# Supplementary Material 1: Details of the Statistical Model

**Data Model**

For individual $i = 1, \ldots, n$ and item $j = 1, \ldots, m$, let $Y_{ij}$ represent the observed response to item j. Since our items include ordinal (e.g., Likert scales) and binary responses (Yes/No responses), we use a latent variable approach to handle the discrete nature of these responses within the continuous factor model framework. The data model depends on the item type:

**Ordinal and Binary Items**: For ordinal and binary items, we assume that the observed discrete responses arise from an underlying continuous latent variable $Z_{ij}$. This latent variable represents the "true" continuous response that gets discretized into the observed categories through a threshold process:

$$Z_{ij} \sim N\left(\mu_{ij}, \tau_z^{(-1)}[j]\right).$$

The observed response $Y_{ij}$ is then determined by which interval the continuous $Z_{ij}$ falls into:

$$Y_{ij} = k \text{ if } \kappa_{j,k} < Z_{ij} \leq \kappa_{j,k+1}$$

where $\kappa_{j,k}$ are threshold parameters (cutpoints) for item j that divide the continuous scale into discrete response categories. For example, for a 5-point Likert item, there would be thresholds $\kappa_{j,1}, \kappa_{j,2}, \kappa_{j,3}, \kappa_{j,4}$ that separate the five response categories. This approach allows us to: (i) model discrete responses within a continuous factor analysis framework, (ii) account for the ordinal nature of Likert-type responses, (iii) handle the relationship between binary and ordinal items in the same model.

**Continuous Items**: For the one continuous item in our model (Medical Stigma Resilience), observed responses are modeled directly without the threshold structure: $Y_{ij} \sim N\left(\mu_{ij}, \tau_z^{(-1)}[j]\right)$

**Measurement Model**

The mean structure for the underlying continuous latent variables $Z_{ij}$ follows a factor model:

$$\mu_{ij} = \beta_0[j] + \lambda[j]\theta_{i,k},$$

where $\beta_0[j]$ is the intercept for item j, $\lambda[j]$ is the factor loading for item j, $\theta_{i,k}$ is individual i's score on factor $k$, and $k$ indicates which factor item j loads on:

- $k = 1$ for items 1-7 (Motivation factor)
- $k = 2$ for items 8-10 (Information factor)
- $k = 3$ for items 11-13 (Behavioral Skills factor)

This distinction is important because:

- For ordinal/binary items: $\mu_{ij}$ is the mean of the continuous $Z_{ij}$, which then gets discretized into the observed $Y_{ij}$ through the threshold process.
- For continuous items: $\mu_{ij}$ directly represents the mean of the observed $Y_{ij}$ since $Z_{ij} = Y_{ij}$.

**Structural Model**

**Factor Structure**: We assume that the motivation and information factors follow a bivariate normal distribution: $(\theta_{i,1}, \theta_{i,2}) \sim N\left((0,0), \Omega^{(-1)}\right)$

where $\Omega$ is the precision matrix corresponding to the covariance matrix $\Sigma$ with correlation $\rho_{12}$. This specification allows for potential correlation between information and motivation constructs, consistent with IMB model where these factors may be.

**Behavioral Skills Factor**: We assume that the third factor is predicted by the first two:

$$\theta_{i,3} \sim N\left(\gamma_1 \theta_{i,1} + \gamma_2 \theta_{i,2}, \tau_{f3}^{-1}\right)$$

**Outcome Model**: LAI-PrEP willingness follows a logistic regression:

$$logit(p_i) = \alpha_0 + \alpha_1 \theta_{i,1} + \alpha_2 \theta_{i,2} + \alpha_3 \theta_{i,3}$$

$$lai_will_i \sim Bernoulli(p_i)$$

**Prior Distributions**

We employed mildly informative priors to provide regularization while allowing the data to drive parameter estimation. Our prior choices were motivated by the need to promote stable estimation and convergence in the complex missing data context while covering reasonable parameter ranges for binary and ordinal model. The use of mildly informative rather than flat priors is appropriate in our setting because small cell sizes in categorical responses can create separation problems, where maximum likelihood estimates do not exist and parameters can diverge toward infinity to perfectly predict empty cells (Albert, 1984). Including prior information helps regularize these extreme estimates while still allowing the data to drive the results (Gelman, 2008).

**Factor Loadings**: For identification, first loadings are fixed at 1. Remaining loadings:

$$\lambda[j] \sim N\left(1, \sigma_\lambda^2\right),$$

where $\sigma_\lambda = 3$ for ordinal items, $\sigma_\lambda = 1$ for binary items.

**Intercepts**:

$$\beta_0[j] \sim N(0, \sigma_\beta^2),$$

where $\sigma_\beta = 3$ for ordinal items, $\sigma_\beta = 1$ for binary items.

**Structural Coefficients**:

$$\alpha_k \sim N(0,4), \quad k = 0,1,2,3,$$

$$\gamma_k \sim N(0,4), \quad k = 1,2$$

.

**Variance Parameters**:

$$\tau_z[j] \sim Gamma(1,0.5)$$

$$\tau_{f3} \sim Gamma(2,0.5)$$

**Factor Covariance**:

$$\sigma_{1,1}, \sigma_{2,2} \sim Gamma(2,1)$$

$$\rho_{12} \sim Uniform(-1,1)$$

**Missing Data**

Missing values are handled within the Bayesian framework by treating them as additional parameters to be estimated. For missing $Y_{ij}$, the corresponding $Z_{ij}$ is sampled from its conditional distribution given the factor model, and the missing $Y_{ij}$ is imputed based on the threshold structure.