

Hierarchical DDM on Relational Task: Final report

Adriana F. Chávez De la Peña

11 June, 2022

1 Introduction

Data for **87 participants** in the pilot of the Human Connectome Project doing the **relational matching task** is analyzed. This task (Smith, Keramatian, & Christoff, 2007) is known to be effective at locating the rostral-lateral prefrontal cortex (RLPFC) by comparing the performance observed across two conditions:

- **Relational matching condition:** Participants are presented with a sample pair of stimuli (Fig 1a, upper half) that differ in one of two possible dimensions (shape or texture) and a test pair of stimuli that could also be different from each other in any of these two dimensions (Fig 1a, bottom half). **Instructions:** Participants have to indicate whether the elements in the test pair differ in the same dimension than they do in the sample pair.
- **Control condition:** Participants are presented with two sample stimuli (Fig 1b, upper half) a test stimuli (Fig 1b, bottom half), and a trial-by-trial indicator of the dimensions they are required to pay to (shape or texture, again). **Instructions:** Participants have to indicate whether the test stimuli matches ANY of the sample stimuli in the signaled dimension.

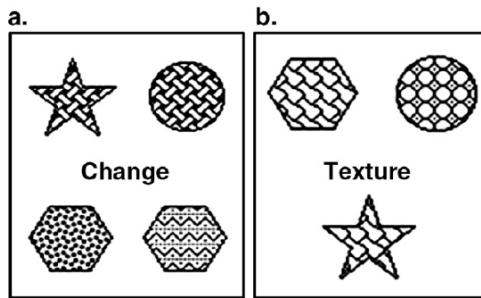


Figure 1: The relational matching task: a) relational matching condition; b) control condition.

Highlights of the experimental design

- Both, the shape and texture dimensions had 6 levels each.
- The task was conducted in blocks (3 per condition)
 - The control condition was run in blocks of 5 trials
 - The relational matching condition was run in blocks of 4 trials
- Instructions were flashed at the beginning of every block to signal the condition: 1) “Match Change” for the Relational matching task, and 2) “Match Shape” and 3) “Match Texture” for the control condition
- The stimuli remained on display until a response was made or a maximum of 3500 ms (for the relational condition) or 2800 ms (for the control condition) had elapsed.

2 Data Description

```
##   trial  subID      cond      dim resp corr acc   rt
## 1     1 100307 relational texture  0    0    1 2548
## 2     2 100307 relational texture  1    1    1 2528
## 3     3 100307 relational shape   0    1    0 1690
## 4     4 100307 relational texture  0    0    1 2438
```

Variables contained in the data set

- *cond*: Task (relational/control)
- *dim*: Dimension signaled during the trial (texture/shape)
- *resp*: Binary response given by participant
- *corr*: Binary correct response
- *acc*: Accuracy (1 when *resp==corr*)
- *rt*: Response time (measured in milliseconds)

2.1 Details about the data to keep in mind for modeling

Note for the reader: This report was done in Rmarkdown. Every value reported from this point onward is being directly printed from R.

- **Missing trials?** We have the same number of observations per participant. All participants completed 27 trials.
- **Dimension frequency.**
 - In the control condition, participants were asked to match the texture on 53% of the trials.
 - On the relational condition, 2/3 of the trials contain sample pairs that differed on their Texture.

```
##           shape texture
## control      575     655
## relational   330     654
```

- **Dimension frequency across participants.** Every participant saw the same number of Shape/Texture trials per condition.

```
##           shape texture
## control      7       8
## relational   4       8
```

- **Proportion of Match/Mismatch trials**

- On the control condition, about 53% of the trials matched.
- On the relational condition, about 42% of the trials matched.

```
##           0   1
## control    574 656
## relational 573 411
```

- **Proportion of Match/Mismatch trials across participants** The exact same proportion od matched trials can be found at the level of every participant in the study.

```
##          0 1
## control    7 8
## relational 7 5
```

- **Proportion of responses observed per condition:** The differences in the proportion of Match/Mismatch trials included per condition is somewhat consistent with the proportion of responses observed across all participants.

- On the control condition, a “Match!” response was found about 48% of the time.
- On the relational condition, a “Match!” response was observed on 36% of the trials.

```
##          0 1
## control   637 593
## relational 628 356
```

- **SDT contingency matrices (across all participants):**

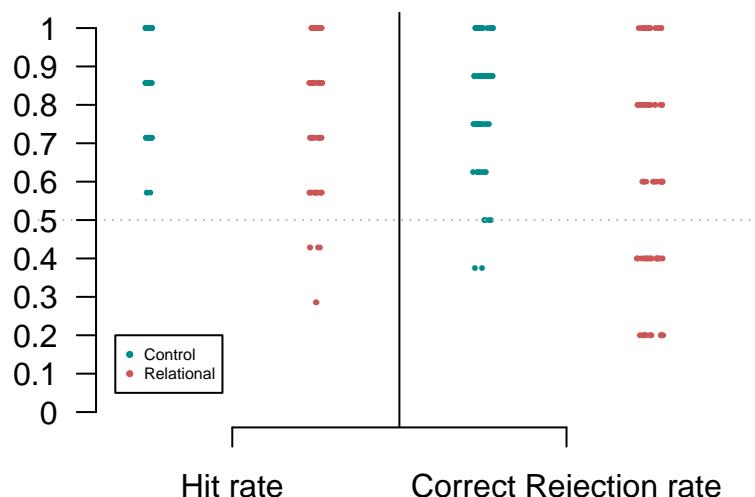
On the **control condition**, a high proportion of Hits (91.8%) and an above-chance proportion of Correct rejections (83.3%) were found.

```
##           Match stimulus Mismatch stimulus
## Match response      0.91986063      0.1661585
## Mismatch response    0.08013937      0.8338415
```

However, on the **relational condition**, a lower proportion of Hits (80.4%) is observed, with the Correct rejection rate being barely above chance (59.3%).

```
##           Match stimulus Mismatch stimulus
## Match response      0.8045375       0.406326
## Mismatch response    0.1954625       0.593674
```

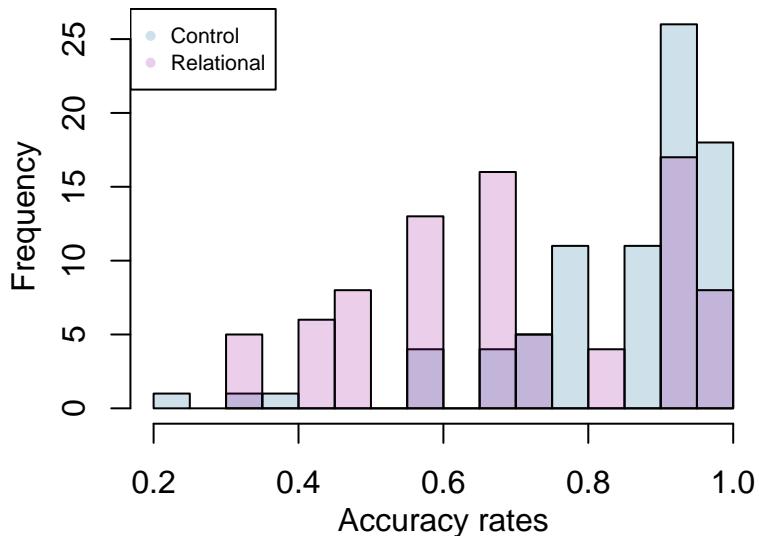
- **Hit and Correct Rejection rates per participant across condition**



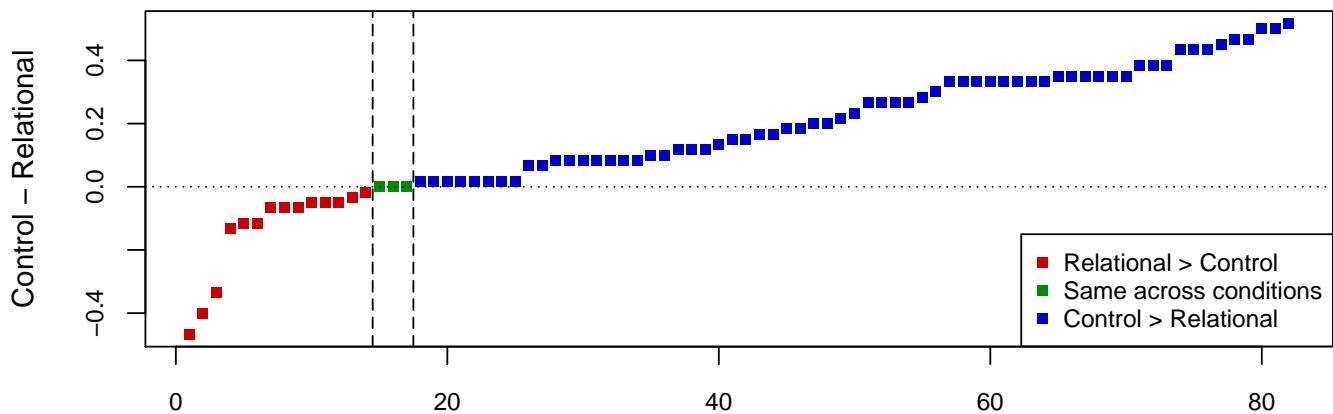
- Condition-blocks were presented in the same order across all participants, except for one of them (*i.e.* ID-735148).

```
##          1  2  3  4  5  6  7  8  9 10 11 12 13 14 15 16 17 18 19 20 21 22
## Control: NA NA NA NA 82 82 82 82 82 1  1  1  1  82 81 81 81 81 81 81 81
## Relational: 82 82 82 82 NA NA NA NA NA 81 81 81 81 NA 1  1  1  1  1  1  1
##          23 24 25 26 27
## Control: 82  1  1  1  1
## Relational: NA 81 81 81 81
```

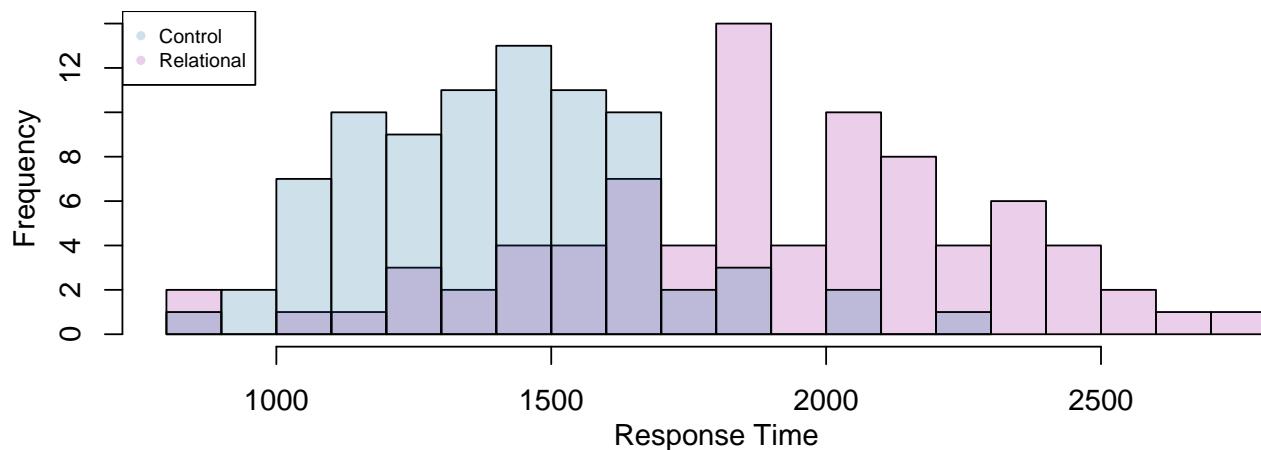
- Individual accuracy rate across condition



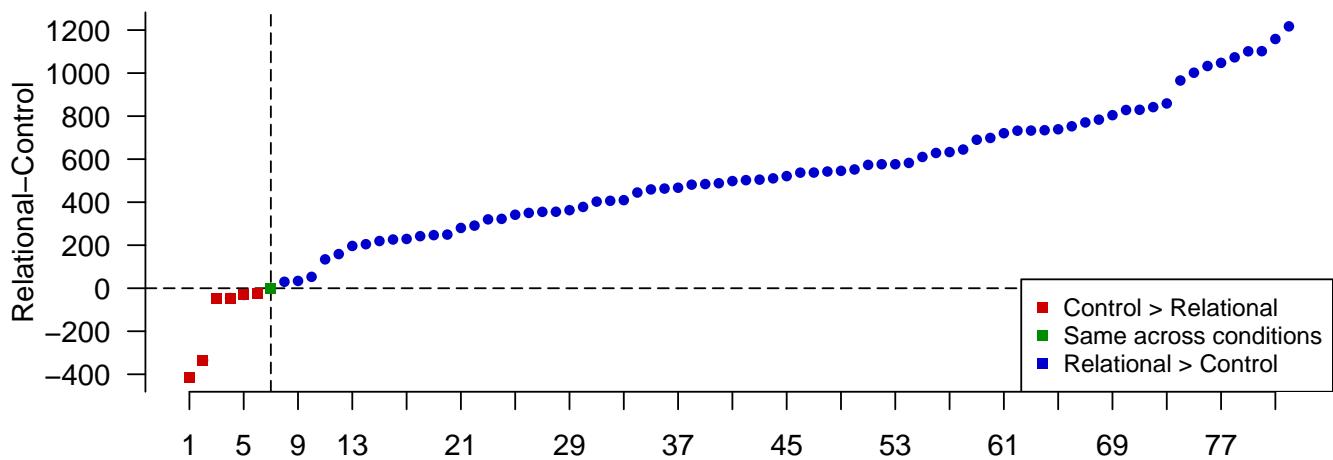
- Difference in the accuracy rates observed across conditions, for each participant.



- Individual mean RT observed across conditions.



- Difference in the accuracy rates observed across conditions, for each participant.



3 The Drift Diffusion Model

The Drift Diffusion Model (DDM) accounts for responses observed in 2-alternative forced choice tasks (2AFC) as the result of a sequential sampling process, characterized by four parameters (See Figure 2). The core assumption is that agents make these binary decisions by accumulating information about the stimulus presented until a response threshold is met.

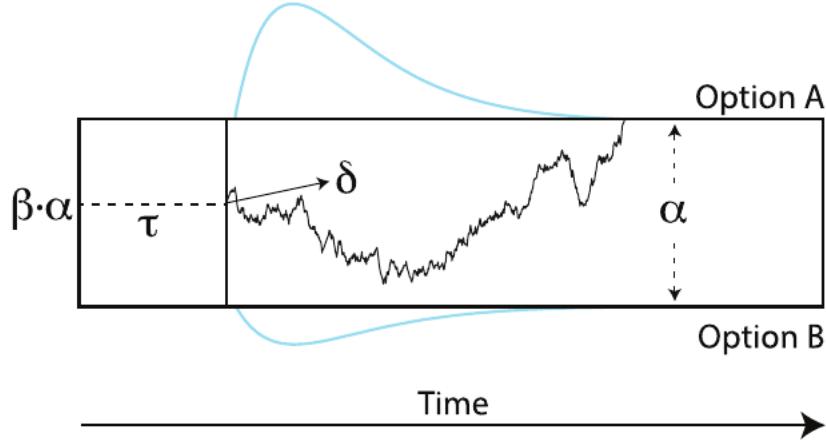


Figure 2: A graphical representation of the accumulation process described by the Drift Diffusion Model. This figure was taken from Johnsen et al. (2017).

The four parameters used in the DDM to describe this sequential sampling process are described below:

1. **Boundary separation (α)**. The distance between the upper and lower threshold (hence, restricted to positive values). This parameter is assumed to capture information about the threshold-accuracy trade-off elicited by the task design. Larger values of α imply that the agent will need to accumulate a greater amount of information before making a judgment, and thus, responses will be slower. This parameter has been found to be sensitive to task manipulations that incentivize either more accurate vs rapid responses.
 - $0 < \alpha$
2. **Bias (β)**. The agent's preference towards the binary choice associated with the upper boundary. It is defined as a value between 0 and 1, that gets multiplied by α to determine the initial state of the participant. For example, in an extreme case where $\beta = 1$, the agent would start their decision process already at the upper boundary (hence, ignoring any information provided by the stimulus and emitting their response right away).
 - $0 < \beta < 1$
 - Starting point = $\beta\alpha$
3. **Non-decision response time (τ)** From the time that elapses between the onset of the stimulus and the response entry, this parameter captures the time that could be attributed to anything else than the decision process. For example, it is assumed to capture the time it takes to produce the motor response, or the time it takes the brain to “read” or “encode” the stimuli they are presented with.
 - $0 < \tau$

4. **Drift rate (δ)**. This is probably the most important parameter in the model. It captures the mean rate of information acquired per unit of time. It represents the speed of the accumulation process that dictates which response threshold is met first. A peculiarity of this parameter, is that it is often assumed to vary across trials:

$$\delta_i \sim \text{Normal}(\mu_\delta, \sigma_\delta)$$

Figure 2 should also serve to illustrate one key feature of this model: it allows us to account for both, the choice observed (i.e. the response threshold met after accumulating information) and the Response Time distributions observed across trials (i.e. how long does it take for the system to reach any of these thresholds, according to the variable drift rate).

Even more, under the DDM framework, one can track back the bivariate probability of observing a specific choice at a given RT with the pdf of the Wiener distribution, which can be defined from the four parameters discussed above. So that:

$$Y = [\text{Choice}, \text{RT}]$$

$$Y \sim \text{Wiener}(\alpha, \beta, \tau, \delta)$$

The Wiener distribution was made available in JAGS by Wabersich & Vandekerckhove (2014)

4 JAGS models

4.1 Model 0: Demo

A simple Wiener process is used to model all the observations Y .

Step 1. We aggregate the information we gathered about the RTs and Choices observed into a single vector Y according to the following rule:

$$Y_i = \begin{cases} +(RT_i) & \text{if resp = "match"} \\ -(RT_i) & \text{if resp = "mismatch"} \end{cases} \quad (1)$$

```
datos <- read.csv("../00_RawMaterial/datos.csv")
datos <- datos[-which(datos$rt==0),]

N <- nrow(datos)
RT <- datos$rt/1000
C <- datos$resp

Y<-RT
Y[C==0] <- -RT[C==0]
```

Step 2. Write the JAGS model with uninformative priors.

```
write('
  model {
    for (i in 1:N) {
      Y[i] ~ dwiener(alpha,tau,beta,delta)
    }
    alpha ~ dunif(-5,5)
    tau ~ dunif(0,1)
    beta ~ dunif(0,1)
    delta ~ dunif(-5,5)
  }', 'wiener00.txt')
```

Step 3. Load variables, model specifications and start sampling with JAGS.

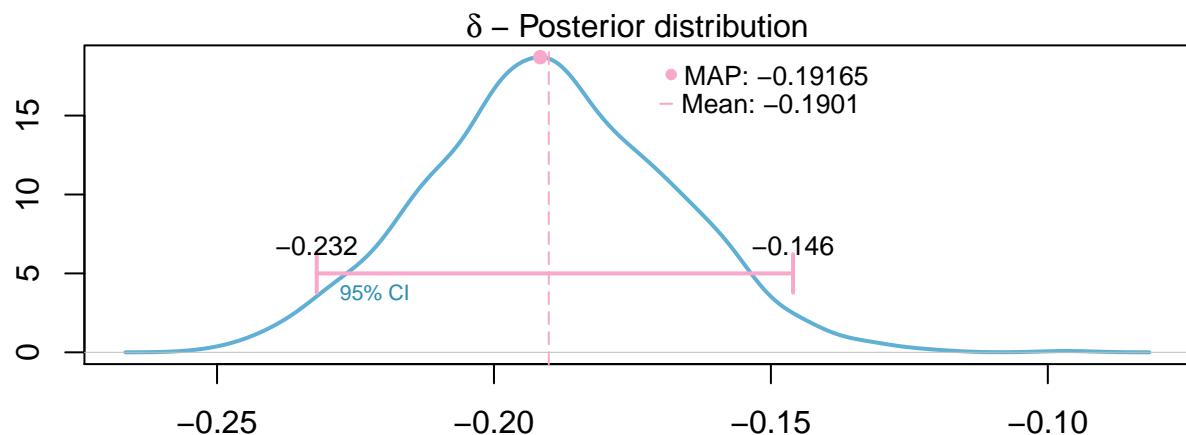
```
dat <- list(Y=Y, N=N)
inits <- list(alpha=1,tau=0.001,beta=0.5,delta=0)

model <- jags.model("./wiener00.txt",data=dat,inits=inits,n.chains=1,n.adapt=0)
samples <- jags.samples(model,c("alpha","tau","beta","delta"),n.iter=1000,thin=1)
```

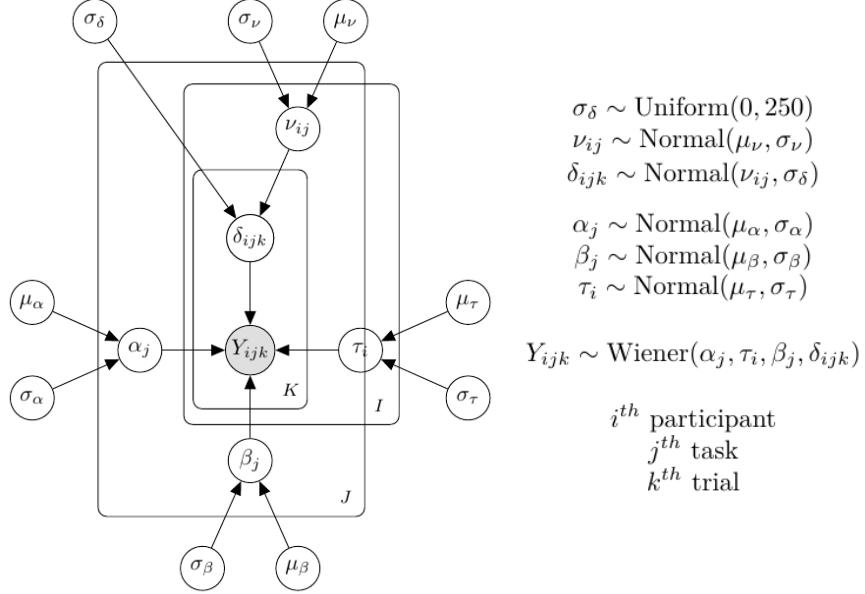
Step 4. Extract sample chains per parameter

```
alpha <- samples$alpha[,]
beta <- samples$beta[,]
delta <- samples$delta[,]
tau <- samples$tau[,]
```

Step 5. Make some informative plots!



4.2 Model 1: Hierarchical Drift Diffusion Model



For this model we used mostly uninformative priors on the parameters characterizing the population-level distributions, (see JAGS code). These priors were set up to take into consideration the constraints of each parameter. The core assumptions are:

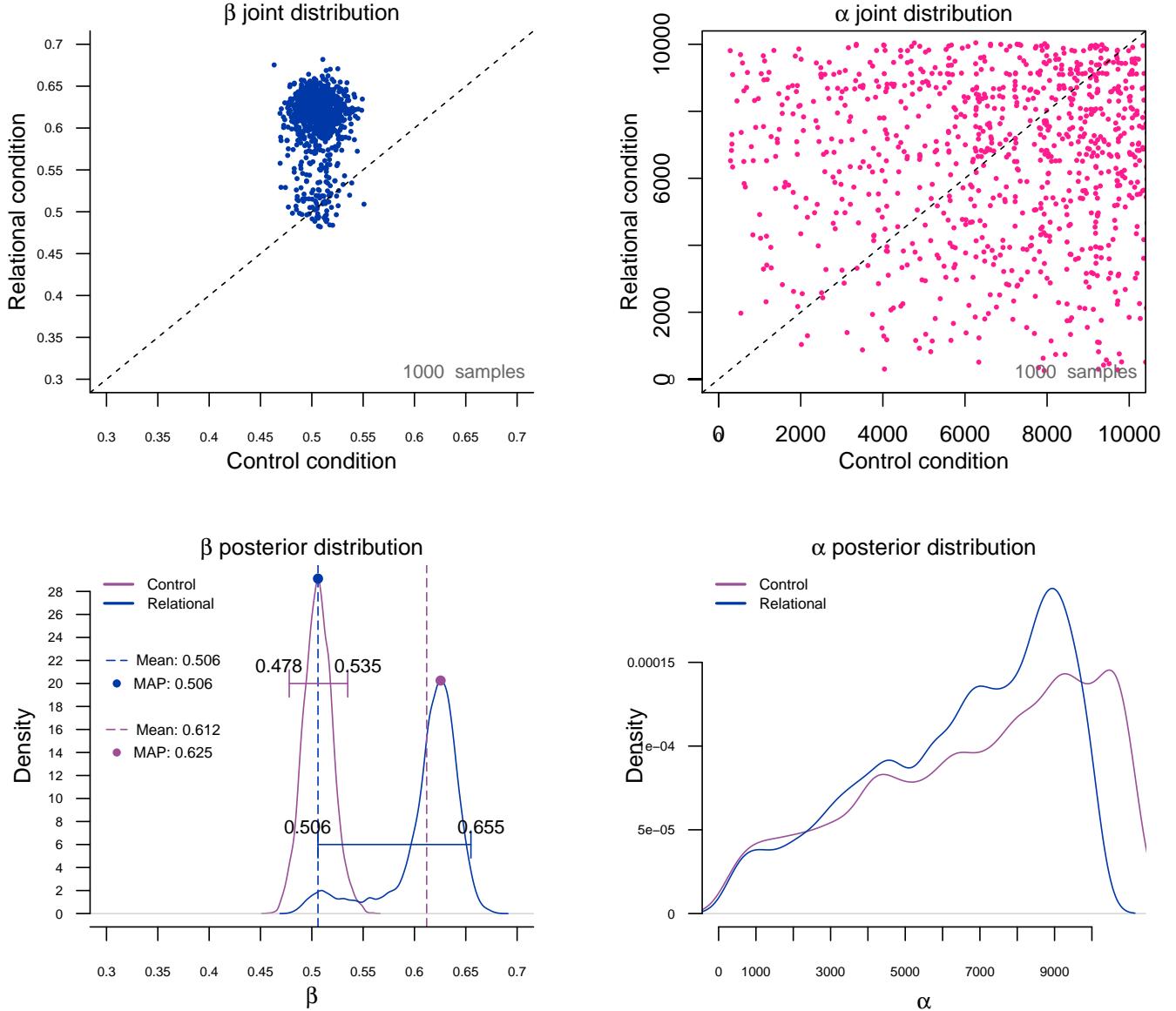
- Bias is task dependent, but it's assumed to be the same across participants (β_j)**. There's an uneven amount of Match/Mismatch trials across tasks which *could* explain a task-dependent preference for one of these judgements. But there's not a reason to assume that different participants would come into the experiment with a personal preference towards any of these judgments.
- Boundary separation is task dependent, but it's assumed to be the same across participants (α_j)**. The relational condition requires a higher-order comparison. In this condition, participants are not making a judgment about features, but rather about the relationship across these features. Boundary separation is usually used to capture information about the speed-accuracy trade-off, which *could* differ by task in this case as the maximum amount of time set-up per trial varies across conditions as well.
- Non-decision time varies per participant, and it's assumed to be the same across tasks (τ_i)**. The non-decision time captures the time that elapses due to any process (i.e. motor response and stimulus encoding) other than the actual time it takes to acquire enough information to make a judgment. The stimuli presented across task is virtually identical, it's just the task instructions that change. We don't expect a difference in non-decision RTs across tasks, but we may expect individual differences across participants (i.e. some individuals may be quicker to press the button than others).
- There's trial-by-trial variations in the drift rate (δ_{ijk})**. This is a classical assumption when using the Drift Diffusion Model, so that the drift rate varies across trials by being sampled from a Normal distribution with some μ_δ and a standard deviation typically fixed at $\sigma_\delta = 0.1$, (Ratcliff & Rouder, 1998). In this model, however, σ_δ is assumed to be unknown, but fixed across conditions and participants.
- We could expect some differences in the mean drift rate across tasks and participants (ν_{ij})**. This is arguably the most important parameter in the model. As described in Section 1, the main reason why these two conditions are presented is because through the comparison of the brain activity observed across them, experimenters are able to locate a specific region of the brain. The higher-order judgment required by the relational task is assumed to activate selective brain regions that are not compromised by the control task.

JAGS code

```
"model {  
    ##### Trial-by-trial variation  
    for (i in 1:N) {  
        Y[i] ~ dwiener(alpha[task[i]],tau[subj[i]],beta[task[i]],delta[i])  
        delta[i] ~ dnorm(nu[subj[i],task[i]],sigmaDelta)  
    }  
    ##### Task-by-Participant mean drift rate  
    for(i in 1:I){  
        for(j in 1:J){  
            nu[i,j] ~ dnorm(muNu,precNu)  
        }  
    }  
    ##### Participant-dependent parameters  
    for(i in 1:I){  
        tau[i] ~ dnorm(muTau,precTau)  
    }  
    ##### Task-dependent parameters  
    for(j in 1:J){  
        alpha[j] ~ dnorm(muAlpha,precAlpha)  
        beta[j] ~ dnorm(muBeta,precBeta)  
    }  
  
    ##### PRIORS  
    muBeta ~ dunif(0,1)  
    muAlpha ~ dunif(0,500)  
    muTau ~ dunif(0,2000)  
    muNu ~ dnorm(0,300)  
  
    sigmaBeta ~ dunif(0.0001,1000)  
    sigmaAlpha ~ dunif(0.0001,1000)  
    sigmaTau ~ dunif(0.0001,1000)  
    sigmaNu ~ dnorm(0.0001,1000)  
    sigmaDelta ~ dnorm(0.0001,1000)  
  
    precBeta <- pow(sigmaBeta,-2)  
    precAlpha <- pow(sigmaAlpha,-2)  
    precTau <- pow(sigmaTau,-2)  
    precNu <- pow(sigmaNu,-2)  
}  
}"
```

```
load("../02_Code/Models/1_Hierarchical/Hierarchical_Base/samples.RData")
```

Task-dependent parameters



Bias per task differs across task (Left panels)

- From the joint posteriors (top-left panel), we can tell that the majority of the random samples plotted show that β is **larger in the Relational condition than in the Control condition**. This same pattern can be found in the marginal posterior distributions shown in the bottom-left panel.
- According to the posterior distributions presented in the bottom-left panel: 1) There doesn't seem to be a bias towards any binary choice in the control condition (with both the MAP and mean of β_{control} falling at 0.5); 2) A slight bias towards the "Match!" response can be found in the relational condition (with a MAP and mean falling close to $\beta_{\text{relational}} = 0.62$)

Boundary distance is... a mess! (Right panels)

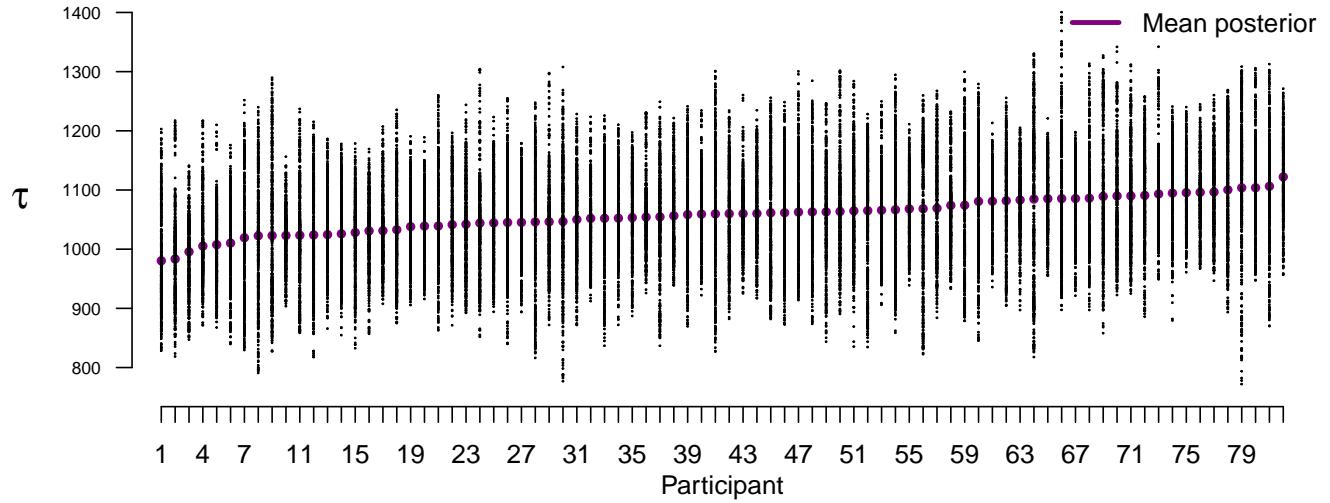
-The interpretation of the α_j parameters is far less straightforward, as the values sampled seemed to be all over the place.

- Both the joint distributions presented in the top-right panel, and the posterior distributions drawn in the bottom-right panel, seem to suggest that the values of α_j extend towards larger values. This may

suggest a problem with the values sampled from the posterior, but changing the priors to change the parameter space did not change the shape nor the “definition” of the posteriors here presented.

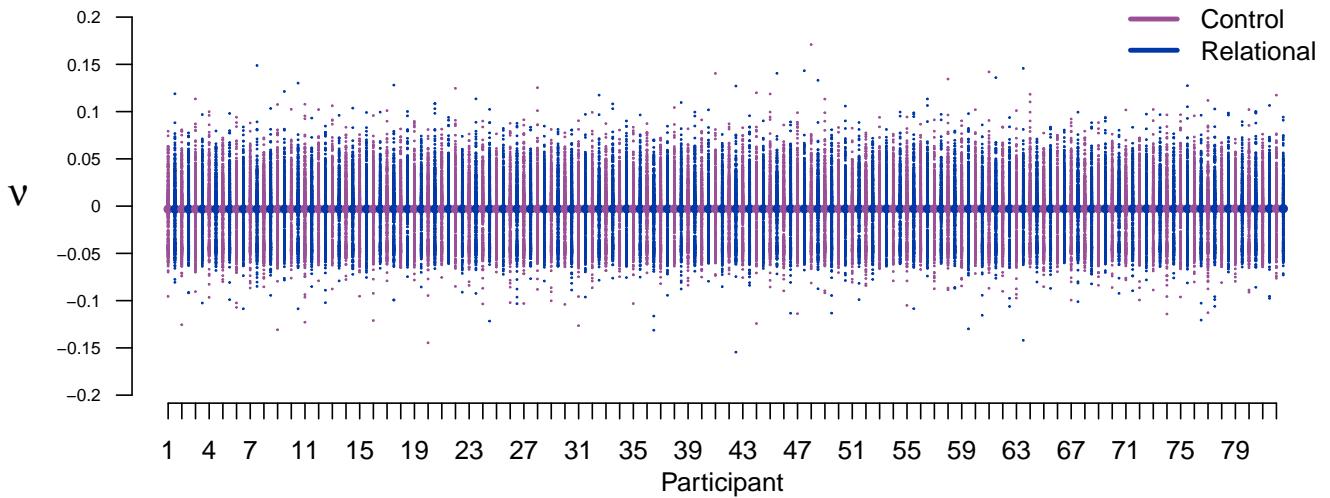
- These ugly priors motivated an extension of this first model, where there’s a single α parameter across conditions and participants.

Participant-dependent parameters



- There seems to be some individual differences in τ_i across participants, although not much. This shouldn’t be necessarily surprising, as we have only 27 observations per participant.

Drift rate

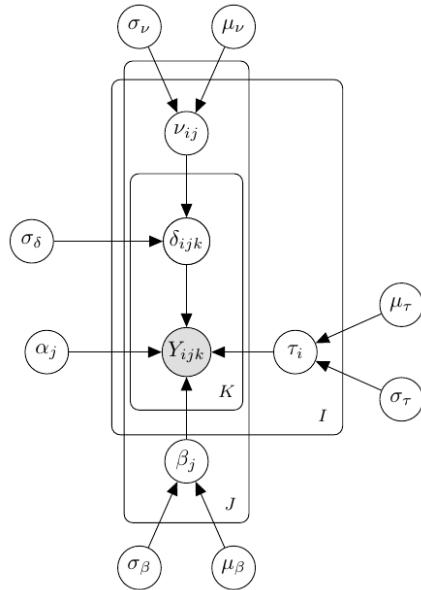


- With only 15/12 trials for the Control/Relational conditions per participant, there is really not a lot of room to estimate mean drift rates across these two levels of variability. However, I am still surprised by how little the variability observed is!

Perhaps we can throw more light into it with our next model...

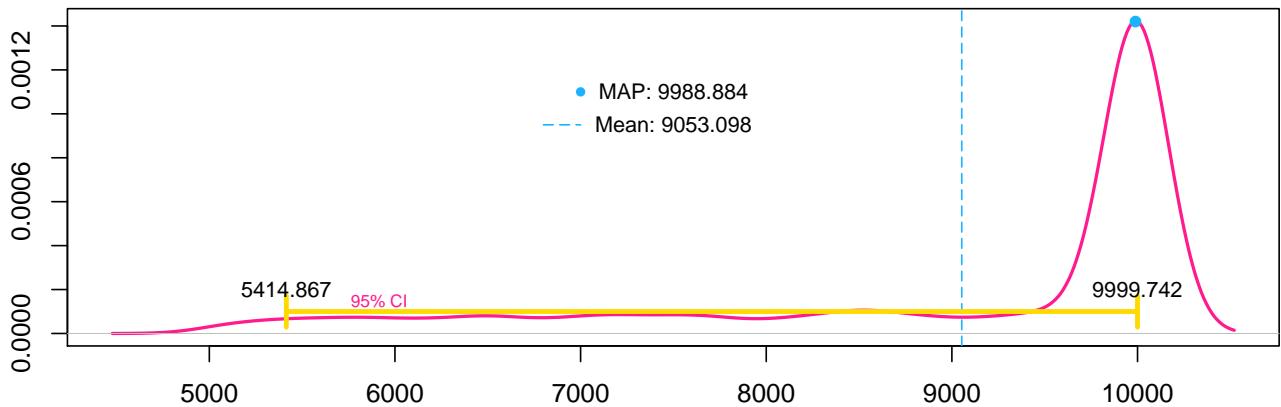
4.3 Model 2: Single alpha.

In this model, the main change is that now α is assumed to be shared across participants and conditions.



```
load("../02_Code/Models/1_Hierarchical/Hierarchical_SingleAlpha/samples.RData")
```

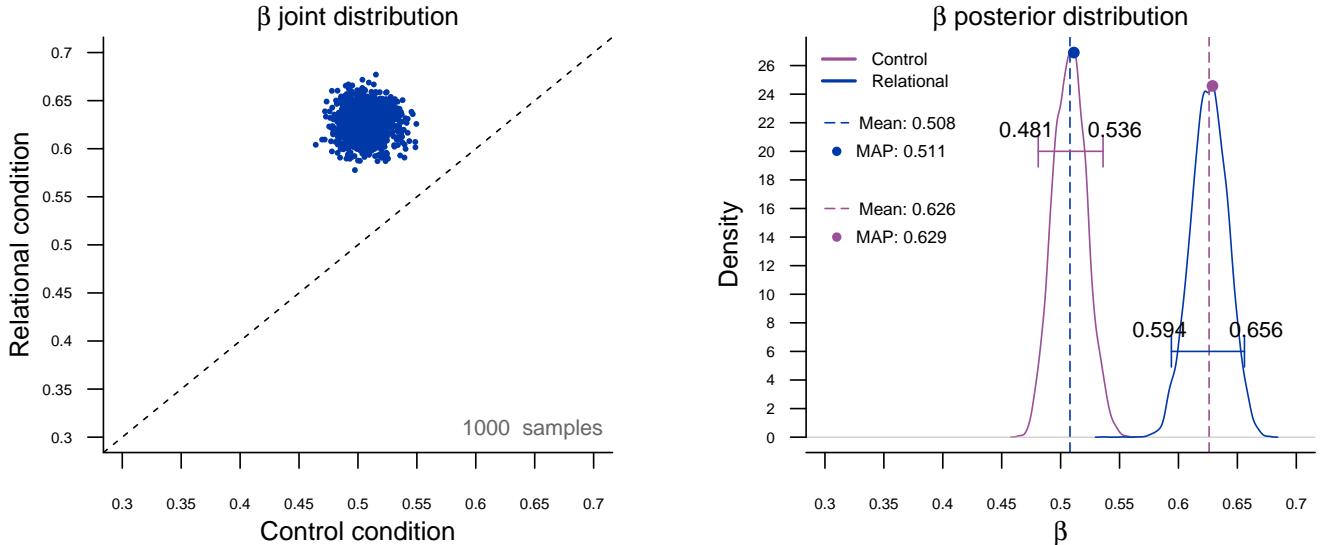
Single α across participants and conditions



Boundary distance is... looking better (?)

- Similar to Model 1, we can see that: 1) The 95% CI covers a wide range of values; 2) The MAP and the mean posterior seem to be much further away than with any other parameter.
- However, if we compare this posterior distribution with those observed in Model 1, we can still note an improvement: There's a clear trend. We can now distinguish easily which values have a higher posterior density.

Task-dependent parameter

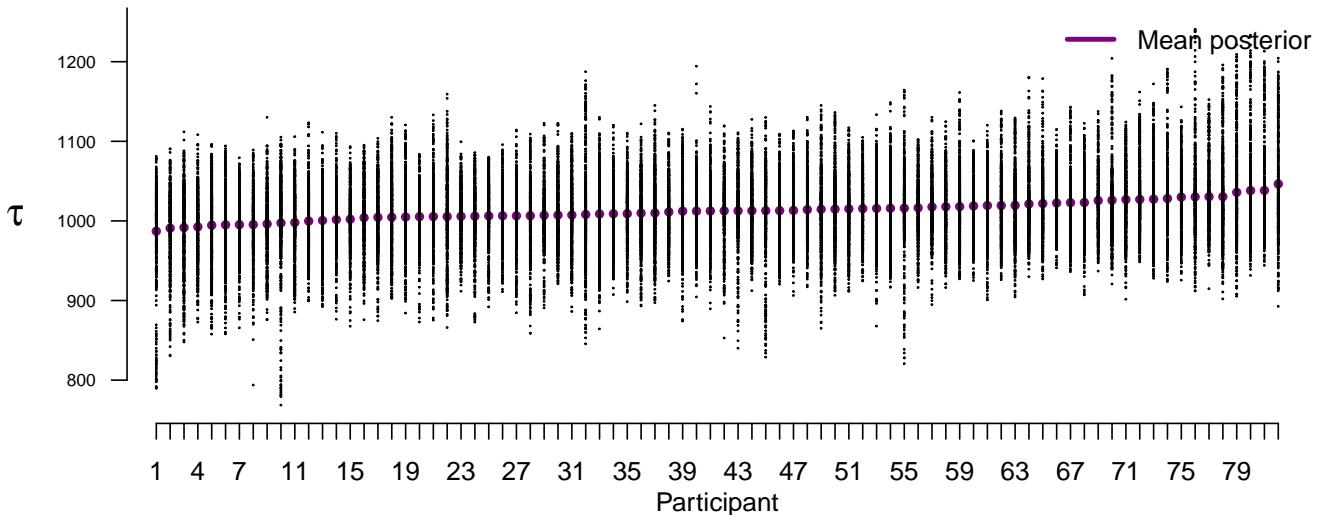


Bias per task differs across task

Overall, the posterior samples for the β_j parameters seem to be consistent with what we had reported on Model 1. If anything, the only difference is that the evidence for these claims is a little bit stronger this time:

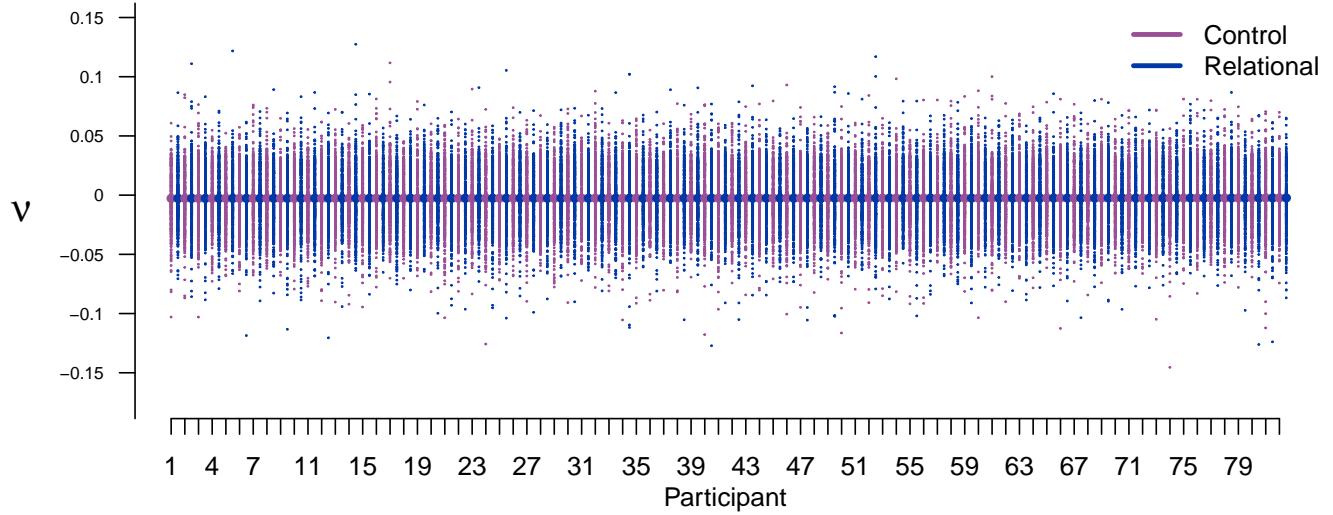
- β is larger in the Relational condition than in the Control condition
- There doesn't seem to be a bias towards any binary choice in the control condition
- A slight bias towards the “Match!” response can be found in the relational condition.

Participant-dependent parameters



- These results are consistent with those found in Model 1.

Drift rate



Oh no! Once again, there doesn't seem to be a clear structure in the mean drift rate observed across conditions and participants.

I've been puzzled by this last result. A few ideas that I have, include:

1. It seems odd that the posteriors of τ have such big values (specially, considering the RT distributions presented in Section 1). It doesn't make sense, to me, that the majority of the RT is attributed to non-decision processes. A first approach would be to try different, more restricted priors on μ_τ and σ_τ
2. Use the information gained from these models to restrict the priors on β_j , α , τ_i and μ_{nu}

References

- Johnson, D. J., Hopwood, C. J., Cesario, J., & Pleskac, T. J. (2017). Advancing Research on Cognitive Processes in Social and Personality Psychology: A Hierarchical Drift Diffusion Model Primer. *Social Psychological and Personality Science*, 8(4), 413–423. doi:10.1177/1948550617703174
- Ratcliff, R., & Rouder, J. N. (1998). Modeling Response Times for Two-Choice Decisions. *Psychol Sci*, 9(5), 347–356. doi:10.1111/1467-9280.00067
- Smith, R., Keramatian, K., & Christoff, K. (2007). Localizing the rostral-lateral prefrontal cortex at the individual level. *NeuroImage*, 36(4), 1387–1396. doi:10.1016/j.neuroimage.2007.04.032
- Wabersich, D., & Vandekerckhove, J. (2014). Extending JAGS: A tutorial on adding custom distributions to JAGS (with a diffusion model example). *Behav Res*, 46(1), 15–28. doi:10.3758/s13428-013-0369-3