A Supplementary material

A.1 MCMC algorithm for the Thurstonian Model

In our estimation procedure, the goal is to draw samples for the latent variables x_{ij} , z_j, μ_0, σ_0, μ_i, and σ_i given the observed orderings $\boldsymbol{y}_{\mathrm{j}}$. For this model, we can estimate all latent variables through Gibbs sampling. We first sample a value for each x_{ij} conditional on all other variables according to:

$$x_{ij} \mid \mu_i,\ \sigma_i,\ \mu_0,\ \sigma_0, z_j, x_{l,j}, x_{u,j} \sim \begin{cases} \mathrm{N}_{truncated}\ (\mu_i,\ \sigma_i, x_{l,j}, x_{u,j}) & z_j = 1 \\ \mathrm{N}_{truncated}\ (\mu_0,\ \sigma_0, x_{l,j}, x_{u,j}) & z_j = 0, \end{cases} \tag{A1}$$

where the sampling distribution is the truncated normal with mean and standard deviation dependent on the latent state z_j; with $z_j = 1$, and $z_j = 0$, the sample comes from the Thurstonian model and the guessing model respectively. The lower and upper bounds for truncated normal are determined by $x_{l,j}$ and $x_{u,j}$ respectively. The values $x_{l,j}$ and $x_{u,j}$ are based on the samples $\boldsymbol{x}_j$ that are ordered right before and after the current value x_{ij} respectively. Specifically, if $\pi(i)$ denotes the rank given to item i and $\pi^{-1}(i)$ denotes the item assigned to rank i, $l = \pi^{-1}(\pi(i) - 1)$, and $u = \pi^{-1}(\pi(i) + 1)$. We also define $x_{l,j} = -\infty$ when $\pi(i) = 1$, and $x_{u,j} = \infty$, when $\pi(i) = N$. With these bounds, the observed data influences the possible locations for the samples. It is guaranteed that the ordering of samples $\boldsymbol{x}_j$ is consistent with the observed ordering $\boldsymbol{y}_j$ for individual j.

To sample μ_i, and σ_i given $\boldsymbol{x}$, we have:

$$\sigma_i^2 \mid \mu_i, s_i^2, \mathbf{z} \sim \text{Inv-}\chi^2\left(M^{(z=1)} - 1, s_i^2\right) \tag{A2}$$

$$\mu_i \mid \ \sigma_i, \bar{x}_i, \mathbf{z} \sim \mathrm{N}\left(\bar{x}_i,\ \sigma_i/\sqrt{M^{(z=1)}}\right), \tag{A3}$$

where s_i^2 and $\bar{x}_i$ are the variance and mean of all samples $\boldsymbol{x}_i$ (restricted to individuals assigned to the Thurstonian model) for item i respectively, and $M^{(z=1)} = \sum_j z_j$, the number of individuals assigned to the Thurstonian model. Similar update equations were used to update μ_0 and σ_0 based on the samples of the individuals assigned to the guessing route:

$$\sigma_0^2 \mid \mu_0, s_0^2, \mathbf{z} \sim \text{Inv-}\chi^2\left(M^{(z=0)} - 1, s_0^2\right) \tag{A4}$$

$$\mu_0 \mid \ \sigma_0, \bar{x}_0, \mathbf{z} \sim \mathrm{N}\left(\bar{x}_0,\ \sigma_0/\sqrt{M^{(z=0)}}\right). \tag{A5}$$

In order to prevent a drift in the item positions during estimation (as there is no natural zero point), we fixed the minimum of μ_i to 0 and the maximum of μ_i to 1, and scaled the other variables accordingly.

Finally, to sample the assignment of individuals to modeling routes, we use

$$p\left(z_j = k \mid \mu_i,\ \sigma_i,\ \mu_0, \sigma_0, x_{l,j}, x_{u,j}\right) \propto \begin{cases} \prod_{i=1}^{N} f\left(\ x_{ij} \mid \mu_i,\ \sigma_i\right) & k = 1 \\ \prod_{i=1}^{N} f\left(\ x_{ij} \mid \mu_0,\ \sigma_0\right) & k = 0. \end{cases} \tag{A6}$$

where $f(x \mid \mu, \sigma)$ is the normal probability density function. In our procedure, we ran 20 chains with a burn-in of 200 iterations. From each chain, we drew 20 samples with an interval of 10 iterations. In total, we collected 400 samples. To construct a single group answer, we analyzed the ordering of the items according to μ_i, separately for each sample, and then picked the mode of this distribution. This corresponds to the most likely order in the distribution over orders inferred by the model.

A.2 MCMC algorithm for Mallows Model

In our MCMC algorithm for Mallows model, we use a combination of Metropolis-Hastings (MH) and Gibbs sampling steps. To estime $\boldsymbol{\omega}$, we use the MH algorithm based on Lebanon and Lafferty (2002). The idea is to move the group estimate $\boldsymbol{\omega}$ by transposing any randomly chosen pair of items. The proposal distribution $q(\boldsymbol{\omega}^*|\boldsymbol{\omega})$ is

$$q(\boldsymbol{\omega}'|\boldsymbol{\omega}) = \begin{cases} 1/\binom{n}{2} & \text{if } S(\boldsymbol{\omega}', \boldsymbol{\omega}) = 1 \\ 0 & \text{otherwise,} \end{cases} \tag{A7}$$

where $S(\boldsymbol{\omega}', \boldsymbol{\omega})$ is the Cayley distance. The Metropolis-Hastings acceptance ratio is

$$\min\left[1, \frac{q(\boldsymbol{\omega}|\boldsymbol{\omega}')}{q(\boldsymbol{\omega}'|\boldsymbol{\omega})} \frac{p(\boldsymbol{y}|\boldsymbol{\omega}', \theta, \boldsymbol{z})}{p(\boldsymbol{y}|\boldsymbol{\omega}, \theta, \boldsymbol{z})}\right]. \tag{A8}$$

Note that the first likelihood ratio for the proposal distribution equals one because of the symmetry in the proposals. Also, in Eq 1., the normalization constant does not depend on $\boldsymbol{\omega}$, which can be used to simplify the acceptance ratio to:

$$\min\left[1, \exp\left(-\theta \sum_{z_j=1} d(\boldsymbol{y}_j, \boldsymbol{\omega}') - d(\boldsymbol{y}_j, \boldsymbol{\omega})\right)\right], \tag{A9}$$

where the sum is taken over all individuals currently assigned to Mallows model. To facilitate the inference for θ, we used a discretized set of 1000 θ values, logarithmically spaced between 10^{-4} and 2. Let v_k refer to the kth value in this set. We use a Gibbs sampling step for θ by sampling from the discrete distribution

$$p(\theta = v_k|\boldsymbol{\omega}, \boldsymbol{z}, \boldsymbol{y}) \propto \exp\left[-v_k \sum_{z_j=1} d(\boldsymbol{y}_j, \boldsymbol{\omega}) - \sum_{z_j=1} \log\Psi(v_k)\right]. \tag{A10}$$

Finally, we use a Gibbs sampling step to estimate the latent state z_j

$$p(z_j = k|\theta, \boldsymbol{\omega}, \boldsymbol{z}_{-j}, \boldsymbol{y}_j) = \begin{cases} 1/\text{N!} & \text{k} = 0 \\ \exp[-\theta d(\boldsymbol{y}_j, \boldsymbol{\omega}) - \log\Psi(\theta)] & \text{k} = 1. \end{cases} \tag{A11}$$

In the MCMC procedure, we ran 20 chains with a burn-in of 200 iterations. From each chain, we drew 20 samples with an interval of 10 iterations. In total, we collected 400 samples. To construct a single group answer, we picked the most frequently occuring sampled ordering $\boldsymbol{\omega}$.