## Supplemental material for

## "A Bayesian Race Model for Recognition Memory"

This supplement includes further details about the experiment, the data, the priors, and the Markov chain Monte Carlo (MCMC) algorithm used to fit the Bayesian models in the article "A Bayesian Race Model for Recognition Memory." All data and software can be downloaded from http://www.stat.osu.edu/~pfc/software/.

## 1 Details of the recognition memory experiment

### 1.1 Subjects

All research subjects came from the pool of introductory psychology students at the Columbus campus of The Ohio State University. A total of 32 subjects participated in the experiment, 12 in a pilot study and 20 in the final study. All subjects were 18 years or older, proficient in English, possessed normal or corrected-to-normal vision, and received partial class credit for participation.

## 1.2 Stimuli and Design

The experiment was conducted using the OpenSesame software (Mathôt et al., 2012) running on a Linux workstation connected to a CRT monitor. Subjects responded by pressing the M and Z keys on a standard QWERTY computer keyboard from which all but the M and Z keys had been pulled. The M and Z keys were replaced with the N and O keytops to indicate the response assignments for "new" (N) and "old" (O) responses.

Figure 1 in the article shows samples of the pictures used as stimuli. Each picture was composed of three geometric elements formed by 1, 2 or 3 line segments connected at a single point with four possible orientations. The connecting point for each element could be placed in one of 25 interior locations in a  $7\times7$  grid within the image (constrained to avoid contact of the elements with each other and with the edge of the grid), resulting in 144 possible arrangements of the elements. This procedure produced a database of 3,456 (6

possible orders of three objects by 4 orientations by 144 arrangements) pictures from which the study and test stimuli were selected.

The pictures included in the study lists were randomly selected from the database. The pictures in each list were inspected after selection to ensure that the pictures within lists were sufficiently dissimilar in terms of the locations and the directions of the three constituent elements. The color of the elements in each list were the same for each study/test trial (red, blue or green), but were different across lists.

#### 1.2.1 Design

The experimental session consisted of 21 study/test blocks.

Each study list was composed of eight unique pictures selected at random from the database. Four of the eight pictures were randomly chosen to begin and end the study list, two at the beginning of the list and two at the end to prevent repeated pictures from appearing too close to the beginning or end of the study list. The remaining four pictures were presented 1, 2, 3, or 4 times, for a total of 14 exposures of pictures during the study phase.

The test list was constructed from the eight studied pictures plus an additional eight pictures selected at random from the database and inspected for dissimilarity from each other and from the studied pictures. This gave a total of 16 pictures in the test list. The 16 test pictures were presented in a random order. The randomizations of both the study list and the test list was done once to generate the study/test trials in the experiment and each subject saw the same study/test lists.

The subjects responded "old" or "new" by pressing the keys labeled O and N with the index fingers of their left and right hands. Ten subjects responded "old" with their right hands and the remaining 11 responded "old" with their left hands. Over 21 study/test trials with 16 responses per trial, each subject made 336 responses: 168 responses to new pictures, 105 responses to pictures presented once, and 21 responses to pictures presented 2, 3 or 4 times.

#### 1.2.2 Procedure

After the informed consent process, we gave the subjects instructions concerning the procedures for the recognition task they were to perform. Each subject was tested individually and observed during the task by a research assistant through a window from an adjoining room.

Subjects read the task instructions on the monitor while the instructions were simultaneously read aloud by the research assistant, who answered the subjects' questions and ensured that the subjects understood what they were to do. The instructions were as follows:

In this experiment, you will complete a recognition memory task. You will view a series of simple line pictures that you will need to study and memorize. You will then view another set of pictures, some of which are ones you previously viewed, and some of which are new pictures which you will not have seen before. In the study series, you will view each picture for a set time. Try to memorize each picture to the best of your ability. You may see some pictures more than once. In the test series, you will view another set of pictures. For each picture you see, press the 'O' key if the picture is one you previously studied, and the 'N' key if the picture is new.

During each study phase, the word "Study" was visible at the top center of the screen, and during each test phase, the word "Test" was visible in the same location. Following the instructions, subjects saw the prompt "Study List" in the center of the screen and initiated the first trial by pressing a key. Immediately after the keypress, each picture in the study list was presented for 3s, followed by a blank screen of 1s. After the last picture in the study list was presented, a blank screen of 1s was presented, followed by a 3s warning that the test phase was about to begin, indicated by the prompt "Trial N: Test series," where N indicated the trial (1-21), which appeared in the middle of the screen. Subjects also saw the reminder, "Press O for previously studied pictures and N for new pictures," which appeared below the warning.

The presentation of the first test picture was not initiated by the subject, but began 1s after the end of the warning prompt. A response to a test picture was not recorded until 200ms after the onset of the picture, and the picture presentation was terminated when a response was recorded. (The 200ms delay prevented subjects holding down a key to escape from the experiment without having to perform the task. This delay was not implemented for the pilot subjects.) The next picture was presented 50ms after the response.

After completing the task, subjects were asked to fill out a questionnaire to determine demographic information, handedness and well-being on the day of the experiment, as well as to collect information about what they believed the purpose of the experiment was and if they had used any particular strategies to perform the task.

## 2 The dataset

The data for the main experiment can be downloaded from http://www.stat.osu.edu/~pfc/software/ and are stored in the comma delimited text file main.csv contained in the folder Data. The data for the pilot experiment are stored in the comma delimited text file pilot.csv. Both files contain the same columns:

- 1. subject: subject ID  $(1, \ldots, 20)$  for the main experiment;  $1, \ldots, 12$  for the pilot).
- 2. block: the block number (0, ..., 20). We excluded the first block (block 0) from the analysis to ensure that the subjects had become familiar with the task.
- 3. exposures: the number of exposures. A value of -1 denotes the responses to the first two and the last two items on the study list (which were exposed only once) that were excluded to avoid primacy and recency effects. A value of 0 exposures indicates a new stimulus, and a value greater than 0 indicates an old stimulus.
- 4. response: The response the subject made: either the m key (new) or the z key (old).
- 5. RT: the response time in milliseconds.

```
\begin{array}{ll} (\mu_{\beta^{O} \text{old}}, \sigma^{2}_{\beta^{O} \text{old}}) \sim N\Gamma^{-1}(\log(400), 2, 20, 8) & (\mu_{T^{0}}, \sigma^{2}_{T^{0}}) \sim N\Gamma^{-1}(100, 1, 3, 20) \\ (\mu_{\beta^{N} \text{old}}, \sigma^{2}_{\beta^{N} \text{old}}) \sim N\Gamma^{-1}(\log(600), 2, 20, 8) & \mu_{G,k} \sim N(\log(50), 0.01^{2}) \\ (\mu_{\beta^{O} \text{new}}, \sigma^{2}_{\beta^{O} \text{new}}) \sim N\Gamma^{-1}(\log(600), 2, 20, 8) & \mu_{H,k} \sim N(\log(7000), 0.01^{2}) \\ (\mu_{\beta^{N} \text{new}}, \sigma^{2}_{\beta^{N} \text{new}}) \sim N\Gamma^{-1}(\log(400), 2, 20, 8) & \mathbf{q}_{jk} \sim \text{Dirichlet}(0.1, 1, 0.1) \\ (\mu_{\beta^{A} \text{old}}, \sigma^{2}_{\beta^{A} \text{old}}) \sim N\Gamma^{-1}(\log(600), 2, 20, 8) & s_{\alpha,k} \sim N(0, 10^{2}) \\ \log(\sigma_{\alpha,k}) \sim N(\log(3), 0.5^{2}) & s_{\alpha,k} \sim N(0, 10^{2}) \\ \log(\sigma_{\alpha,k}) \sim N(\log(0.3), 0.1^{2}) & \phi_{\alpha,k} \sim N_{[-1,1]}(0, 0.15^{2}) \end{array}
```

Table 1: Prior specifications

To read the data from the main experiment into R use

```
main <- read.csv("main.csv", header=TRUE)</pre>
```

To remove the first block and the buffer trials

```
main <- main[main$block!=0 & main$exposures!=-1,]</pre>
```

## 3 Prior distributions

We now describe the prior distributions used for the hyperparameters in our Bayesian models. The notation  $N\Gamma^{-1}(\mu_0, \nu, \alpha_0, \delta_0)$  employed below refers to the Normal-inverse-gamma distribution. This distribution is a conjugate prior distribution for normal likelihoods. The random variables  $(\mu, \sigma^2)$  are distributed as  $N\Gamma^{-1}(\mu_0, \nu, \alpha_0, \delta_0)$ , if

$$1/\sigma^2 \mid \alpha_0, \delta_0 \sim \Gamma(\alpha_0, \delta_0)$$
 and  $\mu \mid \sigma^2, \mu_0, \nu \sim N(\mu_0, \sigma^2/\nu)$ .

The specific priors that we employed are shown in Table 1. The hyperparameters of the prior distributions for  $p_l$  changed as we examined a number of different model structures for

the trace process and as we examined the sensitivity of the results to the choice of the prior for  $p_l$ .

Before examining the prior sensitivity our hyperparameters were as follows:

(a) In the base model, where all  $p_l$  may be different, we assumed that

$$(\mu_{p_l}, \sigma_{p_l}^2) \sim N\Gamma^{-1}(\text{logit}(0.5), 1/4, 3, 5), \quad l = 0, 1, 2, 3.$$

(b) When  $p_0 = 1$  we assumed

$$(\mu_{p_l}, \sigma_{p_l}^2) \sim N\Gamma^{-1}(\text{logit}(0.5), 1/4, 3, 5), \quad l = 1, 2, 3.$$

(c) When  $p_2 = p_3$  we assumed

$$(\mu_{p_l}, \sigma_{p_l}^2) \sim N\Gamma^{-1}(\text{logit}(0.5), 1/4, 3, 5), \quad l = 0, 1, 2.$$

In Section 5.1 of the article we compare the base model (a) with hyperparameters specified as above to the case where the prior mass for  $p_l$  is highly concentrated near 1. For this latter case the hyperparameter prior distributions were specified as

$$(\mu_{p_l}, \sigma_{p_l}^2) \sim N\Gamma^{-1}(\text{logit}(0.99), 1/4, 3, 5), \quad l = 0, 1, 2, 3.$$

For the Weibull model, the following hyperparameter prior distributions were used for those parameters that are not in common with the minimum gamma model:

$$\begin{aligned} &(\mu_{\eta} o_{\text{old}}, \sigma_{\eta}^2 o_{\text{old}}) &\sim N \Gamma^{-1}(\log(1), 2, 20, 8), \\ &(\mu_{\beta} o_{\text{old}}, \sigma_{\beta}^2 o_{\text{old}}) &\sim N \Gamma^{-1}(\log(1500), 2, 20, 8), \\ &(\mu_{\beta} N_{\text{old}}, \sigma_{\beta}^2 N_{\text{old}}) &\sim N \Gamma^{-1}(\log(2000), 2, 20, 8), \\ &(\mu_{\beta} o_{\text{new}}, \sigma_{\beta}^2 O_{\text{new}}) &\sim N \Gamma^{-1}(\log(2000), 2, 20, 8), \end{aligned}$$

and

$$(\mu_{\beta^{N_{\text{new}}}}, \sigma^2_{\beta^{N_{\text{new}}}}) \sim N\Gamma^{-1}(\log(1500), 2, 20, 8).$$

## 4 Markov chain Monte Carlo sampling procedure

We fit our model using an MCMC algorithm. Conditional on the hyperparameters, the posterior distribution of a subject's parameters is independent of the posterior distribution of the other subjects' parameters. Thus, we can sample the subject-specific parameters for each subject first and then sample the hyperparameters. Sampling of the subject-specific parameters could be parallelized to speed up computation. There are six steps in our sampling strategy; steps 2–4 use Metropolis-Hastings (MH) sampling. The six steps are as follows:

For each subject k:

1. For i = 1, ..., n, j = 1, ..., B, draw  $\xi_{ijk}$  independently from a discrete distribution on  $\{0, 1, 2\}$  with probabilities

$$\pi_{ijk} = \left( q_{G,jk} f_G(T_{ijk} \mid \mu_{G,k}, \sigma_{G,k}), \ q_{C,jk} f_{C,R}(T_{ijk} - T_k^0, R_{ijk} \mid T_k^0, E_{ijk}, \mathbf{p}_k, \boldsymbol{\alpha}_k, \boldsymbol{\beta}_k), q_{H,jk} f_H(T_{ijk} \mid \mu_{H,k}, \sigma_{H,k}) \right).$$

2. The posterior distribution of  $(\boldsymbol{\beta}_k, \boldsymbol{\alpha}_k, \mathbf{p}_k)$  conditional on all other parameters and the data is proportional to

$$f_{p}(\mathbf{p}_{k} \mid \boldsymbol{\mu}_{p}, \boldsymbol{\sigma}_{p}) \quad f_{\alpha}(\boldsymbol{\alpha}_{k} \mid \boldsymbol{\mu}_{\alpha,k}, \boldsymbol{\sigma}_{\alpha,k}, \boldsymbol{\phi}_{\alpha,k}) f_{\beta}(\boldsymbol{\beta}_{k} \mid \{\boldsymbol{\mu}_{\beta}^{c}\}, \{\boldsymbol{\sigma}_{\beta}^{c}\}) \quad \times$$

$$\prod_{j=1}^{B} \prod_{i \in \{\xi_{ijk}=1\}} f_{C,R}(T_{ijk} - T_{k}^{0}, R_{ijk} \mid T_{k}^{0}, E_{ijk}, \mathbf{p}_{k}, \boldsymbol{\alpha}_{k}, \boldsymbol{\beta}_{k}).$$

Draw  $(\boldsymbol{\beta}_k, \boldsymbol{\alpha}_k, \mathbf{p}_k)$  using an MH step with a t proposal tuned on the basis of some preliminary runs.

3. The posterior distribution of  $(\mu_{\alpha,k}, s_{\alpha,k}, \sigma_{\alpha,k}, \phi_{\alpha,k}, \mathbf{p}_k)$  conditional on all other parameters and the data is proportional to

$$f_{\alpha}(\boldsymbol{\alpha}_k \mid \mu_{\alpha,k}, \sigma_{\alpha,k}, \phi_{\alpha,k}).$$

Draw  $(\mu_{\alpha,k}, s_{\alpha,k}, \sigma_{\alpha,k}, \phi_{\alpha,k})$  using an MH step with a normal proposal.

4. The posterior distribution of  $T_k^0$  conditional on all other parameters and the data is proportional to

$$f_{T^{0}}\left(T_{k}^{0} \mid \mu_{T^{0},k}, \sigma_{T^{0},k}\right) \prod_{j=1}^{B} \left(\prod_{i \in \{\xi_{ijk}=2\}} f_{H}(T_{ijk} - T_{k}^{0} \mid T_{k}^{0}, \mu_{H,k}, \sigma_{H,k}) \times \prod_{i \in \{\xi_{ijk}=1\}} f_{C,R}(T_{ijk} - T_{k}^{0}, R_{ijk} \mid T_{k}^{0}, E_{ijk}, \mathbf{p}_{k}, \boldsymbol{\alpha}_{k}, \boldsymbol{\beta}_{k})\right).$$

Let  $t_{\min,k}$  be the smallest observed RT that the mixture process identifies as arising from the decision process. Draw  $T_k^0$  using an MH step with a scaled beta(6, 1) proposal supported on  $[0, t_{\min,k}]$ .

5. Draw  $\mu_{G,k}$  from its conditionally conjugate normal posterior

$$N\left(\frac{100^2\log(50) + \sum_{i,j\in\{\xi_{ijk}=0\}}\log(T_{ijk})}{100^2 + \sum_{i,j\in\{\xi_{ijk}=0\}}1}, \left[100^2 + \sum_{i,j\in\{\xi_{ijk}=0\}}1\right]^{-1}\right),$$

and draw  $\mu_{H,k}$  from its conditionally conjugate normal posterior

$$N\left(\frac{100^2\log(7000) + 2.5^2\sum_{i,j\in\{\xi_{ijk}=2\}}\log(T_{ijk} - T_k^0)}{100^2 + 2.5^2\sum_{i,j\in\{\xi_{ijk}=2\}}1}, \left[100^2 + 2.5^2\sum_{i,j\in\{\xi_{ijk}=2\}}1\right]^{-1}\right).$$

6. Draw the mixture parameters  $\mathbf{q}_{jk}$  from a Dirichlet $(n_{G,jk} + 0.1, n_{C,jk} + 1, n_{H,jk} + 0.1)$  distribution, where  $n_{G,jk} = \sum_i I(\xi_{ijk} = 0)$ ,  $n_{C,jk} = \sum_i I(\xi_{ijk} = 1)$ , and  $n_{H,jk} = \sum_i I(\xi_{ijk} = 2)$ .

Sampling of the hyperparameters is straightforward because of the conditional conjugacy of the prior distributions. Specifically,  $\{(\mu_{p_l}, \sigma_{p_l}^2)\}_{i=0,1,2,3}$ ,  $(\mu_{\beta^{N_{\text{old}}}}, \sigma_{\beta^{N_{\text{old}}}}^2)$ ,  $(\mu_{\beta^{O_{\text{new}}}}, \sigma_{\beta^{O_{\text{new}}}}^2)$ ,  $(\mu_{\beta^{O_{\text{new}}}}, \sigma_{\beta^{O_{\text{new}}}}^2)$ ,  $(\mu_{\beta^{N_{\text{new}}}}, \sigma_{\beta^{N_{\text{new}}}}^2)$ ,  $(\mu_{\beta^{N_{\text{new}}}}, \sigma_{\beta^{N_{\text{new}}}}^2)$ ,  $(\mu_{\beta^{N_{\text{new}}}}, \sigma_{\beta^{N_{\text{new}}}}^2)$ , and

 $(\mu_{T^0}, \sigma_{T^0}^2)$  are all drawn from Normal-inverse-gamma distributions with easily-determined parameters. The R, C, and C++ code for our MCMC algorithm is available at http://www.stat.osu.edu/~pfc/publications/.

# 5 Fitting the minimum gamma models and Weibull models using Markov chain Monte Carlo

This code runs on Mac OS and Linux without alteration, as long as the computer can install C and C++ code. (More work is needed to run the code on Windows: find all references to .so and replace them by .dll.)

All the R code needed to fit the model can be found in the R folder. Before running the shell script fit\_main\_models.sh in a terminal to fit the main models used in the paper, make sure that the following is done:

- 1. Install the R library Rcpp.
- 2. Install the R library truncatedNormals. which can be obtained from http://www.stat.osu.edu/~pfc/software/.
- 3. In the folder functions run the makefile to build the RT\_likelihood.so dynamic library.

The folder traces contains all the trace plots generated by the MCMC runs used to fit each Bayesian model, and the folder output contains the MCMC chains, stored in the R object chs:

```
objects(chs)
```

```
[1] "hyper" "log.alpha"
```

[3] "log.alpha.chain" "log.alpha.pars"

[5] "log.alpha.pars.chain" "log.alpha.pars.jumps"

[7] "log.alpha.pars.jumps.chain" "log.beta"

[9] "log.beta.chain" "log.beta.mu"

[11] "log.beta.mu.chain" "log.beta.sigma2"

[13] "log.beta.sigma2.chain" "logit.p"

```
[15] "logit.p.chain"
                                    "logit.p.mu"
[17] "logit.p.mu.chain"
                                    "logit.p.sigma2"
[19] "logit.p.sigma2.chain"
                                    "muG"
                                    "muH"
[21] "muG.chain"
[23] "muH.chain"
                                    "qvec"
                                    "TO"
[25] "qvec.chain"
[27] "T0.chain"
                                    "T0.jumps"
                                    "T0.mu"
[29] "TO.jumps.chain"
[31] "T0.mu.chain"
                                    "T0.sigma2"
[33] "T0.sigma2.chain"
                                    "the.data"
[35] "the.model"
                                    "theta.jumps"
[37] "theta.jumps.chain"
                                    "xi"
[39] "xi.chain"
```

Each chain (e.g., log.alpha.chain) is stored as a list, one list item for each iteration of the MCMC chain. For example, a summary of the posterior distribution of  $T_0$  for each subject can be obtained by executing:

To change the prior distributions modify the content of the function hyperparameters in the file fit\_models.R. For example, to fit a version of the base model (named MG4 in the R code) in which  $p_l$  is concentrated near one we change each occurrence of c(logit(0.5), 1/4, 3, 5) to c(logit(0.99), 1/4, 3, 5) in the else if (the.model=="MG4") clause of the hyperparameters function.

## References

Mathôt, S., Schreij, D., and Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods*, 44:314–324.