Supplementary Materials

Appendix A: Theory of Mind Items

Here is the complete list of the ToM items included in the assessment. The items are a subset of a larger scale developed by Dodell-Feder, Koster-Hale, Bedny, and Saxe (2011).

- False belief: Item 1

  When Lisa left Jacob he was deep asleep on the beach. A few minutes later a wave woke him. Seeing Lisa was gone Jacob decided to go swimming. Lisa now believes that Jacob is sleeping.

- False belief: Item 2

  A window wiper was commissioned by a CEO to wipe an entire building. He finished the right side, but his platform broke before he could do the left side. The next morning the CEO arrived with foreign investors. The CEO comes to work and discovers that all of the walls are cleaned.

- False belief: Item 3

  Amy walked to work today. When George woke up, he saw her car in the drive. Her room was quiet and dark. George knows that when Amy is sick, she lies down in a dark room. In fact, Amy walked to work.

- False belief: Item 4

  Laura didn't have time to braid her horse's mane before going to camp. While she was at camp, William brushed Laura's horse and braided the horse's mane for her. Laura returns assuming that her horse's hair isn't braided.

- False belief: Item 5

  Larry chose a debated topic for his class paper due on Friday. The news on Thursday indicated that the debate had been solved but Larry never read it. When Larry writes his paper he thinks the debate has been solved.

- False belief: Item 6

Sally and Greg called ahead of time to make a reservation for the back-country cabin. The park ranger forgot to write down the reservation and two other hikers got to the cabin first. When the hikers arrive they see no one in their cabin.

- False belief: Item 7

  Susie parked her sportscar in the driveway. In the middle of the night Nathan moved her car into the garage to make room for his minivan. Susie woke up early in the morning. Susie sees the minivan in the driveway.

- False photo: Item 1

  When the picture was taken of the house, it was one story tall. Since then, the renovators added an additional story and a garage. The house is currently one story.

- False photo: Item 2

  Sargent famously painted the south bank of the river in 1885. In 1910 a huge dam was built, flooding out the whole river basin, killing the old forests. Now the whole area is under water. In the painting the south bank of the river is wooded.

- False photo: Item 3

  Accounts of the country's bustling economic success were recorded in both fiction and non-fiction books from the early 1900s. Soon after, a horrible plague hit the country and the country was sent into an economic depression. Early 1900s novels portray the country as experiencing economic wealth.

- False photo: Item 4

  A volcano erupted on a Caribbean island three months ago. Barren lava rock is all that remains today. Satellite photographs show the island as it was before the eruption. Today, the island is covered in lava rock.

- False photo: Item 5

  The girl's middle school pictures showed her wearing a white blouse. Later, a red sock was accidentally washed with the blouse and the blouse turned pink. Today the color of the blouse is white.

- False photo: Item 6

  A biography describes the room as it was in 1965. Originally the walls were covered in dark wallpaper. By 1965 the paper had been stripped and replaced with cream paint. The biography says that the room was light.

- False photo: Item 7

  Old maps of the islands near Titan are displayed in the Maritime museum. Erosion has since taken its toll, leaving only the three largest islands. Near Titan today there are many islands.

Appendix B: MRI Acquisition and Processing

MRI recording was performed using a 12-channel head coil in a Siemens 3T Trio Magnetic Resonance Imaging System with TIM. BOLD functional activations were measured with a T2*-weighted EPI sequence (repetition time = 2000 msec, echo time = 28 msec, flip angle = 72 deg, field of view = 222 $\times$ 222 mm, in-plane resolution = 74 $\times$ 74 pixels, and 38 axial slices with 3-mm thickness covering the entire cerebral cortex and most of the cerebellum). In addition, the anatomical structure of the brain was acquired with a three-dimensional MPRAGE sequence (1 $\times$ 1 $\times$ 1 mm^3 resolution, inversion time = 950 msec, repetition time = 1950 msec, echo time = 4.44 msec, flip angle = 12 deg, matrix size = 256 $\times$ 224, 176 sagittal slices per slab; scan time 7.5 minutes) for each participant.

fMRI image preprocessing was carried out using FEAT (FMRI Expert Analysis Tool) in FSL (FMRIB software library, version 5.0.8, www.fmrib.ox.ac.uk/fsl). The first six volumes were discarded to allow for T1 equilibrium. The remaining images were then realigned to correct for head motion. Data were spatially smoothed using a 6-mm full-width-half maximum Gaussian kernel. The data were filtered in the temporal domain using a non-linear high-pass filter with a 90-s cutoff. A two-step registration procedure was used whereby EPI images were first registered to the MPRAGE structural image, and then into the standard (MNI) space, using affine transformations. Registration from MPRAGE structural image to the standard space was further refined using FNIRT nonlinear registration.

The coordinates of the rTPJ region (region of interest; ROI) were set based on the peak Z-stats from the group analysis contrast between false belief and false photo tasks. The time course of the two ROIs for each participant was extracted from the BOLD time series after preprocessing. The onset and duration of each ToM trial were used to define a box-car function. A double-gamma hemodynamic response function (HRF) was used to convolve the box-car to construct the design matrix to account for the shape of the BOLD response in each trial. The BOLD activation level, i.e., the beta value, in each trial was extracted from the ROI time course and the design matrix using a least square solution.

Appendix C: Conventional Association Analysis

In Section 2.3, we applied three conventional methods that applied researchers might consider for ToM data analysis: correlation, logistic regression, and mixed-effects logistic regression analysis. Here we provide the formulation of the two regression models.

Logistic regression analysis

A_{pi}: a binary response for person p to item i (accuracy).

B_{pi}: a continuous response for person p to item i (fMRI activation).

- Overall association

$$\text{logit}\big(\mathbb{P}(A_{pi} = 1)\big) = \beta_0 + \beta_1 B_{pi}$$

Here β_1 represents the association between fMRI activations and response accuracy.

- Item-specific association

$$\text{logit}\big(\mathbb{P}(A_{pi} = 1)\big) = \beta_1 B_{pi} + \sum_{k=1}^{14} \beta_{k+1} d_{ik}$$

Here β_1 represents the association between fMRI activations and response accuracy after controlling for heterogeneity among test items (i.e., for differences in difficulty). Dummy variables d_{ik} ($k = 1, \cdots, 14$) represents individual items and $d_{ik} = 1$ if $i = k$ and $d_{ik} = 0$ otherwise. β_{k+1} indicates the effect of item i on response accuracy after controlling for fMRI activations levels at the item. The two regression models were fit using the glm function in R.

Logistic mixed effects model analysis

- Overall association

$$\text{logit}\big(\mathbb{P}(A_{pi} = 1|\theta_p)\big) = \beta_0 + \beta_1 B_{pi} + \theta_p, \ \ \theta_p \sim \text{N}(0, \sigma^2)$$

Here β_1 represents the association between fMRI activations and response accuracy after controlling for the respondent's overall accuracy (θ_p).

- Item-specific association

$$\operatorname{logit}\big(\mathbb{P}(A_{pi} = 1|\theta_p)\big) = \beta_1 B_{pi} + \sum_{k=1}^{14} \beta_{k+1} d_{ik} + \theta_p, \quad \theta_p \sim \mathrm{N}(0, \sigma^2)$$

Here β_1 represents the association between fMRI activations and response accuracy after controlling for heterogeneity among test items and among persons, using a random intercept for persons. These two mixed effects models were estimated using the R lme4 package (Bates & Maechler, 2009) with the Laplace approximation.

Appendix D: Interpretation of Expected Value Conditional Dependency

In Section 3.3.2 of the manuscript, we specified the proposed model for within-item conditional dependency as follows:

$$p(A_{pi} = 1|B_{pi}) = g^{-1}\left\{(\lambda_i^{(A|B)} \cdot \{w_i^{(A|B)}\}^{s(B_{pi})})\theta_p^{(A|B)} + \beta_i^{(A|B)} + v_i^{(A|B)} s(B_{pi})\right\}. \tag{1}$$

By applying log-transformation to the item slopes, $(\lambda_i^{(A|B)} \cdot \{w_i^{(A|B)}\}^{s(B_{pi})})$ for $\theta_p^{(A|B)}$, we re-write Equation (1) as

$$g(p(A_{pi} = 1|B_{pi}) = (\lambda_i^{(A|B)^*} + s(B_{pi}) w_i^{(A|B)^*})\theta_p^{(A|B)} + \beta^{(A|B)} + s(B_{pi}) v_i^{(A|B)}, \tag{2}$$

where $\lambda_i^{(A|B)^*} = \ln \lambda_i^{(A|B)}$ and $w_i^{(A|B)^*} = \ln w_i^{(A|B)}$.

With expected values of B_{pi}, $s(B_{pi} = E(B_{pi}|\theta_p^{(B)}) = \lambda_i^{(B)}\theta_p^{(B)} + \beta_i^{(B)}$. By replacing $s(B_{pi})$ with $E(B_{pi}|\theta_p^{(B)})$, we have

$$\begin{aligned}
g(p(A_{pi} = 1|B_{pi}) &= \left(\lambda_i^{(A|B)^*} + w_i^{(A|B)^*}(\lambda_i^{(B)}\theta_p^{(B)} + \beta_i^{(B)})\right)\theta_p^{(A|B)} \\
&\qquad + \beta_i^{(A|B)} + v_i^{(A|B)}(\lambda_i^{(B)}\theta_p^{(B)} + \beta_i^{(B)}), \\
&= \lambda_i^{(A|B)^*}\theta_p^{(A|B)} + w_i^{(A|B)^*}\lambda_i^{(B)}\theta_p^{(A|B)}\theta_p^{(B)} + w_i^{(A|B)^*}\beta_i^{(B)}\theta_p^{(A|B)} \\
&\qquad + \beta_i^{(A|B)} + v_i^{(A|B)}\lambda_i^{(B)}\theta_p^{(B)} + v_i^{(A|B)}\beta_i^{(B)}, \\
&= (\lambda_i^{(A|B)^*} + w_i^{(A|B)^*}\beta_i^{(B)})\theta_p^{(A|B)} + (v_i^{(A|B)}\lambda_i^{(B)})\theta_p^{(B)} + (w_i^{(A|B)^*}\lambda_i^{(B)})\theta_p^{(A|B)}\theta_p^{(B)} \\
&\qquad + (\beta_i^{(A|B)} + v_i^{(A|B)}\beta_i^{(B)}), \\
&= \lambda_i^{(A|B)^\dagger}\theta_p^{(A|B)} + \lambda_i^{(B)^\dagger}\theta_p^{(B)} + \lambda_i^{(AB)^\dagger}\theta_p^{(A|B)}\theta_p^{(B)} + \beta_i^{(A|B)^\dagger},
\end{aligned} \tag{3}$$

where $\lambda_i^{(A|B)^\dagger} = (\lambda_i^{(A|B)^*} + w_i^{(A|B)^*} + w_i^{(A|B)^*}\beta_i^{(B)})$, $\lambda_i^{(B)^\dagger} = (v_i^{(A|B)}\lambda_i^{(B)})$, $\lambda_i^{(AB)^\dagger} = (w_i^{(A|B)^*}\lambda_i^{(B)})$, and $\beta_i^{(A|B)^\dagger} = (\beta_i^{(A|B)} + v_i^{(A|B)}\beta_i^{(B)})$.

It is clear from Equation (3) that the expected value conditional dependency assumes (1) A_{pi} is influenced by both $\theta_p^{(A|B)}$ and $\theta_p^{(B)}$; that is, A_{pi} is cross-loaded on $\theta_p^{(B)}$ as well as on $\theta_p^{(A|B)}$, and (2) the effect of $\theta_p^{(A|B)}$ on A_{pi} is moderated by $\theta_p^{(B)}$. Further, it is important to note that when $\lambda_i^{(B)}$ is non-trivial, meaning that B_{pi} is a good measure of $\theta_p^{(B)}$, conditional dependency associated with item intercepts $v_i^{(A|B)}$ implies the direct effect of $\theta_p^{(B)}$ on A_{pi}, while conditional dependency associated with item slopes $w_i^{(A|B)}$ implies the moderated effect of $\theta_p^{(B)}$ on the effect of $\theta_p^{(A)}$ on A_{pi}. To illustrate this point, suppose we identified positive $w_i^{(A|B)}$ and $v_i^{(A|B)}$ for particular items from the expected value conditional dependency model with the ToM example. In this case, positive $v_i^{(A|B)}$ indicates that overall brain activation levels improved response accuracy for those items, and (2) positive $w_i^{(A|B)}$ indicates that the effect of overall brain activations was additionally boosted for those subjects with a higher-level of overall accuracy (or ability).

Appendix E: Graphical Representation of Three Kinds of Conditional Dependency

Figure 1 provides graphical representations of the three within-item conditional dependency models, compared with the conditional independence model which is described as the baseline model in Section 4.1.1 of the manuscript. For illustration purposes, a within-item conditional dependency is assumed only for the intercept parameter of the first item (out of three items). To additionally accommodate within-item conditional dependency of B_{pi} on the item slope parameter of A_{pi}, we can include additional unidirectional arrows that come from B_{p1}, $\theta_p^{(B)}$, and $r_{p1}^{(B)}$ to go toward $\lambda_1^{(A)}$ in Figures 1(b), (c), and (d). In all figures, the latent variables for A_{pi} and B_{pi} are correlated. Correlations between the item parameters are suppressed in the figures for the clarity of the presentation.

Figure 1(a) illustrates the conditional independence model which means that B_{pi} impacts A_{pi} through correlated latent variables and correlated item parameters only, meaning that B_{pi} is not directly related to A_{pi}; i.e., they are assumed independent given the latent variable correlation. Figure 1(b) presents the raw value conditional dependency model. This model allows for a direct effect of B_{p1} on A_{p1} for the first item. Notice that $\theta^{(B)}$ has now an effect on A_{p1} indirectly through B_{pi}. In other words, the effect of $\theta^{(B)}$ on A_{p1} is mediated by B_{pi}. Figure 1(c) illustrates the expected value conditional dependency model. This model assumes that the latent variable of B_{pi} ($\theta_p^{(B)}$) has a direct influence on A_{p1}, as we showed in Appendix D. Lastly, Figure 1(d) illustrates the residual value conditional dependency model which assumes the residual of B_{pi} ($r_{p1}^{(B)}$) has an effect on A_{p1}. The residuals $r_{p1}^{(B)}$ are independent of latent variable $\theta^{(B)}$, meaning that this model assumes no effect of $\theta^{(B)}$ on A_{pi}.

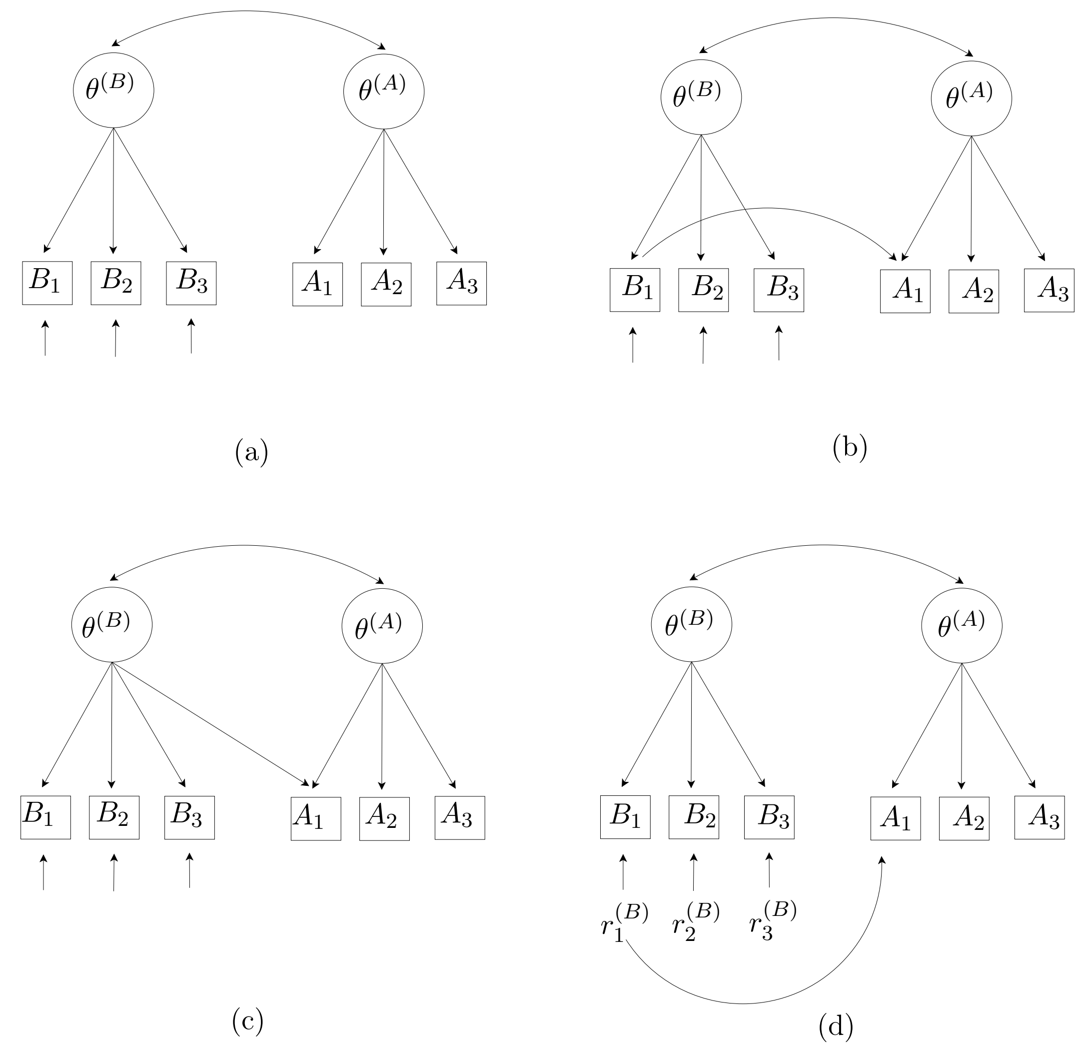


Figure 1: Graphical representation of (a) conditional independence model, (b) raw value conditional dependency model, (c) expected value conditional dependency model, and (d) residual value conditional dependency model.

Appendix F: Model Formulation for Three Extra Scenarios

In Section 3.3.1 of the manuscript, we presented our model to describe the conditional dependency of $A_{pi}|B_{pi}$ where $A_{pi} \in \{0,1\}$ and $B_{pi} \in \mathbb{R}$. We provided three different kinds of conditional dependency scenarios in Section 3.3.4. Here we provide a general formulation of the proposed joint model that can cover those extra conditional dependency scenarios. To this purpose, we utilize two variables X_{pi} and Y_{pi} and define the support of the two variables per scenario. A general joint model for (X_{pi}, Y_{pi}) can be expressed to specify the $(X_{pi}|Y_{pi})$ conditional dependency as follows:

$$f(X_{pi}, Y_{pi}|\theta_p^{(X|Y)}, \theta_p^{(Y)}) = f(Y_{pi}|\theta_p^{(Y)}) f(X_{pi}|Y_{pi}\theta_p^{(X|Y)}), \tag{4}$$

where

$$f(X_{pi}|Y_{pi}, \theta_p^{(X|Y)}) = g^{-1}\left\{(\lambda_i^{(X|Y)} \cdot \{w_i^{(X|Y)}\}^{s(Y_{pi})})\theta_p^{(X|Y)} + \beta_i^{(X|Y)} + v_i^{(X|Y)} s(Y_{pi})\right\}. \tag{5}$$

Here $s(Y_{pi})$ is a function of Y_{pi} and it can take the form of (1) raw values (Y_{pi}), (2) expected values $(E(Y_{pi}))$, or (3) residuals (r_{pi}) of Y_{pi}. With the ToM data, $g(\cdot)$ is the logit link function, $X_{pi} \sim \text{Bernoulli}(p_{pi})$, where $f(X_{pi}|\theta_p^{(X)}) = p_{pi}^{x_{pi}}(1-p_{pi})^{(1-x_{pi})}$, and $Y_{pi} \sim \text{Normal}(\mu_{pi}, \sigma_i^{2(Y)})$, where $\mu_{pi} = f(Y_{pi}|\theta_p^{(Y)})$. In Scenario 1, $X_{pi} \in \mathbb{R}$ and $Y_{pi} \in \{0,1\}$. In this case, $g(\cdot)$ is the identity link function, and $f(Y_{pi}|\theta_p^{(Y)}) = p_{pi}^{y_{pi}}(1-p_{pi})^{(1-y_{pi})}$ where $Y_{pi} \sim \text{Bernoulli}(p_{pi})$. In Scenario 2, $X_{pi} \in \{0,1\}$ and $Y_{pi} \in \{0,1\}$. In this case, $g(\cdot)$ is the logit link function, and $f(Y_{pi}|\theta_p^{(Y)}) = p_{pi}^{y_{pi}}(1-p_{pi})^{(1-y_{pi})}$ where $Y_{pi} \sim \text{Bernoulli}(p_{pi})$. In Scenario 3, $X_{pi} \in \mathbb{R}$ and $Y_{pi} \in \mathbb{R}$. In this case, $g(\cdot)$ is the identity link function, and $Y_{pi} \sim \text{Normal}(\mu_{pi}, \sigma_i^{2(Y)})$, where $\mu_{pi} = f(Y_{pi}|\theta_p^{(Y)})$.

Appendix G: JAGS Code

Here we provide the example JAGS code that can be used to fit the within-item conditional dependency models. The parameter values for the prior distributions were chosen for the ToM bivariate data analysis (response accuracy and fMRI activations) that we used to illustrate the proposed modeling approach.

```
## definitions
# N: number of respondents (j=1,...,N)
# I: number of items (i=1,...,I)
# u: continuous data B (fMRI activations): N by I matrix
# mu: the mean structure for B (N by I)
# prec: precision parameters (1/variance) for B residuals
# r: binary data A (response accuracy) : N by I matrix
# p: success probability for A (N by I)
# ip[i,1]: log slope parameter for B (log lambda_i)
# ip[i,2]: intercept parameter for B (beta_i)
# ip[i,3]: log slope parameter for A (log lambda_i)
# ip[i,4]: intercept parameter for A (beta_i)
# th[j,1]: latent variable for B
# th[j,2]: latent variable for A
# log_w: log dependency parameters for item slopes (log w_i)
# v: dependency parameters for item intercepts (v_i)

model {

for (j in 1:N) {
    for (i in 1:I) {


            # model for variable B (fMRI data)
            # Equation (1)
            u[j,i] ~ dnorm(mu[j,i],prec[i])
            mu[j,i] <- a[i]*th[j,1] + ip[i,2]

            # conditional model for variable A (response accuracy)
            # note that this part changes depending on which modeling option is taken.
            ## with s(B) = raw values of B
            # Equation (4) with log transformation on the item slopes
            r[j,i] ~ dbern(p[j,i])
            logit(p[j,i])  <- exp( ip[i,3] + log_w[i]*u[j,i])*th[j,2] +ip[i,4] + v[i]*u[j,i]
    }
}


## Prior specification

# latent variable distribution
# the mean and the sd are fixed at 0 and 1, respectively for A and B.
# hence, only the correlation is estimated.
# to estimate the correlation parameter, we apply Cholesky decomposition (Cov= LL').
# cc is the off-diagonal element of the lower triangler matrix L.
for (j in 1:N) {
    th[j,1] ~ dnorm(0, 1)
```

```
    mu.th[j]<-cc*th[j,1]
    th[j,2] ~dnorm(mu.th[j],1)
}

cc~dnorm(0,1)


# item parameter distribution
for (i in 1:I) {

    # note that for slope parameters, we work with  log-transformed parameters
    # hence, all four parameters (two log slopes and two intercepts
    # are assumed to follow a multivariate normal prior distribution

    ip[i,1:4] ~ dmnorm(mu, Omega)

# item slopes back to the original scale
a[i] <-  exp(ip[i,1]) # for data B
af[i] <-  exp(ip[i,3]) # for data A

}

# hyper priors for the item parameter distribution
mu ~ dmnorm(mu0, Omega)

# the means are fixed at 0
mu0[1] <- 0
mu0[2] <- 0
mu0[3] <- 0
mu0[4] <- 0

# apply the Wishart prior to the inverse of the covariance matrix
# R is the scale matrix (4 by 4 identity matrix)
# degrees of freedom = 4
Omega ~ dwish(R,4)
Sigma <- inverse(Omega)


# conditional dependency parameters
for (i in 1:I) {

    # for the slopes
    log_w[i] ~ dnorm(0,1)
    # back to the original scale
    w[i] <- exp(log_w[i])

    # for the intercepts
    v[i] ~ dnorm(0, 1)

}


# residual variances for B
# apply gamma prior to the precision parameter (inverse of variance)
for (i in 1:I) {
    prec[i] ~ dgamma(1,1)
    tau[i] <- 1/prec[i]
}


}
```

```
}
```

The code described above is relevant for estimating the raw value conditional dependency model. For estimating the expected value and residual value conditional dependency models, the first part of the code labeled as [conditional model for variable A] needs to be modified as follows:

```
## expected value conditional dependency
# Equation (5) with log transformation on the item slopes
r[j,i] ~ dbern(p[j,i])
logit(p[j,i])  <- exp( ip[i,3] + log_w[i]*mu[j,i])*th[j,2] +ip[i,4] + v[i]*mu[j,i]

## residual value conditional dependency
# Equation (7) with log transformation on the item slopes
r[j,i] ~ dbern(p[j,i])
# define the residual
res[j,i] <-    u[j,i] -    mu[j,i]
logit(p[j,i])  <- exp( ip[i,3] + log_w[i]*res[j,i])*th[j,2] +ip[i,4] + v[i]*res[j,i]
```

The conditional independence model does not contain the conditional dependency parameters. Hence, to estimate the conditional independence model, the code labeled as [conditional dependency parameters] should be removed. In addition, the code [conditional model for variable A] needs to be modified as follows:

```
## conditional independence model
r[j,i] ~ dbern(p[j,i])
logit(p[j,i])  <- af[i]*th[j,2] +ip[i,4]
```

Appendix H: Convergence Evidence

Here we provide the trace plots of the posterior samples for some of the parameters for the three conditional dependency models fitted to the ToM bivariate data.

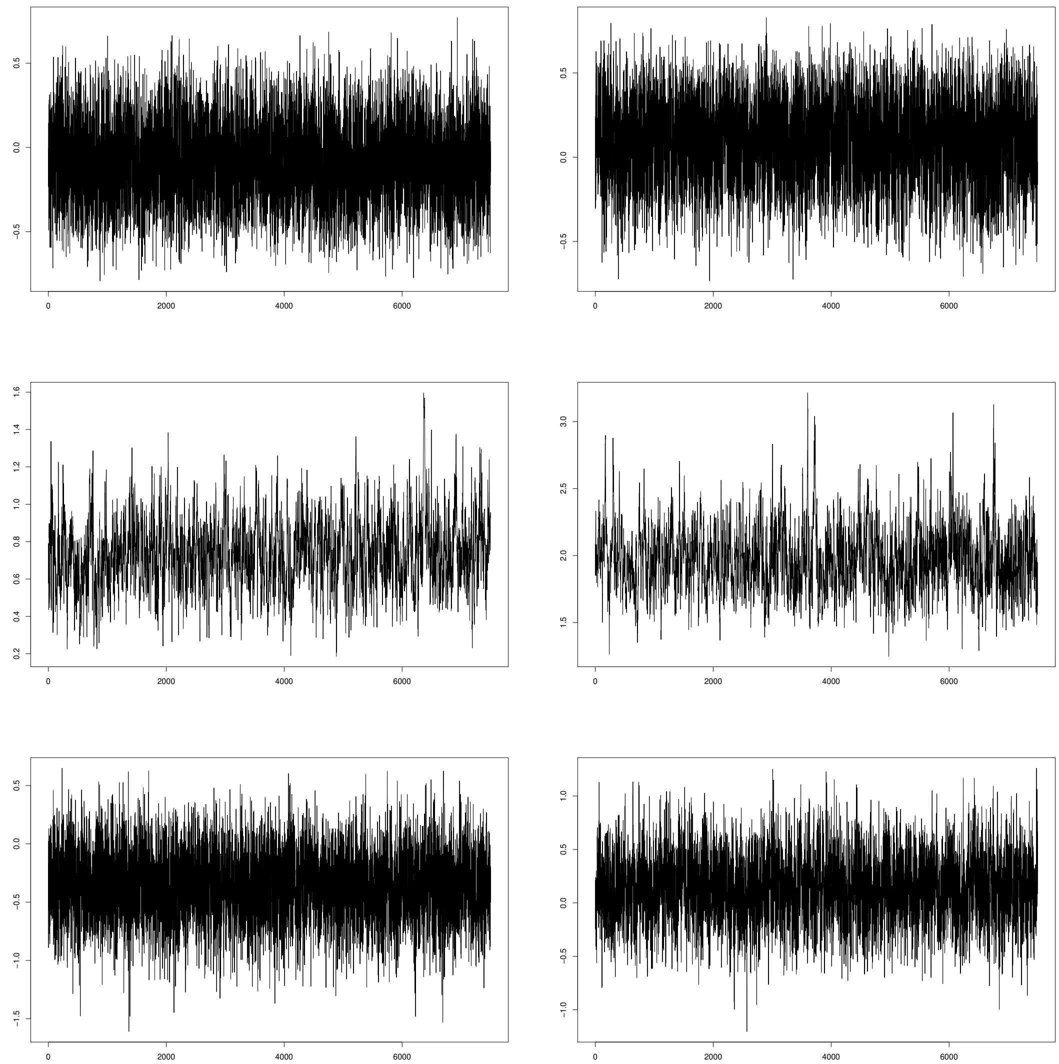


Figure 2: Trace plots of the posterior samples (after burn-in) for the parameters of the raw value conditional dependency model. Top panel: two conditional dependency parameters; Middle panel: two correlation parameters; Bottom panel: two item parameters.

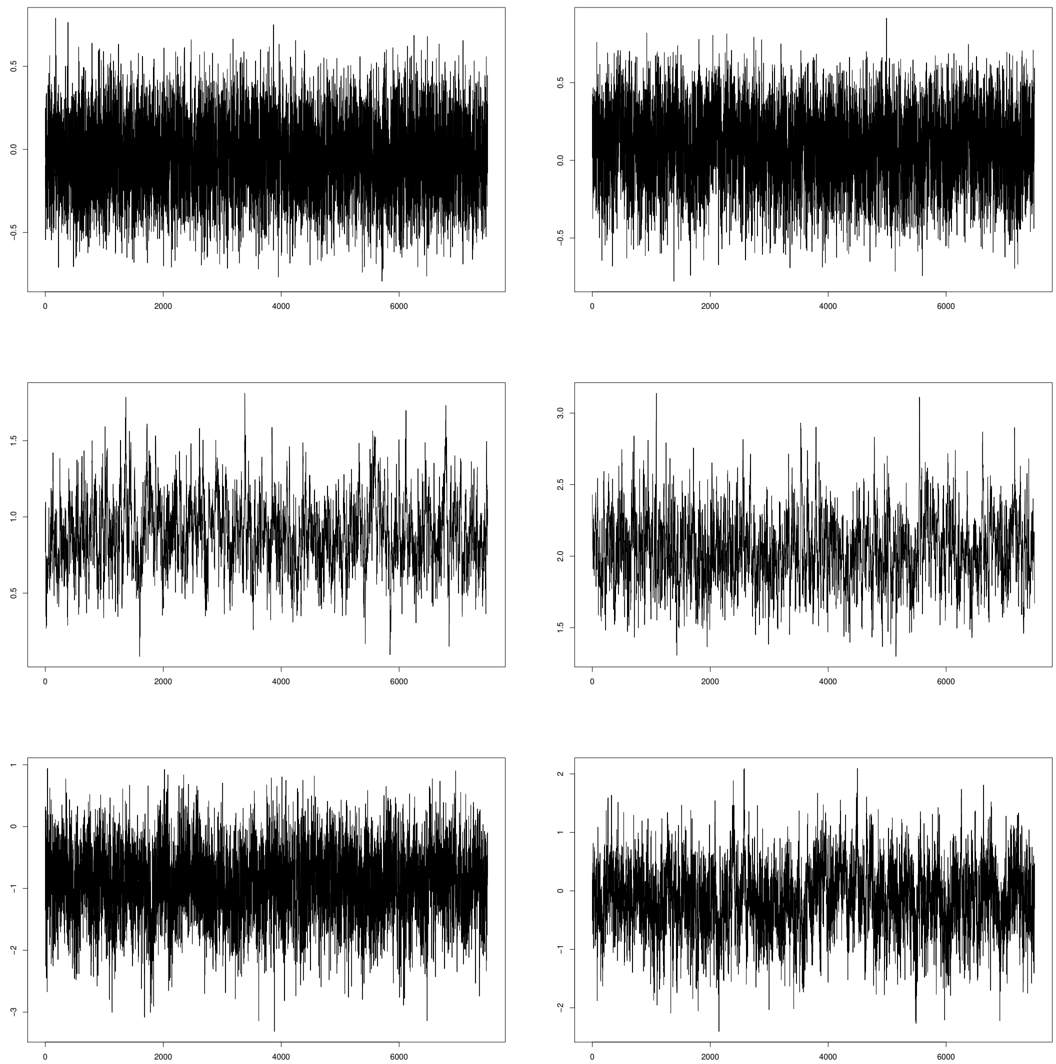


Figure 3: Trace plots of the posterior samples (after burn-in) for the parameters of the expected value conditional dependency model. Top panel: two conditional dependency parameters; Middle panel: two correlation parameters; Bottom panel: two item parameters.

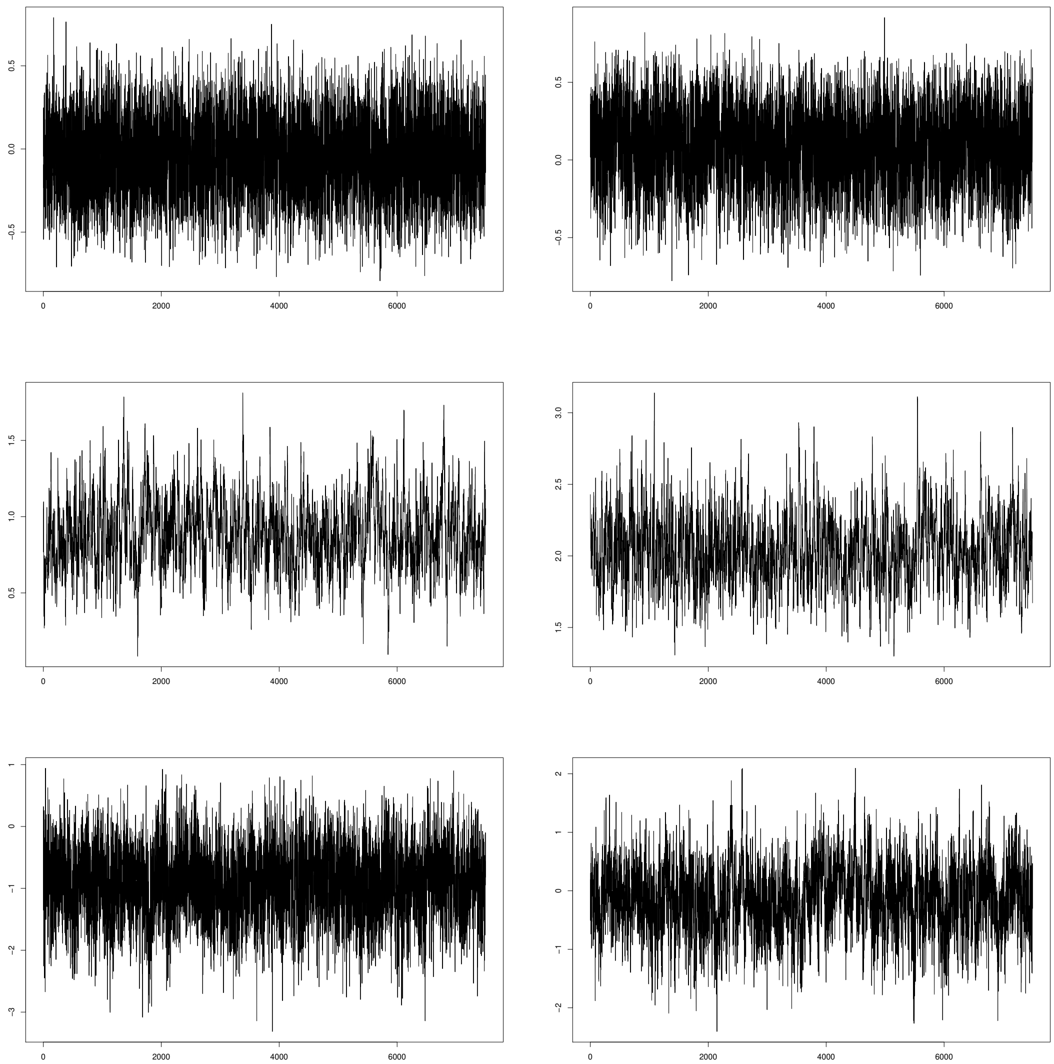


Figure 4: Trace plots of the posterior samples (after burn-in) for the parameters of the residual value conditional dependency model. Top panel: two conditional dependency parameters; Middle panel: two correlation parameters; Bottom panel: two item parameters.

Appendix I: Parameter Estimates from the ToM Data Analysis

Here we provide the posterior means and the 95% credible intervals for the model parameters from the analysis of the ToM data with the three conditional dependency models and the conditional independence model. Within-item conditional dependency parameter estimates were reported in the manuscript.

Table 1: Item parameter estimates (posterior mean and 95% credible interval) from the raw value conditional dependency model. $\lambda_i^{(A)}$ and $\beta_i^{(A)}$ for response accuracy; $\lambda_i^{(B)}$ and $\beta_i^{(B)}$ for fMRI activations

	$\lambda_i^{(A)}$			$\beta_i^{(A)}$		
	Mean	95% CI		Mean	95% CI	
Item		Lower	Upper		Lower	Upper
1	0.425	0.160	0.959	0.737	0.363	1.098
2	1.208	0.513	2.578	3.414	2.430	4.532
3	0.920	0.364	2.037	1.696	1.093	2.437
4	0.572	0.276	1.127	1.303	0.896	1.677
5	0.760	0.345	1.516	2.141	1.573	2.656
6	0.176	0.058	0.486	-0.577	-0.882	-0.263
7	0.552	0.248	1.140	0.977	0.611	1.407
8	0.983	0.381	2.320	3.497	2.603	4.469
9	0.446	0.193	0.927	0.741	0.453	1.044
10	0.899	0.439	1.847	1.443	0.995	1.907
11	0.975	0.452	2.017	2.649	2.007	3.325
12	1.077	0.490	2.299	3.479	2.612	4.389
13	0.180	0.066	0.449	-0.436	-0.697	-0.178
14	0.610	0.273	1.268	1.981	1.505	2.437

	$\lambda_i^{(B)}$			$\beta_i^{(B)}$		
	Mean	95% CI		Mean	95% CI	
Item		Lower	Upper		Lower	Upper
1	0.182	0.142	0.233	0.238	0.194	0.285
2	0.262	0.214	0.322	0.199	0.146	0.256
3	0.269	0.220	0.324	0.317	0.260	0.370
4	0.231	0.187	0.283	0.227	0.177	0.279
5	0.273	0.223	0.328	0.229	0.173	0.284
6	0.220	0.172	0.272	0.281	0.227	0.332
7	0.232	0.186	0.291	0.308	0.255	0.361
8	0.135	0.097	0.180	-0.089	-0.129	-0.046
9	0.190	0.148	0.237	-0.115	-0.159	-0.067
10	0.169	0.128	0.223	-0.195	-0.248	-0.143
11	0.159	0.119	0.211	-0.122	-0.173	-0.076
12	0.171	0.127	0.224	0.004	-0.043	0.054
13	0.208	0.159	0.263	-0.066	-0.124	-0.016
14	0.200	0.151	0.265	-0.149	-0.209	-0.091

Table 2: Item parameter estimates (posterior mean and 95% credible interval) from the expected value conditional dependency model. $\lambda_i^{(A)}$ and $\beta_i^{(A)}$ for response accuracy; $\lambda_i^{(B)}$ and $\beta_i^{(B)}$ for fMRI activations

	$\lambda_i^{(A)}$			$\beta_i^{(A)}$		
	Mean	95% CI		Mean	95% CI	
Item		Lower	Upper		Lower	Upper
1	0.470	0.203	1.032	0.889	0.450	1.374
2	1.291	0.592	2.639	3.440	2.434	4.599
3	1.097	0.537	2.083	2.036	1.298	2.805
4	0.715	0.370	1.339	1.537	1.048	2.086
5	0.802	0.386	1.707	2.204	1.595	2.891
6	0.208	0.067	0.565	-0.489	-0.898	-0.036
7	0.754	0.339	1.556	1.206	0.596	1.795
8	1.185	0.477	2.859	3.553	2.610	4.583
9	0.439	0.207	0.924	0.712	0.386	0.998
10	0.794	0.360	1.664	1.507	0.999	1.988
11	1.025	0.440	2.207	2.614	1.997	3.253
12	1.187	0.511	2.575	3.560	2.616	4.562
13	0.204	0.068	0.507	-0.479	-0.756	-0.214
14	0.686	0.282	1.406	2.043	1.545	2.504
	$\lambda_i^{(B)}$			$\beta_i^{(B)}$		
	Mean	95% CI		Mean	95% CI	
Item		Lower	Upper		Lower	Upper
1	0.168	0.129	0.221	0.238	0.191	0.284
2	0.241	0.187	0.304	0.198	0.143	0.257
3	0.249	0.196	0.313	0.317	0.265	0.372
4	0.213	0.165	0.272	0.228	0.175	0.280
5	0.251	0.197	0.314	0.227	0.173	0.281
6	0.200	0.152	0.256	0.281	0.228	0.333
7	0.215	0.165	0.275	0.307	0.253	0.362
8	0.123	0.089	0.166	-0.089	-0.132	-0.048
9	0.173	0.131	0.227	-0.115	-0.163	-0.068
10	0.154	0.108	0.211	-0.196	-0.248	-0.143
11	0.148	0.104	0.204	-0.122	-0.170	-0.072
12	0.155	0.113	0.208	0.004	-0.048	0.050
13	0.191	0.145	0.252	-0.067	-0.121	-0.012
14	0.182	0.131	0.244	-0.149	-0.208	-0.089

Table 3: Item parameter estimates (posterior mean and 95% credible interval) from the residual value conditional dependency model. $\lambda_i^{(A)}$ and $\beta_i^{(A)}$ for response accuracy; $\lambda_i^{(B)}$ and $\beta_i^{(B)}$ for fMRI activations

	$\lambda_i^{(A)}$			$\beta_i^{(A)}$		
Item	Mean	95% CI		Mean	95% CI	
1	0.422	0.189	0.870	0.793	0.514	1.072
2	1.327	0.634	2.825	3.554	2.546	4.808
3	1.397	0.677	2.728	2.441	1.668	3.342
4	0.658	0.354	1.171	1.429	1.070	1.795
5	0.878	0.458	1.614	2.366	1.857	2.977
6	0.175	0.066	0.417	-0.543	-0.801	-0.270
7	0.711	0.403	1.282	1.330	0.997	1.700
8	0.986	0.417	2.338	3.456	2.563	4.530
9	0.441	0.209	0.907	0.766	0.493	1.037
10	0.948	0.491	1.825	1.381	0.971	1.814
11	0.793	0.376	1.567	2.583	2.006	3.196
12	1.040	0.449	2.366	3.524	2.614	4.551
13	0.169	0.064	0.404	-0.395	-0.655	-0.140
14	0.605	0.277	1.201	1.963	1.548	2.408
	$\lambda_i^{(B)}$			$\beta_i^{(B)}$		
	Mean	95% CI		Mean	95% CI	
1	0.177	0.140	0.224	0.238	0.194	0.285
2	0.257	0.209	0.317	0.199	0.143	0.258
3	0.263	0.216	0.318	0.318	0.261	0.374
4	0.226	0.183	0.279	0.227	0.169	0.278
5	0.268	0.222	0.324	0.228	0.174	0.286
6	0.213	0.168	0.274	0.281	0.229	0.336
7	0.224	0.176	0.280	0.309	0.252	0.363
8	0.132	0.097	0.177	-0.089	-0.131	-0.048
9	0.186	0.146	0.238	-0.115	-0.162	-0.066
10	0.167	0.127	0.223	-0.195	-0.246	-0.143
11	0.151	0.109	0.198	-0.122	-0.171	-0.073
12	0.167	0.126	0.218	0.004	-0.048	0.050
13	0.201	0.154	0.258	-0.067	-0.123	-0.011
14	0.197	0.147	0.266	-0.149	-0.211	-0.089

Table 4: Item parameter estimates (posterior mean and 95% credible interval) from the conditional independence model. $\lambda_i^{(A)}$ and $\beta_i^{(A)}$ for response accuracy; $\lambda_i^{(B)}$ and $\beta_i^{(B)}$ for fMRI activations

	$\lambda_i^{(A)}$			$\beta_i^{(A)}$		
Item	Mean	95% CI		Mean	95% CI	
1	0.376	0.172	0.788	0.769	0.507	1.063
2	1.330	0.686	2.479	3.469	2.646	4.444
3	1.201	0.675	2.048	2.118	1.564	2.732
4	0.628	0.335	1.144	1.374	1.027	1.718
5	0.876	0.445	1.701	2.327	1.781	2.910
6	0.184	0.070	0.401	-0.526	-0.763	-0.265
7	0.794	0.398	1.371	1.326	0.959	1.709
8	1.128	0.578	2.148	3.504	2.737	4.367
9	0.406	0.182	0.780	0.759	0.486	1.019
10	0.944	0.520	1.677	1.352	0.944	1.740
11	0.943	0.484	1.748	2.674	2.022	3.341
12	1.171	0.575	2.422	3.526	2.641	4.515
13	0.157	0.064	0.363	-0.405	-0.657	-0.152
14	0.682	0.335	1.347	1.977	1.549	2.415
	$\lambda_i^{(B)}$			$\beta_i^{(B)}$		
	Mean	95% CI		Mean	95% CI	
1	0.170	0.131	0.216	0.239	0.191	0.282
2	0.248	0.196	0.303	0.201	0.146	0.258
3	0.256	0.211	0.311	0.319	0.261	0.373
4	0.219	0.176	0.269	0.229	0.177	0.280
5	0.257	0.212	0.316	0.230	0.173	0.283
6	0.206	0.162	0.263	0.282	0.227	0.333
7	0.221	0.177	0.278	0.309	0.254	0.361
8	0.127	0.094	0.168	-0.088	-0.129	-0.045
9	0.178	0.138	0.228	-0.114	-0.163	-0.068
10	0.160	0.119	0.214	-0.194	-0.248	-0.143
11	0.150	0.111	0.197	-0.120	-0.166	-0.068
12	0.160	0.120	0.210	0.005	-0.043	0.054
13	0.193	0.148	0.250	-0.065	-0.116	-0.010
14	0.187	0.138	0.247	-0.148	-0.206	-0.089

Table 5: Error variance estimates for fMRI activation (posterior mean and 95% credible interval) from the three conditional dependency models and the conditional independence model.

	Raw			Exp			Res			Ind		
Item	Mean	95% CI		Mean	95% CI		Mean	95% CI		Mean	95% CI	
1	0.10	0.08	0.12	0.10	0.08	0.12	0.10	0.08	0.12	0.10	0.08	0.12
2	0.12	0.10	0.15	0.12	0.10	0.15	0.13	0.10	0.15	0.12	0.10	0.15
3	0.11	0.09	0.13	0.11	0.09	0.13	0.11	0.09	0.13	0.11	0.09	0.13
4	0.11	0.09	0.13	0.11	0.09	0.13	0.11	0.09	0.14	0.11	0.09	0.13
5	0.10	0.08	0.13	0.10	0.08	0.13	0.10	0.08	0.13	0.10	0.08	0.13
6	0.12	0.10	0.14	0.12	0.10	0.14	0.12	0.10	0.15	0.12	0.10	0.14
7	0.13	0.10	0.15	0.13	0.10	0.15	0.13	0.10	0.15	0.13	0.10	0.15
8	0.09	0.07	0.11	0.09	0.07	0.11	0.09	0.07	0.11	0.09	0.07	0.11
9	0.10	0.08	0.12	0.10	0.08	0.12	0.10	0.08	0.12	0.10	0.08	0.12
10	0.14	0.12	0.17	0.14	0.12	0.17	0.14	0.12	0.17	0.14	0.12	0.17
11	0.13	0.10	0.15	0.13	0.10	0.15	0.13	0.10	0.15	0.12	0.10	0.15
12	0.12	0.10	0.14	0.12	0.10	0.14	0.12	0.10	0.14	0.12	0.10	0.14
13	0.14	0.11	0.17	0.14	0.11	0.17	0.14	0.12	0.17	0.14	0.12	0.17
14	0.18	0.15	0.22	0.18	0.15	0.22	0.18	0.15	0.22	0.18	0.15	0.22

Appendix J: Simulation Study 2

Data Generation Details

For creating the two scenarios considered in the Simulation Study 2, we considered the following setting. For Scenario 1, we set $c_\theta = -0.5$ (i.e., $\rho_\theta = -0.5/\sqrt{(-0.5)^2+1} = -0.447$), $w^{(A|B)} = 2.7$ (i.e., $\ln w^{(A|B)} = 1$, $v^{(A|B)} = 2$). For Scenario 2, we set $c_\theta = 0.5$ (i.e., $\rho_\theta = 0.447$), $w^{(A|B)} = 2.7$ (i.e., $\ln w^{(A|B)} = 1$, $v^{(A|B)} = -1$). For other parameters, the same set of data generating values were used in both scenarios. For the item parameters, true parameter values were generated from the multivariate normal prior with means of $\mu_\delta = (-0.12, 0.82, -0.19, -0.34)'$, variances of 0.2, and covariances of 0. We set $\sigma_i^{2(B)} = 1$.

Bias and coverage rates

Table 6: Absolute bias and coverage rates for Scenario 1 of Simulation study 2.

	Raw		Expected		Residual		Independence	
	bias	prop	bias	prop	bias	prop	bias	prop
Raw condition								
$\lambda^{(A\|B)}$	0.13	0.88	0.26	0.73	0.87	0.38	0.52	0.42
$\lambda^{(B)}$	0.03	0.95	0.01	0.96	0.13	0.56	0.02	0.97
$\beta^{(A\|B)}$	0.10	0.94	0.26	0.80	0.98	0.21	0.43	0.47
$\beta^{(B)}$	0.02	0.97	0.01	0.98	0.01	0.98	0.01	0.98
c_θ	0.10	0.92	0.22	0.34	0.25	0.02	0.56	0.00
μ_δ	0.05	0.99	0.10	0.91	0.43	0.50	0.24	0.52
σ_δ	0.03	1.00	0.04	1.00	0.13	0.92	0.05	1.00
$\sigma^{2(B)}$	0.01	0.94	0.02	0.93	0.06	0.90	0.01	0.94
$w^{(A\|B)}$	0.20	0.97	0.30	0.94	0.71	0.98		
$w^{(B)}$	0.23	0.86	1.15	0.05	0.38	0.75		
Expected condition								
$\lambda^{(A\|B)}$	0.77	0.41	0.15	0.97	1.12	0.14	3.23	0.13
$\lambda^{(B)}$	0.01	0.98	0.03	0.97	0.02	0.98	0.17	0.98
$\beta^{(A\|B)}$	0.39	0.61	0.17	0.92	0.42	0.55	1.93	0.54
$\beta^{(B)}$	0.01	0.98	0.01	0.98	0.01	0.97	0.13	0.98
c_θ	0.26	0.00	0.15	0.80	0.29	0.00	4.00	0.00
μ_δ	0.25	0.62	0.06	1.00	0.30	0.63	1.81	0.64
σ_δ	0.09	1.00	0.04	1.00	0.09	0.98	0.77	0.98
$\sigma^{2(B)}$	0.01	0.96	0.01	0.95	0.01	0.96	0.13	0.95
$w^{(A\|B)}$	1.44	0.04	0.42	0.89	1.64	0.01		
$v^{(A\|B)}$	0.89	0.01	0.35	0.78	0.98	0.01		
Residual condition								
$\lambda^{(A\|B)}$	0.53	0.52	0.28	0.75	0.12	0.93	0.17	0.82
$\lambda^{(B)}$	0.28	0.02	0.01	0.96	0.03	0.95	0.02	0.95
$\beta^{(A\|B)}$	1.27	0.08	0.24	0.83	0.12	0.89	0.21	0.65
$\beta^{(B)}$	0.02	0.98	0.01	0.98	0.03	0.97	0.01	0.98
c_θ	0.80	0.00	0.30	0.48	0.11	0.92	0.23	0.26
μ_δ	0.52	0.26	0.12	0.88	0.04	1.00	0.09	1.00
σ_δ	0.12	0.82	0.03	1.00	0.03	1.00	0.02	1.00
$\sigma^{2(B)}$	0.01	0.95	0.01	0.95	0.01	0.94	0.01	0.95
$w^{(A\|B)}$	1.43	0.00	1.37	0.30	0.34	0.92		
$v^{(A\|B)}$	0.47	0.46	2.11	0.00	0.21	0.87		

Table 7: Absolute bias and coverage rates for Scenario 2 of Simulation study 2.

	Raw		Expected		Residual		Independence	
	bias	prop	bias	prop	bias	prop	bias	prop
Raw condition								
$\lambda^{(A\|B)}$	0.12	0.92	0.21	0.80	1.04	0.27	0.72	0.25
$\lambda^{(B)}$	0.04	0.93	0.01	0.97	0.05	0.89	0.05	0.92
$\beta^{(A\|B)}$	0.07	0.95	0.21	0.91	0.34	0.65	0.22	0.72
$\beta^{(B)}$	0.02	0.97	0.01	0.99	0.02	0.98	0.02	0.98
c_θ	0.10	0.98	0.01	1.00	0.08	0.94	0.26	0.02
μ_δ	0.04	1.00	0.09	0.97	0.27	0.73	0.19	0.75
σ_δ	0.03	1.00	0.04	1.00	0.10	0.98	0.05	1.00
$\sigma^{2(B)}$	0.01	0.95	0.01	0.95	0.03	0.94	0.01	0.95
$w^{(A\|B)}$	0.28	0.97	0.34	0.96	0.57	0.98		
$v^{(A\|B)}$	0.05	0.95	0.43	0.67	0.16	0.88		
Expected condition								
$\lambda^{(A\|B)}$	0.74	0.50	0.17	0.92	1.23	0.14	1.18	0.12
$\lambda^{(B)}$	0.04	0.95	0.01	0.98	0.05	0.93	0.05	0.93
$\beta^{(A\|B)}$	0.36	0.63	0.17	0.91	0.36	0.63	0.36	0.61
$\beta^{(B)}$	0.01	0.98	0.01	0.98	0.01	0.98	0.01	0.98
c_θ	0.20	0.24	0.00	1.00	0.24	0.00	0.25	0.00
μ_δ	0.24	0.74	0.07	1.00	0.30	0.75	0.30	0.75
σ_δ	0.10	0.99	0.04	1.00	0.10	0.94	0.10	0.93
$\sigma^{2(B)}$	0.01	0.95	0.01	0.95	0.02	0.95	0.01	0.94
$w^{(A\|B)}$	1.37	0.05	0.54	0.96	1.66	0.01		
$v^{(A\|B)}$	0.88	0.01	0.26	0.85	0.98	0.01		
Residual condition								
$\lambda^{(A\|B)}$	0.24	0.76	0.26	0.75	0.10	0.94	0.11	0.89
$\lambda^{(B)}$	0.18	0.18	0.02	0.97	0.04	0.92	0.02	0.96
$\beta^{(A\|B)}$	0.58	0.30	0.19	0.89	0.04	0.94	0.15	0.75
$\beta^{(B)}$	0.01	0.99	0.02	0.98	0.02	0.99	0.02	0.99
c_θ	0.57	0.00	0.16	0.92	0.09	0.94	0.00	1.00
μ_δ	0.25	0.69	0.12	0.91	0.04	1.00	0.07	1.00
σ_δ	0.06	1.00	0.04	1.00	0.03	1.00	0.03	1.00
$\sigma^{2(B)}$	0.01	0.95	0.01	0.95	0.01	0.95	0.01	0.96
$w^{(A\|B)}$	1.25	0.02	1.33	0.30	0.37	0.95		
$w^{(B)}$	0.16	0.74	1.17	0.03	0.03	0.94		

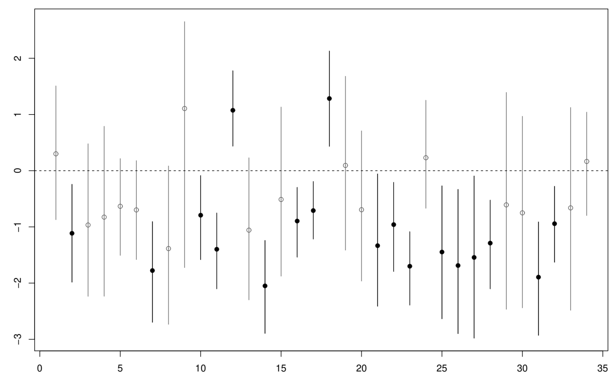
Appendix K: Additional Examples

Here we report within-item conditional dependency estimates from two applications that were remarked in the Discussion section of the manuscript. The data for the first application was obtained from 726 subjects who took a verbal intelligence test with 34 items. This dataset was analyzed in DiTrapani, Jeon, De Boeck, and Partchev (2016), among others. The data from the second application was obtained from 1000 subjects who took a low-stakes educational assessment with 60 items. This dataset was analyzed in Jeon and De Boeck (2019), among others. In both applications, the associations of response accuracy (A) with response times (B) were analyzed. Three conditional dependency models, as well as the conditional independence model, were applied to the two datasets. Table 8 suggests that the residual value conditional dependency model showed the best fit in terms of DIC in both applications. Figure 5 displays the within-item conditional dependency parameter estimates, $v_i^{(A|B)}$ and $w_i^{(A|B)}$ from the residual value conditional dependency model applications. With Data 1, we found 18 items show meaningful $v_i^{(A|B)}$ and 4 items have meaningful $w_i^{(A|B)}$ where 95% credible intervals do not include 0 and 1, respectively. With Data 2, we found 15 items have meaningful $v_i^{(A|B)}$ and 25 items show meaningful $w_i^{(A|B)}$ estimates. The latent variable correlation was different in the two datasets: 0.08 [-0.03, 0.18] and 0.50 [0.12, 0.93] with applications 1 and 2, respectively. With application 1, while the latent variable correlation was near zero, about 50% of the items showed positive within-item conditional dependencies with the intercept parameters. This means that although subjects' overall speed (response time latent variable) and ability (response accuracy latent variable) were unrelated, spending extra time on some items increased the probability of giving a correct response for the majority of items. Interestingly, with application 2, the latent variable correlation was positive, but a number of items showed negative within-item conditional dependencies both for the intercept and slope parameters. This means that while slow respondents tended to have higher ability levels, spending extra time on some items lead to a decrease in the probability of giving a correct response for many items in this low-stake assessment.

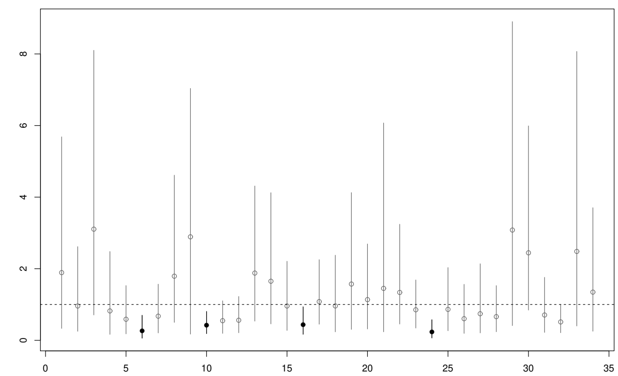
Table 8: Fit statistics (DIC) of the three conditional dependency models and the conditional independence model in Applications 1 and 2

Model	Application 1	Application 2
Within-item conditional dependency model		
Raw value	48568.68	148292.6
Expected value	49432.44	149099.1
Residual value	48495.81	147855.7
Conditional independence	49587.05	149197.7

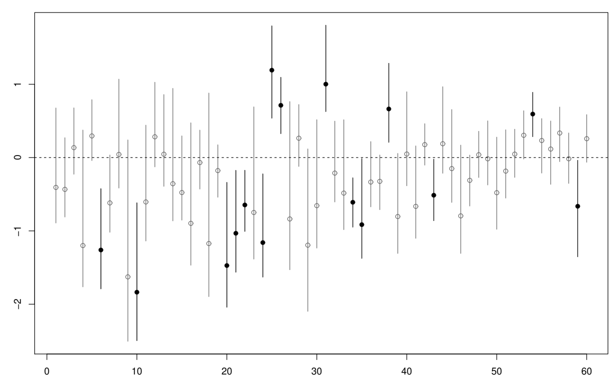
(a) Application 1: $v_i^{(A|B)}$



(b) Application 1: $w_i^{(A|B)}$



(c) Application 2: $v_i^{(A|B)}$



(d) Application 2: $w_i^{(A|B)}$

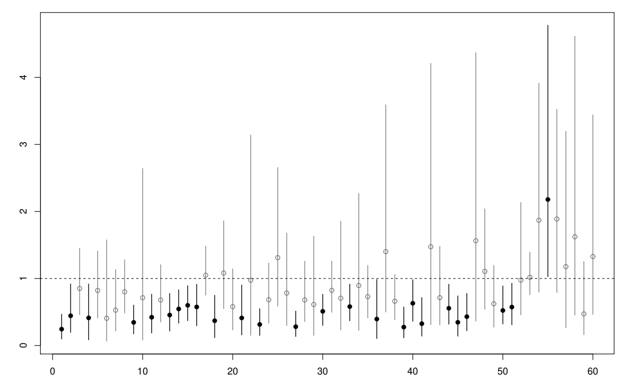


Figure 5: Residual value conditional dependency model: Posterior means (dots) and 95% credible intervals (vertical lines) for the estimated within-item conditional dependency parameters. The cases whose intervals do no include 0 (left panel) and 1 (right panel) are marked with thicker lines and shaded dots. The first row is obtained from the residual value conditional dependency model applied to Application 1 and the second row is the results from the residual value conditional dependency model applied to Application 2.

References

Bates, D., & Maechler, M. (2009). *lme4: Linear mixed-effects models using S4 classes R package version 0.999375-31.* (Downloadable from http://CRAN.Rproject.org/package=lme4)

DiTrapani, J., Jeon, M., De Boeck, P., & Partchev, I. (2016). Attempting to differentiate fast and slow intelligence: Using generalized item response trees to examine the role of speed on intelligence tests. *Intelligence*, *56*, 82-92.

Dodell-Feder, D., Koster-Hale, J., Bedny, M., & Saxe, R. (2011). fMRI item analysis in a theory of mind task. *Neuroimage*, *55*, 705-712.

Jeon, M., & De Boeck, P. (2019). An analysis of an item response strategy based on knowledge retrieval. *Behavior Research Methods*, *51*, 697-719.