PSYCHOMETRIKA
2014
DOI: 10.1007/S11336-013-9396-3

THE LOGNORMAL RACE: A COGNITIVE-PROCESS MODEL OF CHOICE AND LATENCY WITH DESIRABLE PSYCHOMETRIC PROPERTIES

Jeffrey N. Rouder AND Jordan M. Province
UNIVERSITY OF MISSOURI
Richard D. Morey
UNIVERSITY OF GRONINGEN
Pablo Gomez
DEPAUL UNIVERSITY
Andrew Heathcote
UNIVERSITY OF NEWCASTLE

Reporter: Yingshi Huang

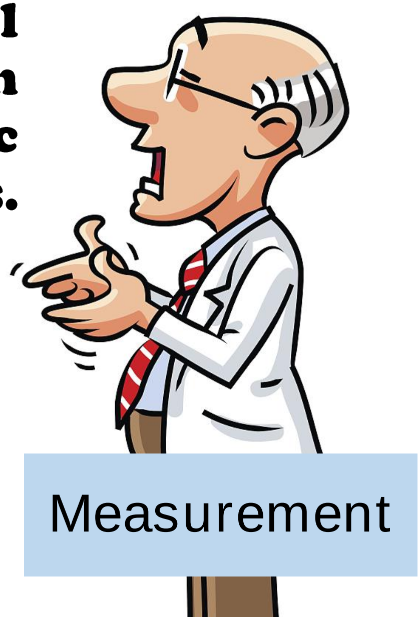
Introduction

- human performance data
  - measure individual abilities
  - measure the effects of covariates
  - assess latent structure

Cognitive

A best model should address the question of how these item responses are generated.

A best model is built on desirable statistical properties with measurement-theoretic assumptions.



Measurement

Introduction

- human performance data
  - measure individual abilities
  - measure the effects of covariates
  - assess latent structure

1. cognitive-process models
   - ✓ incorporate theoretical insights
   - ✓ to draw inferences about the processes and mental representations

finer process, but complex

2. psychometric models
   - ✓ retain desirable statistical properties
   - ✓ to incorporate covariates and understand variation in latent traits and abilities

statistically tractable, but coarse

➡ Purpose: develop a cognitive psychometrics model

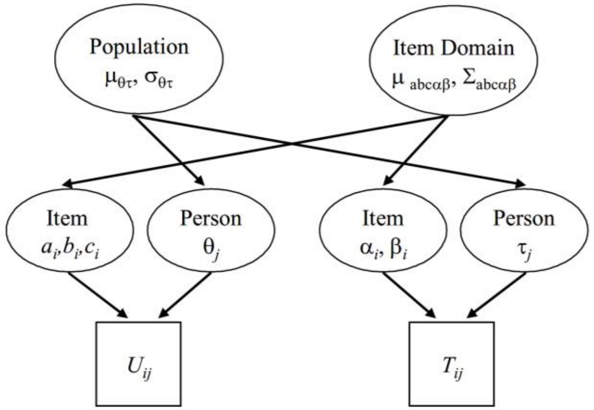
Modeling Response Time and Response Choice Jointly

- psychometric approach:

  1. place separate regression models on choice and RT
  2. allow RT / choice to be a covariate on the other one or share latent parameters

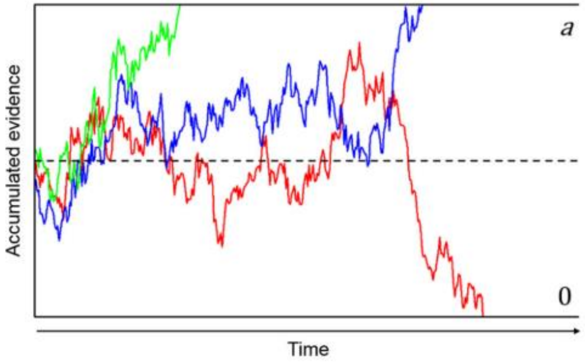
$$p_i(\theta_j) = [1 + \exp(\theta_j - \ln \tau_j - b_i)]^{-\pi_i}$$

$$\ln T_{ij} = \mu + \tau_j + \beta_i - \rho(a_i\theta_j - b_i) + \varepsilon_{ij}$$



- cognitive-process approach:
  - the diffusion model

$$f_{X,T}(x,t) = \frac{\pi\sigma^2}{\alpha^2}\exp\left(\frac{(\alpha x - z)(\theta + \beta)}{\sigma^2} - \frac{(\theta+\beta)^2}{2\sigma^2}(t - T_{er})\right)$$
$$\times \sum_{m=1}^{\infty} m \sin\left(\frac{\pi m(\alpha x - 2zx + z)}{\alpha}\right)\exp\left(-\frac{1}{2}\frac{\pi^2\sigma^2 m^2}{\alpha^2}(t - T_{er})\right)$$



Modeling Response Time and Response Choice Jointly

- adapting the diffusion model for psychometric settings
    - Tuerlinckx and De Boek (2005)

$$P_+ = P(X = 1) = \frac{e^{-2zv} - 1}{e^{-2av} - 1} = \frac{e^{av}}{1 + e^{av}}$$

    - van der Mass et al. (2011)

$$P_+ = \frac{e^{av}}{1 + e^{av}} = \frac{e^{\frac{a_k^p v_k^p}{a_j^i v_j^i}}}{1 + e^{\frac{a_k^p v_k^p}{a_j^i v_j^i}}}$$

    - Boehm et al. (2021)

$$\mathbb{P}(X = 1) = \frac{1}{1 + \exp\left(-\frac{v_{ip}}{cs_i}\right)}$$

➡

☐ test with an arbitrary number of choices

☐ the complexity of a multidimensional diffusion process

☐ the likelihood is not easily analyzed

☐ not clear whether more complex models may be placed on parameters

Specification of the Base Lognormal Race Model

- a simpler and more tractable cognitive process:
  - a race between accumulators

T = D / V

D $\sim B + U(0, A)$

$\sim e^{N(\mu_d, \sigma_d^2)}$

V $\sim N(v, sv)$

$\sim e^{N(\mu_v, \sigma_v^2)}$

  - the fastest accumulator indicates the finishing time

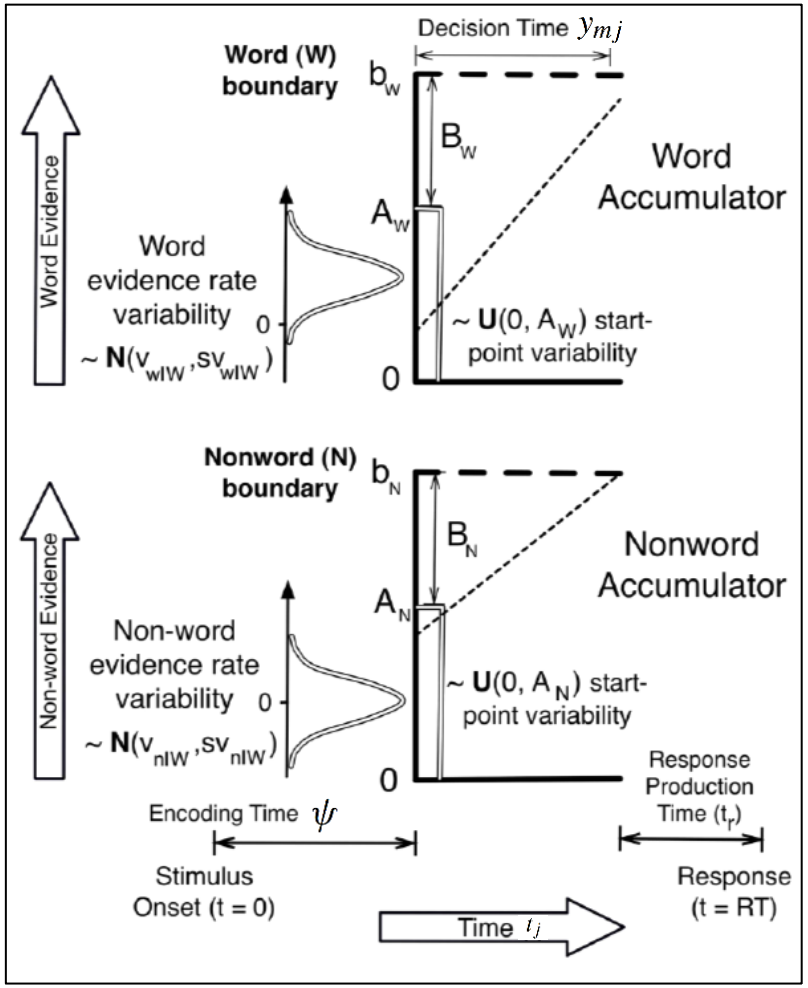
$$x_j = m \quad \Longleftrightarrow \quad y_{mj} = \min_i \boxed{(y_{ij})} \sim e^{N(\mu_d - \mu_v, \sigma_d^2 + \sigma_v^2)} = e^{N(\mu, \sigma^2)}$$

$$t_j = \psi + \min_i (y_{ij}) \qquad \overset{\text{ind}}{\sim} \text{Lognormal}\left(\mu_i, \sigma_i^2\right)$$

the joint density function:

$$f(m, t) = g\left(t - \psi; \mu_m, \sigma_m^2\right) \prod_{i \neq m} \left(1 - G\left(t - \psi; \mu_i, \sigma_i^2\right)\right)$$

Is “ *JUDGE* ” a real word?



Heathcote & Love, 2012 *Front Psychol*

Bayesian Analysis of the Base Model

- MCMC of the joint posterior distribution with Gibbs sampler

  1. prior specification

     $$\pi(\psi) \propto 1$$

     $$\sigma_i^2 \overset{\text{ind}}{\sim} \text{Inverse-Gamma}(a_i, b_i)$$

     $$\mu_i \overset{\text{ind}}{\sim} \text{Normal}(c_i, d_i)$$

  2. sampling from the conditional posterior distribution

     - conveniently expressed when conditioned on latent finishing times y_{ij}
     - let $z_{ij} = \log y_{ij}$ and then $\bar{z}_i = J^{-1} \sum_j z_{ij}$

➡ $$\mu_i \mid \cdots \sim \text{Normal}\big(v_i\big[J\bar{z}_i/\sigma_i^2 + c_i/d_i\big], v_i\big) \qquad v_i = \big(J/\sigma_i^2 + 1/d_i\big)^{-1}$$

$$\sigma_i^2 \mid \cdots \sim \text{Inverse-Gamma}\left(a_i + J/2, b_i + \sum_j (z_{ij} - \mu_i)^2/2\right)$$

$$f(\psi \mid \cdots) \propto \begin{cases} \prod_{ij} g(y_{ij} - \psi; \mu_i, \sigma_i^2), & \psi < \min_{ij}(y_{ij}) \\ 0, & \text{otherwise} \end{cases}$$

does not correspond to a known distribution
(Metropolis step with a symmetric normal random walk)

Bayesian Analysis of the Base Model

  - assume the i th accumulator finished first

$$(z_{ij} \mid x_j = i, \ldots) = \log(t_j - \psi)$$

$$(z_{ij} \mid x_j \neq i, \ldots) \sim \text{Truncated Normal}\big(\mu_i, \sigma_i^2, \log(t_j - \psi)\big)$$

- model comparison
  - the deviance information criterion [DIC]
    1. after each iteration in Gibbs sampling:

$$D^{(g)} = -2\ln\left(f(m,t) = g\big(t - \psi^{(g)}; \mu_m^{(g)}, \sigma_m^{2\,(g)}\big) \prod_{i \neq m} \big(1 - G\big(t - \psi^{(g)}; \mu_i^{(g)}, \sigma_i^{2\,(g)}\big)\big)\right)$$

$$\bar{D} = \frac{\sum_g D^{(g)}}{G}$$

    2. for the posterior mean:

$$D(\bar{\theta}) = -2\ln\left(f(m,t) = g\big(t - \psi; \mu_m, \sigma_m^2\big) \prod_{i \neq m} \big(1 - G\big(t - \psi; \mu_i, \sigma_i^2\big)\big)\right)$$

➡

$$p_D = \bar{D} - D(\bar{\theta})$$

$$\text{DIC} = \bar{D} + p_D$$

Application I: for Testing

- a Rasch-Like IRT Model
  - place backend model on the log finishing times z_{ij}

$$x_{jk} = m \quad \Longleftrightarrow \quad z_{mjk} = \min_i(z_{ijk})$$

$$t_{jk} = \psi_k + \exp\Big(\min_i(z_{ijk})\Big)$$

  - $z_{ijk} = \underbrace{\boxed{\alpha_{ij}}}_{\text{difficulty}} - \delta_{ij}\overbrace{\boxed{\beta_k}}^{\text{ability}} + \epsilon_{ijk}$ if i is the correct choice

$$z_{ijk} = \alpha_{ij} - \beta_k + \epsilon_{ijk}$$

otherwise

$$z_{ijk} = \alpha_{ij} + \epsilon_{ijk}$$

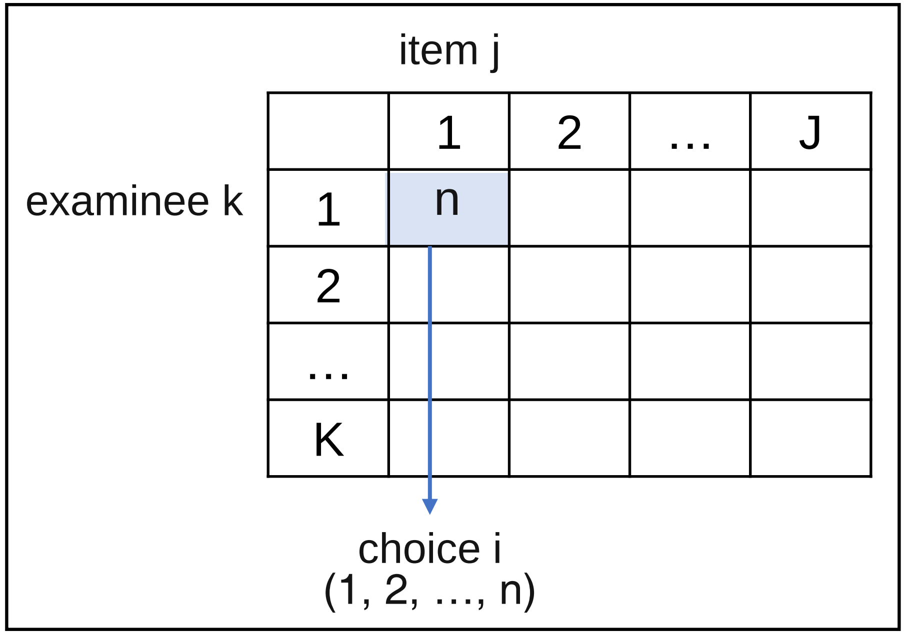
  - for the kth examinee:

$$z_k = (z_{11k}, \ldots, z_{n1k}, z_{12k}, \ldots, z_{nJk})'$$

across all examinees: $\boldsymbol{z} = (z_1', \ldots, z_K')'$

$$\boldsymbol{\alpha} = (\alpha_{11}, \ldots, \alpha_{n1}, \alpha_{1J}, \ldots, \alpha_{nJ})' \text{ and } \boldsymbol{\beta} = (\beta_1, \ldots, \beta_K)'$$

➡ $\boldsymbol{z} = \boldsymbol{X}_\alpha\boldsymbol{\alpha} - \boldsymbol{X}_\beta\boldsymbol{\beta} + \boldsymbol{\epsilon}$



item j

examinee k

	1	2	…	J
1	n			
2				
…				
K				

choice i
(1, 2, …, n)

Model development

- priors and posteriors

$$(z_{ij} \mid x_j = i, \ldots) = \log(t_j - \psi)$$

$$z_{ijk} = \boxed{\alpha_{ij}} - \delta_{ij}\boxed{\beta_k} + \boxed{\epsilon_{ijk}}$$

iid normally distributed zero-centered with variance σ^2

be modeled as random effect (with a hierarchical prior):

$$\boldsymbol{\beta} \mid \sigma_\beta^2 \sim \text{Normal}\left(\mathbf{0}, \sigma_\beta^2 \boldsymbol{I}\right)$$

$$\sigma_\beta^2 \sim \text{Inverse-Gamma}(a, b)$$

  - treat **α** and **β** as a single vector:

$$\boldsymbol{\theta} = (\boldsymbol{\alpha}', \boldsymbol{\beta}')'$$

$$\boldsymbol{z} = \boldsymbol{X\theta} + \boldsymbol{\epsilon} \quad \boldsymbol{X} = (\boldsymbol{X}_\alpha, \boldsymbol{X}_\beta)$$

➡ $\boldsymbol{\theta} \mid \cdots \sim \text{Normal}(\boldsymbol{\phi q}, \boldsymbol{\phi}) \quad \boldsymbol{\phi} = \left(\boldsymbol{X}'\boldsymbol{X}/\sigma^2 + \boldsymbol{B}\right)^{-1}$

$$\boldsymbol{q} = \boldsymbol{X}'\boldsymbol{z}/\sigma^2 + \boldsymbol{B\mu}_0$$

✓ prior precision: $\boldsymbol{B} = \text{diag}(\mathbf{1}_{nJ}/d, \mathbf{1}_K/\sigma_\beta^2)$

✓ prior mean: $\boldsymbol{\mu}_0 = (c\mathbf{1}_{nJ}, \mathbf{0}_K)'$

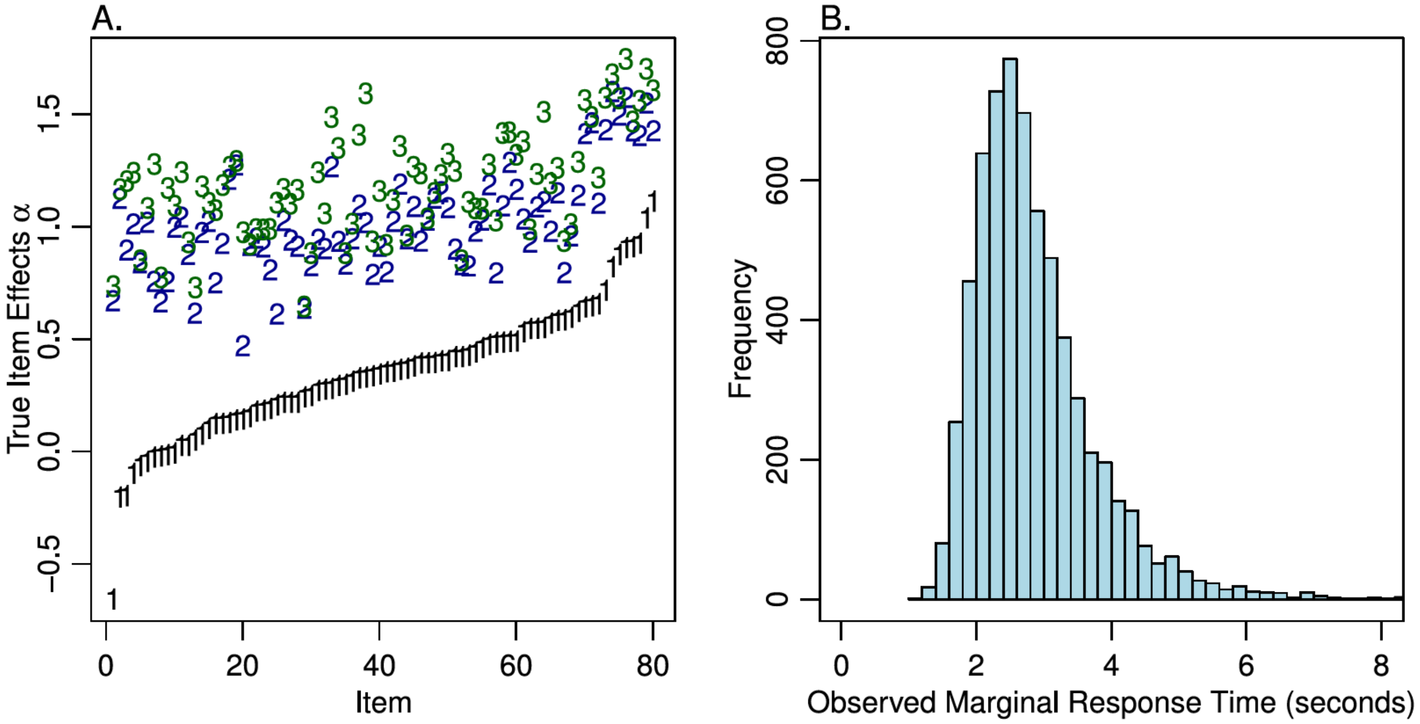
$$\sigma^2 \mid \cdots \sim \text{Inverse-Gamma}\left(a_1 + nJK/2, b_1 + (\boldsymbol{z} - \boldsymbol{X\theta})'(\boldsymbol{z} - \boldsymbol{X\theta})/2\right)$$

$$\sigma_\beta^2 \mid \cdots \sim \text{Inverse-Gamma}\left(a + K/2, b + \boldsymbol{\beta}'\boldsymbol{\beta}/2\right)$$

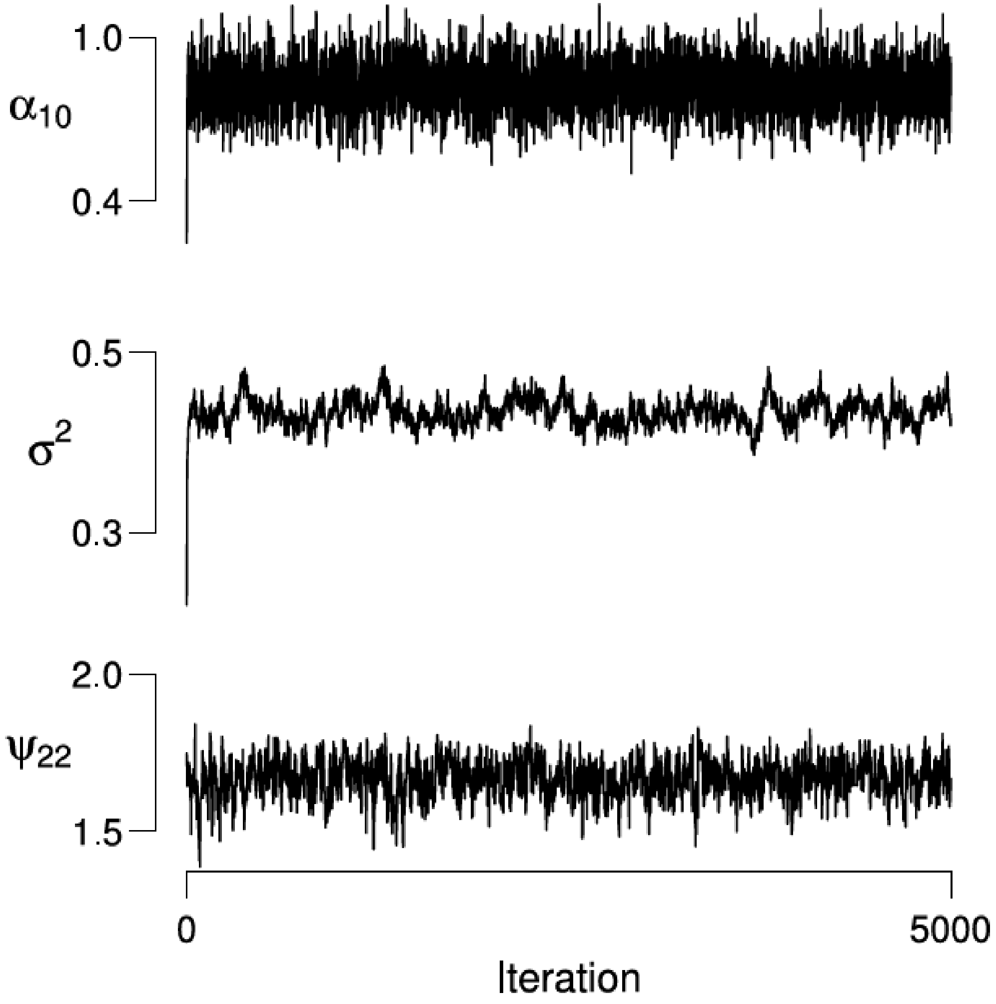
  - set c = 0 and d = 5
  - weakly informative: a = a1 = 1 and b = b1 = 0.1

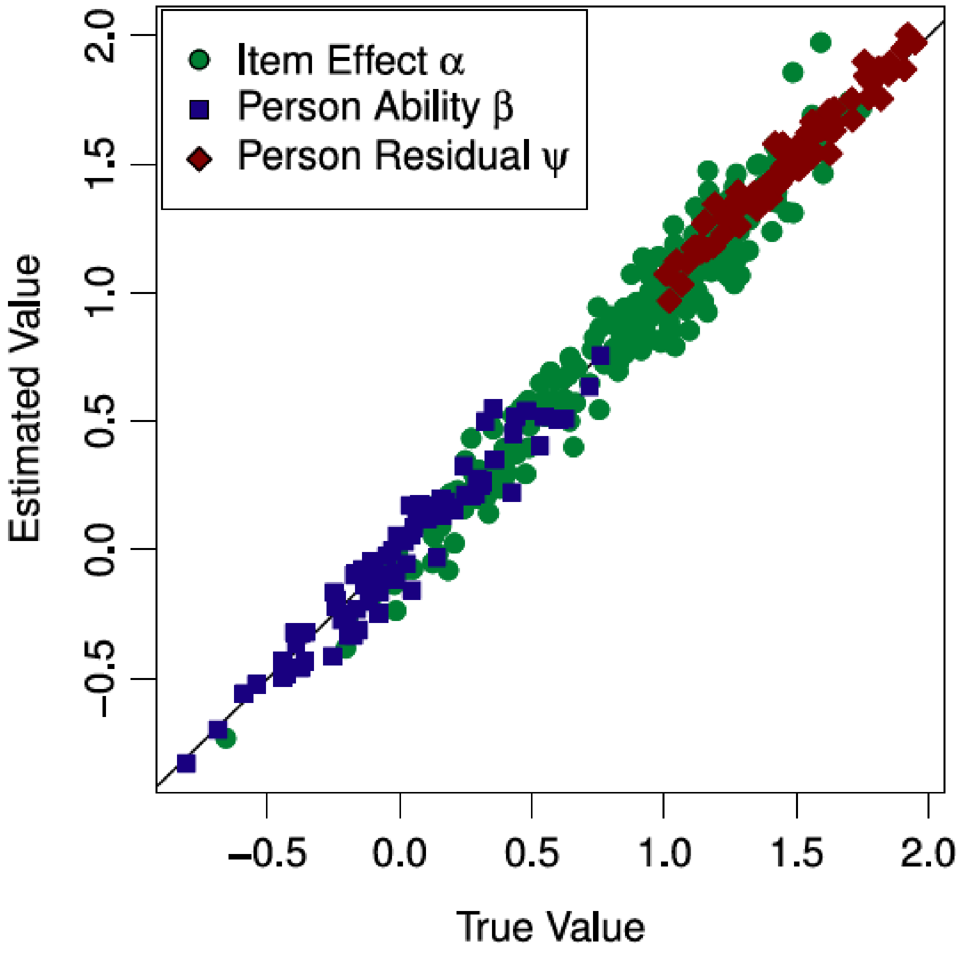
Whether the model can be analyzed efficiently

- simulated data set
  - 80 people
  - 80 items (three response options)
  - true shift values:
    $\psi_k \overset{\text{iid}}{\sim} \text{Unif}(1,2)$
  - MCMC:
    5 runs of 5,000 iterations
    a burn-in period of 500



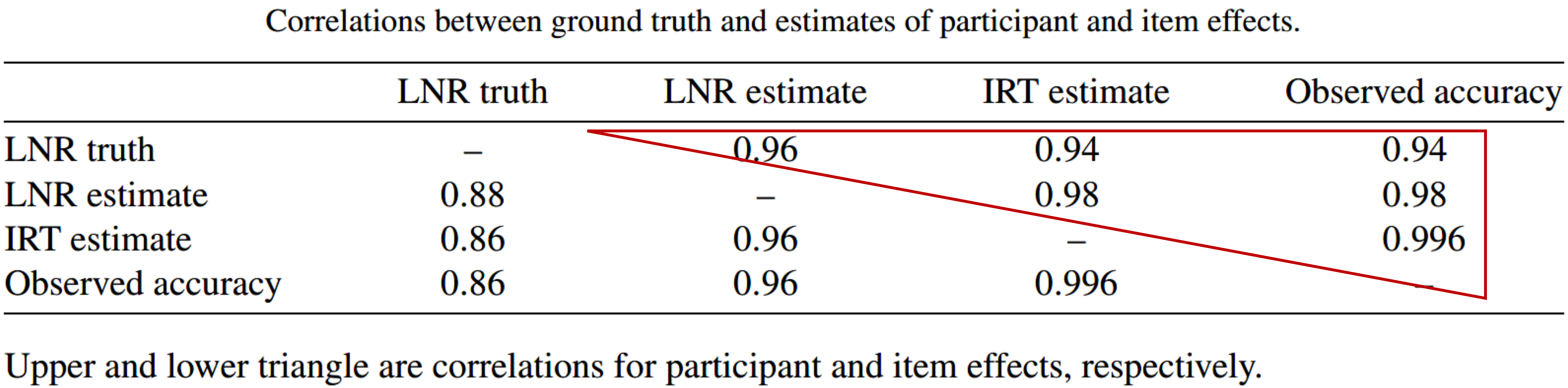
Results


1.0
α_{10}
0.4
0.5
σ^2
0.3
2.0
ψ_{22}
1.5
0
5000
Iteration


● Item Effect α
■ Person Ability β
◆ Person Residual ψ
Estimated Value
–0.5
0.0
0.5
1.0
1.5
2.0
True Value
–0.5
0.0
0.5
1.0
1.5
2.0

Results

- compared with the Rasch model



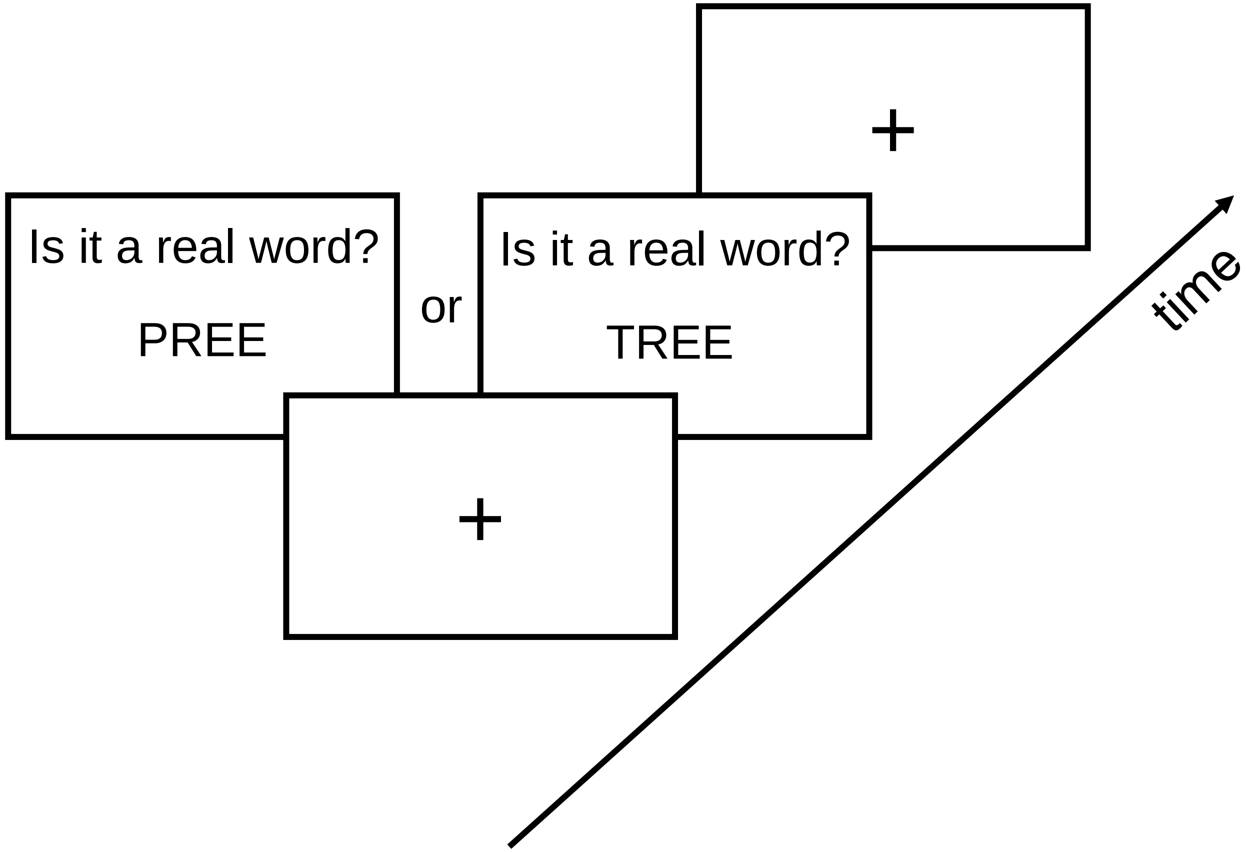
Correlations between ground truth and estimates of participant and item effects.

	LNR truth	LNR estimate	IRT estimate	Observed accuracy
LNR truth	–	0.96	0.94	0.94
LNR estimate	0.88	–	0.98	0.98
IRT estimate	0.86	0.96	–	0.996
Observed accuracy	0.86	0.96	0.996	–

Upper and lower triangle are correlations for participant and item effects, respectively.

Application II: for experimental psychology

- lexical decision task



Is it a real word?

PREE

or

Is it a real word?

TREE

+

+

time

- the data set:
  93 participants each performing 720 trials
- nonword:
  substitution conditions (5)
  e.g., TREE → PREE

  transposition conditions (7)

  e.g., JUDGE → JUGDE
- word:
  frequency of occurrence (3)
  e.g., CITY & AJAR

Model Specification

- autoregressive process: to model the trial order
  - $x_{jk} = 1$ (word) or 2 (nonword)

    z_{1k} and z_{2k} : latent log finishing times for the two accumulators (720 × 1)

    $\boldsymbol{\mu}_{1k}$ and $\boldsymbol{\mu}_{2k}$: participant-specific condition means (15 × 1)

    $\boldsymbol{X}_k$: maps trials into the 15 experimental conditions (720 × 15)

    ➡ residuals: $\boldsymbol{u}_{ik} = \boldsymbol{z}_{ik} - \boldsymbol{X}_k \boldsymbol{\mu}_{ik}$

  - first order autoregressive model

$$u_{ijk} = \rho u_{i,j-1,k} + \epsilon_{ijk}$$

$$\epsilon_{ijk} \overset{\text{iid}}{\sim} \text{Normal}(0, \sigma^2)$$

$$|\rho| < 1$$

$$u_{i,0,k} = 0$$

➡ noise terms:

$$\boldsymbol{\epsilon}_{ik} = \boldsymbol{A}_k \boldsymbol{u}_{ik}$$

(720 × 720)

$$\boldsymbol{A}_k = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ -\rho & 1 & \ddots & \ddots & & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & & \ddots & \ddots & 1 & 0 \\ 0 & 0 & \cdots & 0 & -\rho & 1 \end{bmatrix}$$

Model Specification

- the joint expression

$$\boldsymbol{\epsilon}_{ik} = \boldsymbol{A}_k \boldsymbol{u}_{ik}$$

$$\boldsymbol{A}_k = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ -\rho & 1 & \ddots & \ddots & & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & & \ddots & \ddots & 1 & 0 \\ 0 & 0 & \cdots & 0 & -\rho & 1 \end{bmatrix} \xrightarrow{\text{the inverse matrix}} \boldsymbol{L}_k = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ \rho & 1 & \ddots & \ddots & & 0 \\ \rho^2 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ \rho^{J_k-2} & & \ddots & \ddots & 1 & 0 \\ \rho^{J_k-1} & \rho^{J_k-2} & \cdots & \rho^2 & \rho & 1 \end{bmatrix}$$

$$\boldsymbol{u}_{ik} = \boldsymbol{L}_k \boldsymbol{\epsilon}_{ik}$$

the joint distribution:

$$\boldsymbol{u}_{ik} \sim \text{Normal}(\mathbf{0}, \boldsymbol{\Sigma}_{\mathbf{k}})$$

$$\boldsymbol{\Sigma}_k = \sigma^2 \boldsymbol{L}_k \boldsymbol{L}_k{}'$$

the model on log finishing times:

$$z_{1k} \sim \text{Normal}(\boldsymbol{X}_k \boldsymbol{\mu}_k, \boxed{\boldsymbol{\Sigma}_k})$$

$$z_{2k} \sim \text{Normal}(\boldsymbol{X}_k \boldsymbol{\mu}_k, \boxed{\boldsymbol{\Sigma}_k})$$

nonzero, off-diagonal elements

Analysis

1. sampling the conditional posterior values of log finishing times
   - the winning accumulator:

     $$z_{\bar{i}jk} = \log(t_{jk} - \psi_k)$$

   - the losing accumulator:

     $$z_{ijk} | \cdots \sim \text{Truncated-Normal}\left(\mathrm{E}[z_{ijk}] + \rho u_{i,j-1,k}, \sigma^2, z_{\bar{i}jk}\right)$$

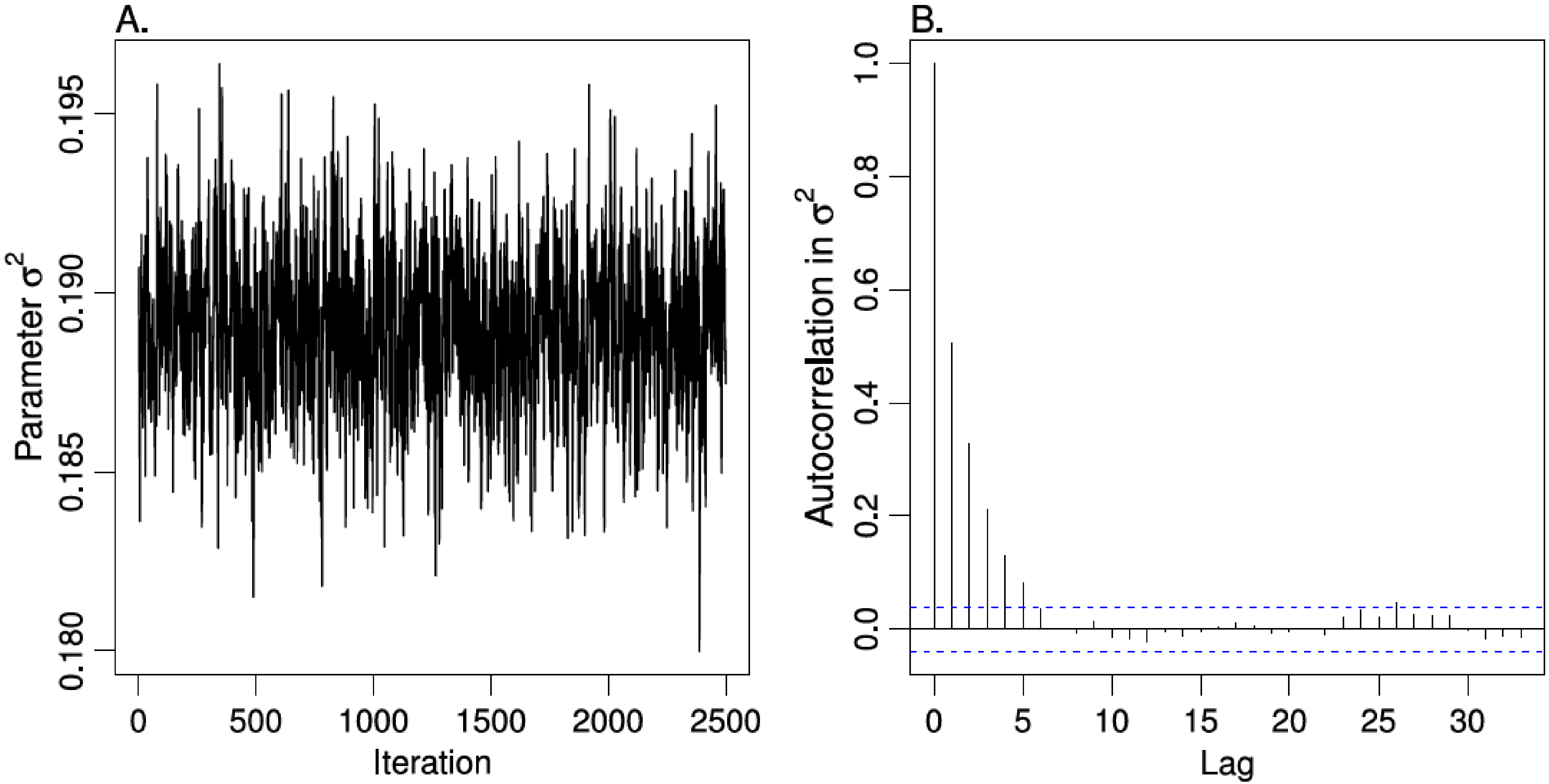
2. sampling the vectors of parameters and hyperparameters given the log finishing times
   - the autoregressive parameter:

     $$\rho \mid \cdots \sim \text{Truncated-Normal}_{(-1,1)}\left(m'/v', \sigma^2/v'\right)$$

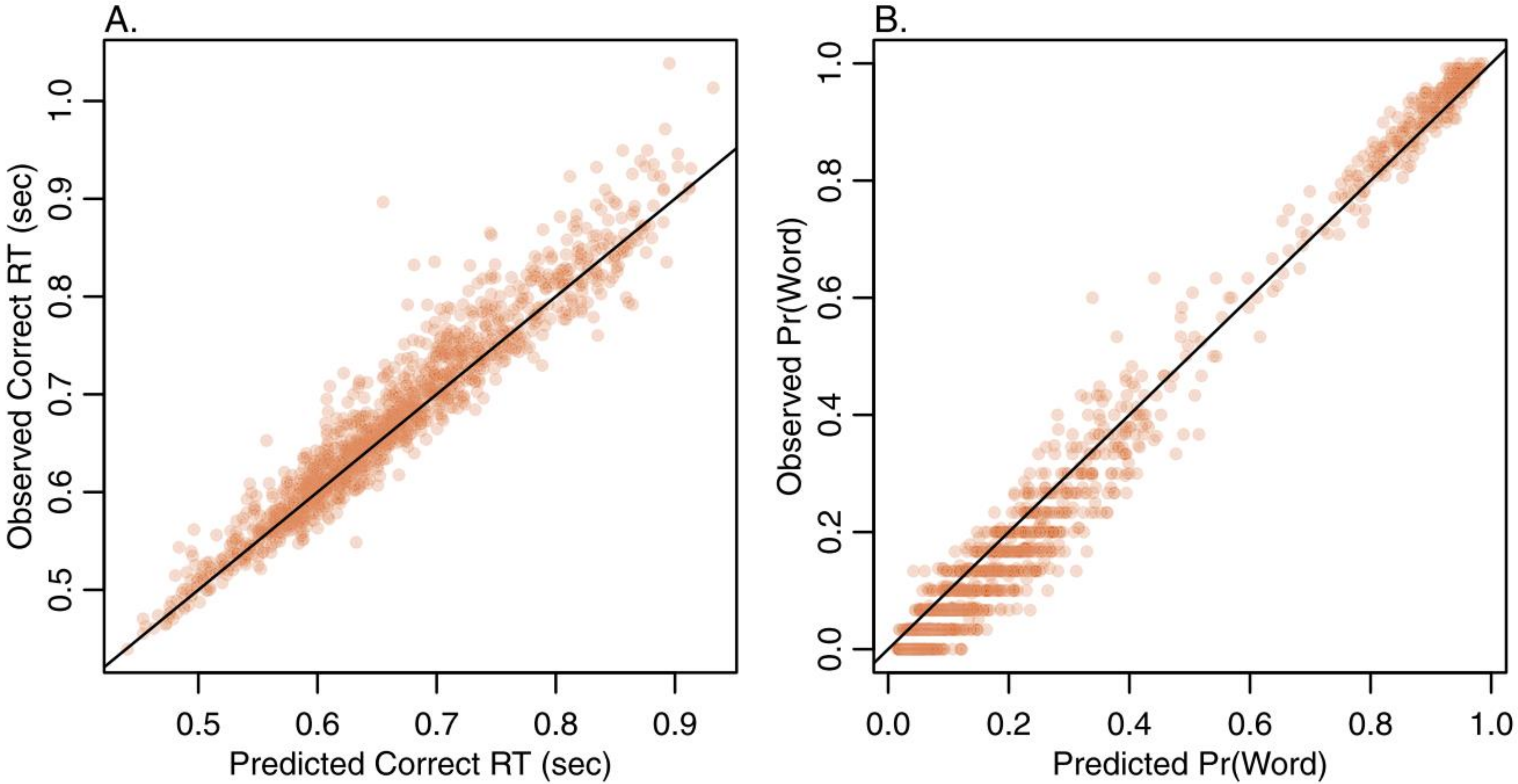
$$m' = \sum_{i=1}^{2} \sum_{k=1}^{K} \sum_{j=1}^{J_k - 1} u_{ijk} u_{i,j+1,k}$$

$$v' = \sum_{i=1}^{2} \sum_{k=1}^{K} \sum_{j=1}^{J_k - 1} u_{ijk}^2$$

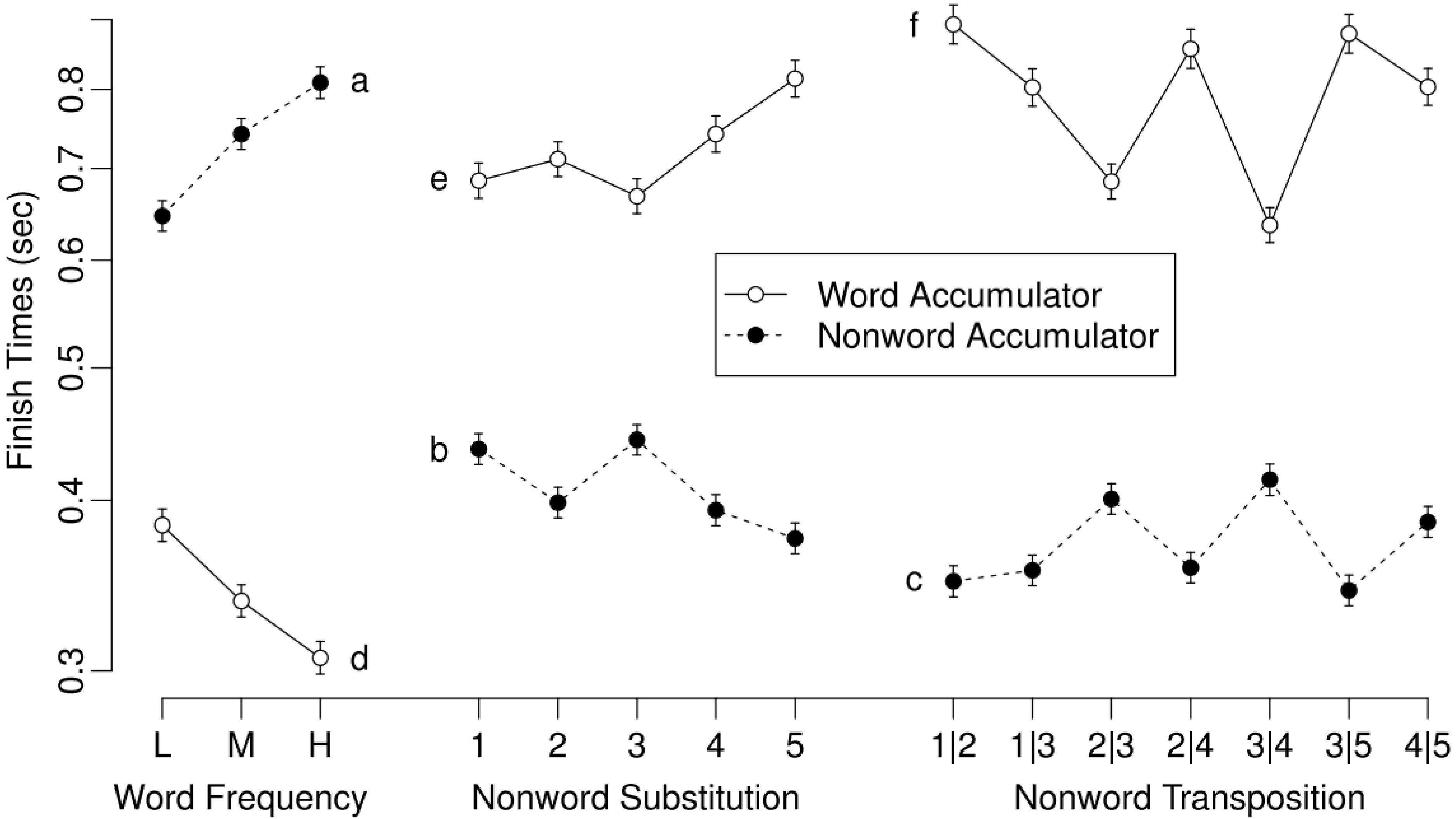
Results: mixing



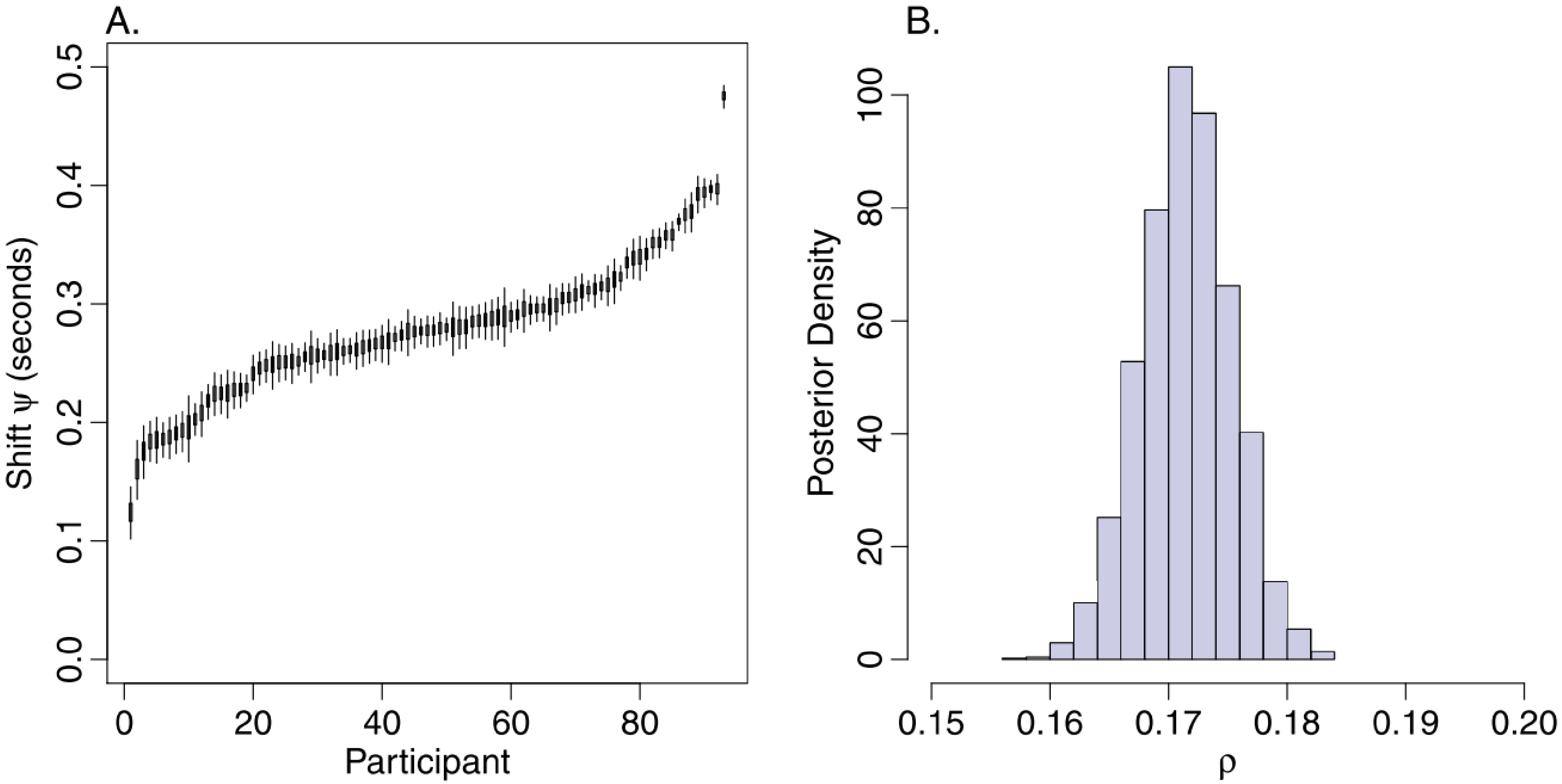
Results: model fit



Results: parameter estimates



Results: parameter estimates



DIC (without shift) = -68,879 > DIC (with shift) = -77,374

Discussion

- decision making process
  - arising from a race between competing evidence-accumulation

- straightforward to place sophisticated model components
  - IRT model & autoregressive model

Limitations

- bounds and accumulation rates cannot be disentangled
  - set bounds to constant 1: $y \sim \text{Lognormal}(\mu, \sigma^2)$

$$\mu = -\mu_\nu \text{ and } \sigma^2 = \sigma_\nu^2$$

  - specific parametric assumptions are needed to identify decision bounds
- highly accurate responses
  - the incorrect accumulators will largely reflect prior assumptions
- the additional development for a shift parameter
  - the inclusion of a Metropolis step & the impact on mixing under certain circumstances

THANKS FOR YOUR ATTENTION!

REPORTER

YINGSHI HUANG