



# Release the BEESTS: Bayesian Estimation of Ex-Gaussian STop-Signal reaction time distributions

Dora Matzke<sup>1\*</sup>, Jonathon Love<sup>1</sup>, Thomas V. Wiecki<sup>2</sup>, Scott D. Brown<sup>3</sup>, Gordon D. Logan<sup>4</sup> and Eric-Jan Wagenmakers<sup>1</sup>

<sup>1</sup> Department of Psychological Methods, University of Amsterdam, Amsterdam, Netherlands

<sup>2</sup> Laboratory of Neural Computation and Cognition, Brown University, Providence, RI, USA

<sup>3</sup> School of Psychology, University of Newcastle, Callaghan, NSW, Australia

<sup>4</sup> Department of Psychology, Vanderbilt University, Nashville, TN, USA

## Edited by:

Holmes Finch, Ball State University, USA

## Reviewed by:

Richard S. John, University of Southern California, USA

Martin Lages, University of Glasgow, UK

## \*Correspondence:

Dora Matzke, Department of Psychology, University of Amsterdam, Weesperplein 4, 1018 XA, Amsterdam, Netherlands  
e-mail: d.matzke@uva.nl

The stop-signal paradigm is frequently used to study response inhibition. In this paradigm, participants perform a two-choice response time (RT) task where the primary task is occasionally interrupted by a stop-signal that prompts participants to withhold their response. The primary goal is to estimate the latency of the unobservable stop response (stop signal reaction time or SSRT). Recently, Matzke et al. (2013) have developed a Bayesian parametric approach (BPA) that allows for the estimation of the entire distribution of SSRTs. The BPA assumes that SSRTs are ex-Gaussian distributed and uses Markov chain Monte Carlo sampling to estimate the parameters of the SSRT distribution. Here we present an efficient and user-friendly software implementation of the BPA—BEESTS—that can be applied to individual as well as hierarchical stop-signal data. BEESTS comes with an easy-to-use graphical user interface and provides users with summary statistics of the posterior distribution of the parameters as well various diagnostic tools to assess the quality of the parameter estimates. The software is open source and runs on Windows and OS X operating systems. In sum, BEESTS allows experimental and clinical psychologists to estimate entire distributions of SSRTs and hence facilitates the more rigorous analysis of stop-signal data.

**Keywords:** stop-signal paradigm, stop-signal RT distribution, ex-Gaussian distribution, hierarchical Bayesian modeling, statistical software

## 1. INTRODUCTION

Response inhibition—the ability to stop an ongoing response—is frequently studied using the stop-signal paradigm. In the stop-signal paradigm (Logan and Cowan, 1984; Lappin and Eriksen, 1966), participants perform a two-choice visual response time (RT) task, such as responding to the color or the shape of the stimuli. This primary task is occasionally interrupted by a stop-signal that instructs participants not to respond on that trial. The goal is to estimate the latency of the unobservable stop response (stop-signal RT; SSRT).

Based on the independent horse-race model (Logan, 1981; Logan and Cowan, 1984), various methods are available to estimate SSRTs (e.g., Logan, 1994; Verbruggen and Logan, 2009; Verbruggen et al., 2009). Over the past decades, the horse-race model has been extensively used to estimate stopping latencies and compare the efficiency of response inhibition between different age groups (e.g., Schachar and Logan, 1990; Kramer et al., 1994; Ridderinkhof et al., 1999; Williams et al., 1999) and clinical populations (Schachar and Logan, 1990; Oosterlaan et al., 1998; Schachar et al., 2000). Unfortunately, most standard methods to estimate SSRTs only provide a summary measure of the latency of the stop process, such as the mean or the median SSRT.

Several researchers have argued, however, that the adequate analysis of RT data should not only focus on mean RT, but should

consider the shape of the entire RT distribution (e.g., Heathcote et al., 1991; Matzke and Wagenmakers, 2009). The shape of SSRT distributions may, for example, differ between different clinical populations, without necessary differences in mean SSRT. Ignoring the shape of SSRT distributions may thus lead to incorrect conclusions about differences in response inhibition between groups.

To allow for a more thorough analysis of stop-signal data, Matzke et al. (2013) have recently developed a Bayesian parametric approach (BPA) that enables researchers to estimate the entire distribution of SSRTs (see Logan et al., *in press*, for an alternative approach). The BPA assumes that SSRTs follow an ex-Gaussian distribution and uses Bayesian parameter estimation to obtain posterior distributions for the model parameters. The BPA enables researchers to compare and evaluate differences in the ex-Gaussian stop parameters between experimental and clinical groups. By doing so, the BPA has the potential to facilitate the interpretation of stop-signal data and contribute to new insights on the nature of response inhibition.

Parameter estimation in the BPA currently relies on the popular Bayesian statistical package WinBUGS (Bayesian inference Using Gibbs Sampling for Windows; Lunn et al., 2012). The practical usefulness of the BPA is, however, severely limited by the disadvantages of the present implementation. The WinBUGS

routine is extremely time consuming and rather user-unfriendly. For instance, WinBUGS requires several hours to produce reliable parameter estimates for a single participant and it requires several days to fit a hierarchical data set. It is therefore all but impossible for experimental and clinical psychologists to take advantage of the theoretical progress offered by the BPA.

In order to overcome this obstacle and promote the broader application of the Bayesian analysis of stop-signal data, we introduce a relatively fast, user-friendly software that allows for the estimation of entire SSRT distributions. BEESTS (**B**ayesian **E**x-Gaussian **E**stimation of **S**TOP-Signal RT distributions) can be applied to individual and hierarchical stop-signal data and comes with an easy-to-use graphical user interface. BEESTS provides users with summary statistics of the posterior distribution of the parameters as well as various diagnostic tools to assess the quality of the parameter estimates.

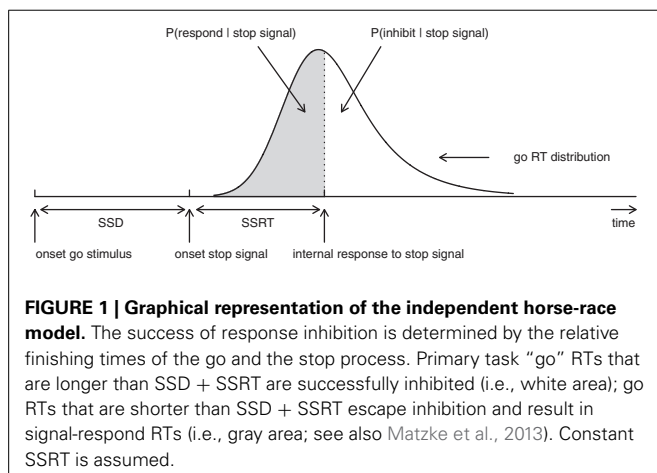
The outline of the paper is as follows. First, we describe the BPA in more detail. Second, we introduce BEESTS, present the installation instructions, and describe the various analysis and output options provided by the software. Third, we illustrate the use of BEESTS with experimental stop-signal data. The last section concludes.

## 2. THE BAYESIAN PARAMETRIC APPROACH

### 2.1. RATIONALE AND ASSUMPTIONS

According to the standard horse-race model (Logan, 1981; Logan and Cowan, 1984), performance in the stop-signal paradigm can be conceptualized as a horse-race between two independent processes that compete against each other: a go-process that is initiated by the primary task “go” stimulus and a stop-process that is generated by the stop-signal. As shown in **Figure 1**, if the go-process finishes before the stop-process, the primary response is executed; if the stop-process finishes before the go-process, the primary response is inhibited. The shorter the time interval between the onset of the go-stimulus and the onset of the stop-signal (i.e., stop-signal delay; SSD), the more likely participants are to inhibit their response on the primary task (see also Matzke et al., 2013).

The BPA is based on the rationale of the standard horse-race model, but it assumes that primary task “go” RTs and SSRTs



are both independent random variables (i.e., complete horse-race model). As shown in **Figure 2**, the BPA assumes that the distribution of RTs that escape inhibition (i.e., signal-respond RTs) can be viewed as a censored go RT distribution. The censoring point is assumed to be drawn from the SSRT distribution and can take on a different value on each stop-signal trial (e.g., SSD + SSRT<sub>1</sub>, SSD + SSRT<sub>2</sub>, and SSD + SSRT<sub>3</sub>). The estimation of the SSRT distribution therefore involves simultaneously estimating the parameters of the go RT distribution and its censoring distribution (see also Matzke et al., 2013).

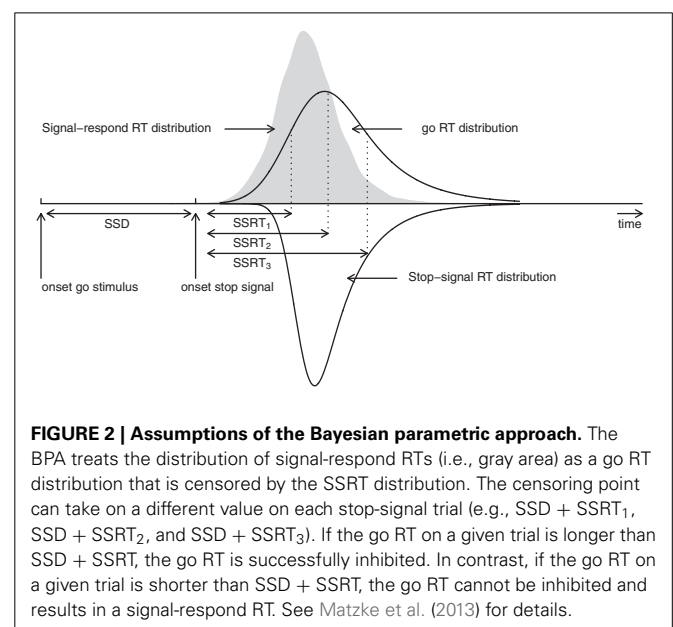
The BPA assumes that the go RTs and SSRTs are ex-Gaussian distributed (Ratcliff and Murdock, 1976; Ratcliff, 1978, 1993; Hockley, 1982, 1984; Heathcote et al., 1991; Matzke and Wagenmakers, 2009). The ex-Gaussian is a three-parameter distribution that is given by the convolution of a Gaussian and an exponential distribution. The  $\mu$  and  $\sigma$  parameters are the mean and the standard deviation of the Gaussian component, respectively, and  $\tau$  is the mean of the exponential component. The  $\mu$  and  $\sigma$  parameters reflect the leading edge and mode of the distribution, whereas  $\tau$  reflects the tail of the distribution. As shown in **Figure 3**, the ex-Gaussian is a positively skewed unimodal distribution that can excellently accommodate the shape of empirical RT data.

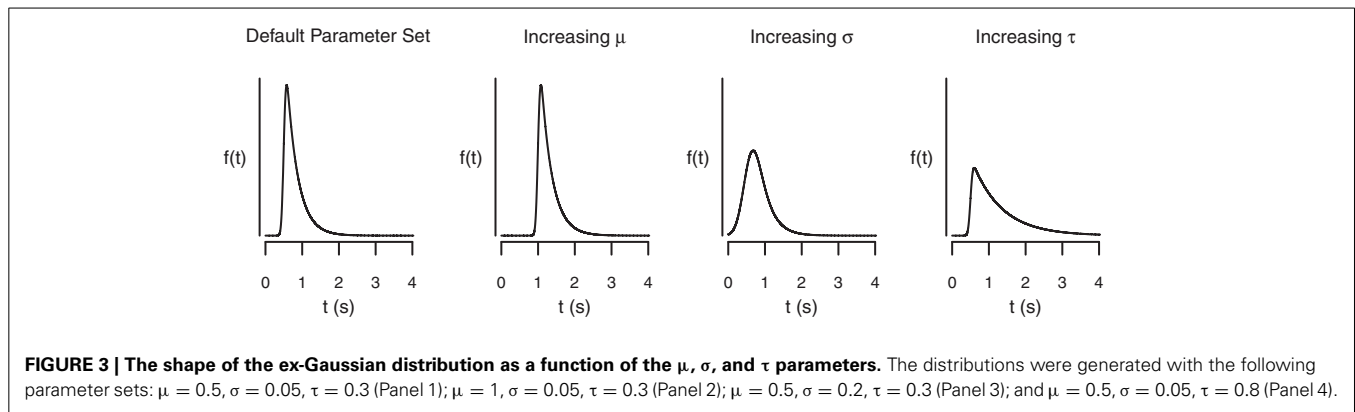
The probability density function of the ex-Gaussian is

$$f(t; \mu, \sigma, \tau) = \frac{1}{\tau} \exp\left(\frac{\mu - t}{\tau} + \frac{\sigma^2}{2\tau^2}\right) \times \Phi\left(\frac{t - \mu}{\sigma} - \frac{\sigma}{\tau}\right) \text{ for } \sigma > 0, \tau > 0, \quad (1)$$

where  $\Phi$  is the standard normal distribution function, given by

$$\Phi\left(\frac{t - \mu}{\sigma} - \frac{\sigma}{\tau}\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{t - \mu}{\sigma} - \frac{\sigma}{\tau}} \exp\left(-\frac{y^2}{2}\right) dy. \quad (2)$$





The mean and variance of the ex-Gaussian distribution equals

$$E(t) = \mu + \tau \quad (3)$$

and

$$\text{Var}(t) = \sigma^2 + \tau^2, \quad (4)$$

respectively. Note that the BPA does not assume that the ex-Gaussian parameters correspond to specific cognitive processes (Matzke and Wagenmakers, 2009); the ex-Gaussian distribution is used as a convenient descriptive model to summarize the distribution of go RTs and SSRTs. As an alternative, one may use, for instance, the ex-Wald distribution (Schwarz, 2001), or “shifted” RT distributions with a parameter-dependent lower bound, such as the shifted Wald, the shifted Weibull or the shifted log normal distribution [e.g., Heathcote, 2004; Heathcote et al., 2004; Rouder, 2005; Rouder et al., 2005; see also Luce (1986) for alternatives].

## 2.2. BAYESIAN PARAMETER ESTIMATION AND PRIORS

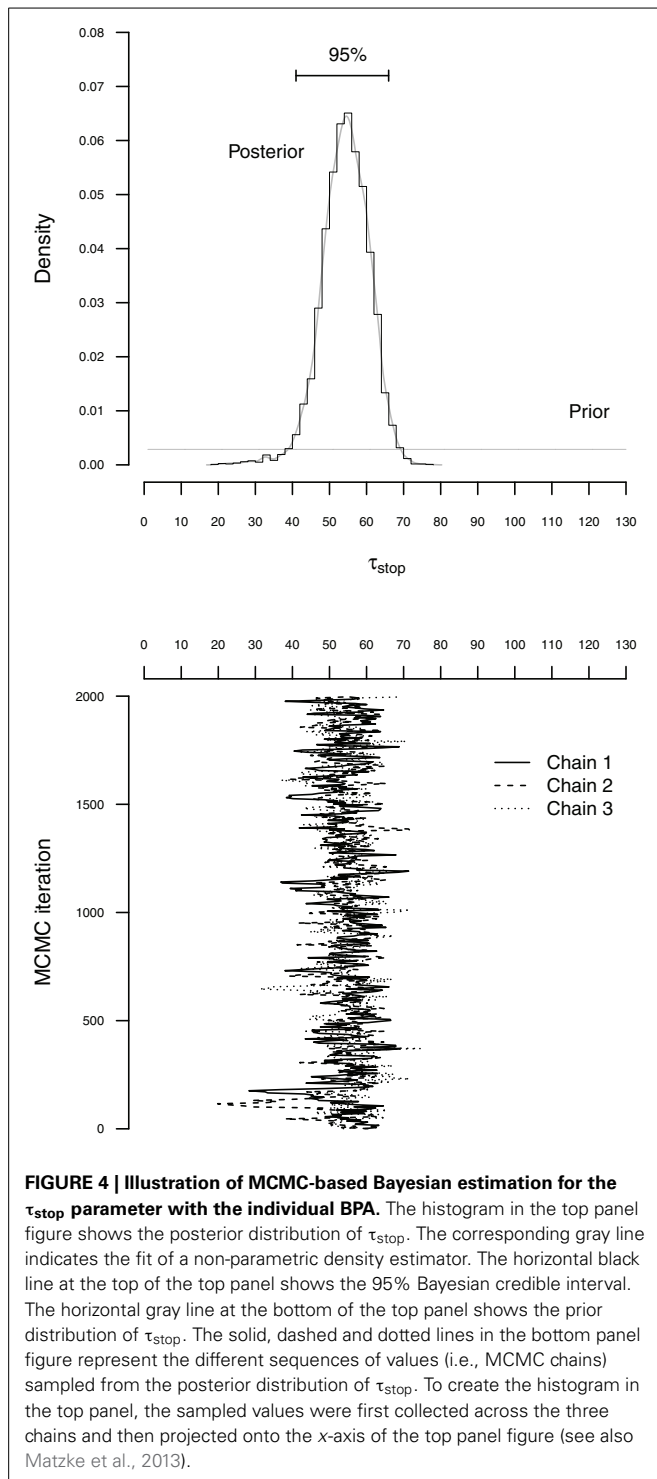
As explained in Matzke et al. (2013), the BPA simultaneously estimates the  $\mu_{\text{go}}$ ,  $\sigma_{\text{go}}$ , and  $\tau_{\text{go}}$  parameters of the go RT distribution and the  $\mu_{\text{stop}}$ ,  $\sigma_{\text{stop}}$ , and  $\tau_{\text{stop}}$  parameters of the SSRT distribution. The BPA relies on Bayesian parameter estimation and therefore involves specifying the prior distribution of the model parameters. BEESTS uses slightly different priors than the WinBUGS implementation of the BPA. Note, however, that Bayesian parameter estimation is insensitive to the choice of the prior as long as sufficiently diagnostic data are available (e.g., Edwards et al., 1963; Gill, 2002; Lee and Wagenmakers, in press). The prior distributions of the model parameters for the BEESTS implementation are listed in the Appendix. The ability of BEESTS to recover underlying true parameter values with the present prior setting has been validated in a series of simulation studies. See the supplemental materials at <http://dora.erbe-matzke.com/publications.html> for a summary of the results of the parameter recoveries.

The BPA relies on Markov chain Monte Carlo sampling (MCMC; Gilks et al., 1996; Gamerman and Lopes, 2006) to obtain posterior distributions for the go and stop parameters. **Figure 4** illustrates the basic concepts of Bayesian parameter estimation using MCMC sampling. The bottom panel of **Figure 4**

shows sequences of values (i.e., MCMC chains) sampled from the posterior distribution of the  $\tau_{\text{stop}}$  parameter. The accuracy of the sampling process can be increased by running multiple chains, discarding the beginning of each chain as burn-in, and by thinning the chains to decrease autocorrelation. In the present illustration, we ran three chains, each with different starting values and retained 2000 iterations per chain, resulting in a total of 6000 samples from the posterior distribution (see also Matzke et al., 2013).

The top panel of **Figure 4** shows the prior and posterior distribution of the  $\tau_{\text{stop}}$  parameter. The horizontal gray line at the bottom of the figure shows the prior distribution of  $\tau_{\text{stop}}$ . The prior is updated by the incoming data to yield the posterior distribution. The histogram and the gray density plot show the distribution of the samples drawn from the posterior distribution of  $\tau_{\text{stop}}$  collapsed over the three MCMC chains. The posterior distribution quantifies the uncertainty about the estimate of  $\tau_{\text{stop}}$ . The central tendency of the posterior, such as the median, is often used as a point estimate of the parameter. The dispersion of the posterior, such as the standard deviation or the percentiles, quantifies the precision of the parameter estimate; the larger the dispersion, the greater the uncertainty in the estimated parameter. For example, the horizontal line at the top of **Figure 4** ranges from the 2.5th to the 97.5th percentile of the posterior (i.e., 95% Bayesian credible interval), indicating that we can be 95% confident that the true value of  $\tau_{\text{stop}}$  lies within this range (see also Matzke et al., 2013).

Before interpreting the parameter estimates, it is crucial to ensure that the chains have converged from their starting values to their stationary distributions. First, we verify that the posterior distributions of the model parameters are unimodal. Second, we run multiple MCMC chains and ascertain that the chains have mixed well. At convergence, the individual MCMC chains should look like “hairy caterpillars” and should be indistinguishable from one another. Lastly, we compute the  $\hat{R}$  (Gelman and Rubin, 1992) convergence diagnostic measure for each model parameter.  $\hat{R}$  compares the between-chain variability to the within-chain variability. As a rule of thumb,  $\hat{R}$  should be lower than 1.1 if the chains have properly converged. In case of convergence problems, we recommend that users increase the number of samples, the length of the burn-in period, and the degree of thinning.



The BPA can be applied to individual as well as hierarchical stop-signal data. See Matzke et al. (2013) for the graphical representation of the individual and hierarchical BPA models. For the individual analysis, the goal is to estimate the ex-Gaussian go and stop parameters for each participant separately. In contrast, for the hierarchical analysis (e.g., Rouder et al., 2005; Gelman and Hill, 2007; Farrell and Ludwig, 2008; Matzke and Wagenmakers,

2009; Lee, 2011), the BPA assumes that the participant-level go and stop parameters are drawn from group-level distributions. The group-level distributions specify the between-subject variability of the participant-level parameters. The group-level distributions are themselves characterized by a set of group-level parameters. The goal is to simultaneously estimate the group-level parameters as well as the participant-level go and stop parameters. As explained in Matzke et al. (2013), hierarchical modeling is particularly valuable in situations with only a small number of observations per participant and moderate between-subject variability in parameter values (Gelman and Hill, 2007). In such situations, Bayesian hierarchical modeling typically yields less variable and more accurate estimates than single-level parameter estimation (Rouder et al., 2005; Farrell and Ludwig, 2008). The advantages of the hierarchical approach are less pronounced in situations with a large number of observations per participant. Similarly, in settings with only a few participants—a typical scenario in psychophysical experiments—the group-level parameters cannot be estimated precisely, a problem that diminishes the benefits of hierarchical modeling. In these cases, the individual approach may perform similarly well as the hierarchical approach.

### 3. RELEASING THE BEESTS

BEESTS is a cross-platform open-source software for the estimation of SSRT distributions with the BPA (Matzke et al., 2013). BEESTS relies on Python for parameter estimation and on R (R Core Team, 2012) for the post-processing of the posterior distribution of the model parameters. Specifically, BEESTS uses the Python-based toolboxes kabuki (Wiecki et al., 2013) and PyMC (Patil et al., 2010) to construct the model and to generate samples from the posterior distribution of the model parameters using Metropolis-within-Gibbs sampling (Tierney, 1994), respectively. For computational efficiency, the likelihood functions are coded in Cython (Behnel et al., 2011). Once the model parameters are estimated, BEESTS relies on R to compute summary statistics for the posterior distribution of the model parameters and to assess the quality of the parameter estimates. As shown in Figure 5, BEESTS is equipped with an easy-to-use graphical user interface (GUI).

### 4. INSTALLATION

BEESTS is a stand-alone and open source software released under the Affero General Public License. BEESTS runs on Windows (Windows XP and Windows 7) and OS X (Mountain Lion) operating systems. The software is freely available at <http://dora.erbe-matzke.com/software.html>. To install BEESTS on Windows, download BEESTS-1.2.zip and unpack the zip file at any desired location on your computer. Start the GUI by clicking on BEESTS.exe. To install BEESTS on OS X, download BEESTS-1.2.dmg, double-click the file, and install it on your computer.

### 5. LOADING DATA

The top panels of Figure 6 show the required data format for the analysis. Data files should be saved as csv (i.e., comma-separated values) files. For the individual analysis, the

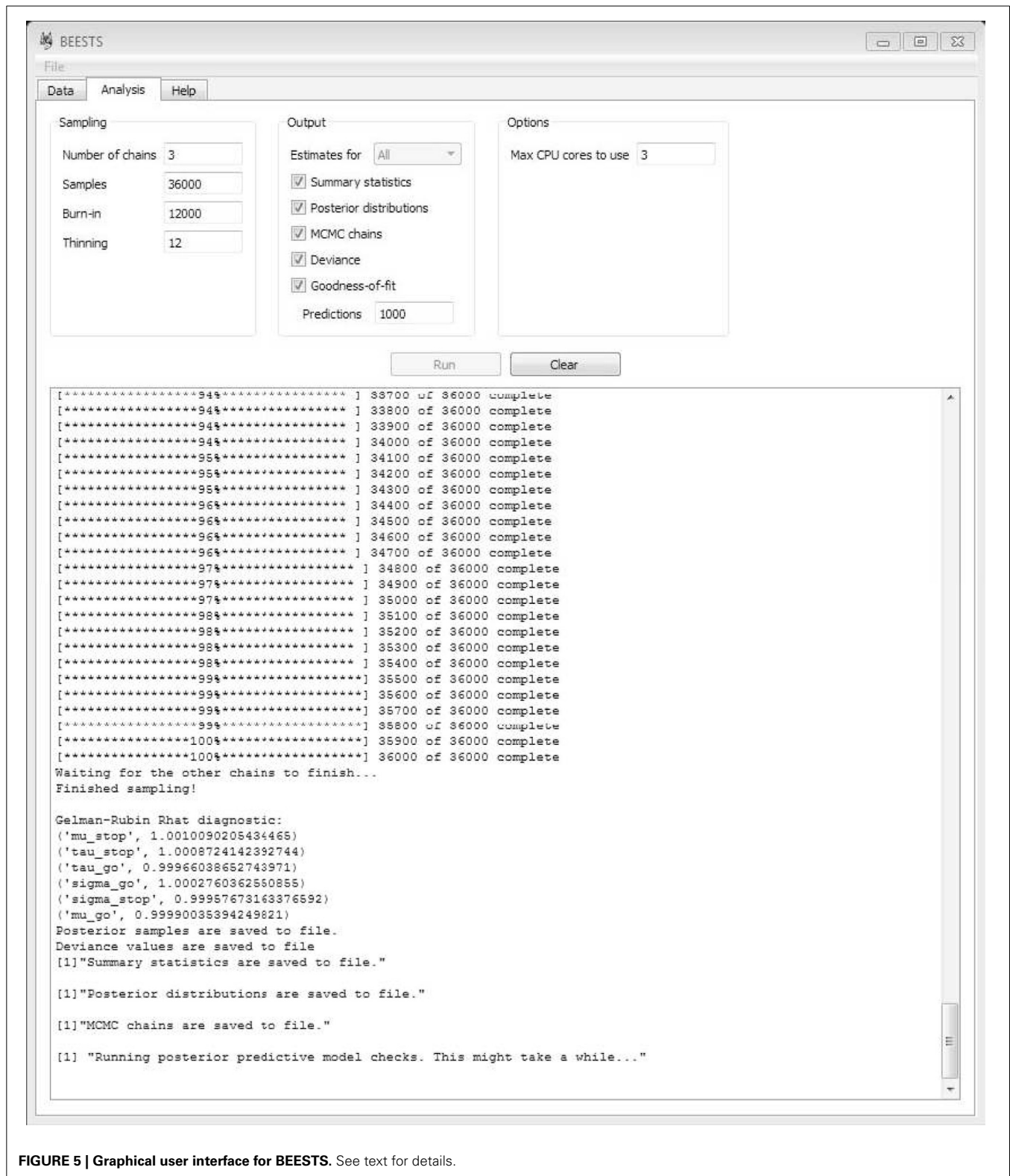
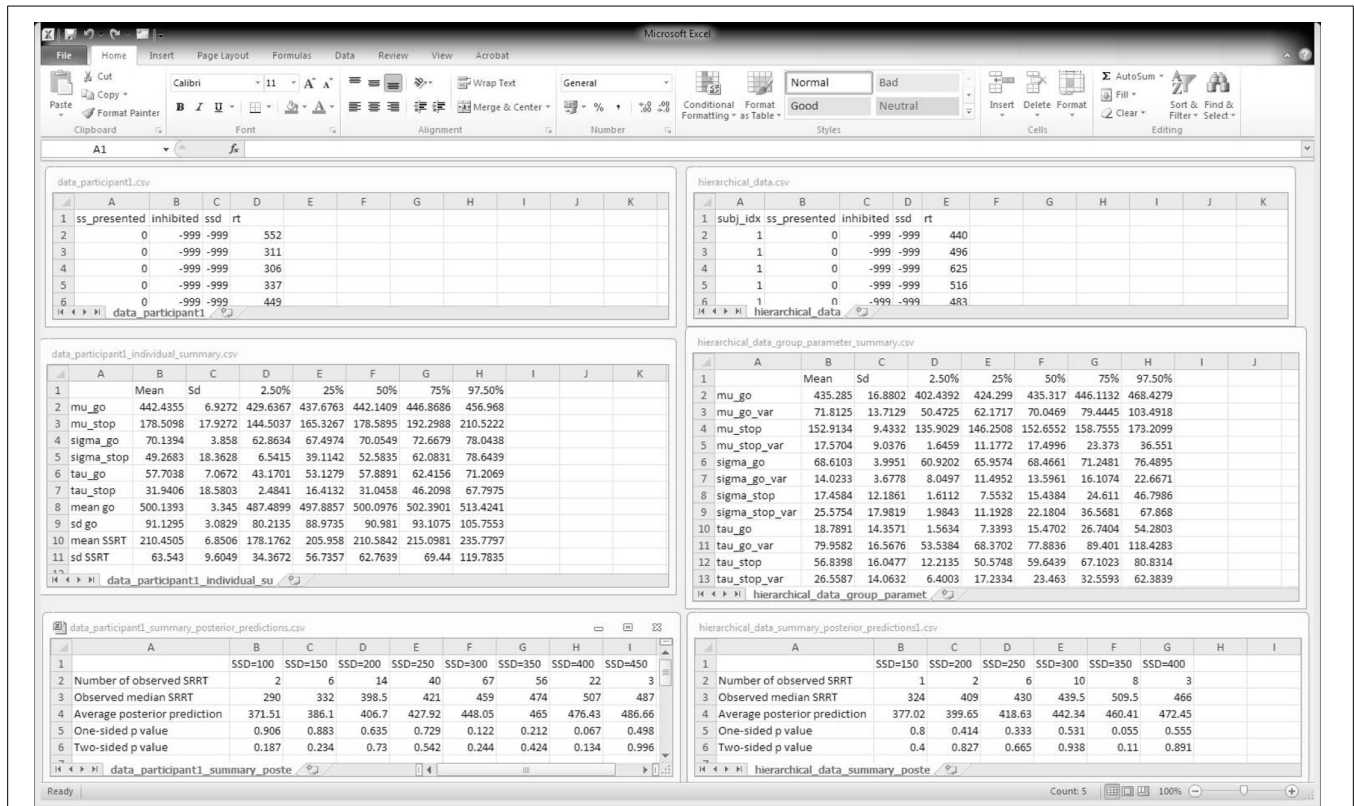


FIGURE 5 | Graphical user interface for BEESTS. See text for details.

first row of the data file must contain the column names "ss\_presented", "inhibited", "ssd", and "rt". The remaining rows contain the data for each go and stop-signal trial. For the hierarchical analysis, the first row of

the data file must additionally contain the column name "subj\_idx". See **Table 1** for instructions on response coding and the examples folder in BEESTS for examples of the data format.



**FIGURE 6 | BEESTS input and output.** The **left panels** show input and output for the individual analysis. The **right panels** show input and output for the hierarchical analysis. The **top panels** show the required data format. The **middle panels** show the output of the Summary statistic option. For the

hierarchical analysis, only the group-level mean and group-level variability (i.e., standard deviation) parameters are shown. The **bottom panels** show partial output for the Goodness-of-fit option for Participant 1 in the Bissett and Logan (2011) experiment. SRRT = signal-respond RT.

**Table 1 | Response coding for the hierarchical BEESTS analysis.**

"subj_idx"	"ss_presented"	"inhibited"	"ssd"	"rt"
1	0	-999	-999	656
1	1	0	300	469
1	1	1	300	-999

The "subj\_idx" column contains the participant number. The "ss\_presented" column contains the trial type, where go trials are coded with 0 and stop-signal trials are coded with 1. The "inhibited" column contains the inhibition data, where signal-respond trials are coded with 0 (i.e., unsuccessful inhibition), signal-inhibit trials are coded with 1 (i.e., successful inhibition), and go trials are coded with -999. The "ssd" column contains the stop-signal delay in ms., where go trials are coded with -999. The "rt" column contains the go RT for go trials and the signal-respond RT for signal-respond trials in ms., where signal-inhibit trials are coded with -999.

To load the data file, click on Open in the File menu and follow the instructions. Based on the data format, BEESTS automatically infers whether an individual or hierarchical analysis is appropriate: data files without the "subj\_idx" column are analyzed with the individual BPA, whereas data files with the "subj\_idx" column are analyzed with the hierarchical BPA.

## 6. ANALYSIS

Once the data are loaded, users can specify the details of the MCMC sampling, the required output, and the preferred number of CPU cores used by BEESTS.

### 6.1. SAMPLING

BEESTS allows users to specify the following aspects of the sampling run. Typical values of the input arguments are shown in Figure 5.

#### 6.1.1. Number of chains

Use the Number of chains option to specify the number of MCMC chains, i.e., sequences of values sampled from the posterior distribution of the parameters. The start values are automatically set to the maximum a posteriori probability (MAP) estimates of the parameters.

#### 6.1.2. Samples

Use the Samples option to specify the total number of MCMC samples per chain.

#### 6.1.3. Burn-in

Use the Burn-in option to specify the number of burn-in samples to discard at the beginning of each chain.

### 6.1.4. Thinning

Use the `Thinning` option to specify the degree of thinning within each chain. For instance, a thinning factor of 12 means that only every 12th MCMC sample will be retained.

## 6.2. OUTPUT

All output will be saved in the directory where the data file is located. BEESTS automatically saves the posterior samples from each chain to a separate csv file (e.g., `name.datafile_parameters1.csv`, `name.datafile_parameters2.csv`, etc.). If multiple chains are run, BEESTS automatically displays the  $\hat{R}$  statistic for each model parameter (see **Figure 5**).

As shown in **Figure 5**, BEESTS allows users to request the following additional output. If `Estimates for` is set to `All` in a hierarchical analysis, BEESTS will provide the selected output options (i.e., summary statistics, density plots of the posterior distributions, and MCMC trace plots) for the group-level parameters *and* for each participant separately. If `Estimates for` is set to `Only-group`, BEESTS will provide the selected output options only for the group-level parameters.

### 6.2.1. Summary statistics

Use the `Summary statistics` option to obtain a csv file with the summary statistics (i.e., mean, standard deviation, and quantiles) of the posterior distribution of the model parameters and of the corresponding mean and standard deviation of the go and SSRT distribution (see Equations 3, 4).

### 6.2.2. Posterior distributions

Use the `Posterior distributions` option to obtain a pdf file with the density plots of the posterior and the prior distribution of the model parameters.

### 6.2.3. MCMC chains

Use the `MCMC chains` option to obtain a pdf file with trace plots for the MCMC chains of the model parameters.

### 6.2.4. Deviance

Use the `Deviance` option to obtain the deviance values from each chain in a separate csv file (e.g., `name.datafile_deviance1.csv`, `name.datafile_deviance2.csv`, etc.). The deviance values may be used to compute the Deviance Information Criterion (DIC, e.g., Spiegelhalter et al., 2002) measure of model selection.

### 6.2.5. Goodness-of-fit

Use the `Goodness-of-fit` option to assess the absolute goodness-of-fit of the model using posterior predictive model checks. As explained in Matzke et al. (2013), the adequacy of the model can be assessed by generating predicted data using the posterior distributions of the parameters. If the model adequately describes the data, the predictions based on the model parameters should closely approximate the observed data. The model checks can be formalized by computing posterior predictive  $p$  values [e.g., Gelman et al., 1996; Gelman and Hill, 2007, but see Bayarri and Berger (1998)]. Extreme  $p$  values close to 0 or 1 indicate that the BPA does not describe the observed data adequately.

For each individual participant, BEESTS uses the median of the observed and predicted signal-respond RTs as test statistics. The `Predictions` option can be used to specify the number of predicted data sets. BEESTS then randomly samples the specified number of parameter vectors from the joint posterior of the individual go and stop parameters. Next, BEESTS generates the specified number of predicted stop-signal data sets for each SSD using the corresponding number of stop-signal trials and the chosen parameter vectors. For each SSD, BEESTS then computes the median signal-respond RT in each predicted data set. Lastly, for each SSD, BEESTS computes the one-sided posterior predictive  $p$  value given by the fraction of times that the predicted median signal-respond RT is greater than the observed median signal-respond RT. Corresponding two-sided  $p$  values can be computed as  $2 \times \min(p, 1 - p)$ . Note, however, that two-sided  $p$  values are well defined only when the test statistic has a symmetric distribution. Note also that BEESTS assesses model fit on all SSDs that contain at least one observed signal-respond RT. In order to obtain stable median signal-respond RTs, however, we advise users to interpret the results only on SSDs with a reasonable number of observed signal-respond RTs.

The output of the posterior predictive model checks consists of (1) a csv file listing for each SSD the number of observed signal-respond RTs, the observed median signal-respond RT, the average of the predicted median signal-respond RTs, and the one-sided and two-sided posterior predictive  $p$  value and (2) a pdf file with a graphical summary of the model checks using violin plots. Violin plots (e.g., Hintze and Nelson, 1998) combine information available from density plots with information about summary statistics in the form of box plots. Note that irrespective of the type of analysis (individual or hierarchical), the goodness-of-fit of the model is assessed on a participant level using the parameter values of the individual participants (see **Figure 8**).

## 6.3. OPTIONS: MAX CPU CORES TO USE

Use the `Max CPU cores to use` option to specify the number of CPU cores to use during the sampling process. If multiple MCMC chains are requested, BEESTS can run the chains in parallel by allocating each chain to a different CPU core in order to increase speed. The default number of CPU cores used by BEESTS is the number of cores available on the computer minus one.

## 6.4. RUNNING THE ANALYSIS

Once the details of the sampling process and the required output are specified, start the analysis by clicking on `Run`. As shown in **Figure 5**, BEESTS automatically displays the progress of the sampling. If multiple MCMC chains are run in parallel, BEESTS displays the progress of only one of the MCMC chains (i.e., the main process). The analysis can be stopped by “killing” the (parallel) processes in the Task Manager. Use the `Clear` command to clear the working space.

## 7. EMPIRICAL DATA EXAMPLES: INDIVIDUAL AND HIERARCHICAL ANALYSIS

In this section, we illustrate the use of BEESTS with the stop-signal data of 20 participants from the 40% stop-signal condition of the first experiment reported in Bissett and Logan (2011). The

data set featured a relatively large number of 720 go trials and 480 stop-signal trials per participant. See Matzke et al. (2013) for the details on the data pre-processing and the model fitting. For all of the participants, the BEESTS implementation yielded parameter estimates that are highly similar to the ones obtained from the WinBUGS routine. For a comparison of the parameter estimates from the BEESTS and the WinBUGS implementation, the reader is referred to the supplemental materials and to the empirical data examples in Matzke et al. (2013).

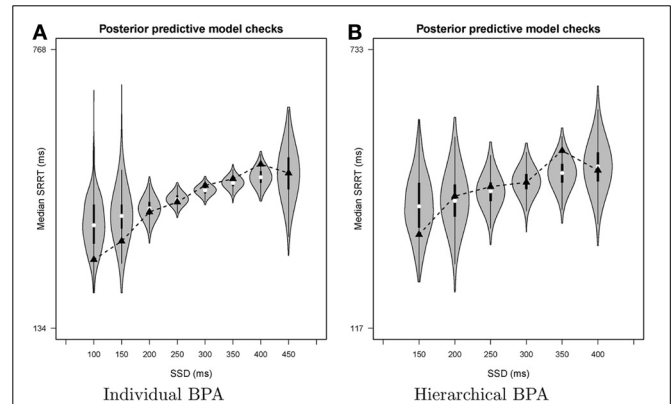
Due to relatively high autocorrelations between the parameters, we ran long chains, discarded the beginning of the chains as burn-in, and thinned each chain. The results reported below are based on 6000 retained samples, using Number of chains = 3, Samples = 36000, Burn-in = 12000, and Thinning = 12.

### 7.1. INDIVIDUAL ANALYSIS

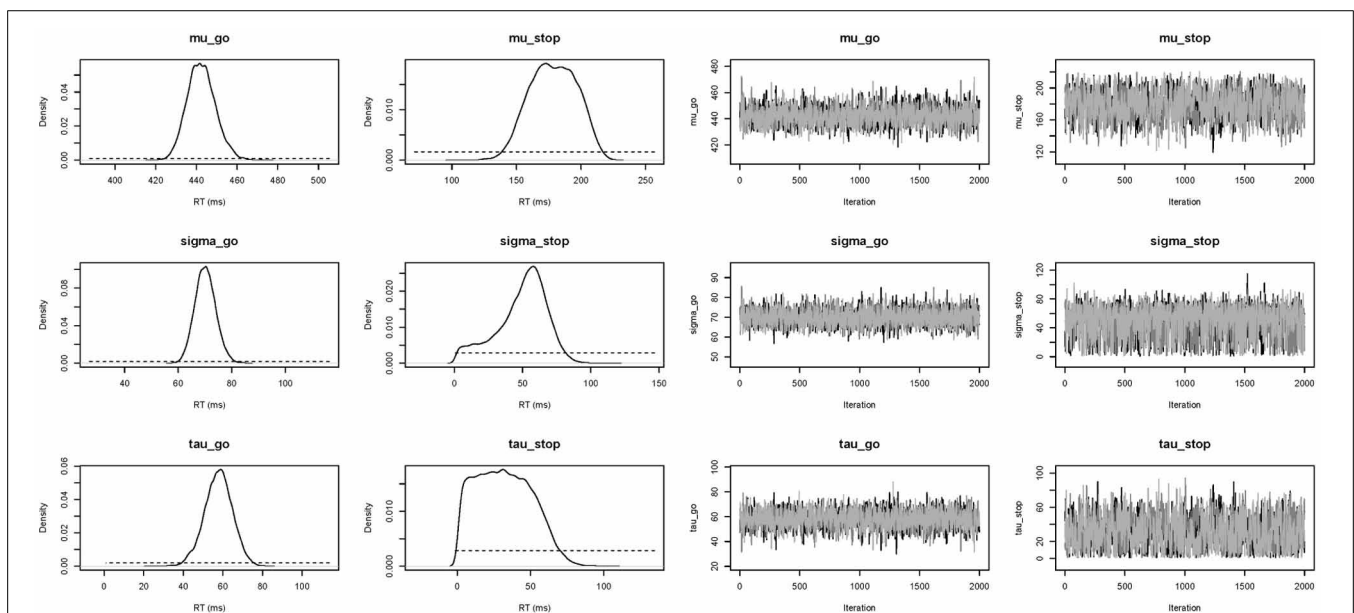
In this section, we present the results of fitting the data of Participant 1 with the individual BPA. See the examples folder for the data set. Using three CPU cores, the sampling took approximately 23 min with BEESTS. The same analysis took about 15 h with WinBUGS. The top left panel of Figure 6 shows the required data format for the individual analysis. Figure 7 shows the posterior and prior distributions (left panel; option Posterior distributions) and the MCMC chains (right panel; option MCMC chains) for the six model parameters. The prior distributions are adequately updated; the posteriors are substantially narrower than the priors. The posterior distributions and the three MCMC chains do not show signs of convergence problems. All  $\hat{R}$  values were lower than 1.05. The middle left panel of Figure 6 shows the summary statistics of the posterior distribution of the model parameters (option Summary statistic).

The posterior distributions are estimated well as evidenced by the relatively small posterior standard deviations. The go parameters are generally estimated more precisely than the stop parameters because the go parameters are estimated based on the go RTs as well as the signal-respond RTs and are therefore better constrained by the data.

The bottom left panel of Figure 6 shows the summary of the posterior predictive model checks (option



**FIGURE 8 | Results of the posterior predictive model checks for Participant 1 in the Bissett and Logan (2011) data set with the individual BPA (panel A) and the hierarchical BPA (panel B).** See text for a detailed description of the posterior predictive analyses. For each SSD, the figures show the observed median signal-respond RT (black triangle), a density plot of the predicted median signal-respond RTs (gray violin plot), a boxplot ranging from the 25th to the 75th percentile of the predicted median signal-respond RTs, and the median of the predicted median signal-respond RTs (white circle). SRRT = signal-respond RT.



**FIGURE 7 | Posterior (black solid lines) and prior distributions (black dotted lines; left panel) and MCMC chains (right panel) of the model parameters for Participant 1 in the Bissett and Logan (2011) data set obtained with the individual BPA.**

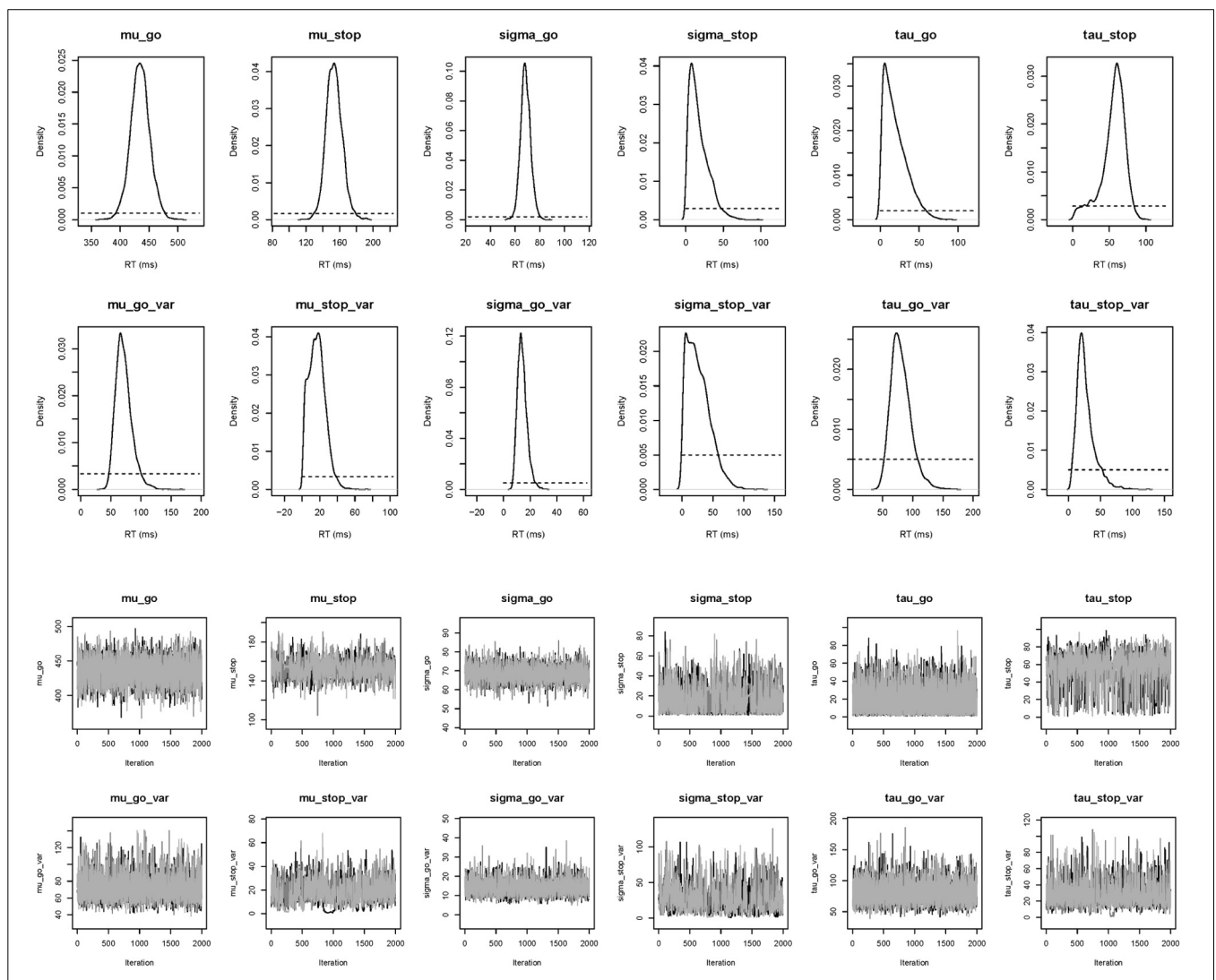


Goodness-of-fit) using 1000 samples from the joint posterior of the model parameters ( $Samples = 1000$ ). As mentioned above, we advise users to assess model fit only on SSDs with a reasonable number of observed signal-respond RTs. For instance, we assessed goodness-of-fit only on SSDs with at least 10 observed signal-respond RTs. The one-sided  $p$  values on these five SSDs (i.e., 200, 250, 300, 350, and 400 ms) are far from 0 or 1 and the two-sided  $p$  values are all above 0.05. The left panel of **Figure 8** shows the corresponding graphical summary for the model checks. For the selected SSDs, the observed median signal-respond RTs (i.e., black triangles) are well within the 2.5th and 97.5th percentile of the predicted median signal-respond RTs (see gray violin plots), and are adequately approximated by the median of the predicted median signal-respond RTs (i.e., white circles). The results of the posterior predictive model checks

indicated thus that the BEESTS analysis appropriately accounted for the observed data.

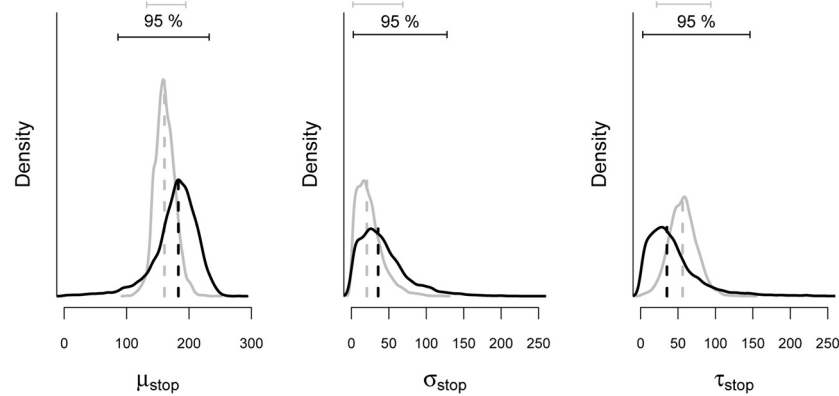
### 7.2. HIERARCHICAL ANALYSIS

As explained above, the hierarchical approach has the potential to provide accurate parameter estimates with relatively few observations per participant. To illustrate the benefits of the hierarchical approach over the individual BPA with scarce data, this section presents the results of fitting a subsample of the observations from the Bissett and Logan (2011) data set with the hierarchical as well as the individual BPA. For each of the 20 participants, we fit a randomly selected 90 go RTs, 30 signal-respond RTs, and 30 successful inhibitions with the hierarchical BPA. We then compared the results from the hierarchical analysis to the results from fitting the same subsample of data with the individual BPA. Using



**FIGURE 9 | Posterior distributions and MCMC chains of the group-level model parameters in the Bissett and Logan (2011) data set obtained with the hierarchical BPA.** The first and third rows show posterior (black solid line) and prior distributions (black dotted line) and MCMC trace plots

for the group-level mean parameters, respectively. The second and fourth rows show posterior and prior distributions and trace plots for the group-level variability (i.e., group-level standard deviation) parameters, respectively.



**FIGURE 10 | Posterior distribution of the stop parameters estimated from a subsample of the data of Participant 1 with the individual and the hierarchical BPA.** The solid black and gray lines show the posterior distribution of the stop parameters and the corresponding 95% Bayesian

credible intervals obtained with the individual and the hierarchical BPA, respectively. The dashed black and gray lines show the median of the posterior distributions obtained with the individual and the hierarchical BPA, respectively (see also Matzke et al., 2013).

three CPU cores, the hierarchical analysis took approximately 3.5 h with BEESTS. The same analysis took about 100 h with WinBUGS.

The top right panel of **Figure 6** shows the required data format for the hierarchical analysis. **Figure 9** shows the posterior and prior distributions (top panel) and the MCMC chains (bottom panel) for the group-level mean and standard deviation parameters. The prior distribution of the group-level parameters are adequately updated; the posteriors are substantially narrower than the priors and the chains have mixed well. The  $\hat{R}$  values for all group-level and individual parameters were lower than 1.05. The middle right panel of **Figure 6** shows the summary statistics of the posterior distribution of the group-level mean and standard deviation parameters. The posterior distributions are estimated relatively precisely. Note that if the *Estimates for All* option is selected, BEESTS also produces output (i.e., density plots of the posteriors, MCMC trace plots, and summary statistics) for the individual go and stop parameters for each participant separately.

The bottom right panel of **Figure 6** shows the summary of the posterior predictive model checks for Participant 1 using 1000 samples from the joint posterior of the participant-level model parameters obtained with the hierarchical BPA. All posterior predictive  $p$  values are well within an acceptable range. Note, however, that the median signal-respond RTs—observed and predicted—are based on only a few observations. The right panel of **Figure 8** shows the corresponding graphical summary of the posterior predictive model checks. All observed median signal-respond RTs are well within the range of the median signal-respond RTs predicted by the joint posterior of the model parameters. Due to the scarcity of the data, however, there is large uncertainty in the predicted median signal-respond RTs. Compare the results of the posterior predictive model checks in the two panels of **Figure 8**. The violin plots in the left panel show the predicted median signal-respond RTs from the individual analysis of the data of Participant 1 based on the full 1200 trials.

The violin plots in the right panel show the predicted median signal-respond RTs from the hierarchical analysis of the data of Participant 1 based on a subsample of only 150 trials. Because the hierarchical analysis is based on substantially fewer observations than the individual analysis of the full data set presented in the previous section, the predicted median signal-respond RTs in the right panel are more spread out than the predicted median signal-respond RTs in the left panel. Posterior predictive  $p$  values resulting from such unstable observed and predicted median signal-respond RTs should be interpreted with caution.

To illustrate the benefits of the hierarchical approach over the individual BPA with scarce data, we compared the parameter estimates from the hierarchical analysis with estimates obtained from the individual analysis of the same subsample of 150 trials. As mentioned above, hierarchical modeling generally results in more accurate and less variable estimates than single-level estimation. **Figure 10** shows the posterior distribution of the stop parameters of Participant 1 obtained with the hierarchical and the individual BPA using the same subsample of 150 observations. The gray density plots show the posterior distribution of the stop parameters from the hierarchical BPA. The black density plots show the posterior distribution of the stop parameters from the individual analysis. The posterior distributions of the stop parameters estimated with the hierarchical approach are less variable (i.e., smaller 95% Bayesian credible interval) than the posteriors estimated with the individual BPA. Also, the posterior medians from the hierarchical analysis are—as expected—shrunk toward their corresponding group mean (see also Matzke et al., 2013).

## 8. DISCUSSION

The horse-race model presents various opportunities to estimate the latency of response inhibition in the stop-signal paradigm. Most methods, however, only focus on deriving a summary measure of SSRT. Recently, Matzke et al. (2013) have developed the BPA that allows for the estimation of the entire distribution of stopping latencies. The goal of the present paper was to promote the widespread application of the Bayesian analysis of stop-signal

data by introducing BEESTS, a relatively fast and user-friendly software implementation of the BPA. BEESTS provides users with a range of output options, such as summary statistics of the posterior distribution of the parameters and various diagnostic tools to assess the quality of the estimates. Importantly, BEESTS is equipped with an easy-to-use graphical user interface.

BEESTS can be applied to individual as well as hierarchical stop-signal data. The advantage of the individual approach lies in its simplicity. The advantage of the hierarchical approach lies in its potential to provide accurate parameter estimates with relatively few observations per participant. The choice between the individual and the hierarchical approach in practical applications depends on a delicate balance between the quality of the data, the number of participants, the number of trials per participant, and whether users are interested in obtaining accurate parameter estimates on the participant level in order to examine individual differences or focus on group comparisons and are satisfied with interpreting only the group-level parameters. Prior to data collection, users are encouraged to generate synthetic data with varying number of trials and participants, fit the data in BEESTS, and inspect the parameter estimates in order to assess the expected uncertainty of the model parameters under the different scenarios and modeling approaches.

BEESTS assumes that go RTs and SSRTs are ex-Gaussian distributed and relies on Bayesian parameter estimation to obtain estimates for the go and stop parameters. Note, however, that the BPA itself does not hinge on the particular parametric form used to summarize the distributions, nor is it heavily influenced by the exact choice of the prior distributions. In our experience, the ex-Gaussian assumption and the corresponding (group-level and hyper) prior distributions implemented in BEESTS provide a reasonable default setting. Nevertheless, interested users may adapt the source code (<https://github.com/twiecki/stopsignal>) to accommodate alternative parametric assumptions or different prior settings. Also, the posterior predictive model check implemented in BEESTS using the median signal-respond RT is only one of many possible approaches to assess the goodness-of-fit of the model. Users may adapt the source code to implement posterior predictive model checks using alternative test statistics (see Matzke et al., 2013).

## 9. CONCLUSION

Here we introduced a user-friendly software package—BEESTS—that allows for the efficient estimation of entire SSRT distributions using MCMC sampling. BEESTS allows researchers to rigorously address important questions about the variability of stopping latencies, such as the relationship between mean SSRT and SSRT variance. Similarly, BEESTS enables investigators to assess differences in the shape of go RT and SSRT distributions between clinical populations or experimental groups. BEESTS therefore facilitates the interpretation of stop-signal data and may open fruitful new avenues for response inhibition research.

## ACKNOWLEDGMENTS

This work was supported by the starting grant “Bayes or Bust” awarded by the European Research Council.

## REFERENCES

- Bayarri, M. J., and Berger, J. O. (1998). “Quantifying surprise in the data and model verification,” in *Bayesian Statistics 6: Proceedings of the Sixth Valencia International Meeting*, 6–10 June 1998, eds J. M. Bernardo, J. O. Berger, A. P. Dawid, and A. F. M. Smith (Oxford: Oxford University Press), 53–82.
- Behnel, S., Bradshaw, R., Citro, C., Dalcin, L., Seljebotn, D. S., and Smith, K. (2011). Cython: the best of both worlds. *Comput. Sci. Eng.* 13, 31–39. doi: 10.1109/MCSE.2010.118
- Bissett, P. G., and Logan, G. D. (2011). Balancing cognitive demands: control adjustments in the stop-signal paradigm. *J. Exp. Psychol. Learn. Mem. Cogn.* 37, 392. doi: 10.1037/a0021800
- Edwards, W., Lindman, H., and Savage, L. J. (1963). Bayesian statistical inference for psychological research. *Psychol. Rev.* 70, 193–242. doi: 10.1037/h0044139
- Farrell, S., and Ludwig, C. J. H. (2008). Bayesian and maximum likelihood estimation of hierarchical response time models. *Psychon. Bull. Rev.* 15, 1209–1217. doi: 10.3758/PBR.15.6.1209
- Gamerman, D., and Lopes, H. F. (2006). *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference*. Boca Raton, FL: Chapman and Hal/CRC.
- Gelman, A., and Hill, J. (2007). *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge: Cambridge University Press.
- Gelman, A., Meng, X., and Stern, H. (1996). Posterior predictive assessment of model fitness via realized discrepancies. *Stat. Sin.* 6, 733–807.
- Gelman, A., and Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Stat. Sci.* 7, 457–472. doi: 10.1214/ss/1177011136
- Gilks, W. R., Richardson, S., and Spiegelhalter, D. J. (1996). *Markov Chain Monte Carlo in Practice*. Boca Raton, FL: Chapman and Hal/CRC. doi: 10.1007/978-1-4899-4485-6
- Gill, J. (2002). *Bayesian Methods: A Social and Behavioral Sciences Approach*. Boca Raton, FL: CRC Press.
- Heathcote, A. (2004). Fitting Wald and ex-Wald distributions to response time data: an example using functions for the S-PLUS package. *Behav. Res. Meth.* 36, 678–694. doi: 10.3758/BF03206550
- Heathcote, A., Brown, S., and Cousineau, D. (2004). QMPE: estimating Lognormal, Wald, and Weibull RT distributions with a parameter-dependent lower bound. *Behav. Res. Meth.* 36, 277–290. doi: 10.3758/BF03195574
- Heathcote, A., Popiel, S. J., and Mewhort, D. J. (1991). Analysis of response time distributions: an example using the Stroop task. *Psychol. Bull.* 109, 340–347. doi: 10.1037/0033-2909.109.2.340
- Hintze, J. L., and Nelson, R. D. (1998). Violin plots: a box plot–density trace synergism. *Am. Stat.* 52, 181–184. doi: 10.1080/00031305.1998.10480559
- Hockley, W. E. (1982). Retrieval processes in continuous recognition. *J. Exp. Psychol. Learn. Mem. Cogn.* 8, 492–512. doi: 10.1037/0278-7393.8.6.497
- Hockley, W. E. (1984). Analysis of response time distributions in the study of cognitive processes. *J. Exp. Psychol. Learn. Mem. Cogn.* 10, 598–615. doi: 10.1037/0278-7393.10.4.598
- Kramer, A. F., Humphrey, D. G., Larish, J. F., Logan, G. D., and Strayer, D. L. (1994). Aging and inhibition: beyond a unitary view of inhibitory processing in attention. *Psychol. Aging* 9, 491–512. doi: 10.1037/0882-7974.9.4.491
- Lappin, J. S., and Eriksen, C. W. (1966). Use of a delayed signal to stop a visual reaction–time response. *J. Exp. Psychol.* 72, 805–811. doi: 10.1037/h0021266
- Lee, M. D. (2011). How cognitive modeling can benefit from hierarchical Bayesian models. *J. Math. Psychol.* 55, 1–7. doi: 10.1016/j.jmp.2010.08.013
- Lee, M. D., and Wagenmakers, E.-J. (in press). *Bayesian Modeling for Cognitive Science: A Practical Course*. Cambridge, MA: Cambridge University Press.
- Logan, G. D. (1981). “Attention, automaticity, and the ability to stop a speeded choice response,” in *Attention and Performance IX*, eds J. Long and A. Baddeley (Hillsdale, NY: Erlbaum), 205–222.
- Logan, G. D. (1994). “On the ability to inhibit thought and action: a users’ guide to the stop signal paradigm,” in *Inhibitory Processes in Attention, Memory and Language*, eds D. Dagenbach and T. H. Carr (San Diego, CA: Academic Press), 189–240.
- Logan, G. D., and Cowan, W. B. (1984). On the ability to inhibit thought and action: a theory of an act of control. *Psychol. Rev.* 91, 295–327. doi: 10.1037/0033-295X.91.3.295
- Logan, G. D., Van Zandt, T., Verbruggen, F., and Wagenmakers, E.-J. (in press). On the ability to inhibit thought and action: general and special theories of an act of control. *Psychol. Rev.*
- Luce, R. D. (1986). *Response Times: Their Role in Inferring Elementary Mental Organization*. New York, NY: Oxford University Press.

- Lunn, D., Jackson, C., Spiegelhalter, D. J., Best, N., and Thomas, A. (2012). *The BUGS Book: A Practical Introduction to Bayesian Analysis*. Boca Raton, FL: CRC Press.
- Matzke, D., Dolan, C. V., Logan, G. D., Brown, S. D., and Wagenmakers, E.-J. (2013). Bayesian parametric estimation of stop-signal reaction time distributions. *J. Exp. Psychol. Gen.* 142, 1047–1073. doi: 10.1037/a0030543
- Matzke, D., and Wagenmakers, E.-J. (2009). Psychological interpretation of the ex-Gaussian and shifted Wald parameters: a diffusion model analysis. *Psychon. Bull. Rev.* 16, 798–817. doi: 10.3758/PBR.16.5.798
- Oosterlaan, J., Logan, G. D., and Sergeant, J. A. (1998). Response inhibition in AD/HD, CD, comorbid AD/HD+ CD, anxious, and control children: a meta-analysis of studies with the stop task. *J. Child Psychol. Psychiatry* 39, 411–425. doi: 10.1111/1469-7610.00336
- Patil, A., Huard, D., and Fonnesbeck, C. J. (2010). PyMC: Bayesian stochastic modelling in Python. *J. Stat. Softw.* 35, 1–81.
- R Core Team (2012). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing. ISBN 3-900051-07-0.
- Ratcliff, R. (1978). A theory of memory retrieval. *Psychol. Rev.* 85, 59–108. doi: 10.1037/0033-295X.85.2.59
- Ratcliff, R. (1993). A method for dealing with reaction time outliers. *Psychol. Bull.* 114, 510–532. doi: 10.1037/0033-2909.114.3.510
- Ratcliff, R., and Murdock, B. B. (1976). Retrieval processes in recognition memory. *Psychol. Rev.* 83, 190–214. doi: 10.1037/0033-295X.83.3.190
- Ridderinkhof, R., Band, G. P. H., and Logan, G. D. (1999). A study of adaptive behavior: effects of age and irrelevant information on the ability to inhibit one's actions. *Acta Psychol.* 101, 315–337. doi: 10.1016/S0001-6918(99)00010-4
- Rouder, J. N. (2005). Are unshifted distributional models appropriate for response time? *Psychometrika* 70, 377–381. doi: 10.1007/s11336-005-1297-7
- Rouder, J. N., Lu, J., Speckman, P., Sun, D., and Jiang, Y. (2005). A hierarchical model for estimating response time distributions. *Psychon. Bull. Rev.* 12, 195–223. doi: 10.3758/BF03257252
- Schachar, R., and Logan, G. D. (1990). Impulsivity and inhibitory control in normal development and childhood psychopathology. *Dev. Psychol.* 26, 710–720. doi: 10.1037/0012-1649.26.5.710
- Schachar, R., Mota, V. L., Logan, G. D., Tannock, R., and Klim, P. (2000). Confirmation of an inhibitory control deficit in attention-deficit/hyperactivity disorder. *J. Abnorm. Child Psychol.* 28, 227–235. doi: 10.1023/A:1005140103162
- Schwarz, W. (2001). The ex-wald distribution as a descriptive model of response times. *Behav. Res. Meth.* 33, 457–469. doi: 10.3758/BF03195403
- Spiegelhalter, D. J., Best, N. G., Carlin, B. P., and van der Linde, A. (2002). Bayesian measures of model complexity and fit (with discussions). *J. R. Stat. Soc. B* 64, 583–616. doi: 10.1111/1467-9868.00353
- Tierney, L. (1994). Markov chains for exploring posterior distributions. *Ann. Stat.* 22, 1701–1728. doi: 10.1214/aos/1176325750
- Verbruggen, F., Chambers, C. D., and Logan, G. D. (2009). Fictitious inhibitory differences: how skewness and slowing distort the estimation of stopping latencies. *Psychol. Sci.* 24, 352–362. doi: 10.1177/0956797612457390
- Verbruggen, F., and Logan, G. D. (2009). Models of response inhibition in the stop-signal and stop-change paradigms. *Neurosci. Biobehav. Rev.* 33, 647–661. doi: 10.1016/j.neubiorev.2008.08.014
- Wiecki, T. V., Sofer, I., and Frank, M. J. (2013). HDDM: hierarchical Bayesian estimation of the Drift-Diffusion model in Python. *Front. Neuroinform.* 7:14. doi: 10.3389/fninf.2013.00014
- Williams, B. R., Ponesse, J. S., Schachar, R. J., Logan, G. D., and Tannock, R. (1999). Development of inhibitory control across the life span. *Dev. Psychol.* 35, 205–213. doi: 10.1037/0012-1649.35.1.205

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 23 September 2013; paper pending published: 29 October 2013; accepted: 19 November 2013; published online: 10 December 2013.

Citation: Matzke D, Love J, Wiecki TV, Brown SD, Logan GD and Wagenmakers E-J (2013) Release the BEESTS: Bayesian Estimation of Ex-Gaussian STop-Signal reaction time distributions. *Front. Psychol.* 4:918. doi: 10.3389/fpsyg.2013.00918  
This article was submitted to *Quantitative Psychology and Measurement*, a section of the journal *Frontiers in Psychology*.

Copyright © 2013 Matzke, Love, Wiecki, Brown, Logan and Wagenmakers. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

## A. APPENDIX

### A.1 PRIOR DISTRIBUTION OF THE MODEL PARAMETERS

This appendix presents the prior distributions of the model parameters in the BEESTS implementation of the Bayesian parametric approach (BPA). The name of each parameter as shown in the BEESTS output is in brackets.

#### A.1.1 Individual BPA

The priors for the go and stop parameters are uniform distributions, spanning a plausible but wide range of values. BEESTS relies on slightly more diffuse priors than the WinBUGS implementation of the BPA (see Matzke et al., 2013):

$$\begin{aligned}
 \mu_{\text{go}} (\text{mu\_go}) &\sim \text{Uniform}(0.001, 1000) \\
 \sigma_{\text{go}} (\text{sigma\_go}) &\sim \text{Uniform}(1, 500) \\
 \tau_{\text{go}} (\text{tau\_go}) &\sim \text{Uniform}(1, 500) \\
 \mu_{\text{stop}} (\text{mu\_stop}) &\sim \text{Uniform}(0.001, 600) \\
 \sigma_{\text{stop}} (\text{sigma\_stop}) &\sim \text{Uniform}(1, 350) \\
 \tau_{\text{stop}} (\text{tau\_stop}) &\sim \text{Uniform}(1, 350).
 \end{aligned}
 \tag{A1}$$

#### A.1.2 Hierarchical BPA

**A.1.2.1 Individual parameters.** The hierarchical BPA assumes that the  $\mu_{\text{go}}$ ,  $\sigma_{\text{go}}$ ,  $\tau_{\text{go}}$ ,  $\mu_{\text{stop}}$ ,  $\sigma_{\text{stop}}$ , and  $\tau_{\text{stop}}$  parameters of each participant  $j = 1, \dots, J$  come from truncated normal group-level distributions. The group-level distributions are themselves characterized by a group mean ( $\mu$ ) and a group standard deviation ( $\sigma$ ) parameter. The WinBUGS implementation relies on normal group-level distributions that are truncated only at the lower end, whereas BEESTS uses normal distributions that are truncated at the lower *and* the upper ends:

$$\begin{aligned}
 \mu_{\text{go}_j} (\text{mu\_go.subj}) &\sim \text{Normal}(\mu_{\mu_{\text{go}}}, \sigma_{\mu_{\text{go}}})[0.001, 1000] \\
 \sigma_{\text{go}_j} (\text{sigma\_go.subj}) &\sim \text{Normal}(\mu_{\sigma_{\text{go}}}, \sigma_{\sigma_{\text{go}}})[1, 500] \\
 \tau_{\text{go}_j} (\text{tau\_go.subj}) &\sim \text{Normal}(\mu_{\tau_{\text{go}}}, \sigma_{\tau_{\text{go}}})[1, 500] \\
 \mu_{\text{stop}_j} (\text{mu\_stop.subj}) &\sim \text{Normal}(\mu_{\mu_{\text{stop}}}, \sigma_{\mu_{\text{stop}}})[0.001, 600] \\
 \sigma_{\text{stop}_j} (\text{sigma\_stop.subj}) &\sim \text{Normal}(\mu_{\sigma_{\text{stop}}}, \sigma_{\sigma_{\text{stop}}})[1, 350] \\
 \tau_{\text{stop}_j} (\text{tau\_stop.subj}) &\sim \text{Normal}(\mu_{\tau_{\text{stop}}}, \sigma_{\tau_{\text{stop}}})[1, 350].
 \end{aligned}
 \tag{A2}$$

**A.1.2.2 Group-level parameters.** The priors for the group mean and group standard deviations are uniform distributions. Note that the WinBUGS implementation uses censored normal priors for the group-level means and relies on slightly less diffuse priors for the group-level standard deviations than BEESTS:

$$\begin{aligned}
 \mu_{\mu_{\text{go}}} (\text{mu\_go}) &\sim \text{Uniform}(0.001, 1000) \\
 \sigma_{\mu_{\text{go}}} (\text{mu\_go\_var}) &\sim \text{Uniform}(0.01, 300) \\
 \mu_{\sigma_{\text{go}}} (\text{sigma\_go}) &\sim \text{Uniform}(1, 500) \\
 \sigma_{\sigma_{\text{go}}} (\text{sigma\_go\_var}) &\sim \text{Uniform}(0.01, 200) \\
 \mu_{\tau_{\text{go}}} (\text{tau\_go}) &\sim \text{Uniform}(1, 500) \\
 \sigma_{\tau_{\text{go}}} (\text{tau\_go\_var}) &\sim \text{Uniform}(0.01, 200) \\
 \mu_{\mu_{\text{stop}}} (\text{mu\_stop}) &\sim \text{Uniform}(0.001, 600) \\
 \sigma_{\mu_{\text{stop}}} (\text{mu\_stop\_var}) &\sim \text{Uniform}(0.01, 300) \\
 \mu_{\sigma_{\text{stop}}} (\text{sigma\_stop}) &\sim \text{Uniform}(1, 350) \\
 \sigma_{\sigma_{\text{stop}}} (\text{sigma\_stop\_var}) &\sim \text{Uniform}(0.01, 200) \\
 \mu_{\tau_{\text{stop}}} (\text{tau\_stop}) &\sim \text{Uniform}(1, 350) \\
 \sigma_{\tau_{\text{stop}}} (\text{tau\_stop\_var}) &\sim \text{Uniform}(0.01, 200).
 \end{aligned}
 \tag{A3}$$