Mean structure analysis from an IRT approach: An application in the context of organizational psychology

Javier Revuelta Menéndez and Carmen Ximénez Gómez
Universidad Autónoma de Madrid

The application of mean and covariance structure analysis with quantitative data is increasing. However, latent means analysis with qualitative data is not as widespread. This article summarizes the procedures to conduct an analysis of latent means of dichotomous data from an item response theory approach. We illustrate the implementation of these procedures in an empirical example referring to the organizational context, where a multi-group analysis was conducted to compare the latent means of three employee groups in two factors measuring personal preferences and the perceived degree of rewards from the organization. Results show that higher personal motivations are associated with higher perceived importance of the organization, and that these perceptions differ across groups, so that higher-level employees have a lower level of personal and perceived motivation. The article shows how to estimate the factor means and the factor correlation from dichotomous data, and how to assess goodness of fit. Lastly, we provide the M-Plus syntax code in order to facilitate the latent means analyses for applied researchers.

METODOLOGÍA

Mean and covariance structure analysis comprises a set of statistical techniques that simultaneously deals with the means of the latent variables and the covariance structure. The factor model is possibly the most popular type of statistical model for this purpose. Its original formulation assumes that the means of the latent variables are null and therefore neglects the mean structure analysis. Additionally, it assumes a linear relation between the observed variables and the factors, which implies that the observed variables are quantitative. However, later developments extended these models for the analysis of categorical data and the mean structure (Jöreskog & Sörbom, 2001).

Adding the mean structure implies that the observed means are summarized into a smaller number of latent means. The practical application of this analysis consists of comparing latent means across several groups of individuals while simultaneously investigating the pattern of covariance in a single statistical framework. This technique is meaningful when several variables measure one common characteristic.
Several theoretical studies have noted the advantages of adding the analysis of the mean structure in the factor model. Yung and Bentler (1999) found that in maximum likelihood factor analysis the reduction of asymptotic variances for factor loadings can be quite substantial when a mean structure is added. Yuan and Bentler (2006) found that summarizing observed variables into latent means conveys an increase of statistical power when null hypothesis significance testing is used to compare means across groups. Finally, another advantage of adding the associated mean structure is that the model accounts for measurement error variances in estimating the latent means whereas this is not possible in the techniques that compare the observed means (e.g., T-test and ANOVA).

Given these advantages, the application of factorial models simultaneously analyzing the mean and covariance structure for quantitative variables is increasingly growing and several empirical studies use this approach. However, the latent means analysis with qualitative data is not as widespread. Particularly, because the use of dichotomous data implies that the linear model is not well suited to approximate the relation between the observed and the latent variables and instead, models are based in a nonlinear function (e.g., an item response theory model, IRT) and the interpretation of the latent means is based on their relation with the probability of the response categories, instead of the relation with the values of the observed variables as in the linear model.

The aim of this article is to explain the procedures to conduct a mean structure analysis of dichotomous data from an IRT approach in the one-dimensional case. Despite IRT models are well known, they are seldom used to compare latent means between groups. This is due to two reasons: First, the mean structure analysis requires the assumption of certain theoretical constraints, both in the item parameters and in the means. Second, the estimation of these models requires and advance use of the computer programs. With this article, we pretend to clarify both issues to facilitate the use of these models for applied researchers.

The article is organized as follows. First, we explain the IRT with mean structure model, its relation with the factor model, and how the interpretation of the latent means is based on the probabilities of the response outcomes. Second, we explain the necessary constraints to be imposed in the IRT model for the estimation of the latent means and the problem of the factorial invariance so that the comparison of the latent means across groups is meaningful. Finally, we illustrate the implementation and interpretation of these procedures in an empirical example taken from the organizational context and provide the M-Plus syntax code for these analyses, which are less known by researchers.

Item response theory model with mean structure

The common factor model with mean structure is defined by (Sörbom, 1981):

\[ x = \tau_x + \Lambda \xi + \delta, \]  

where \( x \) is a vector of \( p \) observed variables, \( \tau_x \) is a vector of \( p \) constant intercept terms, \( \xi \) is a factor or latent variable, \( \Lambda \) is a vector of factor loadings, and \( \delta \) is a vector of \( p \) measurement errors. It is assumed that \( \text{E}(\delta) = 0 \) and \( \text{E}(\xi) = 0 \). However, \( \text{E}(\delta) \) is not 0, it is a parameter denoted by \( \kappa \). By taking the expectations of Equation (1), the mean vector of the observed variables is:

\[ \text{E}(x) = \tau_x + \Lambda \kappa \]  

(2)

Certain constraints need to be imposed for the estimation of parameters. For instance, \( \lambda_i \) is set to 1 and its corresponding \( \tau_x \) to 0, so that the factor has the same measurement scale than \( \xi \) and the same mean: \( \text{E}(\xi) = \kappa \). This constraint makes possible the comparison of the latent means across groups, but only when \( \Lambda \) is invariant across groups, which guaranties that the factors have the same measurement scale for each group.

Social scientists usually do not work with quantitative data. For instance, items from attitude scales are usually scored in a small number of categories, such as right or wrong responses. In these cases, the linearity assumption of the factor model is only an approximation to the relation between factors and scores. The categorical factor analysis, usually referred to as item factor analysis (Bock & Gibbons, 2010; With & Edwards, 2007), assumes that observed data originate from a discretization of an underlying quantitative variable. Consider the linear factor model defined in Equation (1). The item factor model for dichotomous data assumes that \( \xi_i \) is a latent variable and the observed responses arise by categorizing \( \xi_i \) into two response categories using one threshold value. The variable \( x_i \) is converted into the observed variable \( r_i \) according to the transformation:

\[ r_i = \begin{cases} 0, & \text{if } x_i \leq v_i \\ 1, & \text{if } x_i > v_i \end{cases} \]

where \( x_i = \lambda_i \xi_i + \delta_i \) and the threshold \( v_i \) is a new parameter that needs to be estimated. For convenience, the model assumes that the distribution of \( x_i \) conditional on \( \xi_i \) referred to as \( f_j(x_i) \), is normal with mean \( \lambda_i \xi_i \) and variance \( 1-\lambda_i^2 \). Given a fixed value of \( \xi_i \), the probability that \( r_i \) falls in the category 1 is the area under the normal curve that lies above the threshold. If \( \phi(z) \) is a standard normal density function, then:

\[ \pi_i = P(x > v_i) = \int_{v_i}^{\infty} f_j(x_i)dx = \int_{v_i}^{\infty} \phi(z)dz = \int_{v_i-\lambda_i \xi_i}^{\infty} \varphi(z)\varphi(z)dz \]

(3)

This model is also referred to as a normal ogive model.

The two-parameter logistic model from IRT (Hambleton & Swaminathan, 1985) may be seen as a transformation of the item parameters \( \lambda_i \) and \( v_i \) into the \( a_i \) and \( b_i \) parameters to facilitate interpretation, where \( a_i \) is a scale parameter that indicates the ratio of change of \( \pi_i \) in relation to \( \xi_i \), and \( b_i \) is a difficulty parameter that indicates the value of \( \xi_i \) that has a probability 0.50 of endorsing the item. The relation between the parameters of the common factor model and the IRT parameterization is given by (Ferrando, 1996):

\[ a_i = \frac{\lambda}{\sqrt{1-\lambda^2}} \quad b_i = \frac{v_i}{\lambda} \]

Using this transformation and approximating the normal area by a logistic function results in the two parameter logistic model from IRT:
where $D=1.70$. However, we will omit $D$ because it is not necessary to improve model-data fit. When the data are dichotomous, the interpretation of factor means is based on their relation with the probability of the response categories, instead of the relation with the values of the observed variable as in the linear model.

IRT models also have a problem of indeterminacy of the scale of the factors, which is resolved by setting some parameters to constant values so that the model is identified (Revuelta, 2009). More specifically, the difficulty parameter for one of the items is set to zero ($b=0$) so that the factor mean can be estimated without increasing the number of parameters. The value of $\kappa$ is interpreted in relation to the probability of endorsing the first item. Given that $b=0$, then $\pi=\exp(a_1z) / (1 + \exp(a_1z))$ and the factor is function of the odds of endorsing the item:

$$\log\left(\frac{\pi}{1-\pi}\right) = \xi$$

Thus, $\kappa$ is the population mean value of function (5). As this function increases with $\pi$, the high values of $\kappa$ are associated to a higher probability of endorsing Item 1. Additionally, $a_1$ can be fixed to a constant value, say 1, for estimating the variance of the factor.

These constraints allow the estimation of the factor mean and variance in the single group case. In multi-group designs the aim is to compare factor means across groups. This is meaningful if $a_1$ and $b_1$ are invariant across groups. Only if the invariance is achieved, any given trait level has associated the same response probability in all groups and therefore, if the mean in a group is larger than any given trait level has associated the same response probability.

As this constraint is very restrictive, in practice partial invariance models are compared assuming that the same constraints in the latent means comparison. Provided that non-invariance to determine if there is some degree of invariance, in Step 2 the latent means and variances are compared assuming that the same constraints in certain subset of variables.

The practical implementation of a latent mean analysis requires two steps: 1) the analysis of the invariance across groups; and 2) the latent means comparison. Step 1 is carried out by comparing the goodness of fit of models of full invariance, partial invariance and non-invariance to determine if there is some degree of invariance that makes possible the latent means comparison. Provided that there is some degree of invariance, in Step 2 the latent means and variances are compared assuming that the same constraints in $a_1$ and $b_1$ have been imposed across all the groups.

In multi-group studies it is common to assume that the means of the groups are a linear function of the parameters that indicate if there are main or interaction effects, similar as in log-linear analysis (Agresti, 2002) and ANOVA models (Rencher & Schaalje, 2008). For example, in a design with two independent variables, the mean of each cell of the design is denoted by:

$$\kappa_{j\ell} = \mu + \alpha_j + \beta_l + \gamma_{j\ell},$$

where $j$ and $\ell$ are the levels of each independent variable, $\alpha_j$ and $\beta_l$ are the main effects, and $\gamma_{j\ell}$ is the interaction effect. These parameters are constrained to sum 0 and are expressed as linear combinations of a set of parameters, named basic parameters. For example, in a $2 \times 3$ design the cell parameters are expressed as a function of six basic parameters, $\delta_l$ to $\delta_b$, as follows:

$$\begin{align*}
\kappa_1 &= \delta_1 + \delta_2 + \delta_3 + \delta_4 + \delta_5 + \delta_6 \\
\kappa_2 &= \delta_1 - \delta_2 + \delta_3 - \delta_4 + \delta_5 - \delta_6 \\
\kappa_3 &= \delta_1 + \delta_2 - \delta_3 - \delta_4 + \delta_5 + \delta_6 \\
\kappa_4 &= \delta_1 - \delta_2 - \delta_3 + \delta_4 - \delta_5 - \delta_6 \\
\kappa_5 &= \delta_1 - \delta_2 - \delta_3 - \delta_4 + \delta_5 - \delta_6 \\
\kappa_6 &= \delta_1 - \delta_2 + \delta_3 - \delta_4 - \delta_5 + \delta_6 \\
\end{align*}$$

The parameters of model (6) are given by:

$$\begin{align*}
\mu &= \delta_1, \\
a_1 &= \delta_2, a_2 = -\delta_2, \\
b_1 &= \delta_3, b_2 = -\delta_3, b_3 = -\delta_3, \\
\gamma_{11} &= \delta_4, \gamma_{12} = \delta_5, \gamma_{13} = -\delta_4 - \delta_5, \\
\gamma_{21} &= -\delta_4, \gamma_{22} = -\delta_5, \gamma_{23} = \delta_4 + \delta_5. \\
\end{align*}$$

This parameterization implies that the sum of the effects is zero. That is:

$$\sum_j a_j = 0, \sum_l b_l = 0, \sum_j \gamma_{j1} = 0 \mathrm{ and } \sum_j \gamma_{j2} = 0$$

In matrix notation, the vector of factor means is:

$$\kappa = C\theta,$$

where $C$ is a constant matrix that specifies the effects assessed in the model and must have full column rank. For instance, the C matrix for Equation (7) is:

$$C = \begin{pmatrix}
1 & 1 & 1 & 0 & 1 & 0 \\
1 & -1 & 1 & 0 & -1 & 0 \\
1 & 1 & 0 & 1 & 0 & 1 \\
1 & -1 & 0 & 1 & 0 & -1 \\
1 & 1 & -1 & -1 & 1 & -1 \\
1 & -1 & -1 & 1 & 1 & 1
\end{pmatrix}$$

Example

A questionnaire of Person-Organization (P-O) fit was analyzed to illustrate the analysis and interpretation of latent means with dichotomous data.

Participants. A sample of 566 participants was recruited from former university students. 260 were men and 306 were women, and their average age was 35 years (standard deviation: 6.21). Three groups were defined according to the positions of the participants in their organizations. Sample 1 with 216 employees working in low-level positions, Sample 2 with 248 employees working in middle-level positions, and Sample 3 with 102 high-level directors.

Materials and procedure. The factor model is depicted in Figure 1 and includes two factors, $P$ and $O$, each measured with 15 dichotomous items (see Table 1). The items of $P$ assess personal motivations of the employees in their workplaces and are based in the MIQ (Gay et al., 1971), and the items of $O$ assess the importance that the organization attributes to these motivations and are
based in the MJDQ (Borgen et al., 1972). Both questionnaires are independent and were administered separately.

### Table 1

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1/O1</td>
<td>My rewards compare well with those of others</td>
</tr>
<tr>
<td>P2/O2</td>
<td>My group leader provides for my continuing membership</td>
</tr>
<tr>
<td>P3/O3</td>
<td>My group leader backs me up</td>
</tr>
<tr>
<td>P4/O4</td>
<td>My group leader communicates expectations well</td>
</tr>
<tr>
<td>P5/O5</td>
<td>I can make decisions on my own</td>
</tr>
<tr>
<td>P6/O6</td>
<td>I can try out my own ideas</td>
</tr>
<tr>
<td>P7/O7</td>
<td>I can plan things independently</td>
</tr>
<tr>
<td>P8/O8</td>
<td>People at my work are easy to make friends with</td>
</tr>
<tr>
<td>P9/O9</td>
<td>I can do things for other people</td>
</tr>
<tr>
<td>P10/O10</td>
<td>I can be busy all the time</td>
</tr>
<tr>
<td>P11/O11</td>
<td>I can do something different every day</td>
</tr>
<tr>
<td>P12/O12</td>
<td>I can get a feeling of accomplishment</td>
</tr>
<tr>
<td>P13/O13</td>
<td>I can have the opportunity for self-advancement</td>
</tr>
<tr>
<td>P14/O14</td>
<td>I can receive recognition for the things I do</td>
</tr>
<tr>
<td>P15/O15</td>
<td>I can be somebody in the group</td>
</tr>
</tbody>
</table>

1 Items P1 to P15 were answered in terms of “for me it is important that ...”, whereas items O1 to O15 were answered in terms of “for my organization it is important that ...”

Statistical analyses. The estimated parameters are the factor loadings, the thresholds, the latent means and variances, and the factor correlations. Factor loadings and thresholds are constant across groups, whereas the other parameters depend on the model.

Four models were applied that differ in the constraints imposed:

- **Model 1**: no constraints across groups;
- **Model 2**: factor correlations equal across groups;
- **Model 3**: factor means equal across groups; and
- **Model 4**: factor means and correlations equal across groups.

Models were estimated using M-Plus 4.1 (Müthen & Müthen, 2006). This program provides the values of the categorical factor analysis parameters: \( \psi, \lambda, \kappa, \phi \), which were transformed into the IRT parameters \( a_i \) and \( b_i \) by using Equation (3). Table 2 shows the M-plus syntax code for the estimation of Model 1. The equations defined in (7) are implemented in lines 59 to 64. The parameters \( \delta_{6} \) to \( \delta_{15} \) are denoted by C1 to C6, and are defined in line 54. The constraints defined in the other models can be easily implemented by modifying the syntax code. The general procedure consists of assigning a label to the parameters of the factor model (see lines 15 to 50, where all terms in brackets are labels), defining the basic parameters (lines 52 to 54), and specifying the calculation of the parameters of the factorial model from the basic parameters (lines 55 to 64).

The ULS estimation method was used because the other methods are not appropriate for the data of the example (ML uses information of the response patterns and there are very few in the example, WLS uses the asymptotic covariance matrix and it can not be estimated with small sample sizes, and GLS requires normality). We did not calculate the chi-square difference for hierarchical models because it is not possible with ULS (see Technical appendices in the URL: www.statmodel.com to compute this statistic with other estimation methods).

One important concern in this context is to obtain comparable factor scores on the \( P \) and \( O \) scales. This is achieved when both scales have the same thresholds and factor loadings for every pair of items, one for each scale (Ximénez & Revuelta, 2010). As previous analysis showed evidence that this assumption is too restrictive, a partial measurement invariance approach (Byrne, Shavelson, & Muthén, 1989) was taken. Items 1 to 5 were regarded as an anchor test (von Davier, 2010) with the purpose of obtaining commensurate factor scores for the \( P \) and \( O \) factors, and thus their parameters were set to be equal on the two scales. Items 6 to 15 were left free to vary from one scale to the other.

The example corresponds to a 2x3 mixed design with one within-subjects independent variable (with levels \( P \) and \( O \)) and one between-subjects independent variable (with the three levels of employees), that could be solved with an ANOVA if we analyze the effects on the observed means. The difference with the IRT approach is that is the aim of the latter is studying the effects on the latent means.

Results. Table 3 contains the goodness of fit indices chi-square (\( \chi^2 \)) and RMSEA for each model. The results showed that all models fall short of being acceptable and imposing constraints conveys a decrease in model fit compared with the less restrictive model. For these reasons, the best fitting model (Model 1) was selected for interpretation.

Parameter estimates appear in Table 4, which follows the notation of Figure 1. As can be seen, the estimates from every pair of matched items were different when they were allowed to vary (in items 6 to 15), indicating that the reaction towards any given item stem is not the same when it refers to personal or to organizational

---

**Figure 1. Factorial model**
values. The table also shows that factor mean was consistently higher for $P$ than for $O$, indicating that the employees exhibited a lack of fulfillment of their personal values. Moreover, as the level in the organization decreases the difference between the means of personal and organizational values increased. This is reflected in which the interaction parameters are significantly different from zero ($\hat{\delta}_0 = -0.19, \text{Se} = 0.08, Z = -2.40; \text{and } \hat{\delta}_2 = 0.20, \text{Se} = 0.08, Z = 2.63$). Finally, there was a positive correlation between factors, indicating that employees with high expectations scored higher on the scale of organizational values, which increased with the level in the organization.

As a comparison, if this research problem was solved by an ANOVA the independent variable is the score in the $P$ and $O$ scales, and would be obtained by adding their 15 items. This implies a mixed design with Position (low, medium, and high) as a between-subjects factor and Scale ($P$ and $O$) as a within-subjects factor. The means of $P$ for the low, medium and high positions are 11.53, 11.35, and 11.61; the means of $O$ are 4.09, 4.71, and 5.75. Both the main and the interaction effects were significant, and the correlations between $P$ and $O$ in each group were .07, .30, and .35. Thus, the conclusions are equal with the IRT and ANOVA approaches; however, the information that is missed with the ANOVA is the analysis of dimensionality, the relation of factors, and .35. Thus, the conclusions are equal with the IRT and ANOVA approaches; however, the information that is missed with the ANOVA is the analysis of dimensionality, the relation of $P$ and $O$ with the individual scores in each item and the measurement errors.

Figure 2 provides further insight. It contains the item characteristic curves for all the items of $P$ and $O$. Conditional on the factor value, there is a higher probability of endorsing items on the $O$ scale than the matched items on the $P$ scale. The figure also contains the distribution of the factor scores, which showed a shift to the right for the distribution of $P$ because of its higher mean.

Finally, Figure 3 contains the scatter plot for the estimated factor scores. The diagonal line in the figure is the bisector, which corresponds to equality between $P$ and $O$ scores. For almost all individuals the $P$ score was higher than the $O$ score, and thus the points fall in the lower right part of the figure. Two customary measures of $P-O$ fit applied to these data (Edwards, 1993) are $d_2 = \text{Mean}(P-O)$ and $d_2 = \text{Mean}(P-O)^2$. The values of $d_2$ for the total group and the three subgroups were 2.63, 2.78, 2.58, and 2.43; and the values of $d_1$ were 7.72, 8.86, 7.27, and 6.42, indicating that as the level in the organization decreased the degree of mismatch between personal and organizational values increased.

These results lead to two conclusions. First, the individuals feel that their motivations towards their workplaces are not sufficiently fulfilled by the organizations they work in. The mean value of $P$ for the low, medium and high positions were .07, .30, and .35, respectively. Moreover, as the level in the organization decreased the degree of mismatch between personal and organizational values increased.

Finally, Figure 3 contains the scatter plot for the estimated factor scores. The diagonal line in the figure is the bisector, which corresponds to equality between $P$ and $O$ scores. For almost all individuals the $P$ score was higher than the $O$ score, and thus the points fall in the lower right part of the figure. Two customary measures of $P-O$ fit applied to these data (Edwards, 1993) are $d_2 = \text{Mean}(P-O)$ and $d_2 = \text{Mean}(P-O)^2$. The values of $d_2$ for the total group and the three subgroups were 2.63, 2.78, 2.58, and 2.43; and the values of $d_1$ were 7.72, 8.86, 7.27, and 6.42, indicating that as the level in the organization decreased the degree of mismatch between personal and organizational values increased.

These results lead to two conclusions. First, the individuals feel that their motivations towards their workplaces are not sufficiently fulfilled by the organizations they work in. The mean value of $P$ for the low, medium and high positions were .07, .30, and .35, respectively. Moreover, as the level in the organization decreased the degree of mismatch between personal and organizational values increased.

These results lead to two conclusions. First, the individuals feel that their motivations towards their workplaces are not sufficiently fulfilled by the organizations they work in. The mean value of $P$ for the low, medium and high positions were .07, .30, and .35, respectively. Moreover, as the level in the organization decreased the degree of mismatch between personal and organizational values increased.

These results lead to two conclusions. First, the individuals feel that their motivations towards their workplaces are not sufficiently fulfilled by the organizations they work in. The mean value of $P$ for the low, medium and high positions were .07, .30, and .35, respectively. Moreover, as the level in the organization decreased the degree of mismatch between personal and organizational values increased.

These results lead to two conclusions. First, the individuals feel that their motivations towards their workplaces are not sufficiently fulfilled by the organizations they work in. The mean value of $P$ for the low, medium and high positions were .07, .30, and .35, respectively. Moreover, as the level in the organization decreased the degree of mismatch between personal and organizational values increased.

These results lead to two conclusions. First, the individuals feel that their motivations towards their workplaces are not sufficiently fulfilled by the organizations they work in. The mean value of $P$ for the low, medium and high positions were .07, .30, and .35, respectively. Moreover, as the level in the organization decreased the degree of mismatch between personal and organizational values increased.

These results lead to two conclusions. First, the individuals feel that their motivations towards their workplaces are not sufficiently fulfilled by the organizations they work in. The mean value of $P$ for the low, medium and high positions were .07, .30, and .35, respectively. Moreover, as the level in the organization decreased the degree of mismatch between personal and organizational values increased.

These results lead to two conclusions. First, the individuals feel that their motivations towards their workplaces are not sufficiently fulfilled by the organizations they work in. The mean value of $P$ for the low, medium and high positions were .07, .30, and .35, respectively. Moreover, as the level in the organization decreased the degree of mismatch between personal and organizational values increased.
these perceptions differ across groups, so that the employees with higher-level positions have a lower level of personal and a higher level of perceived motivation.

Discussion and conclusion

The analysis of latent means allows summarizing the means of several observed variables in a smaller number of factor means that can be compared across groups. This reduction of information is meaningful when the observed variables measure a common attribute and it provides parsimony in the statistical analysis. Other advantages of the analysis of latent means are: 1) it allows the analysis of the covariance structure and latent mean differences across groups to be carried out simultaneously within a single integrated statistical framework; 2) it accounts for measurement error variances in estimating the latent means whereas ANOVA does not take into account this source of variance; and 3) these models have desirable statistical properties as summarizing observed means into latent means conveys an increase of statistical power when null hypothesis significance testing is used to compare means across groups.

The factorial model is appropriate to compare latent means with quantitative variables. However, the majority of data found in practice are categorical, and this may require an IRT model that allows the analysis of latent means for categorical data. These models present more difficulties because the relation between factors and observed variables is not as straightforward as in the quantitative model as it is interpreted with the probability of the response outcomes associated to each category.

This article has focused in the dichotomous case analyzed by an IRT model. At a theoretical level, it has been shown how to set certain parameters to constant values to estimate and interpret the factor means. More specifically, one difficulty parameter is set to zero and therefore, the factor mean depends on the probability of endorsing that item. Additionally, the article shows how to impose linear constraints in the latent means to assess both main and interaction effects.

At a practical level, the article illustrates the procedure with an example taken from the organizational context, where a multi-group analysis was conducted to compare the latent means of three employees groups in two factors measuring personal preferences and the perceived degree to which the organization rewards them.

The example shows how to estimate the factors working with dichotomous data, how to correlate them, assess the goodness of fit, and compare their latent means. The analysis of latent means from an IRT approach has the problem that the majority of computer software does not allow to estimate them and the programs that do so require a complex syntax. In this article, we have shown how to implement such analyses with a well known program as is the M-Plus and how to define the parameter constraints. An M-Plus syntax code has been included so that readers can adapt it to their own problems.

Acknowledgement

This work was partially supported by grant CCG08-UAM/ESP-3951 of the Comunidad of Madrid (Spain) and grant PSI2009-08264 of the Spanish Ministry of Science and Innovation.

<table>
<thead>
<tr>
<th>Item</th>
<th>$v$</th>
<th>$\hat{a}$</th>
<th>$a$</th>
<th>$b$</th>
<th>$v$</th>
<th>$a$</th>
<th>$b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.00</td>
<td>0.71</td>
<td>1.02</td>
<td>0.00</td>
<td>0.00</td>
<td>0.71</td>
<td>1.02</td>
</tr>
<tr>
<td>2</td>
<td>-0.34</td>
<td>0.40</td>
<td>0.44</td>
<td>-0.84</td>
<td>-0.34</td>
<td>0.40</td>
<td>0.44</td>
</tr>
<tr>
<td>3</td>
<td>-0.24</td>
<td>0.62</td>
<td>0.80</td>
<td>-0.38</td>
<td>-0.24</td>
<td>0.62</td>
<td>0.80</td>
</tr>
<tr>
<td>4</td>
<td>-0.37</td>
<td>0.66</td>
<td>0.88</td>
<td>-0.55</td>
<td>-0.37</td>
<td>0.66</td>
<td>0.88</td>
</tr>
<tr>
<td>5</td>
<td>-0.63</td>
<td>0.74</td>
<td>1.09</td>
<td>-0.85</td>
<td>-0.63</td>
<td>0.74</td>
<td>1.09</td>
</tr>
<tr>
<td>6</td>
<td>-0.32</td>
<td>0.84</td>
<td>1.52</td>
<td>-0.39</td>
<td>-0.41</td>
<td>0.79</td>
<td>1.27</td>
</tr>
<tr>
<td>7</td>
<td>-0.51</td>
<td>0.47</td>
<td>0.54</td>
<td>-1.07</td>
<td>-0.61</td>
<td>0.69</td>
<td>0.96</td>
</tr>
<tr>
<td>8</td>
<td>-0.04</td>
<td>0.24</td>
<td>0.24</td>
<td>-0.16</td>
<td>0.13</td>
<td>0.60</td>
<td>0.76</td>
</tr>
<tr>
<td>9</td>
<td>0.01</td>
<td>0.59</td>
<td>0.73</td>
<td>0.02</td>
<td>-0.43</td>
<td>0.76</td>
<td>1.18</td>
</tr>
<tr>
<td>10</td>
<td>0.14</td>
<td>0.15</td>
<td>0.15</td>
<td>0.95</td>
<td>-0.78</td>
<td>0.48</td>
<td>0.54</td>
</tr>
<tr>
<td>11</td>
<td>0.06</td>
<td>0.04</td>
<td>0.04</td>
<td>1.50</td>
<td>0.10</td>
<td>0.49</td>
<td>0.56</td>
</tr>
<tr>
<td>12</td>
<td>-0.87</td>
<td>0.53</td>
<td>0.63</td>
<td>-1.64</td>
<td>-1.46</td>
<td>0.66</td>
<td>0.87</td>
</tr>
<tr>
<td>13</td>
<td>-0.44</td>
<td>0.35</td>
<td>0.37</td>
<td>-1.24</td>
<td>0.09</td>
<td>0.73</td>
<td>1.08</td>
</tr>
<tr>
<td>14</td>
<td>0.08</td>
<td>0.70</td>
<td>0.99</td>
<td>0.97</td>
<td>0.03</td>
<td>0.80</td>
<td>1.35</td>
</tr>
<tr>
<td>15</td>
<td>0.73</td>
<td>0.53</td>
<td>0.62</td>
<td>1.39</td>
<td>0.18</td>
<td>0.79</td>
<td>1.27</td>
</tr>
</tbody>
</table>

Mean of the factor:

- Low-level employees: 1.61
- Middle-level employees: 1.40
- Directors: 1.35

Variance of the factor: 1.00

Note: The correlation between the $P$ and $O$ factors was .12 for low-level employees, .47 for middle-level employees, and .55 for directors. The variance of the factor is a fixed value equal in the three groups.
Figure 2. Item response functions and distribution of the factor scores for the P and O scales
Note: Solid lines stand for the P scale and dotted lines for the O scale. The model assumes that Items 1 to 5 have the same item response curve for the P and O scales, and so they are superimposed in the figure.
References