

Accompanying Materials for “A Diffusion Model Decomposition of the Practice Effect” A: The Bayesian Ratcliff Diffusion Model

Gilles Dutilh¹, Joachim Vandekerckhove², Francis Tuerlinckx², and Eric-Jan Wagenmakers¹

¹ University of Amsterdam

² Leuven University

Correspondence concerning this article should be addressed to:

Gilles Dutilh

University of Amsterdam, Department of Psychology

Roetersstraat 15

1018 WB Amsterdam, The Netherlands

Ph: (+31) 20-525-6629

Fax: (+31) 20-639-0279

E-mail may be sent to Gilles.Dutilh@gmail.com.

In Appendix A, we provide details regarding the Bayesian implementation of the Ratcliff diffusion model.

Notation. We will use the following notation. The “twiddles” symbol \sim means “is distributed according to”, and the function $W(x, t|a, b, t_{er}, \xi)$ is the bivariate probability density function (PDF) of the first-passage times of a Wiener diffusion process. In this function, a is the boundary separation, t_{er} is the nondecision time, ξ is the drift rate, and b is the ratio of the starting point of the diffusion to the distance between the boundaries, so that the starting point $z_0 = ab$. The full expression for the PDF can be found in Tuerlinckx (2004). We will use indexes p for participants, j for blocks, i for items, and c_i for the word/nonword value of item i . To avoid confusion with other subscripts, indexes are always put between round brackets. In general, therefore, the vector $y_{(pijc_i)}$, which consists of the accuracy score $x_{(pijc_i)}$ and the response time $t_{(pijc_i)}$, adheres to:

$$p(y_{(pijc_i)}) = W(a_{(pijc_i)}, b_{(pijc_i)}, t_{er(pijc_i)}, \xi_{(pijc_i)}).$$

Model. To arrive at the Ratcliff diffusion model, we now add random trial-to-trial variability to the t_{er} , ξ , and b parameters according to the following mixing distributions:

$$t_{er(pijc_i)} \sim N(T_{er(pj)}, s_{t_{er}(pj)}^2),$$

$$b_{(pijc_i)} \sim N(B_{(pj)}, s_b^2(pj)),$$

and

$$\xi_{(pijc_i)} \sim N(v_{(pjc_i)}, \eta_{(pjc_i)}^2).$$

With the addition of these mixing distributions, the model now has nine parameters per block per participant:

- a for the boundary separation
- B for the across-trial mean bias, meaning that the mean starting point $z = aB$

- s_b^2 for the variability in bias, so that the variance in starting point $s_z^2 = a^2 s_b^2$.

- T_{er} for the across-trial mean nondecision time

- s_t^2 for the across-trial variance of the nondecision time

- two v parameters for the mean drift rate of word stimuli and of nonword stimuli

- two η parameters for the trial-to-trial variance in drift rate of word stimuli and of nonword stimuli

Bayesian methods. In Bayesian statistics, one makes inferences about the different model parameters through their posterior distributions (Gelman, Carlin, Stern, & Rubin, 2004), that is, about $p(\text{parameter}|\text{data})$. The method is called Bayesian because the computation of the posterior distribution requires the application of Thomas Bayes’ famous theorem:

$$p(A|B) = p(B|A) p(A) / p(B),$$

where in this case the left hand side is the posterior distribution. Then $p(B|A)$ is the likelihood of the model, $p(A)$ are the prior distributions of parameters, and $p(B)$ is a normalizing constant. The computation of the posterior distribution and associated descriptive measures is often very difficult and computationally expensive, but some software packages exist to assist in the implementation. We have used WinBUGS for this purpose (Lunn, Thomas, Best, & Spiegelhalter, 2000; see also Vandekerckhove, Tuerlinckx, & Lee, 2008).

Priors. In order to initiate a Bayesian analysis, we have to define the prior distribution of each parameter. For parameter estimation, one typically chooses prior distributions that reflect a very vague knowledge of the possible parameter values (uninformative priors). For the present analysis, we have chosen priors that are somewhat informative, reflecting what we know about plausible values of the different parameters, and they are restricted to a specific interval to avoid numerical over- or underflow. The intervals are chosen to reflect

a reasonably wide range of possible parameter values. The priors were defined exactly as follows:

$$a \sim N(0.08, 1/400) I(0.02, 0.18),$$

$$T_{er} \sim U(0.01, 0.60),$$

$$B \sim N(0.50, 1/200) I(0.20, 0.80),$$

$$v \sim U(0.00, 0.90),$$

$$\eta \sim U(0.001, 0.4999),$$

$$s_b \sim U(0.001, 0.4999),$$

$$s_t \sim U(0.001, 0.4999),$$

where the second parameter of the normal distribution $N(\mu, \sigma^2)$ is always its variance, the indicator function $I(\cdot)$ truncates the distribution between its arguments, and $U(L, H)$ is the uniform distribution between L and H . The relatively large overall number of data means that the influence of the shape of the prior distributions is in fact very small. Indeed, we repeated the analysis several times with different priors and encountered no meaningful differences.

Convergence and model fit. We ran 6 independent Markov chains of 10,000 iterations (after a burn-in of 500) and tested their convergence with the \hat{R} criterion (Gelman et al., 2004). Convergence ($\hat{R} < 1.1$) was typically fast for the main parameters of the diffusion model (a , T_{er} , B , and v). Visual inspection of autocorrelation plots showed high dependency at small lags, but none after downsampling the chains by a factor of 200.

References

- Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (2004). *Bayesian data analysis (2nd ed.)*. Boca Raton (FL): Chapman & Hall/CRC.
- Lunn, D., Thomas, A., Best, N., & Spiegelhalter, D. (2000). WinBUGS — a Bayesian modelling framework: concepts, structure, and extensibility. *Statistics and Computing*, *10*, 325–337.
- Tuerlinckx, F. (2004). The efficient computation of the cumulative distribution and probability density functions in the diffusion model. *Behavior Research Methods, Instruments, & Computers*, *36*, 702–716.
- Vandekerckhove, J., Tuerlinckx, F., & Lee, M. D. (2008). Proceedings of the 30th annual conference of the cognitive science society. In B. C. Love, K. McRae, & S. V. M. (Eds.), (pp. 1429–1431). Austin, TX: Cognitive Science Society.