MULTIVARIATE BEHAVIORAL RESEARCH, *42*(2), 323–347

Identification and Small Sample Estimation of Thurstone's Unrestricted Model for Paired Comparisons Data

Alberto Maydeu-Olivares
Faculty of Psychology, University of Barcelona

Adolfo Hernández
Department of Mathematical Sciences, University of Exeter

The interpretation of a Thurstonian model for paired comparisons where the utilities' covariance matrix is unrestricted proved to be difficult due to the comparative nature of the data. We show that under a suitable constraint the utilities' correlation matrix can be estimated, yielding a readily interpretable solution. This set of identification constraints can recover any true utilities' covariance matrix, but it is not unique. Indeed, we show how to transform the estimated correlation matrix into alternative correlation matrices that are equally consistent with the data but may be more consistent with substantive theory. Also, we show how researchers can investigate the sample size needed to estimate a particular model by exploiting the simulation capabilities of a popular structural equation modeling statistical package.

Paired comparison experiments continue to be the most widely used tools for investigating choice behavior. Thurstone proposed in 1927 a model for paired comparisons data that remains to date the most influential model for choice modeling. Thurstone's model is characterized by three assumptions: (a) whenever a pair of stimuli is presented to a respondent it elicits a continuous preference (utility function, or in Thurstone's terminology, *discriminal process*) for each stimulus; (b) within a pair, the stimulus whose preference is larger will be preferred by the respondent; (c) the continuous preferences are normally distributed

Correspondence concerning this article should be addressed to Alberto Maydeu-Olivares, Faculty of Psychology, University of Barcelona, P. Valle de Hebrón, 171, 08035 Barcelona (Spain). E-mail: amaydeu@ub.edu

in the population. Since what Thurstone referred to as discriminal processes are in modern statistical terms latent variables, it seems possible to encompass this model within a more familiar structural equations modeling framework. To do so, we have to consider the sampling scheme used to gather the paired comparisons data.

Different sampling schemes can be used to gather paired comparisons data (Bock & Jones, 1968). The most popular approach is what these authors referred to as multiple judgment sampling. In this sampling scheme, all possible pairs of stimuli are presented to each respondent who is asked to choose one stimulus within each pair. Since each observed paired comparison is a binary variable and with n stimuli there are $\tilde{n} = \binom{n}{2} = \frac{n(n-1)}{2}$ paired comparisons, in a multiple judgment experiment for n stimuli there are $2^{\tilde{n}}$ possible patterns of binary paired comparisons. Therefore, Thurstone's model attempts to model these response patterns using n latent variables. This is because in his model there is one latent variable for each stimulus. These latent variables are *continuous latent preferences* (CLPs), discriminal processes, or utilities, and we use the three terms interchangeably in this article.

Of the possible $2^{\tilde{n}}$ patterns, $n!$ are transitive, meaning that given the binary patterns one can rank order the stimuli, and the rest intransitive. Maydeu-Olivares (1999) showed that Thurstone's model assigns a zero probability to all intransitive patterns. Takane must have been aware of this, as in 1987 he proposed adding a vector of pair specific random errors to Thurstone's model. This extension of Thurstone's model is a proper model for multiple judgment paired comparisons as it assigns non-zero probabilities to all paired comparisons patterns. Takane's (1987) crucial contribution was largely programmatic, and he provided neither identification restrictions nor empirical examples. Yet, it triggered a renewed interest in the field (Böckenholt, 2001a,b, 2004; Böckenholt & Dillon, 1997; Böckenholt & Tsai, 2001; Maydeu-Olivares, 2001, 2002, 2003, 2004; Maydeu-Olivares & Böckenholt, 2005; Tsai, 2000, 2003; Tsai & Böckenholt, 2001, 2002; Tsai & Wu, 2004).

Most applications of Thurstonian models have focused on the Case V and Case III special cases of the model. In the Case III model the continuous latent preferences are assumed to be independent. In the Case V model the CLPs are further assumed to have a common variance. Yet, it is possible to estimate models that do not assume that the CLPs are independent, such as models with unrestricted and factor-analytic covariance structures (Tsai & Böckenholt, 2001; Maydeu-Olivares & Böckenholt, 2005). Indeed, in this paper we focus on a Thurstonian model where the covariance structure of the CLPs is unrestricted. The unrestricted model is the most general model within the class of Thurstonian models and it plays a crucial role in paired comparison modeling (Maydeu-Olivares & Böckenholt, 2005). Should the unrestricted model provide a poor fit to the data,[1] then a model outside this class of models (or a Thurstonian model with latent classes) should be considered.

Thurstonian models for paired comparisons with correlated continuous latent preferences (discriminal processes) such as the unrestricted model have not been considered until recently. This is because to estimate these models by maximum likelihood it is necessary to evaluate high dimensional multivariate normal integrals (since the models involve binary observed variables and normally distributed CLPs). Yet, Thurstonian models are closely related to factor models for binary data (Maydeu-Olivares, 2001; Maydeu-Olivares & Böckenholt, 2005). Factor models for binary data can be estimated very efficiently using a multistage procedure that involves estimating in a first stage tetrachoric correlations among the binary variables. These multi-stage procedures are implemented in several statistical packages for structural equation modeling (SEM) such as Lisrel (Jöreskog & Sörbom, 2001), Mplus (Muthén & Muthén, 2001), and EQS (Bentler, 2004).

Recently, Maydeu-Olivares and Böckenholt (2005) have shown how to embed Thurstonian models for paired comparison data within a SEM framework including the classical Case V and Case III models, the unrestricted model, as well as models with a factor-analytic covariance structures. They also provide details on how to estimate these models using Mplus (Muthén & Muthén, 2001). In this paper, we extend Maydeu-Olivares and Böckenholt's results along three directions: (a) we provide a set of identified parameters for the unrestricted Thurstonian model that, unlike previous proposals, can be easily interpreted, (b) we show how the obtained solution can be transformed to explore the full array of solutions that is consistent with the data, and (c) we show how researchers can investigate the accuracy of the results obtained in an application by exploiting the simulation capabilities of Mplus.

Due to the comparative nature of paired comparisons data, the specification of identification restrictions for Thurstonian models, as well as the interpretation of the identified parameters, is not a trivial task. See Tsai (2003) for a thorough but technical discussion of this topic, and Steiger (2002) for a less technical overview of the problems that may arise when introducing identification constraints in structural equation models in general. For most Thurstonian models although a readily interpretable set of identified parameters can be found (Maydeu-Olivares & Böckenholt, 2005), yet no readily interpretable set of identified parameters had been given for the unrestricted model. In this paper we show how under suitable identification constraints, the correlation matrix among the continuous latent preferences can be estimated, leading to a solution that is readily interpretable.

[1]If multiple groups are available, researchers should first consider fitting a model with a separate covariance structure for each group.

In many important ways the problem of estimating an unrestricted Thurstonian model is analogous to the problem of estimating an unrestricted (a.k.a. exploratory) factor model for binary data. In the same way that the solution of an unrestricted factor model can be rotated to explore alternative solutions that may be more consistent with substantive theory, the solution of an unrestricted Thurstonian model can be transformed to explore alternative solutions that may be more consistent with substantive theory.

Also, in any application it is important to investigate the accuracy of the results obtained. This is particularly important in paired comparisons experiments as the number of binary variables to be modeled grows very fast when the number of stimuli to be compared increases. For instance, if a full paired comparisons design is used, there are 21 binary variables to be modeled when seven stimuli are compared. But when the number of stimuli is 10, the number of binary variables is 45. Yet, most often only small samples are typically collected in paired comparison experiments. Furthermore, n latent variables are used to model the data within a Thurstonian framework. Maydeu-Olivares (2003) reports a small simulation study where as few as 100 observations sufficed to estimate and test some Thurstonian models for seven stimuli. Larger samples were needed to accurately estimate other Thurstonian models for the same number of stimuli. The simulation focused on one of the many different special cases in Thurstone's model. Most importantly, as a reviewer of this manuscript pointed out, it is questionable to present results from one point in the parameter space and draw generalized conclusions from them. The multistage estimation procedures are so computationally efficient that it is possible to perform a simulation study to investigate the accuracy of the parameter estimates, standard errors and goodness of fit tests obtained in any particular application.

The remaining of this paper is organized as follows. First, the Thurstonian unrestricted model for paired comparisons model is discussed following Takane (1987). In this section, we show how the correlation matrix among the continuous latent preferences can be estimated, thus leading to a readily interpretable set of parameter estimates. We also show in this section how to transform the estimated solution into alternative correlation matrices among the CLPs that yield an equivalent fit. This enables researchers to search for equivalent solutions that may be more consistent with substantive theory. In the next section we briefly describe the multistage procedures used in structural equation modeling with binary dependent variables. In this section we go beyond Maydeu-Olivares and Böckenholt (2005) by discussing not only how to test the structural restrictions imposed by the model on the thresholds and tetrachoric correlations, but also how to test the overall restrictions imposed by the model (i.e, the structural restrictions and the dichotomized multivariate normality assumption). We conclude with an application where we briefly review how Thurstonian models can be estimated using Mplus following Maydeu-Olivares and Böckenholt (2005).

In this section we extend their work by showing how to make use of the simulation capabilities of Mplus to investigate the accuracy of the results obtained in any given application, and we illustrate how to identify an alternative solution that may be more consistent with substantive theory.

THURSTONIAN MODELS AS STRUCTURAL EQUATION MODELS WITH BINARY DEPENDENT VARIABLES

In this section Thurstone's (1927) original model is briefly described first. Then, Takane's (1987) extension of the model is introduced. The latter is better suited for modeling multiple judgment paired comparisons data. The section concludes with a discussion of the identification of the unrestricted Thurstonian model.

Thurstonian Models

Consider a set of n stimuli and a random sample of N individuals from the population we wish to investigate. In a multiple judgment paired comparison experiment $\tilde{n}$ pairs of stimuli are constructed and each pair is presented to each individual in the sample. We shall denote by y_k the outcome of each paired comparison. Thus, for each subject we let

$$y_k = \begin{cases} 1 & \text{if stimulus } i \text{ is chosen} \\ 0 & \text{if stimulus } i' \text{ is chosen} \end{cases} \qquad k = 1, \ldots, \tilde{n} \tag{1}$$

where $k \equiv (i, i')$, $(i < i' \leq n)$. Now, let t_i denote a subject's unobserved continuous latent preference for stimulus i. According to Thurstone's (1927) model: (a) the preferences $\mathbf{t} = (t_1, \ldots, t_n)'$ are normally distributed in the population, and (b) a subject will choose stimulus i if $t_i \geq t_{i'}$, otherwise s/he will choose stimulus i'.

Thurstone (1927) proposed performing the following linear transformation on the set of unobserved preferences

$$y_k^* = t_i - t_{i'} \tag{2}$$

Then, (b) may be alternatively expressed as

$$y_k = \begin{cases} 1 & \text{if } y_k^* \geq 0 \\ 0 & \text{if } y_k^* < 0 \end{cases} \tag{3}$$

In matrix notation, Thurstonian models can be expressed as follows: Let $\mathbf{t} \sim N(\boldsymbol{\mu}_t, \boldsymbol{\Sigma}_t)$, where $\boldsymbol{\mu}_t$ and $\boldsymbol{\Sigma}_t$ contain the n means of the unobserved preferences and corresponding (co)variances among the preferences, respectively. Also, $\mathbf{y}^* = \mathbf{At}$. $\mathbf{A}$ is a $\tilde{n} \times n$ design matrix where each column corresponds to one of the stimuli, and each row to one of the paired comparisons. For example, when $n = 4$, $\mathbf{A}$ is

$$\mathbf{A} = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 1 & 0 & -1 & 0 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & -1 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 \end{bmatrix} \tag{4}$$

Therefore the mean and covariance matrix of $\mathbf{y}^*$ are $\boldsymbol{\mu}_{y^*} = \mathbf{A}\boldsymbol{\mu}_t$, and $\boldsymbol{\Sigma}_{y^*} = \mathbf{A}\boldsymbol{\Sigma}_t\mathbf{A}'$, respectively. Notice that the variables $\mathbf{t}$ are latent variables. Also, the variables $\mathbf{y}^*$ are not observed and are multinormally distributed. Only the binary variables $\mathbf{y}$ are observed which under the model are obtained by dichotomizing $\mathbf{y}^*$ using Equation (3).

Maydeu-Olivares (1999) pointed out that because $\mathbf{A}$ is always of rank $n - 1$ (its columns add up to zero), $\boldsymbol{\Sigma}_{y^*}$ has rank $n - 1$ in these models. This in turn implies that Thurstonian models assign zero probabilities to all intransitive paired comparison patterns. Thus, Thurstonian models are better suited for ranking data (which by design are transitive) than for multiple judgment paired comparisons data.

Takane (1987) proposed adding a random error e_k to each paired comparison in expression (2). The addition of these errors enables the model to account for intransitivities that may be observed in the paired comparisons and yields a model that assigns non-zero probabilities to all binary paired comparisons patterns. With the addition of these pair specific errors, we write

$$\mathbf{y}^* = \mathbf{At} + \mathbf{e}. \tag{5}$$

The errors $\mathbf{e}$ are assumed normally distributed, independent of each other, and independent of the continuous latent preferences $\mathbf{t}$. We may write these assumptions as

$$\begin{pmatrix} \mathbf{t} \\ \mathbf{e} \end{pmatrix} \sim N\left(\begin{pmatrix} \boldsymbol{\mu}_t \\ \mathbf{0} \end{pmatrix}, \begin{pmatrix} \boldsymbol{\Sigma}_t & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Omega}^2 \end{pmatrix}\right) \tag{6}$$

where $\boldsymbol{\Omega}^2$ is a diagonal matrix with diagonal element $\omega_1^2, \ldots, \omega_{\tilde{n}}^2$ containing the variances of the random errors $\mathbf{e}$. From Equations (5) and (6), the mean vector and covariance matrix of the unobserved variables $\mathbf{y}^*$ are

$$\boldsymbol{\mu}_{y^*} = \mathbf{A}\boldsymbol{\mu}_t \quad \text{and} \quad \boldsymbol{\Sigma}_{y^*} = \mathbf{A}\boldsymbol{\Sigma}_t\mathbf{A}' + \boldsymbol{\Omega}^2. \tag{7}$$

Under the model, the binary paired comparisons are obtained by dichotomizing the multivariate normal variables $\mathbf{y}^*$ using Equation (3). Yet, dichotomizing the multivariate normal distribution of $\mathbf{y}^*$ with mean and covariance structure using Equation (7) is equivalent to dichotomizing a $\tilde{n}$-dimensional vector of random variables, say $\mathbf{z}^*$, which is multivariate normal with mean zero and correlation structure

$$\mathbf{P}_{z^*} = \mathbf{D}\boldsymbol{\Sigma}_{y^*}\mathbf{D} = \mathbf{D}(\mathbf{A}\boldsymbol{\Sigma}_t\mathbf{A}' + \boldsymbol{\Omega}^2)\mathbf{D} \tag{8}$$

using

$$y_k = \begin{cases} 1 & \text{if } z_k^* \geq \tau_k \\ 0 & \text{if } z_k^* < \tau_k \end{cases}. \tag{9}$$

In Equation (8), $\mathbf{D} = (\text{Diag}(\boldsymbol{\Sigma}_{y^*}))^{-\frac{1}{2}}$ is a diagonal matrix containing the inverse of the standard deviations of $\mathbf{y}^*$ under the model. Further, the vector of thresholds $\boldsymbol{\tau} = (\tau_1, \ldots, \tau_{\tilde{n}})'$ in (9) has the following structure under the model:

$$\boldsymbol{\tau} = -\mathbf{D}\boldsymbol{\mu}_{y^*} = -\mathbf{D}\mathbf{A}\boldsymbol{\mu}_t. \tag{10}$$

As a result of this equivalence, estimating a Thurstonian model for paired comparisons is equivalent to estimating a structural equation model for dichotomous variables where the matrix of tetrachoric correlations is structured as in Equation (8) and the thresholds are structured as in Equation (10).

Equations (8) and (10) define in fact a class of models as $\boldsymbol{\mu}_t$ and $\boldsymbol{\Sigma}_t$ can be restricted in various ways. For an overview of restricted Thurstonian models, see Takane (1987) and Maydeu-Olivares and Böckenholt (2005). The present paper focuses on the unrestricted Thurstone-Takane model where the mean vector $\boldsymbol{\mu}_t$ and the covariance matrix $\boldsymbol{\Sigma}_t$ are left unrestricted.

Identification of the Unrestricted Thurstonian Model

Consider a Thurstonian model for paired comparisons data where $\boldsymbol{\Sigma}_t$ is symmetric positive definite matrix. For $n \geq 3$, it follows from Tsai (2003) that $n + 2$ constraints must be introduced to identify this model. Because of the comparative nature of the data, one constraint needs to be imposed among the elements of $\boldsymbol{\mu}_t$. Also, n constraints need to be imposed on the elements of $\boldsymbol{\Sigma}_t$ such that one constraint is imposed among the elements in each of the rows (columns) of this matrix. These location constraints are needed because $\mathbf{A}$ is of rank $n - 1$. An additional constraint must be introduced among the elements of $\boldsymbol{\Sigma}_t$ to set the scale for the remaining parameters.

Maydeu-Olivares and Böckenholt (2005) gave one set of identification constraints that satisfies these requirements. They suggested: (1) fix the mean preference for one of the stimuli, say $\mu_n = 0$; (2) fix all the covariances involving the last stimuli to 0; and (3) fix the variance of the preferences for the first and last stimuli to 1. For example, when $n = 4$ $\boldsymbol{\mu}_t$ and $\boldsymbol{\Sigma}_t$ are to be specified as

$$\boldsymbol{\mu}_t = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \\ 0^* \end{pmatrix}, \qquad \boldsymbol{\Sigma}_t = \begin{pmatrix} 1^* & \sigma_{21} & \sigma_{31} & 0^* \\ \sigma_{21} & \sigma_2^2 & \sigma_{32} & 0^* \\ \sigma_{31} & \sigma_{32} & \sigma_3^2 & 0^* \\ 0^* & 0^* & 0^* & 1^* \end{pmatrix}, \tag{11}$$

where the parameters fixed for identification are marked with an asterisk. The interpretation of the estimated means relative to a fixed one presents no particular problems. However, interpreting variances and covariances relative to other fixed variances and covariances is rather cumbersome.

An equivalent way to identify the parameters of the covariance structure that avoids the interpretation problem in Equation (11) is obtained as follows: First, all the diagonal elements of $\boldsymbol{\Sigma}_t$ are set equal to one (i.e., $\sigma_i^2 = 1, \forall i$). This set of constraints is convenient from an applied perspective because $\boldsymbol{\Sigma}_t$ becomes a correlation matrix, $\mathbf{P}_t$, which facilitates the interpretation. Yet, an additional linear constraint must be enforced among the elements of $\mathbf{P}_t$, such as $\sum_{i=2}^{n} \rho_{i1} = 1$ (i.e., $\rho_{21} + \rho_{31} + \cdots + \rho_{n1} = 1$).[2] With this constraint we obtain parameter estimates and standard errors for all the elements in $\mathbf{P}_t$, the correlation matrix between the continuous latent preferences.[3] Also, $\mathbf{P}_t$ is at least non-negative definite and therefore its elements $\rho_{ii'}$ are at least admissible, that is, they are bounded between -1 and 1 (see Appendix 1).

Thus, given the identification constraints, the elements of $\mathbf{P}_t$, $\rho_{ii'}$, can be interpreted as follows. A positive correlation implies that strong preferences for stimuli i are associated with strong preferences for stimuli i'. In contrast, when the correlation in negative strong preferences for stimuli i are associated with weak preferences for stimuli i'. In many applications it is important to be able to estimate these correlations as they suggest which stimuli can act as substitutes for other stimuli.

Analogy with Exploratory Factor Analysis

In many ways, the problems of identifying and estimating an unrestricted Thurstonian model for paired comparisons data is similar to the problems of identify-

[2]A simpler constraint is $\rho_{n1} = 0$, see Appendix 1.

[3]A similar constraint can be enforced among the means, say $\sum_{i=1}^{n} \mu_i = 0$, instead of using $\mu_n = 0$.

ing and estimating an unrestricted factor model for binary data with n factors, where n is the number of stimuli being compared. Thus, in the Thurstonian unrestricted model (a) the contrast matrix $\mathbf{A}$ is analogous to the factor loadings matrix, (b) the means of the continuous latent preferences $\boldsymbol{\mu}_t$. are analogous to the factor means, (c) the matrix of correlations among the preferences $\mathbf{P}_t$ is analogous to the matrix of inter-factor correlations, and (d) the diagonal matrix $\boldsymbol{\Omega}^2$ containing the variances of the pair specific errors is analogous to the diagonal matrix containing the variances of the unique factors.

The easiest way to obtain a set of identified parameters in an unrestricted factor model is to (a) set the factors to be uncorrelated, and (b) set the upper triangular part of the factor loading matrix equal to zero (McDonald, 1999: p. 181). The factor means are not identified in an unrestricted factor model and they are set to zero. In an unrestricted Thurstonian model the 'factor means' $\boldsymbol{\mu}_t$. can be estimated because the 'factor loadings matrix' $\mathbf{A}$ is a matrix of contrast. To identify the Thurstonian model, we fix one of the 'factor means' to zero, and we impose a linear constraint among the 'inter-factor correlations'.

For an unrestricted factor model with estimated matrices of factor loadings $\boldsymbol{\Lambda}$ and inter-factor correlations $\boldsymbol{\Phi}$, it is well known that alternative solutions with $\tilde{\boldsymbol{\Lambda}}$ and $\tilde{\boldsymbol{\Phi}}$ that yield an equivalent fit can be obtained by rotating the axes. Analogously, for an unrestricted Thurstonian model with estimated matrices of preferences means $\boldsymbol{\mu}_t$, preferences intercorrelations $\mathbf{P}_t$, and pair specific error variances $\boldsymbol{\Omega}^2$, applied researchers can obtain alternative solutions that yield an equivalent fit with matrices $\tilde{\boldsymbol{\mu}}_t$, $\tilde{\mathbf{P}}_t$, and $\tilde{\boldsymbol{\Omega}}^2$ using the equation

$$\tilde{\boldsymbol{\mu}}_t = \sqrt{c}\boldsymbol{\mu}_t, \qquad \tilde{\mathbf{P}}_t = c\mathbf{P}_t + (1-c)\mathbf{1}\mathbf{1}', \qquad \tilde{\boldsymbol{\Omega}}^2 = c\boldsymbol{\Omega}^2, \tag{12}$$

where c is any positive constant such that $\tilde{\mathbf{P}}_t$ is a correlation matrix (i.e., it is non-negative definite). This is proved in Appendix 2.

Thus, just as when fitting an unrestricted factor analysis, the applied researcher that fits an unrestricted Thurstonian model must bear in mind that the solution obtained is not the only one that is consistent with the data. Other solutions exist that will yield the same fit to the data. For the unrestricted factor model, applied researchers often use rotation algorithms (e.g., Varimax, Oblimin) to aid explore alternative solutions that may be more consistent with substantive theory. Exploring alternative solutions that may be more consistent with substantive theory is easier for the unrestricted Thurstonian model as applied researchers just need to plug in different values of c in Equation (12). However, even though this equation plays the same role as factor rotation in factor analysis, it is not a rotation. Also, just as in factor analysis it is possible to incorporate any substantive knowledge into the model and resort to confirmatory factor analysis, in Thurstonian modeling the applied researcher may wish to incorporate any substantive knowledge into the model, and estimate a model where some elements of the parameter matrices $\boldsymbol{\mu}_t$, $\mathbf{P}_t$, and $\boldsymbol{\Omega}^2$ are fixed at a priori values, or where some parameters are set equal to other parameters in the model.

SEM ESTIMATION OF THURSTONIAN MODELS FOR PAIRED COMPARISONS DATA

Among the most popular statistical packages for structural equation modeling, EQS (Bentler, 2004), Lisrel (Jöreskog & Sörbom, 2001) and Mplus (Muthén & Muthén, 2001) implement multistage procedures for estimating structural equation models with binary dependent variables. In these methods, first the thresholds and tetrachoric correlations are estimated from the univariate and bivariate margins of the contingency table. In a second stage, the structural parameters are estimated from the estimated thresholds and tetrachoric correlations. In this section, the estimation of Thurstonian models for paired comparisons data is described.

Let p_k and $p_{kk'}$ be the sample counterpart of $\pi_k = \Pr(y_k = 1)$ and $\pi_{kk'} = \Pr(y_k = 1, y_{k'} = 1)$, respectively, and let $\Phi_n(\bullet)$ denote a n-variate standard normal distribution function. Then, first each element of $\boldsymbol{\tau}$ is estimated separately using $\hat{\tau}_k = -\Phi_1^{-1}(p_k)$. Next, each tetrachoric correlation is estimated separately given the first stage estimates using $\hat{\rho}_{kk'} = \Phi_2^{-1}(p_{kk'} \mid -\hat{\tau}_k, -\hat{\tau}_{k'})$. This method for estimating the tetrachoric correlations is equivalent to those implemented in Lisrel and Mplus, but not in EQS.

Then, letting $\boldsymbol{\kappa} = (\boldsymbol{\tau}', \boldsymbol{\rho}')'$, the model parameters $\boldsymbol{\theta}$ are estimated by minimizing

$$F = (\hat{\boldsymbol{\kappa}} - \boldsymbol{\kappa}(\boldsymbol{\theta}))'\hat{\mathbf{W}}(\hat{\boldsymbol{\kappa}} - \boldsymbol{\kappa}(\boldsymbol{\theta})) \tag{13}$$

where $\hat{\mathbf{W}} = \hat{\boldsymbol{\Xi}}^{-1}$ (weighted least squares, WLS: Muthén, 1978), $\hat{\mathbf{W}} = (\mathrm{Diag}(\hat{\boldsymbol{\Xi}}))^{-1}$ (diagonally weighted least squares, DWLS: Muthén, du Toit, & Spisic, 1997), or $\hat{\mathbf{W}} = \mathbf{I}$ (unweighted least squares, ULS: Muthén, 1993), and $\boldsymbol{\Xi}$ denotes the asymptotic covariance (Acov) matrix of $\sqrt{N}(\hat{\boldsymbol{\kappa}} - \boldsymbol{\kappa})$.

Standard errors for the parameter estimates are obtained using $\mathrm{Acov}(\hat{\boldsymbol{\theta}}) = \frac{1}{N}\mathbf{H}\boldsymbol{\Xi}\mathbf{H}'$, where $\mathbf{H} = (\boldsymbol{\Delta}'\mathbf{W}\boldsymbol{\Delta})^{-1}\boldsymbol{\Delta}'\mathbf{W}$, and $\boldsymbol{\Delta} = \frac{\partial\boldsymbol{\kappa}}{\partial\boldsymbol{\theta}'}$ (Muthén, 1993). WLS estimation is known to perform poorly in small samples (Muthén, 1993). Also, ULS performs better than DWLS estimation in small samples (Maydeu-Olivares, 2001). Thus, in this paper ULS will be employed.

Goodness of fit tests of the restrictions imposed by the model on the thresholds and tetrachoric correlations for the DWLS and ULS estimators can be obtained (Muthén, 1993; Satorra & Bentler, 1994) by scaling $T := N\hat{F}$ by its mean or adjusting it by its mean and variance so that it approximates a chi-square distribution using $T_s = \frac{r}{\text{Tr}[\mathbf{M}]}T$, and $T_a = \frac{\text{Tr}[\mathbf{M}]}{\text{Tr}[\mathbf{M}^2]}T$, where $\mathbf{M} = \mathbf{W}(\mathbf{I} - \mathbf{\Delta H})\mathbf{\Xi}$. T_s and T_a, denote the scaled (for mean) and adjusted (for mean and variance) test statistics. T_s is referred to a chi-square distribution with $r = \frac{\tilde{n}(\tilde{n}+1)}{2} - q$ degrees of freedom, where q is the number of mathematically independent parameters in $\boldsymbol{\theta}$. T_a is referred to a chi-square distribution with $d = \frac{(\text{Tr}[\mathbf{M}])^2}{\text{Tr}[\mathbf{M}^2]}$ degrees of freedom.

It is important to bear in mind, however, that these test statistics do not assess how well the model reproduces the data. Rather, they assess how well the model reproduces the estimated thresholds and tetrachoric correlations. These statistics are computed under a dichotomized multivariate normality assumption. Muthén and Hofacker (1988) proposed a test of dichotomized normality for triplets of variables. However, this test is not implemented in Mplus, Lisrel, nor EQS. Most importantly, it is not clear what to conclude if the assumption of dichotomized normality is tenable for some triplets of variables, but not for all of them. Rather than testing separately the dichotomized multivariate normality underlying the use of tetrachoric correlations, and testing the structural restrictions imposed by the model on the tetrachoric correlations, Maydeu-Olivares (2001) proposed a test of the overall restrictions (structural and distributional) imposed by a SEM model on the dichotomous data. Let $\mathbf{p}$ be the $\frac{\tilde{n}(\tilde{n}+1)}{2}$ vector of first and second order proportions, and let $\boldsymbol{\pi}$ be its corresponding probabilities. Furthermore, let $(\mathbf{p} - \boldsymbol{\pi}(\hat{\boldsymbol{\theta}}))$ be the vector of residual univariate and bivariate proportions under the model. Then, goodness of fit tests of the overall restrictions imposed by the model on the first and second order marginals of the contingency table $\boldsymbol{\pi}(\boldsymbol{\theta})$ can be obtained by scaling $\tilde{T} = N(\mathbf{p} - \boldsymbol{\pi}(\hat{\boldsymbol{\theta}}))'(\mathbf{p} - \boldsymbol{\pi}(\hat{\boldsymbol{\theta}}))$ using $\tilde{T}_s = \frac{r}{\text{Tr}[\tilde{\mathbf{M}}]}\tilde{T}$, and $\tilde{T}_a = \frac{\text{Tr}[\tilde{\mathbf{M}}]}{\text{Tr}[\tilde{\mathbf{M}}^2]}\tilde{T}$, where $\tilde{\mathbf{M}} = (\mathbf{I} - \tilde{\mathbf{\Delta}}\mathbf{\Delta H}\tilde{\mathbf{\Delta}}^{-1})\mathbf{\Gamma}(\mathbf{I} - \tilde{\mathbf{\Delta}}\mathbf{\Delta H}\tilde{\mathbf{\Delta}}^{-1})'$, $\mathbf{\Gamma}$ denotes the asymptotic covariance matrix of $\sqrt{N}(\mathbf{p} - \boldsymbol{\pi})$, and $\tilde{\mathbf{\Delta}} = \frac{\partial \boldsymbol{\pi}}{\partial \boldsymbol{\kappa}'}$. $\tilde{T}_s$ and $\tilde{T}_a$ are to be referred to a chi-square distribution with r and $d = \frac{(\text{Tr}[\tilde{\mathbf{M}}])^2}{\text{Tr}[\tilde{\mathbf{M}}^2]}$ degrees of freedom, respectively. A small simulation study reported in Maydeu-Olivares (2001) revealed that for some models the mean and variance adjusted statistic yields accurate Type I error rates in models with 21 variables with as few as 100 observations.

AN APPLICATION: MODELING PREFERENCES FOR CELEBRITIES

Kroeger (1992) replicated a classical experiment by Rumelhart and Greeno (1971) in which college students were presented with pairs of celebrities and they were asked to select the celebrity with whom they would rather spend an hour of conversation. Here we shall analyze a subset of Kroeger's data consisting of the females' responses (96 subjects) to the paired comparisons involving the set of former U.S. first ladies (Barbara Bush, Nancy Reagan, and Hillary Clinton) and athletes (Bonnie Blair, Jackee Joyner-Kersee, and Jennifer Capriati). Because no prior knowledge is available, we shall estimate an unrestricted Thurstonian model.

When a single population is involved, Mplus but not current versions[4] of Lisrel and EQS can estimate single population models for categorical dependent variables with mean or threshold structures (as required for Thurstonian modeling of paired comparisons). Consequently, Mplus will be used in this example. The Mplus input file used in this example is provided in Appendix 3.

Fitting an Unrestricted Thurstonian Model

Estimating a Thurstonian unrestricted model for paired comparisons data is similar to estimating a factor model for binary data with n factors, where n is the number of stimuli being compared. Thus, to estimate the Thurstonian unrestricted model the factor loadings must be fixed constants, those of the contrast matrix **A**. Also, the factor means are estimated. This is mean vector of the stimuli preferences $\boldsymbol{\mu}_t$. Finally, the variances of the unique factors are estimated. This is the diagonal matrix $\boldsymbol{\Omega}^2$ containing the variances of the pair specific errors. To identify the model, one of the factor means is fixed to zero. Also, the factor variances are fixed to 1, so that the correlations among the preferences for the stimuli are estimated. This is the matrix $\mathbf{P}_t$. Finally, a linear constraint needs to be imposed among the inter-factor correlations. See Appendix 3 for details on how to implement this constraint in Mplus. For this example, we fixed the mean of the last stimuli to zero, and imposed the constraint that the inter-factor correlations involving the first stimuli added to one.

For ULS estimation, only the mean and variance corrected statistic for testing the restrictions imposed by the model on the thresholds and tetrachoric correlations is available in Mplus. We obtained $T_a = 43.28$ on 32 df,[5] $p = 0.09$. It appears as if the structural restrictions imposed by this model be reasonable. However, an improper solution is obtained as the estimates for two of the diagonal elements of $\boldsymbol{\Omega}^2$ are negative. This is most likely due to the sample size being too small to estimate this model.

As an alternative, we considered a model where the variances of the pair specific errors are equal for all pairs, $\boldsymbol{\Omega}^2 = \omega^2\mathbf{I}$. This restriction constraints

[4] At the time of this writing, the current versions of Mplus, Lisrel and EQS are 3.12, 8.7, and 6.1, respectively.

[5] When mean and variance corrections are used, the number of degrees of freedom is estimated. This is a real number, which in Mplus is rounded to the nearest integer. For this model, the difference between the number of simple statistics and the number of estimated parameters is 86. This is highly constrained model.

the expected number of intransitivities to be approximately equal for all pairs (Maydeu-Olivares & Böckenholt, 2005). For this model, which has 14 parameters less than the previous model, we obtained $T_a = 44.57$ on 33 df, $p = 0.09$. Also, the estimate of the common variance of the pair specific errors is a proper value, 0.24. The more restricted model is clearly to be preferred. Finally, using the test statistic proposed by Maydeu-Olivares (2001) evaluated at the Mplus estimates for this model we can also assess whether the model reproduces the data. We obtained $\tilde{T}_s = 124.62$ on 100 df, $p = 0.05$, and $\tilde{T}_a = 20.05$ on 16.09 df, $p = 0.22$. The model reproduces well the paired comparisons patterns.

The estimates and standard errors for the parameters of this model are shown in Table 1. Taking into account the estimated standard errors, the ordering of the mean preferences for these celebrities under this model is {Hillary Clinton, Jackee Joyner = Barbara Bush = Jennifer Capriati, and Bonnie Blair = Nancy Reagan}. Yet, preferences for the most preferred celebrity, Hillary Clinton, are not significantly correlated with preferences for any other celebrity. On the other hand, preferences for Bonnie Blair are significantly and positively related to preferences for all the other celebrities (but Hillary Clinton). In addition, women that express their preference for Barbara Bush are significantly more likely to prefer Jennifer Capriati as well.

TABLE 1
Parameter Estimates and Standard Errors for an Unrestricted Covariance Structure Model Assuming $\boldsymbol{\Omega}^2 = \omega^2\mathbf{I}$ Applied to the Celebrities Data

$\boldsymbol{P}_t$	*Barbara Bush*	*Nancy Reagan*	*Hillary Clinton*	*Bonnie Blair*	*Jackee Joyner*	*Jennifer Capriati*
Barbara Bush	1					
	(fixed)					
Nancy Reagan	**0.74**	1				
	(0.09)	*(fixed)*				
Hillary Clinton	−0.22	−0.37	1			
	(0.24)	(0.29)	*(fixed)*			
Bonnie Blair	**0.24**	**0.35**	−0.19	1		
	(0.12)	(0.14)	(0.28)	*(fixed)*		
Jackee Joyner	−0.01	0.04	−0.47	**0.37**	1	
	(0.16)	(0.18)	(0.32)	(0.16)	*(fixed)*	
Jennifer Capriati	**0.26**	0.21	−0.56	**0.44**	0.32	1
	(0.13)	(0.19)	(0.33)	(0.15)	(0.20)	*(fixed)*
$\boldsymbol{\mu}_t$	0.09	−0.28	0.43	−0.27	0.16	0
	(0.15)	(0.15)	(0.20)	(0.13)	(0.14)	*(fixed)*

$N = 96$; $\rho_{21} + \rho_{31} + \cdots + \rho_{n1} = 1$; standard errors in parentheses; correlations in bold are significant at $\alpha = 0.05$; $\hat{\omega}^2 = 0.26$ (0.06).

Transformation of the Solution

Just as an unrestricted factor model can be rotated to explore alternative solutions that may be more consistent with substantive theory, the unrestricted Thurstonian model can be transformed to explore alternative solutions that may be more consistent with substantive theory. To transform the solution shown in Table 1 we just need to search through the set of solutions given in Equation (12). In this example, we believe that stronger preferences for Hillary Clinton are associated with weaker preferences for Nancy Reagan. Consequently we shall transform the solution to achieve the maximum negative correlation between these two stimuli. Assigning values greater than 1 to c in Equation (12) the correlation between Nancy Reagan and Hillary Clinton becomes increasingly negative. The maximum value that c can take in this application resulting in a positive definite transformed $\tilde{\mathbf{P}}_t$ is 1.10. When c takes this value, the correlation between the preferences for Nancy Reagan and Hillary Clinton becomes −0.51, this is the largest negative correlation for these two stimuli that results in a proper model. Now, we can obtain standard errors for this transformed solution by simply re-estimating the model using the code given in Appendix 3 where the constraint $\sum_{i=2}^{n} \rho_{i1} = 1$ used to obtain the initial solution is replaced by the constraint corr(Nancy Reagan, Hillary Clinton) = −0.51. The results of this transformed solution are given in Table 2. This solution yields the same goodness of fit as the solution presented in Table 1, but is slightly easier to interpret. According to the solution, for this population of female college students the most preferred celebrity to spend an hour of conversation with is Hillary Clinton (as it has the highest mean preference). The least preferred ones are Bonnie Blair and Nancy Reagan. Mean preferences for the remaining celebrities are roughly equal and in between these extremes. We also observe two significant associations in Table 2. Students who prefer to spend an hour of conversation with Nancy Reagan as more likely to prefer an hour of conversation with Barbara Bush, and students who prefer to spend an hour with Hillary Clinton are less likely to prefer spending an hour with Jennifer Capriati.

Accuracy of the Results

The small sample size (96) relative to the number of binary paired comparisons modeled (15) casts doubts on the reliability of these results. The Monte Carlo capabilities of Mplus can be used to address this concern since a simulation study can be performed using as true parameter values the estimated parameters of the model. Using the Mplus input file provided in Appendix 4, we performed a simulation study to investigate the accuracy of the parameter estimates, standard errors, and goodness of fit tests obtained in our initial model where $\sum_{i=2}^{n} \rho_{i1} = 1$. The simulation of 1000 samples of size 96 took 56 seconds on a 3 Mz machine.

TABLE 2
Estimated Parameter and Standard Errors for the Transformed Solution

$\boldsymbol{P}_t$	*Barbara Bush*	*Nancy Reagan*	*Hillary Clinton*	*Bonnie Blair*	*Jackee Joyner*	*Jennifer Capriati*
Barbara Bush	1					
	(fixed)					
Nancy Reagan	**0.71**	1				
	(0.11)	*(fixed)*				
Hillary Clinton	−0.35	−0.51	1			
	(0.18)	*(fixed)*	*(fixed)*			
Bonnie Blair	0.16	0.28	−0.19	1		
	(0.26)	(0.22)	(0.28)	*(fixed)*		
Jackee Joyner	−0.12	−0.05	−0.62	0.31	1	
	(0.36)	(0.31)	(0.38)	(0.25)	*(fixed)*	
Jennifer Capriati	0.18	0.13	**−0.72**	0.38	0.25	1
	(0.27)	(0.39)	(0.33)	(0.21)	(0.30)	*(fixed)*
$\boldsymbol{\mu}_t$	0.09	−0.29	0.46	−0.28	0.17	0
	(0.15)	(0.16)	(0.21)	(0.15)	(0.15)	*(fixed)*

The solution was transformed so that the correlation between preferences for Nancy Reagan and Hillary Clinton was maximally negative; standard errors in parentheses; correlations in bold are significant at $\alpha = 0.05$; $\hat{\omega}^2 = 0.28$ (0.09).

All replications converged. The bias of the parameter estimates, $\overline{x}_{\hat{\theta}} - \theta_0$, ranged from −0.02 to 0.01 with an average of 0. The bias of the standard errors, $\overline{x}_{SE(\hat{\theta})} - sd_{\hat{\theta}}$, ranged from −0.01 to 0.01, with an average of 0. The coverage of the 95% confidence intervals for the parameter estimates ranged from 0.92 to 0.96, with an average or 0.94. Clearly, the parameter estimates and standard errors we have obtained in our small sample application are to be trusted. As for the behavior of the test statistic implemented in Mplus, the empirical rejection rates at $\alpha = \{0.01, 0.05, 0.10, 0.20\}$ across the 1000 replications were {0.01, 0.03, 0.09, 0.18}. The p-values yielded by the test statistic are reliable. The model indeed reproduces the estimated thresholds and tetrachoric correlations.

DISCUSSION AND CONCLUSIONS

Maydeu-Olivares and Böckenholt (2005) have recently shown how to embed the class of Thurstonian models for paired comparison data within a SEM framework. They also provide details on how to estimate these models using Mplus. Yet, the comparative nature of paired comparisons data causes problems of parameter interpretation in these models. For the Case V, Case III, as well as Thurstonian factor analytic models Maydeu-Olivares and Böckenholt (2005) show how to suitably identify these models to obtain interpretable parameters. Yet, the interpretation of the parameters of the unrestricted Thurstonian model proved difficult. This model plays a crucial role in Thurstonian modeling, as it is the most general model within this class. In this paper, we have extended Maydeu-Olivares and Böckenholt (2005) results by providing a set of identification restrictions for this model that yield an easily interpretable solution. Thus, we have shown that the correlations among the continuous latent preferences can be estimated, if one constraint is imposed among them. With this constraint, the correlation matrix of the CLPs is at least non-negative definitive. Given the close connection between Thurstonian models for paired comparisons and ranking data (see Maydeu-Olivares and Böckenholt, 2005), this specification of the unrestricted Thurstonian model can also be applied to ranking data.

The solution obtained is not the only one that is consistent with the data. Alternative solutions can be obtained by transforming the solution initially obtained that yield an equivalent fit to the data. A formula has been provided that enables searching for alternative solutions that may be more consistent with substantive theory.

Also, we have shown in this paper how applied researchers can use Mplus' simulation capabilities to investigate the accuracy of the parameter estimates, standard errors, and goodness of fit tests obtained using this program for their particular application. Unfortunately, researchers can currently test with Mplus only the restrictions imposed by their structural models on the estimated thresholds and tetrachoric correlations. This test need not be meaningful if the assumption of categorized normality does not hold. It is possible to test the overall restrictions imposed by Thurstonian models on the observed binary data (Maydeu-Olivares, 2001) and existing evidence suggest that these tests perform as well as the structural restrictions tests in small samples. However, the tests of the overall restrictions are yet to be implemented in Mplus.

To sum up, framing the analysis of paired comparisons data within a structural equation modeling framework enables applied researchers to conveniently and reliably estimate and test the full array of Thurstonian models. It is hoped that the widespread availability of software to estimate structural equation modeling will encourage further use of these experimental designs in applications.

ACKNOWLEDGMENTS

The authors are very much indebted to the reviewers for comments that led to substantial improvements to the manuscript. This research was supported by the Dept. of Universities, Research and Information Society (DURSI) of the Catalan Government, and by grants BSO2000-0661 and BSO2003-08507 from the Spanish Ministry of Science and Technology.

REFERENCES

Bentler, P. M. (2004). *EQS*. Encino, CA: Multivariate Software Inc.

Bock, R. D., & Jones, L. V. (1968). *The measurement and prediction of judgment and choice*. San Francisco: Holden-Day.

Böckenholt, U. (2001a). Hierarchical modeling of paired comparison data. *Psychological Methods*, *6*, 49–66.

Böckenholt, U. (2001b). Thresholds and intransitivities in pairwise judgments: A multilevel analysis. *Journal of Educational and Behavioral Statistics*, *26*, 269–282.

Böckenholt, U. (2004). Comparative judgments as an alternative to ratings: Identifying the scale origin. *Psychological Methods*, *4*, 453–465.

Böckenholt, U., & Dillon, W. R. (1997). Modeling within-subject dependencies in ordinal paired comparison data. *Psychometrika*, *62*, 411–434.

Böckenholt, U., & Tsai, R. C. (2001). Individual differences in paired comparison data. *British Journal of Mathematical and Statistical Psychology*, *54*, 265–277.

Dudgeon, P., Bell, R., & Pattison, P. (2003). *Imposing constraints on Cholesky-decomposed covariance matrices*. Paper presented at the International Meeting of the Psychometric Society, Cagliari.

Jöreskog, K.G., & Sörbom, D. (2001). *LISREL 8*. Chicago, IL: Scientific Software.

Kroeger, K. (1992). *Unpublished dataset*. University of Illinois, Urbana-Champaign.

Maydeu-Olivares, A. (1999). Thurstonian modeling of ranking data via mean and covariance structure analysis. *Psychometrika*, *64*, 325–340.

Maydeu-Olivares, A. (2001). Limited information estimation and testing of Thurstonian models for paired comparison data under multiple judgment sampling. *Psychometrika*, *66*, 209–228.

Maydeu-Olivares, A. (2002). Limited information estimation and testing of Thurstonian models for preference data. *Mathematical Social Sciences*, *43*, 467–483.

Maydeu-Olivares, A. (2003). On Thurstone's model for paired comparisons and ranking data. In H. Yanai, A. Okada, K. Shigematu, Y. Kano, & J. J. Meulman (Eds.), *New developments in psychometrics* (pp. 519–526). Tokyo: Springer.

Maydeu-Olivares, A. (2004). Thurstone's Case V model: A structural equation modeling perspective. In K. van Monfort, J. Oud, & A. Satorra (Eds.), *Recent developments on structural equation models* (pp. 41–67). Dordrecht: Kluwer.

Maydeu-Olivares, A., & Böckenholt, U. (2005). Structural equation modeling of paired comparison and ranking data. *Psychological Methods*, *10*, 285–304.

McDonald, R. P. (1999). *Test theory. A unified approach*. Mahwah, NJ: Lawrence Erlbaum.

Muthén, B. (1978). Contributions to factor analysis of dichotomous variables. *Psychometrika*, *43*, 551–560.

Muthén, B. (1993). Goodness of fit with categorical and other non normal variables. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 205–234). Newbury Park, CA: Sage.

Muthén, B., du Toit, S. H. C., & Spisic, D. (1997). Robust inference using weighted least squares and quadratic estimating equations in latent variable modeling with categorical and continuous outcomes. Unpublished manuscript. College of Education, UCLA. Los Angeles, CA.

Muthén, B., & Hofacker, C. (1988). Testing the assumptions underlying tetrachoric correlations. *Psychometrika*, *53*, 563–578.

Muthén, L., & Muthén, B. (2001). *MPLUS*. Los Angeles, CA: Muthén & Muthén.

Rummelhart, D. L., & Greeno, J. G. (1971). Similarity between stimuli: An experimental test of the Luce and Restle choice models. *Journal of Mathematical Psychology*, *8*, 370–381.

Satorra, A., & Bentler, P. M. (1994). Corrections to test statistics and standard errors in covariance structure analysis. In A. von Eye, & C. C. Clogg (Eds.), *Latent variable analysis: Applications to developmental research* (pp. 399–419). Thousand Oaks, CA: Sage.

Searle, S. R. (1982). *Matrix algebra useful for statistics*. New York: Wiley.
Steiger, J. H. (2002). When constraints interact: A caution about reference variables, identification constraints, and scale dependencies in structural equation modeling. *Psychological Methods*, *7*, 210–227.
Takane, Y. (1987). Analysis of covariance structures and probabilistic binary choice data. *Communication and Cognition*, *20*, 45–62.
Tsai, R. C. (2000). Remarks on the identifiability of Thurstonian ranking models: Case V, Case III, or neither? *Psychometrika*, *65*, 233–240.
Tsai, R. C. (2003). Remarks on the identifiability of Thurstonian paired comparison models under multiple judgment. *Psychometrika*, *68*, 361–372.
Tsai, R. C., & Böckenholt, U. (2001). Maximum likelihood estimation of factor and ideal point models for paired comparison data. *Journal of Mathematical Psychology*, *45*, 795–811.
Tsai, R. C., & Böckenholt, U. (2002). Two-level linear paired comparison models: Estimation and identifiability issues. *Mathematical Social Sciences*, *43*, 429–449.
Tsai, R. C., & Wu, T. L. (2004). Analysis of paired comparison data using Mx. *Structural Equation Modeling: An Interdisciplinary Journal*, *11*, 73–91.
Thurstone, L. L. (1927). A law of comparative judgment. *Psychological Review*, *79*, 281–299.

APPENDIX 1

Identification of the Unrestricted Thurstonian Model

To identify $\boldsymbol{\Sigma}_t$ Maydeu-Olivares and Böckenholt (2005) used the following n constraints $\sigma_{n1} = 0, \ldots, \sigma_{n,n-1} = 0$, $\sigma_{nn} = 1$. In contrast, here we use the n constraints $\sigma_{ii} = 1, \forall i$ so that a correlation matrix $\mathbf{P}_t$ is estimated. Letting $\rho_{ii'}$ be an element of $\mathbf{P}_t$ and $\sigma_{ii'}$ be a parameter in Maydeu-Olivares and Böckenholt's $\boldsymbol{\Sigma}_t$, the relationship between both parameterizations is

$$\rho_{ii'} = \begin{cases} \dfrac{1 - \sigma_{i'i'}}{2} & \text{if } i = n \\[2ex] \dfrac{2 + 2\sigma_{ii'} - \sigma_{ii} - \sigma_{i'i'}}{2} & \text{otherwise} \end{cases}. \tag{14}$$

The inverse relationship is

$$\sigma_{ii'} = \begin{cases} 2 - \rho_{ni'} & \text{if } i = i' \\[1ex] \rho_{ii'} - \rho_{ni} - \rho_{ni'} & \text{otherwise} \end{cases}. \tag{15}$$

To show that $\mathbf{P}_t$ is indeed a correlation matrix, notice that we can decompose it using a Cholesky decomposition as $\mathbf{P}_t = \mathbf{CC}'$, where $\mathbf{C}$ is a lower triangular matrix whose diagonal elements are constrained such that $\mathbf{P}_t$ has ones along its diagonal. Dudgeon, Bell and Pattison (2003) have recently provided details on how to construct such matrices $\mathbf{C}$. For instance, for $n = 4$,

$$\mathbf{C} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ c_{21} & \sqrt{1 - c_{21}^2} & 0 & 0 \\ c_{31} & c_{32} & \sqrt{1 - c_{31}^2 - c_{32}^2} & 0 \\ c_{41} & c_{42} & c_{43} & \sqrt{1 - c_{41}^2 - c_{42}^2 - c_{43}^2} \end{pmatrix}. \quad (16)$$

This decomposition ensures that $\mathbf{P}_t$ is at least non-negative definite and therefore that its elements at least admissible (Dudgeon et al., 2003; Searle, 1982: p. 208).

To identify the model, Maydeu-Olivares and Böckenholt's introduce the additional constraint $\sigma_{11} = 1$. Similarly, in our present parameterization, one additional constraint among the elements of $\mathbf{P}_t$ needs to be introduced. Using Equation (14) Maydeu-Olivares and Böckenholt's constraint $\sigma_{11} = 1$ is equivalent to $\rho_{n1} = 0$ when $\mathbf{P}_t$ is estimated. Here, rather than using the constraint $\rho_{n1} = 0$ we use instead the constraint $\sum_{i=2}^{n} \rho_{i1} = 1$ (i.e., $\sum_{i=2}^{n} c_{i1} = 1$) as with this constraint we can obtain parameter estimates and standard errors for all the elements in $\mathbf{P}_t$.

We could identify the model by fixing one element of the diagonal matrix $\boldsymbol{\Omega}^2$ instead of introducing one constraint among the elements of $\mathbf{P}_t$. Fixing one element of $\boldsymbol{\Omega}^2$ is risky, though. To illustrate, suppose for ease of exposition and without loss of generality that $\boldsymbol{\Omega}^2 = \omega^2\mathbf{I}$. For some choices of fixed ω^2 and some 'true' $\boldsymbol{\Sigma}_t$ and ω^2 this alternative set of identification restrictions would break down. That is, if $\mathbf{P}_t$ is estimated without using $\mathbf{P}_t = \mathbf{CC}'$, $\hat{\mathbf{P}}_t$ may become negative definite. Alternatively, if $\mathbf{P}_t$ is estimated using $\mathbf{P}_t = \mathbf{CC}'$ we may not reproduce the true model exactly. To see this, suppose the true model is $\boldsymbol{\Sigma}_t = \begin{pmatrix} 1 & .6 & .4 \\ .6 & 1 & .3 \\ .4 & .3 & 1 \end{pmatrix}$ (a positive definite matrix), and $\omega^2 = .1$. If we were to estimate this model setting $\sigma_{ii} = 1, \forall i$, and $\omega^2 = 1$, we would most likely obtain a negative definite $\hat{\boldsymbol{\Sigma}}_t$ as a model with $\boldsymbol{\Sigma}_t = \begin{pmatrix} 1 & -3 & -5 \\ -3 & 1 & -6 \\ -5 & -6 & 1 \end{pmatrix}$ (a negative definite matrix), and $\omega^2 = 1$ is equivalent to the generating model—i.e., it has the same matrix of tetrachoric correlations $\mathbf{P}_{z*}$. Tsai (2003) discusses in detail the problem of equivalent models in Thurstonian paired comparison models.

APPENDIX 2

Proof of Equation (12)

Tsai (2003) has provided a rule that can be applied to find the full set of models that are equivalent to a given estimated model. Suppose a Thurstonian model for paired comparisons has been estimated. For this specific model, we denote the covariance matrix of the continuous latent preferences as $\mathbf{\Sigma}_t$ and the error covariance matrix as $\mathbf{\Omega}^2$. Both matrices must be positive definite. Then, any other model with $\mathbf{\Sigma}_2$ and $\tilde{\mathbf{\Omega}}^2$ of the form

$$\tilde{\mathbf{\Sigma}}_t = c\mathbf{\Sigma}_t + \mathbf{d1}' + \mathbf{1d}', \quad \text{and} \quad \tilde{\mathbf{\Omega}}^2 = c\mathbf{\Omega}^2, \tag{17}$$

is equivalent to the estimated model (Tsai, 2003; Corollary 1). That is, it yields the same fit to the data. In Equation (17) c is a positive constant and $\mathbf{d}$ is an $n \times 1$ vector of constants. These constants are arbitrary as long as $\tilde{\mathbf{\Sigma}}_t$ and $\tilde{\mathbf{\Omega}}^2$ are positive definite.

Here, the estimated model is a correlation matrix, that is, $\mathbf{\Sigma}_t = \mathbf{P}_t$, and we use (17) to find the set of alternative correlation matrices $\tilde{\mathbf{P}}_t$ that are equally consistent with the data. Thus, we write

$$\tilde{\mathbf{P}}_t = c\mathbf{P}_t + \mathbf{d1}' + \mathbf{1d}', \quad \text{and} \quad \tilde{\mathbf{\Omega}}^2 = c\mathbf{\Omega}^2. \tag{18}$$

The first of these equations yields for the diagonal elements the relationship $1 = c + 2d_i \Rightarrow d_i = \frac{1-c}{2}$, $i = 1, \ldots, n$. That is, for (18) to hold $\mathbf{d} = (\frac{1-c}{2})\mathbf{1}$. Substituting into Equation (18) we obtain $\tilde{\mathbf{P}}_t = c\mathbf{P}_t + (\frac{1-c}{2})\mathbf{11}' + (\frac{1-c}{2})\mathbf{11}' = c\mathbf{P}_t + (1-c)\mathbf{11}'$.

The relationship among the mean vectors can be obtained as follows. Again from Tsai (2003, section 4), two equivalent models should also satisfy

$$\tilde{\mathbf{D}}_t\mathbf{A}\tilde{\boldsymbol{\mu}}_t = \mathbf{D}_t\mathbf{A}\boldsymbol{\mu}_t \tag{19}$$

where $\mathbf{A}$ is the design matrix and $\mathbf{D} = (\mathrm{Diag}(\mathbf{AP}_t\mathbf{A}' + \mathbf{\Omega}^2))^{-\frac{1}{2}}$, $\tilde{\mathbf{D}} = (\mathrm{Diag}(\mathbf{A}\tilde{\mathbf{P}}_t\mathbf{A}' + \tilde{\mathbf{\Omega}}^2))^{-\frac{1}{2}}$. From $\tilde{\mathbf{P}}_t = c\mathbf{P}_t + (1-c)\mathbf{11}'$ it is then easy to verify that

$$\tilde{\mathbf{D}} = (\mathrm{Diag}(\mathbf{A}\tilde{\mathbf{P}}_t\mathbf{A}' + \tilde{\mathbf{\Omega}}^2))^{-\frac{1}{2}} = (\mathrm{Diag}(c(\mathbf{AP}_t\mathbf{A}' + \mathbf{\Omega}^2)))^{-\frac{1}{2}} = \frac{1}{\sqrt{c}}\mathbf{D}$$

and so

$$\tilde{\mathbf{D}}\mathbf{A}\tilde{\boldsymbol{\mu}}_t = \mathbf{D}\mathbf{A}\boldsymbol{\mu}_t \Rightarrow \mathbf{D}_t\mathbf{A}\tilde{\boldsymbol{\mu}}_t = \mathbf{D}_t\mathbf{A}(\sqrt{c}\boldsymbol{\mu}_t) \Rightarrow \mathbf{A}\tilde{\boldsymbol{\mu}}_t = \mathbf{A}(\sqrt{c}\boldsymbol{\mu}_t)$$

If we fix the last component of the mean vector to 0, this implies $\tilde{\boldsymbol{\mu}}_t = \sqrt{c}\boldsymbol{\mu}_t$ as requested.

APPENDIX 3

Mplus Input File for Modeling the Preferences for Celebrities

```
TITLE:  ULS estimation for first ladies and sportwomen
        (females resp only);

DATA: FILE IS 'ladies sport.dat';
! the data contains 96 observations on 15 binary variables

VARIABLE:
    NAMES ARE
    y12 y13 y14 y15 y16
        y23 y24 y25 y26
            y34 y35 y36
                y45 y46
                    y56;
!   it is convenient to assign names that reflect the paired
    comparisons

    CATEGORICAL =    y12-y56;
!   variables are declared as categorical

ANALYSIS:
    TYPE = MEANSTRUCTURE;
!   thresholds and tetrachoric correlations are modeled

    ESTIMATOR = ULS;
!   ULS estimation, mean and variance corrected test
    statistic

    PARAMETERIZATION = THETA;
!   it enables modeling factor variances with categorical
    dependent vars

MODEL:
    f1 BY y12-y16@1;
    f2 by y23-y26@1;
    f3 by y34-y36@1;
    f4 BY y45-y46@1;
    f5 by y56-y56@1;
```

```
    f2 by y12@-1;
    f3 by y13@-1 y23@-1;
    f4 by y14@-1 y24@-1 y34@-1;
    f5 by y15@-1 y25@-1 y35@-1 y45@-1;
    f6 by y16@-1 y26@-1 y36@-1 y46@-1 y56@-1;
!   fixed factor loadings, this is the A matrix

! the factors are
! f1 = Barbara Bush
! f2 = Nancy Reagan
! f3 = Hillary Clinton
! f4 = Bonnie Blair
! f5 = Jackee Joyner
! f6 = Jennifer Capriati

    [y12$1-y56$1@0];
!   thresholds fixed to zero

    [f1-f5* f6@0];
!   factor means free --default is fixed to zero


    y12-y56*.2(1);
!   error specific variances equal
!   for this matrix to be diagonal use instead
!   y12-y56*.2;

    f1-f6@1;
!   factor variances are fixed at 1

    f2 with f3-f6;
    f3 with f4-f6;
    f4 with f6*;
    f5 with f6*;
!   all factors are intercorrelated

    f1 with f2* (p1);
    f1 with f3* (p2);
    f1 with f4* (p3);
    f1 with f5* (p4);
    f1 with f6 (p5);
!   factor intercorrelations with first stimuli
```

```
MODEL CONSTRAINT:
    p5 = 1 - p1 - p2 - p3 -p4;
!   enforces constraint that the sum of factor
!   inter-correlations with first stimuli = 1

OUTPUT: TECH1;
! use this to verify that the A matrix is properly set up

SAVEDATA: ESTIMATES=estimates.dat;
! save estimates for Monte Carlos simulation
```

APPENDIX 4

Mplus Input File for Monte Carlo Simulation

```
TITLE:  simulation for first ladies and sportwomen data,
        ULS estimation;

MONTECARLO:
    NAMES ARE y12 y13 y14 y15 y16 y23 y24 y25 y26 y34 y35
              y36 y45 y46 y56 ;
    NOBS=96;
! same sample size as in the application is used
    NREPS = 1000;
! 1000 replications are requested
    SEED = 4553;
    GENERATE = y12-y56(1);
! multivariate normal data is generated
! each variable is categorized using one threshold
    CATEGORICAL = y12-y56;
! the data is analyzed as categorical
    POPULATION = estimates.dat;
    COVERAGE = estimates.dat;
! the estimated parameter values from the application
! are used as true parameter values to generate the data

MODEL POPULATION:
! same model as in the application
    f1 BY y12-y16@1;
    f2 by y23-y26@1;
```

```
    f3 by y34-y36@1;
    f4 BY y45-y46@1;
    f5 by y56-y56@1;
    f2 by y12@-1;
    f3 by y13@-1 y23@-1;
    f4 by y14@-1 y24@-1 y34@-1;
    f5 by y15@-1 y25@-1 y35@-1 y45@-1;
    f6 by y16@-1 y26@-1 y36@-1 y46@-1 y56@-1;

    f1-f6@1;

    [y12$1-y56$1@0];

    [f1-f5* f6@0];

    y12-y56*.2(1);

    f1 with f2-f6*;
    f2 with f3-f6*;
    f3 with f4-f6*;
    f4 with f5-f6*;
    f5 with f6*;
! the inter-factor correlation matrix is left unconstrained

ANALYSIS:
    TYPE = MEANSTRUCTURE;
    ESTIMATOR = ULS;
!   ULS estimation, mean and variance corrected test statistic
    PARAMETERIZATION = THETA;

MODEL:
! same model as in the application
    f1 BY y12-y16@1;
    f2 by y23-y26@1;
    f3 by y34-y36@1;
    f4 BY y45-y46@1;
    f5 by y56-y56@1;
    f2 by y12@-1;
    f3 by y13@-1 y23@-1;
    f4 by y14@-1 y24@-1 y34@-1;
    f5 by y15@-1 y25@-1 y35@-1 y45@-1;
    f6 by y16@-1 y26@-1 y36@-1 y46@-1 y56@-1;
```

```
    f1-f6@1;

    [y12$1-y56$1@0];

    [f1-f5* f6@0];

    y12-y56*.2(1);

    f2 with f3-f6;
    f3 with f4-f6;
    f4 with f6*;
    f5 with f6*;

    f1 with f2* (p1);
    f1 with f3* (p2);
    f1 with f4* (p3);
    f1 with f5* (p4);
    f1 with f6 (p5);

MODEL CONSTRAINT:
    p5 = 1 - p1 - p2 - p3 -p4;
```