상관행렬의 구조분석에서 집단평균차이의 효과: 요인분석기법을 중심으로[1)]

The effect of group mean differences upon factor analysis

김청택[2)], 이소영[3)]
Cheongtag Kim, Soyoung Lee

요약

이 논문의 목적은 상관분석에서 집단차 변수를 무시하였을 때 자료에 대한 잘못된 해석을 유도할 수 있음을 보여주고, 집단차를 고려한 분석의 중요성과 그 방법을 제시하는 것이었다. 연구 1은 시뮬레이션 연구로 상관구조에 대한 분석인 요인분석에서 집단차를 무시하면 자료가 지니고 있던 요인구조를 파악하지 못함을 보여주었다. 이에 대한 대비책으로 표준점수에 의한 자료의 변환방법과 공분산 구조모형의 집단분석을 이용하는 방법 등이 제시되었다. 연구 2는 사례연구로 실제 자료에서 집단의 평균차에 의한 효과가 발생하는지를 지능검사 자료를 이용하여 예증하고 이러한 문제점을 해결할 수 있는지를 보여주었다.

The purpose of this study is (1) to demonstrate that ignoring group differences may mislead to incorrect conclusion when analyzing correlation data (e.g., factor analysis), (2) to highlight the importance of the data analytic method considering group differences, and (3) to provide ways of incorporating group differences in data analysis. In study 1, ignoring group difference in factor analysis may mislead to incorrect factor structure. To remedy this, z-transform method and group analysis tool in covariance structure models were suggested. In study 2, the group differences effect was illustrated by using real data (IQ test data).

---

* 이 연구는 과학기술부의 뇌과학 연구사업의 지원을 받았음
2) 서울대학교 심리학과, 대학원 인지과학 협동과정
3) 서울대학교 심리학과

I. 서론

본 논문은 상관행렬을 분석하는 요인분석기법에서 자료에 존재하는 집단간의 평균 차이를 무시하였을 때 발생할 수 있는 문제점을 지적하고 이를 해결할 수 있는 방법들을 제시하였다. 심리학을 포함하는 사회과학 연구에서는 많은 경우에 모든 사람들에게 혹은 모든 집단에게 공통적으로 일어나는 현상들에 대하여 관심을 가진다. 예컨대, 부모의 사회경제적인 지위 (Social Economic Status, 이하 SES)가 자녀의 지능에 영향을 미치는지를 조사하고자 하는 경우를 생각하여 보자. 이때에는 남녀에 따른 SES의 효과의 차이를 반영하거나 지역에 따른 SES의 효과의 차이를 반영하는 통계적 기법을 사용하기보다는 지역과 성별을 고려하지 않는 통계적 기법을 주로 사용한다. 회귀분석이나 분산분석들이 이러한 분석방법의 대표적인 예들이 될 것이다. 특히 심리학의 연구에서는 모든 사람들에게 공통적으로 일어나는 심리적인 과정을 연구하는 것을 목적으로 두기 때문에 많은 경우에 집단의 차이 즉 분석단위의 차이는 무시되는 경향이 있다. 단순히 평균의 관계에만 관심을 가지고 우리가 가정하는 통계적인 모형이 선형적일 경우에는 집단의 자료를 집합시켜 평균치를 내면 집단차가 상쇄되어 집단과 무관한 평균 경향성을 찾아낼 수 있다. 그러나 상관이나 공분산 모형과 같이 비선형적인 통계 모형에서는 집단차를 무시하는 분석법은 자료에 대한 잘못된 해석을 유도할 수 있다. 본 연구에서는 조사연구에서 많이 사용되는 요인분석의 기법에서 집단차를 무시하였을 때에 발생할 수 있는 오류와 그 해결책을 제시한다.

I.1. 집단차를 무시하였을 때 발생하는 오류

여러 집단으로 구성된 자료에서 각 개인이 소속된 단위를 무시하고 집단을 모두 통합하여 분석하면 해석의 오류를 낳고 잘못된 결론을 유도할 수 있는 예들로 Simpson의 역설과 상관에서 나타나는 집합의 오류(aggregation error)를 들 수 있다. 먼저, Simpson의 역설(Simpson, 1951)은 발생빈도로부터 특징간 연합을 측정하고자 할 때 집합의 오류가 나타나는 현상을 말한다. 즉 두 개 이상의 수반성 표 (contingency table)를 합치면 원래 표에서 나타난 변수들 간의 관계가 사라지거나 역전될 수 있다. 구체적은 예는 다음과 같다. 표 1은 L과 H의 두 부서가 있는 회사의 고용에 남녀 차별이 있는지를 알아보기 위해 성별에 따른 고용률을 조사한 자료이다(Paik, 1985). 표에서는 부서 L과 부서 H의 남녀별 고용자 수 및 탈락자 수, 그리고 두 부서를 합친 전체 회사의 남녀별 고용자 수 및 탈락자 수가 제시되어 있다.

표 1. 부서 L의 남녀별 고용자/탈락자 수

	부서 L		부서 H		전체 회사	
	남성	여성	남성	여성	남성	여성
고용자 수	550	1250	2950	800	3500	2050
탈락자 수	1450	2750	1050	200	2500	2950
고용률	27.5%	31.25%	73.75%	80%	51.3%	41%

표에서 나타난 바와 같이, 부서 L과 H 모두에서 여성의 고용률이 남성의 고용률보다 통계적으로 유의하게 높은 반면(부서 L: χ^2= 9.18, $p < .01$, 부서 H: χ^2 = 149.06, $p < .01$), 두 부서를 통합한 전체 회사의 자료에서는 오히려 남성고용률이 여성 고용률보다 통계적으로 유의하게 높다(χ^2 = 338.00, $p < .01$). 따라서 이 자료에서는 부서를 무시한 분석은 모든 부서의 특성을 반영하지 않는 방식으로 자료에 대한 결론을 내리게 된다.

상관에서 집합의 오류란 모든 집단들이 동일한 상관관계를 가지고 있는데도 불구하고 집단을 합쳐서 분석하면 각 집단이 가지고 있는 변수들 간의 상관관계가 사라지거나, 혹은 반대의 상관관계를 산출하는 현상을 말한다(Nunnally & Bernstein, 1994). 예컨대, 이질적인 집단을 합쳐서 Pearson의 적률상관계수(r)를 구하면 각 집단에서 나타난 변인관계와 무관한 무의미한 상관계수가 나타날 수 있다 (Pearson, Lee, & Bramley-Moore, 1899; Wendt, 1976). 상관에서 집합의 오류가 발생하면 Simpson의 역설과 마찬가지로 변수들간 관계성을 잘못 해석하게 한다. 즉 집단을 무시하고 두 연속변인 x_1과 x_2의 상관을 구하면 각각의 집단에서 나타나는 상관과 무관하거나 역전되는 상관을 산출할 수가 있는 데 이를 가상관(spurious correlation) 이라고 한다 (Wendt, 1976). 표 2에 x1과 x2를 언어능력과 수리능력을 나타내는 변수라 하고, A와 B의 두집단에 있다고 가정하자. 집단 A, 집단 B에서 x1과 x2의 상관계수는 각각 +.62, +.38로 정적 상관이 나타났으나, 두 집단의 자료를 합쳐서 상관계수를 구하면 -.44로 부적상관을 보인다(Nunnally & Bernstein, 1994). 즉 집단을 무시하고 분석하면, 실제로는 언어능력과 수리능력이 정적인 상관관계에 있는데도 불구하고 부적인 상관이 있다고 결론을 내릴 것이다.

표 2 집단 A와 B에서 x_1, x_2의 분포

집단 A		집단 B	
x1	x2	x1	x2
18	11	10	23
19	14	16	24
32	15	19	27
37	22	21	24
24	12	20	29
34	13	14	21
28	11	14	24
31	19	16	27
25	8	19	21
30	14	20	25

Sockloff(1975)에 의하면, 두 개의 하위집단이 있을 때 변수들간의 집단 평균 차이 패턴에 따라, 두 집단 (하위집단 1 과 하위집단 2)을 합쳤을 때 상관의 왜곡이 3가지 양상으로 나타날 수 있다고 했다. (1) 두 변수의 평균이 모두 하위집단 1에서 더 클 때 ($\overline{X_1} > \overline{X_2}$; $\overline{Y_1} > \overline{Y_2}$)는 평균의 집단차가 커질수록 전체집단의 상관이 양의 방향으로 증가하고, (2) 한 변수의 평균은 하위집단 1에서 크고, 다른 변수의 평균은 두 집단에서 동일할 때($\overline{X_1} > \overline{X_2}$; $\overline{Y_1} = \overline{Y_2}$)는 집단차가 클수록 전체 상관이 0에 접근하며, (3) 한 변수의 평균은 하위집단 1에서 크고, 다른 변수의 평균은 하위집단 2에서 클 때($\overline{X_1} > \overline{X_2}$; $\overline{Y_1} < \overline{Y_2}$)는 집단차가 커질수록 전체상관은 음의 방향으로 증가한다. 앞에서 살펴본 예들은 모두 Sockloff의 분류 가운데 (3)에 해당된다.

특히 모집단의 구조가 위계적일 때, 즉 학생과 학과와 같이 하나의 분석단위에 다른 단위가 내포될 때, 분석단위를 고려하지 않은 상관 분석에서는 집합의 오류가 발생할 확률이 높다. 한 예가 각 대학교에서 수능성적, 면접점수와 학업수행을 반영하는 학점평균간의 관계를 연구할 때 가장 많이 보고되는 면접과 학점의 상관이 수능과 학점의 상관보다 더 높다는 것이다. 그러나 이러한 결과는 분석단위 즉 집단을 무시하여서 나타나는 결과일 가능성이 크다. 표 3은 A, B, C, D 네 학과에 있는 학생 16명의 수능성적, 면접점수, 학점에 대한 가상적인 자료이다. 이 자료는 하위수준의 단위인 학생이 상위수준의 단위인 학과에 속하는 위계적 구조를 갖는다. 이 자료를 통해 학점에 대한 수능시험의 예언 타당도를 조사하기 위해서 예언변수인 수능성적과 준거변인인 학점간의 상관을 구하면 각 학과별 상관은 모두 1.00이지만, 학과라는 분석단위를 무시하고 전 학과를 합쳐서 개인을 단위로 분석하면 전체 학생 16명의 수능성적

과 학점간의 상관, 즉 타당도는 .181로 낮아진다. 반면 각 집단간의 평균차가 존재하지 않는 면접의 경우는 각 학과 단위의 분석에서도 상관이 1.00이고, 집단을 고려하지 않는 분석에서도 상관이 1.00이다. 따라서 집단간의 평균차가 존재할 때 이를 무시한 분석은 잘못된 해석을 내릴 수 있음을 시사하여 준다.

표 3. 네 학과 소속 학생 16명의 수능성적, 면접점수, 학점에 대한 가상적 자료

학과	수능성적	면접	학점
A	300	2	D
A	305	4	C
A	310	6	B
A	315	8	A
B	330	2	D
B	335	4	C
B	340	6	B
B	345	8	A
C	360	2	D
C	365	4	C
C	370	6	B
C	375	8	A
D	380	2	D
D	385	4	C
D	390	6	B
D	395	8	A

또 다른 예로는 심리학에서 오랫동안 진행된 출생순위와 지능간의 상관을 통해서 출생순위효과를 검증하는 연구들을 들 수 있다. 많은 연구들이 출생순위와 지능간의 상관관계에 대하여 서로 상반되는 결과를 내놓았다 (Rogers, Cleveland, van den Oord, & Rowe, 2000; Zajonc & Mullally, 1997). Zajonc와 Mullally는 이러한 상반된 결과가 분석방식에 따른 결과로 해석하고 있는데, 위에서 논의한 바와 같이 가정이라는 상위단위를 무시한 분석에 기인할 가능성이 있다.

요약하면 자료에 있는 위계적 구조를 무시하고 전체 자료를 동질적인 것으로 파악하여 전체 상관(total correlation)을 구하게 되면 무의미한 상관이 생길 수 있고, 이로 인해 변수들 간의 관계에 대하여 잘못된 결론에 도달할 수 있다. 그러나 실제 분석에서는 집단차를 반영하는 분석법이 그리 많이 사용되는 것 같지는 않다. 따라서 집단간의 차이를 반영하는 분석법의 도입이 필요할 것이다. 특히 집단차를 반영하는 분석법은 평균자료보다는 공분산이나 상관의 자료에서 더 요구된다. 평균모형에서는 관찰된 점수는 각 집단의 평균점수와 그 평균에서는 편차점수로 구성된다. 즉

$$X_{i=}\overline{X}_g+e_i$$

여기서 X_i는 관찰 점수, $\overline{X}_g$는 집단의 평균, 그리고 e_i는 관찰점수와 집단평균의 차이를 나타낸다. 이 모형에서 집단의 성질을 무시하고 분석하더라도, X_i들의 평균은 전체 평균이 된다. 즉

$$E(X_i)=E(\ \overline{X}_g+e_i)=E(\ \overline{X}_g)+E(e_i)=\overline{X}$$

그러나 공분산의 경우는 선형적인 모형이 아니기 때문에 집단을 무시하고 분석하면 단순히 전체 집단의 성질을 반영하고 있는 특성이 나타나지 않는다. 이러한 맥락 속에서 이 논문에서는 상관계수 내지는 공분산에 대한 분석에 가장 많이 사용되는 기법 중에 하나인 요인 분석에 중점을 두었다.

I.2. 요인분석: 상관행렬의 구조에 대한 분석

요인분석 (factor analysis) 기법은 심리검사이론에서 지능검사를 개발하는 과정에서 발달된 기법으로 이론적인 요인구조를 가정하고 관찰된 자료가 요인구조에 적합한지를 검정하는 기법이다. Spearman (1904)이 지능을 일반능력 요인과 특수능력 요인으로 구분하고 이를 통계적으로 모형화한 데에서부터 요인분석의 기법이 발달하기 시작하여 Thurstone (1931)의 공통요인 모형의 제안으로 현대적 의미의 요인분석기법이 정립되었다. 요인분석은 주성분 분석과는 달리 요인이라는 잠재변수(latent variable)를 가정하고 있다. 이 잠재변수는 실제로 관찰되지는 않지만 우리의 연구의 대상이 되는 추상적인 개념 (예컨대 지능)을 나타낸다. 이 잠재변수 (지능)은 직접적으로 관찰되지는 않지만 지능점수라는 관찰변수를 예측하는 변수로 작동하고 있다. 이를 구체적으로 요인분석의 자료모형으로 표현하면 다음과 같다.

$$x_{ij}=\mu_j+\lambda_{j1}z_{i1}+\lambda_{j2}z_{i2}+\ldots+\lambda_{jm}z_{im}+u_{ij} \qquad (1)$$

여기서 x_{ij}는 i 번째 개인의 j번째의 점수이고, λ_{jk}는 j번째 변수의 k번째 요인에 대한 부하량이며, z_{ik}는 i번째 개인의 k번째의 요인점수이고 마지막으로 u_{ij}는 독특요인 (unique factor)으로 다른 (공통)요인에 의하여 설명되지 않는 부분을 나타낸다. 자료모형은 회귀분석과 동일한 형태를 띠고 있지만, 요인점수 $z_{i.}$들이 실제로는 관찰되지 않는 가설적인 잠재변수라는 점에서 다르다. 또한 잠재변수가 가설적인 변수라는 점에서 관찰변수들의 선형적 결합으로 성분 (component)을 표시하는 주성분 분석과도 다르다. 요인분석에서는 수식 (1)의 자료모형에서 공분산구조를 유도하고 관찰된 공분산 내지는 상관행렬이 이 공분산 구조를 따르는지를 평가함으로써 요인분석모형의 타당성을 판단한다. 즉 (1)을 행렬식으로 표현하면 다음과 같고,

$$\boldsymbol{x} = \mu + \Lambda \boldsymbol{z} + \boldsymbol{u} \tag{2}$$

요인점수(z)와 독특 요인(u)의 상관이 0이고 u들간의 상관이 0이라는 가정 하에서 x들간의 공분산 행렬(Σ)는 다음과 같이 유도된다.

$$\boldsymbol{\Sigma} = \Lambda \Phi \Lambda' + \boldsymbol{D}_{\Psi} \tag{3}$$

여기서 $\boldsymbol{\Phi}$는 요인들간의 점수를 나타내고 $\boldsymbol{D}_{\Psi}$는 대각행렬로 독특 요인들 간의 공분산을 나타낸다. (3)의 $\boldsymbol{\Sigma}$에 관찰된 공분산 행렬을 적합시키고, 파라미터를 추정함으로써 요인구조를 파악하게 된다.

그러나 각 집단마다 점수의 평균이 다를 경우에는 식 (3)의 모형에 의해 자료가 기술될 수 없다. 즉 집단차가 있을 때와 집단차가 없을 때 자료모형에서 유도되는 요인분석의 통계적 모형, 즉 공분산모형이 동일하지 않다. 바꾸어 말하면, 집단차를 가정하는 모형과 가정하지 않는 모형은 통계적으로 다른 모형이며, 집단차가 존재할 때 그 차이를 무시하는 분석을 하는 것은 잘못된 통계모형을 적용시키는 것이다. 집단을 고려하지 않은 분석에서는 자료의 공분산 구조는 (3)과 동일하나, 집단차를 고려하는 공분산구조 (Σ_M)는 다음과 같다.

$$\begin{aligned} \Sigma_M &= \Sigma + \sum_{g=1}^{G} w_g (\mu_g - \mu_M)(\mu_g - \mu_M)' \\ &= \Lambda(\Psi + \sum_{g=1}^{G} w_g z_g z_g')\Lambda' + D_{\Psi} \end{aligned} \tag{4}$$

여기서 g는 집단을 나타내고 w_g는 각 집단에 주어지는 가중치로 각 집단의 표본크기의 역수가 그 한 예이다. 식 (4)는 식 (3)과는 달리 집단의 평균차이가 공분산구조에 체계적으로 반영되어 있음을 알 수 있다. 따라서 집단차가 존재하는 자료는 식 (4)와 같은 공분산 구조를 발생시킬 것이고, 이때 식 (3)에 자료를 적합시키면, 잘못된 결론을 유도할 가능성이 있다. 따라서 집단차가 존재하는 경우에는 집단차를 고려한 분석이 필요하다.

집단차를 고려하여 분석할 수 있는 방안으로 크게 두 가지를 생각하여 볼 수 있다. 첫째는 각 집단별로 표준점수를 계산한 다음 그 표준점수를 이용하여 집단을 결합하는 것이다. 이때에는 모든 집단의 평균이 일치됨으로 집단간의 평균차이에 의해 발생하는 오류들은 없어질 것이다. 즉 식 (4)에서 각 집단의 평균이 일치하는 경우에 공분산행렬은 다음과 같아질 것이다.

$$\begin{aligned} \Sigma_M &= \Sigma + \sum_{g=1}^{G} w_g (\mu_g - \mu_M)(\mu_g - \mu_M)' \\ &= \Sigma + 0 = \Sigma \end{aligned} \tag{5}$$

따라서 집단차에 의한 문제를 발생하지 않을 것이다.

두 번째는 식 (3)에 해당하는 모형을 사용하는 것이다. 직접적으로 식 (3)에 해당하는 모형은 MPLUS등의 프로그램에 의해 쉽게 적용시킬 수 있으나, 여기에서는 보다 많이 사용되는 기법인 공분산 구조모형에서 집단분석방법을 사용하였다. 공분산 구조모형에서 집단분석을 이용하는 방법은 각 집단별로 요인의 구조를 따로 적합시키고 각 집단별로 적합도(likelihood)를 계산하여 적합도의 합이 최대가 되도록 요인의 구조를 찾아내는 방식이다. 예컨대 두 집단이 있다면, 아래와 같이 각각의 집단에 대하여 모형을 만들고 이 두 모형을 동시에 적합시키는 방법이다.

$$\Sigma_1 = \Lambda_1 \Phi_1 \Lambda_1' + D_{1\psi}$$
$$\Sigma_2 = \Lambda_2 \Phi_2 \Lambda_2' + D_{2\psi} \quad (6)$$

LISREL(Joreskog & Sorbom, 1994), AMOS (Arbuckle & Wothke, 1999), EQS (Bentler, 1989)등의 공분산구조모형의 패키지가 이 방법을 포함하고 있어서 직접적으로 적용시키는 것이 가능하다. 다만 탐색적 요인분석의 모형이 수학적으로 확인되지 않는 모형 (unidentified)이기 때문에 모형을 지정할 때 요령이 필요하다. 모형이 확인되기 위해서는 m(m+1)/2 (m의 요인의 개수)의 추가적인 제약이 필요하다. 따라서 첫 번째 요인에서 나가는 요인부하량 중의 하나가 1로 고정되어야 하고, 두 번째 요인에서 나가는 요인부하량 중 두 개가 1로 고정되어야 하며, 이러한 방식으로 계속 진행하여 m번째 요인에서 나가는 요인부하량 중에 m개가 1로 고정되게 하면 모형이 확인되고 탐색적 요인분석을 공분산 구조모형의 맥락에서 행할 수 있다. 그런 다음, 부하량을 회전시키면 최종적으로 원하는 결과를 얻을 수 있다.

II. 연구 1: 시뮬레이션 연구

이 연구에서는 특성이 다른 집단들을 집합시켜 요인분석의 기법을 적용시킬 때 발생할 수 있는 문제점들을 예시하고, 그 해결책으로 제시된 방법들이 집단차의 문제를 해결할 수 있는지를 시뮬레이션 연구를 통해 검증하였다. 여기에서 사용된 가상적인 자료는 지능검사의 자료로 연령집단이나 성별 등을 무시하고 요인분석을 할 때의 결과를 예증하는 것이었다. 이와 유사한 분석의 예는 심리측정의 영역 뿐만 아니라, 사회과학 연구의 여러 영역에서 쉽게 발견될 수 있다.

본 연구에서는 요인분석에서 많이 사용되어지는 Holzinger자료에 해당하는 원자료

를 인위적으로 생성한 다음, 집단차를 발생시키고, 그 집단차가 요인분석에 어떠한 영향을 미치는지를 살펴보았다. 또한 제안된 집단차를 반영하는 분석법이 원래의 요인구조를 찾아낼 수 있음을 보여주고자 하였다.

II.1. 방법

(1) 자료의 생성

모집단을 형성하는 자료는 9개의 하위 지능검사점수들간의 상관을 보여주는 요인분석 문헌에서 자주 인용되는 Holzinger의 미발표 논문(Harman, 1960)에서 제시된 상관계수행렬을 이용하였다. Kaiser의 방식 (Kaiser & Dickman, 1962)을 사용하여, Holzinger의 상관계수에서 가상적인 원점수를 산출하였다. 먼저 동일한 상관계수 행렬을 가지는 두 세트의 원점수가 생성되었고 (각 세트 당 N=100, 각 세트의 자료는 아래에 제시된 평균과 표준편차를 가지도록 변환되어 각각 집단 1의 자료와 집단 2의 자료로 사용되었다.

	평균 (표준편차는 5로 동일)								
변수	단어 의미	문장 완성	이상한 단어	혼합 대수	나머지 계산	빠진 숫자	장갑	신발	손도끼
집단 1	30	40	50	60	70	60	50	40	30
집단 2	50	40	30	80	70	80	30	40	50

(2) 분석절차

먼저 두 집단 각각에 대하여 요인분석을 실시하여 Holzinger의 요인구조를 산출해 내는지를 검증하였으다. 두 번째로, 집단을 무시하고 두 집단의 자료를 결합(aggregate)시킨 후에 요인분석을 실시하였다. 마지막으로 집단간 차이를 반영하는 분석을 하였다. 이 분석법은 집단별로 표준점수를 계산한 다음 표준점수에 대한 상관행렬을 계산하여 요인분석을 하는 방법과 AMOS의 집단간 분석이 사용되었다. 요인분석은 최대우도법 (Maximum Wishart Likelihood)으로 요인을 적합시킨 다음, VARIMAX 회전법으로 요인부하량을 회전시켰다.

II.2. 결과 및 논의

(1) 각 집단에 대한 분석

각 집단에 대한 요인분석에서, 고유치 (eigen value)가 1보다 큰 개수를 요인의 수로 삼는 기준 (Kaiser의 기준)으로 요인 수를 결정하였을 때 원래 자료가 가지고 있는 3요인 구조를 산출하였다. 표 4에서 제시된 요인 부하량의 결과에서도 알 수 있듯이 해석 가능한 세 요인을 산출하였다. 집단1과 집단2가 동일한 상관행렬을 가지도록 자료가 구성되었으므로 집단에 따라서 요인분석의 결과가 동일하였다.

표 4. 집단 1과 2의 요인부하량 (집단1과 집단2의 요인부하량이 동일함)

	요인 1	요인 2	요인 3
단어의미	.946	-.011	-.023
문장완성	.819	-.038	.160
이상한 단어	.875	.094	.005
혼합 대수	-.010	-.036	.972
나머지찾기	.013	.042	.895
빠진 숫자	.133	.040	.819
장갑	-.150	.744	.175
신발	.105	.839	.055
손도끼	.086	.886	.072

(2) 집단차를 무시한 분석

집단차를 무시하였을 때의 효과를 관찰하기 위하여 두 집단을 결합시킨 자료를 사용하여 요인분석을 하였다. Kaiser의 고유치에 따른 기준에 의해서 요인 수를 결정하면, 하나의 요인만이 산출되었고, 요인부하량이 표 5에 제시되어 있다.

표 5. 집단을 무시한 분석의 요인부하량

	요인
단어의미	.968
문장완성	.971
이상한 단어	.972
혼합 대수	.973
나머지찾기	.974
빠진 숫자	.977
장갑	.952
신발	.959
손도끼	.956

이 결과가 시사하는 바는 명확하다. 만약 집단차가 존재하는 경우 집단차를 무시하고 분석하면, 실제의 요인구조를 밝히는데 실패할 수 있다는 것이다. 위의 예에서 우리가 특정한 지능의 구조를 밝히고자 한다면, 실제로는 지능을 구성하고 있는 요인이 세 개인데, g-factor와 같은 하나의 요인으로 잘못된 결론을 내릴 수 있다.

(3) 집단차를 고려한 분석

위에서 제시한 집단별 표준점수를 활용한 전체 분석 방법과 공분산 구조모형에 의한 결과는 다음과 같다.

집단별 표준점수의 이용

집단별로 표준점수를 내고 이 점수들을 결합시켜 상관계수를 낸 다음 요인분석을 한 경우에는 요인구조를 복구해낼 수 있었다. 결과는 표4와 실질적으로 동일하였다. 이는 각 집단별로 표준점수를 구하면, 전체 상관행렬이 원래 Holzinger의 상관행렬과 동일하므로 그리 놀라운 일이 되지 못한다. 이러한 결과는 각 집단별로 표준점수를 구하여 집단들을 집합하는 것이 집단 간에 걸쳐서 공통적으로 존재하는 요인구조를 찾아낼 수 있음을 보여주고 있다.

(2) 공분산 구조모형에서 집단분석을 이용

공분산 구조모형에서 집단분석을 이용한 경우에도 각 요인 부하량은 원자료의 구조와 동일한 방식으로 산출되었다. 그러나 집단분석을 사용하지 않고 자료를 집합시켜 사용한 경우에는 자료가 3요인 요인분석 모형에 잘 적합되지 않았다. 이를 통계적으로 검증하기 위하여 집단분석을 사용하지 않은 요인분석결과와 집단분석을 사용한 요인분석을 결과를 적합도 지수로 비교하였다 (표 6참조). 경험적 준거에 의하면 GFI (Goodness of Fit Index)는 .9보다 크면 좋은 모형으로 RMSEA(Root Mean Square

Error of Approximation)는 .05보다 작으면 좋은 모형으로 여겨진다. 이 준거에 맞추어서 해석하면, 집단을 무시한 분석에서는 적합도지수는 아주 낮았는데 반하여, 집단별 분석을 한 모형은 적합도지수가 높았다. 따라서 집단을 무시할 때보다 집단을 고려한 분석이나 표준점수에 의한 분석이 적합도가 더 높음을 알 수 있다. 집단을 고려하지 않은 분석은 적합도 지수에 의해서 부적합한 모형으로 판단된다.

표6. 적합도 지수

	집단을 무시한 분석	집단차를 반영한 분석
GFI	.053	.091
RMR	40.63	5.696
RMSEA	.428	.075

이러한 결과는 집단을 고려한 분석의 필요성을 다시 한번 보여준다. 많은 연구자들이 상관의 구조를 파악하는 요인분석 기법 등에서 집단차를 무시하고 분석하고 있고, 집단차를 무시한 분석은 자료에 대한 해석을 오도할 수 있다는 결과를 고려할 때, 집단차를 반영하는 분석은 필수적일 것이다. 비록 연구자가 집단차에 관심을 가지고 있지 않고 모든 집단들에게 공통적으로 일어나는 요인구조를 연구하고자 할 때에도, 집단차를 고려하는 것은 필수적이다. 그렇지 않으면 집단들 사이에 공통적으로 존재하는 요인구조를 발견하는 데 실패할 수 있기 때문이다.

III. 연구 2: 사례연구

본 연구에서는 요인분석에서 결합의 오류가 실제 자료에서 발생할 수 있음을 예증하기 위하여 전국을 대상으로 시행한 지능검사에 대한 자료를 분석하였다. 이 자료에서 관심의 대상이 되는 집단은 연령집단이었다.

III.1. 자료

곽금주, 박혜원, 김청택 (2001)이 K-WISC-Ⅲ를 표준화하기 위하면 수집한 자료가 사용되었다. 원자료는 서울, 부산, 대구, 경기, 전라, 충청, 강원, 경상 지역의 6세이상 16세 이하의 아동 2231명을 대상으로 K-WISC-Ⅲ에 대한 검사점수가 수집되었으나, 본 연구에서는 8세, 13세, 16세의 자료들만 사용되었다. K-WISC-Ⅲ의 13개 소검사(빠진 곳 찾기, 상식, 기호 쓰기, 공통성, 차례 맞추기, 산수, 토막짜기, 어휘, 모양 맞추기, 이해, 동형 찾기, 숫자, 미로)로 구성되어 있으며, 미로 소검사를 제외한 12개의

소검사들은 다음의 네 요인을 지니고 있다(웩슬러, 2000).

언어적 이해 요인: 상식, 공통성, 어휘, 이해 소검사
지각적 조직화 요인: 빠진곳 찾기, 차례 맞추기, 토막짜기, 동형맞추기
주의 집중: 산수, 숫자
처리속도: 기호쓰기, 동형검사

III.2. 분석절차

(1) 집단을 무시한 분석

원점수에 대한 분석은 실제로 피험자들이 각 소검사에서 얻은 점수들에서 집단을 고려하지 않고 공분산을 계산한 다음 그 공분산에 의하여 확인적 요인분석을 하였다. 이때 사용된 프로그램은 AMOS 4.0이었으며, 최대우도 방법으로 모형을 적합시켰다.

(2) 집단분석기법에 의한 분석

집단차를 고려한 분석에서는 AMOS의 집단분석 기법을 사용하여 3가지 연령층의 자료를 동시에 적합시키는 방법을 택하였다. 그 외에는 집단을 무시한 분석법과 동일한 방법으로 분석하였다.

(3) 표준화된 점수에 의한 분석

원점수 대신 표준화된 점수를 사용하였다는 점만 제외하고는 집단을 무시한 분석법과 동일하였다.

III.3. 결과 및 논의

먼저 세 분석법에 대한 적합도 지수들이 표 7에 제시되어 있다. 여기에서 GFI는 클수록 좋은 모형을 나타내고 RMSEA는 작을수록 좋은 모형을 나타낸다. 이 분석에서는 집단분석, 표준점수, 집단을 무시한 분석의 순으로 적합도가 좋았다.

표 7. 적합도 지수

	집단을 무시한 분석	표준점수	집단분석법 (AMOS)
GFI	.950	.963	.928
RMR	2.043	.323	2.370
RMSEA	.069	.056	.039

보다 구체적으로 집단을 무시하였을 때와 그렇지 않을 때의 결과를 비교하기 위하여 실제로 요인 부하량과 요인들간의 상관계수를 검토해 볼 필요가 있다. 표 8-10에서 제시된 요인 부하량에서는 세 분석법에 따라서 차이는 크지 않았다. 다만 집단을 고려하지 않은 분석법이 다른 두 분석법보다 요인 부하량이 높은 특징이 있다. 반면, 요인들간의 상관계수는 분석법에 따라 많은 차이가 관찰되었다. 집단차를 고려하지 않은 분석법에 의한 결과 (표11)는 요인들간의 상관이 .9 정도로 매우 높았으나 표준점수에 의한 분석 (표 12)이나 집단분석기법을 이용한 분석법(표 13)에서는 요인들간의 상관이 그리 높지 않았다. 이러한 결과에 근거하여 요인구조를 해석하면, 집단차를 고려하지 않는 분석에 의한 결과는 네 요인보다는 일 요인을 더 지지한다. 왜냐하면, 요인들간의 상관이 높으므로 이 검사는 네 개의 공통요인으로 설명하기보다는 하나의 공통요인에 의해 설명하는 것이 더 타당하기 때문이다. 반면에 다른 두 분석법은 4요인을 지지하여 준다. KWISC의 지능검사는 이론적으로 그리고 통계적으로 4요인이 지지되는 모형이다. 따라서 집단차를 고려하지 않은 분석은 잘못된 결론을 유도할 수 있으며, 집단을 고려한 분석이 사용되어야만 정확한 자료에 대한 해석이 가능함을 보여준다.

표 8. 집단을 무시한 분석

요인	소검사	요인부하량
언어적 이해	상식	0.930
	공통	0.887
	어휘	0.952
	이해	0.868
지각적 조직화	빠진	0.803
	차례	0.773
	토막	0.892
	모양	0.762
주의 집중	산수	0.874
	숫자	0.783
처리 속도	동형	0.817
	기호	0.892

표 9. 표준점수에 의한 분석

요인	소검사	요인부하량
언어적 이해	상식	0.750
	공통	0.669
	어휘	0.824
	이해	0.722
지각적 조직화	빠진	0.566
	차례	0.547
	토막	0.715
	모양	0.648
주의 집중	산수	0.659
	숫자	0.506
처리 속도	동형	0.598
	기호	0.662

표 10. 집단분석 (AMOS)

요인	소검사	요인부하량		
		8세	13세	16세
언어적 이해	상식	0.742	0.832	0.735
	공통	0.694	0.751	0.631
	어휘	0.833	0.876	0.675
	이해	0.680	0.759	0.713
지각적 조직화	빠진	0.623	0.562	0.525
	차례	0.513	0.634	0.521
	토막	0.712	0.716	0.693
	모양	0.715	0.584	0.694
주의 집중	산수	0.595	0.724	0.718
	숫자	0.481	0.605	0.401
처리 속도	동형	0.572	0.609	0.334
	기호	0.593	0.693	0.563

표 11. 집단차를 무시한 분석의 요인들간의 상관

	언어적 이해	지각적 조직화	주의 집중	처리 속도
언어적 이해	1.000			
지각적 조직화	0.890	1.000		
주의 집중	0.961	0.907	1.000	
처리 속도	0.892	0.896	0.897	1.000

표 12. 표준점수 분석에 의한 요인들간의 상관

	언어적 이해	지각적 조직화	주의 집중	처리 속도
언어적 이해	1.000			
지각적 조직화	0.658	1.000		
주의 집중	0.876	0.743	1.000	
처리 속도	0.423	0.522	0.571	1.000

표 13. 집단분석에 의한 요인들간의 상관

8세

	언어적 이해	지각적 조직화	주의 집중	처리 속도
언어적 이해	1.000			
지각적 조직화	0.735	1.000		
주의 집중	0.987	0.893	1.000	
처리 속도	0.557	0.897	0.885	1.000

13세

	언어적 이해	지각적 조직화	주의 집중	처리 속도
언어적 이해	1.000			
지각적 조직화	0.689	1.000		
주의 집중	0.837	0.673	1.000	
처리 속도	0.431	0.452	0.413	1.000

16세

	언어적 이해	지각적 조직화	주의 집중	처리 속도
언어적 이해	1.000			
지각적 조직화	0.526	1.000		
주의 집중	0.760	0.660	1.000	
처리 속도	0.351	0.385	0.463	1.000

IV. 전체 논의

이 논문에서는 집단차가 존재할 때 이를 무시하는 통계적 기법을 사용하면 자료에 대한 해석이 오도될 수 있음을 보여주었다. 특히 평균에 근거하여 자료를 해석하지 않고 공분산이나 상관계수를 이용하여 자료를 해석하는 경우는 잘못된 해석이 일어날 가능성이 높다. 평균을 이용할 때는 집단차가 존재하더라도 여러 집단의 자료를 결합시키면 그 차이가 상쇄되지만, 상관계수를 사용할 때는 집단을 단순히 결합시키는 것은 변수들간의 관계를 변형시킬 수 있다. 연구 1과 2에서는 상관구조에 대한 분석인 요인분석기법에서 집단의 차이를 무시하면, 원래의 요인구조와는 다른 요인 구조를 산출하였음을 보여 주었으며, 집단차를 고려한 분석방법인 표준점수 변환방법이나 집단분석법을 사용한 공분산 구조 모형은 원래의 요인구조를 찾아낼 수 있음을 보여주었다. 이러한 결과는 상관구조에 대한 분석에서 존재하는 집단차를 고려하는 것은 필수적임을 시사하고 있다.

집단의 차이를 고려하는 분석의 필요성에도 불구하고, 많은 연구들에서 연구관심이 되지 않는 집단차 변수들이 고려하지 않는다. 여러 가지 이유가 있을 수 있으나 가장 큰 이유는 사회과학자들이 집단차를 고려하지 않았을 때 발생할 수 있는 문제점에 대한 인식이 없기 때문일 것이다. 사회과학자들에게 이러한 점을 자각하게 하는 것이 이 논문의 중요한 목적 중에 하나이기도 하다. 또한 이러한 자각이 있다 하더라도, 집단차를 반영할 수 있는 통계적인 기법이 사회과학자들에게 소개되지 않았기 때문에 사용할 기회를 얻지 못하였을 수도 있다. 그러나 공분산 구조모형을 이용하면 상관구조에 대한 분석들에 대한 집단차는 해결할 수 있다. 또한 LISREL, AMOS, MPLUS등과 같은 상용화된 소프트웨어들이 나와 있기 때문에 그 모형을 적용시키는 것은 그리 어려운 일이 아니다.

집단차에 의한 해석의 왜곡은 비선형적 구조를 가지는 모형에서 빈번히 일어나나, 선형모형의 경우에도 집단차를 고려하지 않으면 자료로부터 충분한 정보를 얻어낼 수 없는 경우도 많다 (예, Kim, 1998). 즉 집단차를 고려하지 않는 분석에서는 추정된 평균의 구조는 크게 왜곡되지 않으나, 집단마다 다른 평균들의 분산을 추정할 수 없다든지 상대적으로 검증력이 낮아지는 경향이 있다. 선형적 모형에서 집단차를 반영하는 기법으로 위계적 선형모형이나 다층모형으로 불리는 통계기법들이 있는데 (Bryk & Raudenbush, 1992; Goldstein, 1995; Lee & Nelder, 1996), 이러한 모형들에서는 회귀분석모형이나 일반화된 선형모형에서 각 집단마다 회귀계수가 다름을 가정하고 즉 회귀계수를 고정된 값으로 보지 않고 무선변수로 취급하여 회귀계수에 대한 분포를 추정한다. 예컨대 SES가 성적에 미치는 효과(기울기)가 집단마다 다르며, 평균 SES의 기울기와 집단차의 지표인 기울기의 분산 등을 추정할 수 있다. 따라서 이러

한 선형모형에서도 각 집단의 특성을 기술하는 통계치가 계산될 수 있으며, 또한 오류항의 크기가 작아져서 검정력을 높이는 효과가 있다.

자료에서 집단이 달라짐에 따라 집단차는 반드시 존재한다. 평균 경향성에 주로 관심을 가지는 연구들에서는 이러한 집단차를 무시하여 왔다. 그러나 상관계수와 같은 통계에서는 집단차를 무시하는 것은 잘못된 결론을 유도할 수 있으므로 집단차가 존재하는 경우에는 반드시 이를 고려하는 통계적 기법들이 사용되어야 할 것이다.

참고 문헌

곽금주, 박혜원, 김청택 2001. “한국 웩슬러 아동지능검사 표준화를 위한 예비연구.” <<한국심리학회지:발달>> 14: 43-60.

웩슬러, 2000. 곽금주, 박혜원과 박광배 역, <<WISC-Ⅲ(웩슬러 아동지능검사)>>. 도서출판 특수교육.

Arbuckle, J. L. & Wothke, W. 1999. *AMOS 4.0 user's guide*. Chicago, IL: SmallWaters Corp.

Bentler, P. M. 1989. *Theory and implementation of EQS, a strcutural equations program*. Los Angeles: BMDP Statistical Software.

Bryk, A. S., & Raudenbush, S. W. 1992. *Hierarchical linear models: Applications and data analysis methods*. Newbury Park, CA: Sage.

Joreskog, K. G., Sorbom, D. 1994. *Lisrel 8 user's guide*. Chicago: Scientific Software.

Goldstein, H. 1995. *Multilevel statistical models* (2nd ed.). New York, NY: Halsted.

Harman, M. H. 1960. *Modern Factor Analysis.* Chicago, IL: The University of Chicago Press.

Kaiser, H. & Dickman, K. 1962. "Sample and population matrices and sample correlation matrices from an arbitrary population correlation matrix." *Psychometrica*. 27: 179-182

Kim, C. 1998. *Modeling individual differences in mathematical psychology*. Unpublished doctoral dissertation, Ohio State University.

Lee, Y. and Nelder, J.A. 1996. “Hierarchical generalized linear models (with discussion).” *Journal of Royal Statistics. Society. B* 58: 619-678.

Nunnally, J. C., & Bernstein, I. H. 1994. *Psychometric Theory* (3rd ed.). New York: McGraw-Hill.

Paik, M. (985. “A graphic representation of a three-way contingency table: Simpsons paradox and correlation.” *American Statistician* 39: 53-54.

Pearson, K., Lee, A., & Bramley-Moore, L. 1899. “Genetic (reproductive) selection: Inheritance of fertility in man and of fecundity in thoroughbred racehorse.” *Philosophical Transactions of the Royal Society, Series A* 192: 257-330.

Rogers, J. L. Cleveland, H. H., van den Oord, E., & Rowe, D. C. 2000. “Resolving the debate over birth order, family size, and intelligence.” *American Psychologist* 55: 599-612.

Simpson, E. H. 1951. “The interpretation of interaction in contingency tables.”

Journal of the Royal Statistical Society, Series B 49: 467-492.

Spearman, C. "General intelligence, objectively determined and measured." *American Journal of Psychology* 15: 201-293.

Sockloff, A. L. 1975. "Behavior of the product-moment correlation coefficient when two heterogeneous subgroups are pooled." *Educational and Psychological Measurement* 35: 267-276.

Thurstone, L. L. 1931. "Multiple factor analysis." *Psychological Review* 38: 406-427.

Wendt, H. W. 1976. "Spurious correlation, revisited: A new look at the quantitative outcomes of sampling heterogeneous groups and/or at the wrong time." *Archiv-fur-Psychologie* 128: 292-315.

Zajonc, R. B., & Mullally, P. R. 1997. "Birth Order: Reconciling conflicting effects." *American Psychologist* 52: 685-699.