B

Bi-factor Analysis

Robert Gibbons
Blum-Riese Professor of Biostatistics, Departments of Public Health Sciences, Medicine, Psychiatry and Comparative Human Development, University of Chicago, Chicago, IL, USA

Synonyms

Confirmatory item factor analysis; Inter-battery factor analysis; Multidimensional item response theory

Definition

The bi-factor model is a confirmatory factor analytic model originally proposed for measurement data by Holzinger and Swineford (1937) and then generalized to the case of discrete item-response data by Gibbons and Hedeker (1992). The bi-factor restriction requires that each item load on a primary dimension of interest and no more than one secondary dimension. The secondary dimensions or subdomains can be nuisance variables such as positively and negatively worded questions (i.e., a methodologic factor) or content domains from which the items are sampled (e.g., component dimensions underlying the overall quality of one's life). When appropriate the bi-factor model provides numerous advantages over an unrestricted exploratory item factor analysis model, including rotational invariance and unlimited dimensionality. For categorical item responses, the likelihood of the model can be evaluated using two-dimensional integration regardless of the number of subdomains. The model assumes that all of the intercorrelations among the items are explained by their joint association with the primary dimension and the specific subdomain that they are a part of. The subdomains are assumed to be independent.

Description

In quality of life measurement, interest in bi-factor analysis has increased as the state of the art for scale development shifted from classical test theory to item-response theory (IRT) approaches. Historically, IRT assumed unidimensional data, that is, responses to items could be accounted for by a single attribute or random effect parameter for each subject. However, empirical Bayes and marginal maximum likelihood methods easily extend the theory to item responses that have more than one dimension. Such an extension was sketched by Bock and Aitkin (1981) and

F. Maggino (ed.), *Encyclopedia of Quality of Life and Well-Being Research*,
https://doi.org/10.1007/978-3-319-69909-7_207-2

presented more formally in Bock et al. (1988). The basic ideas follow:

Following Thurstone (1947) assume that an individual's response to a test item j is controlled by a latent variable

$$y_{ij} = \sum_{k}^{m} \alpha_{jk}\theta_{ki} + \epsilon_{ij}$$

where α_{jk} is the loading of item j on factor k, θ_{ki} is the attribute value, or score, of individual i on factor k, and ϵ_{ij} is an independent residual. According to the conventions of factor analysis, assume that all variables are standardized and that scores for different factors are uncorrelated. Assuming normality, the distributions of these parameters are y_j ~N(0, 1), θ_k ~NID(0, 1), and $\epsilon_j \sim \text{NID}\left(0, \sigma_j^2\right)$, where $\sigma_j^2 = 1 - \sum_k^m \alpha_{jk}^2$ is the "uniqueness" (i.e., unique item variance).

Individual i is assumed to respond positively to item j when y_{ij} is greater than the item threshold γ_j. Thus, the probability that an individual with factor score vector $\boldsymbol{\theta}_i$ will respond positively to item j, as indicated by the item score $x_{ij} = 1$, is given by the item-response function,

$$\begin{aligned}
\Phi_j(\boldsymbol{\theta}_{ij}) &= P(x_{ij} = 1|\boldsymbol{\theta}_i) \\
&= P(y_{ij} = \gamma_j|\boldsymbol{\theta}_i) \\
&= \frac{1}{2\pi\sigma_j}\int_{\gamma_j}^{\infty} \exp\left[-\frac{1}{2}\left(y_{ij} - \sum_k^m \alpha_{jk}\theta_{ki}\right)^2 / \sigma_j^2\right] dy_j, \\
&= \Phi\left[-\frac{\gamma_j - \sum_k^m \alpha_{jk}\theta_{ki}}{\sigma_j}\right]
\end{aligned}$$

and the probability that the individual will respond incorrectly, indicated by $x_{ij} = 0$, is the complement,

$$P\left(x_{ij} = 0\,|\boldsymbol{\theta}_i\right) = 1 - \Phi(\boldsymbol{\theta}_i).$$

Since the multiple factor model implies conditional independence (i.e., the items are uncorrelated conditional on the underlying factors $\boldsymbol{\theta}$), the conditional probability of the item score vector $\boldsymbol{x}_i$ is

$$P(\boldsymbol{x} = \boldsymbol{x}_i|\boldsymbol{\theta}, \boldsymbol{\gamma}, \boldsymbol{\alpha}) = \prod_j^{n_i} [\Phi_i(\boldsymbol{\theta}_i)]^{x_{ij}} [1 - \Phi_j(\boldsymbol{\theta}_i)]^{1-x_{ij}}.$$

For computational purposes it is convenient to express the argument of the response function in terms of an intercept,

$$c_j = -\gamma_j/\sigma_j$$

and factor slopes

$$a_{jk} = \alpha_{jk}/\sigma_j$$

rather than threshold and factor loadings.

In the context of Bayes estimation, the equation above is the likelihood of $\boldsymbol{\theta}_i$, and the prior, which is multivariate normal, is completely specified. However, because of the nature of this likelihood function, this is an example of a model outside the exponential family for which no closed form of the posterior mean or covariance matrix is available. Note, however, that the unconditional probability of score pattern $\boldsymbol{x}_i$ can be expressed as

$$h(\boldsymbol{x}_i) = \int_{-\infty}^{\infty} P(\boldsymbol{x} = \boldsymbol{x}_i \mid \boldsymbol{\theta}, \boldsymbol{\gamma}, \boldsymbol{\alpha}) g(\boldsymbol{\theta}) d\boldsymbol{\theta}$$

As such, the integral in the preceding equation is numerically approximated by m-fold Gauss-Hermite product quadrature (Bock and Aitkin 1981; Bock et al. 1988).

The Bi-factor IRT Model

The bi-factor restriction for IRT models (Gibbons and Hedeker 1992) was the first example of a confirmatory multidimensional IRT model. The bi-factor model is based on the idea that in many cases multidimensionality is produced by the sampling of items from multiple domains of an overall psychological construct. For example, in the measurement of fatigue impact, a measure could include items that assess cognitive, physical, and social impact of fatigue. It is quite natural for responses to such items to appear to be multidimensional, when in fact, the items measure a unidimensional construct, that is, fatigue impact.

However, the items within subdomains are more highly correlated than items between domains. For example, responses to two items measuring cognitive fatigue will be more highly correlated than responses to a pair of items, one measuring cognitive fatigue and the other measuring physical fatigue. This leads to violation of the conditional independence assumption of a unidimensional IRT model and results in dimensionality equal to the number of domains from which the items were sampled. However, a plausible s-factor solution for many types of psychological and educational tests is one that exhibits a general factor and $s - 1$ group- or method-related factors. The bi-factor solution constrains each item j to have a nonzero loading α_{j1} on the primary dimension and a second loading (α_{jk}, $k = 2, \ldots, s$) on not more than one of the $s - 1$ group factors. For four items, the bi-factor pattern matrix might be

$$\alpha = \begin{bmatrix} \alpha_{11} & \alpha_{12} & 0 \\ \alpha_{21} & \alpha_{22} & 0 \\ \alpha_{31} & 0 & \alpha_{33} \\ \alpha_{41} & 0 & \alpha_{43} \end{bmatrix}$$

This structure, which Holzinger and Swineford (1937) termed the "bi-factor" solution, also appears in the inter-battery factor analysis of Tucker (1958) and is one confirmatory factor analysis model considered by Jöreskog (1969). In these applications, the model is restricted to test scores assumed to be continuously distributed. But it is easy to conceive of situations where the bi-factor pattern might also arise at the item level (Muthén 1989). It is plausible for paragraph comprehension tests, for example, where the primary dimension describes the targeted process skill and additional factors describe content area knowledge within paragraphs. Similarly, in the context of mental health measurement, symptom items are often selected from measurement domains and can be related to the primary dimension of interest (e.g., depression) and one subdomain (e.g., anxiety). In these contexts, items would be conditionally independent between paragraphs or domains but conditionally dependent within paragraphs or domains.

The bi-factor restriction leads to a major simplification of likelihood equations that (a) permits analysis of models with large numbers of group factors (e.g., domains), (b) permits conditional dependence among identified subsets of items, (c) is rotationally invariant in contrast to the unrestricted item factor model, (d) always reduces the likelihood to a two-dimensional integral that is easily evaluated using traditional numerical methods, and (e) in many cases provides a more parsimonious factor solution than an unrestricted full-information item factor analysis (Bock and Aitkin 1981). Furthermore, in the context of computer adaptive testing (CAT), the bi-factor model provides a single endpoint (i.e., the core dimension) by which to adaptively select items from a potentially large bank of symptom items.

Recently, Gibbons et al. (2007) extended the bi-factor model to the case of ordinal symptom items, making the methodology even more useful in the context of psychological measurement problems (e.g., quality of life) where Likert-type rating scales are often used. The bi-factor model can also be used in conjunction with CAT to measure mental health constructs (Gibbons et al. 2008). Cai (2010) has further generalized the bi-factor model to the case in which there are multiple intercorrelated primary domains in addition to subdomains that nested within each of the multiple primary domains. There are numerous interesting applications of this even more general model. The IRTPRO computer program (Cai et al. 2011) can fit unidimensional and multidimensional IRT models in addition to the bi-factor model for binary and ordinal response data.

Recently, there have been several applications of bi-factor models in the area of personality research (Patrick 2007; Reininghaus et al. 2011; Reise et al. 2010; Rijmen 2010; Yang et al. 2009). In the area of life quality assessment, unidimensional IRT models have been considered (Fryback et al. 2010) as well as IRT-based CAT (Rebollo et al. 2010). Gibbons et al. (2007) have presented a bi-factor analysis of a quality of life scale, as described in the following.

Quality of Life Illustration

As an illustration of the bi-factor model for graded response data, (Gibbons et al. 2007) Gibbons and colleagues analyzed the "Quality of Life Interview for the Chronically Mentally Ill" (Lehman 1988). Their analysis was based on item responses of 586 chronically mentally ill patients. Analyses were performed using the freely available POLYBIF software (Gibbons and Hedeker 2007). The scale consists of seven subdomains (Family, Finance, Health, Leisure, Living, Safety, and Social), each with 4–6 items for a total of 34 items. In addition, there is one global life satisfaction item, which was allowed to load on its own subdomain in the event that it had a unique contribution to the residual variation above and beyond its contribution to the primary dimension. Each item is rated on a 7-point scale with the following response categories: 1 = terrible; 2 = unhappy; 3 = mostly dissatisfied; 4 = mixed, about equally satisfied and dissatisfied; 5 = mostly satisfied; 6 = pleased; and 7 = delighted.

Item intercepts, primary factor loadings, and factor loadings on the eight subdomains are displayed in Table 1 based on the polytomous rating scale model (Gibbons et al. 2007). Table 1 shows that all items had substantial loadings on the primary dimension (factor 1), indicating that the scale was well designed and that all items were related to overall life satisfaction. The three most discriminating items were "global life satisfaction," factor loading (FL) = 0.694; satisfaction with "free time," FL = 0.611 (subdomain 4); and "emotional well-being," FL = 0.609 (subdomain 3). The three least discriminating items were satisfaction with "people in general," FL = 0.385 (subdomain 7); "amount you pay for basic needs," FL = 0.391 (subdomain 2); and "pleasure from TV," FL = 0.414 (subdomain 4). The unique "life as a whole" item loaded heavily on the primary dimension, but not at all on the subdomain, indicating that the primary dimension is a good measure of overall life satisfaction. The item intercepts permit items to be positioned relative to the global life satisfaction item to determine at what point on the scale a person would report global life satisfaction. Table 1 shows that the Health (subdomain 3), Living (subdomain 5), and Social domains (subdomain 7) were typically reported at lower levels of satisfaction than the global item, whereas Financial (subdomain 2) and Leisure (subdomain 4) items had, on average, higher intercepts than the global satisfaction item. The domains of Family (subdomain 1) and Safety (subdomain 6) items were located at similar levels to the global item.

In terms of subdomains, items within domains had a high degree of residual association, with an average loading of 0.406. Consistent with this finding was a significant likelihood ratio test for improvement in fit of the bi-factor model over the unidimensional graded response model ($\chi^2 = 2188$, $df = 35, p < 0.0001$). Table 2 displays the observed and expected (in italics) category proportions for each item. In general, there is close agreement between observed and expected response proportions. The root mean square error (RMSE) between observed and expected proportions (over all items and categories) was 0.026, indicating that the model with common category parameters (Muraki 1990) fit these data extremely well. The six category parameters were as follows:

$-1.395, -.858, -.449, .044, .866, 1.793$

A model with unique item category parameters (Samejima 1969) produced a significant likelihood ratio test for improvement in fit over the rating scale model ($\chi^2 = 1637$, $df = 169$, $p < 0.0001$), with a decrease in RMSE between observed and expected proportions to 0.010. Factor loadings were almost identical between the two models. Furthermore, there were only minor changes in the estimated item thresholds between the two models, despite the fact that the rating scale model has only one item-specific threshold (and six general thresholds) and Samejima's model has six unique thresholds per item. For example, estimated item thresholds for the first ten quality of life items for both models are presented in Table 3. Table 3 shows that the estimated thresholds are quite similar for the two models. Although the fit of the model is significantly improved when estimating category parameters separately for each item (presumably due to the large number of subjects, items, and categories), the model with common category parameters may

Bi-factor Analysis, Table 1 Nine-dimensional bi-factor solution for the Lehman quality of life rating scale data ($N = 586$) item intercepts and factor loadings

			Factor								
Scale	Item	Intept	1	2	3	4	5	6	7	8	9
Global	Life as a whole	−0.402	0.694	0.001							
1	Family	−0.768	0.499		0.566						
1	Amount of family contact	−0.349	0.534		0.443						
1	Family with interaction	−0.282	0.548		0.518						
1	General family stuff	−0.350	0.597		0.491						
2	Total money you get	0.209	0.435			0.568					
2	Amt pay for basic needs	−0.136	0.391			0.477					
2	Financial well-being	0.319	0.503			0.562					
2	Money for fun	0.242	0.491			0.568					
3	Health in general	−0.482	0.458				0.270				
3	Medical care	−0.701	0.475				0.419				
3	How often see doctor	−0.441	0.441				0.397				
3	Talk to therapist	−0.621	0.478				0.378				
3	Physical condition	−0.582	0.553				0.299				
3	Emotional well-being	−0.284	0.609				0.185				
4	Way spend free time	−0.139	0.611					0.262			
4	Amount of free time	−0.292	0.509					0.342			
4	Chance to enjoy time	−0.552	0.578					0.386			
4	Amount of fun	−0.270	0.597					0.430			
4	Amount of relaxation	−0.306	0.525					0.393			
4	Pleasure from TV	−0.776	0.414					0.163			
5	Living arrangements	−0.435	0.493						0.493		
5	Food	−0.982	0.449						0.468		
5	Privacy	−0.709	0.478						0.610		
5	Amount of freedom	−1.090	0.478						0.649		
5	Prospect of staying	−0.100	0.469						0.630		
6	Neighborhood safety	−0.298	0.511							0.445	
6	Safe at home	−0.666	0.542							0.416	
6	Police access	−0.062	0.487							0.429	
6	Protect robbed/attack	−0.214	0.517							0.465	
6	Personal safety	−0.533	0.531							0.326	

(continued)

Bi-factor Analysis, Table 1 (continued)

			Factor								
Scale	Item	Intept	1	2	3	4	5	6	7	8	9
7	Do things with others	−0.614	0.494								0.326
7	Time with others	−0.411	0.519								0.257
7	Social interactions	−0.604	0.472								0.346
7	People in general	−0.835	0.385								0.220

Bi-factor Analysis, Table 2 Observed and expected (*in italics*) proportions from the nine-dimensional graded bi-factor analysis of Lehman quality of life rating scale data ($N = 586$)

	Category						
Item	1	2	3	4	5	6	7
Life as a whole	0.063	0.087	0.080	0.176	0.224	0.212	0.159
	0.098	*0.084*	*0.088*	*0.128*	*0.232*	*0.211*	*0.158*
Family	0.046	0.068	0.049	0.160	0.203	0.232	0.241
	0.078	*0.065*	*0.069*	*0.105*	*0.208*	*0.224*	*0.251*
Amount of family contact	0.061	0.097	0.114	0.133	0.244	0.210	0.140
	0.104	*0.088*	*0.090*	*0.131*	*0.232*	*0.206*	*0.149*
Family with interaction	0.067	0.125	0.094	0.167	0.200	0.217	0.131
	0.135	*0.092*	*0.089*	*0.123*	*0.211*	*0.190*	*0.160*
General family stuff	0.072	0.108	0.087	0.159	0.229	0.186	0.160
	0.134	*0.088*	*0.084*	*0.117*	*0.205*	*0.192*	*0.180*
Total money you get	0.138	0.155	0.137	0.128	0.235	0.143	0.063
	0.204	*0.121*	*0.108*	*0.137*	*0.203*	*0.145*	*0.081*
Amt pay for basic needs	0.077	0.121	0.106	0.145	0.276	0.195	0.080
	0.114	*0.103*	*0.106*	*0.149*	*0.246*	*0.187*	*0.096*
Financial well-being	0.174	0.152	0.133	0.131	0.201	0.142	0.067
	0.240	*0.122*	*0.104*	*0.128*	*0.187*	*0.136*	*0.083*
Money for fun	0.147	0.171	0.148	0.109	0.208	0.135	0.082
	0.223	*0.119*	*0.104*	*0.129*	*0.193*	*0.143*	*0.090*
Health in general	0.048	0.063	0.051	0.113	0.392	0.215	0.118
	0.056	*0.072*	*0.087*	*0.140*	*0.272*	*0.239*	*0.133*
Medical care	0.043	0.039	0.055	0.135	0.258	0.311	0.160
	0.052	*0.061*	*0.073*	*0.119*	*0.245*	*0.250*	*0.199*
How often see doctor	0.049	0.061	0.099	0.125	0.309	0.242	0.114
	0.070	*0.078*	*0.089*	*0.138*	*0.259*	*0.228*	*0.138*
Talk to therapist	0.036	0.041	0.085	0.123	0.292	0.280	0.143
	0.055	*0.065*	*0.078*	*0.126*	*0.253*	*0.247*	*0.176*
Physical condition	0.034	0.072	0.084	0.119	0.261	0.283	0.174
	0.062	*0.069*	*0.080*	*0.127*	*0.249*	*0.240*	*0.173*
Emotional well-being	0.065	0.087	0.104	0.157	0.273	0.195	0.119
	0.098	*0.091*	*0.097*	*0.141*	*0.246*	*0.205*	*0.122*
Way spend free time	0.077	0.113	0.126	0.159	0.225	0.201	0.099
	0.126	*0.102*	*0.102*	*0.142*	*0.235*	*0.185*	*0.108*
Amount of free time	0.060	0.077	0.119	0.154	0.273	0.208	0.109
	0.091	*0.090*	*0.097*	*0.143*	*0.252*	*0.207*	*0.118*
Chance to enjoy time	0.053	0.082	0.087	0.130	0.218	0.241	0.189
	0.081	*0.075*	*0.081*	*0.122*	*0.232*	*0.225*	*0.186*
Amount of fun	0.077	0.118	0.114	0.126	0.218	0.196	0.150
	0.130	*0.093*	*0.091*	*0.186*	*0.217*	*0.192*	*0.151*
Amount of relaxation	0.077	0.080	0.108	0.131	0.259	0.225	0.119
	0.100	*0.090*	*0.095*	*0.137*	*0.212*	*0.205*	*0.131*
Pleasure from tv	0.020	0.034	0.055	0.143	0.275	0.282	0.188
	0.026	0.016	0.065	0.120	0.276	0.287	0.181
Living arrangements	0.073	0.070	0.085	0.131	0.271	0.210	0.189
	0.095	*0.082*	*0.086*	*0.127*	*0.232*	*0.214*	*0.163*
Food	0.041	0.032	0.056	0.072	0.234	0.304	0.261
	0.035	*0.045*	*0.057*	*0.100*	*0.227*	*0.267*	*0.269*

(continued)

Bi-factor Analysis, Table 2 (continued)

	Category						
Item	1	2	3	4	5	6	7
Privacy	0.087	0.051	0.080	0.097	0.186	0.258	0.241
	0.092	*0.069*	*0.071*	*0.105*	*0.202*	*0.214*	*0.247*
Amount of freedom	0.065	0.051	0.049	0.067	0.195	0.230	0.343
	0.071	*0.054*	*0.057*	*0.087*	*0.179*	*0.214*	*0.339*
Prospect of staying	0.130	0.119	0.094	0.150	0.160	0.140	0.186
	0.177	*0.099*	*0.090*	*0.119*	*0.196*	*0.170*	*0.147*
Neighborhood safety	0.077	0.080	0.082	0.143	0.294	0.218	0.106
	0.106	*0.091*	*0.094*	*0.135*	*0.236*	*0.202*	*0.136*
Safe at home	0.55	0.043	0.061	0.111	0.280	0.280	0.171
	0.066	*0.067*	*0.075*	*0.117*	*0.233*	*0.237*	*0.205*
Police access	0.137	0.061	0.131	0.172	0.217	0.174	0.108
	0.134	*0.108*	*0.107*	*0.146*	*0.235*	*0.176*	*0.094*
Protect robbed/attack	0.094	0.073	0.118	0.147	0.254	0.203	0.111
	0.124	*0.097*	*0.096*	*0.135*	*0.229*	*0.191*	*0.125*
Personal safety	0.048	0.048	0.070	0.130	0.309	0.276	0.119
	0.66	*0.072*	*0.083*	*0.130*	*0.252*	*0.235*	*0.162*
Do things with others	0.031	0.032	0.067	0.142	0.341	0.254	0.133
	0.053	*0.065*	*0.078*	*0.127*	*0.257*	*0.249*	*0.171*
Times with others	0.036	0.063	0.080	0.167	0.317	0.247	0.089
	0.070	*0.080*	*0.091*	*0.141*	*0.262*	*0.225*	*0.130*
Social interactions	0.036	0.039	0.067	0.159	0.285	0.278	0.137
	0.053	*0.065*	*0.079*	*0.128*	*0.259*	*0.248*	*0.168*
People in general	0.019	0.027	0.032	0.128	0.302	0.321	0.171
	0.023	*0.042*	*0.060*	*0.114*	*0.272*	*0.294*	*0.195*

be a useful alternative for applications in which the items have the same number of categories.

While the bifactor model improves the precision of measurement, it does not in and of itself reduce the burden of measurement. Gibbons et al. (2012, 2014, 2017, Gibbons et al. 2016; Gibbons and deGruy 2019) have demonstrated that the bifactor model can be used to provide adaptive administration of items to measure complex (i.e. multidimensional) traits (e.g., depression, anxiety, suicidality) that increases the precision of measurement over traditional short-form tests, but at the same time completely eliminates clinician burden and minimizes patient/subject burden. Applying this methodology to the current example of a 34-item test yields a computerized adaptive quality of life test that requires an average of eight items (range 4–13), yet maintains an almost perfect correlation of r = 0.95 with the 34-item QOL score. The CAT-QOL can be administered in less than 2 min, potentially on any internet capable device. Of course, it could be further improved by expanding the item bank to a much larger set of items (hundreds of items) covering additional domains of life quality.

Summary

In summary, the bi-factor model provides an excellent modern psychometric approach to life quality measurement problems. The general model provides unlimited dimensionality under the restriction that the subdomains of interest are

Bi-factor Analysis, Table 3 Estimated category thresholds for Rating Scale Model (RSM) and Samejima Model (SM) for the first ten quality of life items

		Category					
Item		1–2	2–3	3–4	4–5	5–6	6–7
Life as a whole	RSM	−1.452	−1.005	−0.673	−0.278	0.361	1.079
	SM	−1.450	−0.994	−0.714	−0.217	0.374	1.072
Family	RSM	−1.578	−1.171	−0.868	−0.508	0.074	0.728
	SM	−1.666	−1.184	−0.950	−0.408	0.110	0.722
Family contact	RSM	−1.435	−0.981	−0.642	−0.241	0.409	1.140
	SM	−1.484	−0.953	−0.575	−0.210	0.409	1.115
Family interaction	RSM	−1.236	−0.831	−0.530	−0.172	0.408	1.059
	SM	−1.441	−0.836	−0.536	−0.085	0.416	1.156
Family stuff	RSM	−1.240	−0.489	−0.558	−0.213	0.347	0.976
	SM	−1.397	−0.876	−0.588	−0.154	0.418	1.023
Total money	RSM	−0.952	−0.521	−0.199	0.182	0.799	1.493
	SM	−1.084	−0.528	−0.134	0.206	0.870	1.552
Basic needs	RSM	−1.395	−0.895	−0.523	−0.082	0.634	1.437
	SM	−1.432	−0.840	−0.486	−0.084	0.643	1.424
Financial well-being	RSM	−0.819	−0.415	−0.114	0.243	0.821	1.471
	SM	−0.913	−0.417	−0.064	0.281	0.869	1.533
Money for fun	RSM	−0.873	−0.467	−0.165	0.194	0.775	1.428
	SM	−1.032	−0.439	−0.046	0.236	0.829	1.430
General health	RSM	−1.876	−1.322	−0.908	−0.418	0.376	1.268
	SM	−1.659	−1.189	−0.945	−0.555	0.465	1.224

known in advance and that each item taps the primary domain and no more than one subdomain. In general, this information is readily available in life quality research in that the items are sampled from unique subdomains as illustrated in the example. A further benefit of the bi-factor model over unrestricted exploratory item factor analysis is that the solution is rotationally invariant and therefore easily interpretable. As shown here, the model is easily extended to the case of ordinal response items, which characterize many if not most applications in life quality research. When the focus is on the primary domain of interest (e.g., overall life quality) as in the example, the addition of the subdomains resolves problems of conditional dependence that invalidates unidimensional IRT models and allows one to develop large item banks that can be used for CAT (Gibbons et al. 2012).

Cross-References

▶ Confirmatory Factor Analysis (CFA)
▶ Exploratory Factor Analysis
▶ Factor Analysis
▶ Item Response Theory (IRT)

References

Bock, R. D., & Aitkin, M. (1981). Marginal maximum likelihood estimation of item parameters: Application of an EM algorithm. *Psychometrika, 46*, 443–459.

Bock, R. D., Gibbons, R. D., & Muraki, E. (1988). Full-information item factor analysis. *Applied Psychological Measurement, 12*, 261–280.

Cai, L. (2010). A two-tier full-information item factor analysis model with applications. *Psychometrika, 75*, 581–612.

Cai, L., du Toit, S. H. C., & Thissen, D. (2011). *IRTPRO: Flexible, multidimensional, multiple categorical IRT modeling*. Chicago: Scientific Software International.

Fryback, D. G., Palta, M., Cherepanov, D., Bolt, D., & Kim, J. S. (2010). Comparison of 5 health-related quality of life indexes using item response theory analysis. *Medical Decision Making, 30*, 5–15.

Gibbons, R. D., & deGruy, F. V. (2019). Without wasting a word: Extreme improvements in efficiency and accuracy using computerized adaptive testing for mental health disorders (CAT-MH). *Current Psychiatry Reports, 21*, 1053–1059.

Gibbons, R. D., & Hedeker, D. (1992). Full-information item bi-factor analysis. *Psychometrika, 57*, 423–436.

Gibbons, R. D., & Hedeker, D. (2007). *POLYBIF [Computer software]*. Chicago: Center for Health Statistics, University of Chicago. Available from http://www.healthstats.org

Gibbons, R. D., Bock, R. D., Hedeker, D., Weiss, D. J., Segawa, E., Bhaumik, D. K., et al. (2007). Full-information item bifactor analysis of graded response data. *Applied Psychological Measurement, 31*, 4–19.

Gibbons, R. D., Weiss, D. J., Kupfer, D. J., Frank, E., Fagiolini, A., Grochocinski, V. J., et al. (2008). Using computerized adaptive testing to reduce the burden of mental health assessment. *Psychiatric Services, 59*, 361–368.

Gibbons, R. D., Weiss, D. J., Pilkonis, P. A., Frank, E., Moore, T., et al. (2012). Development of a computerized adaptive test for depression. *Archives of General Psychiatry, 69*, 1104–1112.

Gibbons, R. D., Weiss, D. J., Pilkonis, P. A., Frank, E., Moore, T., Kim, J. B., & Kupfer, D. J. (2014). Development of the CAT-ANX: A computerized adaptive test for anxiety. *American Journal of Psychiatry, 171*, 187–194.

Gibbons, R. D., Weiss, D. J., Frank, E., & Kupfer, D. (2016). Computerized adaptive diagnosis and testing of mental health disorders. *Annu Rev Clin Psychol, 12*, 83–104. https://doi.org/10.1146/annurev-clinpsy-021815-093634. Epub 2015 Nov 20. PMID: 26651865.

Gibbons, R. D., Kupfer, D., Frank, E., Moore, T., & Boudreaux, E. (2017). Development of a computerized adaptive suicide scale. *Journal of Clinical Psychiatry, 78*, 1376–1382.

Holzinger, K. J., & Swineford, F. (1937). The bi-factor method. *Psychometrika, 2*, 41–54.

Jöreskog, K. G. (1969). A general approach to confirmatory maximum likelihood factor analysis. *Psychometrika, 34*, 183–202.

Lehman, A. F. (1988). A quality of life interview for the chronically mentally ill. *Evaluation and Program Planning, 11*, 51–62.

Muraki, E. (1990). Fitting a polytomous item response model to Likert-type data. *Applied Psychological Measurement, 14*, 59–71.

Muthén, B. O. (1989). Latent variable modeling in heterogeneous populations. *Psychometrika, 54*, 557–585.

Patrick, C. J. (2007). A bifactor approach to modeling the structure of the psychopathy checklist revised. *Journal of Personality Disorders, 21*, 118–141.

Rebollo, P., Castejon, I., Cuervo, J., Villa, G., Garcia-Cueto, E., Diaz-Cuervo, H., et al. (2010). Validation of a computer-adaptive test to evaluate generic health-related quality of life. *Health and Quality of Life Outcomes, 8*, 1–8.

Reininghaus, U., McCabe, R., Burns, T., Croudace, T., & Priebe, S. (2011). Measuring patients' views: A bifactor model of distinct patient-reported outcomes in psychosis. *Psychological Medicine, 41*, 277–289.

Reise, S. P., Moore, T. M., & Haviland, M. G. (2010). Bifactor models and rotations: Exploring the extent to which multidimensional data yield univocal scale scores. *Journal of Personality Assessment, 92*, 544–559.

Rijmen, F. (2010). Formal relations and an empirical comparison among the bi-factor, the testlet, and a second-order multidimensional IRT model. *Journal of Educational Measurement, 47*, 361–372.

Samejima, F. (1969). Estimation of latent ability using a response pattern of graded scores. *Psychometrika Monograph Supplement, 17*, 1–68.

Thurstone, L. L. (1947). *Multiple factor analysis*. Chicago: University of Chicago Press.

Tucker, L. R. (1958). An inter-battery method of factor analysis. *Psychometrika, 23*, 111–136.

Yang, F. M., Tommet, D., & Jones, R. N. (2009). Disparities in self-reported geriatric depressive symptoms due to sociodemographic differences: An extension of the bi-factor item response theory model for use in differential item functioning. *Journal of Psychiatric Research, 43*, 1025–1035.