



UvA-DARE (Digital Academic Repository)

Bayesian explorations in mathematical psychology

Matzke, D.

Publication date

2014

Document Version

Final published version

[Link to publication](#)

Citation for published version (APA):

Matzke, D. (2014). *Bayesian explorations in mathematical psychology*. [Thesis, fully internal, Universiteit van Amsterdam].

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, P.O. Box 19185, 1000 GD Amsterdam, The Netherlands. You will be contacted as soon as possible.

Bayesian Parametric Estimation of Stop-Signal Reaction Time Distributions

This chapter has been published as:
Dora Matzke, Conor V. Dolan, Gordon D. Logan, Scott D. Brown. and Eric-Jan Wagenmakers
(2013).

Bayesian parametric estimation of stop-signal reaction time distributions.
Journal of Experimental Psychology: General, 142, 1047-1073, doi: 10.1037/a0030543.¹

Abstract

The cognitive concept of response inhibition can be measured with the stop-signal paradigm. In this paradigm, participants perform a two-choice response time (RT) task where, on some of the trials, the primary task is interrupted by a stop signal that prompts participants to withhold their response. The dependent variable of interest is the latency of the unobservable stop response (stop-signal reaction time or SSRT). Based on the horse race model (Logan & Cowan, 1984), several methods have been developed to estimate SSRTs. None of these approaches allow for the accurate estimation of the entire distribution of SSRTs. Here we introduce a Bayesian parametric approach that addresses this limitation. Our method is based on the assumptions of the horse race model and rests on the concept of censored distributions. We treat response inhibition as a censoring mechanism, where the distribution of RTs on the primary task (go RTs) is censored by the distribution of SSRTs. The method assumes that go RTs and SSRTs are ex-Gaussian distributed and uses Markov chain Monte Carlo sampling to obtain posterior distributions for the model parameters. The method can be applied to individual as well as hierarchical data structures. We present the results of a number of parameter recovery and robustness studies and apply our approach to published data from a stop-signal experiment.

¹This chapter may not exactly replicate the final version published in the *Journal of Experimental Psychology: General*. It is not the copy of record.

3.1 Introduction

The stop-signal task (Lappin & Eriksen, 1966; Logan & Cowan, 1984) is frequently used to investigate response inhibition. Response inhibition refers to the ability to stop an ongoing action that is no longer appropriate: for example, driving your car and rapidly hitting the break when you notice that the traffic light turned red. The stop-signal paradigm can be used to investigate the operation of such simple type of inhibitory control in a carefully controlled laboratory setting.

In the standard stop-signal paradigm, participants perform a two-choice response time (RT) task, such as responding to the orientation of the visually presented stimuli. On some of the trials, this primary task is interrupted by an auditory stop signal that prompts participants to withhold their response on that trial. One of the primary dependent variables is the time required to inhibit the ongoing response (stop-signal RT [SSRT]). However, unlike the latency of the overt primary response, SSRTs cannot be observed directly.

To formally account for performance in the stop-signal paradigm, Logan (1981) and Logan and Cowan (1984) proposed that response inhibition can be viewed as a horse race between two competing processes: a go process that is set into motion by the primary task and a stop process that is initiated by the stop signal. If the go process wins the race, the primary response is executed; if the stop process wins the race, the primary response is successfully inhibited.

Since its development, the horse race model (Logan, 1981; Logan & Cowan, 1984) has successfully accounted for stop signal data in various settings and has facilitated the interpretation of numerous stopping experiments. For instance, the stop signal task has been used extensively to investigate response inhibition in different age groups (e.g., Kramer, Humphrey, Larish, Logan, & Strayer, 1994; Ridderinkhof, Band, & Logan, 1999; Schachar & Logan, 1990; Williams, Ponesse, Schachar, Logan, & Tannock, 1999) and clinical populations, such as children with Attention Deficit Hyperactivity Disorder (ADHD; Oosterlaan, Logan, & Sergeant, 1998; Schachar & Logan, 1990; Schachar, Mota, Logan, Tannock, & Klim, 2000).

The horse race model owes its popularity to the ability to quantify the otherwise unobservable latency of stopping. Various methods are available to estimate SSRTs. The standard analysis methods for the horse race model only yield a summary measure of the latency of inhibition, such as the mean SSRT; they do not reveal the shape of the entire SSRT distribution. It is well known that important features of the data may be missed in focusing only on the mean (e.g., Heathcote et al., 1991). A growing number of researchers therefore rely on distributional models, like the ex-Gaussian distribution (e.g., Balota & Yap, 2011; Matzke & Wagenmakers, 2009) to estimate the shape of entire RT distributions. For instance, Leth-Steensen et al. (2000) reported that children with ADHD differed from age-matched controls only in the ex-Gaussian parameter that quantifies the tail (i.e., very long RTs) of the RT distribution. Similarly, the RT distribution of schizophrenia patients is more variable and follows a markedly different shape than the RT distribution of controls, without necessarily differing in the mean (Belin & Rubin, 1995).

In the context of the stop-signal paradigm, focusing only on mean SSRT may likewise mask crucial features of the data and result in erroneous conclusions about the nature of response inhibition. Consider, for instance, the two SSRT distributions shown in Figure 3.1. The distributions have the same mean, but have clearly different shapes. The distribution drawn in solid line is more peaked, whereas the distribution drawn in dashed line is more spread out, with a faster leading edge and a longer tail. Ignoring such differences in the shapes of SSRT distributions may lead to the incorrect conclusion that two clinical groups or experimental conditions do not differ in SSRT. Unfortunately, the existing methods for obtaining SSRT estimates do not enable researchers to accurately estimate and evaluate differences in the shape of SSRT distributions.

The goal of this article therefore is to introduce a method that allows for the estimation of

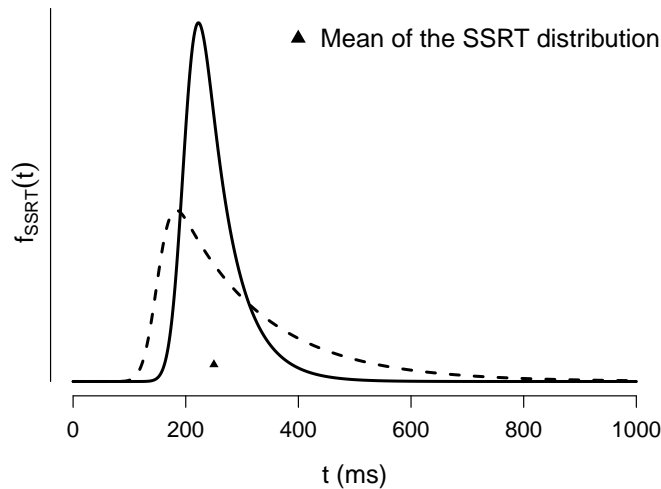


Figure 3.1 *Examples of stop-signal reaction time (SSRT) distributions with synthetic data.* The solid line shows an SSRT distribution with a slow leading edge and a short tail. The dashed line shows an SSRT distribution with a fast leading edge and a long tail. Despite the differences in their shapes, the means (i.e., black triangle) of the two distributions are equal.

the entire distribution of SSRTs, such as those shown in Figure 3.1. Our approach is based on the assumptions of the horse race model. The new method rests on the concept of censored distributions, where response inhibition is treated as a mechanism for censoring observed RTs. In order to quantify the shape of the distributions, the method assumes that the go RTs and SSRTs follow a parametric form, namely an ex-Gaussian distribution. Note, however, that our method does not hinge on this choice of parametric form; almost any other choice of distribution would do just as well. The ex-Gaussian distribution is purely used as a convenient choice to summarize the go RTs and the SSRTs. The ex-Gaussian is a commonly used distributional model, and it typically produces excellent fit to empirical RT distributions. (Heathcote et al., 1991; Hockley, 1982, 1984; Ratcliff, 1978, 1993; Ratcliff & Murdock, 1976). Our approach relies on Markov chain Monte Carlo (MCMC) sampling (Gamerman & Lopes, 2006; Gilks et al., 1996) and calculates posterior distributions for the model parameters.

An important advance of our Bayesian parametric method is that it makes it relatively easy to conduct both individual and hierarchical analyses. In individual analysis, the parameters of the SSRT distribution are estimated separately for each participant. In contrast, the hierarchical analysis (e.g., Gelman & Hill, 2007) recognizes that participants share some similarities and uses information available from the entire group to improve parameter estimation for the individual participants. The hierarchical approach has the potential to provide accurate parameter estimates with relatively few observations. Hierarchical modeling is therefore especially valuable in developmental and clinical stop-signal studies that typically use a very small number of trials per participant.

The outline of the article is as follows. In the first section, we describe the stop-signal paradigm

in more detail and discuss existing methods for estimating SSRTs. In the second section, we introduce the individual and the hierarchical Bayesian parametric approach to the estimation of SSRT distributions. In the third section, we report the results of various parameter recovery studies and show that our method accurately recovered the parameters of the generating SSRT distributions. In the fourth section, we apply the Bayesian parametric approach to an existing stop-signal data set. The fifth section concludes our investigation.

3.2 The Stop-Signal Paradigm

In the standard stop-signal paradigm (Lappin & Eriksen, 1966; Logan & Cowan, 1984), participants perform a two-choice RT task (i.e., the go task), such as responding to the orientation of the visually presented stimuli (e.g., press the right button for a right-pointing arrow and press the left button for a left-pointing arrow). Occasionally, the go stimulus is followed by an auditory stop signal (e.g., a high-pitched tone) that prompts participants to withhold their response on that trial. Typically, the stop signal is presented on a random 25-30% of the trials. The probability of successful inhibition can be experimentally manipulated by varying the time interval between the onset of the go stimulus and the onset of the stop signal (i.e., stop-signal delay [SSD]). The shorter the SSD, the more likely participants are to inhibit their response to the go stimulus.

To facilitate the interpretation of stop-signal data, Logan (1981) and Logan and Cowan (1984) introduced the horse race model. The horse race model conceptualizes response inhibition as a horse race between a go and a stop process. If the go process finishes before the stop process, the response is an error of commission. If the stop process finishes before the go process, the response is successfully inhibited. According to the horse race model, response inhibition is thus determined by the relative finishing times of the go and the stop process. Figure 3.2 illustrates how the probability of responding to the go stimulus (i.e., gray area) and the probability of inhibiting the response to the go stimulus (i.e., white area) are determined by the SSD, the SSRT, and the go RT distribution. Go RTs that are longer than $SSD + SSRT$ are successfully inhibited. In contrast, go RTs that are shorter than $SSD + SSRT$ cannot be inhibited and result in signal-respond RTs.

The standard horse race model depicted in Figure 3.2 assumes that, conditional on SSD, SSRT is constant (Logan & Cowan, 1984). This assumption is implausible, as SSRTs are certainly variable. Also, estimated SSRTs tend to decrease as SSD increases, a common finding that is explained in terms of the variability in SSRT. At short SSDs, almost all SSRTs are fast enough to win the race against the go RTs. The estimated mean SSRT therefore closely approximates the mean of the entire SSRT distribution. At long SSDs, only very fast SSRTs can win the race against the go RTs. The estimated mean SSRT is therefore lower than the mean of the entire SSRT distribution. As a result, SSRT estimates are longer at short SSDs than at long SSDs (de Jong, Coles, Logan, & Gratton, 1990; Logan & Burkell, 1986; Logan & Cowan, 1984).

To account for variability in SSRT, Logan and Cowan (1984) introduced the complete version of the horse race model. The complete race model treats both go RTs and SSRTs as independent random variables. To formalize the model, Logan and Cowan made the following simplifying assumptions about the independence of the go and the stop process. According to the context independence assumption, the distribution of go RTs is the same for go trials and for stop-signal trials. According to the stochastic independence assumption, the finishing times of the go and the stop process are uncorrelated. These two independence assumptions allow one to treat the go RT distribution on go trials as the underlying distribution of go RTs on stop-signal trials.

The formulation of the complete race model is closely connected to the concept of inhibition functions: functions that describe the relationship between the $P(\text{respond} \mid \text{stop signal})$ and SSD.

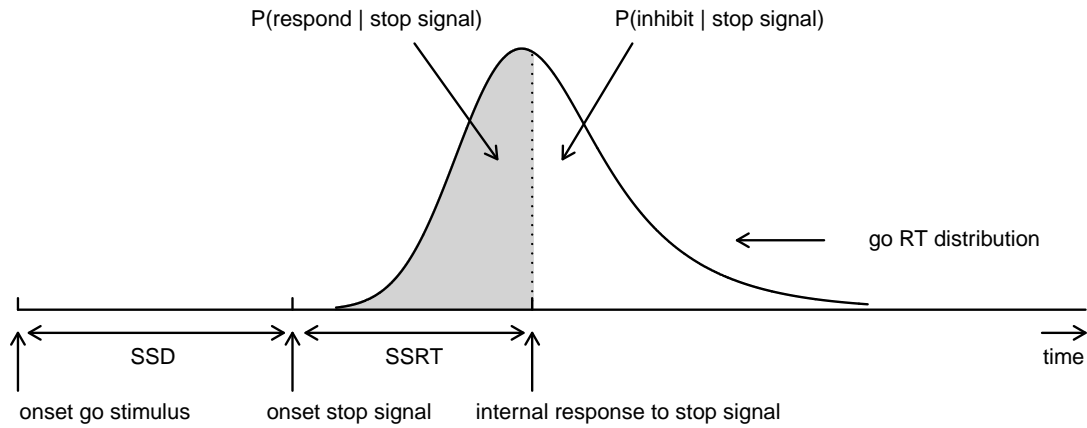


Figure 3.2 *Graphical representation of the horse race model.* RT = response time; SSD = stop-signal delay; SSRT = stop-signal reaction time.

As shown in Figure 3.3, the $P(\text{respond} \mid \text{stop signal})$ typically increases with increasing SSD. Logan and Cowan (1984) treated the inhibition function as a cumulative distribution and showed that its mean equals the difference between the mean go RT and the mean SSRT:

$$E(I) = E(\text{goRT}) - E(\text{SSRT}). \quad (3.1)$$

Further, they showed that the variance of the inhibition function equals the sum of the variances of the go RTs and the SSRTs:

$$\text{Var}(I) = \text{Var}(\text{goRT}) + \text{Var}(\text{SSRT}). \quad (3.2)$$

Note that the derivation of the complete horse race model is not based on any specific distribution shapes for the go RT and SSRT distributions.

Estimating SSRTs

One of the major advantages of the horse race model is that it allows for the estimation of the otherwise unobservable SSRT. Various methods are available for estimating SSRTs. The choice of method depends on the way SSDs are set in a particular experiment.

The SSD can be set according to the fixed-SSDs procedure or according to the staircase tracking procedure (e.g., Logan, 1994). The fixed-SSDs procedure requires a number of a priori chosen delays to be presented to the participants (e.g., SSDs of 80, 160, 240, 320, 400, and 480 ms; Logan & Burkell, 1986). Stop signals at the different SSDs are presented with equal frequencies at a random order. The challenge is to find a set of SSDs that span the entire range of the inhibition function. For the fixed-SSDs procedure, the integration method (Logan, 1981; Logan & Cowan, 1984) is the most popular approach to estimate SSRTs. The integration method assumes that SSRT is constant. SSRTs are estimated from the observed go RT distribution and the $P(\text{respond} \mid \text{stop}$

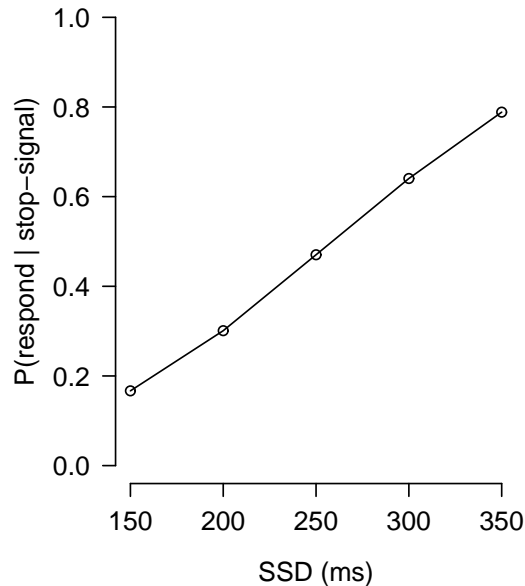


Figure 3.3 *Example of an inhibition function based on synthetic data from the recovery study for the individual Bayesian parametric approach.* The figure shows how the probability of responding on a stop-signal trial increases with increasing stop-signal delay (SSD).

signal) by finding the point (i.e., $SSRT + SSD$) at which the integral of the go RT distribution equals the $P(\text{respond} \mid \text{stop signal})$,

$$P(\text{respond} \mid \text{stop signal}) = \int_{-\infty}^{SSRT+SSD} f_{go}(t) dt. \quad (3.3)$$

In terms of Figure 3.2, the integration method involves deriving the time point at which the internal response to the stop signal occurs and subtracting SSD to obtain the SSRT. In practice, the following procedure is used: Go RTs are collapsed into a single distribution and are rank ordered. Subsequently, the n th go RT is selected, where n is obtained by multiplying the number of go RTs by the $P(\text{respond} \mid \text{stop signal})$ at a given SSD. Lastly, the SSD is subtracted to arrive at the SSRT. The integration method yields SSRT estimates for each SSD. As estimated SSRTs tend to decrease with increasing SSD (Logan & Burkell, 1986; Logan & Cowan, 1984), SSRTs at different SSDs are often averaged to yield a summary score for each participant.

The integration method has several drawbacks. It assumes that SSRT is constant, an assumption that is certainly incorrect. Moreover, the integration method requires a relatively large number of observations to produce accurate estimates of average SSRT. Researchers are advised to present participants with at least 900 go trials and 60 stop-signal trials on each of five different SSDs (Band, van der Molen, & Logan, 2003).

The second method for presenting SSDs, the staircase tracking procedure, sets SSDs dynamically, contingent on participants' performance. A typical staircase procedure will increase SSD by, say, 50 ms after successful inhibition, and decrease SSD by 50 ms after unsuccessful inhibi-

tion (see e.g., Bissett & Logan, 2011; Logan, Schachar, & Tannock, 1997; Osman, Kornblum, & Meyer, 1986; Verbruggen, Logan, & Stevens, 2008). This tracking procedure results in an overall $P(\text{respond} \mid \text{stop signal})$ of 0.50 for each participant.

For the staircase tracking procedure, the mean method is the easiest approach to estimate SSRTs. The mean method originates from Logan and Cowan’s (1984) treatment of SSRT as a random variable and is based on the following relationship:

$$E(\text{SSRT}) = E(\text{goRT}) - E(\text{I}). \quad (3.4)$$

Mean SSRT is thus given by the difference between the mean go RT and the mean of the inhibition function. Several approaches are available to compute the mean of the inhibition function (see, e.g., Logan, 1994; Logan & Cowan, 1984). The simplest way is to exploit the fact that when the staircase tracking procedure yields an overall $P(\text{respond} \mid \text{stop signal})$ of 0.50, the mean of the inhibition function equals the mean of the SSDs. As shown in Equation 3.4, the mean SSRT can be obtained by subtracting the mean SSD from the mean of the go RTs (Logan & Cowan, 1984; Logan et al., 1997).

The mean method can be used with a relatively small amount of data. Stop-signal experiments with healthy young adults typically include a total of 500-1,000 trials. Developmental and clinical studies generally include 250-500 trials, but investigations with as few as 100-250 trials are also common. Note, however, that contrary to the integration method, the mean method cannot be used to calculate SSRTs for each SSD separately.

Several variants of the integration and the mean method are available for the fixed-SSDs as well as the staircase tracking procedure (for a summary, see Verbruggen & Logan, 2009). Band et al. (2003) used simulations to show that SSRT estimates for which the $P(\text{respond} \mid \text{stop signal})$ equals 0.50, such as the mean method, are the most reliable. The mean method therefore has become the dominant method for estimating SSRTs.

Estimating Variability in SSRT

Logan and Cowan’s (1984) treatment of SSRT as a random variable provides a method for estimating the variability in SSRT. Logan and Cowan showed that the variance of the inhibition function can be calculated from its slope at the median. Once the variance of the inhibition function is known, the variance of SSRTs can be obtained from Equation 3.2. Logan and Cowan’s method is based on the observation that in a symmetrical distribution, the variance is proportional to the slope of the cumulative distribution at the median. If we treat the inhibition function as a cumulative distribution and assume a particular parametric form, say normal, the slope of the inhibition function at the median is given by

$$B_{0.5} = \frac{1}{\sqrt{2\pi} \times \text{SD}(\text{I})}. \quad (3.5)$$

It then follows from Equation 3.2 that the variance of SSRTs can be obtained by

$$\text{Var}(\text{SSRT}) = \left(\frac{1}{B_{0.5}\sqrt{2\pi}} \right)^2 - \text{Var}(\text{goRT}). \quad (3.6)$$

In contrast to the generality of the horse race model, the Logan and Cowan method for estimating SSRT variability assumes a particular parametric form of the inhibition function. Most importantly, Band et al. (2003) showed with simulations that the Logan and Cowan method overestimates the true variability in SSRT.

An Existing Method for Estimating SSRT Distributions

Up to now, the only existing approach for estimating the entire distribution of SSRTs was developed by Colonius (1990, see also de Jong et al., 1990, p. 181). Colonius showed that the survival distribution of SSRTs can be recovered using the distribution of go RTs, the distribution of signal-respond RTs, and the $P(\text{respond} \mid \text{stop signal})$ at a given SSD. Formally,

$$P(\text{SSRT} + \text{SSD} > t \mid \text{SSD}) = P(\text{respond} \mid \text{stop signal}, \text{SSD}) \times \frac{f_{SR}(t \mid \text{SSD})}{f_{go}(t)}, \quad (3.7)$$

where $f_{go}(t)$ and $f_{SR}(t \mid \text{SSD})$ are the probability density functions of the go RTs and the signal-respond RTs, respectively. Colonius' method does not depend on the specific parameterization of the go RT and the signal-respond RT distributions. The densities $f_{go}(t)$ and $f_{SR}(t \mid \text{SSD})$ can be estimated with various nonparametric density estimation methods (e.g., Silverman, 1986). Once the survival distribution of SSRTs is obtained, measures of location (e.g., median) and dispersion (e.g., interquartile distance) can be calculated easily.

Although Colonius' (1990) method is straightforward and elegant, it requires a very large number of observations to perform adequately (Logan, 1994). Band et al. (2003) used simulations to show that the Colonius method underestimates SSRT and overestimates its variability. In our implementation, over 250,000 stop-signal trials per SSD were required to obtain relatively accurate estimates of SSRT distributions. Using a more realistic number of stop-signal trials (e.g., 200 per SSD) resulted in inaccurate estimates, especially in the tails of the SSRT distribution. These problems are typical of nonparametric methods that estimate distribution tails from data (Luce, 1986).

To summarize, the stop-signal paradigm offers various methods to estimate the otherwise unobservable latency of stopping. Most methods only provide a summary measure of SSRT and are unable to accurately estimate the variability in SSRT. The only existing method for estimating entire SSRT distributions requires an unrealistically large number of observations to produce accurate estimates, particularly in the tail of the SSRT distribution. In what follows, we present a novel approach that relies on a parametric assumption to quantify the shape of the go RT and the SSRT distributions. As a result, the new method can provide accurate estimates of SSRT distributions even with relatively few observations.

3.3 Bayesian Parametric Approach for the Estimation of SSRT Distributions

Here we introduce a novel approach that allows for the estimation of the entire distribution of SSRTs. The method assumes that the go RTs and SSRTs follow an ex-Gaussian distribution. The ex-Gaussian distribution is purely used as a convenient choice to describe the go RTs and SSRTs. The ex-Gaussian is a frequently used distributional model that typically produces excellent fit to empirical RT distributions (Heathcote et al., 1991; Hockley, 1982, 1984; Ratcliff, 1978, 1993; Ratcliff & Murdock, 1976). The new approach may be applied to individual as well as hierarchical data structures and relies on MCMC sampling to obtain estimates of the parameters of the ex-Gaussian SSRT distribution.

We first introduce the rationale behind the Bayesian parametric approach (BPA), with special focus on the ex-Gaussian distribution and the assumptions of the method. We then introduce the basic concepts of Bayesian parameter estimation. Lastly, we present the individual and hierarchical BPA models for estimating SSRT distributions.

Introducing the Bayesian Parametric Approach

Rationale

The BPA rests on the concept of right-censored distributions. In right-censored distributions, observations to the right of a cutoff point (i.e., the censoring point) are omitted, but the number of censored observations is known. Censoring is a type of missing data problem that is frequently encountered in survival analysis (e.g., Elandt-Johnson & Johnson, 1980). In most applications, the censoring point is known and the focus is on estimating the parameters of the censored distribution. For instance, imposed censoring has been considered as a method to accommodate outliers in estimating the parameters of RT distributions (Ulrich & Miller, 1994).

As shown in Figure 3.2, the estimation of SSRT using the standard horse race model with constant SSRT can be viewed as a censoring problem. Specifically, the signal-respond RT distribution (i.e., gray area) can be treated as a right-censored go RT distribution with a constant censoring point that is given by the finishing time of the stop process (i.e., $SSD + SSRT$). On a given SSD, go RTs that are shorter than the finishing time of the stop process are observed. In contrast, go RTs that are longer than the finishing time of the stop process are successfully inhibited and therefore cannot be observed. Note that contrary to typical censoring problems, the censoring point of the go RT distribution is unknown. The estimation of SSRT therefore involves estimating the censoring point of the go RT distribution.

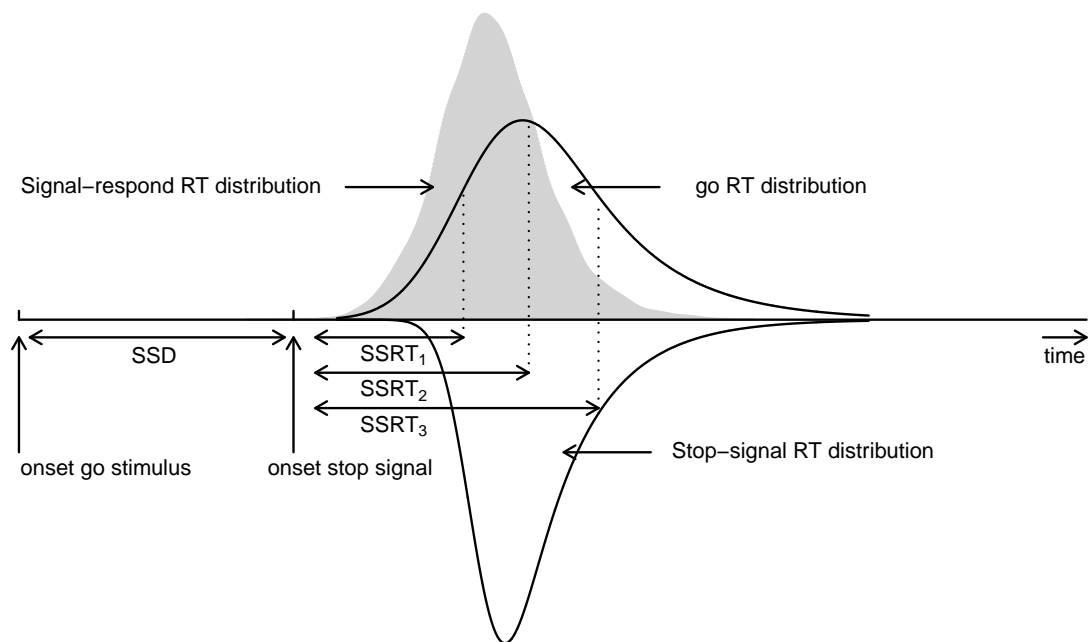


Figure 3.4 *Graphical representation of the complete horse race model.* RT = response time; SSD = stop-signal delay; SSRT = stop-signal reaction time.

The same reasoning can be extended to the estimation of the entire SSRT distribution using the complete horse race model. The censoring problem is, however, complicated by the fact that both go RTs and SSRTs are treated as random variables. As shown in Figure 3.4, the censoring

point on a given SSD takes on different values on each stop-signal trial (i.e., $SSD + SSRT_1$, $SSD + SSRT_2$, and $SSD + SSRT_3$). The signal-response RT distribution (i.e., gray area) can be viewed as a censored go RT distribution with censoring points drawn from the SSRT distribution that is shifted with the SSD on the time axis to longer RTs. The estimation of the SSRT distribution therefore involves estimating the finishing time distribution of the stop process that censors the go RT distribution.

The BPA is a parametric approach and as such involves choosing a parametric form for the go RT and the SSRT distribution. In what follows, we assume that go RTs and SSRTs—and therefore the finishing times of the stop process—are ex-Gaussian distributed and focus on simultaneously estimating the parameters of the two distributions.

Ex-Gaussian Distribution

The BPA assumes that the go RTs and the SSRTs are ex-Gaussian distributed. The ex-Gaussian distribution is given by the convolution of a Gaussian and an exponential distribution. The ex-Gaussian has three parameters. The μ and σ parameters give the mean and the standard deviation of the Gaussian component and reflect the leading edge and mode of the distribution. The τ parameter gives the mean of the exponential component and reflects the tail of the distribution.

The ex-Gaussian distribution has a positively skewed unimodal shape that typically fits empirical RT distributions well (Heathcote et al., 1991; Hockley, 1982, 1984; Ratcliff, 1978, 1993; Ratcliff & Murdock, 1976). Figure 3.5 shows changes in the ex-Gaussian distribution as a result of changes in the μ , σ and τ parameters. Increasing the μ parameter shifts the entire distribution to longer RTs and increases only the mean. Increasing σ influences the shape of the distribution and increases only the variance. Lastly, increasing τ influences both the location and the shape of the distribution and therefore increases both the mean and the variance (see Equation 3.11 and Equation 3.12).

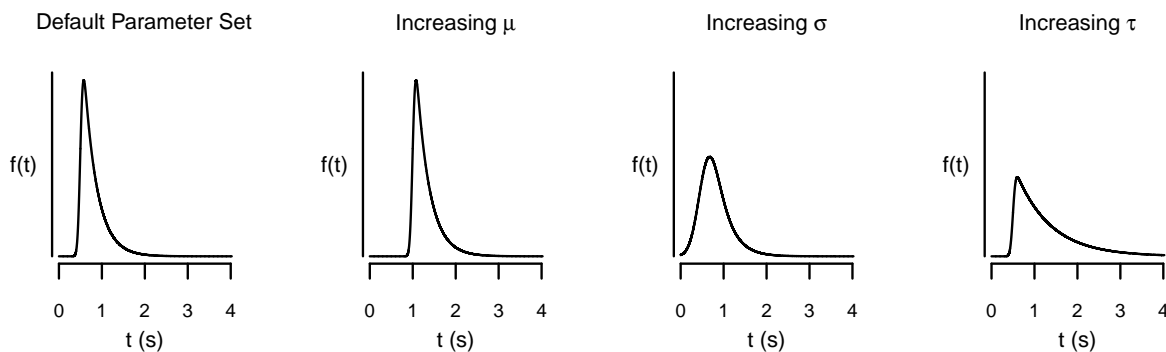


Figure 3.5 *Changes in the shape of the ex-Gaussian distribution as a result of changes in the ex-Gaussian parameters μ , σ , and τ . The parameter sets used to generate the distributions are $\mu = 0.5, \sigma = 0.05, \tau = 0.3$ (Panel 1); $\mu = 1, \sigma = 0.05, \tau = 0.3$ (Panel 2); $\mu = 0.5, \sigma = 0.2, \tau = 0.3$ (Panel 3); and $\mu = 0.5, \sigma = 0.05, \tau = 0.8$ (Panel 4).*

The probability density function of the ex-Gaussian is

$$f(t; \mu, \sigma, \tau) = \frac{1}{\tau} \exp\left(\frac{\mu - t}{\tau} + \frac{\sigma^2}{2\tau^2}\right) \Phi\left(\frac{t - \mu}{\sigma} - \frac{\sigma}{\tau}\right) \text{ for } \sigma > 0, \tau > 0, \quad (3.8)$$

where Φ is the standard normal distribution function, given by

$$\Phi\left(\frac{t - \mu}{\sigma} - \frac{\sigma}{\tau}\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{t - \mu}{\sigma} - \frac{\sigma}{\tau}} \exp\left(-\frac{y^2}{2}\right) dy. \quad (3.9)$$

The distribution function of the ex-Gaussian is

$$F(t; \mu, \sigma, \tau) = \Phi\left(\frac{t - \mu}{\sigma}\right) - \exp\left(\frac{\sigma^2}{2\tau^2} - \frac{t - \mu}{\tau}\right) \Phi\left(\frac{t - \mu}{\sigma} - \frac{\sigma}{\tau}\right), \quad (3.10)$$

and its mean and variance equal

$$E(t) = \mu + \tau \quad (3.11)$$

and

$$\text{Var}(t) = \sigma^2 + \tau^2, \quad (3.12)$$

respectively. Equation 3.11 and Equation 3.12 show how two SSRT distributions with the same mean or variance may have very different shapes, as illustrated in Figure 3.1.

We use the ex-Gaussian distribution purely as a descriptive tool to summarize the go RT and the SSRT distributions (see also Band et al., 2003; Heathcote et al., 1991; Ratcliff, 1978; Wagenmakers, van der Maas, et al., 2008). We do not assume that changes in the ex-Gaussian parameters map onto changes in specific cognitive processes (Matzke & Wagenmakers, 2009). Nevertheless, the ex-Gaussian can excellently accommodate the shape of RT distributions and is easy to fit to data. Moreover, as will be discussed later, sensitivity analyses indicated that the ex-Gaussian based BPA is robust to misspecification of the parametric form of the go RT and SSRT distributions. Note that other distributional assumptions can easily be made within our method.

Assumptions of the BPA

Similar to the complete horse race model, the BPA assumes that go RTs and SSRTs are independent random variables. The independence of the go and the stop process allows one to treat the go RT distribution on go trials as the underlying distribution of go RTs on stop-signal trials. The BPA assumes that the go RTs and the SSRTs follow ex-Gaussian distributions, with parameters μ_{go} , σ_{go} , and τ_{go} for the go RT distribution, and μ_{stop} , σ_{stop} , and τ_{stop} for the SSRT distribution. The log-likelihood of the $g = 1, \dots, G$ go RTs is given by

$$\ln L(\mu_{go}, \sigma_{go}, \tau_{go})_{go} = \sum_{g=1}^G \ln f_{go}(t_g; \mu_{go}, \sigma_{go}, \tau_{go}), \quad (3.13)$$

where $f_{go}(t; \mu_{go}, \sigma_{go}, \tau_{go})$ is the probability density of the ex-Gaussian go RT distribution given in Equation 3.8.

The log-likelihood of the data on the $s = 1, \dots, S$ stop-signal trials on a given SSD consists of the sum of the log-likelihoods of the $r = 1, \dots, R$ signal-respond RTs and the $i = 1, \dots, I$ successful inhibitions. According to the race model, signal-respond RTs are obtained on stop-signal trials

where the finishing time of the go process is shorter than the finishing time of the stop process (i.e., go RT < SSD + SSRT). The log-likelihood of a given signal-respond RT, t_r , can therefore be computed by (1) evaluating the probability density function of the go RT distribution at t_r and (2) evaluating the probability of obtaining an SSD + SSRT that is longer than t_r with the distribution function of the finishing time distribution of the stop process, that is, the distribution function of the SSRTs shifted with the SSD.

Similar reasoning can be extended to the log-likelihood of the successful inhibitions on signal-inhibit trials. According to the race model, successful inhibitions are obtained on stop-signal trials where the finishing time of the go process is longer than the finishing time of the stop process (i.e., go RT > SSRT + SSD). The log-likelihood of a given SSD + SSRT, t_i , can be computed by (1) evaluating the probability of obtaining a signal-respond RT that is longer than t_i with the distribution function of the go RT distribution and (2) evaluating the probability density function of the finishing time distribution of the stop process (i.e., SSRT distribution shifted with SSD) at t_i . Note, however, that SSRTs are by definition unobservable. Obtaining the log-likelihood on signal-inhibit trials therefore involves integrating out t_i from the go RT and the stop process finishing time distributions. Formally,

$$\begin{aligned} \ln L(\mu_{go}, \sigma_{go}, \tau_{go}, \mu_{stop}, \sigma_{stop}, \tau_{stop})_{stop} &= \\ &= \sum_{r=1}^R \left\{ \ln f_{go}(t_r; \mu_{go}, \sigma_{go}, \tau_{go}) + \ln \left[1 - F_{stop}(t_r; \mu_{stop}, \sigma_{stop}, \tau_{stop}, SSD) \right] \right\} \\ &+ \sum_{i=1}^I \ln \int_{-\infty}^{\infty} \left[1 - F_{go}(t_i; \mu_{go}, \sigma_{go}, \tau_{go}) \right] \times f_{stop}(t_i; \mu_{stop}, \sigma_{stop}, \tau_{stop}, SSD) dt, \end{aligned} \quad (3.14)$$

where $f_{go}(t; \mu_{go}, \sigma_{go}, \tau_{go})$ and $F_{go}(t; \mu_{go}, \sigma_{go}, \tau_{go})$ are the probability density and the distribution function of the ex-Gaussian go RT distribution given in Equation 3.8 and Equation 3.10, respectively. Similarly, $f_{stop}(t; \mu_{stop}, \sigma_{stop}, \tau_{stop}, SSD)$ and $F_{stop}(t; \mu_{stop}, \sigma_{stop}, \tau_{stop}, SSD)$ are the probability density and the distribution function of the ex-Gaussian finishing time distribution of the stop process, that is, the SSRT distribution shifted with the SSD.

The goal is to simultaneously estimate the μ_{go} , σ_{go} , and τ_{go} parameters of the go RT distribution and the μ_{stop} , σ_{stop} , and τ_{stop} parameters of the SSRT distribution. Parameter estimation may proceed by means of standard maximum likelihood estimation (Dolan, van der Maas, & Molenaar, 2002; Myung, 2003). However, the BPA is intended to handle individual as well as hierarchical data structures. Maximum likelihood estimation can become practically difficult for hierarchical problems, so we chose to use Bayesian parameter estimation instead. This also confers the typical benefits of Bayesian estimation, such as a coherent inferential framework.

Bayesian Parameter Estimation

In Bayesian parameter estimation, we start with a prior probability distribution for the parameter of interest. The prior distribution quantifies the existing knowledge about the parameter. The prior distribution is then updated by the incoming data (i.e., likelihood) to yield a posterior probability distribution under Bayes' rule:

$$\text{posterior} = \frac{\text{likelihood} \times \text{prior}}{\text{marginal likelihood}}. \quad (3.15)$$

The marginal likelihood is the probability of the observed data and does not involve the parameters of interest. Equation 3.15 can hence be expressed as

$$\text{posterior} \propto \text{likelihood} \times \text{prior}. \tag{3.16}$$

The top panels of Figure 3.6 illustrate the basic concepts of Bayesian estimation for the parameters of the SSRT distribution with simulated data from a synthetic participant. For each parameter, we start with a uniform prior distribution reflecting the assumption that all values of the parameter within some wide range are equally likely a priori. The prior distributions are then updated by the data to yield the posterior distributions. The posterior distributions quantify all the available information about the parameters. The central tendency of the posterior distribution can be expressed by its mean, median or mode. The central tendency of the posterior is often used as a point estimate of the parameter (e.g., with a uniform prior, the mode corresponds to the maximum-likelihood estimator). The dispersion of the posterior distribution can be quantified by the standard deviation or the percentiles. The dispersion of the posterior conveys important information about the precision of the parameter estimates: The larger the posterior standard deviation, the greater the uncertainty of the estimated parameter.

In many applications, the posterior distribution cannot be derived analytically. Fortunately, the posterior can be approximated using numerical sampling techniques such as MCMC sampling (Gelman & Lopes, 2006; Gilks et al., 1996). The BPA currently relies on WinBUGS (Bayesian inference Using Gibbs Sampling for Windows; Lunn et al., 2000; see Kruschke, 2010b for an introduction) to obtain the posterior distributions of the model parameters. WinBUGS is a general-purpose statistical software for Bayesian analysis that uses MCMC techniques to sample from the posterior distribution of the model parameters.

Figure 3.6 gives a simple illustration of Bayesian parameter estimation with MCMC sampling. The bottom panels show sequences of values (i.e., MCMC chains) sampled from the posterior distribution of the parameters of the SSRT distribution. More accurate sampling from the posterior distribution can be obtained by running multiple chains and discarding the beginning of each chain as burn-in. For each parameter, we ran three chains, each with different starting values (i.e., overdispersed starting values). The starting values were randomly generated from uniform distributions covering a wide range of possible parameter values. Per chain, we collected 2,000 iterations, resulting in a total of 6,000 samples from the posterior distributions. The chains converged successfully from the starting values to their stationary distributions; the individual chains look like “hairy caterpillars” and they seem identical to one another. Formal diagnostic measures of convergence are available. For instance, \hat{R} (Gelman & Rubin, 1992) 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. For the present example, \hat{R} was lower than 1.05 for all of the parameters.

The top panels of Figure 3.6 show histograms and density estimates of the posterior samples of the stop parameters. The histograms were plotted by collecting the sampled values across the three chains and projecting them onto the x -axis of the top panel figures. The median of the posterior distribution equals 186.80 for μ_{stop} , 32.76 for σ_{stop} , and 57.43 for τ_{stop} . The region extending from the 2.5th to the 97.5th percentile of the posterior distribution gives the so-called 95% Bayesian confidence interval. For example, the 95% Bayesian confidence interval for μ_{stop} ranges from 178.30 to 195.70, indicating that we can be 95% confident that the true value of μ_{stop} lies within this range. The Bayesian confidence interval is the narrowest for the μ_{stop} parameter, indicating that μ_{stop} is estimated the most precisely among the stop parameters.

3. BAYESIAN PARAMETRIC ESTIMATION OF STOP-SIGNAL REACTION TIME DISTRIBUTIONS

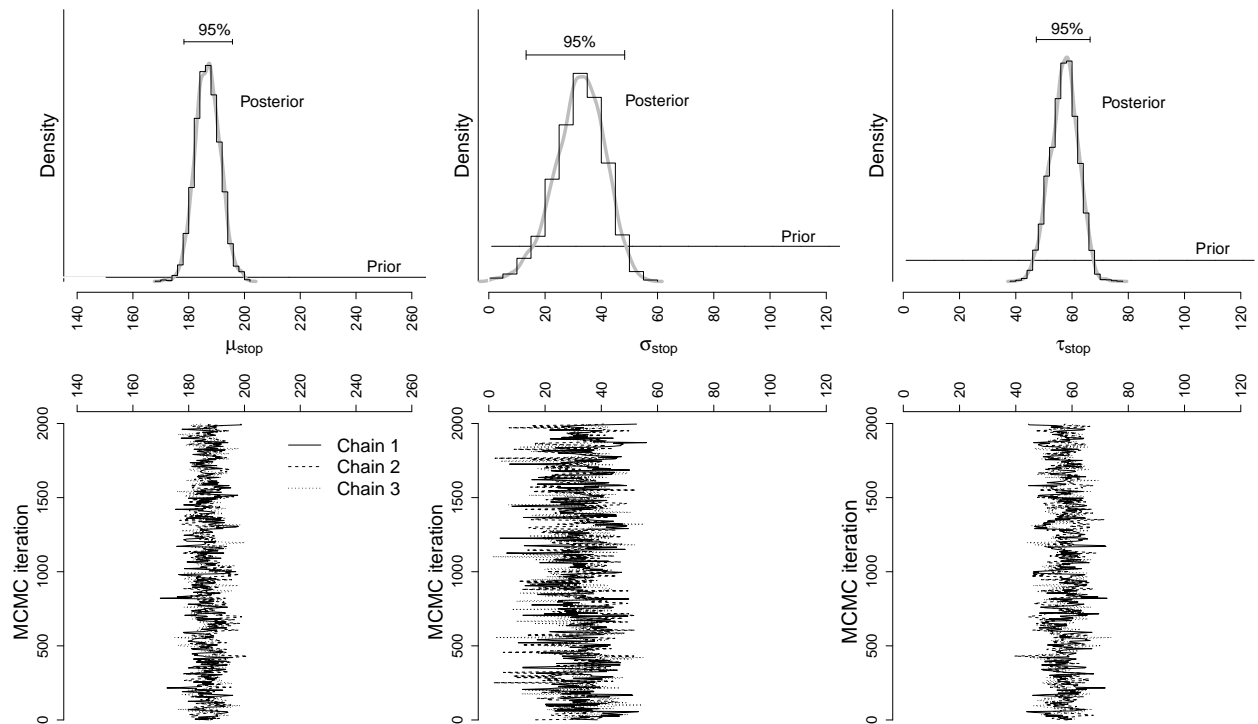


Figure 3.6 *Illustration of Markov chain Monte Carlo (MCMC)-based Bayesian estimation for the ex-Gaussian parameters μ_{stop} , σ_{stop} , and τ_{stop} for a synthetic data set with the individual Bayesian parametric approach.* The histograms in the top panels show the posterior distribution of the parameters. The corresponding thick gray lines indicate the fit of a nonparametric density estimator to the posterior samples. The horizontal black lines at the bottom show the prior distribution of the parameters. The horizontal black lines at the top show the 95% Bayesian confidence interval. The solid, dashed and dotted lines in the bottom panels represent the different sequences of values (i.e., MCMC chains) sampled from the posterior distribution of the parameters.

The Bayesian approach can be applied to hierarchical as well as individual data. In individual estimation, the parameters of the SSRT distribution are estimated separately for each participant. In the hierarchical approach (e.g., Gelman & Hill, 2007; M. D. Lee, 2011; Lindley & Smith, 1972; Rouder et al., 2005, 2003), the estimation of the individual stop parameters is supported by information from the entire group. In the next section, we introduce the individual and the hierarchical BPA models for estimating SSRT distributions.

Individual BPA

Figure 3.7 shows the graphical model for the individual BPA. Observed variables are represented by shaded nodes and unobserved variables are represented by unshaded nodes. The graph structure indicates dependencies between the nodes, and the plates represent independent replications of the different types of trials (e.g., M. D. Lee, 2008).

The individual BPA assumes that the $g = 1, \dots, G$ go RTs come from an ex-Gaussian distribution, with parameters μ_{go} , σ_{go} , and τ_{go} (see Equation 3.13). On the $s = 1, \dots, S$ stop-signal trials, the $r = 1, \dots, R$ signal-respond RTs (i.e., SR-RT) and the $i = 1, \dots, I$ successful inhibitions (i.e., NA)

3.3. Bayesian Parametric Approach for the Estimation of SSRT Distributions

Model parameters:

$$\mu_{go} \sim \text{Uniform}(1, 1000)$$

$$\sigma_{go} \sim \text{Uniform}(1, 300)$$

$$\tau_{go} \sim \text{Uniform}(1, 300)$$

$$\mu_{stop} \sim \text{Uniform}(1, 600)$$

$$\sigma_{stop} \sim \text{Uniform}(1, 250)$$

$$\tau_{stop} \sim \text{Uniform}(1, 250)$$

Data:

$$goRT_g \sim \text{ExGaussian}(\mu_{go}, \sigma_{go}, \tau_{go})$$

$$SR - RT_r \sim \text{CensoredExGaussian} - \text{SR}(\mu_{go}, \sigma_{go}, \tau_{go}, \mu_{stop}, \sigma_{stop}, \tau_{stop}, SSD_s)$$

$$NA_i \sim \text{CensoredExGaussian} - \text{I}(\mu_{go}, \sigma_{go}, \tau_{go}, \mu_{stop}, \sigma_{stop}, \tau_{stop}, SSD_s)$$

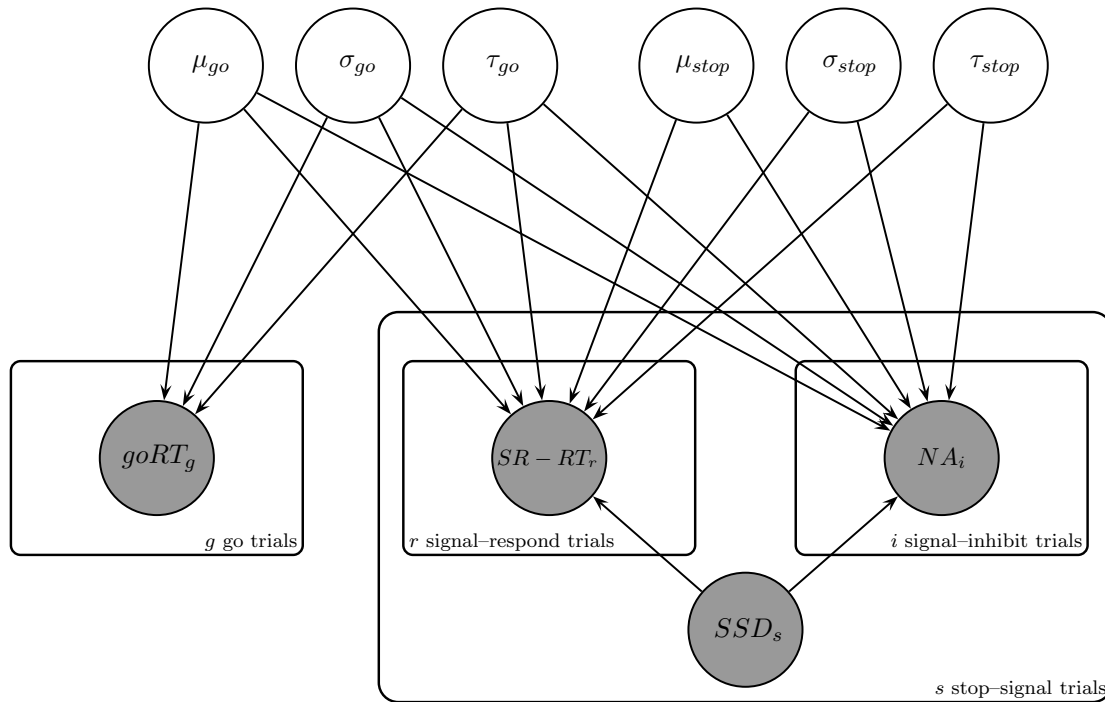


Figure 3.7 *Graphical model for the individual Bayesian parametric approach.* Observed variables are represented by shaded nodes and unobserved variables are represented by unshaded nodes. The plates represent independent replications of the different types of trials. The go response times (RTs) come from an ex-Gaussian distribution, with parameters μ_{go} , σ_{go} , and τ_{go} . The signal-respond RTs (i.e., SR-RT) and the successful inhibitions (i.e., NA) come from censored ex-Gaussian distributions, with parameters μ_{go} , σ_{go} , τ_{go} , μ_{stop} , σ_{stop} , τ_{stop} , and stop-signal delay (SSD). The priors for the model parameters are uniform distributions.

come from censored ex-Gaussian distributions, with parameters μ_{go} , σ_{go} , τ_{go} , μ_{stop} , σ_{stop} , τ_{stop} , and SSD_s (see Equation 3.14). The priors for the model parameters are uniform distributions, spanning a plausible but wide range of parameter values. The range of the uniform prior distributions is loosely based on the results of a life-span study of stop-signal performance reported in Williams et al. (1999) and the corresponding ex-Gaussian parameter values used in the simulation studies of Band et al. (2003).

The individual BPA makes no connections between participants; it assumes that they are

completely independent. The goal is to estimate the μ_{go} , σ_{go} , τ_{go} , μ_{stop} , σ_{stop} , and τ_{stop} parameters for each participant separately. To this end, we use WinBUGS to sample from the posterior distribution of the model parameters. The WinBUGS script for the individual BPA is available in Appendix B.1. The median of the posterior distributions can be used as a point estimate of the model parameters. SSRT distributions, such as those shown in Figure 3.1, can be obtained by evaluating the ex-Gaussian probability density function (Equation 3.8) with the posterior median of the parameters. The mean and the variance of the SSRT distribution can be computed from Equation 3.11 and Equation 3.12 with the posterior median of the parameters. Alternatively, we can quantify the uncertainty of the estimated SSRT distribution by drawing random parameter vectors from the joint posterior of the stop parameters and evaluating the ex-Gaussian probability density function using the chosen parameter vectors.

Hierarchical BPA

A particularly useful application of the Bayesian hierarchical approach (Farrell & Ludwig, 2008; Gelman & Hill, 2007; M. D. Lee, 2011; Nilsson, Rieskamp, & Wagenmakers, 2011; Rouder & Lu, 2005; Rouder et al., 2005, 2003; Shiffrin et al., 2008; Wagenmakers, Lodewyckx, Kuriyal, & Grasman, 2010) explicitly models individual differences in the parameter values, but at the same time recognizes that participants share some similarities. Hierarchical modeling is a compromise between the assumption that participants are completely independent (i.e., individual BPA) and the assumption that all participant are identical (Gelman & Hill, 2007). Rather than estimating the parameters separately for each individual, hierarchical modeling assumes that the individual parameters are drawn from group-level distributions. The group-level distributions specify how the individual parameters are distributed in the population and thus define the between-subject variability in the model parameters. The goal is to obtain individual parameter estimates as well as estimates for the parameters of the group-level distributions.

Hierarchical methods have the potential to provide more accurate and less variable parameter estimates than individual Bayesian and maximum likelihood estimation (Farrell & Ludwig, 2008; Rouder et al., 2005). The advantages of hierarchical modeling are the most pronounced in situations with only moderate between-subject variability and a small number of observations per participant (Gelman & Hill, 2007). The benefits of hierarchical modeling arise from using information available from the whole group to improve parameter estimation for the individual participants. Hierarchical modeling uses the group-level distributions as priors to adjust poorly estimated extreme parameter values to more moderate ones. As a result, outlying individual estimates —especially the ones that are estimated with a great degree of uncertainty— are “shrunk” towards the group mean. The hierarchical approach is especially valuable in situations with relatively few observations per participant, as is often the case in stop-signal experiments.

Figure 3.8 shows the graphical model for the hierarchical BPA. The hierarchical BPA assumes that the $g = 1, \dots, G$ go RTs of each participant, $j = 1, \dots, J$, come from ex-Gaussian distributions, but with different values of μ_{go} , σ_{go} , and τ_{go} . On the $s = 1, \dots, S$, stop-signal trials, the $r = 1, \dots, R$ signal-respond RTs (i.e., SR-RT) and the $i = 1, \dots, I$ successful inhibitions (i.e., NA) of each participant come from censored ex-Gaussian distributions, but again with different values of μ_{go} , σ_{go} , τ_{go} , μ_{stop} , σ_{stop} , τ_{stop} , and SSD_s . The individual μ_{go_j} , σ_{go_j} , τ_{go_j} , μ_{stop_j} , σ_{stop_j} , and τ_{stop_j} parameters are in turn assumed to come from truncated normal group-level distributions that are characterized by group-level parameters. For example, the μ_{stop} parameters codetermine the location of the individual SSRT distributions. As SSRTs are by definition positive, the μ_{stop} parameters must be positive as well. The μ_{stop} parameters are therefore assumed to come from a normal group-level distribution truncated at 0 ms, with mean $\mu_{\mu_{stop}}$ and standard deviation

$\sigma_{\mu_{stop}}$. Similarly, the σ_{stop} parameters are the standard deviations of the Gaussian component of the individual SSRT distributions and are by definition positive (see Equation 3.8). The σ_{stop} parameters are assumed to come from a normal group-level distribution truncated at one ms, with mean $\mu_{\sigma_{stop}}$ and standard deviation $\sigma_{\sigma_{stop}}$.²

The use of normal group-level distributions is a common choice in Bayesian hierarchical modeling (e.g., Gelman & Hill, 2007; M. D. Lee & Wagenmakers, 2013). In real data, however, the distribution of individual parameters may deviate from normality, especially in clinical populations. As will be described later, sensitivity analyses indicated that the hierarchical BPA is robust to misspecification of the group-level distribution of the individual go parameters. In contrast, the BPA with misspecified group-level distributions results in biased stop parameter estimates. Fortunately, the bias decreases substantially as the number of participants and especially as the number of trials increase. The reader is referred to the Discussion for some suggestions on examining the validity of the hierarchical assumptions of the BPA.

The priors for the mean and standard deviation of the group-level distributions are normal and uniform distributions, respectively. For example, the $\mu_{\sigma_{stop}}$ parameter is the mean of the group-level distribution of the individual σ_{stop} parameters and as such it must be positive. The group mean $\mu_{\sigma_{stop}}$ parameter is assumed to come from a normal distribution censored to be positive, with mean 40 and standard deviation $1/\sqrt{0.001}$. The group standard deviation $\sigma_{\sigma_{stop}}$ parameter is assumed to be uniformly distributed between 0 and 100. The parameters of the priors for the group-level means and standard deviations are loosely based on the results reported in Williams et al. (1999) and the corresponding ex-Gaussian parameter values used in Band et al. (2003).

In the hierarchical BPA, the goal is to estimate the group-level means and standard deviations as well as the individual go and stop parameters. The WinBUGS script for the hierarchical BPA is available in Appendix B.1. The median of the posterior distributions can be used as a point estimate for the parameters. The SSRT distribution of each participant can be obtained by evaluating the ex-Gaussian probability density function with the posterior median of the individual parameters. The mean and the variance of the individual SSRT distributions can be computed from Equation 3.11 and Equation 3.12 with the posterior median of the individual parameters. Also, we can quantify the uncertainty of the individual SSRT distributions by drawing random parameter vectors from the joint posterior of the individual stop parameters and evaluating the ex-Gaussian probability density function using the chosen parameter vectors.

3.4 Parameter Recovery Studies

Individual BPA

We conducted two simulation studies to investigate the ability of the individual BPA to recover underlying true parameter values. The first recovery study examined the asymptotic properties of the parameter estimates. The second recovery study investigated the number of stop-signal trials necessary to obtain accurate parameter estimates.

Methods

We generated stop-signal data from the horse race model, where the go RTs and the SSRTs were drawn from ex-Gaussian distributions, with parameters $\mu_{go} = 440$, $\sigma_{go} = 80$, $\tau_{go} = 60$, $\mu_{stop} = 190$,

²For computational reasons, the truncated normal group-level distributions of the σ_{go} , τ_{go} , σ_{stop} , and τ_{stop} parameters are truncated at 1 ms instead of 0.

3. BAYESIAN PARAMETRIC ESTIMATION OF STOP-SIGNAL REACTION TIME DISTRIBUTIONS

<i>Group-level parameters:</i>	<i>Individual parameters:</i>
$\mu_{\mu_{go}} \sim \text{Normal}(500, 0.0001)I[0, \infty]$	$\mu_{go_j} \sim \text{TruncatedNormal}(\mu_{\mu_{go}}, \frac{1}{\sigma_{\mu_{go}}^2}, 0)$
$\sigma_{\mu_{go}} \sim \text{Uniform}(0, 300)$	$\sigma_{go_j} \sim \text{TruncatedNormal}(\mu_{\sigma_{go}}, \frac{1}{\sigma_{\sigma_{go}}^2}, 1)$
$\mu_{\sigma_{go}} \sim \text{Normal}(100, 0.001)I[0, \infty]$	$\tau_{go_j} \sim \text{TruncatedNormal}(\mu_{\tau_{go}}, \frac{1}{\sigma_{\tau_{go}}^2}, 1)$
$\sigma_{\sigma_{go}} \sim \text{Uniform}(0, 200)$	$\mu_{\mu_{stop_j}} \sim \text{TruncatedNormal}(\mu_{\mu_{stop}}, \frac{1}{\sigma_{\mu_{stop}}^2}, 0)$
$\mu_{\tau_{go}} \sim \text{Normal}(80, 0.001)I[0, \infty]$	$\sigma_{\sigma_{stop_j}} \sim \text{TruncatedNormal}(\mu_{\sigma_{stop}}, \frac{1}{\sigma_{\sigma_{stop}}^2}, 1)$
$\sigma_{\tau_{go}} \sim \text{Uniform}(0, 200)$	$\tau_{\tau_{stop_j}} \sim \text{TruncatedNormal}(\mu_{\tau_{stop}}, \frac{1}{\sigma_{\tau_{stop}}^2}, 1)$
$\mu_{\mu_{stop}} \sim \text{Normal}(200, 0.0001)I[0, \infty]$	<i>Data:</i>
$\sigma_{\mu_{stop}} \sim \text{Uniform}(0, 200)$	$goRT_{gj} \sim \text{ExGaussian}(\mu_{go_j}, \sigma_{go_j}, \tau_{go_j})$
$\mu_{\sigma_{stop}} \sim \text{Normal}(40, 0.001)I[0, \infty]$	$SR - RT_{rj} \sim \text{CensoredExGaussian} - \text{SR}(\mu_{go_j}, \sigma_{go_j}, \tau_{go_j}, \mu_{stop_j}, \sigma_{stop_j}, \tau_{stop_j}, SSD_{sj})$
$\sigma_{\sigma_{stop}} \sim \text{Uniform}(0, 100)$	$NA_{ij} \sim \text{CensoredExGaussian} - \text{I}(\mu_{go_j}, \sigma_{go_j}, \tau_{go_j}, \mu_{stop_j}, \sigma_{stop_j}, \tau_{stop_j}, SSD_{sj})$
$\mu_{\tau_{stop}} \sim \text{Normal}(30, 0.001)I[0, \infty]$	
$\sigma_{\tau_{stop}} \sim \text{Uniform}(0, 100)$	

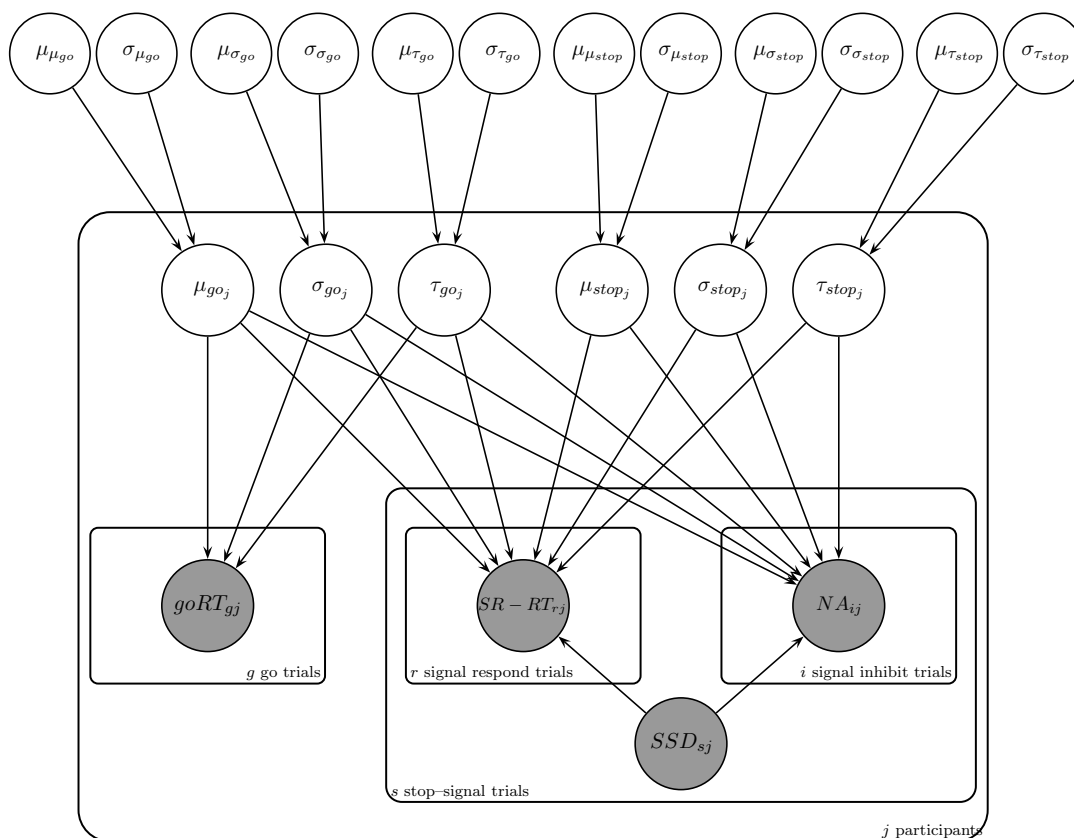


Figure 3.8 *Graphical model for the hierarchical Bayesian parametric approach.* The go response times (RTs) of each participant come from ex-Gaussian distributions, with different values of μ_{go} , σ_{go} , and τ_{go} . The signal-respond RTs (i.e., SR-RT) and the successful inhibitions (i.e., NA) of each participant come from censored ex-Gaussian distributions, with different values of μ_{go} , σ_{go} , τ_{go} , μ_{stop} , σ_{stop} , τ_{stop} , and stop-signal delay (SSD). The individual go and stop parameters come from truncated normal group-level distributions that are characterized by group-level parameters. In order to maintain consistency with the WinBUGS syntax, the group-level normal and truncated normal distributions are parameterized in terms of their precision (i.e., inverse variance) rather than their variance. The $I[0, \infty]$ construct denotes distributional censoring with lower bound equal to 0 and upper bound equal to infinity.

$\sigma_{stop} = 40$, and $\tau_{stop} = 50$.³ These parameters made the mean and the standard deviation of the go RT distribution 500 and 100 ms, respectively, and the mean and the standard deviation of the SSRT distribution 240 and 64 ms, respectively. The SSDs were set to 150, 200, 250, 300, and 350 ms. The above parameter values and SSDs resulted in $P(\text{respond} \mid \text{stop signal, SSD} = 150) = 0.17$, $P(\text{respond} \mid \text{stop signal, SSD} = 200) = 0.30$, $P(\text{respond} \mid \text{stop signal, SSD} = 250) = 0.47$, $P(\text{respond} \mid \text{stop signal, SSD} = 300) = 0.65$, and $P(\text{respond} \mid \text{stop signal, SSD} = 350) = 0.79$. The resulting inhibition function for a randomly chosen data set is shown in Figure 3.3.

In the first recovery study, we generated a single data set containing 200,000 go trials and 5 (SSD) \times 100,000 stop-signal trials. The estimated μ_{stop} , σ_{stop} , and τ_{stop} parameters were free to vary across the five SSDs. In the second recovery study, we conducted four sets of simulations that varied the number of trials, with 100 data sets for each set. For the first set, each data set contained 4,500 go trials and 5 (SSD) \times 300 stop-signal trials. For the second set, each data set contained 2,250 go trials and 5 \times 150 stop-signal trials. For the third set, each data set contained 750 go trials and 5 \times 50 stop-signal trials. For the fourth set, each data set contained 375 go trials and 5 \times 25 stop-signal trials. In contrast to the first recovery study, the estimated μ_{stop} , σ_{stop} , and τ_{stop} parameters were constrained to be equal across the five SSDs.

We fit the data sets with the individual BPA using WinBUGS. We ran three MCMC chains and used overdispersed starting values to confirm that the chains have converged to the stationary distribution ($\hat{R} \approx 1$). The first 500 samples of each MCMC chain were discarded. The reported parameter estimates are based on 3 \times 4,000 recorded samples.

Results

The parameters of the go RT distribution were excellently recovered in both recovery studies. As the go RT distribution is of little theoretical interest, the remainder of this section focuses exclusively on results related to the SSRT distribution.

The results of the first recovery study are shown in Figure 3.9. For all SSDs, the posterior median recovered the generating parameter values, and the mean and the standard deviation of the true SSRT distributions very well. Across the five SSDs, the posterior standard deviations ranged from 0.92 to 1.84 for μ_{stop} , from 1.50 to 2.2 for σ_{stop} , and from 0.91 to 2.09 for τ_{stop} . The posterior standard deviations were small, indicating that the parameters were estimated precisely. In contrast to the integration method, the mean SSRT estimated with the BPA did not decrease with increasing SSD. Theoretically, one may obtain accurate estimates for the stop parameters using stop-signal data on a single SSD.

The results of the second recovery study are shown in Figure 3.10 and Figure 3.11. Figure 3.10 shows the mean of the posterior medians of the stop parameters across the 100 replications and the mean and standard deviation of the SSRT distribution. Figure 3.11 shows the estimated SSRT distributions based on the posterior medians for the 100 replications. As shown in the figures, the BPA recovered the generating parameter values and the shape of the true SSRT distribution with little bias even with relatively few (i.e., 25) stop-signal trials per SSD. Naturally, as the number of trials increased, the bias, the standard error, and the posterior standard deviation of the estimates decreased. The mean posterior standard deviation of μ_{stop} across the 100 replications decreased from 27.40 for the simulation set with 25 stop-signal trials per SSD to 9.51 for the set with 300 stop-signal trials per SSD. The mean posterior standard deviation of σ_{stop} decreased from 26.05 to 12.51. The mean posterior standard deviation of τ_{stop} decreased from 26.64 to 9.98. Note also that the BPA parameter estimates as well as the average of the integration method estimates across the

³We conducted several recovery studies using alternative true parameter values. The results were essentially the same as the ones reported here.

3. BAYESIAN PARAMETRIC ESTIMATION OF STOP-SIGNAL REACTION TIME DISTRIBUTIONS

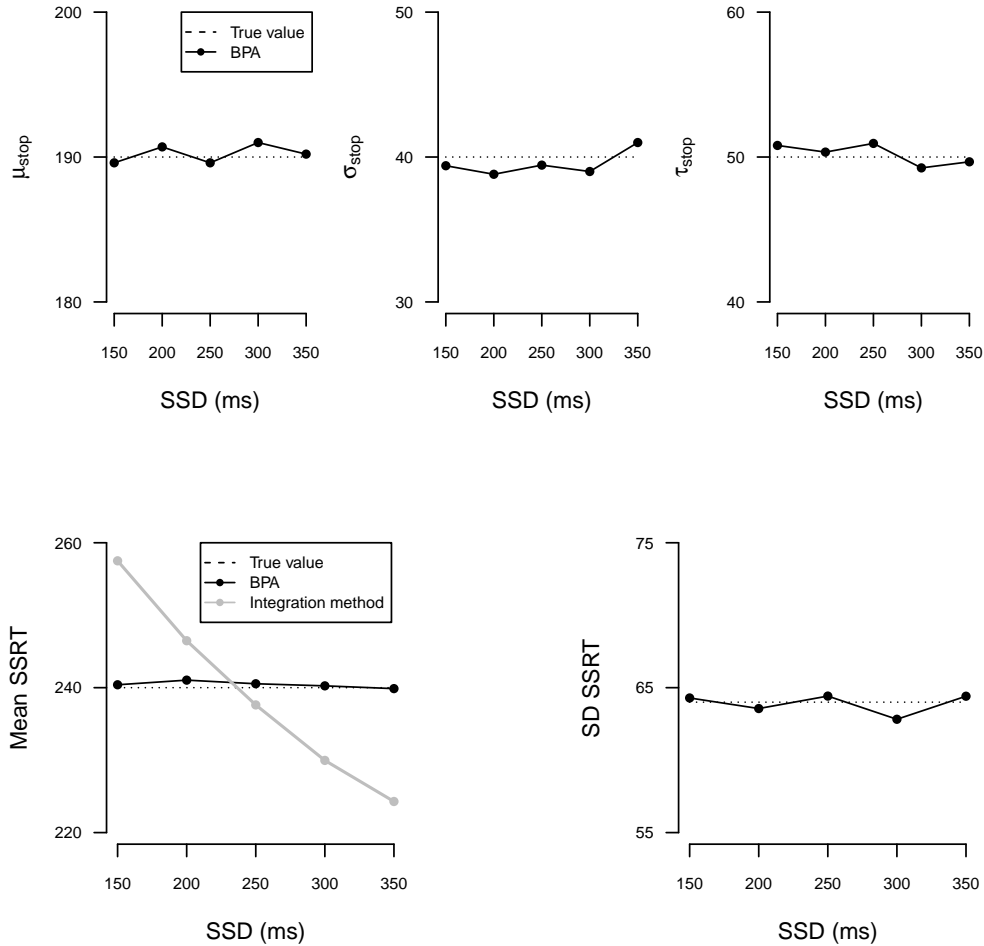


Figure 3.9 *Posterior medians of the stop parameters and the mean and standard deviation of the stop-signal reaction time (SSRT) distribution from the first recovery study for the individual Bayesian parametric approach.* The results are based on one data set containing 200,000 go trials and 100,000 stop-signal trials per stop-signal delay (SSD). The estimated μ_{stop} , σ_{stop} , and τ_{stop} parameters were free to vary across the five SSDs. The dashed lines give the true value of the stop parameters and the true mean and standard deviation of the SSRT distribution. In the top panels, the black bullets show the posterior median of the estimated stop parameters. In the bottom panels, the black bullets show the estimated mean and standard deviation of the SSRT distribution computed with the the posterior median of the stop parameters. The gray bullets show SSRT estimates computed with the traditional integration method.

SSDs recovered the mean of the generating SSRT distribution very accurately even with only 25 stop-signal trials per SSD.

To summarize, the results of the two simulation studies indicated that the individual BPA accurately recovered the parameters of the generating SSRT distribution. Also, similar to the integration method, the BPA recovered the mean of the generating SSRT distribution very accurately. As the number of stop-signal trials increased, the stop parameters and the mean SSRT were estimated more precisely. Nevertheless, the individual BPA was able to provide reasonable estimates even with relatively scarce data (i.e., 5×25 stop-signal trials) often encountered in stop-signal studies.

Hierarchical BPA

Methods

We generated 100 data sets from the horse race model, each containing the stop-signal data of $j = 1, \dots, 25$ participants. The individual parameters μ_{go_j} , σ_{go_j} , τ_{go_j} , μ_{stop_j} , σ_{stop_j} , and τ_{stop_j} were drawn from truncated normal distributions. The individual parameters were then used to generate 300 go trials and 100 stop-signal trials for each participant, using the ex-Gaussian distribution. The generating parameter values are shown in Figure 3.12. For computational efficiency, we used a single SSD per participant that produced a $P(\text{respond} \mid \text{stop signal})$ equal to 0.50.

We fit the 100 data sets with the hierarchical BPA using WinBUGS. We ran three MCMC chains and used overdispersed starting values. The first 3,000 samples of each MCMC chain were discarded. The reported parameter estimates are based on $3 \times 7,750$ recorded samples.

Results

In this section, we focus exclusively on results related to the group-level parameters of the go RT and the SSRT distribution. The individual parameter estimates from the hierarchical BPA is discussed in the next section with experimental data.

Figure 3.12 shows the posterior median of the group-level parameters averaged over the 100 replications. The hierarchical BPA recovered the group-level parameters quite accurately. The posterior standard deviations and the standard errors are typically larger for the stop parameters than for the go parameters. This result is not surprising because the go parameters are estimated based on the go RTs as well as the signal-respond RTs. The go parameters are therefore better constrained by the data than the stop parameters.

In sum, the results of the simulation studies indicate that the individual and the hierarchical BPA accurately recovered the true individual stop parameters and the generating group-level parameters, respectively. In contrast to Colonius' (1990) method, the BPA resulted in accurate estimates with a reasonable amount of data. The individual BPA provided accurate estimates for the stop parameters with only 125 stop-signal trials per participant. The hierarchical BPA yielded precise group-level stop parameters with a modest sample size of only 25 participants, each performing as few as 100 stop-signal trials.

3.5 Fitting Experimental Data

The aim of this section is to illustrate the application of the BPA using the stop-signal data set reported by Bissett and Logan (2011). Bissett and Logan presented participants with two sessions

3. BAYESIAN PARAMETRIC ESTIMATION OF STOP-SIGNAL REACTION TIME DISTRIBUTIONS

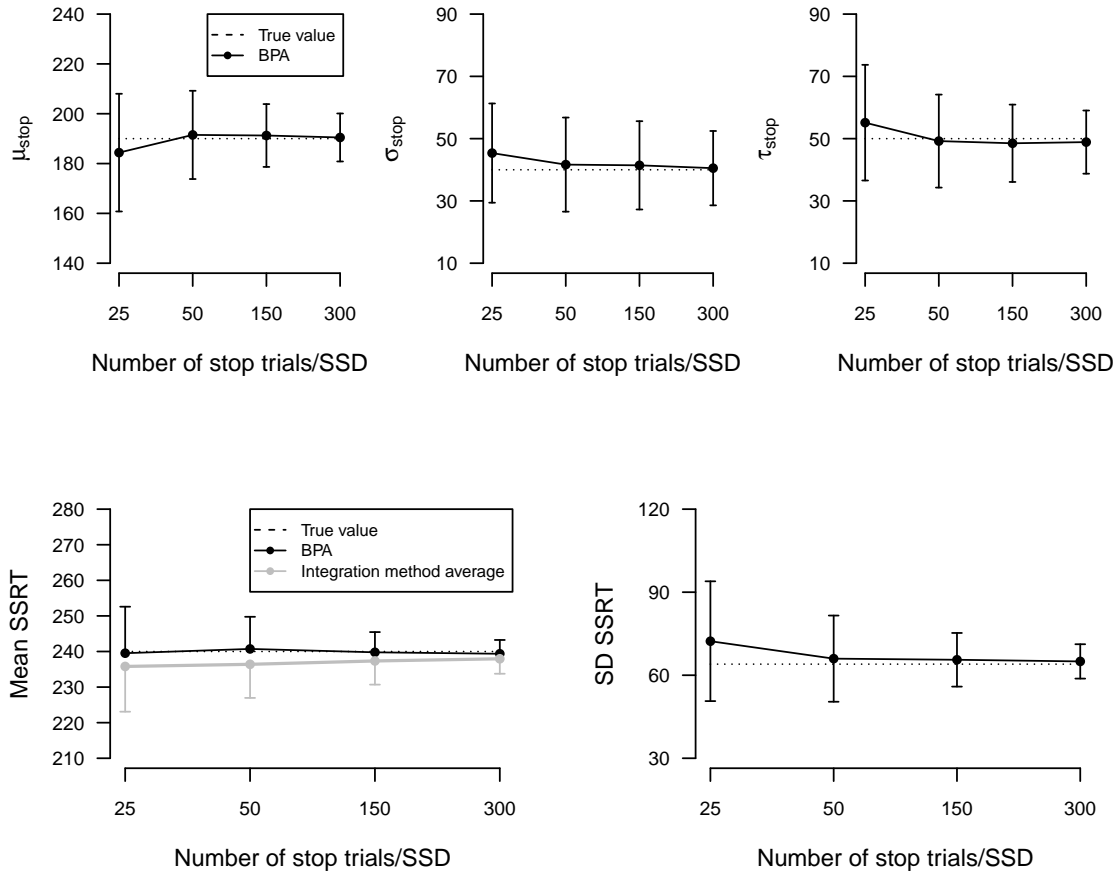


Figure 3.10 *Posterior medians of the stop parameters and the mean and standard deviation of the stop-signal reaction time (SSRT) distribution from the second recovery study for the individual Bayesian parametric approach.* We conducted four sets of simulations that varied the the number of go and stop-signal trials, with 100 data sets for each set. The estimated μ_{stop} , σ_{stop} , and τ_{stop} parameters were constrained to be equal across the five stop-signal delays (SSDs). The dashed lines give the true value of the stop parameters and the true mean and standard deviation of the SSRT distribution. In the top panels, the black bullets show the mean of the posterior medians of the estimated stop parameters across the 100 replications. In the bottom panels, the black bullets show the mean of the estimated mean and standard deviation of the SSRT distribution computed with the posterior median of the stop parameters across the 100 replications. The gray bullets show SSRT estimates computed by averaging the integration method SSRT estimates over the five SSDs. The vertical lines indicate the size of the standard error across the 100 replications.

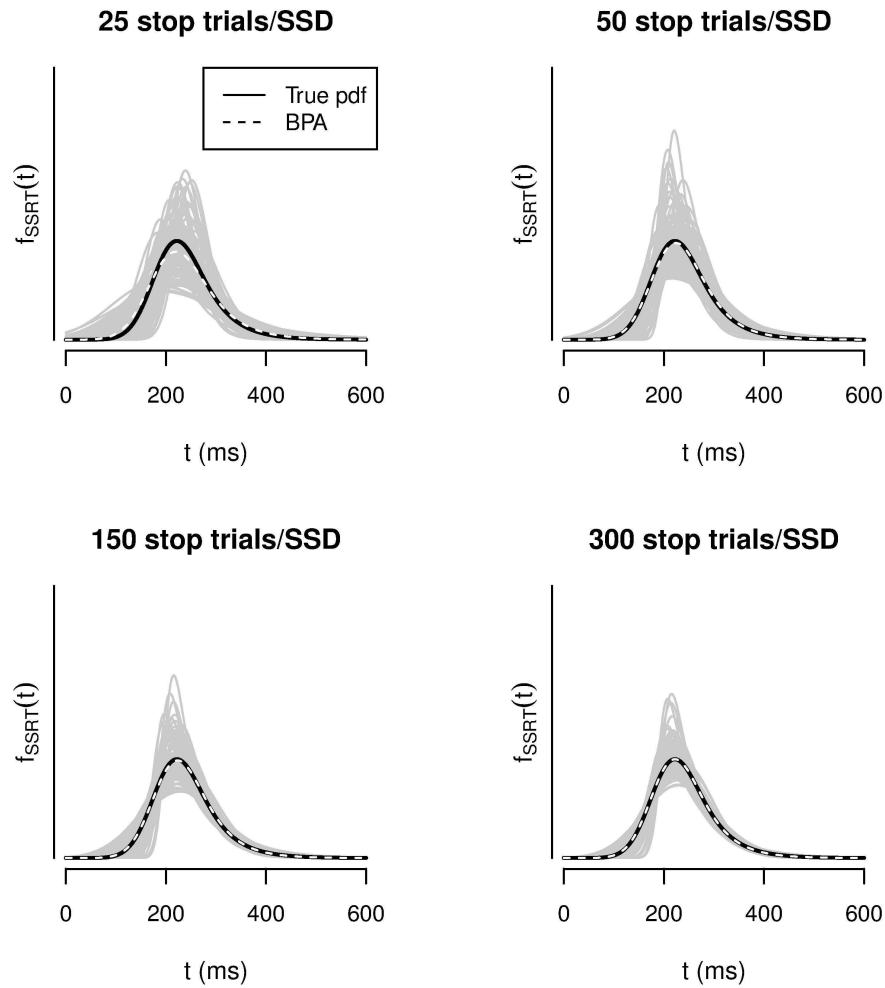


Figure 3.11 *Estimated stop-signal reaction time (SSRT) distributions from the second recovery study for the individual Bayesian parametric approach.* We conducted four sets of simulations that varied the number of go and stop-signal trials, with 100 data sets for each set. The estimated μ_{stop} , σ_{stop} , and τ_{stop} parameters were constrained to be equal across the five stop-signal delays (SSDs). The solid black line shows the true SSRT distribution. The gray lines show the SSRT distributions based on the posterior medians of the 100 replications. The dashed white line shows the SSRT distribution based on the mean of the posterior medians of the stop parameters across the 100 replications. pdf = probability density function.

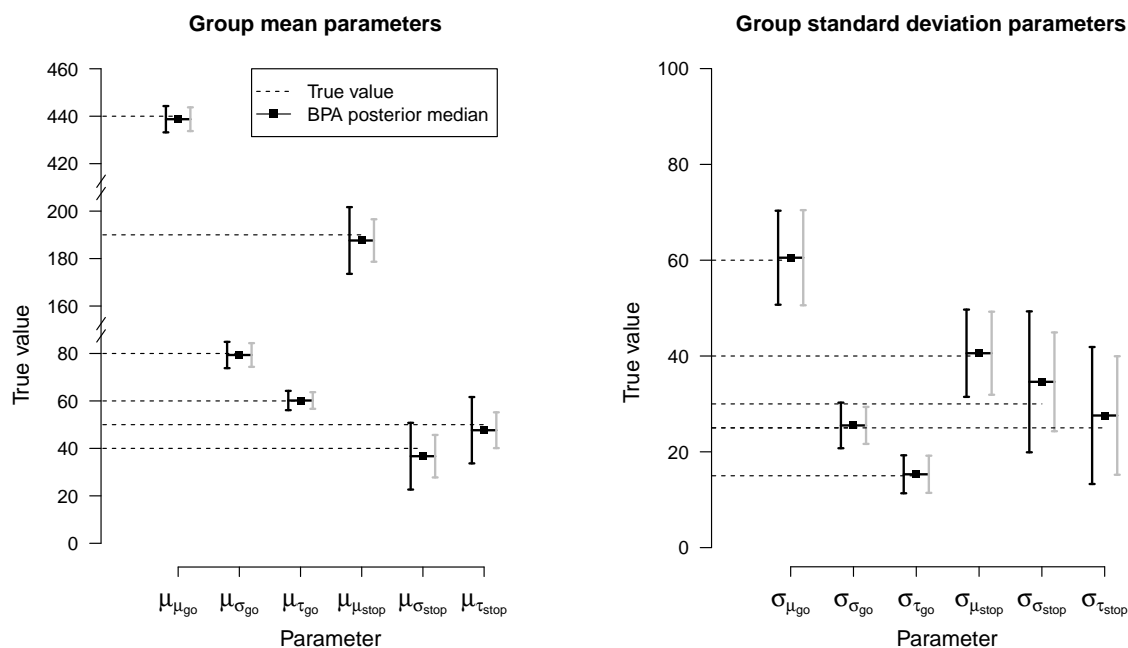


Figure 3.12 *Posterior medians of the group-level parameters from the hierarchical parameter recovery study.* We generated 100 data sets from the horse race model, each containing the stop-signal data of 25 participants responding to 300 go trials and 100 stop-signal trials. The dashed lines give the true value of the parameters. The black bullets indicate the mean of the posterior medians across the 100 replications. The black vertical lines show the size of the posterior standard deviations averaged across the 100 replications. The gray vertical lines indicate the size of the standard error. BPA = Bayesian parametric approach.

of the stop-signal task in order to investigate the adjustment of speed and caution in a dual-task environment. Here we focus on the first experiment of the Bissett and Logan study that manipulated the percentage of stop-signal trials across two sessions. The authors concluded that the two experimental sessions did not differ significantly in mean SSRT.

The Data Set

The go task required the 24 participants to respond to the shape of the presented stimuli. For instance, participants responded by pressing the “1” key on the computer keyboard when presented with a triangle or a circle, and by pressing the “0” key when presented with a square or a diamond. Each participant performed two sessions of the task. The first session featured 960 go trials and 240 stop-signal trials, resulting in 20% stop-signal trials. The second session featured 720 go trials and 480 stop-signal trials, resulting in 40% stop-signal trials. The SSD was set using the staircase tracking procedure; SSD was lengthened by 50 ms after successful inhibitions and SSD was shortened by 50 ms after incorrect responses, yielding 50% inhibition for each participant.

Incorrect RTs and RTs shorter than 200 ms and longer than 1,850 ms were excluded from all

subsequent analyses (see Bissett & Logan, 2011). As the ex-Gaussian distribution is sensitive to outliers, we also removed RTs that were slower or faster than a given participant’s mean RT plus or minus $2 \times$ the standard deviation. For comparison, we report the results of fitting the raw data with the individual BPA and the results of fitting the data without the outliers. Moreover, we excluded four participants with erratic stop-signal performance, such as extremely long and variable go RTs and a very large number of SSDs.

Individual BPA

The individual BPA was fit to the Bissett and Logan (2011) data set with WinBUGS. The estimated μ_{stop} , σ_{stop} , and τ_{stop} parameters of each participant were constrained to be equal across the different SSDs. We ran three MCMC chains and used overdispersed starting values. The analyses were based on $3 \times 5,000$ recorded samples.

The parameters of the go RT distribution were estimated precisely for all of the 20 participants. The remainder of this section focuses exclusively on the parameters of the SSRT distribution. With the exception of four participants in the first session, the posterior distributions of the stop parameters were estimated well. For the four exceptions, we obtained unrealistically large posterior medians for τ_{stop} and/or very large posterior standard deviations for σ_{stop} and τ_{stop} .⁴ In these cases, the stop parameters were thus estimated with great uncertainty, resulting in uninformative SSRT distributions and uninterpretable mean SSRT estimates.

The results from fitting the Bissett and Logan (2011) data set with the individual BPA are shown in Figure 3.13 and Figure 3.14. Figure 3.13 compares the mean SSRT of each participant computed with the mean method to the mean SSRT computed with the BPA posterior medians of the stop parameters. The BPA produced mean SSRT estimates very similar to those obtained by the mean method. The correspondence between the two methods further improved after the outliers were removed; this is not surprising because the two methods are affected to different degrees by the presence of outliers. Also, the agreement between the two sets of estimates is better for the second session than for the first session. Again, this is to be expected because the second session featured twice as many stop-signal trials than the first session, resulting in more accurate estimates for both methods.

The circles in the left panels of Figure 3.13 mark the three data points with the largest discrepancy between the two methods. The three estimates are clustered together at very high values of BPA mean SSRT. Note that these mean SSRTs belonged to three of the participants with uninformative posterior distributions with high medians and very large standard deviations for τ_{stop} . The high posterior median for τ_{stop} resulted in unusually high BPA mean SSRTs (see Equation 3.11). However, due to the large posterior uncertainty of τ_{stop} , the resulting mean SSRT estimates are uninterpretable. Lastly, consider the data point marked with “A” in the bottom right panel of Figure 3.13. For this mean SSRT, the mean method resulted in an unrealistic estimate of 118 ms. The BPA, however, yielded a more reasonable estimate of 209 ms.

The first published estimates of entire SSRT distributions are shown in Figure 3.14. The gray SSRT distributions are based on the posterior medians of the individual stop parameters. There is considerable between-participant variability in the shape of the SSRT distributions. Some distributions are very peaked, whereas others are more spread out indicating substantial within-participant variability in SSRT. Note the few extremely flat distributions with very large variance and long tail in the left panels of Figure 3.14. These flat distributions belonged to the four

⁴For these four participants, we used a uniform prior distribution ranging from 1 to 450 for τ_{stop} to accommodate the extreme parameter estimates. Note also that these participants are not the same as the four participants who were previously excluded from the analyses.

3. BAYESIAN PARAMETRIC ESTIMATION OF STOP-SIGNAL REACTION TIME DISTRIBUTIONS

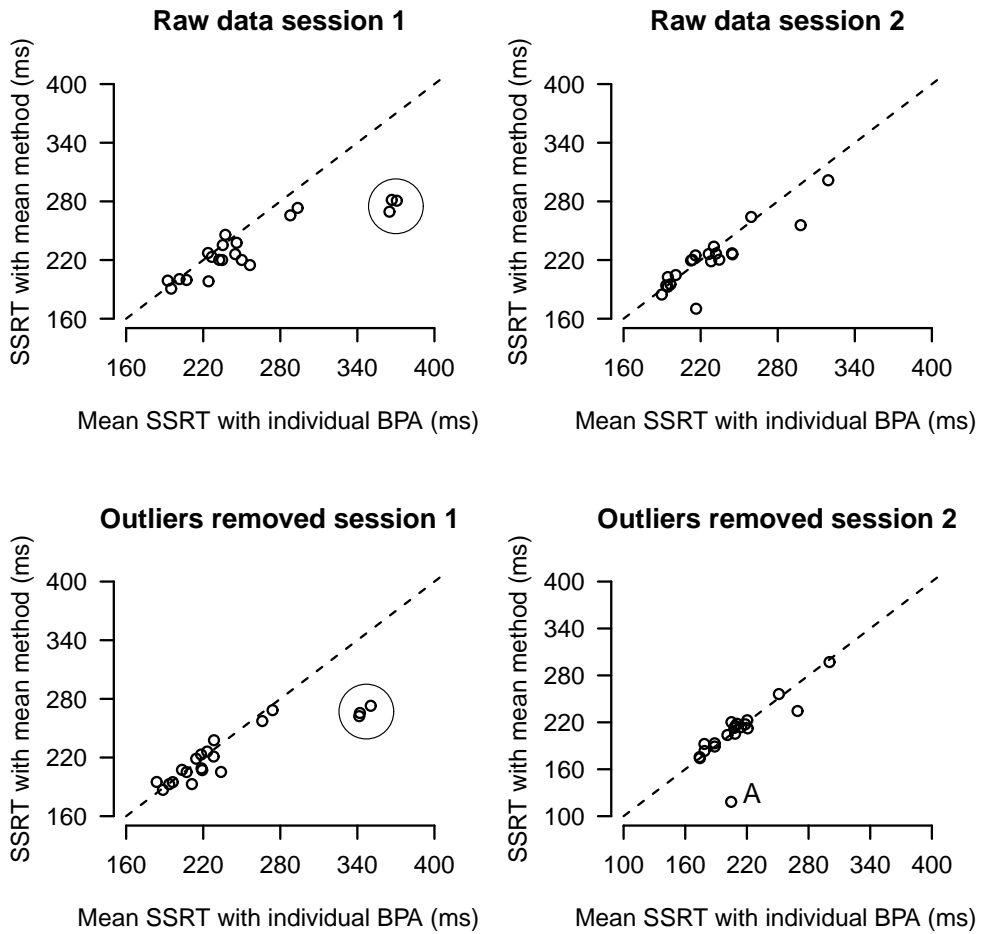


Figure 3.13 Comparison of mean stop-signal reaction times (SSRTs) computed using the mean method and the individual Bayesian parametric approach (BPA) posterior medians in the Bissett and Logan (2011) data set. The data points marked with circles represent mean SSRTs that are based on the imprecise and therefore uninterpretable posterior distributions. The data point marked with “A” in the bottom right panel represents a mean SSRT estimate for which the individual BPA resulted in a more reasonable estimate than the mean method.

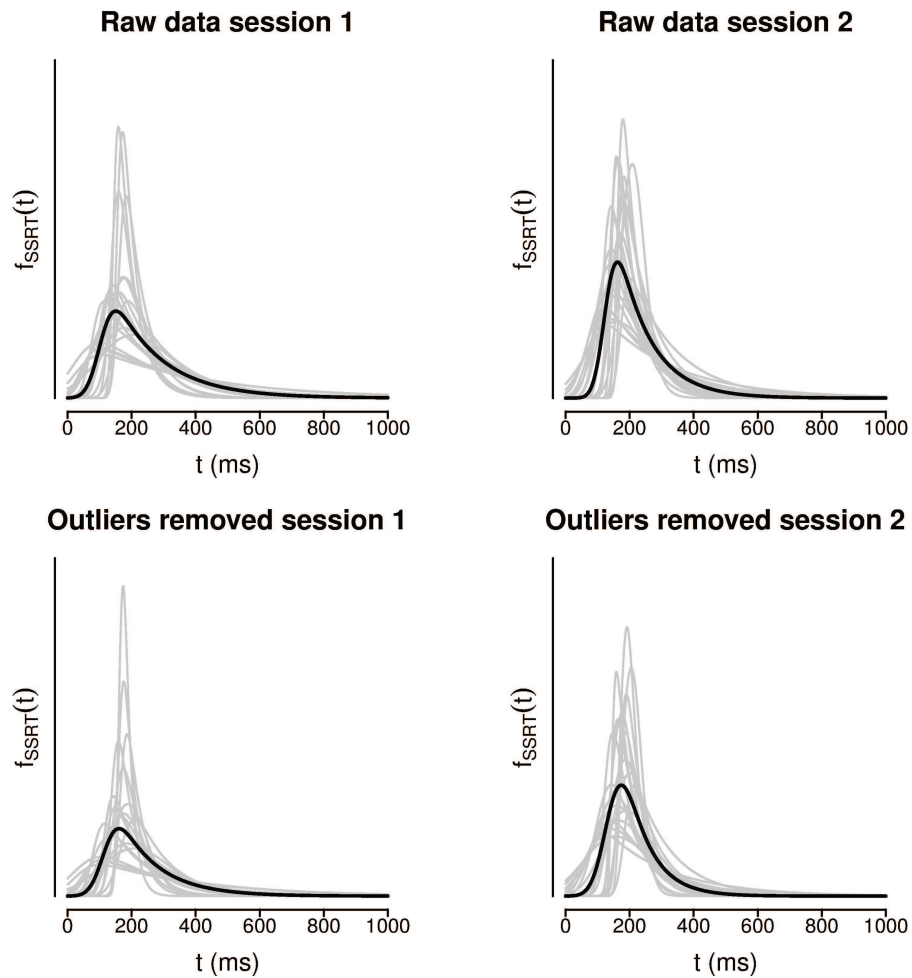


Figure 3.14 *Estimated stop-signal reaction times (SSRT) distributions for the Bissett and Logan (2011) data set.* The gray lines show the SSRT distributions based on the posterior medians of the stop parameters of each individual participant. The black line shows the SSRT distribution based on the mean of the posterior medians of the stop parameters across the 20 participants.

participants with the uninformative posterior distributions. The resulting SSRT distributions are therefore also uninformative.

The solid black line in Figure 3.14 shows the average SSRT distribution created using the mean of the posterior medians of the individual μ_{stop} , σ_{stop} and τ_{stop} parameters across participants. There are substantial differences between the shape of the average SSRT distributions in the two sessions of the experiment. The average SSRT distribution for the first session is spread out and has a fast leading edge and a long tail. In contrast, the average SSRT distribution for the second session is more peaked, with a slower leading edge and a shorter tail. Despite these differences, consistent with the results of Bissett and Logan (2011), the means of the two distributions are roughly equal. Similar to the example shown in Figure 3.1, ignoring the differences in the shape of these SSRT distributions would lead to the incorrect conclusion that the two experimental sessions do not differ with respect to SSRT.

In conclusion, the individual BPA provided well-behaved posterior distributions for most par-

ticipants. Moreover, the mean SSRTs computed with the BPA posterior medians accurately approximated the mean SSRTs obtained by the traditional mean method. In the next section, we formally investigate whether the individual BPA adequately described the observed data.

Assessing Model Fit Using Posterior Predictive Model Checks

We used posterior predictive model checks to determine whether the individual BPA produced parameter estimates that adequately describe the Bissett and Logan (2011) data. Posterior predictive model checks are frequently used procedures in Bayesian inference to assess the absolute goodness of fit of a proposed model (e.g., Gelman & Hill, 2007; Gelman, Meng, & Stern, 1996). In posterior predictive checks, we assess the adequacy of the model by generating new data (i.e., predictions) using the posterior distributions of the parameters obtained from fitting the model. If the model adequately describes the data, the predictions based on the model parameters should closely approximate the observed data.

We formalized the model checks by computing posterior predictive p values (e.g., Gelman & Hill, 2007; Gelman et al., 1996). We first defined a test statistic T , and computed its value for the observed data: $T(data)$. For each of the $i = 1, \dots, N$ draws from the posterior distribution of the parameters, we sampled new stop-signal data, $data^* = (data_1^*, data_2^*, \dots, data_N^*)$, using the ex-Gaussian assumption. Lastly, we calculated the test statistic T for each $data_i^*$: $T(data_i^*)$. The posterior predictive p value is given by the fraction of times that $T(data^*)$ is greater than $T(data)$. The posterior predictive p value compares thus the observed value of the test statistic to its sampling distribution under the assumptions of the BPA. Extreme p values close to 0 or 1 (e.g., lower than 0.05 or higher than 0.95) indicate that the BPA does not describe the observed data adequately. For each participant we conducted two posterior predictive analyses using different test statistics.

In the first posterior predictive analysis, we compared the observed signal-respond RT distribution to the signal-respond RTs predicted by the posterior distribution of the model parameters. The model check was performed only for the SSD with the highest number of observed signal-respond RTs in order to obtain stable observed and predicted signal-respond RT distributions. For each participant, we randomly selected $N = 1,000$ parameter vectors from the joint posterior of μ_{go} , σ_{go} , τ_{go} , μ_{stop} , σ_{stop} , and τ_{stop} . Then, we generated 1,000 stop-signal data sets using the 1,000 parameter vectors, the chosen SSD and the corresponding number of stop-signal trials. We used the median of the signal-respond RTs of the observed and the predicted distributions as test statistic. For each participant, the 1,000 predicted signal-respond RT distributions were compared to the observed signal-respond RT distribution, using posterior predictive p values and visual inspection of the distributions.

Figure 3.15 shows the observed go RT and signal-respond RT distributions, and 100 randomly chosen predicted signal-respond RT distributions for six participants with satisfactory model fit. The predicted signal-respond RT distributions (i.e., gray lines) adequately followed the shape of the observed signal-respond RT distribution. Also, the predicted signal-respond RTs were generally faster than the observed go RTs (i.e., dashed line), a common finding that follows from the architecture of the horse race model (Logan & Cowan, 1984). Lastly, the average of the medians of the predicted signal-respond distributions closely matched the observed median. This result is also evident from the posterior predictive p values listed in the second column of Table 3.1. The posterior predictive p values for these six participants are well within the 0.05-0.95 range, indicating that the BPA adequately accounted for the median of the observed signal-respond RTs.

Figure 3.16 shows the observed go RT and signal-respond RT distributions, and 100 randomly chosen predicted signal-respond RT distributions for three participants with unsatisfactory model fit. Note that Participant 16 and Participant 20 were among the few cases that produced un-

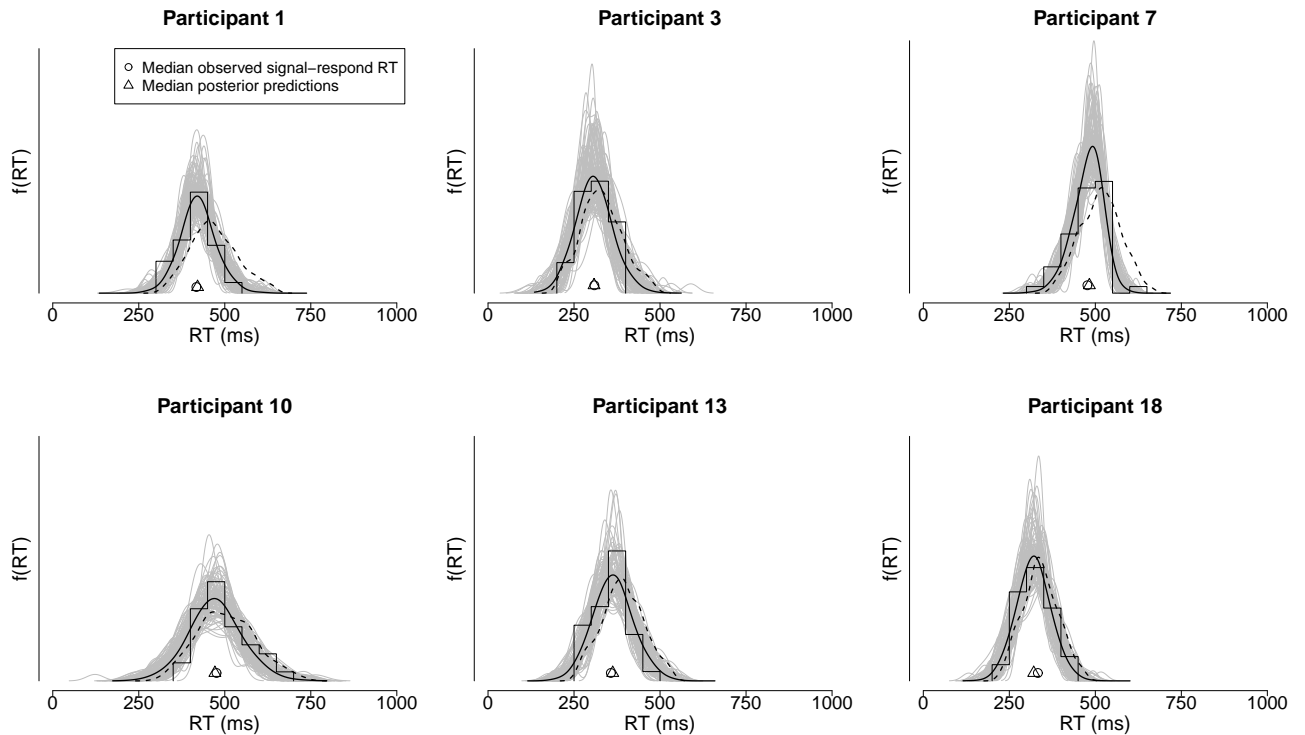


Figure 3.15 *Examples of satisfactory model fit with the individual Bayesian parametric approach (BPA): Predicted and observed signal-respond response time (RT) distributions for the first session of the Bissett and Logan (2011) experiment.* See text for a detailed description of the posterior predictive analyses. The histogram shows the observed signal-respond RT distribution. The gray lines show 100 randomly chosen predicted signal-respond RT distributions. The solid black line gives a predicted signal-respond RT distributions based on the mean of the posterior predictions. The circle indicates the median of the observed signal-respond RTs. The triangle indicates the median of the predicted signal-respond RTs. The median of the predicted signal-respond RTs is computed as the mean of the medians of the predicted signal-respond RT distributions. The dashed line shows the observed go RT distribution.

informative posterior distributions. The location and shape of the observed signal-respond RT distributions were not well approximated by the predicted signal-respond RT distributions. For Participant 8 and Participant 16, the predicted signal-respond RTs are shifted to the right. For Participant 20, the predicted signal-respond RT distribution fails to capture the bimodality of the observed signal-respond RTs. For Participant 16 and 20, the predicted signal-respond RTs are less variable than the observed signal-respond RTs. Moreover, the predicted signal-respond RTs are not substantially faster than the observed go RTs. For Participant 8 and Participant 16, the median of the predicted signal-respond RTs overestimated the median of the observed signal-respond RTs. In contrast, for Participant 20, the median of the predicted signal-respond RTs underestimated the observed median. These latter results are also shown in Table 3.1. The posterior predictive p values for these three participants are very close to or are equal to 0 or 1, indicating that the BPA failed to account for the median of the observed signal-respond RTs.

In the second posterior predictive analysis, we compared the observed response rates to the

3. BAYESIAN PARAMETRIC ESTIMATION OF STOP-SIGNAL REACTION TIME DISTRIBUTIONS

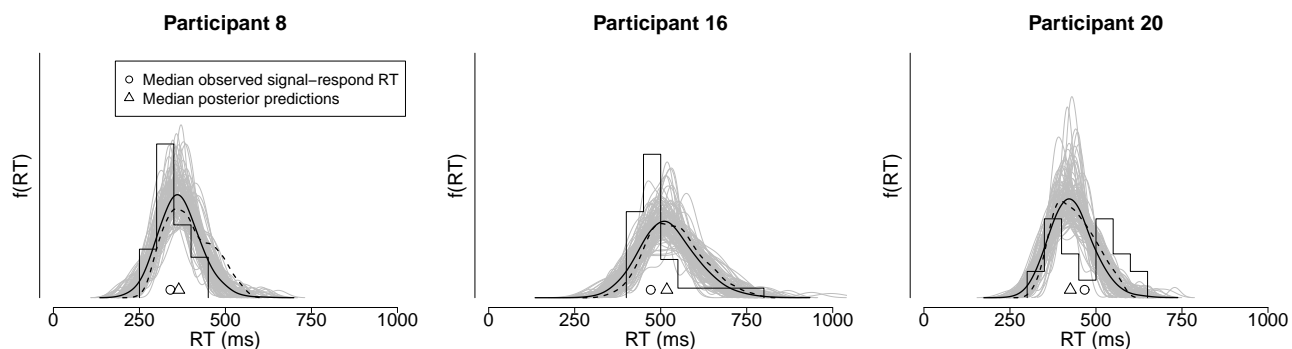


Figure 3.16 *Examples of unsatisfactory model fit with the individual Bayesian parametric approach (BPA): Predicted and observed signal-respond response time (RT) distributions for the first session of the Bissett and Logan (2011) experiment. See Figure 3.15 for details.*

Table 3.1 Posterior Predictive p Values for the Median of the Signal-Respond Response Time Distribution and the Response Rate for the First Session of the Bissett and Logan (2011) Experiment Computed From the Parameter Estimates From the Individual Bayesian Parametric Approach.

Participant	p value median	Minimum p value RR	Maximum p value RR
1	0.64	0.11	0.91
3	0.44	0.10	0.79
7	0.77	0.13	0.70
8	0.98	0.03	0.95
10	0.41	0.28	0.86
13	0.69	0.08	0.92
16	1.00	0.30	0.96
18	0.11	0.25	0.61
20	0.00	0.33	0.96

Note. Posterior predictive p values that indicate unsatisfactory model fit are shown in bold. p value median = posterior predictive p value for the median of the signal-respond response time distribution on the stop-signal delay (SSD) with the highest number of observed signal-respond trials; minimum p value RR = the lowest posterior predictive p value for the response rate (RR) computed for the SSDs that contained at least 10% of the trials; maximum p value RR = the highest posterior predictive p value for the RR computed for the SSDs that contained at least 10% of the trials.

response rates predicted by the posterior distribution of the model parameters. The model check was performed for the SSDs that featured at least 10% of the total number of stop-signal trials. For each participant, we generated 1,000 stop-signal data sets using the 1,000 parameter vectors selected for the first posterior predictive analysis, the chosen SSDs and the corresponding number of stop-signal trials. We computed posterior predictive p values for each participant on each SSD separately, where we used the observed and predicted response rates as test statistic.

Table 3.1 shows the minimum and the maximum of the posterior predictive p values for the response rates across the various SSDs of nine participants. For Participant 1, 3, 7, 10, 13 and 18, the minimum and maximum of the p values all lie between 0.05 and 0.95, corroborating our previous conclusion of satisfactory model fit with the median of the signal-respond RTs. In contrast, for Participant 8, 16, and 20, the minimum or maximum of the p values are very close to 0 or 1, supporting our previous finding that the BPA failed to account for the data of these participants.

In sum, the results of the posterior predictive model checks indicated that for most participants the BPA provided plausible parameter estimates that adequately describe the observed data. Additionally, the posterior predictions supported our earlier conclusion that the BPA resulted in uninterpretable parameter estimates for participants with imprecise posterior distributions. The results of the posterior predictive model checks for the remaining participants in the first as well as the second session of the Bissett and Logan (2011) experiment are available in the supplemental materials at <http://dora.erbe-matzke.com/publications.html>.

Hierarchical BPA

The hierarchical approach has the potential to provide accurate parameter estimates with relatively few observations per participant. We therefore did not analyze the complete Bissett and Logan (2011) data set, but used only a subsample of the available go RTs and signal-respond RTs from the second experimental session. Per participant, we fit a randomly selected 90 go RTs, 30 signal-respond RTs, and 30 successful inhibitions after removing the outliers.

The hierarchical BPA was fit to the data with WinBUGS. The estimated individual μ_{stop} , σ_{stop} , and τ_{stop} parameters were constrained to be equal across the different SSDs. We ran three MCMC chains and used overdispersed starting values. The hierarchical analysis was based on $3 \times 23,750$ samples.

The hierarchical BPA resulted in informative posterior distributions for the group-level parameters of the go RT as well as the SSRT distribution. As before, this section focuses exclusively on the parameters of the SSRT distribution. Figure 3.17 shows the prior and the posterior distributions of the group-level stop parameters. As for the hierarchical recovery study, the $\mu_{\mu_{stop}}$ and $\sigma_{\mu_{stop}}$ parameters were estimated the most precisely as indicated by the small posterior standard deviations. Also in line with the simulations, the $\sigma_{\sigma_{stop}}$ parameter was estimated with the largest posterior uncertainty. Nevertheless, the group-level stop parameters were estimated relatively well given the scarce data and the small sample size.

For most participants, the individual μ_{stop} , σ_{stop} , and τ_{stop} parameters were estimated adequately as evidenced by the well-behaved posterior distributions. For the same subsample of the data, the posterior distributions estimated with the hierarchical BPA were less variable than the posteriors estimated with the individual BPA. In fact, the 60 stop-signal trials were occasionally insufficient to obtain informative posterior distributions with the individual BPA. Figure 3.18 illustrates the benefits of hierarchical modeling for a representative participant. For the same subsample of the data, the 95% Bayesian confidence intervals are smaller for the posterior distributions estimated with the hierarchical BPA (i.e., gray line) than for the posteriors estimated with the individual BPA (i.e., black line). Also, the posterior medians from the hierarchical analysis

3. BAYESIAN PARAMETRIC ESTIMATION OF STOP-SIGNAL REACTION TIME DISTRIBUTIONS

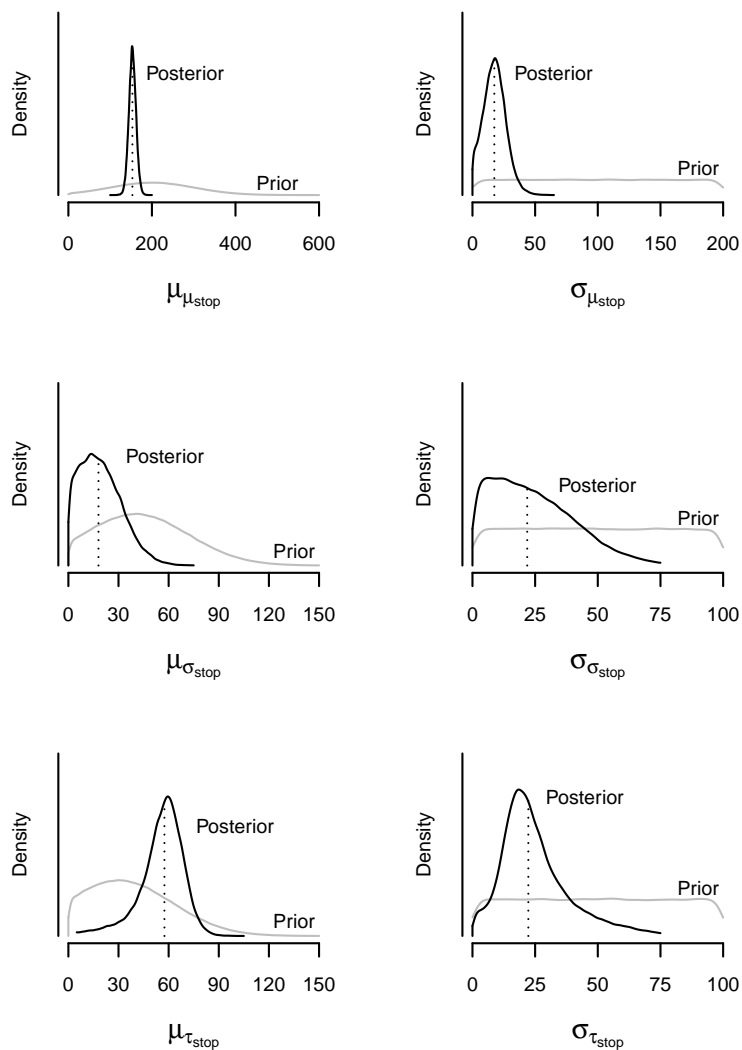


Figure 3.17 *Posterior distributions for the group-level stop parameters from a subsample of the Bissett and Logan (2011) data set.* The black lines show the posterior distributions and the gray lines show the prior distributions of the group-level parameters. The dashed lines give the posterior medians.

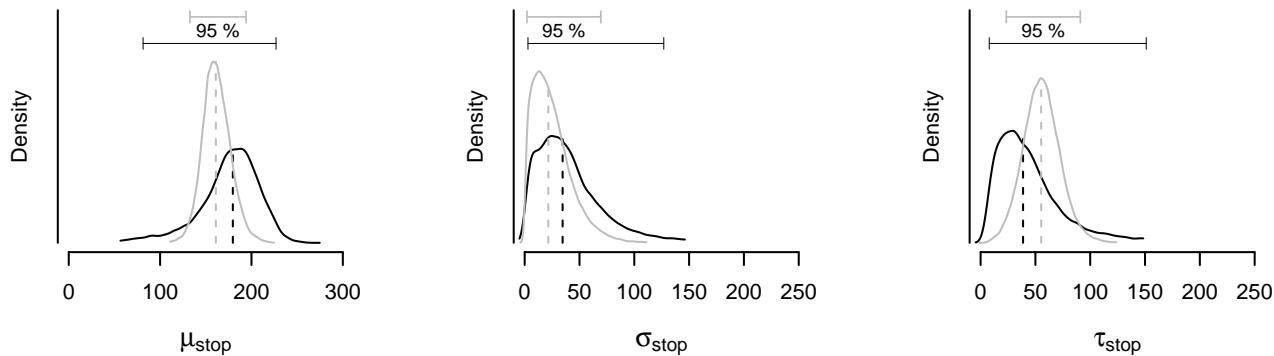


Figure 3.18 *Posterior distribution of the stop parameters for Participant 1 from the Bissett and Logan (2011) data set estimated using the individual and the hierarchical Bayesian parametric approach (BPA). The solid black and gray lines show the posterior distribution of the stop parameters and the corresponding 95% Bayesian confidence 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.*

are slightly pulled towards their corresponding group mean, a typical consequence of hierarchical modeling.

As pointed out above, hierarchical Bayesian estimation has the potential to reduce the variability in the estimated parameters compared to individual-level estimation. Figure 3.19 compares the mean SSRTs computed with the traditional mean method, the posterior medians from the individual BPA, and the posterior medians from the hierarchical BPA with the same subsample of the data. As shown in Figure 3.19a, the individual BPA provided mean SSRTs that are slightly less variable than the mean SSRTs obtained using the mean method. More importantly, Figure 3.19b and Figure 3.19c show that the hierarchical BPA resulted in mean SSRTs that are substantially less variable than the mean SSRTs obtained with either the mean method or the individual BPA.

Assessing Model Fit Using Posterior Predictive Model Checks

We used posterior predictive model checks to determine whether the individual parameter estimates from the hierarchical BPA adequately describe the Bissett and Logan (2011) data. As the data of most participants featured fewer than 10 observed signal-respond RTs even on the SSD with the highest number of observations, posterior predictive model checks using the signal-respond RT distributions are not sensible for the present data set. For each participant, we compared the observed response rates to the response rates predicted by the posterior distribution of the model parameters. The posterior predictive analyses followed the procedure described for the individual BPA, using $N = 1,000$ parameter vectors from the joint posterior of the individual μ_{go} , σ_{go} , τ_{go} , μ_{stop} , σ_{stop} , and τ_{stop} parameters.

Table 3.2 shows the minimum and the maximum of the posterior predictive p values for the response rates across the various SSDs. With the exception of Participant 12 and 19, the minimum and maximum value of the posterior predictive p values all lie well within the 0.05 - 0.95 range,

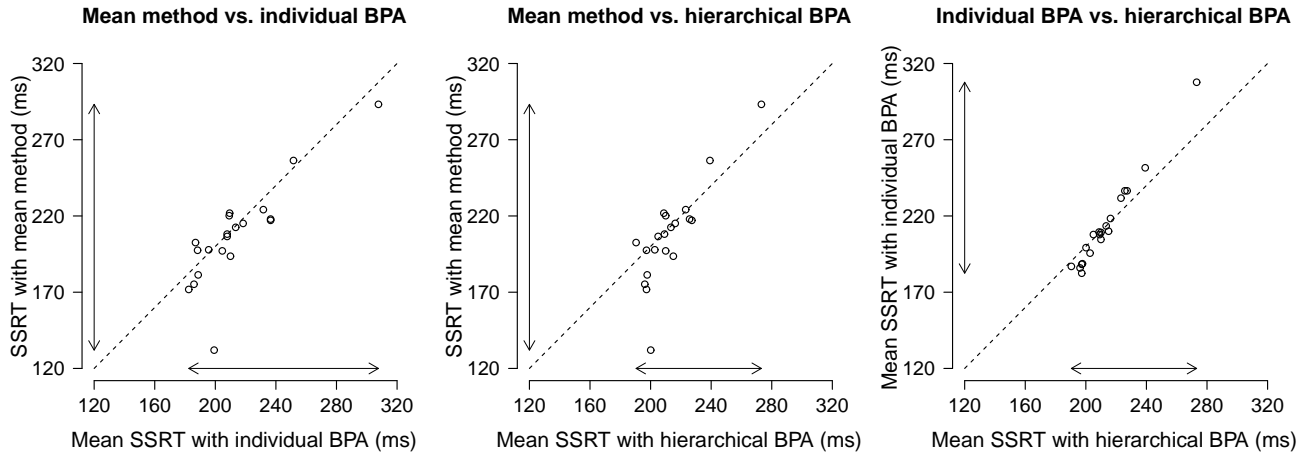


Figure 3.19 *Comparison of mean stop-signal reaction times (SSRTs) obtained with the mean method, with the posterior medians of the stop parameters from the individual Bayesian parametric approach (BPA), and with the posterior medians of the individual stop parameters from the hierarchical BPA in the Bissett and Logan (2011) data set. The arrows indicate the range of the estimates. Note that the range is smallest for the hierarchical BPA.*

indicating that the BPA adequately accounted for the response rates of most participants.

The above section illustrated that the hierarchical BPA provided sensible group-level parameters with relatively well-calibrated posterior distributions even with a small sample size and only 60 stop-signal trials per participant. Posterior predictive model checks indicated that for most participants the hierarchical BPA provided plausible individual parameter estimates that adequately describe the observed data. Moreover, the individual parameter estimates yielded sound mean SSRTs and demonstrated the characteristic benefits of hierarchical modeling.

3.6 Discussion

The stop-signal task is a frequently used experimental measure of response inhibition. Over the past 30 years, the horse race model (Logan, 1981; Logan & Cowan, 1984) has successfully accounted for stop-signal data in different settings and has facilitated the interpretation of stopping experiments with various age groups and clinical populations (e.g., Kramer et al., 1994; Oosterlaan et al., 1998; Ridderinkhof et al., 1999; Schachar & Logan, 1990; Schachar et al., 2000; Williams et al., 1999). The horse race model offers numerous methods to estimate the otherwise unobservable latency of stopping.

The existing methods to estimate SSRT are unable to adequately estimate the shape of entire SSRT distributions. Ignoring the shape of SSRT distributions, and focusing only on the mean SSRT, may mask crucial features of the data and result in erroneous conclusions about the nature of response inhibition. The goal of this paper was therefore to introduce a novel method—a Bayesian parametric approach—that enables researchers to estimate the entire distribution of SSRTs. The BPA is based on the assumptions of the horse race model and treats response inhibition as a censoring mechanism. The method assumes that go RTs and SSRTs are ex-Gaussian distributed and relies on MCMC sampling to obtain posterior distributions for the model parameters. Note

Table 3.2 Posterior Predictive p Values for the Response Rate for the Second Session of the Bissett and Logan (2011) Experiment Computed From the Individual Parameter Estimates From the Hierarchical Bayesian Parametric Approach.

Participant	Minimum p value RR	Maximum p value RR
1	0.31	0.51
2	0.40	0.61
3	0.14	0.56
4	0.36	0.64
5	0.17	0.54
6	0.16	0.40
7	0.14	0.69
8	0.15	0.35
9	0.34	0.65
10	0.27	0.60
12	0.01	0.30
13	0.38	0.59
15	0.16	0.52
16	0.33	0.72
18	0.10	0.32
19	0.04	0.64
20	0.28	0.55
21	0.23	0.67
22	0.19	0.30
23	0.26	0.36

Note. Posterior predictive p values that indicate unsatisfactory model fit are shown in bold. Minimum p value RR = the lowest posterior predictive p value for the response rate (RR) computed for the stop-signal delays (SSDs) that contained at least 10% of the trials; maximum p value RR = the highest posterior predictive p value for the RR computed for the SSDs that contained at least 10% of the trials.

that we could have carried out parameter estimation by means of standard maximum likelihood estimation (Dolan et al., 2002; Myung, 2003). However, our goal was to develop a method for estimating SSRT distributions that may be applied to individual as well as hierarchical data structures. As maximum likelihood estimation can become practically difficult for hierarchical models, we chose to use Bayesian parameter estimation instead. This also brings along the typical benefits of Bayesian estimation, such as easy-to-use estimation software (e.g., WinBUGS) and a coherent inferential framework.

We demonstrated using simulations that the BPA adequately recovers the parameters of the generating SSRT distributions in individual and hierarchical data structures. We showed that the individual BPA can provide accurate estimates of SSRT distributions in experimental stop-signal data featuring a realistic number of trials. Similarly, we demonstrated using real data that the hierarchical BPA resulted in interpretable individual and group-level stop parameters with a small sample size and as few as 60 stop-signal trials per participant.

The BPA enables researchers to evaluate differences in the shape of SSRT distributions between clinical or experimental groups. SSRT distributions obtained from the individual BPA can be compared visually, as illustrated in the introduction and with the Bissett and Logan (2011) data set. Similarly, the group-level go and stop parameters obtained from the hierarchical BPA may be compared visually by inspecting the overlap—or the lack of overlap—between the posterior

distribution of the parameters in the different groups. Alternatively, differences in the ex-Gaussian individual go and stop parameters between clinical groups or experimental conditions may be evaluated formally using Bayesian hypothesis testing. Various user-friendly options are now available to perform, for instance, Bayesian t tests (Rouder et al., 2009; Wetzels, Raaijmakers, Jakab, & Wagenmakers, 2009) and analyses of variance (Masson, 2011; Wetzels et al., 2012) using Bayes factors (e.g., Berger & Pericchi, 1996; Dickey, 1971; Gamerman & Lopes, 2006; Kass & Raftery, 1995; Klugkist, Laudy, & Hoijtink, 2005; O’Hagan & Forster, 2004). Moreover, easy-to-use Bayesian hypothesis tests for correlations (Wetzels & Wagenmakers, 2012) and regression analyses (Liang, Paulo, Molina, Clyde, & Berger, 2008)⁵ are also available.

Although the BPA offers the advantages of estimating entire SSRT distributions, it also comes with a number of costs. One drawback is related to the amount of data that is required to obtain precise stop parameter estimates. The individual BPA may fail to provide precise estimates with the very small amount of data that is sometimes collected in developmental and clinical stop-signal studies. For a number of participants, the sampled 150 trials (i.e., 90 go and 60 stop-signal trials) of the Bissett and Logan (2011) data set were in fact insufficient to obtain informative posterior distributions for the stop parameters using the individual BPA. Nevertheless, we illustrated that the hierarchical BPA may provide a solution in such situations.

Another drawback of the BPA is related to the present implementation in WinBUGS. First, the BPA requires some basic programming skills to obtain the necessary data format for the WinBUGS script. Second, the fitting algorithm is time-consuming. Running on a fast personal computer, WinBUGS required an average of about 5 hr to reach convergence per participant in the Bissett and Logan (2011) experiment. Likewise, it took several days to fit the hierarchical BPA to the subsample of the Bissett and Logan data set. We are currently working on a user-friendly implementation of the BPA that will increase the speed of the fitting routine.

Parametric Assumptions

In contrast to the Colonius method (1990, see also de Jong et al., 1990) for estimating SSRT distributions, the BPA requires a parametric form to describe the go RTs and the SSRTs. In the abstract sense, as the Colonius method does not require any assumptions about the distribution of the go RTs and the SSRTs, it may be preferable to the BPA. In the practical sense, however, the applicability of the Colonius method is limited by the amount of data that is available per participant. In contrast to the BPA, the Colonius method requires an unrealistically large amount of data to perform adequately. For the analysis of experimental stop-signal data, the BPA is therefore preferable to the Colonius method. Of course, the practical advantage of the BPA comes at a price: We need parametric assumptions to quantify the shape of the go RT and the SSRT distributions.

Here we assumed that the go RTs and the SSRTs are ex-Gaussian distributed. Note, however, that the BPA does not hinge on the particular parametric form used to summarize the distributions. The ex-Gaussian distribution is used as a convenient choice to quantify the go RTs and the SSRTs. The ex-Gaussian is a frequently used distributional model that can excellently accommodate the shape of RT distributions and is easy to fit to data (Heathcote et al., 1991; Hockley, 1982, 1984; Ratcliff, 1978, 1993; Ratcliff & Murdock, 1976). Moreover, sensitivity analyses indicated that the BPA is robust to misspecification of the parametric form of the go RT and SSRT distributions. Even when the go RTs and the SSRTs were drawn from shifted log-normal distributions, the ex-Gaussian based BPA excellently approximated the shape of their distribution.

⁵For software implementation, see <http://pcl.missouri.edu/bf-reg>. Retrieved May 23, 2012.

The above result is not surprising because the ex-Gaussian distribution is flexible enough to accommodate a wide range of distributions observed in RT data. Unless the go RTs and SSRTs are left skewed or bimodal—an unlikely scenario for RT distributions—the ex-Gaussian is likely to provide adequate description of their distributions. The interested reader is referred to the supplemental materials for the detailed results of the sensitivity analyses for the individual BPA. Note also that the posterior predictive model checks indicated that the ex-Gaussian distribution provided an excellent description of the go RTs and SSRTs in experimental stop-signal data.

Nevertheless, the ex-Gaussian distribution comes also with some disadvantages. Specifically, the ex-Gaussian has a number of characteristics that are atypical of empirical RT data. First, the ex-Gaussian has a monotonically increasing hazard function, while empirical hazard functions are typically peaked (e.g., W. Schwarz, 2001; Van Zandt, 2002). Second, the ex-Gaussian distribution assigns probability to unrealistically short and negative RTs. As an alternative, one may use “shifted” RT distributions with a parameter-dependent lower bound. For example, the ex-Gaussian distribution may be replaced by the shifted Wald, the shifted Weibull or the shifted log-normal distributions (e.g., Heathcote, 2004; Heathcote, Brown, & Cousineau, 2004). However, shifted distributions are notoriously difficult to fit. Moreover, in our implementation the shifted log-normal distribution resulted in somewhat less accurate estimates than the ex-Gaussian. Another alternative is to use the ex-Wald distribution (W. Schwarz, 2001) to describe the go RTs and the SSRTs. Heathcote (2004) showed, however, that the ex-Wald requires at least 400 observations to produce adequate parameter estimates, a requirement that is often not satisfied in stop-signal experiments.

Process Models

Process models of response inhibition provide further possibilities to model performance in the stop-signal paradigm. A prominent alternative to the BPA is the interactive race model (Boucher, Palmeri, Logan, & Schall, 2007), a neurally plausible instantiation of the horse race model. The interactive race model conceptualizes the go and the stop process as two noisy accumulators that race towards a fixed response threshold and may interact via inhibitory links. The interactive race model assumes constant rates of rise to the threshold and noise terms with standard deviations σ_{go} and σ_{stop} that reflect the amount of noise added in each step of the rise. Boucher et al. (2007) showed that the go and the stop process are for the most part independent. The inhibitory effect of the stop process on the go process is very brief and is much stronger than the inhibitory effect of the go process on the stop process. Note that the interactive race model applies specifically to saccadic inhibition (Verbruggen et al., 2008).

Another alternative is the Hanes-Carpenter model (Hanes & Carpenter, 1999; Hanes & Schall, 1995; Hanes, Patterson, & Schall, 1998) for saccade inhibition. The Hanes-Carpenter model is based on the Linear Approach to Threshold with Ergodic Rate (LATER; Carpenter, 1981; Carpenter & Williams, 1995). The model assumes that the competing go and the stop process rise in a linear fashion to a fixed response threshold. If the stop process reaches the threshold before the go process, the response is inhibited. If the go process reaches the threshold before the stop process, the response is executed.

The Hanes-Carpenter model is equivalent to the horse race model with specific distributional assumptions about the rate of information accumulation of the go and the stop process (Colonius, Özyurt, & Arndt, 2001). Specifically, the Hanes-Carpenter model assumes that the rates of rise are normally distributed, resulting in the following parameters: the means and the standard deviations of the rates of rise of the go and the stop process, μ_{go} , σ_{go} , μ_{stop} , and σ_{stop} , respectively. The model parameters can be estimated using Monte Carlo simulations (e.g., Asrress & Carpenter,

2001; Colonijs et al., 2001; Hanes & Carpenter, 1999) or maximum likelihood estimation (e.g., Corneil & Elsley, 2005; Kornyló, Dill, Saenz, & Krauzlis, 2003). Using the properties of the stop process accumulator, one can obtain the distribution of the finishing times of the stop process (SSRT distribution). Note, however, that in typical applications of LATER, the goal is to use the estimated rate parameters to test and compare the predictions of competing models of response inhibition and not to explicitly estimate SSRT distributions. For yet another alternative to model inhibitory control in the stop-signal task, see Shenoy, Rao, and Yu (2010) and Shenoy and Yu (2011).

LATER and the BPA constitute different perspectives on modeling response inhibition. LATER is focused on the nature of the (neural) processes underlying response inhibition and thereby makes particular assumptions of the shape for the finishing time distribution of the stop processes. In contrast, the BPA constitutes a more statistical approach. The BPA is not concerned with the nature of the underlying go and stop process; it rather focuses on how the SSRT distribution can be estimated irrespective of the particular parametric choice — be it ex-Gaussian or shifted Wald — used to quantify its shape.

Prior Distributions

The BPA uses Bayesian parameter estimation and therefore involves choosing prior distributions for the ex-Gaussian go and stop parameters. With respect to the individual BPA, the priors for the go and stop parameters are informative in the sense that they cover a wide but realistic range of values informed by results from the stop-signal literature (Williams et al., 1999; Band et al., 2003). We feel that using informative priors is justified since there is a large body of past research that provides valuable information about the plausible range of parameter values. Also, with increasing opportunity to apply the BPA to empirical data sets, we will be able to make even better informed choices about the prior distribution of the parameters. Note also that as long as sufficiently informative data are available, the data readily overwhelm the prior (e.g., M. D. Lee & Wagenmakers, 2013). Whereas Bayesian parameter estimation can be robust to changes in priors, Bayesian hypothesis testing using Bayes factors (e.g., Berger & Pericchi, 1996; Dickey, 1971; Gamerman & Lopes, 2006; Kass & Raftery, 1995; Klugkist et al., 2005; O’Hagan & Forster, 2004) can be relatively sensitive to prior inputs. The shape of the prior distribution can greatly influence the Bayes factor and the resulting inferences (e.g., Bartlett, 1957; Liu & Aitkin, 2008; but see Vanpaemel, 2010). Fortunately, various user-friendly approaches to Bayesian hypothesis testing are now available that rely on principled choices of prior distributions (e.g., Rouder et al., 2009; Wetzels et al., 2009).

With respect to the hierarchical approach, the BPA assumes that the individual go and stop parameters come from truncated normal group-level distributions. The use of normal group-level distributions is a common choice in Bayesian hierarchical modeling (e.g., Gelman & Hill, 2007; M. D. Lee & Wagenmakers, 2013). Also, sensitivity analyses indicated that the hierarchical BPA is relatively robust to misspecification of the group-level distribution of the individual go parameters. Even when the true go parameters were drawn from uniform or bimodal group-level distributions, the hierarchical BPA with truncated normal group-level distributions provided accurate individual go parameter estimates. Unfortunately, the hierarchical BPA is less robust to misspecification of the group-level distribution of the individual stop parameters. When the true stop parameters were drawn from uniform or bimodal group-level distributions, the hierarchical BPA with truncated normal group-level distributions resulted in biased parameter estimates, particularly for the σ_{stop} and τ_{stop} parameters. Fortunately, the bias decreased as the number of participants and especially as the number of trials increased. The finding that the go parameters are more robust to misspeci-

fication of the group-level distributions is not surprising. The go parameters are estimated based on the go RTs as well as the signal-respond RTs. Also, the sensitivity analyses—similar to typical stop-signal studies—featured three times as many go trials as stop-signal trials. As a result, the go parameters are more strongly constrained by the data and are less strongly influenced by their group-level distribution than the stop parameters.

Moreover, the sensitivity analyses indicated that misspecification of the group-level distributions often results in convergence problems. We therefore recommend researchers to carefully monitor the convergence of the individual parameter estimates. If there are reasons to suspect that the hierarchical assumptions are violated, we advise users to inspect the distribution of the individual go and stop parameters obtained either from the individual BPA or from the hierarchical BPA with very weak priors for the group-level parameters. If these preliminary analyses indicate that the distribution of the individual parameters substantially deviates from normality, one may use the unconstrained individual go and stop parameters. Alternatively, if substantive knowledge of the form of the group-level distributions is available, the hierarchical BPA may be adapted to accommodate the desired (mixture) distribution. The reader is referred to the supplemental materials for the detailed results of the sensitivity analyses for the hierarchical BPA.

Conclusion

Here we introduced a novel Bayesian parametric method that provides for the estimation of entire distribution of SSRTs. The new method enables researchers to evaluate differences in the shapes of SSRT distributions between various clinical populations or experimental groups. In doing so, our Bayesian parametric approach aids the interpretation of stop-signal data and may reveal some hitherto unknown aspects of response inhibition.