 때 나빠진다(Kenny, 2010). 그러나 '우리' 참조 자료의 RMSEA 값이 높게 나타난 것은 이 경우에 해당되지 않는다. 그보다는 '나' 참조 자료의  $\chi^2$  값에 비하여 '우리' 참조 자료의  $\chi^2$  값이 크게 나타난 점으로 볼 때, '우리' 참조 자료에서 모형의 개선 여지가 있기 때문일 수 있다. 그러나 본 연구의 목적상 모형의 개선은 불필요하여 그대로 다음의 분석을 진행하였다.

#### 측정단위 동일성 검증

측정단위 동일성 검증은 두 모형 간 요인계수 동일성을 검증하는 것으로서, 구성개념이 어떻게 양적으로 구체화되며, 표준화될 경우 상대적으로 강조되는 측면이 무엇인지를 나타낸다. 두 집단에 대한 확인적 요인분석 결과(표 4 참조), 집단 간 요인계수가 유사한 문항이 많으므로 전반적 제약방식을 사용하여, 측정단위 동일성을 검증하였다. 이론변수에 대한 척도 제공을 위해 Yoon과 Millsap(2007)의

방식을 따라 '나' 참조 모형의 요인분산을 1.0으로 고정하고, '우리' 참조 모형의 요인분산을 자유모수로 설정한 후 모든 문항의 요인계수에 동일화 제약을 부과하였다. 참조점이 서로 다른 두 모형에서 큰 수정지수는 나타나지 않아 모형수정을 하지 않았다.

표 6의 측정단위 동일화 제약 결과 전반적 합치도가 양호하고( $\chi^2$ 는 표본크기에 민감하므로 참고로만 제시), 표 7의 측정단위 동일성 검증 결과에서 Chen(2007)이 제시한  $\Delta CFI$  및  $\Delta RMSEA$ 가 기준을 만족하므로, 참조점이 다른 두 모형의 측정단위 동일성 가정이 지지되었다. 즉, '나'를 참조로 한 각 문항들과 분배공정성 개념과의 관계 강도는 '우리'를 참조로 한 각 문항들과 분배공정성 개념과의 관계와 동일하다고 볼 수 있다. 이것은 참조점이 다르더라도 각 문항들이 분배공정성을 측정하는 지표로서 동일한 가중치를 지님을 의미한다.

## 논 의

본 연구에서는 다수준분석 시 집단참조 일

표 6. 측정단위 동일화 제약 결과

제약	$\chi^2$	df	p	CFI	TLI	RMSEA (90% C.I.)	SRMR
모형1: 측정틀 동일화제약	24.752	10	.006	.989	.979	.073(.037~.109)	.023
모형2: 측정단위 동일화제약	36.890	14	.001	.983	.976	.076(.047~.107)	.055

표 7. 측정단위 동일성 검증 결과

검증대상	비교	$\Delta\chi^2$	$\Delta df$	p	$\Delta CFI$	$\Delta RMSEA$	결과
측정단위 동일성	모형1 대 모형2	12.138	4	.02	-.006	.003	유지

주. Chen(2007)이 제시한 기준치에 따라  $\Delta CFI = -.01$  이상,  $\Delta RMSEA = .015$  이하일 경우 측정동일성이 유지되는 것으로 판단할 수 있다.

치모형에 의한 동질상향 개념이 개인 및 집단 수준 모두에서 측정모형으로 사용될 때, 참조점이 다른 '나(self)' 및 '우리(we)' 자료가 모두 필요한지, 둘 중 하나로도 충분한지 여부를 판단하기 위한 방법으로 다집단 분석을 소개하였다. 먼저 충분한 이론적 근거나 경험적 검증 절차 없이 연구자에 의해 전제되는 동질상향 가정이 내포하는 문제점들을 지적하고, 다수준분석에서 동질상향의 의미와 방법론적 조건을 명확히 하였다. 다수준적 구성개념에 대한 동질상향의 근거는 교차수준적 구성개념 타당화 관점에서 확보되어야 하며, 특히 교차수준적 측정동일성의 성립이 중요함을 조명하였다. 또한 상향적 개념의 측정치를 산출하고자 개인참조 일치모형이 아닌 집단참조 일치모형을 따를 때, 참조점을 달리한 '나' 및 '우리' 자료가 모두 필요한지 여부를 판단하기 위해, 두 자료 간에 측정동일성 가정을 통계적으로 검증하는 방법을 제시하였다. 선행연구에서 측정대상인 개념에 대해 참조점을 달리한 두 자료간에 측정동일성이 성립한다는 검증이 있으면, 추후의 연구들은 그것을 근거로 다수준분석 시 '개인' 또는 '조직' 참조의 한 번의 측정만으로 자료를 분석에 사용할 수 있다.

측정동일성 검증은 서구에서 이미 1990년대부터 많은 활용이 있었고 국내에서도 일찍이 소개되어(예: 이순목, 1993) 적용된 사례도 많으므로 특별히 새로운 방법이 아니다. 그러나 본 연구는 산업 및 조직심리 분야에서 앞으로의 활용이 더욱 기대되는 다수준이론 및 방법론에 있어서 동질상향 가정이 내포한 측정동일성의 의미와 중요성을 강조하고, 그에 대한 수량적 검증 방법을 제시하였다는 점에서 의의를 찾을 수 있다. 본 연구의 방법을 통해

다수준분석 시, 참조점을 달리한 '나' 자료와 '우리' 자료 간에 동일성 가정을 수량적으로 검증할 수 있는 기반이 마련되었다. 이로써 구성개념의 이론적 상향 모형이 집단참조 일치모형을 따를 때, 동질상향 개념의 측정치를 산출하기 위해, 참조점을 달리한 두 번의 측정이 필요한지 여부를 판단할 수 있다. 즉, 참조점을 달리한 두 자료간 측정동일성이 성립하면 다수준적 구성개념에 대해 굳이 참조점을 바꾸어가며 두 번의 측정을 하지 않고도 개인수준의 '자기' 참조 문항으로 자료를 수집한 후 평균하여 집단수준의 동일 변수로 사용할 수 있는 정당성이 확보된다. 마찬가지로 상층을 참조점으로 한 자료가 있다면, 그것으로 하층의 측정모형 및 상층 모형을 위한 자료로 사용할 수 있는 근거가 확보될 수 있다. 그러나 이러한 측정동일성이 성립하지 않으면 연구자는 반드시 참조점을 바꾸어가며 '나'와 '우리'를 각각 측정하여 자료를 수집해야 한다.

본 연구의 예시 결과로 조직분배공정성은 '나'와 '우리' 측정치간에 자료동일성, 측정의 틀 및 단위 동일성이 상당한 정도로 지지되어 강한 동질상향의 개념으로 나타났다. 이러한 결과는 '우리' 참조 문항의 구성에 일부 기인한다. 표 1에서 보면 '우리' 참조 문항이, '나' 참조와 '우리' 참조가 혼합되어 내용이 구성됨을 알 수 있다. 엄밀히 말하여 이것은 Chan이 제시한 완전한 '집단' 참조(referent-shift)와는 다소 차이가 있다. 예컨대, 혹자는 '우리' 참조 문항이 "내가 맡은 책임의 양에 비추어 볼 때, 우리 복지관은 공정한 보상을 해준다" 보다는, "우리 팀이 맡은 책임의 양에 비추어 볼 때, 우리 복지관은 공정한 보상을 해준다" 라고 하는 것이 조직수준의 분배공정성에 대한 개

인의 인식을 보다 분명하게 측정한다고 비평할 수 있다. 이러한 지적은 본 연구의 제한점으로서 향후, 완전히 '집단' 참조(referent-shift)된 문항에 근거한 반복 연구의 필요성을 제기한다. 한편 '나' 참조와 '우리' 참조가 혼합된 문항이 분석에 사용된 본 연구의 결과는 다수준 모형 분석을 위한 문항설계에 있어 새로운 통찰을 제공해준다. 다수준 분석시, 완전 개인 참조 문항은 상층에서 요인이 수렴하지 않을 가능성이 높고, 완전 집단 참조 문항은 하층에서 요인이 수렴하지 않을 가능성이 높다. 다수준 분석에서 동질상향적 구성개념은 개인의 지각을 바탕으로 집단 간 차이를 보이는 개념을 의미하기 때문에, 개인의 지각이 뒷받침 되지 않는 상층에서의 개념은 의미가 없다. 또한, 개인의 지각만 있고 공유된 지각과 인식에 의한 상층 수준의 개념이 식별되지 않으면 아무 소용이 없다. 그러므로 측정대상인 개념에 대해 참조점을 달리한 두 자료간에 측정동일성이 성립한다는 검증이 있으면, 추후의 연구들은 집단 참조 문항을 '나' 참조와 '우리' 참조를 혼합하여 내용을 구성하는 것이 합리적이며 보다 유용할 수 있다.

그러나 본 예시의 결과를 확증하고 적용하는 데는 한계점이 있음을 밝혀둔다. 예시의 자료는 사회복지기관의 사회복지사를 표본으로 한 조직분배공정성에 대한 인식이므로, 결과를 일반화하기 위해서는 일반 기업체 조직에 근무하는 직장인의 경우에도 측정동일성이 성립하는지는 확인해볼 필요가 있다.

### 참고문헌

김아영, 차정은, 이채희, 서애리, & 최기연

(2004). 학교급간 학업적 자기조절척도의 구인동등성 검증 및 잠재평균분석. *교육심리연구*, 18(2), 227-244.

이순목. (1993). *중급 LISREL*. 한국심리학회 동계연수회 자료집.

이순목. (2010). 역량과 역량관련 프로그램의 타당화를 위한 제안. *한국 심리학회지: 산업 및 조직*, 23(3), 551-573.

이순목. (2014). 구조방정식 모형의 일반화와 차별화: 다집단 분석. *미발간*

이순목, 김한조. (2011). 구조방정식 모형의 일반화 또는 집단차 연구를 위한 다집단 분석의 관행과 문제점. *사회과학(성균관대)*, 43(1), 63-112.

이순목, 윤수철, 차정은, 김종남, & 여성철. (2012). 한국판 CPGI 와 원본척도 (CPGI) 간 측정동등성 및 점수연계 가능성. *한국 심리학회지: 임상*, 31(2), 401-425.

이희영. (2012). 심리적 자본과 조직공정성이 직무만족과 조직몰입에 미치는 영향. *성균관대학교 석사학위 논문*.

이희영, 김종규, & 이순목 (2013). 포스터 발표: 3 분과 산업 및 조직; 심리적 자본과 조직공정성이 직무만족과 조직몰입에 미치는 영향: 다수준적 (multi-level) 고찰. *한국 심리학회 연차 학술발표논문집*, 2013(1), 314.

최은하 (2012). 팀 유연성에 대한 타당도 분석: 다수준적 관점의 적용. *성균관대학교 석사학위 논문*.

최은하, & 구병모 (2013). 포스터 발표: 3 분과 산업 및 조직; 상향적으로 정의되는 팀 유연성의 타당화.

한정원, 이경수, 박찬신, & 손영우 (2009). 조중사의 안전행동을 예측하는 조직의 안전문

- 화와 개인의 안전태도 및 안전동기 간의 관계. 한국심리학회지: 산업 및 조직, 22 (1), 109-129.
- 한태영, & 김원형 (2006). 권한위임과 조직공정성이 직무효과성에 미치는 영향에 대한 다수준적(multilevel) 고찰. 인사조직연구, 14, 183-216.
- 홍세희, 황매향, & 이은설 (2005). 청소년용 여성 진로장벽 척도의 잠재평균분석. 교육심리연구, 19(4), 1159-1177.
- Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological bulletin*, 107(2), 238.
- Bentler, P. M. (1995). *EQS structural equations program manual*. Multivariate Software.
- Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological bulletin*, 88(3), 588.
- Bliese, P. D. (2000). Within-group agreement, non-independence, and reliability: Implications for data aggregation and analysis. In K. J. Klein & Steve W. J. Kozlowski(Eds.), *Multilevel theory, research, and methods in organizations: Foundations, extensions, and new directions*(pp. 349-381). San Francisco, CA, US: Jossey-Bass, xxix, 605 pp.
- Bovaird, J. A. & Koziol, N. A. (2012). Measurement Models for Ordered-Categorical Indicators. In Hoyle, R. H. (Ed.) *Handbook of Structural Equation Modeling*. NY: Guilford Press.
- 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.
- Chan, D. (1998). Functional relations among constructs in the same content domain at different levels of analysis: A typology of composition models. *Journal of Applied psychology*, 83, 234-246.
- Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling*, 14(3), 464-504.
- Chen, G., Bliese, P. D., & Mathieu, J. E. (2005). Conceptual framework and statistical procedures for delineating and testing multilevel theories of homology. *Organizational Research Methods*, 8(4), 375-409.
- Chen, G., Mathieu, J. E., & Bliese, P. D. (2004). A framework for conducting multilevel construct validation. *Research in Multi Level Issues*, 3, 273-303.
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural equation modeling*, 9(2), 233-255.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum.
- Cronbach, L. J. (1976). *Research on classrooms and schools: Formulation of questions, design and analysis*. Occasional Paper of the Stanford Evaluation Consortium, Stanford University.
- DeShon, R. P., Kozlowski, S. W., Schmidt, A. M., Milner, K. R., & Wiechmann, D. (2004). A multiple-goal, multilevel model of feedback effects on the regulation of individual and team performance. *Journal of Applied Psychology*,

- 89(6), 1035.
- Dyer, N. G., Hanges, P. J., & Hall, R. J. (2005). Applying multilevel confirmatory factor analysis techniques to the study of leadership. *The leadership quarterly, 16*(1), 149-167.
- Goldstein, I. L. (1993). *Training in organizations: Needs assessment, development, and evaluation*. Thomson Brooks/Cole Publishing Co.
- Griffin, M. A., & Mason, C. M. (2002). Group task Satisfaction : Applying the Construct of Job Satisfaction to Groups. *Small Group Research, 33*, 271-312.
- Guzzo, R. A., Yost, P. R., Campbell, R. J., & Shea, G. P. (1993). Potency in groups: Articulating a construct. *British journal of social psychology, 32*(1), 87-106.
- Heck, R. H. (2001). Multilevel modeling with SEM. New developments and techniques in structural equation modeling, 89-127.
- James, L. R. (1982). Aggregation bias in estimates of perceptual agreement. *Journal of Applied Psychology, 67*, 219-229.
- James, L. R., Demaree, R. G., & Hater, J. J. (1980). A statistical rationale for relating situational variables and individual differences. *Organizational Behavior and Human Performance, 25*(3), 354-364.
- James, L. R., Demaree, R. G., & Wolf, G. (1984). Estimating within-group interrater reliability with and without response bias. *Journal of applied psychology, 69*(1), 85.
- James, L. R., & Jones, A. P. (1974). Organizational climate: A review of theory and research. *Psychological bulletin, 81*(12), 1096.
- Jones, A. P., & James, L. R. (1979). Psychological climate: Dimensions and relationships of individual and aggregated work environment perceptions. *Organizational behavior and human performance, 23*(2), 201-250.
- Jöreskog, K. G. (1971). Simultaneous factor analysis in several populations, *Psychometrika, 36*(4), 409-426.
- Kenny, D. A. (2010). Fit measures. <http://davidakenny.net/cm/fit.htm>.
- Kozlowski, S.W. & Hattrup, K. 1992. A disagreement about within-group agreement: Disentangling issues of consistency versus consensus. *Journal of Applied Psychology, 77*: 161-167.
- Kozlowski, S. W. J., Gully, S. M., Nason, E. R., Ford, J. K., Smith, E. M., Smith, M. R., & Futch, C. J. (1994). A composition theory of team development: Levels, content, process, and learning outcomes. In *JE Mathieu (Chair), Developmental views of team process and performance. Symposium conducted at the ninth annual conference of the Society for Industrial and Organizational Psychology, Nashville, TN.*
- Kozlowski, S. W., & Hults, B. M. (1987). An exploration of climates for technical updating and performance. *Personnel Psychology, 40*(3), 539-563.
- Kozlowski, S. W., & Klein, K. J. (2000). A Multilevel approach to theory and research in organizations: Contextual, Temporal, and Emergent Processes. In *K. J. Klein & S. W. J. Kozloeski (Eds.), Multilevel Theory, Research, and Methods in Organizations*. San Francisco: Jossey-Bass.

- Kozlowski, S. W., & Salas, E. (1997). A multilevel organizational systems approach for the implementation and transfer of training. *Improving training effectiveness in work organizations*, 247, 287.
- Kozlowski, S.W.J. (1998, March). Extending and elaborating models of emergent phenomena. Presentation at MESO Organization Studies Group, Arizona State University, Tempe.
- Kozlowski, S.W.J. (1999, April). A typology of emergence: Theoretical mechanisms understanding bottom-up phenomena in organizations. In F. P. Morgeson & D.A. Hofmann (Chairs), *New perspectives on higher-level phenomenon in industrial/organizational psychology*. Symposium conducted at the fourteenth annual conference of the Society for Industrial and Organizational Psychology, Atlanta, GA.
- LeBreton, J. M., & Senter, J. L. (2008). Answers to 20 questions about interrater reliability and interrater agreement. *Organizational Research Methods*, 11, 815-852. doi: 10.1177/1094428106296642
- Lee, B. H. (2003). An Empirical Study of Organizational Commitment: A Multi-Level Approach. *Journal of Behavioral and Applied Management*, 4(3), 176-188.
- Maas, C. J., & Hox, J. J. (2005). Sufficient sample sizes for multilevel modeling. *Methodology*, 1(3), 86-92.
- Marsh, H. W., Lüdtke, O., Robitzsch, A., Trautwein, U., Asparouhov, T., Muthén, B., & Nagengast, B. (2009). Doubly-latent models of school contextual effects: Integrating multilevel and structural equation approaches to control measurement and sampling error. *Multivariate Behavioral Research*, 44(6), 764-802.
- McGehee, W., & Thayer, P. W. (1961). Training in business and industry.
- Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika*, 58(4), 525-543.
- Moore, D. S. (2010). *The basic practice of statistics*. (5th Ed.). NY: Freeman & Company.
- Moorman, R. H. (1991). Relationship between organizational justice and organizational citizenship behaviors: do fairness perceptions influence employee citizenship?. *Journal of applied psychology*, 76(6), 845.
- Morgeson, F. P., & Hofmann, D. A. (1999). The structure and function of collective constructs: Implications for multilevel research and theory development. *Academy of Management Review*, 24(2), 249-265.
- Muthén, B. O. (1991). Multilevel factor analysis of class and student achievement components. *Journal of Educational Measurement*, 28(4), 338-354.
- Muthén, B. O. (1994). Multilevel covariance structure analysis. *Sociological methods & research*, 22(3), 376-398.
- Muthén, L. K., & Muthén, B. O. (1998). *Mplus [computer software]*. Los Angeles, CA: Muthén & Muthén.
- Muthén, B. O., & Satorra, A. (1989). Multilevel aspects of varying parameters in structural models. *Multilevel analysis of educational data*, 87-99.
- Muthén, B. O., & Satorra, A. (1995). Technical

- aspects of Muthén's LISCOMP approach to estimation of latent variable relations with a comprehensive measurement model. *Psychometrika*, 69(4), 489-503.
- Myers, N. D., Feltz, D. L., & Short, S. E. (2004). Collective Efficacy and Team Performance: A Longitudinal Study of Collegiate Football Teams. *Group Dynamics: Theory, Research, and Practice*, 8(2), 126.
- Myers, N. D., Payment, C. A., & Feltz, D. L. (2004). Reciprocal Relationships Between Collective Efficacy and Team Performance in Women's Ice Hockey. *Group Dynamics: Theory, Research, and Practice*, 8(3), 182.
- Niehoff, B. P., & Moorman, R. H. (1993). Justice as a mediator of the relationship between methods of monitoring and organizational citizenship behavior. *Academy of Management Journal*, 36(3), 527-556.
- Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (Vol. 1). Sage.
- Schmitt, N., & Kuljanin, G. (2008). Measurement invariance: Review of practice and implications. *Human Resource Management Review*, 18(4), 210-222.
- Schweig, J. (2013). *Testing the assumption of cross-level measurement invariance in multilevel models: Evidence from school and classroom environment surveys* (CRESST Report 829). Los Angeles, CA: University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST).
- Sirotnik, K. A. (1980). Psychometric implications of the unit of analysis problem (with examples from the measurement of organizational climate). *Journal of Educational Measurement*, 17(4), 245-282.
- Snijders T. A. B., & Bosker RJ (1999). *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling*. Sage Publications, London.
- Sörbom, D. (1974). A general method for studying differences in factor means and factor structure between groups. *British Journal of Mathematical and Statistical Psychology*, 27(2), 229-239.
- Stapleton, L. M. (2006). Using multilevel structural equation modeling techniques with complex sample data. In Hancock, G. R. & Mueller, R. O(Eds.) , *Structural equation modeling: A second course*. pp.345-383. Greenwich, CON: Information age publishing.
- Steiger, J. H., & Lind, J. C. (1980, May). Statistically based tests for the number of common factors. In *annual meeting of the Psychometric Society, Iowa City, IA* (Vol. 758, pp. 424-453).
- Tay, L., Woo, S. E., & Vermunt, J. K. (2014). A Conceptual and Methodological Framework for Psychometric Isomorphism Validation of Multilevel Construct Measures. *Organizational Research Methods*, 17(1), 77-106.
- 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.
- Yoon, M., & Millsap, R.E. ( 2007). Detecting violations of factorial invariance using data-based specification searches: A Monte Carlo study. *Structural Equation Modeling*, 14.

435-463.

Zyphur, M. J., Kaplan, S. A., & Christian, M. S. (2008). Assumptions of cross-level measurement and structural invariance in the analysis of multilevel data: Problems and solutions. *Group Dynamics: Theory, Research, and Practice*, 12(2), 127.

1차 원고접수 : 2018. 04. 24

2차 원고접수 : 2019. 01. 08

최종게재결정 : 2019. 01. 19



## Multigroup analysis in the context of multilevel analysis to judge if a referent-shift model is needed for measurement of composition emergence

Jungwon Ahn<sup>1)</sup>

Soonmook Lee<sup>2)</sup>

Hyunjoo Jung<sup>3)</sup>

Hoeyeong Lee<sup>4)</sup>

<sup>1)</sup>Fordham University

<sup>2)</sup>Sungkyunkwan University

<sup>3)</sup>University of Massachusetts Amherst

<sup>4)</sup>Sungkyunkwan University

This study demonstrates how a multigroup analysis approach is used in the analysis of multilevel data to judge if a referent-shift consensus model is needed to measure a compositional property. A compositional property in multilevel context means that the forms of emergence from individual levels to group levels are isomorphic as individuals interact, communicate perspectives, and iteratively construct a common interpretation, so that all individuals in the collective are similar. The measurement principle for conceptualization of multilevel compositional properties is to use a referent-shift consensus model proposed by Chan(1998). However, if the researcher wants to use the same construct in individual levels as well as in group levels, she needs to administer the same items to the same individuals again with a change of reference from “group” to “individual”. It sounds bothersome and creates difficulties in reality. For that reason, researchers often collect data from individuals using self-referenced items, aggregate, and then use the aggregate scores as measures of group level variables. However in these cases, measurement invariance is tacitly assumed across the individual and group levels. We pointed out the problems of this unjustified assumption in analyses of multilevel data, and presented an analytic procedure to test the assumption using multigroup analysis framework. In sum, if measurement invariance across levels is established, researchers can use either a self-referent or a referent-shift data at individual levels and aggregate data at group levels without dual measurement. Moreover, in such a case, using a referent-shift data(‘we’ data) is more appropriate in light of construct validity because of its higher possibility to reveal group effects. If the measurement invariance across levels is not supported, researchers should collect data separately for individual level variables with a self-reference items and for group level variables with group-referenced items.

*Key words* : multilevel analysis, emergent property, referent-shift consensus model, multigroup analysis, measurement invariance, multilevel structural equation modeling



This work is licensed under a Creative Commons Attribution 4.0 International License.